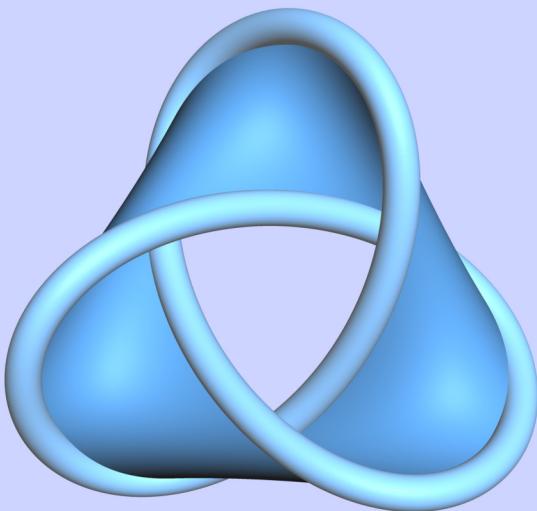


# Mathematics and Physics



*Ryan Maguire*

May 14, 2020

---

# Contents

---

<b>List of Figures</b>	xvii
<b>List of Tables</b>	xxi
<b>Preface</b>	xxiii
<b>Acknowledgements</b>	xxv

## Book One: Foundations

<b>Part I: Logic and Sets</b>	1
<b>Chapter 1: Propositional Logic</b>	3
1.1 What is Logic? . . . . .	3
1.1.1 Truth . . . . .	7
1.1.2 Sets . . . . .	12
1.1.3 Predicates and Propositions . . . . .	17
1.1.4 Rules of Inference . . . . .	19
1.2 Hilbert Systems . . . . .	27
1.2.1 Connectives . . . . .	27
1.2.2 Conjunction . . . . .	29
1.2.3 Disjunction . . . . .	29
1.2.4 Implication . . . . .	30
1.2.5 Misc . . . . .	30
<b>Chapter 2: Predicate Calculus</b>	33
2.1 Quantifiers . . . . .	33
2.1.1 Negating Quantifiers . . . . .	34
<b>Chapter 3: Zermelo-Fraenkel Set Theory</b>	35
3.1 The Axioms of Zermelo and Fraenkel . . . . .	35
3.1.1 Subsets and Equality . . . . .	38
3.1.2 Ordered Pairs and Unions . . . . .	46
3.1.3 The Axiom of the Power Set . . . . .	60
3.1.4 Cartesian Products and Functions . . . . .	63
3.1.5 The Axiom of Choice and Diaconescu's Theorem . . . . .	77
3.2 The Structure of Sets . . . . .	80

3.2.1	Basic Theorems . . . . .	80
3.2.2	Operations on Sets . . . . .	84
3.3	Relations . . . . .	101
<b>Chapter 4:</b>	<b>The Real Numbers</b>	<b>113</b>
<b>Chapter 5:</b>	<b>Function Theory</b>	<b>115</b>
5.1	Basic Definitions and Theorems . . . . .	115
5.1.1	Surjections . . . . .	122
5.1.2	Injections . . . . .	124
5.1.3	Bijections . . . . .	125
5.1.4	Compositions . . . . .	126
5.2	Binary Operations . . . . .	131
5.3	Boolean Algebras . . . . .	144
5.4	Sequences and Matrices . . . . .	150
<b>Chapter 6:</b>	<b>Arithmetic</b>	<b>151</b>
<b>Chapter 7:</b>	<b>Cardinality</b>	<b>153</b>
<b>Part II:</b>	<b>Lattice Theory</b>	<b>155</b>
<b>Chapter 8:</b>	<b>Relations of Order</b>	<b>157</b>
<b>Chapter 9:</b>	<b>Graph Theory</b>	<b>159</b>
<b>Part III:</b>	<b>Categories and Models</b>	<b>161</b>
<b>Chapter 10:</b>	<b>Category Theory</b>	<b>163</b>

## Book Two: Algebra

<b>Part IV: Group Theory</b>	<b>165</b>
<b>Chapter 11: Elementary Group Theory</b>	<b>167</b>
11.1 Group-Like Structures . . . . .	167
11.1.1 Semigroups and Monoids . . . . .	168
11.1.2 Subsemigroups . . . . .	178
11.2 Groups . . . . .	183
11.2.1 Direct Product . . . . .	202
11.2.2 Subgroups . . . . .	202
11.2.3 Cayley Diagrams . . . . .	208
11.2.4 Comments from Meeting with David . . . . .	211
11.3 Group Morphisms . . . . .	212
11.3.1 Homomorphisms . . . . .	212
<b>Chapter 12: Finite Groups</b>	<b>215</b>
12.1 Permutation Groups . . . . .	215
12.1.1 Finite Permutations . . . . .	217
12.2 Dihedral Groups . . . . .	220
12.3 Polyhedra Groups . . . . .	222
<b>Part V: Ring Theory</b>	<b>223</b>
<b>Chapter 13: Rings</b>	<b>225</b>
13.1 Definitions . . . . .	225
13.1.1 Rngs . . . . .	225
13.1.2 Rings . . . . .	227
13.2 Ring Morphisms . . . . .	228
<b>Chapter 14: Fields</b>	<b>231</b>
14.1 Definitions . . . . .	231
<b>Part VI: Modules</b>	<b>237</b>
<b>Chapter 15: Elementary Module Theory</b>	<b>239</b>
15.1 Definitions . . . . .	239
15.2 Old . . . . .	250
<b>Part VII: Galois Theory</b>	<b>253</b>
<b>Chapter 16: Elementary Field Theory</b>	<b>255</b>
16.1 What's the Point? . . . . .	255
16.1.1 Cubic Equations and Higher . . . . .	258
16.1.2 Some Reminders . . . . .	258
16.1.3 Polynomials . . . . .	259
16.1.4 Review of Previous Lecture . . . . .	261
16.2 Stuff . . . . .	261
16.3 Rings . . . . .	265
16.4 Ideals . . . . .	272
16.5 Polynomial Rings . . . . .	279
16.5.1 Roots and Irreducibility . . . . .	295

16.5.2	Reduction Modulo a Prime . . . . .	297
16.5.3	Review of Previous Lecture . . . . .	298
16.6	Gauss's Lemma . . . . .	299
16.7	Field Extensions . . . . .	300
16.8	Minimal Polynomial . . . . .	301
16.9	Review of Previous Lectures . . . . .	302
16.10	Compass and Straight Edge . . . . .	303
16.11	Review . . . . .	305
16.12	Splitting Fields . . . . .	306
16.12.1	Review from Previous Lecture . . . . .	306
16.13	Separability . . . . .	307
16.14	Subfields and Automorphisms . . . . .	310
16.15	More on Symmetric Polynomials . . . . .	311
16.16	Normal Extension . . . . .	313
16.17	Radical Stuff . . . . .	315
16.18	Quartic Polynomials . . . . .	315
16.19	Stuff . . . . .	316
16.20	Orthogonal Transformations . . . . .	317
16.21	Sesquilinear Geometry . . . . .	318
16.22	Unitary Transformations . . . . .	320
16.23	Spectral Theorem . . . . .	320
16.24	Tensor Products . . . . .	322
16.25	Bonus Stuff . . . . .	324
<b>Part VIII:</b>	<b>Unsorted Stuff</b>	<b>325</b>
<b>Chapter 17:</b>	<b>Combinatorics</b>	<b>327</b>
17.1	Introduction . . . . .	327
17.2	Counting Techniques . . . . .	328
17.2.1	Basic Numbers . . . . .	328
17.2.2	Basic Counting Principles . . . . .	329
17.3	Posets . . . . .	330
17.4	Binomial Coefficients and Multi-Sets . . . . .	331
17.4.1	Lattice Paths . . . . .	332
17.4.2	The Involution Principle . . . . .	332
17.4.3	Diagonal Lattice Paths . . . . .	334
17.5	q-Analogues . . . . .	334
17.6	Lecture 6 . . . . .	335
17.6.1	Lattice Paths and Gaussian Polynomials . . . . .	338
17.7	Lecture 7 (I Think) . . . . .	338
17.8	Lecture 9 . . . . .	342
17.8.1	q-Catalan Analogue . . . . .	343
17.9	Generating Functions . . . . .	345
17.10	Euler's Theorem . . . . .	347
17.11	Generating Function for Multisets . . . . .	350

17.12	Symmetric Functions . . . . .	351
17.12.1	Symmetric Polynomials . . . . .	351
17.12.2	Misc Problems . . . . .	353
<b>Chapter 18:</b>	<b>Discrete Structure</b>	<b>355</b>
18.1	Combinatorics . . . . .	355
18.2	Exams . . . . .	355
18.2.1	Exam I . . . . .	358
18.2.2	Practice Exam II . . . . .	359
18.2.3	Exam II . . . . .	361
18.2.4	Practice Exam III . . . . .	363
18.2.5	Exam III . . . . .	367
18.2.6	Final Exam . . . . .	369
18.3	Quizzes . . . . .	370
18.3.1	Quiz I . . . . .	370
18.3.2	Quiz II . . . . .	370
18.3.3	Quiz III . . . . .	371
<b>Chapter 19:</b>	<b>Linear Algebra</b>	<b>373</b>
19.1	Linear Algebra . . . . .	373
19.2	Miscellaneous Lecture Notes . . . . .	378
19.2.1	Orthogonal Projections . . . . .	378
19.2.2	Reflections . . . . .	379
19.2.3	Lecture Notes on Orthogonal Matrices . . . . .	379
19.2.4	Rotations . . . . .	382
19.2.5	The Matrix Exponential . . . . .	384
19.2.6	Linear Systems of Ordinary Differential Equations .	385
19.3	Problem Sets . . . . .	386
19.3.1	Problem Set I . . . . .	386
19.3.2	Problem Set II . . . . .	388
19.3.3	Problem Set III . . . . .	390
19.3.4	Problem Set IV . . . . .	392
19.3.5	Problem Set V . . . . .	393
19.3.6	Problem Set VI . . . . .	396
19.3.7	Problem Set VII . . . . .	398
19.3.8	Problem Set VIII . . . . .	400
<b>Chapter 20:</b>	<b>Algebraic Geometry</b>	<b>403</b>
20.1	Notes on Cox, Little, and O'Shea . . . . .	403
20.1.1	Geometry, Algebra, and Algorithms . . . . .	404
20.1.2	Elimination Theory . . . . .	414
20.1.3	Groebner Bases . . . . .	417
20.1.4	The Algebra-Geometry Dictionary . . . . .	420
20.1.5	Polynomials and Rational Functions on a Variety . .	423
20.2	Miscellaneous Notes . . . . .	425
20.2.1	Groebner Bases . . . . .	425

20.2.2	Elimination Theory . . . . .	427
20.2.3	Étale Cohomology . . . . .	428
20.2.4	The Zariski Topology . . . . .	431
20.2.5	Notes on Varieties . . . . .	434
20.3	Lie Algebras . . . . .	437
20.4	Review of Differentiation . . . . .	444
20.5	Lie Groups . . . . .	445
20.6	Solvability and Semisimplicity . . . . .	446
20.7	Semidirect Products . . . . .	448
20.8	Abstract Jordan-Chevelay Decomposition . . . . .	449
20.9	Representations of $sl_2(\mathbb{C})$ . . . . .	450
20.10	Root Space Decomposition . . . . .	451
20.11	Geometric Properties of Root Space Decomposition . . . . .	453
20.12	Root Systems . . . . .	454

## Book Three: Topology

<b>Part IX: Point-Set Topology</b>	<b>455</b>
<b>Chapter 21: Topological Spaces</b>	<b>457</b>
21.1 Topologies . . . . .	457
21.1.1 Separation Axioms . . . . .	468
21.2 Old Notes . . . . .	472
<b>Chapter 22: The Product Topology</b>	<b>475</b>
22.1 Product Topology . . . . .	475
22.2 Hocking and Young (Chapter 1) . . . . .	484
22.3 A Review of Topology . . . . .	495
<b>Part X: Metric Spaces</b>	<b>503</b>
<b>Part XI: Homotopy</b>	<b>505</b>
<b>Chapter 23: Stuff</b>	<b>507</b>
23.1 Lecture 7-ish, Maybe . . . . .	507
23.2 Van Kampen's Theorem . . . . .	508
23.2.1 Examples . . . . .	509
23.3 Covering Spaces . . . . .	510
23.4 Universal Cover . . . . .	510
23.5 Group Actions . . . . .	512
23.5.1 Lifting Criterion . . . . .	514
23.6 Homology . . . . .	514
23.6.1 History . . . . .	514
23.6.2 Singular Homology . . . . .	515
23.6.3 Simplicial Homology . . . . .	516
23.7 More Homology . . . . .	517
23.8 Mayer-Vietoris Sequence . . . . .	518
23.9 Cohomology . . . . .	519
23.10 Len's Spaces . . . . .	520
23.11 Cohomology Rings . . . . .	521
23.12 Cap Product . . . . .	522
<b>Part XII: Homology</b>	<b>525</b>
<b>Part XIII: Cohomology</b>	<b>527</b>

## Book Four: Analysis

<b>Part XIV: Measure Theory</b>	<b>529</b>
<b>Chapter 24: Infinite Series</b>	<b>531</b>
<b>Chapter 25: Real Analysis</b>	<b>533</b>
25.1 Old stuff . . . . .	542
25.1.1 Continuity . . . . .	542
25.1.2 Sequences of Functions . . . . .	548
25.1.3 Inequalities . . . . .	553
25.2 Notes from Rosenlicht . . . . .	555
25.2.1 Sets . . . . .	555
25.2.2 The Real Number System . . . . .	556
25.3 Old Notes . . . . .	569
25.3.1 Definitions . . . . .	569
25.3.2 Theorems . . . . .	570
25.3.3 Metric Spaces . . . . .	571
25.3.4 The Real Numbers . . . . .	578
25.3.5 Vector Spaces and Euclidean Spaces . . . . .	586
25.4 Definitions and Theorems . . . . .	591
25.4.1 Definitions . . . . .	591
25.4.2 Theorems . . . . .	593
25.5 Cheat Sheet . . . . .	594
25.5.1 Series . . . . .	594
25.5.2 Complex Variables . . . . .	596
25.5.3 Matrices . . . . .	598
25.5.4 Vectors . . . . .	599
<b>Chapter 26: Measurable Spaces</b>	<b>601</b>
26.1 Set Rings . . . . .	601
26.2 Set Algebras . . . . .	604
26.3 $\sigma$ -Rings . . . . .	605
26.4 $\sigma$ -Algebras . . . . .	606
26.4.1 Dynkin System . . . . .	606
26.4.2 Borel $\sigma$ -Algebra . . . . .	607
<b>Chapter 27: Measurable Functions</b>	<b>609</b>
27.1 Definitions and Properties . . . . .	609
27.1.1 Measurable Functions . . . . .	611
27.1.2 Sequences of Measurable Functions . . . . .	613
27.2 Convergence of Measurable Functions . . . . .	613
<b>Chapter 28: Measures</b>	<b>619</b>
28.1 Measures . . . . .	619
28.1.1 A Review Infinite Series . . . . .	619
28.1.2 Measure Functions . . . . .	620
28.1.3 Properties of Measure . . . . .	621
28.2 Lebesgue-Stieltjes Measures . . . . .	623

<b>Chapter 29: Product Measures</b>	<b>627</b>
29.1 Product Measures . . . . .	627
<b>Part XV: Probability Theory</b>	<b>629</b>
29.2 Product Measures . . . . .	631
29.3 Probablity Spaces . . . . .	632
29.4 Random Variables . . . . .	632
29.5 Lecture 8-ish Maybe . . . . .	638
29.5.1 Covariance . . . . .	639
29.6 Laws of Large Numbers . . . . .	641
29.6.1 Borel Numbers . . . . .	648
29.7 Central Limit Theorem . . . . .	650
29.7.1 Convergence of Distributions . . . . .	653
<b>Part XVI: Complex Analysis</b>	<b>657</b>
<b>Chapter 30: Complex Numbers</b>	<b>659</b>
30.1 Complex Numbers . . . . .	659
30.1.1 Polar Representation of Complex Numbers . . . . .	668
30.1.2 Analytic Functions . . . . .	676
30.1.3 Contour Integrals . . . . .	683
30.2 Complex Variables . . . . .	690
<b>Part XVII: Calculus on Normed Spaces</b>	<b>693</b>
<b>Chapter 31: Calculus on Normed Spaces</b>	<b>695</b>
31.1 Gateaux Derivative . . . . .	695
31.2 Frechet Derivative . . . . .	695
31.3 Malliavin Calculus . . . . .	695
<b>Part XVIII: Functional Analysis</b>	<b>697</b>
<b>Chapter 32: Functional Analysis</b>	<b>699</b>
32.1 Metric Spaces . . . . .	699
32.1.1 Basic Definitions . . . . .	699
32.1.2 Topology . . . . .	705
32.1.3 Completeness . . . . .	711
32.1.4 Banach's Fixed Point Theorem . . . . .	714
32.2 Normed and Inner Product Spaces . . . . .	718
32.2.1 Basic Definitions . . . . .	718
32.2.2 Lecture 7: October 22, 2018 . . . . .	722
32.2.3 Lecture 8: October 29, 2018 . . . . .	725
32.2.4 Lecture 9: November 5, 2018 . . . . .	729
32.2.5 Lecture 10: November 19, 2018 . . . . .	733
32.2.6 Lecture 11: November 26, 2018 . . . . .	736
32.3 More Stuffs . . . . .	738
32.3.1 Lecture 12: December 3, 2018 . . . . .	738
32.3.2 Lecture 13: December 10, 2018 . . . . .	744
32.4 Old Notes . . . . .	746

---

32.4.1	Summary of Lectures	746
32.5	Metric Spaces	748
32.5.1	Basic Definitions	748
32.5.2	Completeness	757
32.5.3	Compactness	763
32.5.4	Lebesgue Spaces	767
32.5.5	Equicontinuity	768
32.5.6	Baire Spaces	769
32.6	Normed Vector Spaces	773
32.6.1	Basic Definitions	773
32.6.2	Banach Spaces	773
32.7	Zorn's Lemma	780
32.7.1	Brushing Up on Topology	785
32.8	Stuff	788
32.9	More Normed Vector Space Stuff	791
32.10	Hilbert Spaces	792
32.11	Even More Stuff	796
32.12	Even MORE Stuff!	799
<b>Chapter 33:</b>	<b>Homeworks</b>	<b>805</b>
33.1	Homework I	805
33.2	Homework II	815
<b>Part XIX:</b>	<b>Fourier Analysis</b>	<b>821</b>
<b>Chapter 34:</b>	<b>Fourier Analysis</b>	<b>823</b>
34.0.1	Basic Notions	823
34.0.2	Fourier Series	825
34.0.3	The Fourier Transform	825
34.0.4	Convolutions	829
34.0.5	Sampling	829
<b>Chapter 35:</b>	<b>Chaos Theory</b>	<b>831</b>
35.1	A Review of Differential Equations	831
35.1.1	First Order Equations	831
<b>Part XX:</b>	<b>Special Functions</b>	<b>839</b>
<b>Chapter 36:</b>	<b>Numerical Analysis</b>	<b>841</b>
36.1	Power Series	841
36.2	Asymptotic Expansions	841
36.3	Stationary Phase Approximation	841
36.3.1	Root Finding	843
36.4	Special Functions	845
36.4.1	The Fresnel Integrals	846
36.4.2	Bessel Functions	854
36.4.3	Lambert's W Function	854
36.4.4	Legendre Polynomials	854

## Book Five: Geometry

<b>Part XXI: Manifolds</b>	<b>855</b>
<b>Chapter 37: Euclidean Spaces</b>	<b>857</b>
37.1 Topology . . . . .	857
37.1.1 Basic Definitions . . . . .	857
37.1.2 Homeomorphisms . . . . .	871
37.1.3 Subspace Topology . . . . .	873
37.1.4 Basis for a Topology . . . . .	873
37.1.5 Continuous Maps and Products . . . . .	874
37.1.6 Countability Properties . . . . .	876
37.1.7 Algebraic Topology . . . . .	889
37.2 Homogeneous Spaces . . . . .	892
37.3 Locally Euclidean Spaces . . . . .	904
37.4 Topological Manifolds . . . . .	910
37.5 Reading From Lee (Chapter 1) . . . . .	911
37.5.1 The Topology of Manifolds . . . . .	924
37.5.2 Differentiable Manifolds . . . . .	928
37.6 Smooth Manifolds . . . . .	929
37.6.1 Smooth Mappings . . . . .	934
37.7 Smooth Manifolds with Boundary . . . . .	941
37.8 Diffeomorphisms . . . . .	942
37.9 Partitions of Unity . . . . .	943
37.10 Vector Fields . . . . .	945
37.11 Immersions, Submersions, and Embeddings . . . . .	945
37.12 Notes from O'Neill (Chapter 1) . . . . .	946
37.12.1 Smooth Functions . . . . .	948
37.12.2 Tangent Vectors . . . . .	952
37.12.3 The Differential Pushforward . . . . .	954
37.12.4 Curves . . . . .	956
37.12.5 Vector Fields . . . . .	957
37.12.6 One Forms . . . . .	962
37.12.7 Submanifolds . . . . .	963
37.12.8 Immersions and Submersions . . . . .	965
37.12.9 Partitions of Unity . . . . .	967
37.12.10 Orientability . . . . .	968
37.12.11 Special Manifolds . . . . .	968
37.12.12 Vector Spaces as Manifolds . . . . .	969
37.12.13 The Tangent Bundle . . . . .	969
37.12.14 Integral Curves . . . . .	970
37.13 O'Neill Problems Chapter 7 . . . . .	970
37.14 Manifolds Review . . . . .	971
37.15 Connections . . . . .	977
37.16 Pendulums . . . . .	978

37.17 Spivak Calculus on Manifolds . . . . .	978
37.18 Spivak (Chapter 2) . . . . .	984
37.19 Homework I . . . . .	984
37.19.1 Sequential Spaces . . . . .	984
37.19.2 Various Types of Compactness . . . . .	988
37.19.3 Connectedness . . . . .	993
37.19.4 Topological Manifolds . . . . .	994
37.19.5 Problems: Part A . . . . .	998
37.19.6 Problems: Part B . . . . .	1004
37.20 Homework II . . . . .	1006
37.21 Homework III . . . . .	1010
37.21.1 Weierstrass Approximation Theorem . . . . .	1010
37.21.2 Homework III . . . . .	1011
37.22 Homework 4 . . . . .	1013
37.23 Homework 5 . . . . .	1015
<b>Part XXII: Euclidean Geometry</b>	<b>1019</b>
<b>Chapter 38: Convex Geometry</b>	<b>1021</b>
38.1 Convexity: Part I . . . . .	1021
38.1.1 Convex Sets . . . . .	1021
38.1.2 Convexity in the Euclidean Plane . . . . .	1025
38.2 On Uniform Convergence . . . . .	1035
38.3 On Analyticity . . . . .	1041
38.4 On Infinite Order O.D.E.'s . . . . .	1043
38.5 Other Results . . . . .	1045
38.6 An Almost Group . . . . .	1047
38.7 On Sequences . . . . .	1048
38.8 A Class of Differentiability . . . . .	1050
38.9 Degenerate Fredholm Equations of the First Kind . . . . .	1051
<b>Part XXIII: Riemannian Geometry</b>	<b>1053</b>
<b>Chapter 39: Semi-Riemannian Geometry</b>	<b>1055</b>
39.1 Definitions . . . . .	1055

**Book Six: Geometric Topology**

<b>Part XXIV: Surgery Theory</b>	<b>1057</b>
<b>Chapter 40: Surgery Theory</b>	<b>1059</b>
40.1 Lecture 2: Surgery Structure Sets . . . . .	1059
40.1.1 Lecture 3: Vector Bundles . . . . .	1064
40.1.2 Lecture 4: Principal G-Bundles . . . . .	1069
40.1.3 Lecture 5: The Wall L-Groups . . . . .	1076
40.1.4 Lecture 6: The Brown Representation Theorem . . .	1079
40.1.5 Lecture 1: Singular and Simplicial Homology . . . .	1081
<b>Part XXV: Knot Theory</b>	<b>1085</b>
<b>Chapter 41: Knot Theory</b>	<b>1087</b>

## Book Seven: Physics

<b>Part XXVI: Classical Mechanics</b>	<b>1091</b>
41.1 Notes . . . . .	1093
41.1.1 Old Notes on Lagrangians . . . . .	1093
<b>Part XXVII: Diffraction Theory</b>	<b>1095</b>
<b>Chapter 42: Diffraction Theory</b>	<b>1097</b>
42.1 Maxwell's Equations . . . . .	1097
42.2 Fresnel-Fraunhofer Theory . . . . .	1097
42.3 Fresnel's Approximation . . . . .	1097
42.4 Fresnel Inversion . . . . .	1097
42.4.1 Diffraction Through a Square Well . . . . .	1097
42.4.2 Diffraction Through an Inverted Square Well . . . . .	1099
<b>Chapter 43: Geometry</b>	<b>1101</b>
43.1 Titan Geometry . . . . .	1101
43.2 Ring Geometry . . . . .	1105
43.3 Derivations of the Fresnel Kernel . . . . .	1107
<b>Chapter 44: Occultation Observations</b>	<b>1111</b>
44.1 Diffraction Theory for Occultations . . . . .	1111
44.1.1 Reduction to a Single Integral . . . . .	1111
44.1.2 The Inversion Approximation . . . . .	1112
44.1.3 Window Functions . . . . .	1114
44.2 Problems with Fresnel Inversion . . . . .	1115
44.2.1 Radial Shift from a Linear Phase Offset . . . . .	1115
44.2.2 Notes on the Fresnel Approximation . . . . .	1117
<b>Part XXVIII: Electromagnetism</b>	<b>1121</b>
<b>Chapter 45: Electromagnetism I</b>	<b>1123</b>
45.1 Homework Sets . . . . .	1123
45.1.1 Homework I . . . . .	1123
45.1.2 Homework II . . . . .	1128
45.1.3 Homework III . . . . .	1134
45.1.4 Homework IV . . . . .	1140
45.1.5 Homework V . . . . .	1143
45.1.6 Homework VI . . . . .	1146
45.1.7 Homework VII . . . . .	1150
45.1.8 Homework VIII . . . . .	1153
45.1.9 Homework IX . . . . .	1154
45.1.10 Homework X . . . . .	1155
45.1.11 Homework XI . . . . .	1157
45.1.12 Homework XII . . . . .	1159
45.1.13 Homework XIII . . . . .	1160
45.2 Exams . . . . .	1161
45.2.1 Exam I . . . . .	1161

45.2.2	Exam II . . . . .	1162
45.2.3	Exam III . . . . .	1163
45.2.4	Practice Final Exam . . . . .	1164
45.2.5	Final Exam . . . . .	1165
<b>Chapter 46:</b>	<b>Electromagnetism II</b>	<b>1169</b>
46.1	Homework Sets . . . . .	1169
46.1.1	Homework I . . . . .	1169
46.1.2	Homework II . . . . .	1174
46.1.3	Homework III . . . . .	1178

<b>Acronyms</b>	<b>1181</b>
<b>Notation</b>	<b>1183</b>
<b>Glossary</b>	<b>1185</b>
<b>Bibliography</b>	<b>1197</b>
<b>Index</b>	<b>1201</b>

---

## List of Figures

---

1.1	Sketch of the Intermediate Value Theorem . . . . .	4
1.2	Lines for Galileo’s Paradox . . . . .	16
1.3	Solution to Galileo’s Paradox . . . . .	17
3.1	Visualizing Subsets as Blobs . . . . .	41
3.2	Venn Diagram for Union . . . . .	53
3.3	The Union of Three Sets . . . . .	54
3.4	Venn Diagram for Intersection . . . . .	56
3.5	The Intersection of Three Sets . . . . .	58
3.6	The Cartesian Plane $\mathbb{R}^2$ . . . . .	66
3.7	The Lattice $\mathbb{N}^2$ . . . . .	67
3.8	Cartesian Product of Two Sets . . . . .	68
3.9	Example of a Function $f : \mathbb{R} \rightarrow \mathbb{R}$ . . . . .	70
3.10	Example of a Non-Function . . . . .	71
3.11	Visual for Abstract Functions . . . . .	72
3.12	Non-Functions . . . . .	72
3.13	Image of a Subset and of a Point under a Function . . . . .	75
3.14	Example of a Mapping from Projective Geometry . . . . .	77
3.15	Figure for Thm. 3.2.20 . . . . .	85
3.16	Visual for Thm. 3.2.37. . . . .	89
3.17	Venn Diagram for Distributive Law of Unions . . . . .	93
3.18	Venn Diagram for the Distributive Law of Intersections . . . . .	94
3.19	Figure for Thm. 3.2.54 . . . . .	94
3.20	Venn Diagram for Set Difference . . . . .	96
3.21	Venn Diagram for Symmetric Difference . . . . .	98
3.22	Diagram for a Transitive Relation . . . . .	106
3.23	The Intersection of Transitive and Non-Transitive Relations . . . . .	107
3.24	Commutative Diagram for the Quotient Set . . . . .	111
5.1	A Commutative Diagram . . . . .	126
5.2	A Slightly Complicated Commutative Diagram . . . . .	127

5.3	A Very Complicated Commutative Diagram . . . . .	127
5.4	Associativity of Function Composition . . . . .	128
11.1	The Dihedral Group $D_6$ . . . . .	194
11.2	Restraints on the Dihedral Group $D_6$ . . . . .	195
11.3	Associativity of the Dihedral Group $D_6$ . . . . .	196
11.4	Reflections on a Square . . . . .	199
11.5	Cayley Diagram of $D_6$ . . . . .	209
11.6	Cayley Diagram of $\mathbb{Z}_2 \times \mathbb{Z}_2$ . . . . .	210
11.7	Cayley Diagram for $S_3$ . . . . .	210
11.8	Cayley Diagrams for $D_8$ . . . . .	211
12.1	Cycle Diagram for a Permutation . . . . .	218
16.1	Lagrange Discriminant . . . . .	316
21.1	The Open Interval $(a, b)$ is an Open Subset of $\mathbb{R}$ . . . . .	457
21.2	The Closed Interval $[a, b]$ is Not Open. . . . .	458
21.3	The Intersection of Open Intervals is Open. . . . .	458
21.4	The Sierpinski Topology . . . . .	460
21.5	The Union of Topologies Need Not be Closed to Finite Intersections . . . . .	463
22.1	Examples of Open Subsets of $\mathbb{R}^2$ . . . . .	477
22.2	Tiling of the Open Unit Disc by Rectangles. . . . .	478
22.3	Strips in the Plane. . . . .	483
22.4	Blocks in Space. . . . .	484
22.5	How to construct $L(p, q)$ . . . . .	497
25.1	Dirichlet's Popcorn Function . . . . .	544
30.1	Cartesian Representation of Complex Numbers . . . . .	660
30.2	Modulus and Conjugate of a Complex Number . . . . .	662
30.3	Visual Representation of the Triangle Inequality . . . . .	666
30.4	Polar Representation of a Complex Number . . . . .	674
30.5	Roots of Unity for Degrees 3 to 6. . . . .	676
30.6	A Smooth Function That is Not Analytic at the Origin . . . . .	678
30.7	A Smooth and Nowhere Analytic Function . . . . .	679
30.8	Jordan Curves in the Complex Plane . . . . .	683
30.9	Examples of Simple Regions . . . . .	684
32.1	Figures for Example ???. . . . .	706
34.1	Fourier Transform of the Hat Function. . . . .	827

35.1	Phase line for $\dot{x}(t) = kx(t)$ when $k < 0$	833
36.1	Newton Fractal for $z^3 - 1 = 0$	843
36.2	Newton Fractal for $z^4 - 1 = 0$	844
36.3	Mandelbrot Set	845
36.4	Graph of the Fresnel Cosine Function	847
36.5	Graph of the Fresnel Sine Function	848
36.6	Jordan Curve Used to Evaluate the Fresnel Integrals.	851
37.1	Interior Point of a Set	859
37.2	Interior of a Set	860
37.3	Exterior Point of a Set	865
37.4	Exterior a Set	866
37.5	The Quotient of $\mathbb{S}^2$ by $\mathbb{S}^1$	897
37.6	Construction of the Bug-Eyed Line	901
37.7	Open Subsets of the Bug-Eyed Line	902
37.8	Construction of the Branching Line	903
37.9	Homomorphism from the Square to the Circle	912
37.10	A Chart in a Topological Space	915
37.11	Orthographic Chart for the Sphere	917
37.12	Near-Sided Projection from Geosynchronous Orbit	918
37.13	Far-Sided Projection from Geosynchronous Orbit	919
37.14	Stereographic Projection of the Sphere	920
37.15	Smoothly Overlapping Charts	930
37.16	Smooth Real-Valued Function on a Manifold	935
37.17	Smooth Function Between Manifolds	949
37.18	Creating a Manifold Structure on an Arbitrary Set	951
37.19	The Pushforward of a Tangent Vector	955
38.1	Drawing for Thm. 38.1.28.	1026
38.2	Drawing for Thm. 38.1.29.	1027
38.3	Drawing for Thm. 38.1.30.	1028
38.4	Caption	1030
38.6	Triangle	1034
40.1	Example of a Commutative Diagram.	1059
40.2	Simple Surgery Example.	1060
40.3	Example of a Zero Surgery.	1061
40.4	More Complicated Surgery Example.	1062
40.5	Partition of $S^{Cat}(\mathcal{M})$ .	1064
40.6	Example of a Vector Bundle: $(D^1, D^1 \times \mathbb{R}, p)$ .	1066
40.7	Möbius Strip.	1067
40.8	$\mathbb{R}$ is the Universal Cover of $S^1$ .	1068

---

40.9 Examples of Simplices.	1070
40.10 Diagram for the Surgery Exact Sequence of $S^5$ .	1070
40.12 Turning a Vector Bundle into a Sphere Bundle.	1080
40.13 Diagrams for the Lifting Property.	1081
41.1 A Trefoil Knot	1087
41.2 A Colorful Trefoil Knot	1088
41.3 Tricoloring of the Trefoil	1088
41.4 Mobius Strip	1088
41.5 A Sea Shell	1089
41.6 Seifert Surface for a Trefoil Knot	1089
41.7 Seifert Surface for a Hopf Link	1090
41.8 Non-Oriented Surface With Trefoil Boundary	1090
43.1 Various Geometries for Titan	1101
45.1 Figures for Wangsness 1-11 and 1-12	1129
45.2 Figures for Wangsness 1-13 and 1-14	1132
45.3 Figure for Wangsness 1-15	1134
45.4 Figures for Wangsness 1-22 and 1-23	1138
45.5 Figures for problems 45.1.19 and 45.1.21	1141
45.6 Drawing for Wangsness 4-3	1144
45.7 Drawing for Wangsness 4-3	1144
45.8 Drawings for problems 45.1.31 and 45.1.32	1148
45.9 Circuits for problem 45.1.37	1151
45.10 Infinite Cylinders for problem 45.1.39	1152
45.11 Drawing for Wangsness 8-5	1154
45.14 Drawing for Wangsness 13-4	1158
45.15 Drawing for Wangsness 14-15	1159
45.16 Drawing for Wangsness 15-7	1159
45.17 Drawing for Wangsness 17-4	1161
46.1 Figures for Problem 46.1.6.	1175
46.2 Solution to Problem 46.1.7.	1176
46.3 Figure for Problem 46.1.8.	1177
46.4 Diagram for Problem 46.1.10.	1179

---

## List of Tables

---

1.1	Gauss' Sum of 1 to 100 . . . . .	5
1.2	Grandi's Series $1 - 1 + 1 - 1 + \dots$ . . . . .	6
1.3	The Sum of $1 - 2 + 3 - 4 + \dots$ . . . . .	6
1.4	The Sum of $1 + 2 + 3 + 4 + \dots$ . . . . .	6
1.5	The Sum of $1 + 1 + 1 + 1 + \dots$ . . . . .	7
1.6	Truth Table for Implication . . . . .	20
1.7	Alternate Table for Implication . . . . .	20
1.8	Truth Table for Negation . . . . .	23
1.9	Truth Table for the Converse . . . . .	24
1.10	Truth Table for the Contrapositive . . . . .	24
1.11	Truth Table for the Inverse . . . . .	25
1.12	Fallacy of the Undistributed Middle . . . . .	26
1.13	Truth Table for Conjunction . . . . .	29
1.14	Truth Table for Disjunction . . . . .	30
1.15	Truth Table for Equivalence . . . . .	30
1.16	Truth Table for $(p \vee \neg q) \wedge r$ . . . . .	30
5.1	Simple Binary Operation on $\mathbb{Z}_2$ . . . . .	132
11.1	Cayley Table for $\mathbb{Z}_2$ . . . . .	174
11.2	The Group Structure of $\mathbb{Z}_4$ . . . . .	174
11.3	Example of a Latin Square . . . . .	175
11.4	Cayley Table for the Hyperbolic Quaternion Quasigroup . . . . .	176
11.5	Cayley Table of $D_6$ . . . . .	196
11.6	Cayley Table for Reflection on a Square . . . . .	198
11.7	Cayley Table of Group Formed on Two Coins . . . . .	200
14.1	The Arithmetic of $\mathbb{F}_2$ . . . . .	232
16.1	Monic Irreducible Quadratics in $\mathbb{Z}/3\mathbb{Z}$ . . . . .	263
16.2	Monic Irreducible Cubics in $\mathbb{Z}/3\mathbb{Z}$ . . . . .	264

16.3	Classifying Quartics . . . . .	316
17.1	Caption . . . . .	339
18.1	Truth Table for Problem 18.2.10 . . . . .	359
18.2	Truth Table for Problem 18.2.18 . . . . .	361
18.3	Truth Table for Problem 18.2.21 . . . . .	362
18.4	The Minsets of $A$ , $B$ , and $C$ . . . . .	364
18.5	Minsets of $A$ and $B$ . . . . .	364
25.1	Addition in $\mathbb{F}_2$ . . . . .	559
25.2	Multiplication in $\mathbb{F}_2$ . . . . .	559
25.3	Addition in $\mathbb{F}_3$ . . . . .	560
25.4	Multiplication in $\mathbb{F}_3$ . . . . .	560
46.1	Quadrupole Moments for Problem 46.1.1. . . . .	1170
46.2	Quadrupole Moment for Problem 46.1. . . . .	1170
46.3	Caption . . . . .	1171

---

## Preface

---

This work contains mathematics and physics. It starts from scratch and develops, in a rather lengthy fashion, all of the mathematics that I have come across and decided to write down. An attempt was made (but almost certainly failed) to mimic the style of Euclid's text *The Elements*, in which he proclaims a few postulates and definitions, and then proceeds from there in developing over 400 propositions and theorems in a logical order. This is not a complete mimicry since there are discussions, examples, and many figures to give intuition whereas the elements is simply theorem-proof, with figures only drawn to show a construction described in a proof. Unlike most textbooks there are no exercises, but rather an abundance of worked out examples and an attempt was made to prove every claim in a logical and consistent order. The goal is to justify every step by a definition, axiom, or previously proved theorem. As such, there are no logical prerequisites to read the theorems and proofs, but the examples that are used to build intuition often presume a belief in the existence of real numbers (in particular, the non-negative integers and rational numbers), and some of the motivating examples also use calculus and the elementary algebra of a polynomial in one real variable. Both of these concepts are, eventually, formally developed, but for pedagogical reasons many examples use these notions beforehand. Theorems and proofs do not rely on examples, and in this sense there are no prerequisites. A reader lacking a background in more rigorous mathematics may fail to see the point in laboriously developing logic and set theory, and might not find any motivation for certain definitions, but nonetheless should be able to follow along the proofs in the order presented.

This is not a textbook (or collection thereof) in the usual sense in that, as mentioned previously, all claims are worked out in full. There are no steps that are *left as an exercise to the reader*. This can still be used as a textbook if the reader simply reads the claim of a theorem and tries to prove it first before reading onwards. Since it is all too tempting to allow ones eyes to wander all of 2 inches to the solution, many excellent textbooks for various topics are cited in the bibliography. Thus this work can be seen as a supplement to these,

or as a standalone. The existence of such a work is to give those eager to see mathematics presented in a single logical order a source to work with. Knowing the troubles of Gödel's Incompleteness Theorems (Discussed in Book One), this is merely an attempt at doing so. The advanced mathematician will find that having all of the details spelled out for them to be superfluous, and the beginner will not have the time to read such a large volume, nonetheless I feel such a text should *exist*.

The first book deals with logic and set theory and is perhaps the most contentious. Although I've tried to find consensus about what the definitions of various primitive notions, such as *set*, *proposition*, *predicate*, *truth*, etc., there seems to be no such universally agreed upon definitions and many arguments started to feel circular. Thus an intuitive approach was taken, defining various things in somewhat of a dictionary style. This may appall the logician, and corrections and advice are more than welcome, but for most of mathematics this seems to work well.

This project is very much a work in progress and will remain so for many years, do in part to the sheer scope of the project. Any and all suggestions, corrections, and improvements are welcome and the source code is hosted on GitHub<sup>1</sup> under the GNU GPL 3 license. My only wish is that this material is not *stolen* in the sense that one claims the work their own, but all of the code is freely available and may be used by anyone. This includes all of the tikz code for reproducing figures. Figures produced via the use of the C programming language are compatible with the C99 standard, and the asymptote code is not too innovative either.<sup>2</sup>

Ryan Maguire,  
Lowell, MA

---

<sup>1</sup> <https://github.com/ryanmaguire/Mathematics-and-Physics>

<sup>2</sup> <https://github.com/ryanmaguire/Mathematics-and-Physics/tree/master/tikz>

---

## Acknowledgements

---

I'd like to thank the many people who have helped me better understand mathematics and physics over the years. This includes Stanley Chang, Enrique Gonzalez-Valesco, Tim Cook, James Egan, Richard French, Sam Gomez, Sam Fingerman, Alexander "Sasha" Kheifets, Tibor Beke, Justin Jozokos, and Mike. A particular thanks must go out to Dan Klain who was both infinitely patient and helpful throughout the years, and, of course, to James "Kiwi" Graham-Eagle who was both a mentor and friend, and whom first showed me how amazing mathematics truly is. Personal thanks go to Molly, Kyle, Asa, Kaileigh, and Nayeeb for their years (and decades) of friendship.

# Book One

# Foundations



**Part I**

# **Logic and Sets**



# CHAPTER 1

---

## Propositional Logic

---

We'll begin our discussion of logic with the most primitive kind: propositional.<sup>1</sup> This deals with the structure of sentences, how the English language is used to formulate arguments and deduce new facts. It is for this reason one should start with the foundations of logic for at the heart of mathematics is the concept of *definition-theorem-proof*. The word definition, it is hoped, is understood as a primitive of the English language (much like the word *the*) but theorem and proof require an explanation if we are to use them consistently. We'll discuss propositions, predicates, implications, negations, and overall what we are to consider valid reasoning. We return to logic later in Books [One](#) and [Three](#) when discussing Boolean algebras and Stone spaces, culminating in Stone's representation theorem.

### 1.1 What is Logic?

It may seem strange to begin a study of mathematics with the development of logic as one might think such conversations should reside in philosophy. Indeed, most of classical logic was developed by philosophers rather than mathematicians. Many problems, which we will discuss in Chapt. 3, arose in the early 1900s with the very core of mathematics. Arguments once considered sound were shattered and contradictions were discovered. On the other hand other methods of proof that are very intuitive were shown to be able to prove the existence of non-intuitive and almost impossible objects. This motivates us to develop the *axioms* of logic and explore what valid arguments should look like.

**Example 1.1.1** A student of calculus has likely heard of the intermediate value theorem. Fear not those who haven't, we shall draw pictures. Given a

---

<sup>1</sup> Also called *sentential logic*.

*continuous*<sup>2</sup> function  $f$  of real numbers, if 0 evaluates to a negative number and 1 to a positive, then there is some point in the middle which evaluates to zero. The proof is quite simple: We first look at what happens to the point  $\frac{1}{2}$ . If  $f$  is zero here we are done, otherwise if  $f$  is positive then we may suspect there's a value in between 0 and  $\frac{1}{2}$  which evaluates to zero and if  $f$  is negative at this point, then there's probably a zero between  $\frac{1}{2}$  and 1. In either case we divide the range of possibilities in half and see what happens at  $\frac{1}{4}$  in the first case and  $\frac{3}{4}$  in the latter. We continue *inductively* (whatever this means) and obtain a *sequence* of real numbers which we then show *converges*. Invoking continuity,  $f$  then evaluates to zero at this limit and we are done (see Fig. 1.1).

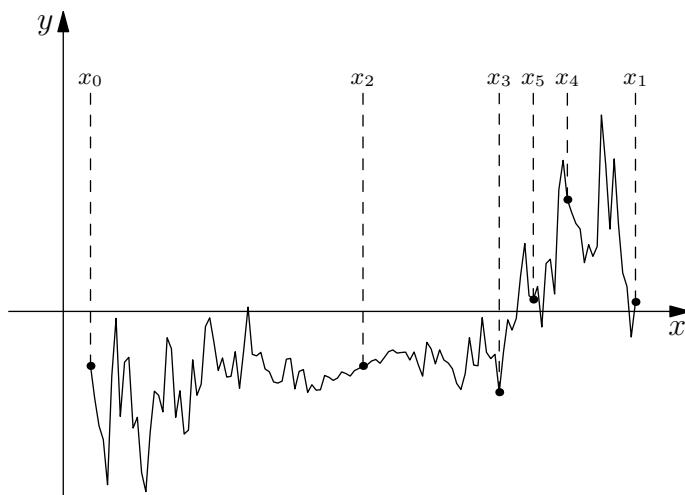


Fig. 1.1: Sketch of the Intermediate Value Theorem

We can see why this may work. After a few iterations we've narrowed down a zero point to a small range between  $x_3$  and  $x_5$  and this is a nice algorithm we can tell a computer to execute to arbitrary precision<sup>3</sup> but what went into the proof? If we are to phrase this with absolute rigor, what definitions, assumptions, and previous theorems are we relying on? For starters, the existence of *real numbers*, a notion of *continuity*, and the definition of a *sequence*. Our exposition of logic is to make clear what is required for valid proofs.

**Example 1.1.2** It has been alleged, though as always one should exhibit skepticism, that in 1784 at the age of seven the great mathematician Carl Friedrich Gauss (1777-1855 C.E.) demonstrated that  $1 + 2 + \dots + 99 + 100 = 5050$

<sup>2</sup> Intuitively a *curve* one can draw from left to right without lifting up their pencil.

<sup>3</sup> This is known as the *bisection* method of root finding.

[vW56, p. 12-13]. It has also been claimed he had revelations about the normal distribution while counting the number of steps on his way to school, and that he corrected his fathers mathematical calculations at the age of three.<sup>4</sup> Alas, mathematicians are a few tales away from forming the religion of Gauss. Nevertheless, let's see what the argument is. We rearrange this sum as follows:

$$\begin{array}{cccccccccc}
 & 1 & + & 2 & + & \cdots & + & 99 & + & 100 \\
 \hline
 = & 1 & + & 2 & + & \cdots & + & 49 & + & 50 \\
 & + & 100 & + & 99 & + & \cdots & + & 52 & + & 51 \\
 \hline
 = & 101 & + & 101 & + & \cdots & + & 101 & + & 101 \\
 \hline
 = & 50 & \times & 101 \\
 \hline
 = & 5050
 \end{array}$$

Table 1.1: Gauss' Sum of 1 to 100

While this seems to be a concrete proof of the claim, is it valid? Can we generalize it? What assumptions about integers and arithmetic are we making?

**Example 1.1.3** Gauss is often called the prince of Mathematics, and the king is the great Leonhard Euler (1707-1783 C.E.). He too studied sums and considered the bizarre series  $1 + 2 + 3 + 4 + \cdots$  arriving at the answer  $-\frac{1}{12}$ .<sup>5</sup> Anyways, let's see how we can arrive at this answer. First, we consider Grandi's series which was studied by Luigi Guido Grandi (1671-1742 C.E.). We have:

$$G = 1 - 1 + 1 - 1 + 1 - 1 + \cdots \quad (1.1.1)$$

Is there a meaningful number to assign to  $G$ ? Suppose there is and write:

<sup>4</sup> Many such claims were made by Gauss' biographer Wolfgang Sartorius von Walterhausen, born 30 years after Gauss, publishing these stories in *Gauss: zum Gedächtnis*. Gauss apparently told some of these stories to von Waltershausen at old age in great excitement. It would be great to accept the tale as a feel-good story, but shortly after János Bolyai (1802-1860 C.E.) published his account of absolute geometry in 1831 creating a common foundation for Euclidean and hyperbolic geometry, Gauss wrote that he was unable to give praise since he had developed these results 30 years prior but never published. With such comments it is hard to tell what is factual about Gauss.

<sup>5</sup> While this sum was studied in the 18<sup>th</sup> century, there seems to be ambiguity as to whether or not Euler earns the credit on this one. He writes this sum and others, such as  $1 + 2 + 4 + 8 + \cdots$  which he claims sums to  $-1$  and  $1 + 3 + 9 + 27 + \cdots$  arriving at  $-\frac{1}{2}$ , in his text *De Seriebus Divergentibus* (English: *On Divergent Series*). He also computes Grandi's summation using the geometric series [Eul60, p. 206-208]. He does not sum  $1 + 2 + 3 + \cdots$  to  $-\frac{1}{12}$  like we do, only mentions it is probably negative.

$$\begin{array}{rcl}
 1 - G & = & 1 - \left[ 1 - 1 + 1 - \dots \right] \\
 \\ 
 & = & 1 - 1 + 1 - 1 + \dots \\
 \\ 
 & = & G
 \end{array}$$

Table 1.2: Grandi's Series  $1 - 1 + 1 - 1 + \dots$ 

And hence we have  $1 - G = G$ . Rearranging we get  $2G = 1$  and therefore  $G = \frac{1}{2}$ . Next, we consider the series:

$$T = 1 - 2 + 3 - 4 + 5 - 6 + \dots \quad (1.1.2)$$

Again, suppose  $T$  has a legal value and apply the following manipulations:

$$\begin{array}{rcl}
 2T & = & 1 - 2 + 3 - 4 + \dots \\
 & & + 1 - 2 + 3 - \dots \\
 \\ 
 & = & 1 - 1 + 1 - 1 + \dots \\
 \\ 
 & = & \frac{1}{2}
 \end{array}$$

Table 1.3: The Sum of  $1 - 2 + 3 - 4 + \dots$ 

And hence  $T = \frac{1}{4}$ . We now return to Euler's sum which we'll denote  $S$ . We subtract the series  $T$  which we computed above, obtaining:

$$\begin{array}{rcl}
 S - T & = & 1 + 2 + 3 + 4 \dots \\
 & - & \left[ 1 - 2 + 3 - 4 \dots \right] \\
 \\ 
 & = & 0 + 4 + 0 + 8 \dots \\
 \\ 
 & = & 4 \left[ 1 + 2 + 3 + 4 \dots \right] \\
 \\ 
 & = & 4S
 \end{array}$$

Table 1.4: The Sum of  $1 + 2 + 3 + 4 + \dots$ 

So  $S - T = 4S$ . But  $T = \frac{1}{4}$  and hence  $3S = -\frac{1}{4}$ . Thus  $S = -\frac{1}{12}$ .<sup>6</sup>

---

<sup>6</sup> The method of summation presented here is an augmentation of Srinivasa Ramanujan's (1887-1920 C.E.) [Ram57, Chapt. VIII p. 3].

We now contemplate the previous examples and ask which are valid. It is tempting to say that Ex. 1.1.1 and Ex. 1.1.2 were presented with accurate proofs, whereas Ex. 1.1.3 is garbage, but why? We “proved” Gauss’ and Euler’s sum in the same manner: Rearranged terms, used *dot dot dot* notation to indicate some pattern, added things together in a convincing way, and then simplified. One might suggest that Gauss’ proof involved a finite scheme and that if asked one could laboriously write out the numbers indicated by the dots, whereas Euler’s sum is infinite. But if we were to perform Gauss’ problem with a number so large that it would take longer than the age of the universe to write down, even with the aid of computer, would we reject his method of proof then? There is some solace, and we can show that the Euler sum is invalid if we accept that  $1 \neq 0$ . Consider the sum  $1 + 1 + 1 + 1 + \dots$  which we shall denote  $B$  for bad. Using Grandi’s series we subtract and obtain:

$$\begin{aligned}
 B - G &= \frac{1 + 1 + 1 + 1 + \dots}{- [1 - 1 + 1 - 1 \dots]} \\
 &= \frac{0 + 2 + 0 + 2 \dots}{=} \\
 &= \frac{2[1 + 1 + 1 + 1 \dots]}{=} \\
 &= 2B
 \end{aligned}$$

Table 1.5: The Sum of  $1 + 1 + 1 + 1 + \dots$

And so we obtain  $B - G = 2B$ , so  $B = -G$ . But we know Grandi’s series is  $G = \frac{1}{2}$ , and hence  $B = -\frac{1}{2}$ . We can also do the following:

$$1 + B = 1 + (1 + 1 + 1 + 1 + \dots) = 1 + 1 + 1 + 1 + \dots = B \quad (1.1.3)$$

And hence  $1 + B = B$ , so  $1 = 0$  which is a contradiction. For the rest of the chapter we wish to hone in on what we are to consider as valid mathematics.

### 1.1.1 Truth

Since the aim of mathematics is to prove the validity of mathematical statements, we should start with a definition of truth. We run into a wall instantly since this is essentially an impossible task. Any definition will be circular, and it is a theorem of Alfred Tarski (1901-1983 C.E.) that if one has defined arithmetic, then one cannot use arithmetic to define truth.<sup>7</sup> That is, if we take upon the assumption of the existence of the natural numbers  $0, 1, 2, \dots$  with

---

<sup>7</sup> This is known as Tarski’s Undefinability Theorem.

the familiar notion of addition<sup>8</sup> then *arithmetic* truth cannot be defined using this arithmetic. Let us consult the dictionary. The Oxford English dictionary defines truth to mean *in accordance with fact or reality* [OED20], Merriam-Webster states that truth is *the body of real things, events, and facts* [Mer20], and Cambridge claims it is *the quality of being true* [Cam20]. As hypothesized these definitions are circular and rely on other predefined terms. We propose the following work around: Truth is a primitive notion that needs no definition. We can then define false to mean *not true*.

Tarski's result came about in the 1930's when he tried to mathematically work out the *liar's paradox* [Tar36]. Consider the following sentence:

$$\text{This sentence is false.} \quad (1.1.4)$$

Similar statements have been considered throughout the ages, including the variant known as Epimenides' paradox. Epimenides of Cnossos (*c.* 600 B.C.E.), who was from Crete, proclaims *Cretans are always liars* [Kin11]. The question was considered again 200 years later when Eubulides of Miletus (*c.* 400 B.C.E.) considered the sentence *I am lying*. Further still in the Book of Psalms king David says *I said in my haste, all men are liars* [Kin11]. Needless to say, the paradox is quite old and well studied. Now we ask, is the statement *true* or *false* (assuming such notions are defined)? Let's work through it and suppose truth. If this sentence is false is true, then the sentence is false even though we just claimed it to be true. Hence, it must be false. But if this sentence is false is false, then the sentence is true, but we just showed it cannot be true. So, which one is it? There are two interpretations: The statement is *neither* true nor false, and the sentence is *both* true and false. Suppose we accept that the statement is neither true nor false. This leads to another sentence where we cannot make such a conclusion:

$$\text{This sentence is not true.} \quad (1.1.5)$$

If this is neither true nor false, then it is not true, and hence true, bringing us back to the paradox. Now we claim it is both true and false, leading us to:

$$\text{The sentence is false and not true.} \quad (1.1.6)$$

The problem intensifies if we consider pairs of sentences:

$$\text{Statement 1.1.7b is true.} \quad (1.1.7a) \quad \text{Statement 1.1.7a is false.} \quad (1.1.7b)$$

and now we go round and round in an endless circle. As Alfred Tarski pointed out, the problem arises in languages in which statements are allowed to be

---

<sup>8</sup> i.e.  $1 + 1 = 2$  and other mathematical gems

self-referential. To see this is indeed a self referencing claim we write  $P$  for the proposition and arrive at the equation:

$$P = P \text{ is false} \quad (1.1.8)$$

if we substitute  $P$ , we obtain:

$$P = (P \text{ is false}) \text{ is false} = ((P \text{ is false}) \text{ is false}) \text{ is false} \quad (1.1.9)$$

While it may seem like this is an unnecessary discussion, the liar's paradox plays a role in mathematics. For one it motivates Tarski's theorem on the defineability of truth, and perhaps more famously it allowed Kurt Gödel (1906-1978 C.E.) to prove his *incompleteness theorems*, which really shook most of modern mathematics. Indeed, this theorem allegedly made Albert Einstein believe there could be no *theory of everything*.<sup>9</sup> <sup>10</sup> For the sake of moving on to mathematics we accept truth to be a primitive notion and acknowledge that the foundations of this concept are very shaky.

Let us examine a few more paradoxes of the English language. The main paradoxes of set theory (Russell's, Cantor's, and Burali-Forti's) will have to wait until we've developed more vocabulary. The first to discuss is *Berry's paradox*, named after G. G. Berry (1867-1928 C.E.), a junior librarian at one of Oxford's library who relayed the paradox to Bertrand Russell [Gri03, p. 63], who published it in 1906.<sup>11</sup> The paradox arises from the following sentence:

*The smallest positive integer not defineable in less than 60 letters*

This statement is itself only 57 characters long. The English language has 26 letters, a space bar, and 10 numerical symbols (in addition to grammatical symbols like commas), so let's suppose there are 50 distinct characters allowed in a sentence. There are then a total of  $50^{60} \approx 8 \times 10^{101}$  combinations. Just about all of these sentences are absolute gibberish, for example:

*qjasneofiq923m woasmd fd/'?maojs 3m ansdjf aia sdf iquer sj*

One of my more poetic works, called *bashing my keyboard and counting to 60*. Some of these combinations of characters do indeed correspond to integers. For example, *The smallest positive integer* corresponds to 1. In just about all of the systems of arithmetic that are studied there exists a *well-ordering* property of the integers. If you are given a collection of positive integers and told there's at least some number in this collection, then there is a *smallest* one. The proof is quite simple, ask yourself *is 1 in the collection?* If yes, you are done since 1

<sup>9</sup> A theory of physics that could solve all problems great and small.

<sup>10</sup> While no direct quotation from Einstein could be found, Stephen Hawking wrote that Gödel's theorem convinced him no such theory can exist [Haw02].

<sup>11</sup> Berry's original paradox dealt with Cantor's theory of *ordinal* numbers, not integers.

is the smallest positive integer, if not proceed. Then ask *is 2 in the collection?* Again, if yes then you are done since 2 is the smallest positive integer greater than 1, but you've already checked that 1 is not in your collection. You continue until finally you hit an integer where the answer is *yes*. You are guaranteed this process will stop since by hypothesis the collection has some integer  $n$ , and hence you need at most  $n$  iterations of this procedure.

In mathematics we describe collections of objects with sentences. This need not be with English, but the problem will exist in just about any human language. For example, *the collection of all integers which are divisible by two* is a description of the *even* integers. So now we propose *the set of all integers not defineable in less than 60 letters*. Since there are at most  $\approx 10^{102}$  such integers that are defineable in less than 60 characters, and since most presume there are infinitely many numbers, we conclude the collection of integers defined by this sentence is non-empty. Then by the well ordering principle there is a least such element. But then this least element satisfies the criterion *The smallest positive integer not defineable in less than 60 letters*, which is less than 60 letters. But we've described it in fewer than 60 characters, a contradiction.<sup>12 13</sup>

The resolution has already been alluded to. There is an ambiguity with the word *defineable*. Does this sentence indeed define an integer? It seems the paradox merely gives a proof that it does not. One proposed alternative to this solution is to create a hierarchy. Hence *The smallest positive integer that is not defineable<sub>0</sub> in less than 60 letters* is a number that is *defineable<sub>1</sub>* in less than 60 letters. Like the liar's paradox, Berry's has a role in mathematics. In 1989 the American mathematician George Boolos<sup>14</sup> used the paradox to prove Gödel's incompleteness theorem in a different manner.

Next is the *Grelling-Nelson Paradox*. This is less mathematical than the previous one but has a familiar ring to it. It is named after the German logicians Kurt Grelling (1886-1943 C.E.) and Leonard Nelson (1882-1927 C.E.).<sup>15</sup> Label an adjective of the English language as *autological* if whatever the adjective is describing also holds for the adjective itself. For example, *polysyllabic* describes words many with syllables, of which *polysyllabic* is such a word. Even more creative, *pentasyllabic* words have five syllables, which *pentasyllabic* happens to have. The word *word* is autological, as is the word *English* (so long as it's

<sup>12</sup> Some number theorists use this paradox to prove all integers are interesting. If not, then there is a least such integer that is not interesting. But being the smallest boring integer is pretty interesting! Hence, all integers are interesting.

<sup>13</sup> There is a semantical equivalent more familiar to most (though almost all are unbothered by it). In the classic Disney movie *Aladdin*, whilst singing the song *A Whole New World*, princess Jasmine proclaims *indescribable feeling*, and yet she just described it.

<sup>14</sup> Note to be confused with George Boole.

<sup>15</sup> Nelson's great grandfather is the mathematician Johann Peter Gustav Lejeune Dirichlet.

read in English). Label an adjective *heterological* otherwise. The words *monosyllabic* and *long* are heterological. Now we consider the word heterological. Since this is an adjective it is valid to ask if it is autological or heterological. If we suppose it is heterological, then it describes itself and is thus autological, a contradiction. If it is autological then it describes itself, but heterological words do not describe themselves, a contradiction.

The game being played here is similar to the circularity of the liar's paradox. Since this problem is purely semantical, we move on to *Curry's paradox*.<sup>16</sup> Like the liar's paradox, this problem has mathematical use and is often seen as a simplification of the Kleene-Rosser paradox which was used to prove certain formal systems are inconsistent.<sup>17</sup> The problem is named after the American mathematician Haskell Curry<sup>18</sup> (1900-1982 C.E.).<sup>19</sup> Let  $P$  be any sentence, and let  $Q$  be the sentence *if this sentence is true, then  $P$  is true*. If  $Q$  is true, then *if this sentence is true, then  $P$  is true* is a true statement. Hence,  $P$  is true. From this *anything* can be proven. Like the liar's paradox the problem is in the allowance of self-referencing sentences.  $Q$  can be written as the statement *if  $Q$ , then  $P$* , which is certainly self-referential.

We've quite a lot of cleaning up to do before we can delve into the core of mathematics. One might think this is a waste of time, but an inconsistent theory is truly horrible. The *principle of explosion*<sup>20</sup> allows one to prove, given an inconsistent set of assumptions, that *everything* is provably true and false. This is rather boring and something one would hope to avoid. While it will only play a small role throughout our investigations, we should discuss Bertrand Russell's *type theory*, invented in collaboration with Alfred North Whitehead (1861-1947 C.E.).<sup>21</sup> This was one of the first attempts at resolving these paradoxes. Here everything has a *type*, which may be thought of as a non-negative integer. When discussing containment, for example *the collection A is contained in the collection B*, we only give this meaning if  $B$  is of type 1 greater than  $A$ . In general operations can only be applied to objects of the correct type. This has applications in computer science and programming. Programming languages like Python allow one to define functions that take undefined inputs. The program will crash if an input is given that the function cannot safely handle. For example, consider the following code which is meant to take in a real number and add 1 to it:

---

<sup>16</sup> Also known as Löb's paradox

<sup>17</sup> Much the way Russell's paradox proved naïve set theory to be inconsistent.

<sup>18</sup> Programming enthusiasts should note the language *Haskell* is named after Curry.

<sup>19</sup> Curry's paradox shows the inconsistency of naïve set theory, the original *lambda calculus*, and Curry's *combinatory logic*.

<sup>20</sup> We will prove the principle later in this chapter.

<sup>21</sup> Not to be confused with the great topologist J. H. C. Whitehead.

## Python Code with Ambiguous Input

```

1 def f(x):
2     return x+1

```

If we enter  $x = 1$ , then type  $f(x)$ , this will return 2 without error. However Python allows strings and if we enter  $x = "Bob"$  the program will crash. This implementation of type checking is called *Duck typing*<sup>22</sup> in which the type of the input is checked at run time and if the type looks correct, the program will attempt to execute it. Contrast this with languages which require types to be declared prior to compilation or execution. Writing this in C we have.<sup>23</sup>

## C Code with Declared Input

```

1 double f(double x){
2     return x+1;
3 }

```

If we try to use this function on anything that is not a real number the program will refuse to compile.<sup>24</sup> The entirety of type theory is laid out in the three volume treatise *Principia Mathematica*. We will adopt Zermelo-Fraenkel set theory in addition to the logical axioms of Hilbert to formulate mathematics, both of which were introduced 20 years after Russell and Whitehead's efforts.

### 1.1.2 Sets

The main objects in mathematics are *sets*. This development came about in the 1800's with figures like Georg Cantor, Augustus De Morgan, and Bernard Bolzano making the first strides in the theory. The early history is vague and intuitive, but the obscurity led to Russell's Paradox which showed the naïvity of set theory to be inconsistent. We'll discuss this in Chapt. 3, for now we just need a definition. Georg Cantor (1845-1918 C.E.) wrote [Can95]:

*A set is a gathering together into a whole of definite distinct objects of our perception or of our thought, which are called the elements of the set.*

*Beiträge zur Begründung der Transfiniten Mengenlehre*<sup>25</sup>  
Georg Cantor, 1985 C.E.

<sup>22</sup> If your input looks like a duck and sounds like a duck, it is probably a duck.

<sup>23</sup> For simplicity, *double* in C simply means a real number.

<sup>24</sup> Unless you're using a *really* outdated compiler.

<sup>25</sup> English: *Contributions in Support of Transfinite Set Theory*.

Felix Hausdorff (1868-1942 C.E.) posits [Hau27, p. 11]:

*A set is formed by the grouping together of single objects into a whole. A set is a plurality thought of as a unit.*

*Mengenlehre<sup>26</sup>*  
*Felix Hausdorff, 1927 C.E.*

Both are beautifully phrased, but circular since the terms *gathering*, *objects*, and *grouping* are not defined. This form of circularity was addressed by Alfred Tarski in his 1946 book *Introduction to Logic and the Methodology of the Deductive Science*. He expresses the need for primitive undefined notions that we take for granted and use freely. We collect the smallest number of primitives possible, motivated by intuition, and then define other terms in our theory by means of sentences involving these primitives and previously defined terms.

We do not wish to imply the circularity of the foundations of mathematics was created with the advent of set theory. In what is perhaps the most important textbook ever written, *The Elements* by Euclid of Alexandria (c. 300 B.C.E), we find the first known work that employs the *axiomatic method*. It starts with definitions, *postulates*, and *common notions*, and proceeds to prove a plethora of important theorems in a logical manner deriving results from these primitives and previously proved theorems. In modern language postulates and common notions are known as *axioms*, which are statements that we accept as true without evidence or proof. It is not without flaw since his primitive definitions are circular. For example, the first definition is of a point:

*A point is that which has no part.*

*The Elements*  
*Euclid of Alexandria, c. 300 B.C.E.*

The word *part* is never defined. Similarly, a line is defined as *breadthless length*. This is not to detract from his efforts but to show the problem of *infinite regress* is unavoidable unless we assert that certain terms need no definition. For us, the word *set* will have no real definition. Nevertheless, we write the following:

### Definition 1.1.1: Set

A **set** is a collection of objects called the elements of the set.

The circularity we pointed out in Cantor's definition arises here since neither *collection* nor *object* have been defined. To begin doing mathematics we need a *thing*. Sets act as our thing. We know they exist<sup>27</sup> but cannot define them

<sup>26</sup> English: *Set Theory*.

<sup>27</sup> In the sense of Plato's realism. Whether sets exist in any real sense is another question.

very well. Instead we describe how they behave and how to obtain new sets from pre-existing ones via *axioms*. Before doing so we should first get familiar with the notation. We cannot build set theory just yet since we've yet to develop logic. In the most elementary systems such as Peano arithmetic there is a notion of set, and hence we need to define this first. Pedagogically it is poor to proceed without examples, so we provide some now.

### Notation 1.1.1: Element Notation

If  $A$  is a *set* and if  $x$  is an element of  $A$ , then we denote this by writing  $x \in A$ . If  $x$  is not an element of  $A$ , we write  $x \notin A$ .

**Example 1.1.4** The first three letters of the Latin alphabet can be expressed in set notation as follows. If let the symbol  $A$  denote this set, we may write:

$$A = \{a, b, c\} \quad (1.1.10)$$

If we let  $B$  denote the first three positive integers, we obtain:

$$B = \{1, 2, 3\} \quad (1.1.11)$$

Using element notation (Not. 1.1.1) we have  $1 \in B$ , but  $4 \notin B$ . That is,  $B$  contains the number 1 but does not contain the number 4. Similarly,  $a \in A$  and  $d \notin A$ . The symbol  $\in$  reads *is in*, or *is an element of*, or *is contained inside of*. Thus  $a \in A$  reads  $a$  is an element of  $A$ , or simply  $a$  is in  $A$ . The notation  $b \in A$  also reads as  $b$  is contained inside of  $A$ . The notation  $\notin$  is the negation of this: *not in* or *not an element of*, so  $4 \notin B$  reads 4 is not an element of  $B$ .

**Example 1.1.5** For the working mathematician sets are allowed to contain almost anything they like. For example, we could consider the set of all cities of Earth and label this  $C$ . Then Boston is an element of  $C$ , but Massachusetts is not since Boston is a city, but Massachusetts is not (it's a state). Similarly, London would be an element of  $C$ , but England would not be.

**Example 1.1.6** In the set theory that we will be working with, Zermelo-Fraenkel set theory, *everything* is a set. This will be explained later, but we quite literally mean everything. The integers will be defined via John von Neumann's construction. We start with the empty set  $\emptyset$  which is the set that contains nothing, often denoted  $\emptyset = \{\}$ , and this will be our zero. We proceed and define  $1 = \{\emptyset\}$ ,  $2 = \{\emptyset, 1\}$ ,  $3 = \{\emptyset, 1, 2\}$ , and so on. Moreover *functions* are defined as sets, as are *ordered pairs*  $(a, b)$ , and even *orderings*.

**Example 1.1.7** The collection of all non-negative integers  $0, 1, 2, \dots$  constitute a set which is often denoted  $\mathbb{N}$ . If we include the negatives we also get a set,

labelled  $\mathbb{Z}$ . The rational numbers form a set, as do the real numbers, and these are written  $\mathbb{Q}$  and  $\mathbb{R}$ , respectively. Lastly, the complex numbers also create a set  $\mathbb{C}$ . It is not obvious how to make these familiar things into sets without circularity, something we wish to avoid in foundations. Such constructions are another aim of Book One.

Elaborating on the discussion in Ex. 1.1.6, there are other theories for the foundations of mathematics that allow for primitive notions such as classes and universes. Some of these theories are extremely weak (cannot prove much) but very safe (there is likely no contradiction), whereas some are very user friendly but almost certainly fallacious. Peano's axioms are an example of a weak system of which the axioms are so basic and obvious that no one is likely to ever find a contradiction, but they cannot assert the existence of negative integers, let alone the reals. On the other hand, any theory that allows one to say the *collection* of all sets whether the collection is a class, or a universe, or whatever, is one that should be treated with skepticism. The theory of Zermelo and Fraenkel is a healthy middle ground. Strong enough to do most mathematics, and no contradiction found yet, though much of the 20<sup>th</sup> century was spent searching to no avail.

## A Brief History

Considering collections of things and naming them accordingly dates back to antiquity. Aggregates of points are defined in Euclid's elements and large assemblages have been used in mathematics since. The term *set* seems to appear first in the works of Bernard Placidus Johann Nepomuk Bolzano (1781-1848 C.E.), a Bohemian who coined the German term *Menge*, which translates to set in English. The phrase appears in his *Paradoxien des Unendlichen*<sup>28</sup> which was posthumously published in 1851. At the time Bolzano was known mostly for his philosophical works, including his 1837 *Wissenschaftslehre*<sup>29</sup> which attempts, much like this book, to provide a logical foundation to many of the natural sciences. It is a shame, then, that he did not enjoy the fame as a mathematician accredited to him today since his groundbreaking works in real analysis were not published until the 1880's when the Austrian mathematician Otto Stolz (1842-1905 C.E.) happened across Bolzano's journals.

Alluded to in Bolzano's text is *Galileo's Paradox* put forward by the famed Italian astronomer Galileo Galilei in his work *Discorsi e Dimostrazioni Matematiche Intorno a Due Nuove Scienze*.<sup>30</sup> This paradox is formally resolved in set theory, most notably in the works of Georg Cantor in the late 1800's. Quite

<sup>28</sup> English: *The Paradoxes of the Infinite*.

<sup>29</sup> English: *Theory of Science*

<sup>30</sup> English: *Discourses and Mathematical Demonstrations Relating to Two New Sciences*, though commonly shortened to *Two New Sciences*.

impressive, then, that Galileo pondered such problems two centuries ahead of his time. Let's examine his paradoxes.

Consider the two lines drawn in Fig. 1.2. One is longer than the other but both contain infinitely many points. Since one is greater we are forced to conclude, as Galileo writes, that we have something greater than infinity since the infinity of points in the longer is greater than the infinity of points in the shorter [Gal38].



Fig. 1.2: Lines for Galileo's Paradox

The conclusion is somewhat correct, there are different infinities in set theory. The infinity of the real numbers is strictly greater than the infinity of the natural numbers, for example. This will be made very clear when we discuss bijective functions and cardinalities, but for now it may seem as mathematical gibberish.<sup>31</sup> Galileo is wrong in claiming one line has more points than the other since both have the same infinity of points. That is, we can match up every point of the long line to a unique point in the short line, showing the longer one can't be greater in quantity.

To construct this one-to-one correspondence requires a bit of perspective. Suppose we place an observer at the point  $O$  shown in Fig. 1.3. The observer would not see the longer line, but if he or she were able to then both would appear to be the same length. To get our one-to-one correspondence we draw a straight line from this observer to the first line and then continue outwards to the longer one. This takes a point in the short line uniquely to a point in the long one, and every point in the long line is hit by some point in the short one. So we've matched every point in the short line to every point in the long line and hence the claim cannot be made that one has *more* elements than the other. What sets them apart (their *length*) is not a set-theoretic concept, but rather a *measure* theoretic one.<sup>32</sup>

Galileo's second paradox is presented in a similar manner. He writes that there are infinitely many integers, and infinitely many square integers. Every square integer is also an integer, for example 1, 4, 9, 16, 25, and so on. The converse

<sup>31</sup> These conclusions are due to Cantor. Some, like Hilbert, defended his work while many, like Kronecker and Wittgenstein, lambasted his writings and his character. Some solace: Most of Cantor's works were accepted as true by the mid 20<sup>th</sup> century and his results have become standard in 21<sup>st</sup> century analysis, topology, and measure theory courses.

<sup>32</sup> In a late night musing I stumbled across this same problem. I learned of the solution from my mathematical mentor James *Kiwi* Graham-Eagle who presented me with Fig. 1.3.

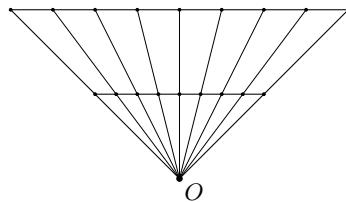


Fig. 1.3: Solution to Galileo's Paradox

is false, not every integer is a square integer since 2, 3, 5, 6, 7, 8, 10, and so on are not squares. From this argument Galileo concluded there are *more* integers than square integers. The following observation then troubled him. For every integer<sup>33</sup> there is a unique square number, namely given  $n$  there is  $n^2$ . Given this there cannot be more integers than squares, a contradiction.

The amazing part about this paradox is that it already contains its own solution. The *function* which sends  $n$  to  $n^2$  is a *bijection* between the integers and the squares. In set theory if such bijection exists, then the two sets have the same size. The problem can be reduced significantly if we note that not every integer is even, but every even integer is an integer. One might wrongly conclude then that there are more integers than even integers. But for every integer there is precisely one even integer, namely given  $n$  we have  $2n$ . Even easier, all positive integers are integers, but not all integers are positive since zero is not considered positive. We might then conclude that the set of positive integers has infinity points, whereas the set of all integers has infinity +1 points. But for every integer there is precisely one positive integer, namely given  $n$  we have  $n + 1$ .

All three examples have the same problem. We take an infinite set, remove points<sup>34</sup> but arrive at a set that is the same *size* as what we started with. As stated, this problem is handled in set theory. By definition two sets are of the same size if one-to-one correspondence, much like the three we have demonstrated, exists between them. What's even more troubling<sup>35</sup> is that the set of all *rational* numbers is the same size as all of these.

### 1.1.3 Predicates and Propositions

Propositions and predicates will be two more of our primitive notions which we will vaguely define, but mostly rely on intuition.

<sup>33</sup> Here we loosely use integer to mean non-negative natural number: 0, 1, 2, 3, etc.

<sup>34</sup> Perhaps infinitely many, as Galileo has done.

<sup>35</sup> Unbeknownst to Galileo but knowst to Cantor.

**Definition 1.1.2: Predicate**

A **predicate**  $P$  on a **set** of variables is a sentence such that for any valid input one may state that  $P$  is either true or false.

We did not define *variable*, but this primitive is understood as a symbol or placeholder that may represent *things*. Predicates are the main tool used in set theory for defining and building new sets via the *axiom schema of specification*. Many describe predicates as *functions* from a set  $A$  to the Boolean-valued set  $\{\text{True}, \text{False}\}$  and this is a fine way of doing this provided the notion function has been defined. Many take function as another primitive, and thus one has the following decision to make: Do we accept *predicate* as a primitive, or *function*? We'll adopt predicate and later *define* functions using elementary notions from the axioms of set theory.

**Example 1.1.8** Let  $A$  be the set  $A = \{1, 2, 3\}$  and let  $P$  be the predicate  $x$  is greater than 2. The set of values in  $A$  that satisfy this claim is  $B = \{3\}$ . If we let  $Q$  represent  $x$  is negative, then there are no elements in  $A$  that satisfy  $Q$  and hence the resulting set is the empty set  $\emptyset$ .

Moving on to propositions, we first express the naïve Aristotelian definition put forward by Aristotle (384-322 B.C.E.).

A proposition is a sentence which affirms or denies a predicate. (1.1.12)

The classic example is the so-called *Socrates syllogism*.

**Example 1.1.9** We wish to conclude the ancient greek philosopher Socrates was mortal. We start with the following proposition: *All men are mortal*. We assert this sentence is true and we may be used later in the proofs of other claims. Second, we state *Socrates is a man*. This is another proposition, a statement that we accept as true. To conclude Socrates is mortal we need some *rule of inference* that allows us to tie these two propositions together. One such rule, known as *modus ponens* does exactly what we need. It states that if  $P$  and  $Q$  are propositions, if  $P$  implies  $Q$ , and if  $P$  is true, then  $Q$  is true. Using this, since all men are mortal, and since Socrates is a man, we conclude that Socrates is mortal.

An easier proof that Socrates is mortal goes as follows: He's dead. Nevertheless, we are starting to see what is needed to construct valid proofs.

**Definition 1.1.3: Proposition**

A **proposition** is a **predicate** evaluated at a particular **set** of variables, taken in a particular order, which is then affirmed or denied to be true.

The order of the input of variables is important. Sets do not have order, and to give rise to such notions requires the bulk of the axioms of set theory. Axioms are themselves propositions which we assert to be true without proof, and hence we stumble upon circularity. To define ordered sets we need the definition of proposition, and to define proposition we need ordered sets. The solution is to accept proposition as a primitive which needs no real definition.

**Example 1.1.10** Let's use the previous example of Socrates to motivate what we mean. Let  $P$  be the predicate  $x \text{ is a man}$ . If we input Socrates we obtain  $P(\text{Socrates}) = \text{True}$ . If we input Hatshepsut, we get  $P(\text{Hatshepsut}) = \text{False}$ . This is how we distinguish a predicate from a proposition.

**Example 1.1.11** Suppose we have the predicate  $x \text{ is a man}$ ,  $y \text{ is a woman}$  and let us label this  $P(x, y)$ . The *order* of the inputs now matters. For example,  $P(\text{Socrates}, \text{Hatshepsut}) = \text{True}$ , but  $P(\text{Hatshepsut}, \text{Socrates}) = \text{False}$ .

#### 1.1.4 Rules of Inference

Given a collection of propositions we often wish to derive new ones. Indeed, that is the entirety of mathematics: Proving new theorems. We must precisely state which rules we accept as valid and then attempt to stay consistent with them. The first we have already discussed, *modus ponens*. This rule applies to *implications*, one of the two primitives we adopt in our language of deducing new claims, the second being negation which we will discuss soon enough.

##### Definition 1.1.4: Implication

An **implication** on a **proposition**  $Q$  by a proposition  $P$  is the sentence that if  $P$ , then  $Q$ . We denote this  $P \Rightarrow Q$ .

The proposition  $P$  is called the *hypothesis*, and  $Q$  is the *conclusion*. We can describe implication via *truth tables*. Truth tables for a finite collection of propositions exhaust all possible combinations of True and False, and then apply these combinations to the logical question at hand. Implication is depicted in Tab. 1.6. Given  $n$  propositions there are  $2^n$  possible scenarios and so these truth tables get very big very fast. We can see that there are  $2^n$  possibilities since that are  $n$  choices to be made from 0 and 1, so there are  $2 \cdot 2 \cdots 2$  possibilities with  $n$  2's (of course, we've yet to define what  $2^n$  means).

$P$	$P$	$P \Rightarrow Q$
False	False	True
False	True	True
True	False	False
True	True	True

Table 1.6: Truth Table for Implication

$P$	$Q$	$P \Rightarrow Q$
0	0	1
0	1	1
1	0	0
1	1	1

Table 1.7: Alternate Table for Implication

It is often easier and more useful to write these tables using zeros and ones. For one this hints at a means for computers to be able to understand and manipulate logical statements and proofs, but it is also less cumbersome and allows us to systematically run through the possibilities. That is, we start with all propositions set to 0 and then flick the right-most one to 1, then the second right-most, and so on. The symbol 0 denotes false, and 1 represents truth. This alternative truth table for implication is shown in Tab. 1.7. This shows that the only possible way for  $P \Rightarrow Q$  to be false is if  $P$  is true, yet  $Q$  is false. This may be strange according to everyday language and there is a common tendency to confuse the order of implication. We spell this out in an example.

**Example 1.1.12** Consider the proposition  $P = I \text{ am late for work}$  together with  $Q = I \text{ will be fired}$ , and let's exhaust the four possible scenarios of if  $P$ , then  $Q$ . That is, we consider the claim *if I am late for work, then I will be fired*. Suppose I was not late for work, and I was not fired. Is  $P \Rightarrow Q$  true? Well, the criterion for  $P$  was not satisfied and hence the statement is *not* false, and so we claim it is true. Next, I was not late for work yet I was still fired (harsh). Again, the criterion for  $P$  was not satisfied and so the claim is not false, and hence we accept it is true. Third, I was late for work and I was not fired (nice boss). Here we see that  $P \Rightarrow Q$  is *false*. The criterion for  $P$  was satisfied, yet  $Q$  was not and therefore  $P \Rightarrow Q$  is a false statement. Lastly, I was late for work and I was fired. This is perhaps the easiest one to handle since it is verbatim what one thinks of when they hear *if, then* claims. Here,  $P \Rightarrow Q$  is true.

Given this definition of implication the axiom of *modus ponens*, short for *modus ponendo ponens* which in Latin means *mode that by affirming affirms*, seems redundantly obvious. Nevertheless, we must write it out. First, we must define *axiom* and *proof*. Proof is another primitive, but axiom is not. We may define axiom in terms of previous notions.

### Definition 1.1.5: Proof

A **proof** of a **proposition** is a valid argument that rigorously affirms the proposition.

We've yet to see what a valid argument is. These are formed by combining previously proved propositions, together with definitions, to arrive at a conclusion using the allowed rules of inference. As of yet we have not claimed what rules of inference we will accept, nor do we have any proposition to build from. Both of these issues are addressed by axioms.

### Definition 1.1.6: Axiom

An **axiom** is a **proposition** that is affirmed to be true without **proof**.

In a good system the axioms should be intuitively obvious and not too controversial. In *the Elements* there are 10 axioms (five are called postulates and the others common notions), most of which are not too awe inspiring. The five postulates are phrased as follows:

1. *To draw a straight line from any point to any point.*
2. *To produce a finite straight line continuous in a straight line.*
3. *To describe a circle with any centre and distance.*
4. *That all right angles are equal to one another.*
5. *That if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the straight lines, if produced indefinitely, meet on that side on which are the angles less than two right angles.*

*The Elements,  
Euclid of Alexandria, c. 300 B.C.E.*

Only the fifth postulate, known famously as *Euclid's Fifth Axiom*, is non-obvious. These five constitute the study of *Euclidean geometry*, the first four of which make up *absolute geometry*. For almost two thousand years a handful of mathematicians attempting to prove that absolute geometry and Euclidean geometry are the same. That is, the fifth postulate can be proved from the other four. Many minds of antiquity such as Claudius Ptolemy (c. 100-170 C.E.), Proclus Lycaeus (c. 412-485 C.E.), Omar Khayyám (1050-1123 C.E.), and others made the attempts, but all merely introduced an alternative axiom either explicitly or implicitly which was equivalent to Euclid's. Euclid himself was hesitant in his use of the postulate, only using it nearly 30 theorems into book one when he needed it. So the question remains, can the fifth be proven from the first four? The answer is *no*. We know this because there are *models* of the first four where the fifth one fails.

In the 18<sup>th</sup> century the Swiss mathematician Johann Heinrich Lambert (1728-1777 C.E.) made investigations into a quadrilateral where three of the angles are right. If one could prove the fourth angles must also be right, then with this Euclid's fifth could be proven (Omar Khayyám did precisely this but claimed the fourth angle being right is self-evident). Lambert discarded the possibility of the fourth angle to be obtuse, but then made investigations into the acute case. He was never able to prove the acute case was impossible. This leads to another *model* of absolute geometry, *hyperbolic geometry* in which the fourth angle is allowed to be acute.

The revelation that Euclid's fifth is not provable from the other four is made more intuitive by the work of John Playfair (1748-1819 C.E.) who described the following axiom, now known as *Playfair's axiom*:

*In a plane, given a line and a point not on it, at most one line parallel to the given line can be drawn through the point*

*Elements of Geometry  
John Playfair, 1795 C.E.*

In combination with the first four axioms this is equivalent to Euclid's fifth. The first four can prove parallel lines exist, and hence we cannot claim refute this, but we may change Playfair's axiom to state that *many* such lines exist, again giving us the model of hyperbolic geometry. Perhaps we wish to claim there are no parallel lines, in which case we must do away with Euclid's second axiom. This leads to *elliptic geometry*, which is the geometry on a sphere. This model shows that Euclid's second cannot be proved from the others as well.

### Axiom 1.1.1: Modus Ponens

If  $P$  and  $Q$  are propositions, if  $P \Rightarrow Q$ , and if  $P$  is true, then  $Q$  is true.

**Example 1.1.13** Let  $P$  be the proposition *bears are mammals* and  $Q$  be the proposition *mammals are animals*. If  $P$ , then  $Q$  reads if bears are mammals, then bears are animals. Since bears are mammals, we infer that bears are animals.

The next two commonly accepted rules of inference, known as *modus tollens* and *contraposition* are widely used in the mathematical world and relate to a statements contrapositive. The contrapositive is related to the *negation* of a proposition, and so we define this now.

### Definition 1.1.7: Negation

The negation of a proposition  $P$  is the proposition *not*  $P$ , denoted  $\neg P$ .

Negation is a *unary* operation acting on a single variable. The truth table is short.

$P$	$\neg P$
0	1
1	0

Table 1.8: Truth Table for Negation

That is, negation takes true to false and maps false to true. Negation and implication form our two primitive logical operations, and the other familiar terms (disjunction, conjunction, equivalence) can be expressed using these two. With negation defined, we now present a *logical fallacy*, a form of reasoning that is invalid and leads to contradiction. This is the fallacy of *affirming the consequent*.

### Example 1.1.14: Affirming the Consequent

The fallacy of affirming the consequent is also known as the inverse fallacy, and in mathematics it is called the *converse* fallacy. *Modus ponens* tells us that if  $P$  and  $Q$  are propositions, if  $P \Rightarrow Q$ , and if  $P$  is true, then  $Q$  is true. The *converse* of the statement *if  $P$ , then  $Q$*  is the sentence *if  $Q$ , then  $P$* . The validity of  $P \Rightarrow Q$  does **not** verify the converse, much to the dismay of mathematicians. Theorems are often held in higher regard if they express the *equivalence* of two propositions:  $P \Rightarrow Q$  and  $Q \Rightarrow P$ . That is, both the statement and its converse are true. There are everyday and mathematical examples showing that this line of reasoning is faulty. Previously we discussed the classification of bears: *if an animal is a bear, then it is a mammal*. The converse states if an animal is a mammal, then it is a bear. Humans are a counterexample to this claim since humans are mammals but are not bears. In the mathematical world we can consider the proposition *if  $n$  is an odd integer, then  $n$  is not 2*. The converse states that if  $n$  is not 2, then  $n$  is not an odd integer. But 1 is not 2, yet 1 is indeed an odd integer. ■

That affirming the consequent is invalid can be seen from truth tables (see Tab. 1.9).

$P$	$Q$	$P \Rightarrow Q$	$Q \Rightarrow P$
0	0	1	1
0	1	1	0
1	0	0	1
1	1	1	1

Table 1.9: Truth Table for the Converse

Since the columns for  $P \Rightarrow Q$  and  $Q \Rightarrow P$  are different, we see that these are different propositions. Negation further allows us to define the *contrapositive* of the proposition  $P \Rightarrow Q$ , which is the new proposition  $\neg Q \Rightarrow \neg P$ . As it turns out, this is not new proposition at all and is equivalent to  $P \Rightarrow Q$ . Note  $P \Rightarrow Q$  is false only when  $P$  is true, yet  $Q$  is false. Similarly,  $\neg Q \Rightarrow \neg P$  is false only if  $\neg Q$  is true and  $\neg P$  is false. But if  $\neg Q$  is true, then  $Q$  is false (Def. 1.1.7) and if  $\neg P$  is false, then  $P$  is true (Def. 1.1.7). Thus  $\neg Q \Rightarrow \neg P$  is only false when  $P$  is true and  $Q$  is false. We can further examine this via truth tables (see Tab. 1.10).

$P$	$Q$	$\neg P$	$\neg Q$	$P \Rightarrow Q$	$\neg Q \Rightarrow \neg P$
0	0	1	1	1	1
0	1	1	0	1	1
1	0	0	1	0	0
1	1	0	0	1	1

Table 1.10: Truth Table for the Contrapositive

**Example 1.1.15** Suppose  $a$  and  $b$  are variables representing real numbers and  $P$  is the proposition  $a < 1/2$  and  $b < 1/2$ , and let  $Q$  be the proposition  $a+b < 1$ . What is the contrapositive of  $P \Rightarrow Q$ ? This would be  $\neg Q \Rightarrow \neg P$ , where  $\neg Q$  is the negation of  $Q$  which reads  $a+b \geq 1$ . Similarly,  $\neg P$  is the statement  $a \geq 1/2$  or  $b \geq 1/2$ . Thus, the contrapositive says that if  $a+b \geq 1$ , then either  $a \geq 1/2$  or  $b \geq 1/2$  (or both). While the contrapositive of a statement is always equivalent to the original statement, the converse need not be. Indeed, this statement is true (once one knows the order structure of real numbers), but the converse is not. The converse states that if  $a+b < 1$ , then  $a < 1/2$  and  $b < 1/2$ , but letting  $a = 2$  and  $b = -3$  contradicts this claim.

The axiom of *modus tollens* states that if  $P \Rightarrow Q$ , and if  $\neg Q$  is true, then  $\neg P$  is true. We are not directly adopting this axiom since it is provable from the axiom of *modus ponens* if one accepts the standard axioms of set theory. Indeed, *modus tollens* is implied by the *law of the excluded middle*, which is in turn implied by the *axiom of choice*, two topics that will be discussed in Chapt. 3. Similar to *modus tollens*, the axiom of contraposition states that  $P \Rightarrow Q$  is equivalent to  $\neg Q \rightarrow \neg P$ . That is, if the statement  $P \Rightarrow Q$  is true,

then the contrapositive  $\neg Q \rightarrow \neg P$  is also true.

**Example 1.1.16** Consider once again the description of Socrates. We start with the proposition *all humans are mammals*. If Socrates is a human, then Socrates is a mammal. Therefore if Socrates is *not* a mammal, then Socrates is *not* a human. We could also say that all bears are mammals, and hence if you come across an animal that is *not* a mammal, then this animal is *not* a bear. Reasoning like this follows from contraposition.

Before moving to connectives Hilbert systems it is worth while mentioning a few *invalid* forms of reasoning. We do not include these as rules of inference since they lead to contradiction, although students often make these fallacious mistakes when exploring proofs for the first time. We've discussed *affirming the consequent* and now identify *denying the antecedent*, which is another type of converse fallacy and is the most common to make. Denying the antecedent goes as follows, if  $P$  implies  $Q$ , and  $\neg P$ , then  $\neg Q$ . This is false, and we will provide plenty of examples to indicate this.

**Example 1.1.17** Harking back to a previous example, consider the statement *if I am late to work, then I will be fired*. Now suppose I was not late to work. Does this mean I was not fired? No! Perhaps I was lazy on the job, or uttered too many vulgarities (I do have a sailor's mouth). Knowing that I was not late tells us nothing about whether or not I was fired. It is only if I *was* late that we can then appropriately apply *modus ponens* and conclude that I was fired.

**Example 1.1.18** Consider the proposition *If  $n$  is an odd integer, then  $n$  is not 2*. Now suppose we are told than  $n$  is *not* an odd integer. Can we conclude that  $n$  is 2? No! It may be 4, or 6, or any other even integer.

We can consider  $P \Rightarrow Q$  and  $\neg P \Rightarrow \neg Q$  by means of truth table. The statement  $\neg P \Rightarrow \neg Q$  is called the *inverse* of  $P \Rightarrow Q$ .

$P$	$Q$	$\neg P$	$\neg Q$	$P \Rightarrow Q$	$\neg P \Rightarrow \neg Q$
0	0	1	1	1	1
0	1	1	0	1	0
1	0	0	1	0	1
1	1	0	0	1	1

Table 1.11: Truth Table for the Inverse

The next fallacy is known as the *fallacy of the undistributed middle*. This is the first argument that takes three propositions. It falsely concludes that if  $P_1 \Rightarrow Q$ , and if  $P_2 \Rightarrow Q$ , then  $P_1 \Rightarrow P_2$ . This is false as the truth table below demonstrates.

$P_1$	$P_2$	$Q$	$P_1 \Rightarrow Q$	$P_2 \Rightarrow Q$	$P_1 \Rightarrow P_2$
0	0	0	1	1	1
0	0	1	1	1	1
0	1	0	1	0	1
0	1	1	1	1	1
1	0	0	0	1	0
1	0	1	1	1	0
1	1	0	0	0	1
1	1	1	1	1	1

Table 1.12: Fallacy of the Undistributed Middle

There's a column on this table where  $P_1 \Rightarrow Q$  is true, and  $P_2 \Rightarrow Q$  is true, yet  $P_1 \Rightarrow P_2$  is false. Namely, choose  $P_1 = \text{True}$ ,  $P_2 = \text{False}$ , and  $Q = \text{True}$ .

**Example 1.1.19** Consider the claim *all mathematicians love geometry* which one could only hope is true. Consider also *all physicists love geometry*. Given that these two statements are true, it would be wrong to conclude that all mathematicians are physicists.

**Example 1.1.20** We can consider a mathematical proposition. If  $n$  is divisible by 2, then  $n$  is even, and if  $n$  is divisible by 4, then  $n$  is even. We cannot conclude that if  $n$  is divisible by 2, then  $n$  is divisible by 4 since the number 6 serves as a counterexample. That is,  $6 = 2 \cdot 3$  and hence 6 is divisible by 2, but it is not divisible by 4.

Lastly, we discuss the difference between a *valid* argument and a *sound* one. A valid argument is one that proves a claim from hypothesized propositions by correctly using the rules of inference. A sound argument is a valid argument of which the hypothesized propositions are true.

**Example 1.1.21** Consider the proposition *all birds can fly*. We invoke *modus ponens* and arrive at the following absurdity:

$$\text{All birds can fly.} \tag{1.1.13a}$$

$$\text{Penguins are birds.} \tag{1.1.13b}$$

$$\text{Therefore, penguins can fly.} \tag{1.1.13c}$$

It is currently believed that penguins are incapable of flight no matter how hard they may try, and hence we have used our rules of inference correctly, but we've arrived at a false claim. That is, our argument is valid, but it cannot be sound. And indeed, the flaw is that the proposition *all birds can fly* is false since penguins serve as a counterexample.

Another example is known as the *masked-man fallacy*.

### Example 1.1.22: Masked-Man Fallacy

---

Suppose I have a friend *Bob*. Since he is my friend, the proposition *I know Bob* is true. Suppose further that there is a man wearing a mask. Since he is wearing a mask, I do not know who he is and hence the proposition *I do not know the masked man* is true. We obtain the following:

$$\text{I know Bob.} \tag{1.1.14a}$$

$$\text{I do not know that masked man.} \tag{1.1.14b}$$

$$\text{Therefore, Bob is not the masked man.} \tag{1.1.14c}$$

Our argument is valid and follows from *modus tollens*. That is, we have the proposition *if the man is Bob, then I know him*. Hence, if I don't know the man, then it is not Bob. However this argument is not sound since it is perfectly possible for Bob to be wearing the mask. The flaw comes from the proposition *I do not know the masked man*. In truth, it may be possible that I do. This knowledge is not available to me and cannot be used or refuted in the argument. ■

## 1.2 Hilbert Systems

In the set theory we will be working with there are a few words and symbols that are left undefined. As stated, this is unavoidable since defining everything would be circular, and we try to use the fewest number of *primitive*. The main undefined symbol in set theory is that of *containment* ( $\in$ ) (see Not. 1.1.1), a type of *predicate* of the form *x is in A*. Other common symbols such as subset ( $\subseteq$ ) and equality ( $=$ ) are then defined in terms of this. In a similar manner there are other commonly used symbols in mathematical logic such as *disjunction* ( $\vee$ ), *conjunction* ( $\wedge$ ), and *equivalence* ( $\Leftrightarrow$ ) that we need not accept as primitives, but rather can define in terms of implication ( $\Rightarrow$ ) and negation ( $\neg$ ). We start with conjunction, which gives meaning the logical term *and*.

### 1.2.1 Connectives

The symbol  $\wedge$  is used to represent the word *and* in a mathematical way.

**Definition 1.2.1: Conjunction**

The conjunction of propositions  $P$  and  $Q$  is the statement  $P$  and  $Q$  defined by the formula:

$$P \wedge Q \equiv \neg(P \Rightarrow \neg Q)$$

Before justifying this definition, we wish to get an idea as to what *and* should mean. Given two propositions  $P$  and  $Q$ ,  $P$  and  $Q$  should be considered if and only if both  $P$  is true and  $Q$  is true. That is, both are true simultaneously.

**Definition 1.2.2: Disjunction**

The disjunction of propositions  $P$  and  $Q$  is the statement  $P$  or  $Q$  defined by the formula:

$$P \vee Q \equiv (P \Rightarrow Q) \Rightarrow P$$

We have relied on the word *statement* being already defined, and similarly for the words *parameter* or *variable*. For most this is not an issue, but it may irk others. From our undefined symbol  $\in$  we build new symbols by expressing them in terms of a *formula*, which is simply a finite sequence of symbols. Here the word *sequence* is meant to imply that the *order* in which we combine these symbols is important and that rearranging said order may create a different inequivalent formula. We build formulas by defining a few symbols that stand as placeholders for standard words in English. There are four symbols, called *connectives*, that we use. From this we see that we have introduced 6 new words that are undefined but require comment. The words are *and*, *or*, *if*, *then*, *true*, and *false*. There are other symbols we could adopt, such as *equivalence*:

$$a \Leftrightarrow b$$

But from how we shall define these notions, this new symbol is equivalent to a combination of the previous ones:

$$\left( (a \Leftrightarrow b) \Rightarrow ((a \Rightarrow b) \wedge (b \Rightarrow a)) \right) \wedge \left( ((a \Rightarrow b) \wedge (b \Rightarrow a)) \Rightarrow (a \Leftrightarrow b) \right)$$

That is,  $a \Leftrightarrow b$  if and only if  $a$  is true if and only if  $b$  is true. Similarly, we could define *does not imply*:

$$a \not\Rightarrow b$$

But this is the same as:

$$a \not\Rightarrow b \iff \neg(a \Rightarrow b) \iff a \wedge \neg b$$

The words true and false are assumed to be well defined. They are also assumed to be opposites of each other (which we will define in terms of negation). We will use truth tables to define what various connectives mean when it is known that certain propositions are true or false. In such tables the symbol 0 represents that a proposition is false and the symbol 1 represents truth.

The other four words can be ambiguous in their everyday usage which we cannot allow for in mathematics. As such we must specify what we mean when we use these words and rid of any such ambiguity.

### 1.2.2 Conjunction

The conjunction connective ( $\wedge$ ) is used to denote the word *and*. Given two propositions  $P$  and  $Q$ ,  $P \wedge Q$  is a true statement if and only if both  $P$  and  $Q$  are true. That is, we associate to  $\wedge$  the following truth table:

$P$	$Q$	$P \wedge Q$
0	0	0
0	1	0
1	0	0
1	1	1

Table 1.13: Truth Table for Conjunction

There are several *axioms* of conjunctions that are intuitively obvious, but must be stated since their use is wide spread.

#### Axiom 1.2.1: Axioms of Conjunction

If  $P$  and  $Q$  are propositions, then the following are true:

$$P \wedge Q \iff Q \wedge P \quad (\text{Commutativity of Conjunction})$$

$$P \wedge \text{True} \iff P \quad (\text{Identity of Conjunction})$$

### 1.2.3 Disjunction

The disjunction connective ( $\vee$ ) represents the word *or*. Given two propositions  $P$  and  $Q$ ,  $P \vee Q$  is true if and only if  $P$  is true, or  $Q$  is true, or both  $P$  and  $Q$  are true. There is an unfortunate ambiguity in English as to whether  $P$  or  $Q$  means  $P$  is true or  $Q$  is true, but not both, or whether it means  $P$  is true or  $Q$  is true, or *both* are true. The convention is to adopt the latter definition. That is,  $P \vee Q$  has the following truth table:

$P$	$Q$	$P \vee Q$
0	0	0
0	1	1
1	0	1
1	1	1

Table 1.14: Truth Table for Disjunction

There is another connective called the *exclusive* or, which is defined to be false if both  $P$  and  $Q$  are true. The symbol  $\vee$  is strictly used to denote the inclusive or. That is, the word or as represented by the truth table in Tab. 1.14.

### 1.2.4 Implication

Examining, we see that there are scenarios where  $P \Rightarrow Q$  is true and  $Q \Rightarrow P$  is false, and similarly where  $P \Rightarrow Q$  is false and  $Q \Rightarrow P$  is true. Propositions  $P$  and  $Q$  such that  $P \Rightarrow Q$  and  $Q \Rightarrow P$  are called *equivalent*, and great deal of mathematics is devoted to the search for equivalencies of statements. This is denoted by the connective  $P \Leftrightarrow Q$ . Equivalence has the following truth table:

$P$	$Q$	$P \Rightarrow Q$	$Q \Rightarrow P$	$P \Leftrightarrow Q$	$(P \Rightarrow Q) \wedge (Q \Rightarrow P)$
0	0	1	1	1	1
0	1	1	0	0	0
1	0	0	1	0	0
1	1	1	1	1	1

Table 1.15: Truth Table for Equivalence

### 1.2.5 Misc

$p$	$q$	$r$	$\neg q$	$p \vee \neg q$	$(p \vee \neg q) \wedge r$
0	0	0	1	1	0
0	0	1	1	1	1
0	1	0	0	0	0
0	1	1	0	0	0
1	0	0	1	1	0
1	0	1	1	1	1
1	1	0	0	1	0
1	1	1	0	1	1

Table 1.16: Truth Table for  $(p \vee \neg q) \wedge r$ 

**Theorem 1.2.1.** If  $a \rightarrow b$ , if  $\neg c \rightarrow \neg b$ , and if  $\neg c$ , then  $\neg a$ .

*Proof.* For if  $a \rightarrow b$ , then  $\neg b \rightarrow \neg a$ . But  $\neg c \rightarrow \neg b$ . But if  $\neg c \rightarrow \neg b$  and  $\neg b \rightarrow \neg a$ , then  $\neg c \rightarrow \neg a$ . Thus  $a \rightarrow b$ ,  $\neg c \rightarrow \neg b$ , and thus  $\neg c \Rightarrow \neg a$ .  $\square$

**Problem 1.2.1** If  $a \rightarrow b$ , if  $\neg c \rightarrow \neg b$ , and if  $\neg c$ , then  $\neg a$ .

*Proof.* For if  $\neg c \rightarrow \neg b$ , then  $b \rightarrow c$ . But if  $a \rightarrow b$  and  $b \rightarrow c$ , then  $a \rightarrow c$ . Therefore  $a \rightarrow c$ . But if  $a \rightarrow c$ , then  $\neg c \rightarrow \neg a$ . Therefore,  $a \rightarrow b$ ,  $\neg c \rightarrow \neg b$ ,  $\neg c \Rightarrow \neg a$ .  $\square$

### Theorem 1.2.2: Law of Syllogism

If  $P$ ,  $Q$ , and  $R$  are propositions, if  $P \Rightarrow Q$ , and if  $Q \Rightarrow R$ , then  $P \Rightarrow R$ .



# CHAPTER 2

---

## Predicate Calculus

---

### 2.1 Quantifiers

There are two more symbols called *quantifiers*.

$$\forall_x \quad \text{For all } x \qquad \exists_x \quad \text{There exists } x$$

Quantifiers, together with connectives, the word *set*, and the  $\in$  symbol are combined to define new terms and new symbols. The rest of mathematics rests on trusting ones intuition behind these notions.

**Example 2.1.1** Let  $\mathbb{R}$  denote the set of real numbers. The symbols  $\forall_{R \in \mathbb{R}}(n^2 \geq 0)$  can then be read in English as *For all real numbers  $x$ , the square of  $x$  is non-negative*, which is indeed a true statement. We can combine quantifiers to create more complicated statements, such as:

$$\forall_{x \in \mathbb{R}}(x \neq 0) \exists_{y \in \mathbb{R}}(xy = 1) \tag{2.1.1}$$

This reads that for all non-zero real numbers  $x$ , there exists a real numbers  $y$  such that the product  $xy$  is equal to 1. This is also a true statement.

**Example 2.1.2** The order of quantifiers is very important and often can not be interchanged. Considering the previous example, if we switch the order of the quantifiers we get:

$$\exists_{y \in \mathbb{R}}(xy = 1) \forall_{x \in \mathbb{R}}(x \neq 0) \tag{2.1.2}$$

This states that there exists a real number  $y$  such that, for every non-zero real number  $x$ , it is true that  $xy = 1$ . But this is certainly not true because if  $x = 1$  and  $y = -1$ , we obtain  $(1)y = 1$  and  $(-1)y = 1$ , and from this we conclude that  $-1 = 1$ , which is false. Hence, the order of the quantifiers is important.

**Example 2.1.3** Quantifiers can be combined with connectives to make longer and more complicated statements. For example, suppose  $P$  is the proposition *true if  $n$  is an even integer, false otherwise*. Furthermore, let  $Q$  be the proposition *true if  $n$  is a square integer, false otherwise*. Lastly, let  $r$  be the proposition *true if  $n$  is divisible by 4, false otherwise*. Consider then the following statement:

$$\forall_{n \in \mathbb{Z}}(p(n) \wedge q(n) \Rightarrow r(n)) \quad (2.1.3)$$

This reads in English as *for all integers  $n$ , if  $n$  is an integer, and if  $n$  is a square, then  $n$  is divisible by 4.*

### 2.1.1 Negating Quantifiers

The negation of the statement *for all  $x$ ,  $P(x)$  is true* implies this is false. Thus there must exist one  $x$  such that  $P(x)$  is false, and from this we see that negating the  $\forall$  quantifier produces the  $\exists$  quantifier.

**Example 2.1.4** Let  $P$  be the proposition *true if  $x^2 = 2$*  and consider the following statement:

$$\exists_{x \in \mathbb{Q}}(P(x)) \quad (2.1.4)$$

This reads in plain English as the statement *there exists a rational number  $x$  whose square is equal to 2*. This has been known to be false since the ancient Greeks, and thus it's negation is true. We can write the negation as follows:

$$\neg(\exists_{x \in \mathbb{Q}}(P(x))) \iff \forall_{x \in \mathbb{Q}}(\neg P(x)) \quad (2.1.5)$$

This now says that for all rational numbers  $x$ , the square of  $x$  is not equal to 2.

# CHAPTER 3

---

## Zermelo-Fraenkel Set Theory

---

We'll develop mathematics from an axiomatic view built on set theory, adopting as truths the few postulates of Zermelo and Fraenkel. We'll then add the axiom of choice and proceed from there to define many familiar concepts and prove some basic results that are often taken for granted. The existence of many types of sets will be proven, rather than accepting these things as trivial truths.

### 3.1 The Axioms of Zermelo and Fraenkel

We do not yet know that sets exist. Pedagogically it seems poor to wait for examples, so we'll speak loosely for the moment so we may familiarize ourselves with the notation.

### Example 3.1.1: Using Element Notation

Given a set  $A$  that contains only a few objects, we can represent  $A$  by listing out the elements, separated by commas, and enclosing them in braces. Suppose  $A$  is the set that contains three distinct objects labelled  $a$ ,  $b$ , and  $c$ . We then write:

$$A = \{ a, b, c \} \quad (3.1.1)$$

If we are told that there is a fourth object  $d$  that is different from  $a$ ,  $b$ , and  $c$ , then we can use the notation defined in Not. 1.1.1 to write the following:

$$a \in A \quad (3.1.2a) \qquad d \notin A \quad (3.1.2b)$$

The notation  $a \in A$  should be read as  $a$  is an element of  $A$ , or  $a$  is contained in  $A$ , or simply  $a$  is in  $A$ . Similarly, the notation  $d \notin A$  should be read as  $d$  is not an element of  $A$ , or  $d$  is not contained in  $A$ .

$A$  is an example of a *finite* set, moreover it contains only three elements. For larger sets we rely on other methods to write them down. One such means is to indicate a pattern and use an ellipses to show that it goes on. Such a description is vague and lacks rigor, but can be helpful when the pattern is obvious. The set of all *natural* numbers, or non-negative integers (denoted  $\mathbb{N}$ ) can be loosely represented by writing:

$$\mathbb{N} = \{ 0, 1, 2, 3, 4, 5, \dots \} \quad (3.1.3)$$

Using our developed notation, we can write:

$$23 \in \mathbb{N} \quad (3.1.4a) \qquad -4 \notin \mathbb{N} \quad (3.1.4b)$$

Letting  $\mathbb{Z}_n$  denote all non-negative integers between 0 and  $n - 1$ , we have:

$$\mathbb{Z}_n = \{ 0, 1, 2, \dots, n - 1 \} \quad (3.1.5)$$

Thus  $17 \in \mathbb{Z}_{18}$  but  $19 \notin \mathbb{Z}_{18}$ . Lastly, we present the integers ( $\mathbb{Z}$ ).

$$\mathbb{Z} = \{ \dots, -3, -2, -1, 0, 1, 2, 3, \dots \} \quad (3.1.6)$$

It is common to append to the definition of sets (Def. 1.1.1) the requirement that a set cannot contain itself. That is, if  $A$  is a set, then  $A \notin A$ . This requirement was introduced to avoid paradoxes discovered by Bertrand Russell

in 1901. Allow us to neglect this requirement for a moment and reveal why it is essential. Recall from logic that a system of mathematics is inconsistent if one can prove a contradiction within the theory. In Naive Set Theory we allow the *axiom of unrestricted comprehension*. This allows us to construct sets as any definable collection. That is, if we have a proposition  $P$ , then we can define a set  $A$  as the set of all objects that satisfy  $P$ . We can write:

$$A = \{ x \mid P(x) \} \quad (3.1.7)$$

Problems with such a loose definition arise instantly. Let  $P$  be the proposition *true if  $x$  is a set, false otherwise*. Then  $A = \{ x \mid P(x) \}$  can be read in plain English as the *set of all sets*. A natural question would be whether or not  $A$  then contains itself. That is, is  $A \in A$ ? Russell's paradox arises by defining proper sets to be sets  $B$  such that  $B \notin B$ , and improper sets to be sets  $B$  such that  $B \in B$ . Using the *Law of the Excluded Middle* (which we will prove later), one has that every set is either proper or improper.

### Theorem 3.1.1: Russell's Paradox

Naive Set Theory is inconsistent.

*Proof.* For let  $P$  be the proposition *true if  $x \notin x$ , false otherwise*. Let  $A$  be the set defined by this proposition:

$$A = \{ x \mid P(x) \} \quad (3.1.8)$$

That is,  $A$  is the set of all sets that do not contain themselves. Suppose  $A \in A$ . If  $A \in A$  then  $P(A)$  is true. That is,  $A$  is a proper set. But proper sets do not contain themselves and  $A \in A$ , a contradiction. Thus  $A \notin A$ . But if  $A \notin A$  then  $P(A)$  is false. But if  $P(A)$  is false, than  $A$  is an improper set. But then  $A \in A$ , a contradiction as  $A \notin A$ . Thus  $A \in A$  if and only if  $A \notin A$ , a contradiction. Therefore, Naive Set Theory is inconsistent.  $\square$

Our development of Zermelo-Fraenkel Set Theory is to avoid this paradox and attempt to develop a consistent system of mathematics. The proof of Russell's Paradox (Thm. 3.1.1) relied on the *Law of the Excluded Middle* which states that, given a proposition  $P$ , either  $P$  is true or its negation is true. Thus we have shown that the axiom of unrestricted comprehension and the law of the excluded middle are not compatible. This is quite unfortunate as the law of the excluded middle is essential in mathematics as it allows one to prove things via contradiction. That is, given some statement we assume the opposite is true and arrive at a contradiction thus showing the negation of our statement is false. We then invoke the law of the excluded middle to show that our original statement is true. The axioms of Zermelo and Fraenkel, together with the

axiom of choice (a system commonly abbreviated as [ZFC](#)) are able to prove the validity of the law of the excluded middle. That is, if ZFC is consistent, then so is the law of the excluded middle. This is one of the reasons for studying ZFC in detail.

The first collection of axioms were proposed in 1908 by Ernst Zermelo. Subtle problems were pointed out by Abraham Fraenkel in 1920, and in 1921 the system of Zermelo-Fraenkel Set Theory came to be. The requirement that a set does not contain itself, which is equivalent to the *axiom of regularity*, is sufficient to avoid Russell's paradox. We will prove the equivalence of this axiom with our definition once we have obtained the law of the excluded middle.

### 3.1.1 Subsets and Equality

To delve more into set theory it would be convenient to know that at least *one* set exists. The axiom of the empty set gives us such an existence.

#### Axiom 3.1.1: Axiom of the Empty Set

There exists a set  $\emptyset$  (the empty set) such that for all  $x$  it is true that  $x \notin \emptyset$ .

$$\exists_{\emptyset} : \forall_x (\neg(x \in \emptyset))$$

The empty set is the set that contains no elements. As such some choose to write  $\emptyset = \{\}$ . Note that this is different from the set  $\{\emptyset\}$  since the empty set contains no elements whereas  $\{\emptyset\}$  contains one elements (it contains the empty set). Indeed, the equality of  $\emptyset$  and  $\{\emptyset\}$  would violate our requirement that sets do not contain themselves. Any set that contains *something* is called non-empty.

#### Definition 3.1.1: Non-Empty Set

A [non-empty set](#) is a [set](#)  $A$  such that there exists an  $x$  such that  $x \in A$ .

The terminology is somewhat redundant, and essentially every set we deal with is non-empty. Indeed, there is only one empty set (see Thm. [3.2.2](#)). Thus, every other set one thinks of is non-empty.

**Example 3.1.2** Using the notation from Ex. [3.1.1](#), the set of all natural numbers ( $\mathbb{N}$ ) and the set of all integers ( $\mathbb{Z}$ ) are non-empty since  $0 \in \mathbb{N}$  and  $0 \in \mathbb{Z}$ . If  $n$  is a positive integer, then the set of integers between  $0$  and  $n - 1$  ( $\mathbb{Z}_n$ ) is non-empty as well since  $0 \in \mathbb{Z}_n$ . Note that there is some ambiguity behind the

meaning of  $\mathbb{Z}_0$ . This stems from our *dot dot dot* definition and it is unclear what this should mean for  $n = 0$ . When we rigorously define this notation we will see that  $\mathbb{Z}_0$  is empty. That is, we will say  $k \in \mathbb{Z}_n$  if  $k \in \mathbb{N}$  and  $k < n$ , a description that will be justified by the axiom schema of specification. Thus, for  $\mathbb{Z}_0$  we seek an integer  $k \in \mathbb{N}$  such that  $k < 0$ . But there are no such integers, and thus the set is empty.

**Example 3.1.3** It's possible to write down some formula for a set that ultimately leads to the empty set. For consider the *set of all rational numbers whose square is two*. This set turns out to be empty since there is no rational that satisfies this criterion. That is,  $\sqrt{2}$  is known to be an irrational number. Thus, the set specified by our proposition is the empty set.

**Example 3.1.4** Going in the other direction, it is possible to write a formula for a set that appears empty, but is indeed not. The set of all  $p$ -Sylow subgroups of a non-empty finite group (Discussed in Book Two) is a non-empty set, but there's no reason to believe so from the start.

A set is entirely determined by its elements, and as such repetition and order cannot be accounted for. Thus the sets  $\{a, b\}$  and  $\{a, a, b\}$  must be considered the same since they contain precisely the same elements. This will be made clear once equality has been defined. In a similar manner, sets have no sense of order and thus  $\{a, b\}$  and  $\{b, a\}$  are equivalent. It then becomes a task to invent some new object that does have a notion of order. To do this requires the concept of a *function*, and it is our current aim to develop this topic.

To rigorously show that the examples in the previous paragraph are equal requires a definition of equality. This is the *axiom of extensionality*. First, we define the familiar symbol for equality ( $=$ ) in terms of containment ( $\in$ ).

### Notation 3.1.1: Equality

If  $A$  and  $B$  are sets, then  $A = B$  if and only if for all sets  $C$ ,  $C \in A$  if and only if  $C \in B$ , and for all sets  $D$ ,  $A \in D$  if and only if  $B \in D$ .

$$\forall_A \forall_B \left( (A = B) \iff (\forall_C (C \in A \iff C \in B) \wedge \forall_D (A \in D \iff B \in D)) \right)$$

**Example 3.1.5** Consider the set of all planets in the solar system, and consider the set of the eight largest objects in the solar system other than the sun. These two sets are equal since the eight largest objects (other than the sun) are the eight planets (sorry Pluto), and the set of planets form the eight largest objects. The tricky part is to check that for any set one can name, it is true that if the set of planets lies in the set, then the set of the eight largest objects not equal

to the sun lie in this set as well, and vice versa. This is almost impossible, and seemingly redundant, and so we rely on the *axiom of extensionality* to ease the demonstration of equality.

The axiom of extensionality says that to check for equality it suffices to show that for all  $C$ ,  $C \in A$  if and only if  $C \in B$ . That is, there is no need to check that for all  $D$ ,  $A \in D$  if and only if  $B \in D$ . For simplicity, the axiom of extensionality may be taken as the definition of equality.

### Axiom 3.1.2: Axiom of Extensionality

If  $A$  and  $B$  are sets, and if for all  $x$  it is true that  $x \in A$  if and only if  $x \in B$ , then  $A = B$ . That is,  $A$  and  $B$  are equal sets.

$$\forall_A \forall_B (\forall_x (x \in A \Leftrightarrow x \in B) \Leftrightarrow (A = B))$$

**Example 3.1.6** Returning to our example of planets, we have seen that the set of all planets and the set of the eight largest objects other than the sun contain precisely the same elements. By the axiom of extensionality, we thus have equality amongst these two.

**Example 3.1.7** Let  $\mathbb{R}$  denote the real numbers, let  $>$  and  $\leq$  denote the usual notions of *greater than* and *less than or equal to*, respectively. Define  $A$  and  $B$  by:

$$A = \{ x \in \mathbb{R} \mid x > 0 \} \quad (3.1.9a) \quad B = \{ y \in \mathbb{R} \mid y \leq 0 \} \quad (3.1.9b)$$

This notation will be justified by the specification axiom (see Ax. 3.1.3). Then for any real number  $x \in \mathbb{R}$ , we have that  $x \in A$  if and only if  $x > 0$ . That is,  $A$  is the set of all positive real numbers. But if  $x > 0$ , then  $x \neq 0$  and  $x$  is non-negative, so  $x \leq 0$ . Thus  $x$  satisfies the criterion for membership of  $B$ , and therefore  $x \in B$ . Similarly,  $y \in B$  if and only if  $y \leq 0$  and this is just another way of stating that  $y > 0$ , and hence  $y \in A$ . By the axiom of extensionality,  $A = B$ .

We'll restate the definition of equality using the language of subsets, lessening the effort required in proving various things are equal. The notions are equivalent. Subsets are sets that are defined in terms of another given set by simply removing some (or none, or all) of the elements.

**Definition 3.1.2: Subsets**

A **subset** of a **set**  $B$  is a set  $A$  such that for all  $x \in A$  it is true that  $x \in B$ . If  $A$  is a subset of  $B$  we write  $A \subseteq B$ . Otherwise, we write  $A \not\subseteq B$ .

$$\forall_A \forall_B ((A \subseteq B) \iff \forall_x (x \in A \Rightarrow x \in B))$$

We can often visualize sets as blobs in the plane. Using such a visual, we can envision subsets as well (Fig. 3.1). Given a blob  $B$ , a subset of  $B$  is another blob  $A$  that is entirely contained within  $B$ .

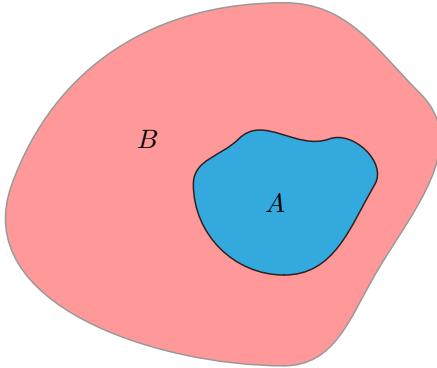


Fig. 3.1: Visualizing Subsets as Blobs

**Example 3.1.8** Consider the set of natural numbers  $\mathbb{N}$  and the set of integers  $\mathbb{Z}$  (loosely defined in Eqn. 3.1.3 and Eqn. 3.1.6, respectively). It can be seen that every natural number is also an integer, and thus we have:

$$\mathbb{N} \subseteq \mathbb{Z} \tag{3.1.10}$$

Letting  $\mathbb{Q}$  denote the rational numbers  $p/q$ , where  $p, q \in \mathbb{Z}$  and  $q$  is non-zero, we can see that  $\mathbb{Q}$  contains  $\mathbb{Z}$  as a subset. That is, setting  $q = 1$  and allowing  $p$  to vary over  $\mathbb{Z}$  gives us every integer. Thus:

$$\mathbb{Z} \subseteq \mathbb{Q} \tag{3.1.11}$$

We can continue with the real numbers ( $\mathbb{R}$ ) and the complex numbers ( $\mathbb{C}$ ) as well, creating a chain of subsets:

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C} \tag{3.1.12}$$

**Example 3.1.9** Let  $\mathbb{Z}_n$  be the set of all integers  $k \in \mathbb{N}$  such that  $k < n$ . If  $m, n \in \mathbb{N}$  and  $m < n$  we see that:

$$\mathbb{Z}_m \subseteq \mathbb{Z}_n \tag{3.1.13}$$

This is because  $\mathbb{Z}_m$  is the set of all  $k \in \mathbb{N}$  such that  $k < m$ . But since  $\mathbb{Z}_n$  is the set of all  $k \in \mathbb{N}$  such that  $k < n$ , and since  $m < n$ , if  $k \in \mathbb{Z}_m$  then  $k < m$ , and thus  $k < n$ , which implies that  $k \in \mathbb{Z}_n$ .

It is important to note the distinction between the symbols for containment ( $\in$ ) for subset ( $\subseteq$ ). The symbol  $\in$  is used to denote that some object  $x$  is an *element* of some set. That is,  $x \in A$  indicates that  $x$  is an element of  $A$ . This does not necessarily imply  $x \subseteq A$ , but this *does* imply that  $\{x\} \subseteq A$ . That is, if  $x \in A$ , then the set that contains only  $x$  is a subset of  $A$ . Moreover, the notions are not mutually exclusive. It is possible for  $A$  to be a set such that  $x \in A$  and  $x \subseteq A$ . For let  $A = \{\emptyset\}$ . For any set  $A$  it is true that  $\emptyset \subseteq A$  (see Thm. 3.2.1). But from how  $A$  is defined, we have that  $\emptyset \in A$ . Thus it is true that both  $\emptyset \in A$  and  $\emptyset \subseteq A$ .

### Example 3.1.10: Elementary Examples of Subsets

Let  $A$  and  $B$  be the sets defined by:

$$A = \{a, b, c\} \quad (3.1.14a) \qquad B = \{a, b, c, d\} \quad (3.1.14b)$$

where we assume that  $a$ ,  $b$ ,  $c$ , and  $d$  are distinct objects. From the definition of subsets (Def. 3.1.2):

$$A \subseteq B \quad (3.1.15a) \qquad B \not\subseteq A \quad (3.1.15b)$$

This is true since from the definition of  $A$  and  $B$ , every element of  $A$  is also an element of  $B$ . The converse of this is not true since there is an element of  $B$  that is not an element of  $A$  (namely, the element  $d$ ). That is,  $d \in B$  but  $d \notin A$  and therefore  $B \not\subseteq A$ .

The example shown in Ex. 3.1.10 hints at how we can redefine equality of sets. We see that  $A \subseteq B$ , but  $B \not\subseteq A$ . If we have two sets  $A$  and  $B$  such that  $A \subseteq B$  and  $B \subseteq A$ , then it would be impossible to discern between the two. This gives us our new definition of equality. We now prove this equivalence with the axiom of extensionality (Ax. 3.1.2).

**Theorem 3.1.2.** *If  $A$  and  $B$  are sets, then  $A = B$  if and only if  $A \subseteq B$  and  $B \subseteq A$ .*

*Proof.* By the axiom of extensionality (Ax. 3.1.2),  $A = B$  if and only if for all  $x$  it is true that  $x \in A$  if and only if  $x \in B$ . But then  $x \in A$  implies that  $x \in B$ , and thus  $A \subseteq B$  (Def. 3.1.2). But also  $x \in B$  implies  $x \in A$ , and therefore  $B \subseteq A$ . Therefore if  $A = B$ , then  $A \subseteq B$  and  $B \subseteq A$ . Now if  $A \subseteq B$  and  $B \subseteq A$ , then for all  $x \in A$  it is true that  $x \in B$  and for all  $x \in B$  it is true that

$x \in A$  (Def. 3.1.2), and therefore  $x \in A$  if and only if  $x \in B$ . Thus if  $A \subseteq B$  and  $B \subseteq A$ , then  $A = B$ . But it was just proved that if  $A = B$ , then  $A \subseteq B$  and  $B \subseteq A$ . Therefore  $A = B$  if and only if  $A \subseteq B$  and  $B \subseteq A$ .  $\square$

With this, we can redefine the notion of *equal sets*.

### Definition 3.1.3: Equal Sets

Equal sets are sets  $A$  and  $B$ , denoted  $A = B$ , such that  $A \subseteq B$  and  $B \subseteq A$ .

$$\forall_A \forall_B ((A = B) \iff ((A \subseteq B) \wedge (B \subseteq A)))$$

Def. 3.1.3 is justified by Thm. 3.1.2, and thus there is no contradiction with the axiom of extensionality (Ax. 3.1.2). If  $A$  and  $B$  are not equal, we write  $A \neq B$ .

### Example 3.1.11: More Examples of Subsets

Using the notation from Ex. 3.1.1, for all  $n \in \mathbb{N}$  we have:

$$\mathbb{Z}_n \subseteq \mathbb{N} \quad (3.1.16)$$

Let's define  $\mathbb{N}_e$  and  $\mathbb{N}_o$  to be the sets of even and odd non-negative integers, respectively:

$$\mathbb{N}_e = \{0, 2, 4, 6, 8, \dots\} \quad (3.1.17a) \quad \mathbb{N}_o = \{1, 3, 5, 7, 9, \dots\} \quad (3.1.17b)$$

From this we see the following two expressions are true:

$$\mathbb{N}_o \subseteq \mathbb{N} \quad (3.1.18a) \quad \mathbb{N}_e = \mathbb{N} \quad (3.1.18b)$$

Moreover we see that  $\mathbb{N}_o$  and  $\mathbb{N}_e$  have no elements in common. That is, they are *disjoint*. From this, we can write:

$$\mathbb{N}_o \not\subseteq \mathbb{N}_e \quad (3.1.19a) \quad \mathbb{N}_e \not\subseteq \mathbb{N}_o \quad (3.1.19b)$$

We can also think of trivial examples:

$$\mathbb{Z}_3 \subseteq \mathbb{Z}_4 \quad (3.1.20a) \quad \mathbb{Z}_4 \not\subseteq \mathbb{Z}_3 \quad (3.1.20b)$$

This is because every element of  $\mathbb{Z}_3$  is contained in  $\mathbb{Z}_4$ , but  $3 \in \mathbb{Z}_4$  and  $3 \notin \mathbb{Z}_3$ .  $\blacksquare$

It may seem like bad notation to write  $3 \notin \mathbb{Z}_3$ , but since we want  $\mathbb{Z}_n$  to have

$n$  elements, and since we started counting at zero, we have that  $n \notin \mathbb{Z}_n$  for all  $n \in \mathbb{N}$ . Such counting schemes are common in computer science, but there's disagreement in mathematics as to whether  $0 \in \mathbb{N}$  or not. We will use the *axiom of infinity* to prove the existence of  $\mathbb{N}$ , and in doing so it will be natural to define  $\mathbb{N}$  as a set that contains 0.

While Def. 3.1.3 is indeed equivalent to the axiom of extensionality, this definition creates a few problems. As discussed previously, sets have no notion of order and cannot account for repetition. For let  $A$ ,  $B$ , and  $C$  be the sets defined by:

$$A = \{a, b\} \quad (3.1.21a) \quad B = \{a, a, b\} \quad (3.1.21b) \quad C = \{b, a\} \quad (3.1.21c)$$

All three of these sets are equal by both the definition of equality (Def. 3.1.3) and the axiom of extensionality. It seems clear that  $A \subseteq B$ , but it is also true that  $B \subseteq A$ . This is because  $B$  contains only the elements  $a$  and  $b$ . While  $a$  is included twice, repetition cannot be accounted for and  $B$  is entirely determined by  $a$  and  $b$ . But  $A$  also contains  $a$  and  $b$ , and therefore  $B \subseteq A$ . By the definition of equality (Def. 3.1.3), we have that  $A = B$ . In a similar manner,  $A = C$ . From the definition of subsets, for any set  $A$  we see that  $A \subseteq A$  (see Thm. 3.2.6). It would be nice to distinguish between subsets that aren't the entire set itself. These are called proper subsets, and we can define them in terms of equality.

#### Definition 3.1.4: Proper Subsets

A **proper subset** of a **set**  $B$  is a set  $A$  such that  $A \subseteq B$  and  $A \neq B$ . We write  $A \subsetneq B$  to denote that  $A$  is a proper subset of  $B$ .

$$\forall_A \forall_B (A \subsetneq B) \iff ((A \subseteq B) \wedge (A \neq B))$$

The symbols  $\subseteq$  and  $\subsetneq$  are analogous to the notations of inequalities that one finds in calculus:  $\leq$  and  $<$ . In many texts, the two symbols  $\subseteq$  and  $\subset$  are taken to be identical, which may cause confusion. In an attempt to reduce confusion,  $\subseteq$  will denote any subset,  $\subsetneq$  denotes a proper subset, and the symbol  $\subset$  will be avoided.

#### Example 3.1.12: Proper Subsets

Let  $A$  and  $B$  be sets defined as follows:

$$A = \{a, b, c\} \quad (3.1.22a) \quad B = \{a, b, c, d\} \quad (3.1.22b)$$

Then  $A \subseteq B$ , since every element of  $A$  is an element of  $B$ , but  $B \not\subseteq A$  since  $d \in B$  and  $d \notin A$ . Therefore  $A \neq B$ , and thus  $A$  is a proper subset of  $B$ . We can denote this by writing  $A \subsetneq B$ . ■

Proper subsets are subsets that are missing at least one element (see Thm. 3.2.12). Returning to our claim that  $\emptyset \subseteq A$  for any set  $A$ , for any non-empty set  $A$  we have that  $\emptyset \subsetneq A$ . This is because for non-empty sets there is at least one  $x$  such that  $x \in A$  (Def. 3.1.1), whereas for all  $x$  it is true that  $x \notin \emptyset$ . Thus equality cannot occur, and the empty set must be a proper subset. It is also true that the empty set contains no proper subsets. The only subset of  $\emptyset$  is itself.

**Example 3.1.13** Returning to more concrete examples,  $\mathbb{N}$  is a proper subset of  $\mathbb{Z}$ . To see this, note that  $-1 \in \mathbb{Z}$  but  $-1 \notin \mathbb{N}$ . Indeed, none of the negative integers are natural numbers, but they are integers. We can write this by:

$$\mathbb{N} \subsetneq \mathbb{Z} \quad (3.1.23)$$

Similarly,  $\mathbb{Q}$  contains numbers that are not integers, for example  $1/2$ . Thus,  $\mathbb{Z}$  is also a proper subset of  $\mathbb{Q}$ . Lastly, since  $\sqrt{2}$  is not a rational number, the set of rational numbers must then be a proper subset of the set of real numbers.

We now introduce the *axiom schema of specification*.

### Axiom 3.1.3: Axiom Schema of Specification

If  $A$  is a set and if  $P$  is a proposition, then there exists a set  $B$  such that  $x \in B$  if and only if  $x \in A$  and  $P(x)$  is true. We can write this as:

$$B = \{ x \in A \mid P(x) \}$$

Using our formal language, we have:

$$\forall_A \forall_P \exists_B : \forall_x ((x \in B) \Leftrightarrow ((x \in A) \wedge P(x)))$$

Ax. 3.1.3 is different from the inconsistent axiom of unrestricted comprehension in that we can only speak of elements that are already defined and contained in some other set. That is, this new axiom does not allow us to talk about the *set of all sets*, and so we have avoided the crux of Russell's paradox.

This allows us to use the Set-Builder method of constructing sets. We described the natural numbers  $\mathbb{N}$  and integers  $\mathbb{Z}$  (From the German *Zahl*) using Eqns. 3.1.3 and 3.1.6, respectively. It would be more difficult (but not impossible) to describe the set of rational numbers in such a way. Instead, we use

set-builder notation if it is known that  $\mathbb{Q}$  is contained in some larger set  $\mathbb{R}$  (the *real* numbers).

$$\mathbb{Q} = \left\{ \frac{p}{q} \in \mathbb{R} \mid p, q \in \mathbb{Z} \text{ and } q \neq 0 \right\} \quad (3.1.24)$$

That is, the rational numbers are the set of all real numbers which can be written as the ratios of integers with non-zero denominator. The Axiom Schema of Specification states that this is a valid method of describing sets. It is also known as the axiom of separation.

**Example 3.1.14** We can describe the sets  $\mathbb{Z}$ ,  $\mathbb{N}$ ,  $\mathbb{N}_e$ , and  $\mathbb{N}_o$  using set-builder notation if we assume these belong to some larger set  $\mathbb{R}$ . We define  $\mathbb{Z}$  by:

$$\mathbb{Z} = \{ n \in \mathbb{R} \mid n \text{ is an integer} \} \quad (3.1.25)$$

From here we can define  $\mathbb{N}$  by:

$$\mathbb{N} = \{ n \in \mathbb{Z} \mid n \geq 0 \} \quad (3.1.26)$$

Furthermore,  $\mathbb{N}_e$  and  $\mathbb{N}_o$  can be described as follows:

$$\mathbb{N}_e = \{ n \in \mathbb{N} \mid n \text{ is even} \} \quad (3.1.27a) \quad \mathbb{N}_o = \{ n \in \mathbb{N} \mid n \text{ is odd} \} \quad (3.1.27b)$$

Such notation is justified by the axiom schema of specification.

We are not adopting these definitions since they lack rigor. These examples build intuition behind the notation and the axioms, but we will develop arithmetic axiomatically using the *axiom of infinity*.

**Example 3.1.15** Concrete examples of set builder notation can be made if we examine propositions that lead to finite sets. Consider the set  $A$  defined by:

$$S = \{ n \in \mathbb{N} \mid \text{There exists } k \in \mathbb{N} \text{ such that } k < 5 \text{ and } n = 7k - 3 \} \quad (3.1.28)$$

It may be easier if we rewrite this as follows:

$$S = \{ 7k - 3 \mid k \in \mathbb{N} \text{ and } k < 5 \} \quad (3.1.29)$$

So we just need to loop through this equation from  $k = 0$  to  $k = 4$ . We obtain:

$$S = \{ -3, 4, 11, 18, 24 \} \quad (3.1.30)$$

### 3.1.2 Ordered Pairs and Unions

We now wish to solve the issue previously raised that sets do not have order. We'll develop a new object, called ordered pairs, that can distinguish such

things. The definition we'll adopt is due to Kuratowski and defines  $(a, b)$  as follows:

$$(a, b) = \{ \{ a \}, \{ a, b \} \} \quad (3.1.31)$$

We now prove such a set exists within the framework of [ZFC](#).

#### Axiom 3.1.4: Axiom of Pairing

If  $A$  and  $B$  are sets, then there exists a set  $\mathcal{C}$  such that  $A \in \mathcal{C}$  and  $B \in \mathcal{C}$ .

$$\forall_A \forall_B \exists_{\mathcal{C}} : ((A \in \mathcal{C}) \wedge (B \in \mathcal{C}))$$

The set hypothesized to exist in this axiom may be very large, we have no way of knowing. What we want from this is a set that contains two elements  $A$  and  $B$ , and only those elements. We obtain this by combining pairing with specification.

**Theorem 3.1.3.** *If  $A$  and  $B$  are sets, then there exists a set  $D$  such that for all  $x$  it is true that  $x \in D$  if and only if  $x = A$  or  $x = B$ . That is:*

$$D = \{ A, B \} \quad (3.1.32)$$

*Proof.* By the axiom of pairing (Ax. 3.1.4) there exists a set  $\mathcal{C}$  such that  $A \in \mathcal{C}$  and  $B \in \mathcal{C}$ . Let  $P$  be the proposition *true if  $x = A$  or  $x = B$ , false otherwise*. By the axiom schema of specification (Ax. 3.1.3), there is a set  $D$  such that:

$$D = \{ x \in \mathcal{C} \mid P(x) \} \quad (3.1.33)$$

That is,  $x \in D$  if and only if  $x \in \mathcal{C}$  and  $P(x)$  is true. But then  $x \in D$  if and only if  $x \in \mathcal{C}$  and  $x = A$  or  $x \in \mathcal{C}$  and  $x = B$ . But  $A \in \mathcal{C}$  and  $B \in \mathcal{C}$ , and thus  $P(x)$  implies  $x \in \mathcal{C}$ . Thus,  $x \in D$  if and only if  $P(x)$  is true. That is,  $x \in D$  if and only if  $x = A$  or  $x = B$ .  $\square$

By the axiom of extensionality (Ax. 3.1.2), the set hypothesized in Thm. 3.1.3 is unique, and thus there is no trouble in *defining* the symbol  $\{A, B\}$  to be the unique set that contains the elements  $A$  and  $B$  and only those elements. That is, we develop the new notation:

#### Notation 3.1.2: Finite Set Notation

If  $A$  and  $B$  are sets, then  $\{A, B\}$  is the unique set such that for all  $x$ ,  $x \in \{A, B\}$  if and only if  $x = A$  or  $x = B$ .

$$\forall_x \left( (x \in \{A, B\}) \iff ((x = A) \vee (x = B)) \right)$$

**Theorem 3.1.4.** *If  $A$  is a set, then there is a set  $B$  such that  $x \in B$  if and only if  $x = A$ . That is, there exists a set  $B$  such that:*

$$B = \{A\} \quad (3.1.34)$$

*Proof.* For since  $A$  is a set, by Thm. 3.1.3 there exists a set  $B = \{A, A\}$ . But then  $x \in B$  if and only if  $x = A$ .  $\square$

We can apply Not. 3.1.2 to a single set  $A$  and similarly define what the notation  $\{A\}$  means. With this, we can now prove the existence of ordered pairs.

### Theorem 3.1.5: Existence of Ordered Pairs

If  $A$  and  $B$  are sets, then there is a set  $(A, B)$  such that for all  $x$  it is true that  $x \in (A, B)$  if and only if  $x = \{A\}$  or  $x = \{A, B\}$ .  $\blacksquare$

*Proof.* For by Thm. 3.1.4, there is a set  $\{A\}$  such that  $x \in \{A\}$  if and only if  $x = A$ . But by Thm. 3.1.3, there is a set  $\{A, B\}$  such that  $x \in \{A, B\}$  and if and only if  $x = A$  or  $x = B$ . But again by Thm. 3.1.3, since  $\{A\}$  and  $\{A, B\}$  are sets, there is a set  $(A, B)$  such that  $x \in (A, B)$  if and only if  $x = \{A\}$  or  $x = \{A, B\}$ .  $\square$

#### Definition 3.1.5: Ordered Pairs

The **ordered pair** of a **set**  $x$  with respect to a set  $y$  is the set:

$$(x, y) = \{\{x\}, \{x, y\}\}$$

Using our formal language:

$$\forall_x \forall_y \forall_z ((z \in (x, y)) \iff ((z = \{x\}) \vee (z = \{x, y\})))$$

Thm. 3.1.5 asserts the existence of ordered pairs, as defined by Kuratowski, and allows us to present Def. 3.1.5 in a way that is consistent with ZFC. Kuratowski first put forward this definition in 1921 and it does precisely what we want it to do and orders elements. That is, if  $x$  and  $y$  are distinct, then  $(x, y) \neq (y, x)$ . The caveat with this definition is the following reduction:

$$(x, x) = \{\{x\}, \{x, x\}\} = \{\{x\}, \{x\}\} = \{\{x\}\} \quad (3.1.35)$$

Prior to Kuratowski there existed a definition due to Norbert Wiener, put forward in 1914. His definition grew out of Bertrand Russell's Type Theory

which was an attempt to rid set theory of the paradoxes he discovered. Wiener writes:

$$(x, y)_W = \left\{ \{ \{ x \}, \emptyset \}, \{ \{ y \} \} \right\} \quad (3.1.36)$$

Returning to Kuratowski's definition (Def. 3.1.5), consider the ordered pair  $(1, 2)$ , where we take for granted that  $1 \neq 2$ . We have:

$$(1, 2) = \{ \{ 1 \}, \{ 1, 2 \} \} \quad (3.1.37)$$

Swapping and computing  $(2, 1)$ , we obtain:

$$(2, 1) = \{ \{ 2 \}, \{ 2, 1 \} \} \quad (3.1.38)$$

We know that sets cannot distinguish order, so  $\{ 1, 2 \} = \{ 2, 1 \}$ . Thus:

$$(1, 2) = \{ \{ 1 \}, \{ 1, 2 \} \} \quad (3.1.39a) \quad (2, 1) = \{ \{ 2 \}, \{ 1, 2 \} \} \quad (3.1.39b)$$

Combining these equations, we now have that:

$$(1, 2) \neq (2, 1) \quad (3.1.40)$$

To see this, note that both sets contain the element  $\{ 1, 2 \}$ , but  $\{ 1 \}$  is an element of  $(1, 2)$  and not an element of  $(2, 1)$ , and thus  $(1, 2) \not\subseteq (2, 1)$ . Similarly,  $\{ 2 \}$  is an element of  $(2, 1)$  but not an element of  $(1, 2)$ , and therefore  $(2, 1) \not\subseteq (1, 2)$ . From the definition of equality (Def. 3.1.3), we have that these sets are not equal.

There's is an unfortunate doubling of notation that occurs in mathematics, and  $(a, b)$  has two common meanings. The first meaning is the ordered pair which we've just defined, and the second is the *open interval* defined in the context of a *partially ordered set*. The most common example is when discussing the real numbers  $\mathbb{R}$ ,  $(a, b)$  denotes the set of all real numbers  $x$  such that  $a < x$  and  $x < b$ . Hopefully it will be clear what the notation means when a theorem or example is being presented, but we will be explicit when ambiguity can arise.

The natural thing from here is to construct the *Cartesian Product* of two sets. This is the set of all ordered pairs  $(a, b)$  where  $a$  belongs to some set  $A$  and  $b$  belongs to another set  $B$ . To prove such a set exists requires two more axioms.

**Axiom 3.1.5: Axiom of Union**

If  $\mathcal{O}$  is a set, then there exists a set  $\mathcal{F}$  such that, for all  $A$  such that  $A \in \mathcal{O}$  and for all  $x$  such that  $x \in A$ , it is true that  $x \in \mathcal{F}$ .

$$\forall_{\mathcal{O}} \exists_{\mathcal{F}} : \forall_x \left( (\exists_{A \in \mathcal{O}} : x \in A) \Rightarrow x \in \mathcal{F} \right)$$

This states that, given a collection of sets  $\mathcal{O}$ , there exists a larger set which contains the elements of the constituent sets of  $\mathcal{O}$ . Similar to the axiom of pairing,  $\mathcal{F}$  may be much larger than desired and we must invoke the axiom schema of specification to arrive at the *union* over a collection.

**Theorem 3.1.6: Existence of the Union of Sets**

If  $\mathcal{O}$  is a set, then there exists a set  $\bigcup \mathcal{O}$  such that for all  $x$  it is true that  $x \in \bigcup \mathcal{O}$  if and only if there is a set  $A \in \mathcal{O}$  such that  $x \in A$ . ■

*Proof.* For by the axiom of union (Ax. 3.1.5), there exists a set  $\mathcal{F}$  such that for all  $A \in \mathcal{O}$  and for all  $x \in A$  it is true that  $x \in \mathcal{F}$ . Let  $P$  be the proposition *true if there exists a set  $A \in \mathcal{O}$  such that  $x \in A$ , false otherwise*. Then, by the axiom schema of specification (Ax. 3.1.3) there exists a set  $\bigcup \mathcal{O}$  such that:

$$\bigcup \mathcal{O} = \{ x \in \mathcal{F} \mid P(x) \} \quad (3.1.41)$$

But then  $x \in \bigcup \mathcal{O}$  if and only if  $x \in \mathcal{F}$  and  $P(x)$  is true. But if  $P(x)$  is true, then  $x \in \mathcal{F}$ , and thus  $x \in \bigcup \mathcal{O}$  if and only if there is a set  $A \in \mathcal{O}$  such that  $x \in A$ . □

One question that arises is *what happens if our collection is empty?* That is, if  $\mathcal{O} = \emptyset$ , is there any meaning behind the equation:

$$\mathcal{F} = \bigcup \emptyset \quad (3.1.42)$$

There is, and  $\mathcal{F}$  will be the empty set. That is,  $\mathcal{F} = \emptyset$ . This is true in a vacuous sense and can be proved via contradiction with the law of the excluded middle. We define the set  $\mathcal{F}$  described in Thm. 3.1.6 as the *union* over the set  $\mathcal{O}$ . The set  $\mathcal{O}$  is often called the index set for which we take the union over.

**Definition 3.1.6: Union over a Set**

The [union over a set](#)  $\mathcal{O}$  is the set:

$$\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} = \{x \mid \text{There exists a set } \mathcal{U} \in \mathcal{O} \text{ such that } x \in \mathcal{U}\}$$

Using our formal language:

$$\forall_{\mathcal{O}} \forall_x \left( \left( x \in \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \right) \iff (\exists_{\mathcal{U} \in \mathcal{O}} : x \in \mathcal{U}) \right)$$

There are two ways to write unions for arbitrary collections, and we will make use of both depending on scenario. The first manner we have already seen in Thm. 3.1.6 where, given a collection  $\mathcal{O}$ , we wrote  $\bigcup \mathcal{O}$  to denote the union over  $\mathcal{O}$ . The second is depicted in Def. 3.1.6. That is, given  $\mathcal{O}$  we write:

$$\bigcup \mathcal{O} = \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \tag{3.1.43}$$

This alternative notation can be useful when we are using various indexing tricks to solve problems, or when combining unions with the various other set operations such as differences and intersections.

The notion of union is very convenient if we already have a collection of sets defined, but it would be nice to form the union over two given sets without considering them as part of a larger collection. This can be done by combining the axiom of union with pairing.

**Theorem 3.1.7.** *If  $A$  and  $B$  are sets, then there exists a set  $A \cup B$  such that  $x \in A \cup B$  if and only if either  $x \in A$  or  $x \in B$ .*

*Proof.* For by Thm. 3.1.3 there exists a set  $\mathcal{O}$  such that  $y \in \mathcal{O}$  if and only if  $y = A$  or  $y = B$ . That is,  $\mathcal{O} = \{A, B\}$ . But by Thm. 3.1.6 there exists a set  $A \cup B$  such that  $x \in A \cup B$  if and only if there exists a set  $\mathcal{U} \in \mathcal{O}$  such that  $x \in \mathcal{U}$ . But then  $x \in A \cup B$  if and only if either  $x \in A$  or  $x \in B$ .  $\square$

This allows us to define our first *operation* of two sets.

**Definition 3.1.7: Union of Two Sets**

The [union of two sets](#)  $A$  and  $B$  is the set  $A \cup B$  defined by:

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

That is:

$$\forall_A \forall_B \forall_x ((x \in A \cup B) \iff ((x \in A) \vee (x \in B)))$$

In our definition of the union over a collection and the union of two sets we have slightly abused our set-builder notation. The axiom schema of specification allows us to write a set as  $A = \{x \in B \mid P(x)\}$  given some set  $B$  that is already known to exists, and some proposition  $P$ . These two definitions (Defs. 3.1.6 and 3.1.7) are justified by the theorems we have proven, and so there is no contradiction.

**Example 3.1.16** Again using the notation found in Eqn. 3.1.5, if we let  $\mathbb{Z}_n$  denote the integers between 0 and  $n - 1$ , we have the following: If  $m$  is less than  $n$ , then:

$$\mathbb{Z}_m \cup \mathbb{Z}_n = \mathbb{Z}_n \quad (3.1.44)$$

This is because every element of  $\mathbb{Z}_m$  is already an element of  $\mathbb{Z}_n$ , and thus taking the union adds nothing new to  $\mathbb{Z}_n$ , so the resulting set is  $\mathbb{Z}_m$ .

**Example 3.1.17** Denoting the even and odd non-negative integers by  $\mathbb{N}_e$  and  $\mathbb{N}_o$ , respectively, we see that:

$$\mathbb{N}_e \cup \mathbb{N}_o = \mathbb{N} \quad (3.1.45)$$

This is because every non-negative integer  $n \in \mathbb{N}$  is either even or odd, and thus either  $n \in \mathbb{N}_e$  or  $n \in \mathbb{N}_o$ . Taking the union therefore gives the entire set  $\mathbb{N}$ . The union does not add anything more than  $\mathbb{N}$  since  $\mathbb{N}_e \subseteq \mathbb{N}$  and  $\mathbb{N}_o \subseteq \mathbb{N}$ .

### Example 3.1.18: Union of Two Sets

Let  $A$  and  $B$  be the sets defined by:

$$A = \{a, b, c\} \quad (3.1.46a) \qquad B = \{c, 1, 2\} \quad (3.1.46b)$$

The union of  $A$  and  $B$  is the set that contains all of the elements of  $A$  and all of the elements of  $B$ , and only such elements. That is:

$$A \cup B = \{a, b, c, 1, 2\} \quad (3.1.47)$$

Even though  $c \in A$  and  $c \in B$ ,  $c$  only appears once in the union. This is because sets cannot account for repetition, so including  $c$  twice would be redundant.

The union of two sets can again be visualized by considering blobs in the plane. Let  $A$  and  $B$  be two circles that overlap somewhere in the middle. The union  $A \cup B$  can then be represented by shading in the region covered by either  $A$  or  $B$  (see Fig. 3.2). Such a drawing is called a *Venn diagram*.

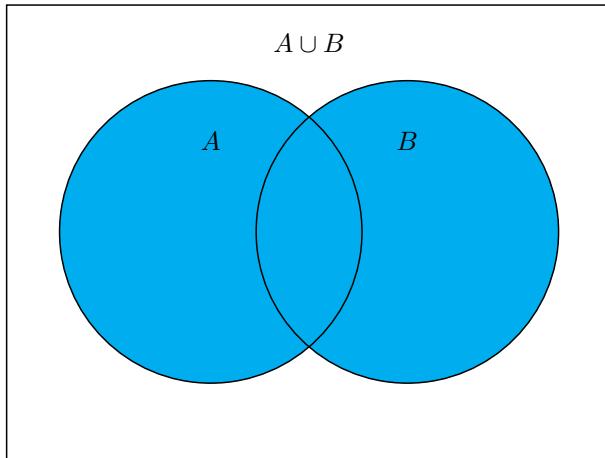


Fig. 3.2: Venn Diagram for Union

Fig. 3.2 can be extended to an arbitrary collection of sets. For the sake of simplicity, a Venn diagram for the union of three sets is shown in Fig. 3.3.

**Example 3.1.19** Suppose we have  $A = \{1, 3, 5\}$  and  $B = \{3, 5, 6, 7\}$ . We compute the union as follows:

$$A \cup B = \{1, 3, 5, 6, 7\} \quad (3.1.48)$$

Again, even though 3 and 5 occur in both  $A$  and  $B$ , sets have no notion of repetition so we need only include them once in the union.

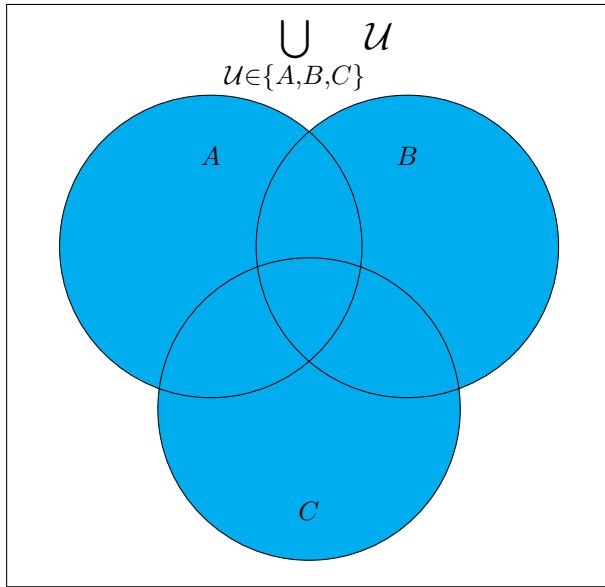


Fig. 3.3: The Union of Three Sets

We can combine the axiom schema of specification (Ax. 3.1.3) with the existence of the union of two sets to define intersections. The intersection of two sets, denoted  $A \cap B$ , is the set consisting of all elements that lie in both  $A$  and  $B$  simultaneously.

**Theorem 3.1.8.** *If  $A$  and  $B$  are sets, then there exists a set  $A \cap B$  such that for all  $x$  it is true that  $x \in A \cap B$  if and only if  $x \in A$  and  $x \in B$ .*

*Proof.* For by Thm. 3.1.7, there exists a set  $A \cup B$  such that for all  $x$  it is true that  $x \in A \cup B$  if and only if  $x \in A$  or  $x \in B$ . Let  $P$  be the proposition *True if  $x \in A$  and  $x \in B$ , false otherwise*. Then by the axiom schema of specification (Ax. 3.1.3) there is a set  $A \cap B$  such that:

$$A \cap B = \{ x \in A \cup B \mid P(x) \} \quad (3.1.49)$$

But then  $x \in A \cap B$  if and only if  $x \in A \cup B$  and  $x \in A$  and  $x \in B$ . But if  $x \in A$  and  $x \in B$ , then  $x \in A$ , and thus  $x \in A \cup B$  (Def. 3.1.7). That is,  $P(x)$  implies that  $x \in A \cap B$ . Therefore,  $x \in A \cap B$  if and only if  $P(x)$  is true. That is,  $x \in A \cap B$  if and only if  $x \in A$  and  $x \in B$ .  $\square$

**Definition 3.1.8: Intersection of Two Sets**

The [intersection of two sets](#)  $A$  and  $B$ , denoted  $A \cap B$ , is the set:

$$A \cap B = \{ x \in A \cup B \mid a \in A \text{ and } b \in B \}$$

That is:

$$\forall_A \forall_B \forall_x ((x \in A \cap B) \Leftrightarrow ((x \in A) \wedge (x \in B)))$$

**Example 3.1.20** Using Eqn. 3.1.5 to represent  $\mathbb{Z}_n$ , we can see that if  $m < n$  then:

$$\mathbb{Z}_m \cap \mathbb{Z}_n = \mathbb{Z}_m \quad (3.1.50)$$

Since  $m < n$ , every element of  $\mathbb{Z}_m$  is contained in  $\mathbb{Z}_n$ . But every element of  $\mathbb{Z}_n$  that is not contained in  $\mathbb{Z}_m$  will not be in the intersection. This is the opposite of the pattern we saw in Ex. 3.1.16 when we considered the union of  $\mathbb{Z}_m$  and  $\mathbb{Z}_n$ . This spells out a general theorem: If  $A \subseteq B$ , then  $A \cap B = A$  and  $A \cup B = B$  (see Thm. 3.2.37 and 3.2.20, respectively).

**Example 3.1.21: Intersections of Two Sets**

If we let  $A$  and  $B$  be the sets defined by:

$$A = \{ a, b, c \} \quad (3.1.51a) \qquad B = \{ c, 1, 2 \} \quad (3.1.51b)$$

we have that the intersection is then:

$$A \cap B = \{ 1 \} \quad (3.1.52)$$

This is because 1 is the only element that appears in both  $A$  and  $B$ , and is hence the only member of  $A \cap B$ .

**Example 3.1.22** Recalling our comment in Ex. 3.1.11, we claimed that the set of even integers ( $\mathbb{N}_e$ ) and the set of odd integers ( $\mathbb{N}_o$ ) are *disjoint*. We can now be precise about what this means. Since even numbers are of the form  $2n$  and odd numbers are of the form  $2n + 1$ , there are no natural numbers  $k \in \mathbb{N}$  that are both even and odd. Thus:

$$\mathbb{N}_e \cap \mathbb{N}_o = \emptyset \quad (3.1.53)$$

This is our definition of disjoint sets: Those with empty intersection.

**Definition 3.1.9: Disjoint Sets**

Disjoint sets are sets  $A$  and  $B$  such that  $A \cap B = \emptyset$ .

We'll need one brief theorem about intersections to allow us to prove that certain sets are not equal.

**Theorem 3.1.9.** *If  $A$  and  $B$  are sets, if  $x \in B$ , and if  $x \notin A \cap B$ , then  $x \notin A$ .*

*Proof.* For if  $x \notin A \cap B$  then either  $x \notin A$  or  $x \notin B$  (Def. 3.1.8). But  $x \in B$ , and therefore  $x \notin A$ .  $\square$

Similar to how unions can be visualized with Venn diagrams (Fig. 3.2), so can the intersection of two sets. We draw two circles that overlap slightly, and consider the region contained in both (see Fig. 3.4).

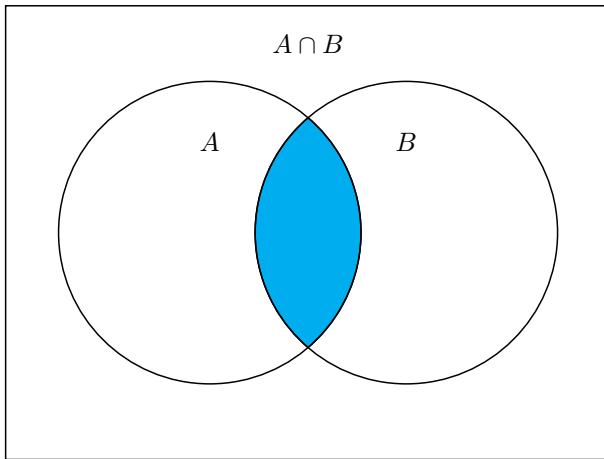


Fig. 3.4: Venn Diagram for Intersection

We can extend this further and define the intersection over any collection of sets.

**Theorem 3.1.10: Existence of the Intersection of Sets**

If  $\mathcal{O}$  is a set, then there exists a set  $\bigcap \mathcal{O}$  such that for all  $x$  it is true that  $x \in \bigcap \mathcal{O}$  if and only if  $x \in \bigcup \mathcal{O}$  and for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $x \in \mathcal{U}$ .  $\blacksquare$

*Proof.* For by Thm. 3.1.6 there is a set  $\bigcup \mathcal{O}$  such that for all  $x$  it is true that  $x \in \bigcup \mathcal{O}$  if and only if there exists a  $\mathcal{U} \in \mathcal{O}$  such that  $x \in \mathcal{U}$ . Let  $P$  be the

proposition *True if for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $x \in \mathcal{U}$ , false otherwise.* Then by the axiom schema of specification (Ax. 3.1.3), there exists the set:

$$\bigcap \mathcal{O} = \left\{ x \in \bigcup \mathcal{O} \mid P(x) \right\} \quad (3.1.54)$$

But then  $x \in \bigcap \mathcal{O}$  if and only if  $x \in \bigcup \mathcal{O}$  and  $P(x)$  is true. That is,  $x \in \bigcap \mathcal{O}$  if and only if  $x \in \bigcup \mathcal{O}$  and for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $x \in \mathcal{U}$ .  $\square$

It is common to consider some *universal* set, of which all other sets of current consideration are drawn from. Using this the definition of arbitrary intersection is defined as the subset of this universal set such that every element of this subset is contained in every element of the arbitrary collection. One may then ask what would happen if the collection is empty. Using this definition the intersection would be the entire universal set in a vacuous sense. That is, there would be no  $x$  in the universe that fails the definition of the intersection over an empty collection, and thus the intersection is everything. Letting  $X$  denote our universe, we obtain:

$$\emptyset = \bigcup_{\mathcal{U} \in \emptyset} \mathcal{U} \subseteq \bigcap_{\mathcal{U} \in \emptyset} \mathcal{U} = X \quad (3.1.55)$$

It seems like unions should always be bigger. Indeed, for any non-empty collection, the intersection over the collection is a subset of the union over the collection. Because of this we do not adopt this definition of the intersection over a collection, but rather require in our construction the use of the union over the collection, and then use the axiom schema of specification to pick the subset of all elements of the union that belong to every element of the collection. Thus:

$$\bigcap_{\mathcal{U} \in \emptyset} \mathcal{U} \subseteq \bigcup_{\mathcal{U} \in \emptyset} \mathcal{U} = \emptyset \quad (3.1.56)$$

And from this we conclude the intersection is empty as well.

### Definition 3.1.10: Intersection Over a Collection

The **intersection over a set**  $\mathcal{O}$  of sets is the set  $\bigcap \mathcal{O}$  defined by:

$$\bigcap_{\mathcal{U} \in \mathcal{O}} \mathcal{U} = \left\{ x \in \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \mid x \in \mathcal{U} \text{ for all } \mathcal{U} \in \mathcal{O} \right\}$$

That is:

$$\forall \mathcal{O} \forall x \left( \left( x \in \bigcap_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \right) \iff \left( (x \in \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}) \wedge (\forall_{\mathcal{U} \in \mathcal{O}} (x \in \mathcal{U})) \right) \right)$$

We can extend our Venn diagram for larger collections as well (see Fig. 3.5).

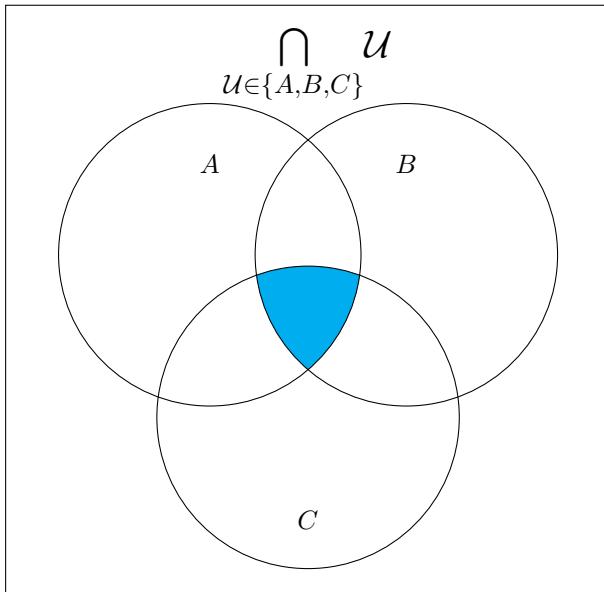


Fig. 3.5: The Intersection of Three Sets

Much the way we've defined what it means for two sets to be disjoint, we can extend this to arbitrary collections. We'll say that a collection of sets is mutually disjoint if all distinct elements of the collection have empty intersection. That is, all distinct pairs are disjoint.

#### Definition 3.1.11: Mutually Disjoint Collection of Sets

A mutually disjoint collection of sets is a set  $\mathcal{O}$  such that for all  $A, B \in \mathcal{O}$  such that  $A \neq B$  it is true that  $A \cap B = \emptyset$ . That is, for all distinct elements  $A, B \in \mathcal{O}$  it is true that  $A$  and  $B$  are disjoint.

The term pairwise disjoint is also frequently used in measure theory and probability. As such, one might guess that the notion has a fair amount of use in these subjects. Next on the list of axioms is that of *regularity*.

**Axiom 3.1.6: Axiom of Regularity**

If  $A$  is a non-empty set, then there is an element  $B \in A$  such that  $A \cap B = \emptyset$ .

$$\forall_A (\exists_{x \in A}) \Rightarrow \exists_y : ((y \in A) \wedge ((y \cap A) = \emptyset))$$

This axiom is often seen as unnecessary by many working mathematicians and indeed it's use seems to only lie in set theory and foundations. That is, unlike the axioms of choice and union which are widely applicable to analysis and topology, regularity seems to only be useful to set theorists. This is not entirely true if one really pays attention to the details. Often we are presented some object  $X$ , perhaps a topological space, or an algebraic structure like a group, and we need to extend this object to a larger collection. One concrete example comes from topology when we have a *non-compact* space  $(X, \tau)$  and we want to find a *compact* space  $(\tilde{X}, \tilde{\tau})$  that contains our original space. In the construction we add a single point to  $(X, \tau)$  called *infinity*, and label it  $\infty$ . But what if the symbol  $\infty$  already belongs to  $X$ ? How do we find a new object that is guaranteed not to lie in  $X$ ? The axiom of regularity allows us to show that  $\{\infty\}$  does not lie in  $X$ , for any set  $X$ , and thus we can take this to be our new point. This, and many similar constructions, rely on the axiom of regularity to guarantee that our reasoning is not ultimately circular.

Regardless of the axioms application, its existence is vital to support the claim that ZFC is a good system to base mathematics on. We will combine this with pairing to prove that for any set  $A$  it is true that  $A \notin A$ . That is, Zermelo-Fraenkel set theory is free of Russell's paradox.

**Theorem 3.1.11.** *If  $A$  is a set, then  $A \notin A$ .*

*Proof.* For if  $A$  is a set, then  $\{A\}$  is a set (Thm. 3.1.4). But since  $A \in \{A\}$ ,  $\{A\}$  is a non-empty set (Def. 3.1.1). Thus by the axiom of regularity (Ax. 3.1.6) there is a set  $B \in \{A\}$  such that  $B \cap \{A\} = \emptyset$ . But  $B \in \{A\}$  if and only if  $B = A$ , and therefore  $A \cap \{A\} = \emptyset$ . Thus, by the axiom of the empty set (Ax. 3.1.1), for all  $x$  it is true that  $x \notin A \cap \{A\}$  and therefore  $A \notin A \cap \{A\}$ . But  $A \in \{A\}$  and therefore  $A \notin A$  (Thm. 3.1.9).  $\square$

**Theorem 3.1.12.** *If  $A$  and  $B$  are sets and if  $A \in B$ , then  $A \neq B$ .*

*Proof.* For  $A \notin A$  (Thm. 3.1.11) and  $A \in B$  and therefore it is not true that for all  $x$ ,  $x \in A$  if and only if  $x \in B$ . Therefore, by the Axiom of Extensionality (Ax. 3.1.2),  $A \neq B$ .  $\square$

**Theorem 3.1.13.** *If  $A$  is a set, then  $A \neq \{A\}$ .*

*Proof.* For if  $A$  is a set, then  $A \notin A$  (Thm. 3.1.11). But if  $A$  is a set, then  $\{A\}$  is a set (Thm. 3.1.4). But  $A \in \{A\}$ , and thus  $A \neq \{A\}$  (Thm. 3.1.12).  $\square$

These quick theorems will eventually prove the well known result that  $0 \neq 1$ . It also shows us that there is no set of all sets.

### 3.1.3 The Axiom of the Power Set

Continuing in our goal of constructing order, we move on to the Cartesian product of two sets  $A$  and  $B$ . This is the collection of all ordered pairs  $(a, b)$  such that  $a \in A$  and  $b \in B$ . To prove such a set exists requires the *axiom of the power set*.

#### Axiom 3.1.7: Axiom of the Power Set

If  $A$  is a set, then there exists a set  $\mathcal{P}$  such that for all  $x \subseteq A$  it is true that  $x \in \mathcal{P}$ .

$$\forall_A \exists_{\mathcal{P}} : \forall_x ((x \subseteq A) \Rightarrow (x \in \mathcal{P}))$$

Again, much like the axiom of union and the axiom of pairing, this set may be bigger than we would like. We wish to find a set, called the *power set*, that contains all of the subsets of a given set  $A$  and nothing else. Combining the axiom of the power set with the axiom schema of specification gives us such existence.

#### Theorem 3.1.14: Existence of the Power Set

If  $A$  is a set, then there exists a set  $\mathcal{P}(A)$  such that for all  $x$  it is true that  $x \in \mathcal{P}(A)$  if and only if  $x \subseteq A$ .  $\blacksquare$

*Proof.* For by the axiom of the power set (Ax. 3.1.7) there is a set  $\mathcal{P}$  such that for all  $x \subseteq A$  it is true that  $x \in \mathcal{P}$ . Let  $P$  be the proposition *true if*  $x \subseteq A$ , *false otherwise*. By the axiom schema of specification (Ax. 3.1.3), there is a set  $\mathcal{P}(A)$  such that:

$$\mathcal{P}(A) = \{ x \in \mathcal{P}(A) \mid P(x) \} \quad (3.1.57)$$

But if  $P(x)$  is true, then  $x$  is a subset of  $A$ , and therefore  $x \in \mathcal{P}(A)$ . Thus  $x \in \mathcal{P}(A)$  if and only if  $x \subseteq A$ .  $\blacksquare$

With this we now define the *power set* of a given set.

**Definition 3.1.12: Power Set**

The **power set** of a **set**  $A$  is the set  $\mathcal{P}(A)$  defined by:

$$\mathcal{P}(A) = \{ x \mid x \subseteq A \}$$

That is, the set of all subsets of  $X$ .

$$\forall_A \forall_B (B \in \mathcal{P}(A) \iff B \subseteq A)$$

Again, there is some abuse of our set-builder notation, but Thm. 3.1.14 justifies such a definition. The power set of a set is a crucial construction for when one discusses the *cardinality* of sets, denoted  $\text{Card}(A)$ . This describes the *size* of a set in a very precise manner. A theorem that will eventually be proved known as *Cantor's Theorem* shows that the power set of a set is always strictly *larger* than the original set. That is:

$$\text{Card}(A) < \text{Card}(\mathcal{P}(A)) \quad (3.1.58)$$

This will be made precise soon enough. The axiom of the power set allows us to build *larger* sets from a given set. This creates a paradoxical heirarchy of infinities. Starting with the *smallest* infinity, the natural numbers  $\mathbb{N}$ , we can create a significantly larger set by considering  $\mathcal{P}(\mathbb{N})$ . We can continue and consider  $\mathcal{P}(\mathcal{P}(\mathbb{N}))$ , and there's no reason to stop there. At each step we create a new, massively larger set. This is both unintuitive and paradoxical and as such some may choose to reject it. This axiom is vital in the discussion of topology and measure theory, and so we choose to accept it as true.

**Example 3.1.23** If  $A = \{1, 2\}$ , then the power set is:

$$\mathcal{P}(A) = \{ \emptyset, \{1\}, \{2\}, \{1, 2\} \} \quad (3.1.59)$$

We must consider the empty set since  $\emptyset \subseteq A$ . Now suppose  $A = \{1, 2, 3\}$ :

$$\mathcal{P}(\{1, 2, 3\}) = \{ \emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\} \} \quad (3.1.60)$$

We see that a set with 2 elements has a power set with 4 elements and a set with 3 elements has a power set with 8. This pattern continues for finite sets and if  $A$  has  $n$  elements, then  $\mathcal{P}(A)$  has  $2^n$  elements.

**Example 3.1.24** Suppose we have two variables  $a, b$  and the set  $A = \{a, b\}$ . We can compute the power set of  $A$  as follows:

$$\mathcal{P}(A) = \{ \emptyset, \{a\}, \{b\}, \{a, b\} \} \quad (3.1.61)$$

much like the previous example, we see that the power set of a set with 2 elements has  $2^2 = 4$  elements.

**Example 3.1.25** Let  $A = \{a_1, \dots, a_n\}$ , where all of the elements  $a_k$  are distinct. To count the total number of subsets we first note that there is one set that contains zero elements, the empty set. Next, there are  $n$  subsets that contain one element, these are the sets  $\{a_k\}$ . There are  $n(n - 1)/2$  sets that contain two elements,  $\{a_i, a_j\}$ , such that  $i \neq j$ . Continuing, we see that there are  $\binom{n}{k}$  subsets that contain  $k$  elements, where  $\binom{n}{k}$  is the *binomial coefficient*. This is defined in terms of the factorial function:

$$\binom{n}{k} = \frac{n!}{k!(n - k)!} \quad (3.1.62a)$$

$$= \frac{n \cdot (n - 1) \cdots 2 \cdot 1}{k \cdot (k - 1) \cdots 2 \cdot 1 \cdot (n - k) \cdot (n - k - 1) \cdots 2 \cdot 1} \quad (3.1.62b)$$

$$= \frac{n \cdot (n - 1) \cdots (n - k + 1)}{k \cdot (k - 1) \cdots 2 \cdot 1} \quad (3.1.62c)$$

To avoid undefined ratios, we define  $0! = 1$ . Note then that  $\binom{n}{n} = 1$ . This says that the number of ways to choose  $n$  element subsets from  $A$  is 1. This makes sense since the only  $n$  element subset of  $A$  is the entirety of  $A$ ! To compute the size of  $\mathcal{P}(A)$  it now suffices to sum over all of these binomial coefficients. Such a task can be achieved by invoking the *binomial theorem*. Given a positive integer  $n$  and two real numbers  $x$  and  $y$ , the binomial theorem states that:

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \quad (3.1.63a)$$

$$= \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \cdots + \binom{n}{n-1} x y^{n-1} + \binom{n}{n} y^n \quad (3.1.63b)$$

Here, the notation  $\Sigma$  is simply shorthand for denoting a long sum. For example:

$$\sum_{n=1}^3 n = 1 + 2 + 3 = 6 \quad (3.1.64a) \quad \sum_{n=1}^3 n^2 = 1 + 2^2 + 3^2 = 14 \quad (3.1.64b)$$

Setting  $x = y = 1$ , we obtain:

$$2^n = \sum_{k=0}^n \binom{n}{k} \quad (3.1.65)$$

and this is precisely the number of elements of  $\mathcal{P}(A)$ .

**Example 3.1.26** When we consider the case of an *infinite* set  $A$  we have that  $\mathcal{P}(A)$  is a strictly larger set and this creates a paradoxical heirarchy of infinities. The smallest heirarchy is that of the *countable* infinite sets, like  $\mathbb{N}$ . Everything larger is called *uncountable*. It will be shown that the following is true:

$$\text{curl}(\mathcal{P}(\mathbb{N})) = \text{curl}(\mathbb{R}) \quad (3.1.66)$$

where again  $\mathbb{N}$  denotes the non-negative integers and  $\mathbb{R}$  denotes the set of all *real* numbers. We can loosely show this by using the binary representation of real numbers. A real number may be thought of as an infinite decimal. For example,  $\pi = 3.1415926\dots$  and  $1 = 1.000\dots$  We can also represent real numbers as a sequence of zeroes and ones and this is the *binary* representation. For  $A \subseteq \mathbb{N}$  and let  $r_A = 0.n_1n_2\dots$  where:

$$n_i = \begin{cases} 0, & i \notin A \\ 1, & i \in A \end{cases} \quad (3.1.67)$$

Thus for each  $A \in \mathcal{P}(\mathbb{N})$  there is a real number  $r_A$  such that  $0 \leq r_A \leq 1$  that is associated with it, and moreover to every real number between zero and one there is a subset of  $\mathbb{N}$  associated with it. The tricky numbers to see are zero and one, but note that  $r_\emptyset$  is associated to 0 and  $r_{\mathbb{N}}$  gets paired with 1. To show that  $\mathbb{R}$  and  $\mathcal{P}(\mathbb{N})$  are the same size requires us to refine this association so that every element of  $\mathcal{P}(\mathbb{N})$  uniquely corresponds to an element of  $\mathbb{R}$ , and vice-versa.

### 3.1.4 Cartesian Products and Functions

Previously we've introduced ordered pairs and the notion of the power set. We can use both of these concepts to define and prove the existence of *Cartesian products*. Intuitively we want to define  $A \times B$  to be the set of all ordered pairs  $(a, b)$  where  $a \in A$  and  $b \in B$ :

$$A \times B = \{ (a, b) \mid a \in A \text{ and } b \in B \} \quad (3.1.68)$$

But recalling Def. 3.1.5, ordered pairs are sets of the form  $\{\{a\}, \{a, b\}\}$ . Thus elements of  $A \times B$  are contained in the power set of the power set of  $A \cup B$ :

$$A \times B \subseteq \mathcal{P}(\mathcal{P}(A \cup B)) \quad (3.1.69)$$

We can combine the axiom of the power set with the axiom schema of specification to obtain the existence of the Cartesian product of two sets.

**Theorem 3.1.15.** *If  $A$  and  $B$  are sets, if  $a \in A$  and  $b \in B$ , then  $(a, b) \subseteq \mathcal{P}(A \cup B)$ .*

*Proof.* For if  $a \in A$  and  $b \in B$ , then  $(a, b) = \{\{a\}, \{a, b\}\}$  (Def. 3.1.5). But if  $a \in A$ , then  $a \in A$  or  $a \in B$ , and thus  $a \in A \cup B$  (Def. 3.1.7). But then  $\{a\} \subseteq A \cup B$  (Def. 3.1.2). But if  $b \in B$ , then  $b \in A$  or  $b \in B$ , and thus  $b \in A \cup B$  (Def. 3.1.7). But then  $\{a, b\} \subseteq A \cup B$  (Def. 3.1.2). But then  $\{a\} \subseteq A \cup B$  and  $\{a, b\} \subseteq A \cup B$ , and thus  $(a, b) \subseteq \mathcal{P}(A \cup B)$  (Def. 3.1.12).  $\square$

**Theorem 3.1.16: Existence of the Cartesian Product**

If  $A$  and  $B$  are sets, then there exists a set  $A \times B$  such that, for all  $z$ ,  $z \in A \times B$  if and only if there is an  $a \in A$  and  $b \in B$  such that  $z = (a, b)$ . ■

*Proof.* For if  $A$  and  $B$  are sets, then  $A \cup B$  is a set (Thm. 3.1.7). But if  $A \cup B$  is a set, then  $\mathcal{P}(A \cup B)$  is a set (Thm. 3.1.14), where  $\mathcal{P}(X)$  denotes the power set of  $X$ . But if  $\mathcal{P}(A \cup B)$  is a set, then  $\mathcal{P}(\mathcal{P}(A \cup B))$  is a set (Thm. 3.1.14). But then  $z \in \mathcal{P}(\mathcal{P}(A \cup B))$  if and only if  $z \subseteq \mathcal{P}(A \cup B)$  (Def. 3.1.12). But if  $a \in A$  and  $b \in B$ , then  $(a, b) \subseteq \mathcal{P}(A \cup B)$  (Thm. 3.1.15), and therefore  $(a, b) \in \mathcal{P}(\mathcal{P}(A \cup B))$ . Let  $P$  be the proposition *True if there exists  $a \in A$  and  $b \in B$  such that  $z = (a, b)$ , false otherwise*. Then by the axiom schema of specification (Ax. 3.1.3), there exists a set  $A \times B$  such that:

$$A \times B = \{ z \in \mathcal{P}(\mathcal{P}(A \cup B)) \mid P(z) \} \quad (3.1.70)$$

But it was proved that  $P(z)$  implies that  $z \in \mathcal{P}(\mathcal{P}(A \cup B))$ . Thus  $z \in A \times B$  if and only if there exists  $a \in A$  and  $b \in B$  such that  $z = (a, b)$ . □

**Definition 3.1.13: Cartesian Product of Two Sets**

The **Cartesian product** of two **sets**  $A$  and  $B$  is the set:

$$A \times B = \{ (a, b) \mid a \in A \text{ and } b \in B \}$$

Formally:

$$\forall_A \forall_B \forall_z \left( (z \in A \times B) \iff (\exists_{x \in A} \wedge \exists_{y \in B} : z = (x, y)) \right)$$

Note that since, in general,  $(a, b) \neq (b, a)$ , it is generally true that  $A \times B \neq B \times A$ . Indeed, equality occurs if and only if  $A = B$  (or if either set is empty).

**Example 3.1.27: Basic Cartesian Products**

Let  $A$  and  $B$  be sets defined as follows:

$$A = \{1, 2, 3\} \quad (3.1.71a)$$

$$B = \{a, b\} \quad (3.1.71b)$$

Let's compute  $A \times B$  and  $B \times A$ . From the definition (Def. 3.1.13) we have:

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\} \quad (3.1.72)$$

Using this, we can compute:

$$A \times B = \{(1, a), (2, a), (3, a), (1, b), (2, b), (3, b)\} \quad (3.1.73)$$

Computing  $B \times A$ , we have:

$$B \times A = \{(a, 1), (a, 2), (a, 3), (b, 1), (b, 2), (b, 3)\} \quad (3.1.74)$$

Now if we suppose that  $a$  is not equal to 1, then we see that  $(a, 1)$  is a different element than  $(1, a)$ , and thus  $A \times B$  is not equal to  $B \times A$ . Next, compute  $A \times A$ :

$$A \times A = \left\{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\right\} \quad (3.1.75)$$

And finally  $B \times B$ :

$$B \times B = \{(a, a), (a, b), (b, a), (b, b)\} \quad (3.1.76)$$

Equality of  $A \times B$  and  $B \times A$  is achieved if and only if  $A = B$ , or if either set is the empty set.

Note that in Ex. 3.1.27, the *size* of the Cartesian product of two sets was simply the product of the number of elements of the constituent sets. That is, we see that  $A$  has three elements and  $B$  has two elements, but also that  $A \times B$  has six elements. Moreover,  $A \times A$  has nine elements and  $B \times B$  has four. This pattern holds for the Cartesian products of any two *finite* sets.

It is common to consider the Cartesian product of a set with itself. That is, given a set  $A$ , we are often interested in  $A \times A$ . We denote this by writing  $A^2$ . One such example is when we consider the set of real numbers  $\mathbb{R}$ . The Cartesian product  $\mathbb{R}^2$  is called the *Euclidean Plane*, or the *Cartesian Plane*, after Euclid of Alexandria and René Descartes. This is because  $\mathbb{R}^2$  is used

to model both planar geometry and analytical geometry, of which Euclid and Descartes were pioneers of, respectively. The term Cartesian products is in honor of René Descartes, as well. Let  $\mathbb{R}$  denote the set of real numbers, and let  $A = \mathbb{R}$  and  $B = \mathbb{R}$ . Then we have:

$$A \times B = \mathbb{R} \times \mathbb{R} \equiv \mathbb{R}^2 \quad (3.1.77)$$

Where the symbol  $\equiv$  means that  $\mathbb{R}^2$  is defined by this expression. Using the definition of Cartesian products (Def. 3.1.13), we obtain:

$$\mathbb{R}^2 = \{ (x, y) : x \in \mathbb{R} \text{ and } y \in \mathbb{R} \} \quad (3.1.78)$$

That is,  $\mathbb{R}^2$  is the set of all ordered pairs of real numbers. The first term is called the  $x$  coordinate, and similarly the second term is called the  $y$  coordinate. We envision this as a *plane* of points, each one corresponding to an ordered pair  $(x, y)$ . This is depicted in Fig. 3.6.

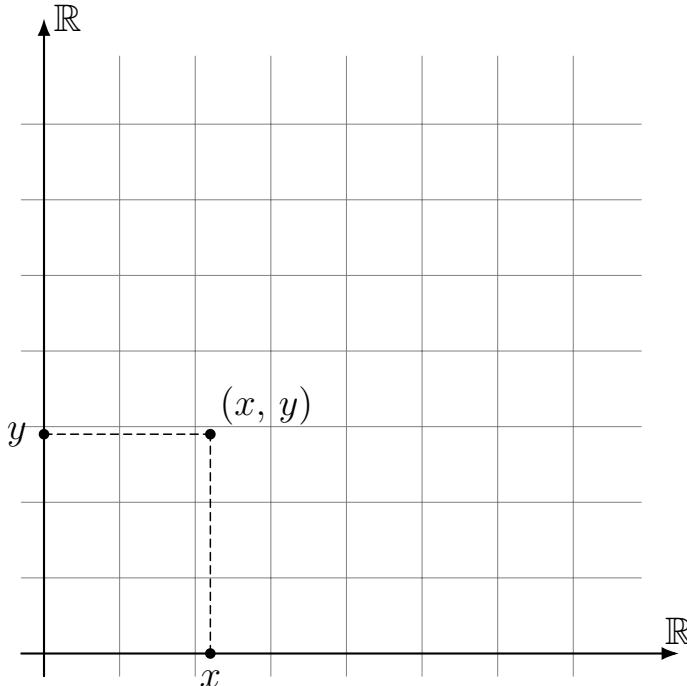


Fig. 3.6: The Cartesian Plane  $\mathbb{R}^2$

Consider further the set  $\mathbb{N}^2$ . That is, letting  $\mathbb{N}$  denote the set of natural numbers (Eqn. 3.1.3), letting  $A = \mathbb{N}$  and  $B = \mathbb{N}$  we have:

$$A \times B = \mathbb{N} \times \mathbb{N} \equiv \mathbb{N}^2 \quad (3.1.79)$$

Again using the definition of Cartesian products (Def. 3.1.13), we have:

$$\mathbb{N}^2 = \{ (n, m) \mid n \in \mathbb{N} \text{ and } m \in \mathbb{N} \} \quad (3.1.80)$$

We can visualize this as a subset of  $\mathbb{R}^2$  by drawing a lattice of points in the Cartesian plane (Fig. 3.7).

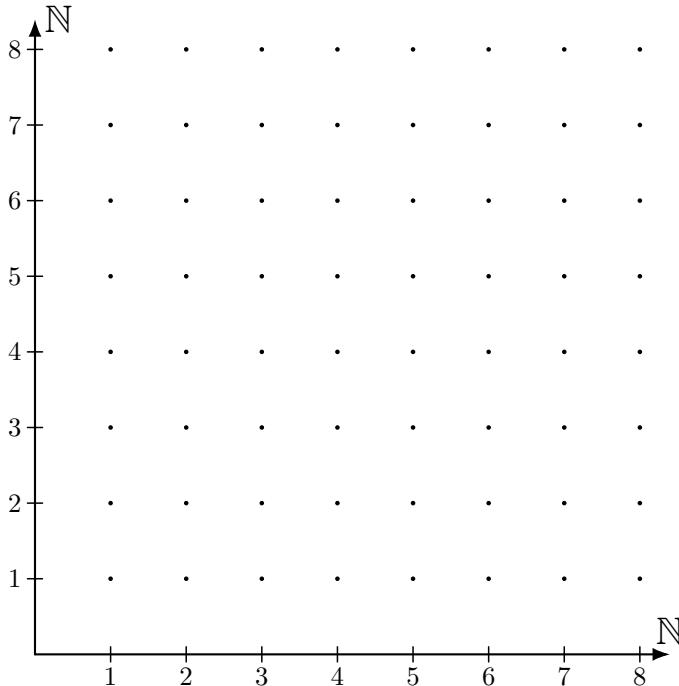


Fig. 3.7: The Lattice  $\mathbb{N}^2$

This can then be considered a subset of the Euclidean plane  $\mathbb{R}^2$ . That is, at every ordered pair of integers  $(m, n)$ , we place a point in the Euclidean plane whose  $x$  coordinate is  $m$  and whose  $y$  coordinate is  $n$ . We can also be more abstract and general in our examples. Consider the following sets:

$$A = \{ \text{Point, Line 1, Line 2} \}$$

$$(3.1.81a) \qquad B = \{ \text{Point, Line} \} \quad (3.1.81b)$$

We can visually represent the Cartesian product  $A \times B$  by drawing  $A$  in green and  $B$  in red, as shown in Fig. 3.8. The Cartesian Product  $A \times B$  is the set formed by connecting all of the points from  $A$  and  $B$  in the plane. This is shown in blue.

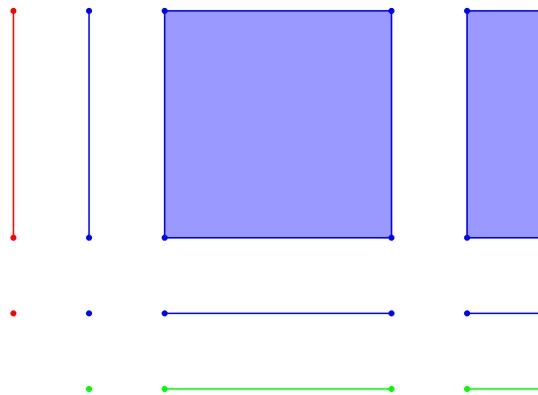


Fig. 3.8: The Cartesian Product of Two Sets. \$A\$ is in Green, \$B\$ is in red, and \$A \times B\$ is in blue.

Cartesian products are not *associative*. That is, given three sets \$A\$, \$B\$, and \$C\$, there is no clear way to take the Cartesian product of these since:

$$A \times (B \times C) \neq (A \times B) \times C \quad (3.1.82)$$

To see this, note that the elements of \$A \times (B \times C)\$ are ordered pairs of the form \$(a, (b, c))\$, whereas elements of \$(A \times B) \times C\$ are of the form \$((a, b), c)\$. When we write \$A \times B \times C\$ we really want ordered *triples* of the form \$(a, b, c)\$. Much the way ordered pairs have been defined, we can modify Kuratowski's approach and define ordered triples and ordered \$n\$ tuples. Rather than doing this we will use the language of functions to define higher order Cartesian products.

#### Definition 3.1.14: Functions

A **function** from a **set** \$A\$ to a set \$B\$ is a **subset** \$f \subseteq A \times B\$, denoted \$f : A \rightarrow B\$, such that for all \$x \in A\$ there is a unique \$y \in B\$ such that \$(x, y) \in f\$. \$A\$ is called the **domain** of \$f\$ and \$B\$ is called the **codomain**.

We're used to hearing that a function is a rule that assigns to an input value \$x\$ some output value \$f(x)\$. It may seem hard to justify, then, why we've defined a function as a subset of the Cartesian product. But note the requirement that for each \$x \in A\$ there is a *unique* \$y \in B\$ such that \$(x, y) \in f\$. We call this unique element the *image* of \$x\$ under the function \$f\$ and write \$y = f(x)\$. The condition that there is a unique such value \$y\$ to each \$x\$ is called the *vertical line test* when graphing functions of the form \$f : \mathbb{R} \rightarrow \mathbb{R}\$ (Fig. 3.9). Simply, given such a function, if one draws a vertical line in the plane, then it must intersect the graph of \$f\$ once and only once. This provides a quick means of discerning functions from non-functions.

**Example 3.1.28: The Square Function**

If we can come up with some rule that assigns to every element  $a \in A$  a unique element of  $B$ , then we can use this rule to define a function  $f : A \rightarrow B$ . Such a rule often comes in the form of a *formula*. We write the unique element that  $a$  corresponds to as  $f(a)$ . For example, let  $A = \mathbb{R}$  and let  $B = \mathbb{R}$ . We can define a function by the squaring formula:

$$f(x) = x \cdot x = x^2 \quad (3.1.83)$$

Once we know that  $x^2$  gives a unique number (which will require some notion of arithmetic), we can define the function  $f : \mathbb{R} \rightarrow \mathbb{R}$  by:

$$f = \{ (x, x^2) \in \mathbb{R}^2 \mid x \in \mathbb{R} \} \quad (3.1.84)$$

Usually we'll define functions by their formula's, rather than expressing them explicitly as subsets of the Cartesian product. ■

In the field of mathematical analysis we are often concerned with functions involving real numbers. For the sake of intuition, let us consider functions of the form  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Any curve that we draw left-to-right, without picking up the pencil, will be a valid function. (See Fig. 3.9).

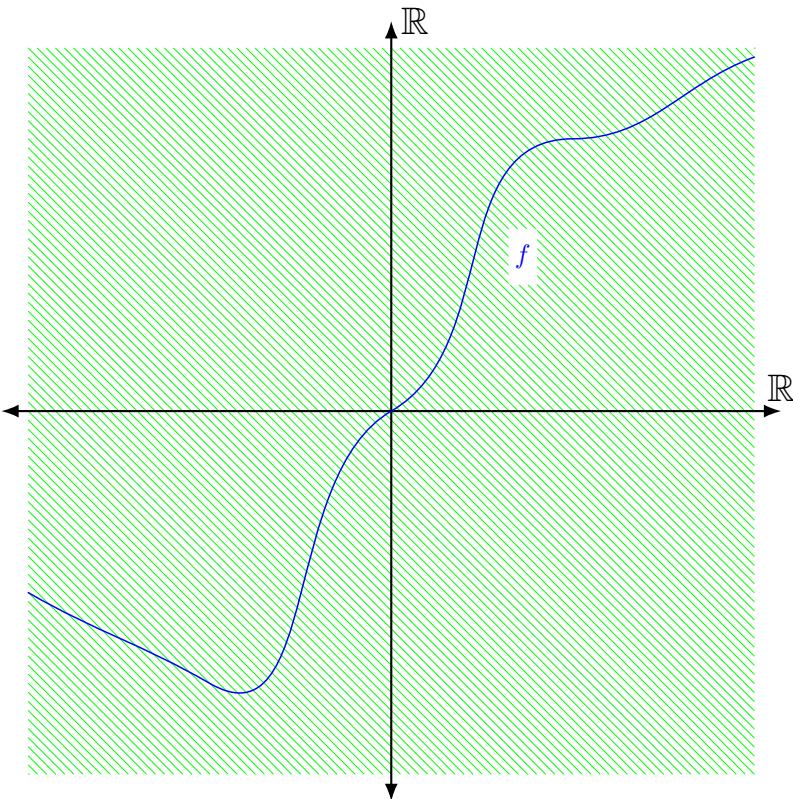


Fig. 3.9: Example of a function  $f : \mathbb{R} \rightarrow \mathbb{R}$ . The Cartesian product  $\mathbb{R} \times \mathbb{R}$  is shown in green, and the function  $f \subseteq \mathbb{R} \times \mathbb{R}$  is shown in blue.

Let  $g \subseteq \mathbb{R} \times \mathbb{R}$  be defined as follows:

$$g = \{ (x, y) \in \mathbb{R}^2 \mid y^2 = x \} \quad (3.1.85)$$

It is tempting to label  $g$  by writing  $g(x) = \sqrt{x}$ , but  $g$  is not a function for it fails two of the requirements of a function. Firstly, for any  $x > 0$ , there are two values  $y_1$  and  $y_2$  whose square is equal to  $x$ . Indeed, if  $y_1$  is one such value, then setting  $y_2 = -y_1$  will result in a second distinct value. Thus  $g$  does not have the uniqueness property required for functions. Moreover, if  $x < 0$ , then there is no such value  $y \in \mathbb{R}$  such that  $(x, y) \in g$ , and thus  $g$  also lacks the existence property. In terms of the vertical line test, there are points  $x$  such that the vertical line through  $(x, 0)$  intersects  $g$  twice, and there are points such that the vertical line does not intersect at all. The graph of  $g$  is shown in Fig. 3.10.

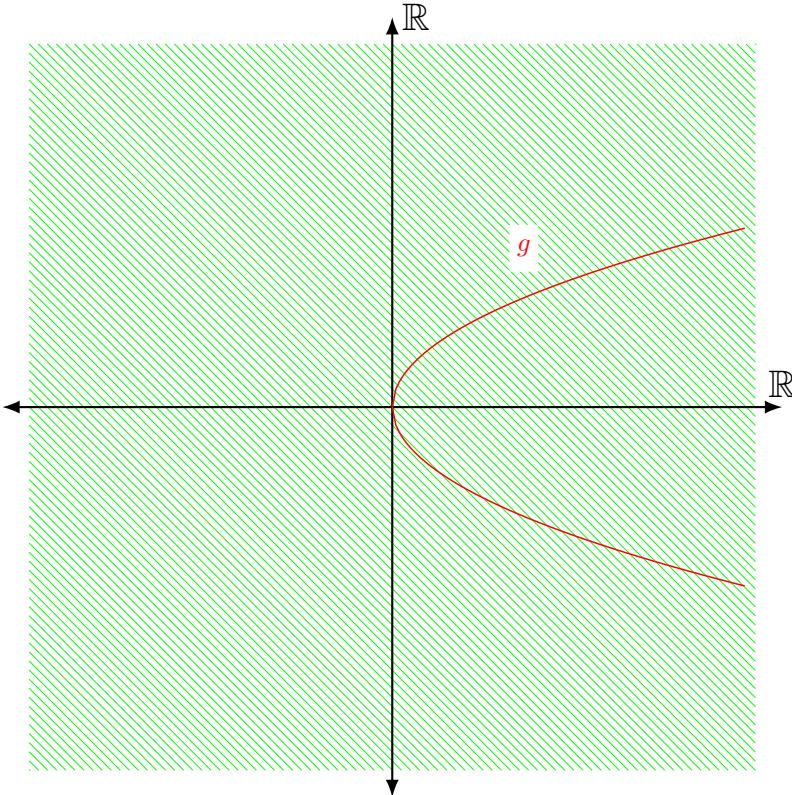
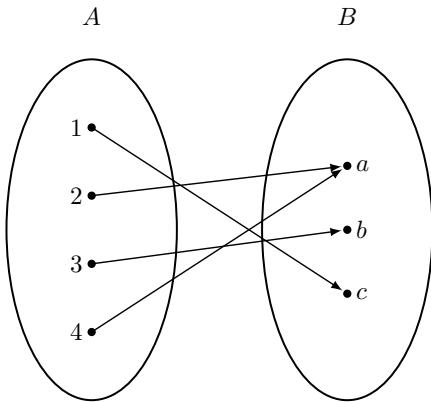


Fig. 3.10:  $g \subseteq \mathbb{R} \times \mathbb{R}$  is not a function since it fails the vertical line test.

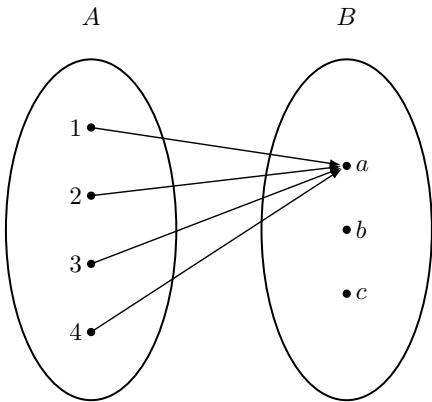
We need not only consider functions of the form  $f : \mathbb{R} \rightarrow \mathbb{R}$ , nor are we restricted to function like  $f : \mathcal{U} \rightarrow \mathcal{V}$  where  $\mathcal{U}$  and  $\mathcal{V}$  are subsets of  $\mathbb{R}$ , and we can allow for arbitrary abstract functions. Let  $A$  and  $B$  be defined as follows:

$$A = \{1, 2, 3, 4\} \quad (3.1.86a) \qquad B = \{a, b, c\} \quad (3.1.86b)$$

Similar to the vertical line test, we can devise a visual to discerning functions from non-functions for abstract sets. We represent the elements of  $A$  and  $B$  as points in some blob in the plane, and then draw arrows between the points  $x \in A$  and  $y \in B$  indicating that  $(x, y) \in f$ . This allows us to determine if a given  $f \subseteq A \times B$  is a functions ore not. Every point in  $A$  must be mapped to a unique point in  $B$ . That is, every point in  $A$  must have one and only one arrow connecting it to some point in  $B$ . Examples of valid functions are shown in Fig. 3.11, and non-functions are shown in Fig. 3.12.

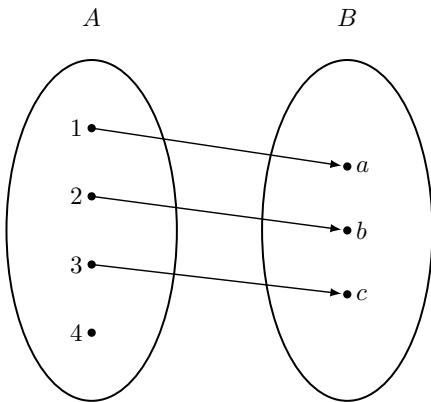


3.11.1: A Valid Function.

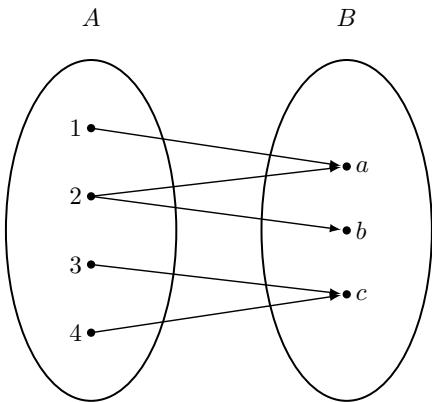


3.11.2: Another Valid Function.

Fig. 3.11: Visual for Abstract Functions



3.12.1: Fails Existence.



3.12.2: Fails Uniqueness.

Fig. 3.12: Non-Functions

It is possible to count the total number of functions from  $A$  to  $B$ . Since every element of  $A$  needs to be mapped to some element of  $B$ , and since there are 4 elements in  $A$  and 3 elements in  $B$ , the total number of functions  $f : A \rightarrow B$  is  $4^3 = 64$ . On the other hand, the total number of subsets of  $A \times B$  is  $2^{12} = 4096$  (we will justify this when we discuss the *cardinality* of sets). Thus, if we were to randomly pick a subset of  $A \times B$ , the odds are that it is almost certainly *not* a function (1.5625%). Thus, functions are very special subsets. There is a frequent need to discuss the *set of all functions* from a given set  $A$  into another set  $B$ . To ensure we don't create a function version of Russell's paradox, we prove such a set exists.

**Theorem 3.1.17.** *If  $A$  and  $B$  are sets, then there exists a set  $\mathcal{F}$  such that, for all  $f$  it is true that  $f \in \mathcal{F}$  if and only if  $f$  is a function from  $A$  to  $B$ ,  $f : A \rightarrow B$ .*

*Proof.* For if  $A$  and  $B$  are sets, then by Thm. 3.1.16 the set  $A \times B$  exists. But by Thm. 3.1.14, the power set of  $A \times B$ ,  $\mathcal{P}(A \times B)$ , exists. Let  $P$  be the proposition *True if  $f$  is a function from  $A$  to  $B$ , false otherwise*. Then by axiom schema of specification (Ax. 3.1.3), there is a set  $\mathcal{F}$  such that:

$$\mathcal{F} = \{ f \in \mathcal{P}(A \times B) \mid P(f) \} \quad (3.1.87)$$

But then for all  $f$ ,  $f \in \mathcal{F}$  if and only if  $f \in \mathcal{F}$  and  $P(f)$  is true. But if  $P(f)$  is true then  $f$  is a function from  $A$  to  $B$ , and thus by the definition of a function (Def. 3.1.14)  $f \subseteq A \times B$ . But then by the definition of the power set (Def. 3.1.12) we have that  $f \in \mathcal{P}(A \times B)$ . Thus  $P(f)$  implies  $f \in \mathcal{F}$ . Therefore  $f \in \mathcal{F}$  if and only if  $P(f)$ . That is,  $f \in \mathcal{F}$  if and only if  $f$  is a function from  $A$  to  $B$ .  $\square$

There is non-standard notation when discussing the set of all functions from a given set  $A$  to a set  $B$ :

### Notation 3.1.3: Set of All Functions

If  $A$  and  $B$  are sets, the set of all functions from  $A$  to  $B$ ,  $f : A \rightarrow B$ , is denoted as either  $\mathcal{F}(A, B)$  or  $B^A$ .

The notation  $B^A$  is common in many areas such as topology and algebra, especially when  $A = B$ . The *topological space*  $I^I$ , which is the set of all functions from the *closed unit interval* to itself, is often used to construct examples and counterexamples. In analysis the notation  $\mathcal{F}(A, B)$  seems to be more common, in particular  $\mathcal{C}(A, B)$  is often used to denote the set of all *continuous* functions from  $A$  to  $B$ , provided the word continuous has meaning. Since the notation is not universal nor standard across the various disciplines, an attempt will be made to specify what  $B^A$  or  $\mathcal{F}(A, B)$  means before using it in a theorem or counterexample.

### Definition 3.1.15: Image of a Point

The *image* of an element  $x$  in a set  $A$  under a *function*  $f : A \rightarrow B$  is the unique value  $y \in B$  such that  $(x, y) \in f$ . We write  $y = f(x)$ .

This allows us to define functions by simply specifying what the image of each  $x \in A$  is. Restating our previous claim, if we can define some formula such

that for each  $x \in A$  there is a unique  $f(x) \in B$  such that the formula takes  $x$  to  $f(x)$ , then we can define  $f$  as the set of all such ordered pairs  $(x, f(x))$ , and this will be a function.

#### Notation 3.1.4: Image Notation

If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, if  $x \in A$  and if  $y = f(x) \in B$ , then we denote this by writing  $x \xrightarrow{f} y$  or just  $x \mapsto y$ .

Throughout we will almost exclusively use the notation  $y = f(x)$  rather than  $x \mapsto y$ . The reasons are purely aesthetic and both notations are common in mathematics. In a similar manner, we can define the image of an entire subset.

**Theorem 3.1.18.** *If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, and if  $\mathcal{U} \subseteq A$ , then there is a set  $\mathcal{V} \subseteq B$  such that for all  $y$  it is true that  $y \in \mathcal{V}$  if and only if  $y \in B$  and such that there is an  $x \in \mathcal{U}$  such that  $y = f(x)$ .*

*Proof.* For let  $P$  be the proposition *True if there exists  $x \in \mathcal{U}$  such that  $y = f(x)$ , false otherwise*. By the axiom schema of specification (Ax. 3.1.3) there is a set  $\mathcal{V}$  such that, for all  $y$  it is true that  $y \in \mathcal{V}$  if and only if  $y \in B$  and  $P(y)$  is true. That is,  $y \in \mathcal{V}$  if and only if  $y \in B$  and if there is an  $x \in \mathcal{U}$  such that  $y = f(x)$ .  $\square$

#### Definition 3.1.16: Image of a Subset

The **image** of a **subset**  $\mathcal{U}$  of a **set**  $A$  under a **function**  $f : A \rightarrow B$  is the set:

$$f(\mathcal{U}) = \{ y \in B \mid \text{There exists } x \in \mathcal{U} \text{ such that } y = f(x) \}$$

That is, the set of all points in  $B$  that are the image of points in  $\mathcal{U}$ . Formally:

$$\forall_A \forall_B \forall_{f:A \rightarrow B} \forall_{\mathcal{U} \subseteq A} \forall_y \left( (y \in f(\mathcal{U})) \iff (\exists_{x \in \mathcal{U}} : y = f(x)) \right)$$

This definition of the image of a subset was given in such a manner so that it only relies on the axiom schema of specification to justify its existence. We could also use the notation:

$$f(\mathcal{U}) = \{ f(x) \in B \mid x \in \mathcal{U} \} \tag{3.1.88}$$

Writing the definition of the image of a subset in such a way is justified by the

*axiom schema of replacement*, but we've not yet included this axiom in our system.

**Example 3.1.29** If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is the function  $f(x) = x^2$ , then:

$$f(\mathbb{R}) = [0, \infty) \equiv \{x \in \mathbb{R} \mid x \geq 0\} \quad (3.1.89)$$

This is because every non-negative real number  $y$  gets mapped to by at least one real number (the positive square root  $\sqrt{y}$ ). None of the negative numbers are the image of any element of  $\mathbb{R}$  since the square of a real number is always non-negative.

We can visualize functions and images by using blobs in the plane. Given some sub-blob of a set  $A$ , the image of this will be another sub-blob of  $B$ . Note that if  $f : A \rightarrow B$  is a function, it does **not** need to be true that  $f(A) = B$ . These are special functions that are called *surjective* and are discussed in Chapt. 5. Such a drawing of the general case is shown in Fig. 3.13.

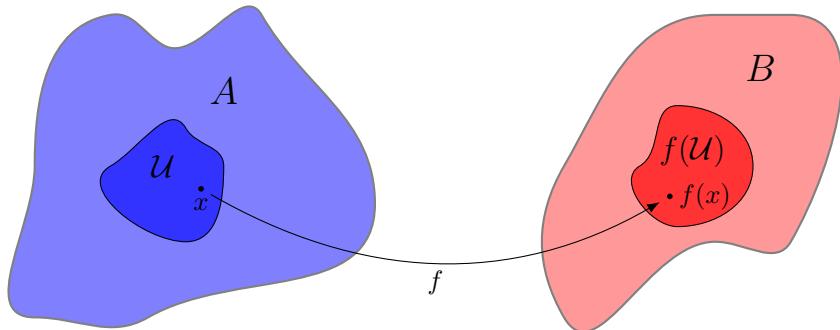


Fig. 3.13: Image of a Subset and of a Point under a Function

If we consider a function  $f : A \rightarrow B$  and the image of the entire set  $A$  we obtain the *range* of  $f$ . That is, the range is the set  $f(A) \subseteq B$ . In a similar manner to the forward image of a function, we can define the pre-image.

**Theorem 3.1.19.** *If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function from  $A$  to  $B$ , and if  $\mathcal{V} \subseteq B$ , then there is a set  $\mathcal{U} \subseteq A$  such that for all  $x$  it is true that  $x \in \mathcal{U}$  if and only if  $x \in A$  and  $f(x) \in \mathcal{V}$ .*

*Proof.* For let  $P$  be the proposition *True if  $x \in A$  and  $f(x) \in \mathcal{V}$ , false otherwise*. Then by the axiom schema of specification (Ax. 3.1.3) there is a set  $\mathcal{U}$  such that:

$$\mathcal{U} = \{x \in A \mid P(x)\}$$

But  $P(x)$  implies  $x \in A$  and thus  $x \in \mathcal{U}$  if and only if  $x \in A$  and  $f(x) \in \mathcal{V}$ .  $\square$

### Definition 3.1.17: Pre-Image of a Subset

The **pre-image** of a **subset**  $\mathcal{V} \subseteq B$  under a **function**  $f : A \rightarrow B$  is the set:

$$f^{-1}(\mathcal{V}) = \{x \in A \mid f(x) \in \mathcal{V}\} \quad (3.1.90)$$

Using our formal language:

$$\forall_A \forall_B \forall_{f:A \rightarrow B} \forall_{\mathcal{V} \subseteq B} \forall_x ((x \in f^{-1}(\mathcal{V})) \iff ((x \in A) \wedge (f(x) \in \mathcal{V})))$$

The pre-image of a set behaves a lot differently than the image, and this will be explored in detail when functions are discussed. The cause of the discrepancy is the requirement that elements of  $A$  map uniquely to elements of  $B$ , but a single element in  $B$  can be the image of many different points in  $A$ . This gives rise to the notion of the **fiber** of a point in  $B$ .

**Theorem 3.1.20.** *If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, and if  $b \in B$ , then there is a set  $\mathcal{U} \subseteq A$  such for all  $x \in A$  it is true that  $x \in \mathcal{U}$  if and only if  $f(x) = b$ .*

*Proof.* For by Thm. 3.1.4,  $\{b\}$  is a set and  $\{b\} \subseteq B$  (Def. 3.1.2). But if  $\{b\}$  is a subset of  $B$ , then there is a set  $\mathcal{U} \subseteq A$  such that for all  $x \in A$  it is true that  $x \in \mathcal{U}$  if and only if  $f(x) \in \{b\}$  (Thm. 3.1.19). But  $f(x) \in \{b\}$  if and only if  $f(x) = b$  (Thm. 3.1.4). Thus, for all  $x \in A$ ,  $x \in \mathcal{U}$  if and only if  $f(x) = b$ .  $\square$

### Definition 3.1.18: Fiber of an Element

The **fiber** of an element  $b$  in a **set**  $B$  under a **function**  $f : A \rightarrow B$  from a set  $A$  to a set  $B$  is the pre-image of the set  $\{b\}$ . That is:

$$f^{-1}(\{b\}) = \{a \in A \mid f(a) = b\}$$

Before we end our introduction to the concept of a function, we should first note the other synonomous terminology that is used and give some historical background. In many texts, including this one, authors use the word *map* or *mapping* to denote a function, rather than simply use the word function. This language has its roots in geometry, and particularly in the study of projective geometry. One can consider *perspectives*, which are functions that map one line in the Euclidean plane to another. Consider two non-parallel lines  $\overline{AB}$  and  $\overline{CD}$

and a point  $P$  in the plane that lies on neither of these two. Given a point  $x$  which lies on  $\overline{AB}$ , we map this to the line  $\overline{CD}$  by drawing a straight line from  $P$  to  $x$  and marking where this intersects  $\overline{CD}$  (see Fig. 3.14).

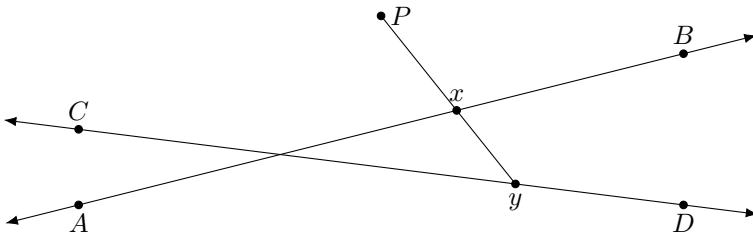


Fig. 3.14: Example of a Mapping from Projective Geometry

Such constructions play a central role in geometry, and particularly projective geometry, and culminate in a beautiful theorem known as *Desargues's Theorem*, named after Girard Desargues. From this historical motivation, the word function, map, and mapping are synonymous in the realm of mathematics. In select fields such as analysis and geometry the words *operator* and *transformation* are used, and occasionally the word *graph* is used. We will try to be clear in our usage of this vocabulary to rid of any potential ambiguity.

### 3.1.5 The Axiom of Choice and Diaconescu's Theorem

The next two axioms to be introduced are the most controversial of those listed in ZFC: The *axiom of infinity* and the *axiom of choice*. While the axiom of infinity only has a small number of critics, the axiom of choice is far more contentious. Choice is equivalent to many other statements that come across in almost all forms of mathematics (analysis, algebra, topology, etc.). Many of which are theorems we would *want* to be true, and so accepting the axiom of choice allows us to prove them. In particular, the axioms presented thus far can be combined with the axiom of choice to prove the *Law of the Excluded Middle*, a result known as Diaconescu's theorem, and this is our current goal.

#### Axiom 3.1.8: Axiom of Choice

If  $\mathcal{O}$  is a non-empty set such that for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $\mathcal{U}$  is non-empty, and if  $\bigcup \mathcal{O}$  is the union over  $\mathcal{O}$ , then there is a function  $f : \mathcal{O} \rightarrow \mathcal{F}$  such that for all  $x \in \mathcal{O}$  it is true that  $f(x) \in x$ . Formally:

$$\forall_{\mathcal{O}} ((\mathcal{O} \neq \emptyset) \wedge (\emptyset \notin \mathcal{O})) \exists_{f: \mathcal{O} \rightarrow \bigcup \mathcal{O}} \left( \forall_{x \in \mathcal{O}} (f(x) \in x) \right)$$

Such a function is called a *choice function*. The axiom can be made obviously true or obviously false depending on how we word it. To convince one of its validity requires talking about products. The Cartesian product has been defined using ordered pairs as defined by Kuratowski and allows us to order two elements. Given two sets  $A$  and  $B$  we can define an equivalent notion using the set of all functions from  $\mathbb{Z}_2 = \{0, 1\}$  into  $A \cup B$  with a particular property:

$$A \times B = \{ f : \mathbb{Z}_2 \rightarrow A \cup B \mid f(0) \in A \text{ and } f(1) \in B \} \quad (3.1.91)$$

To see why this is equivalent note that  $A \times B$  is the set of all ordered pairs whose first entry lies in  $A$  and whose second entry lies in  $B$ . Given  $a \in A$  and  $b \in B$ , let  $f : \mathbb{Z}_2 \rightarrow A \cup B$  be the function such that  $f(0) = a$  and  $f(1) = b$ . Then we can identify the ordered pair  $(a, b)$  with  $f$ . Indeed,  $(a, b) = (f(0), f(1))$  making our identification very explicit. We can now generalize to a collection of  $n$  different sets and define the ordered  $n$  tuple over a collection of  $n$  sets to be the set of all functions from  $\mathbb{Z}_n$  into the union over this collection with a similar property:

$$\prod_{k \in \mathbb{Z}_n} A_k = \left\{ f : \mathbb{Z}_n \rightarrow \bigcup_{k \in \mathbb{Z}_n} A_k \mid f(k) \in A_k \text{ for all } k \in \mathbb{Z}_n \right\} \quad (3.1.92)$$

Given the function  $f$  that maps  $k$  to  $a_k \in A_k$ , we identify this by:

$$f = (a_0, a_1, \dots, a_k, \dots, a_{n-1}) = (f(0), f(1), \dots, f(k), \dots, f(n-1)) \quad (3.1.93)$$

And thus we have a more general way of defining products. Note that swapping the order of the product is equivalent to changing functions, so two  $n$  tuples are equal if and only if all of their entries are equal. What's nice about our function definition is that it allows one to define products over *arbitrary* collections. This is crucial for topology and analysis as we often wish to speak of *infinite dimensional* spaces that are constructed using these abstract products. Given a set  $I$ , often called the *index set*, such that for all  $\mathcal{U} \in I$  it is true  $\mathcal{U}$  is a set, we can form the product over  $I$  by defining this to be the collection of all functions from  $I$  into the union over  $I$ .

$$\prod_{i \in I} A_i = \left\{ f : I \rightarrow \bigcup_{i \in I} A_i \mid f(i) \in A_i \text{ for all } i \in I \right\} \quad (3.1.94)$$

The axiom of choice is equivalent to the statement *The infinite product of non-empty sets is non-empty*. These functions that identify  $k$  with the  $k^{\text{th}}$  set are precisely choice functions. Phrasing it like this we see that the axiom of choice is somewhat obvious. The infinite product of non-empty sets is most likely enormous! Claiming it's non-empty seems trivial. It is then unfortunate that this claim can not be proven with the other axioms we've developed. As

stated before, the axiom of choice is equivalent to many other statements such as *Zorn's lemma*, *Tychonoff's theorem*, *the well ordering theorem*, *every vector space has a basis*, *every set has a group structure*, and countless more. Many of these theorems have many applications to algebra, analysis, and topology, and since we would like to use them to prove other things we are forced to accept the axiom of choice. Many theorems in real analysis hide the use of the axiom of choice by constructing sequences *by induction*. An attempt will be made to be clear whenever the axiom of choice is used in a proof.

We conclude this section by presenting Diaconescu's theorem.

**Theorem 3.1.21: Diaconescu's Theorem**

If  $P$  is a proposition on sets and if  $x$  is a set, then either  $P(x)$  is true or the negation of  $P(x)$  is true. That is,  $P \vee \neg P$  is true.

*Proof.* Let  $0 = \emptyset$ . By Thm. 3.1.4, we have that the set  $\{0\}$  exists. Let  $1 = \{0\}$ . Then since  $0 \in 1$ ,  $0 \neq 1$  (Thm. 3.1.12). Since  $0$  and  $1$  are sets, by Thm. 3.1.3 we have that the set  $\{0, 1\}$  exists. Let  $Q$  be the proposition *true if  $P(x)$  or  $x = 0$ , false otherwise*. By the axiom schema of specification (Ax. 3.1.3) there exists a set  $\mathcal{U}$  such that:

$$\mathcal{U} = \{ x \in \{0, 1\} \mid Q(x) \} \quad (3.1.95)$$

Similarly, let  $R$  be the proposition *true if  $P(x)$  or  $x = 1$ , false otherwise*. By the axiom schema of specification we have that the following set exists:

$$\mathcal{V} = \{ x \in \{0, 1\} \mid R(x) \} \quad (3.1.96)$$

By Thm. 3.1.3, we have that the set  $\{\mathcal{U}, \mathcal{V}\}$  exists. By the axiom of choice (Ax. 3.1.8), there exists a function  $f : \{\mathcal{U}, \mathcal{V}\} \rightarrow \bigcup\{\mathcal{U}, \mathcal{V}\}$  such that  $f(\mathcal{U}) \in \mathcal{U}$  and  $f(\mathcal{V}) \in \mathcal{V}$ . But then, by the definition of  $\mathcal{U}$ , either  $f(\mathcal{U}) = 0$  or  $P(x)$  is true. Similarly, either  $f(\mathcal{V}) = 1$  or  $P(x)$  is true. But since  $0 \neq 1$ , either  $f(\mathcal{U}) \neq f(\mathcal{V})$  or  $P(x)$  is true. Again by the axiom of extensionality (Ax. 3.1.2), and by the definition of  $\mathcal{U}$  and  $\mathcal{V}$ , if  $P(x)$  is true then  $\mathcal{U} = \mathcal{V}$ . But then  $f(\mathcal{U}) = f(\mathcal{V})$ . But then, by the contrapositive,  $\neg P(x)$  implies that  $f(\mathcal{U}) \neq f(\mathcal{V})$ . But by extensionality, either  $f(\mathcal{U}) = f(\mathcal{V})$  or  $f(\mathcal{U}) \neq f(\mathcal{V})$ , and thus either  $P(x)$  or  $\neg P(x)$ . That is,  $P \vee \neg P$  is true.  $\square$

We can now prove things via *proof by contradiction*. While we have made great efforts to justify every step of a proof thus far, we will often omit mention of Diaconescu's theorem as a justification for the law of the excluded middle and simply use it freely. We may rest easy knowing that we've proved its validity

within the framework of ZFC. There are two more axioms remaining, that of infinity and replacement. The axiom of infinity is best introduced when we construct the natural numbers, and from there build the real numbers, and thus we shall delay its development briefly. The axiom of replacement needs a notion of class and thus will also be postponed.

## 3.2 The Structure of Sets

We've developed two *operations* on sets thus far, those of union and intersection. Several of their properties give rise to the notion of a *Boolean algebra*. Many theorems about sets can then be proven algebraically and it is our current goal to prove the basics about unions and intersections so we may transition to algebra.

### 3.2.1 Basic Theorems

With the law of the excluded middle established we may now rapidly prove many basic results about sets. First, we prove the empty set is unique.

**Theorem 3.2.1.** *If  $A$  is a set, then  $\emptyset \subseteq A$ .*

*Proof.* For if not, then there is an  $x \in \emptyset$  such that  $x \notin A$  (Def. 3.1.2). But for all  $x$  it is true that  $x \notin \emptyset$  (Def. 3.2.1), a contradiction. Therefore  $\emptyset \subseteq A$ .  $\square$

**Theorem 3.2.2.** *If  $\emptyset'$  is a set with no elements, then  $\emptyset = \emptyset'$ .*

*Proof.* For suppose not. But  $\emptyset'$  is a set, and thus  $\emptyset \subseteq \emptyset'$  (Thm. 3.2.1). By the definition of equality if  $\emptyset \neq \emptyset'$ , then  $\emptyset' \not\subseteq \emptyset$  (Def. 3.1.3). But then there is an  $x$  such that  $x \in \emptyset'$  and  $x \notin \emptyset$  (Def. 3.1.2). But by hypothesis  $\emptyset'$  contains no elements, a contradiction. Thus  $\emptyset' \subseteq \emptyset$ . Therefore,  $\emptyset = \emptyset'$  (Def. 3.1.3).  $\square$

With this we can define *the* empty set.

#### Definition 3.2.1: The Empty Set

The *empty set* is the unique *set*  $\emptyset$  such that for all  $x$  it is true that  $x \notin \emptyset$ .

Often times when encountering problems with sets, or in any other branch of mathematics, one creates a chain of equalities or inequalities and then invokes the *transitive law* to conclude the proof. Such examples are found in elementary arithmetic with real numbers: If  $a < b$  and  $b < c$ , for real numbers  $a, b, c \in \mathbb{R}$ , then it must be true that  $a < c$ . Similar claims are found in Euclidean geometry. Indeed, if one were to peruse book one of Euclid's *Elements*, common

notion 1 reads: *things which are equal to the same thing are also equal to one another.* That is, if  $A = B$  and if  $B = C$ , then  $A = C$ . Euclid adopted this as a fundamental truth, or *axiom*, the word *equality* being one of his primitive undefined words. We've used the language of sets and thus the transitivity of equality must be proved as a theorem using the definitions adopted. To do so first requires the transitivity of inclusion, since equality is defined in terms of subsets.

### Theorem 3.2.3: Transitivity of Inclusion

---

If  $A$ ,  $B$ , and  $C$  are sets, if  $A \subseteq B$ , and if  $B \subseteq C$ , then  $A \subseteq C$ . ■

*Proof.* For suppose not. Then by the definition of subset there is an  $x \in A$  such that  $x \notin C$  (Def. 3.1.2). But  $A$  is a subset of  $B$  and thus  $x \in B$  (Def. 3.1.2). Similarly,  $B$  is a subset of  $C$  and therefore  $x \in C$  (Def. 3.1.2), a contradiction. □

**Example 3.2.1** Consider the sets we've discussed so far:  $\mathbb{N}$ ,  $\mathbb{Z}$ , and  $\mathbb{Q}$ . We can see from Eqns. 3.1.3 and 3.1.6 that  $\mathbb{N} \subseteq \mathbb{Z}$ , and we've also claimed that  $\mathbb{Z} \subseteq \mathbb{Q}$ . By the transitivity of inclusion (Thm. 3.2.3) we then have  $\mathbb{N} \subseteq \mathbb{Q}$ .

Containment ( $\in$ ) is not transitive. That is, if  $A \in B$  and  $B \in C$ , it may not be true that  $A \in C$ , nor is it necessarily true that  $\{A\} \in C$ , or even  $\{A\} \subseteq C$ . For a simple example, let  $A = \emptyset$ ,  $B = \{A, \{A\}\}$ , and  $C = \{B\}$ . Then by definition  $A \in B$  and  $B \in C$ , but  $A \notin C$ . This is because for all  $x$  it is true that  $x \in C$  if and only if  $x = B$ , and since  $A \in B$  it is necessarily true that  $A \neq B$  (see Thm. 3.1.12), and therefore  $A \notin C$ . Moreover, since  $\{A\} \notin \{A\}$  (Thm. 3.1.11) and  $\{A\} \in B$ , by the axiom of extensionality (Ax. 3.1.2) we conclude that  $\{A\} \neq B$ . Thus  $\{A\} \notin C$  and  $\{A\} \not\subseteq C$ . This last claim stems from the fact that the only subsets of  $C$  are  $\emptyset$  and the entirety of  $C$  itself. But  $\emptyset \neq \{\emptyset\}$  and  $\{\emptyset\} \neq C$ , and thus  $\{A\} \not\subseteq C$ .

It is possible to find sets  $A$ ,  $B$ , and  $C$  such that  $A \in B$ ,  $B \in C$ , and  $A \in C$ , we need only consider a nested chain of sets. Let  $A = \emptyset$ ,  $B = \{\emptyset\}$ , and  $C = \{\emptyset, \{\emptyset\}\}$ . We can write this more clearly as follows:

$$A = \emptyset \quad (3.2.1a) \quad B = \{A\} \quad (3.2.1b) \quad C = \{A, B\} \quad (3.2.1c)$$

Then by construction,  $A \in B$ ,  $B \in C$ , and  $A \in C$ . Such chains are used in the von Neumann construction of the integers.

With the transitivity of inclusion we can present a few brief corollaries.

**Theorem 3.2.4.** *If  $A$ ,  $B$ , and  $C$  are sets, if  $A = B$ , and if  $C \subseteq A$ , then  $C \subseteq B$ .*

*Proof.* For if  $A = B$ , then  $A \subseteq B$  (Def. 3.1.3). But by the transitivity of inclusion, if  $C \subseteq A$  and  $A \subseteq B$ , then  $C \subseteq B$  (Thm. 3.2.3).  $\square$

This theorem relates subsets and equal sets, and in a similar manner we can relate *supersets*. A superset of a set  $A$  is simply a set  $B$  that contains  $A$ . That is,  $A \subseteq B$ . This is the dual notion of a subset.

**Theorem 3.2.5.** *If  $A, B$ , and  $C$  are sets, if  $A = B$ , and if  $A \subseteq C$ , then  $B \subseteq C$ .*

*Proof.* For if  $A = B$  then  $B \subseteq A$  (Def. 3.1.3). But if  $B \subseteq A$  and  $A \subseteq C$ , then by the transitivity of inclusion,  $B \subseteq C$  (Thm. 3.2.3).  $\square$

More than just being transitive, inclusion is also reflexive.

### Theorem 3.2.6: Reflexivity of Inclusion

If  $A$  is a set, then  $A \subseteq A$ . ■

*Proof.* For if not, then by the definition of subset there is an  $x \in A$  such that  $x \notin A$  (Def. 3.1.2), a contradiction. Therefore  $A \subseteq A$ .  $\square$

The notion of inclusion ( $\subseteq$ ) is also antisymmetric. That is to say, if  $A \subseteq B$  and  $B \subseteq A$ , then  $A = B$ . This is simply the definition of equality (Def. 3.1.3). A relation that is transitive, reflexive, and antisymmetric is known as a partial ordering. The partial ordering of inclusion on the power set of some given set is the quintessential example of a partial order. On the other hand, the notion of containment is not reflexive. That is, for any set  $A$  it is true that  $A \notin A$  (Thm. 3.1.11). Indeed, this is desired to avoid Russell's Paradox. We now turn towards proving the familiar structure behind the notion of equality.

### Theorem 3.2.7: Reflexivity of Equality

If  $A$  is a set, then  $A = A$ . ■

*Proof.* For if  $A$  is a set then  $A \subseteq A$  (Thm. 3.2.6). Thus,  $A = A$  (Def. 3.1.3).  $\square$

And equally trivial theorem is the symmetry of equality, a direct corollary of the commutativity of the disjunction connective ( $\wedge$ ).

### Theorem 3.2.8: Symmetry of Equality

If  $A$  and  $B$  are sets and if  $A = B$ , then  $B = A$ . ■

*Proof.* For if  $A = B$ , then  $A \subseteq B$  and  $B \subseteq A$  (Def. 3.1.3). But then  $B \subseteq A$  and  $A \subseteq B$ , and thus  $B = A$  (Def. 3.1.3).  $\square$

An important but non-obvious statement is that containment ( $\in$ ) is *not* symmetric, and indeed is the exact opposite, it is antisymmetric. That is, for any two sets  $A$  and  $B$  it impossible for both  $A \in B$  and  $B \in A$ , for this would violate the axiom of regularity (Ax. 3.1.6). We now prove this claim rigorously.

### Theorem 3.2.9: Antisymmetry of Containment

---

If  $A$  and  $B$  are sets, and if  $A \in B$ , then  $B \notin A$ . ■

*Proof.* For suppose not, and suppose  $B \in A$ . But by Thm. 3.1.3, if  $A$  and  $B$  are sets, then there is a set  $\{A, B\}$  such that for all  $x$  it is true that  $x \in \{A, B\}$  if and only if  $x = A$  or  $x = B$ . But then  $\{A, B\}$  is a non-empty set, and thus by the axiom of regularity (Ax. 3.1.6) there is an  $x \in \{A, B\}$  such that  $x \cap \{A, B\} = \emptyset$ . But by hypothesis,  $A \in B$ , and thus  $A \in B \cap \{A, B\}$  (Def. 3.1.8), and therefore  $x \neq B$ . But similarly, if  $B \in A$ , then  $B \in A \cap \{A, B\}$  and thus  $x \neq A$ . But  $x \in \{A, B\}$  if and only if  $x = A$  or  $x = B$ , a contradiction. Therefore,  $B \notin A$ . □

An instant corollary of this is that  $\{A\} \notin A$ .

**Theorem 3.2.10.** *If  $A$  is a set, then  $\{A\} \notin A$ .*

*Proof.* For  $A \in \{A\}$  (Thm. 3.1.4). But if  $A \in \{A\}$ , then  $\{A\} \notin A$  (Thm. 3.2.9). □

We now continue developing the structure of equality.

### Theorem 3.2.11: Transitivity of Equality

---

If  $A$ ,  $B$ , and  $C$  are sets, if  $A = B$ , and if  $B = C$ , then  $A = C$ . ■

*Proof.* For if  $B = C$ , then  $C \subseteq B$  (Def. 3.1.3). But if  $A = B$ , then  $B = A$  (Thm. 3.2.6). But if  $B = A$  and  $C \subseteq B$ , then  $C \subseteq A$  (Thm. 3.2.4). And if  $A = B$ , then  $A \subseteq B$  (Def. 3.1.3). But if  $B = C$  and  $A \subseteq B$ , then  $A \subseteq C$  (Thm. 3.2.4). But it was proved that  $C \subseteq A$ , and thus  $A = C$  (Def. 3.1.3). □

The three properties we've proved thus far, that of reflexivity (Thm. 3.2.7), symmetry (Thm. 3.2.8), and transitivity (Thm. 3.2.11), are the key ingredients in defining *equivalence relations*. These are relations which are used to model the notion of equality in more abstract setting. We'll discuss these more in § 3.3.

**Theorem 3.2.12.** *If  $A$  and  $B$  are sets, and if  $A \subsetneq B$ , then there is an  $x \in B$  such that  $x \notin A$ .*

*Proof.* For suppose not. Then for all  $x \in B$  it is true that  $x \in A$ . But then  $B \subseteq A$  (Def. 3.1.2). But  $A \subseteq B$  and thus  $A = B$  (Def. 3.1.3). But  $A \subsetneq B$  and therefore  $A \neq B$ , a contradiction. □

**Theorem 3.2.13.** *If  $A$  is a set, then  $A$  is not a proper subset of  $\emptyset$ .*

*Proof.* For suppose not. Then there is an  $x \in \emptyset$  such that  $x \notin A$  (Thm. 3.2.12), a contradiction (Def. 3.2.1).  $\square$

### 3.2.2 Operations on Sets

Similar to the arithmetic of real numbers, there are standard operations that can be performed on sets to obtain new sets. The four most common operations are union, intersection, set difference, and symmetric difference. As stated before, we wish to build the structure of sets in an algebraic manner. To do this requires the notion that the operations of intersection and unions are *commutative*, *distributive*, have *identities*, and have *complements*.

#### Theorem 3.2.14: Commutative Law of Unions

If  $A$  and  $B$  are sets, then  $A \cup B = B \cup A$ . ■

*Proof.* For if  $x \in A \cup B$ , then either  $x \in A$  or  $x \in B$ , or both (Def. 3.1.7). But then either  $x \in B$  or  $x \in A$ , or both, and therefore  $x \in B \cup A$  (Def. 3.1.7). But then for all  $x \in A \cup B$  it is true that  $x \in B \cup A$ , and therefore  $A \cup B \subseteq B \cup A$  (Def. 3.1.2). Similarly,  $B \cup A \subseteq A \cup B$ , and thus  $A \cup B = B \cup A$  (Def. 3.1.3).  $\square$

Taking unions tends to produce *larger* sets, a property that is analogous to the addition of non-negative real numbers. Given any such values  $a, b \geq 0$  it is true that  $a \leq a + b$ . The same is true of unions. Similarly, if  $b \leq c$ , then  $a + b \leq a + c$ .

**Theorem 3.2.15.** *If  $A$  and  $B$  are sets, then  $A \subseteq A \cup B$ .*

*Proof.* For suppose not. Then there is an  $x \in A$  such that  $x \notin A \cup B$ . But if  $x \in A$ , then  $x \in A$  or  $x \in B$  and thus  $x \in A \cup B$  (Def. 3.1.7), a contradiction.  $\square$

**Theorem 3.2.16.** *If  $A$  and  $B$  are sets, then  $B \subseteq A \cup B$ .*

*Proof.* For  $A \cup B = B \cup A$  (Thm. 3.2.14) and  $B \subseteq B \cup A$  (Thm. 3.2.15). But if  $B \subseteq B \cup A$  and  $B \cup A = A \cup B$ , then  $B \subseteq A \cup B$  (Thm. 3.2.4).  $\square$

**Theorem 3.2.17.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $B \subseteq C$ , then  $A \cup B \subseteq A \cup C$ .*

*Proof.* For if  $x \in A \cup B$ , then either  $x \in A$ , or  $x \in B$ , or both (Def. 3.1.7). But  $B$  is a subset of  $C$ , and therefore if  $x \in B$ , then  $x \in C$  (Def. 3.1.2). Thus, if  $x \in A$  or  $x \in B$ , then  $x \in A$  or  $x \in C$ , and therefore  $x \in A \cup C$  (Def. 3.1.7). Therefore,  $A \cup B \subseteq A \cup C$  (Def. 3.1.2).  $\square$

**Theorem 3.2.18.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $B \subseteq C$ , then  $B \cup A \subseteq C \cup A$ .*

*Proof.* For  $B \cup A = A \cup B$  (Thm. 3.2.14). But if  $B \subseteq C$ , then  $A \cup B \subseteq A \cup C$  (Thm. 3.2.17). And if  $B \cup A = A \cup B$  and  $A \cup B \subseteq A \cup C$ , then  $B \cup A \subseteq A \cup C$  (Thm. 3.2.5). But  $A \cup C = C \cup A$  (Thm. 3.2.14) and if  $B \cup A \subseteq A \cup C$  and  $A \cup C = C \cup A$ , then  $B \cup A \subseteq C \cup A$  (Thm. 3.2.4).  $\square$

**Theorem 3.2.19.** *If  $A$ ,  $B$ ,  $C$ , and  $D$  are sets, if  $A \subseteq C$ , and if  $B \subseteq D$ , then  $A \cup B \subseteq C \cup D$ .*

*Proof.* For if  $B \subseteq D$ , then  $A \cup B \subseteq A \cup D$  (Thm. 3.2.17). But if  $A \subseteq C$ , then  $A \cup D \subseteq C \cup D$  (Thm. 3.2.18). But if  $A \cup B \subseteq A \cup D$  and  $A \cup D \subseteq C \cup D$ , then  $A \cup B \subseteq C \cup D$  (Thm. 3.2.3).  $\square$

Taking the union of subsets is redundant, as we simply obtain the larger set. This breaks down the analogy with arithmetic since there is only one *zero*. That is, there is only one number  $b$  such that  $a + b = a$  and that is  $b = 0$ . While any subset acts as a *zero* of a given set, the empty set has the property that it acts as a zero for *every* set. It is the only set with this property.

**Theorem 3.2.20.** *If  $A$  and  $B$  are sets, and if  $A \subseteq B$ , then  $A \cup B = B$ .*

*Proof.* For if  $A$  and  $B$  are sets, then  $B \subseteq A \cup B$  (Thm. 3.2.16). But if  $A \subseteq B$ , then for all  $x \in A$  it is true that  $x \in B$  (Def. 3.1.2). Thus if  $x \in A$  or if  $x \in B$ , then  $x \in B$ . But then, for all  $x \in A \cup B$ , it is true that  $x \in B$ , and therefore  $A \cup B \subseteq B$  (Def. 3.1.2). Thus,  $A \cup B = B$  (Def. 3.1.3).  $\square$

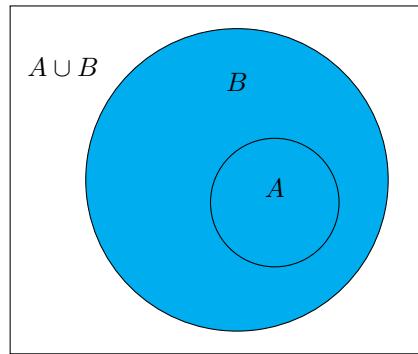


Fig. 3.15: Figure for Thm. 3.2.20

**Theorem 3.2.21.** *If  $A$ ,  $B$ , and  $C$  are sets, if  $A \subseteq C$ , and if  $B \subseteq C$ , then  $A \cup B \subseteq C$ .*

*Proof.* For if  $B \subseteq C$ , then  $A \cup B \subseteq A \cup C$  (Thm. 3.2.17). But if  $A \subseteq C$ , then  $A \cup C = C$  (Thm. 3.2.20). And if  $A \cup B \subseteq A \cup C$  and  $A \cup C = C$ , then  $A \cup B \subseteq C$  (Thm. 3.2.4).  $\square$

The converse of Thm. 3.2.20 can be proved as well.

**Theorem 3.2.22.** *If  $A$  and  $B$  are sets, and if  $A \cup B \subseteq A$ , then  $A \cup B = A$ .*

*Proof.* For  $A \subseteq A \cup B$  (Thm. 3.2.15). But by hypothesis,  $A \cup B \subseteq A$ . But then  $A = A \cup B$  (Def. 3.1.3).  $\square$

**Theorem 3.2.23.** *If  $A$  and  $B$  are sets, and if  $A \cup B \subseteq A$ , then  $B \subseteq A$ .*

*Proof.* For if  $A \cup B \subseteq A$ , then  $A \cup B = A$  (Thm. 3.2.22). And also,  $B \subseteq A \cup B$  (Thm. 3.2.16). But if  $A \cup B = A$  and  $B \subseteq A \cup B$ , then  $B \subseteq A$  (Thm. 3.2.4).  $\square$

We now prove our claim that the empty set is the only set that acts as a zero for *every* set one can name.

### Theorem 3.2.24: Identity Law of Unions

---

If  $A$  is a set, then  $A = \emptyset \cup A$  and  $A = A \cup \emptyset$ . ■

*Proof.* For  $\emptyset \subseteq A$  (Thm. 3.2.1) and therefore  $\emptyset \cup A = A$  (Thm. 3.2.20). By the commutative law (Thm. 3.2.14),  $\emptyset \cup A = A \cup \emptyset$  and thus by the transitivity of equality,  $A = A \cup \emptyset$  (Thm. 3.2.11).  $\square$

**Theorem 3.2.25.** *If  $A$  is a set such that, for any set  $B$  it is true that  $A \cup B = B$ , then  $A$  is the empty set.*

*Proof.* For suppose not. If  $A \neq \emptyset$ , then there is an  $x \in A$  (Def. 3.2.1). But if  $A$  is a set, then  $B = \{A\}$  is a set (Thm. 3.1.4). But then  $x \in A \cup B$  (Def. 3.1.7). But if  $x \in A$  then  $x \notin B$  (Thm. 3.1.12), and thus  $x \notin B$ . But then  $A \cup B \neq B$  (Def. 3.1.3), contradicting our hypothesis. Thus,  $A$  is the empty set.  $\square$

### Theorem 3.2.26: Idempotent Law of Unions

---

If  $A$  is a set, then  $A \cup A = A$ . ■

*Proof.* For  $A \subseteq A$  (Thm. 3.2.6) and thus  $A \cup A = A$  (Thm. 3.2.20).  $\square$

Logically, and pedagogically, it would seem appropriate to demonstrate the fact that union is an *associative* operation on sets. We will do this at the risk of being repetitive since similar results were proved for connectives in Chapt. 1, and will again be proven in the more general setting of a Boolean Algebra (see § 5.3)). Boolean algebras consist of two operations that are commutative, distributive, contain identities and complements. For the case of sets we will use union and intersection as our operations, and the power set of some given set as the set which these operations act on. Any system satisfying these properties can then prove that the associative law is true for both operations. For the sake of completeness, we present the classic arguments that unions and

intersections are associative, and save the algebraic proofs for the more general setting later.

**Theorem 3.2.27.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $A = B$ , then  $A \cup C = B \cup C$ .*

*Proof.* For if  $A = B$ , then  $A \subseteq B$  (Def. 3.1.3). But if  $A \subseteq B$ , then  $A \cup C \subseteq B \cup C$  (Thm. 3.2.18). But if  $A = B$ , then  $B \subseteq A$  (Def. 3.1.3), and therefore  $B \cup C \subseteq A \cup C$  (Thm. 3.2.18), and therefore  $A \cup C = B \cup C$  (Def. 3.1.3).  $\square$

**Theorem 3.2.28.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $B \subseteq A$ , then  $A \cup (B \cup C) = A \cup C$ .*

*Proof.* For since  $C \subseteq B \cup C$  (Thm. 3.2.16), we have that  $A \cup C \subseteq A \cup (B \cup C)$  (Thm. 3.2.17). But since  $B \subseteq A$ , it is true that  $B \cup C \subseteq A \cup C$  (Thm. 3.2.18). But then  $A \cup (B \cup C) \subseteq A \cup (A \cup C)$  (Thm. 3.2.17). But  $A \subseteq A \cup C$  (Thm. 3.2.15), and thus  $A \cup (A \cup C) = A \cup C$  (Thm. 3.2.20). But if  $A \cup (B \cup C) \subseteq A \cup (A \cup C)$  and  $A \cup (A \cup C) = A \cup C$ , then  $A \cup (A \cup B) \subseteq A \cup C$  (Thm. 3.2.4). But it was proved that  $A \cup C \subseteq A \cup (B \cup C)$ , and thus  $A \cup C = A \cup (B \cup C)$  (Def. 3.1.3).  $\square$

**Theorem 3.2.29.** *If  $A$ ,  $B$ , and  $C$  are sets, if  $B \subseteq C$ , then  $(A \cup B) \cup C = A \cup C$ .*

*Proof.* For by the commutative law of unions,  $A \cup B = B \cup A$ . But if  $A \cup B = B \cup A$ , then  $(A \cup B) \cup C = (B \cup A) \cup C$  (Thm. 3.2.27). But again by the commutative law,  $(B \cup A) \cup C = C \cup (B \cup A)$  (Thm. 3.2.14). But if  $B \subseteq C$ , then  $C \cup (B \cup A) = C \cup A$  (Thm. 3.2.28). But  $C \cup A = A \cup C$  (Thm. 3.2.14), and thus by the transitivity of equality,  $(A \cup B) \cup C = A \cup C$  (Thm. 3.2.11).  $\square$

With this set of rather trivial results, we can now quickly prove the associative law of unions.

### Theorem 3.2.30: Associative Law of Unions

---

If  $A$ ,  $B$ , and  $C$  are sets, then  $A \cup (B \cup C) = (A \cup B) \cup C$ . ■

*Proof.* For since  $B \subseteq A \cup B$  (Thm. 3.2.16), by Thm. 3.2.28 we have:

$$(A \cup B) \cup (B \cup C) = (A \cup B) \cup C \tag{3.2.2}$$

But since  $B \subseteq B \cup C$  (Thm. 3.2.15), by Thm. 3.2.29 we have:

$$(A \cup B) \cup (B \cup C) = A \cup (B \cup C) \tag{3.2.3}$$

Thus, by the transitivity of equality,  $A \cup (B \cup C) = (A \cup B) \cup C$  (Thm. 3.2.11).  $\square$

There is a notion called the *principle of duality* that applies to unions and intersections. Notably, any theorem that applies to unions also applies to intersections if we appropriately rearrange the symbols  $\subseteq$ ,  $\emptyset$ , and whatever universe

set  $U$  we are currently considering. In a given proposition we can take the universe to be the union of all of the sets hypothesized to exist in the statement of the theorem. This duality is made explicit by the distribute laws of unions and intersections, as well as the De Morgan Laws.

### Theorem 3.2.31: Commutative Law of Intersections

If  $A$  and  $B$  are sets, then  $A \cap B = B \cap A$ . ■

*Proof.* For if  $x \in A \cap B$ , then  $x \in A$  and  $x \in B$  (Def. 3.1.8). But then  $x \in B$  and  $x \in A$ , and therefore  $x \in B \cap A$  (Def. 3.1.8). But then for all  $x \in A \cap B$  it is true that  $x \in B \cap A$ , and therefore  $A \cup B \subseteq B \cup A$  (Def. 3.1.2). Similarly,  $B \cap A \subseteq A \cap B$ , and thus  $A \cap B = B \cap A$  (Def. 3.1.3). □

**Theorem 3.2.32.** *If  $A$  and  $B$  are sets, then  $A \cap B \subseteq A$ .*

*Proof.* If  $x \in A \cap B$ , then  $x \in A$  and  $x \in B$  (Def. 3.1.8), and thus  $x \in A$ . □

**Theorem 3.2.33.** *If  $A$  and  $B$  are sets, then  $A \cap B \subseteq B$ .*

*Proof.* For  $A \cap B = B \cap A$  (Thm. 3.2.31) and  $B \cap A \subseteq B$  (Thm. 3.2.32). But if  $B \cap A \subseteq B$  and  $B \cap A = A \cap B$ , then  $A \cap B \subseteq B$  (Thm. 3.2.5). □

**Theorem 3.2.34.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $B \subseteq C$ , then  $A \cap B \subseteq A \cap C$ .*

*Proof.* For if  $x \in A \cap B$ , then  $x \in A$  and  $x \in B$  (Def. 3.1.8). But  $B$  is a subset of  $C$  and thus if  $x \in B$ , then  $x \in C$  (Def. 3.1.2). But then  $x \in A$  and  $x \in C$ , and therefore  $x \in A \cap C$  (Def. 3.1.8). But then  $A \cap B \subseteq A \cap C$  (Def. 3.1.2). □

**Theorem 3.2.35.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $B \subseteq C$ , then  $B \cap A \subseteq C \cap A$ .*

*Proof.* For  $B \cap A = A \cap B$  (Thm. 3.2.31). But if  $B \subseteq C$ , then  $A \cap B \subseteq A \cap C$  (Thm. 3.2.34). But if  $B \cap A = A \cap B$  and  $A \cap B \subseteq A \cap C$ , then  $B \cap A \subseteq A \cap C$  (Thm. 3.2.5). But  $A \cap C = C \cap A$  (Thm. 3.2.31), and if  $B \cap A \subseteq A \cap C$  and  $A \cap C = C \cap A$ , then  $B \cap A \subseteq C \cap A$  (Thm. 3.2.4). □

**Theorem 3.2.36.** *If  $A$ ,  $B$ ,  $C$ , and  $D$  are sets, if  $A \subseteq C$ , and if  $B \subseteq D$ , then  $A \cap B \subseteq C \cap D$ .*

*Proof.* For if  $B \subseteq D$ , then  $A \cap B \subseteq A \cap D$  (Thm. 3.2.34). But if  $A \subseteq C$ , then  $A \cap D \subseteq C \cap D$  (Thm. 3.2.35). But if  $A \cap B \subseteq A \cap D$  and  $A \cap D \subseteq C \cap D$ , then  $A \cap B \subseteq C \cap D$  (Thm. 3.2.3). □

**Theorem 3.2.37.** *If  $A$  and  $B$  are sets, and if  $A \subseteq B$ , then  $A \cap B = A$ .*

*Proof.* For  $A \cap B \subseteq A$  (Thm. 3.2.32). But if  $A \subseteq B$ , then for all  $x \in A$  it is true that  $x \in B$  (Def. 3.1.2). Thus if  $x \in A$ , then  $x \in A$  and  $x \in B$ , and therefore  $x \in A \cap B$  (Def. 3.1.8), hence  $A \subseteq A \cap B$ . Thus,  $A = A \cap B$  (Def. 3.1.3). □

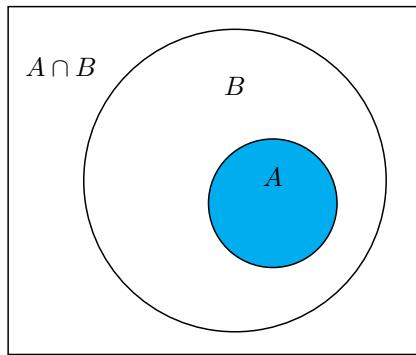


Fig. 3.16: Visual for Thm. 3.2.37.

**Theorem 3.2.38.** If  $A$ ,  $B$ , and  $C$  are sets, if  $A \subseteq B$ , if  $B \subseteq C$ , and if  $A \cap C = B$ , then  $A = B$ .

*Proof.* For if  $A \subseteq B$  and  $B \subseteq C$ , then  $A \subseteq C$  (Thm. 3.2.3). But if  $A \subseteq C$ , then  $A \cap C = A$  (Thm. 3.2.37). But  $A \cap C = B$ , and therefore  $A = B$  (Thm. 3.2.11).  $\square$

**Theorem 3.2.39.** If  $A$ ,  $B$ , and  $C$  are sets, if  $A \subseteq B$ , and if  $A \subseteq C$ , then  $A \subseteq B \cap C$ .

*Proof.* For if  $A \subseteq B$ , then  $A \cap C \subseteq B \cap C$  (Thm. 3.2.35). But if  $A \subseteq C$ , then  $A \cap C = A$  (Thm. 3.2.37). But if  $A = A \cap C$  and  $A \cap C \subseteq B \cap C$ , then  $A \subseteq A \cap C$  (Thm. 3.2.5).  $\square$

**Theorem 3.2.40.** If  $A$  and  $B$  are sets, and if  $A \subseteq A \cap B$ , then  $A = A \cap B$ .

*Proof.* For  $A \cap B \subseteq A$  (Thm. 3.2.32). But by hypothesis,  $A \subseteq A \cap B$ , and thus  $A = A \cap B$  (Def. 3.1.3).  $\square$

**Theorem 3.2.41.** If  $A$  and  $B$  are sets, and if  $A = A \cap B$ , then  $A \subseteq B$ .

*Proof.* For if  $A = A \cap B$ , then  $A \subseteq A \cap B$  (Def. 3.1.3). But  $A \cap B \subseteq B$  (Thm. 3.2.33) and thus by the transitivity of inclusion,  $A \subseteq B$  (Thm. 3.2.3).  $\square$

In Thm. 3.2.24 we proved that the empty set is an identity for unions. That is, for any set  $A$  it is true that  $A = \emptyset \cup A$  and  $A = A \cup \emptyset$ . There is no legal identity law of intersections. We proved that for any two sets  $A$  and  $B$  such that  $A \subseteq B$ , it is then true that  $A \cap B = A$ . If there were some *universe set*  $U$ , then  $A \cap U = A$  and thus  $U$  acts as the identity of intersections. In many contexts there is such a universe, notably the power set of some given set that is of current consideration, but there is no general universe set since this amounts to a set of all sets, which we've proved is impossible under the axioms of ZFC.

**Theorem 3.2.42.** *If  $A$  is a set, then  $\emptyset \cap A = \emptyset$ .*

*Proof.* For  $\emptyset \subseteq A$  (Thm. 3.2.1), and therefore  $\emptyset \cap A = \emptyset$  (Thm. 3.2.37).  $\square$

**Theorem 3.2.43.** *If  $A$  is a set such that for any set  $B$  it is true that  $A \cap B = A$ , then  $A = \emptyset$ .*

*Proof.* For  $A \cap \emptyset = \emptyset \cap A = \emptyset$  (Thm. 3.2.31) and  $\emptyset \cap A = \emptyset$  (Thm. 3.2.42). Thus, by the transitivity of equality,  $A \cap \emptyset = \emptyset$  (Thm. 3.2.11). But by hypothesis,  $A \cap \emptyset = A$ . Again by transitivity,  $A = \emptyset$  (Thm. 3.2.11).  $\square$

### Theorem 3.2.44: Idempotent Law of Intersections

---

If  $A$  is a set, then  $A \cap A = A$ . ■

*Proof.* For  $A \subseteq A$  (Thm. 3.2.6) and thus  $A \cap A = A$  (Thm. 3.2.37).  $\square$

**Theorem 3.2.45.** *If  $A$ ,  $B$ , and  $C$  are sets, if  $A \subseteq C$ , and if  $B \subseteq C$ , then  $A \cap B \subseteq C$ .*

*Proof.* For if  $A \subseteq C$  and  $B \subseteq C$ , then  $A \cap B \subseteq C \cap C$  (Thm. 3.2.36). But  $C \cap C = C$  (Thm. 3.2.44). But if  $A \cap B \subseteq C \cap C$  and  $C \cap C = C$ , then  $A \cap B \subseteq C$  (Thm. 3.2.4).  $\square$

**Theorem 3.2.46.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $B = C$ , then  $A \cap B = A \cap C$ .*

*Proof.* For if  $B = C$ , then  $B \subseteq C$  (Def. 3.1.3). But if  $B \subseteq C$ , then  $A \cap B \subseteq A \cap C$  (Thm. 3.2.34). But if  $B = C$ , then  $C \subseteq B$  (Def. 3.1.3). But if  $C \subseteq B$ , then  $A \cap C \subseteq A \cap B$  (Thm. 3.2.34). But it was just proved that  $A \cap B \subseteq A \cap C$ , and thus  $A \cap B = A \cap C$  (Def. 3.1.3).  $\square$

**Theorem 3.2.47.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $A \subseteq B$ , then  $A \cap (B \cap C) = A \cap C$ .*

*Proof.* For since  $B \cap C \subseteq C$  (Thm. 3.2.33) it is true that  $A \cap (B \cap C) \subseteq A \cap C$  (Thm. 3.2.34). But since  $A \subseteq B$  it is true that  $A \cap C \subseteq B \cap C$  (Thm. 3.2.35). But then  $A \cap (A \cap C) \subseteq A \cap (B \cap C)$  (Thm. 3.2.34). But  $A \cap C \subseteq A$  (Thm. 3.2.32) and thus  $A \cap (A \cap C) = A \cap C$  (Thm. 3.2.37). But if  $A \cap C = A \cap (A \cap C)$  and  $A \cap (A \cap C) \subseteq A \cap (B \cap C)$ , then  $A \cap C \subseteq A \cap (B \cap C)$  (Thm. 3.2.5). Thus,  $A \cap (B \cap C) = A \cap C$  (Def. 3.1.3).  $\square$

**Theorem 3.2.48.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $C \subseteq B$ , then  $(A \cap B) \cap C = A \cap C$ .*

*Proof.* For  $A \cap B = B \cap A$  (Thm. 3.2.31), and thus  $(A \cap B) \cap C = (B \cap A) \cap C$  (Thm. 3.2.46). But  $(B \cap A) \cap C = C \cap (B \cap A)$  (Thm. 3.2.31). But if  $C \subseteq B$ , then  $C \cap (B \cap A) = C \cap A$  (Thm. 3.2.47). But  $C \cap A = A \cap C$  (Thm. 3.2.31), and thus by transitivity,  $(A \cap B) \cap C = A \cap C$  (Thm. 3.2.11).  $\square$

And now we completely mimic the proof of the associative law of unions, and derive the associativity of intersections.

### Theorem 3.2.49: Associative Law of Intersections

---

If  $A$ ,  $B$ , and  $C$  are sets, then  $A \cap (B \cap C) = (A \cap B) \cap C$ . ■

*Proof.* For since  $A \cap B \subseteq B$  (Thm. 3.2.33), by Thm. 3.2.47 we have:

$$(A \cap B) \cap (B \cap C) = (A \cap B) \cap C \quad (3.2.4)$$

But since  $B \cap C \subseteq B$  (Thm. 3.2.32), by Thm. 3.2.48 we have:

$$(A \cap B) \cap (B \cap C) = A \cap (B \cap C) \quad (3.2.5)$$

Thus, by the transitivity of equality,  $A \cap (B \cap C) = (A \cap B) \cap C$  (Thm. 3.2.11). □

We now move on to the distributive laws of unions and intersections. Thus far the theorems proved have dealt solely with unions or with intersections, but have not mixed the two. Associativity now tells us that an expression of the form  $A \cup B \cup C$  is no longer ambiguous since  $A \cup (B \cup C) = (A \cup B) \cup C$ , and thus we may rid ourselves of the burden of writing parentheses. Similarly,  $A \cap B \cap C$  is well-defined now. There is still an air of mystery surrounding an expression such as  $A \cap B \cup C$ . Should this mean  $A \cap (B \cup C)$  or  $(A \cap B) \cup C$ , and are these equal? The distributive laws tell us how to simplify such a thing, but unfortunately these theorems tell us that parentheses are essential for when mixing intersections and unions. That is, as we will see, an expression of the form  $A \cap B \cup C$  is indeed undefined without the introduction of parentheses:  $A \cap (B \cup C)$  need not be equal to  $(A \cap B) \cup C$ .

**Theorem 3.2.50.** *If  $A$ ,  $B$ , and  $C$  are sets, then  $B \cap C \subseteq (A \cup B) \cap (A \cup C)$ .*

*Proof.* For by Thm. 3.2.16 we have that  $B \subseteq A \cup B$  and  $C \subseteq A \cup C$ , and therefore  $B \cap C \subseteq (A \cup C) \cap (A \cup C)$  (Thm. 3.2.36). □

**Theorem 3.2.51.** *If  $A$ ,  $B$ , and  $C$  are sets, then  $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$ .*

*Proof.* For  $A \subseteq A \cup B$  and  $A \subseteq A \cup C$  (Thm. 3.2.15), and therefore it is true that  $A \cap A \subseteq (A \cup B) \cap (A \cup C)$  (Thm. 3.2.19). But  $A = A \cap A$  (Thm. 3.2.26), and if  $A = A \cap A$  and  $A \cap A \subseteq (A \cup B) \cap (A \cup C)$ , then  $A \subseteq (A \cup B) \cap (A \cup C)$  (Thm. 3.2.5). But  $B \cap C \subseteq (A \cup B) \cap (A \cup C)$  (Thm. 3.2.50), and therefore:

$$A \cap (B \cup C) \subseteq ((A \cup B) \cap (A \cup C)) \cap ((A \cup B) \cap (A \cup C)) = (A \cup B) \cap (A \cup C) \quad (3.2.6)$$

Where this last equality is by the idempotent law of intersections (Thm. 3.2.44). But then  $A \cap (B \cup C) \subseteq (A \cup B) \cap (A \cup C)$  (Thm. 3.2.4). □

This proves half of the distributive law of unions which states that  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ . That is, unions distribute over intersections. To prove the opposite direction of this equality seems to resist algebraic efforts and one must rely directly on the use of the containment symbol ( $\in$ ). However, once one has one of the distributive laws, the latter may be proved with purely algebraic arguments. We will do so presently.

**Theorem 3.2.52: Distributive Law of Unions**

If  $A$ ,  $B$ , and  $C$  are sets, then:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

*Proof.* For suppose not. But  $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$  (Thm. 3.2.51), and therefore  $A \cup (B \cap C) \subsetneq (A \cup B) \cap (A \cup C)$  (Def. 3.1.4). But then there is an  $x \in (A \cup B) \cap (A \cup C)$  such that  $x \notin A \cup (B \cap C)$  (Thm. 3.2.12). But if  $x \notin A \cup (B \cap C)$ , then  $x \notin A$  and  $x \notin B \cap C$  (Def. 3.1.7). But  $x \in (A \cup B) \cap (A \cup C)$ , and therefore  $x \in A \cup B$  and  $x \in A \cup C$  (Def. 3.1.8). But  $x \notin A$ , and thus if  $x \in A \cup B$  then it must be true that  $x \in B$ . And similarly, since  $x \in A \cup C$ , it must be true that  $x \in C$ . But then  $x \in B$  and  $x \in C$ , and therefore  $x \in B \cap C$  (Def. 3.1.8), a contradiction. This completes the proof.  $\square$

We can visualize Thm. 3.2.52 with Venn diagrams (Fig. 3.17).

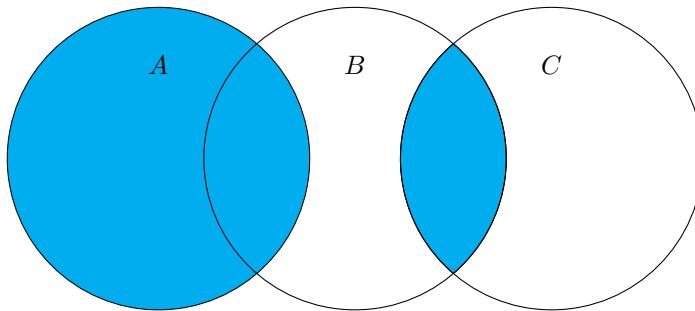


Fig. 3.17: Venn Diagram for Distributive Law of Unions

**Theorem 3.2.53: Distributive Law of Intersections**

If  $A$ ,  $B$ , and  $C$  are sets, then:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

*Proof.* For  $(A \cap B) \cup (A \cap C) = ((A \cap B) \cup A) \cap ((A \cap B) \cup C)$  (Thm. 3.2.52). But  $A \cap B \subseteq A$  (Thm. 3.2.32) and therefore  $(A \cap B) \cup A = A$  (Thm. 3.2.20). But then  $((A \cap B) \cup A) \cap ((A \cap B) \cup C) = A \cap ((A \cap B) \cup C)$  (Thm. 3.2.46). But  $(A \cap B) \cup C = C \cup (A \cap B)$  (Thm. 3.2.14) and  $C \cup (A \cap B) = (C \cup A) \cap (C \cup B)$  (Thm. 3.2.52). But then  $A \cap ((A \cap B) \cup C) = A \cap ((C \cup A) \cap (C \cup B))$

(Thm. 3.2.46). But  $A \cap ((C \cup A) \cap (C \cup B)) = (A \cap (C \cup A)) \cap (C \cup B)$  (Thm. 3.2.30). And since  $A \subseteq C \cup A$  (Thm. 3.2.16), we have that  $A \cap (C \cup A) = A$  (Thm. 3.2.37). But then  $(A \cap (C \cup A)) \cap (C \cup B) = A \cap (C \cup B)$  (Thm. 3.2.46). But  $C \cup B = B \cup C$  (Thm. 3.2.14), and thus  $A \cap (C \cup B) = A \cap (B \cup C)$  (Thm. 3.2.46). Thus by transitivity,  $(A \cap B) \cup (A \cap C) = A \cap (B \cup C)$  (Thm. 3.2.11).  $\square$

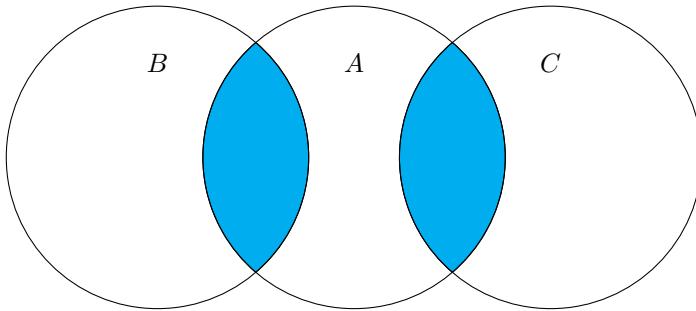


Fig. 3.18: Venn Diagram for the Distributive Law of Intersections

**Theorem 3.2.54.** If  $A$ ,  $B$ , and  $C$  are sets, then  $A \cap (B \cup C) \subseteq A \cup (B \cap C)$ .

*Proof.* For  $A \cap (B \cup C) \subseteq A$  (Thm. 3.2.32) and  $A \subseteq A \cup (B \cap C)$  (Thm. 3.2.15), and thus  $A \cap (B \cup C) \subseteq A \cup (B \cap C)$  (Thm. 3.2.3).  $\square$

We can use Venn diagrams to visualize this more clearly (see Fig. 3.19).

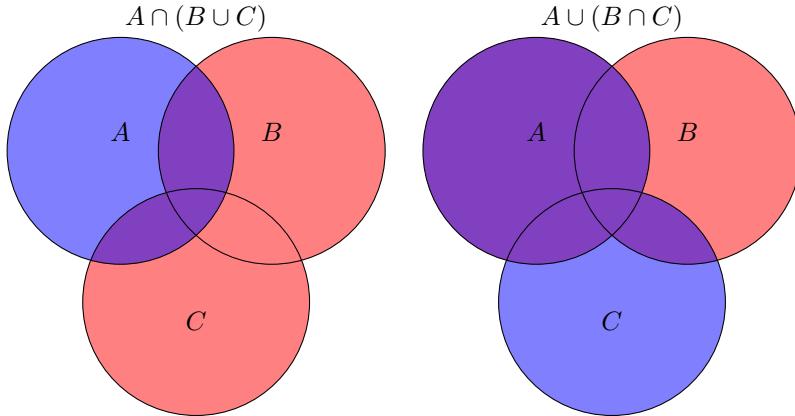


Fig. 3.19: Figure for Thm. 3.2.54

We now introduce a new set operation called *set difference*. In many respects this is analogous to the operation of subtraction one finds in arithmetic. One

should be careful since there are many situations in which the analogy breaks down.

**Theorem 3.2.55.** *If  $A$  and  $B$  are sets, then there is a set  $C$  such that for all  $x$  it is true that  $x \in C$  if and only if  $x \in A$  and  $x \notin B$ .*

*Proof.* For let  $P$  be the proposition *true if  $x \notin B$ , false otherwise*. Then by the axiom schema of specification (Ax. 3.1.3), there is a set  $C$  such that:

$$C = \{ x \in A \mid P(x) \} \quad (3.2.7)$$

But then  $x \in C$  if and only if  $x \in A$  and  $P(x)$  is true. But  $P(x)$  is true if and only if  $x \notin B$ . Thus  $x \in C$  if and only if  $x \in A$  and  $x \notin B$ .  $\square$

The set described in Thm. 3.2.55 is known as the set difference of  $B$  with respect to  $A$ . We take a moment to discuss some of its properties, as well as the related notion of *complement*

### Definition 3.2.2: Set Difference

The set difference of a set  $B$  with respect to a set  $A$  is the set:

$$A \setminus B = \{ x \in A \mid x \notin B \}$$

**Example 3.2.2** Let  $\mathbb{Z}$  denote the integers, and  $\mathbb{N}$  denote the set of natural numbers. Then  $\mathbb{Z} \setminus \mathbb{N}$  is the set of all negative integers. Flipping this, we see that  $\mathbb{N} \setminus \mathbb{Z}$  is the empty set since there are no non-negative integers that are not also integers.

**Example 3.2.3** Letting  $\mathbb{R}$  denote the real numbers and  $\mathbb{Q}$  denote the rationals,  $\mathbb{R} \setminus \mathbb{Q}$  is the set of all *irrational* numbers. Famous examples include  $\sqrt{2}$ ,  $\pi$ , and  $e$  (sometimes known as Euler's constant).

**Example 3.2.4** If we let  $A$  and  $B$  be defined as:

$$A = \{ a, b, c \} \quad (3.2.8a) \qquad B = \{ b, c, d \} \quad (3.2.8b)$$

Then we can compute directly the set difference between the two:

$$A \setminus B = \{ a \} \quad (3.2.9a) \qquad B \setminus A = \{ d \} \quad (3.2.9b)$$

Thus set difference between sets is not a *commutative* set operation.

Set difference can be visualized, much like unions and intersections, via the use of Venn diagrams (Fig. 3.20).

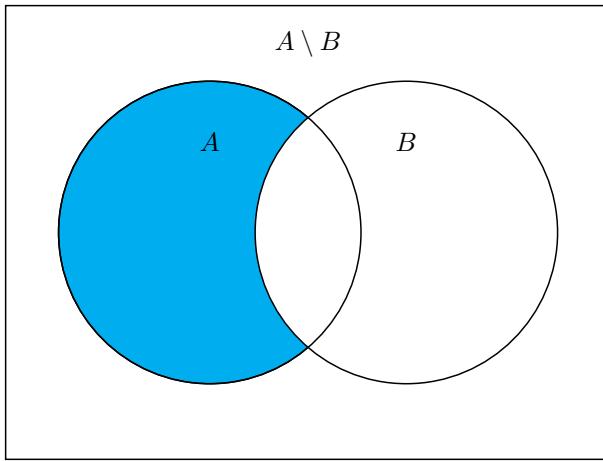


Fig. 3.20: Venn Diagram for Set Difference

**Theorem 3.2.56.** *If  $A$  is a set, then  $A \setminus A = \emptyset$ .*

*Proof.* For suppose not. Then there is an  $x \in A \setminus A$  (Def. 3.2.1). But then  $x \in A$  and  $x \notin A$  (Def. 3.2.2), a contradiction.  $\square$

**Theorem 3.2.57.** *If  $A$  and  $B$  are sets, then  $A \setminus B \subseteq A$ .*

*Proof.* For suppose not. Then there is an  $x \in A \setminus B$  such that  $x \notin A$ . But  $x \in A \setminus B$  if and only if  $x \in A$  and  $x \notin B$  (Def. 3.2.2), and thus  $x \in A$ , a contradiction.  $\square$

**Theorem 3.2.58.** *If  $A$  is a set, then  $A \setminus \emptyset = A$ .*

*Proof.* For suppose not. Since  $A \setminus \emptyset \subseteq A$  (Thm. 3.2.57), there is an  $x \in A$  such that  $x \notin A \setminus \emptyset$  (Def. 3.1.3). But if  $x \in A$  and  $x \notin A \setminus \emptyset$ , then  $x \in \emptyset$  (Def. 3.2.2) a contradiction since  $x \notin \emptyset$  (Def. 3.2.1).  $\square$

While set difference appears similar to subtraction, the two have their differences. For any two real numbers  $a$  and  $b$ , it is always true that  $b = a - (a - b)$ . For sets this is not true. For let  $A$  be the empty set, and let  $B$  be non-empty. Then  $A \setminus (A \setminus B) = \emptyset$ , which is not  $B$ . Set differences can not be easily simplified. The notion is not associative, nor is it commutative. If there is a larger *universe* set, then set difference can be related to intersection.

**Example 3.2.5** More than being a non-commutative operation, set difference is not *associative* either. For let  $A$  be any non-empty set and let  $A = B = C$ . Then:

$$A \setminus (A \setminus A) = A \setminus \emptyset = A \quad (3.2.10)$$

Flipping this around, we have:

$$(A \setminus A) \setminus A = \emptyset \setminus A = \emptyset \quad (3.2.11)$$

But since  $A$  is a non-empty set,  $A \neq \emptyset$ , thus showing that set difference is not associative.

**Theorem 3.2.59.** *If  $A$  and  $B$  are sets, and if  $A \setminus B = B \setminus A$ , then  $A = B$ .*

*Proof.* For suppose not. Then either  $A \not\subseteq B$  or  $B \not\subseteq A$  (Def. 3.1.3). Suppose  $A \not\subseteq B$ . Then there is an  $x \in A$  such that  $x \notin B$  (Def. 3.1.2). But then  $x \in A \setminus B$  (Def. 3.2.2). And by hypothesis  $A \setminus B = B \setminus A$ , and thus  $x \in B \setminus A$  (Def. 3.1.3). But then  $x \in B$  and  $x \notin A$ , a contradiction since  $x \in A$ . Therefore  $A \subseteq B$ , and similarly  $B \subseteq A$ . Hence, equality (Def. 3.1.3).  $\square$

It then follows that if  $A \setminus B = B \setminus A$ , then both are equal to the empty set.

**Theorem 3.2.60.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $B \subseteq C$ , then  $B \setminus A \subseteq C \setminus A$ .*

*Proof.* For If  $x \in B \setminus A$ , then  $x \in B$  and  $x \notin A$  (Def. 3.2.2). But  $B \subseteq C$ , and thus if  $x \in B$ , then  $x \in C$  (Def. 3.1.2). But then  $x \in C$  and  $x \notin A$ , and therefore  $x \in C \setminus A$  (Def. 3.2.2). Thus,  $B \setminus A \subseteq C \setminus A$  (Def. 3.1.2).  $\square$

If one is afforded the language of a universe set, then one can define the complement of a set  $A$  as:

$$A^C = U \setminus A \quad (3.2.12)$$

Such notation is particularly attractive when one discusses the DeMorgan laws. We can relate complements back to set difference by combining these notions with intersections:

$$B \setminus A = B \cap A^C \quad (3.2.13)$$

We'll now prove this claim, but avoid the use of a universal set.

**Theorem 3.2.61.** *If  $A$ ,  $B$ , and  $U$  are sets, if  $A \subseteq U$ , and if  $B \subseteq U$ , then  $B \setminus A = B \cap (U \setminus A)$ .*

*Proof.* For if  $B \subseteq U$ , then  $B \setminus A \subseteq U \setminus A$  (Thm. 3.2.60). But  $B \setminus A \subseteq B$  (Thm. 3.2.57), and thus by Thm. 3.2.36 and by the idempotent law of intersections (Thm. 3.2.44), we have:

$$B \setminus A = (B \setminus A) \cap (B \setminus A) \subseteq B \cap (U \setminus A) \quad (3.2.14)$$

And therefore  $B \setminus A \subseteq B \cap (U \setminus A)$  (Thm. 3.2.5).  $\square$

**Definition 3.2.3:** Symmetric Difference

The symmetric difference of  $A$  and  $B$ , denoted  $A \ominus B$ , is the set:

$$A \ominus B = (A \cup B) \setminus (A \cap B)$$

**Theorem 3.2.62.** If  $A$ ,  $B$ , and  $C$  are sets, and if  $A \subseteq C$  and  $B \subseteq C$ , then:

$$B \setminus A = B \cap (C \setminus A) \quad (3.2.15)$$

*Proof.* For if  $x \in B \setminus A$ , then  $x \in B$  and  $x \notin A$ . But  $B \subseteq C$ , and thus if  $x \in B$ , then  $x \in C$ . But if  $x \notin A$ , then  $x \in C \setminus A$ . Therefore  $B \setminus A \subseteq B \cap (C \setminus A)$ . Similarly,  $B \cap (C \setminus A) \subseteq B \setminus A$ , and therefore  $B \setminus A = B \cap (C \setminus A)$ .  $\square$

We can draw a Venn diagram for the symmetric difference, see Fig. 3.21.

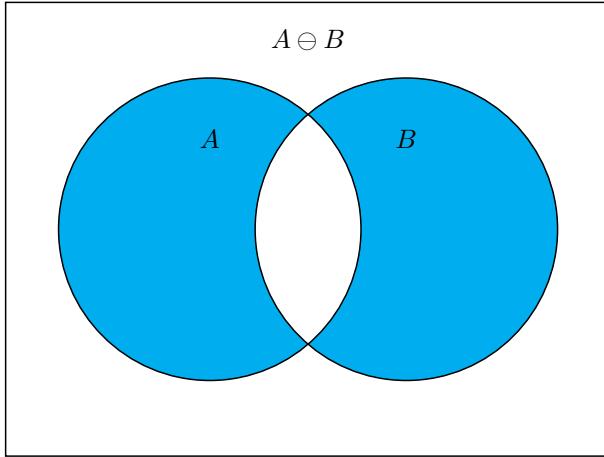


Fig. 3.21: Venn Diagram for Symmetric Difference

The concept of set difference can then be used to define the concept of complement. Thm. 3.2.62 can then be translated into the notation of complements as follows:

**Theorem 3.2.63.** If  $A$ ,  $B$ , and  $\Omega$  are sets,  $A, B \subseteq \Omega$ , and if  $A^C$  is the complement of  $A$  with respect to  $\Omega$ , then:

$$B \setminus A = B \cap A^C \quad (3.2.16)$$

*Proof.* By the definition of complement,  $A^C = \Omega \setminus A$ . As  $A \subseteq \Omega$  and  $B \subseteq \Omega$ , by Thm. 3.2.62,  $B \setminus A = B \cap (\Omega \setminus A)$ , and therefore  $B \setminus A = B \cap A^C$ .  $\square$

The main result about complements are known as DeMorgan's Laws. The laws relate unions and intersections by means of complements. The general laws hold for arbitrary unions and arbitrary intersections, as will be shown later.

**Theorem 3.2.64: De Morgan's Laws**

If  $A$ ,  $B$ , and  $\Omega$  are sets, if  $A \subseteq \Omega$  and  $B \subseteq \Omega$ , then:

$$(A \cap B)^C = A^C \cup B^C \quad (3.2.17a)$$

$$(A \cup B)^C = A^C \cap B^C \quad (3.2.17b)$$

With this, we can prove some results about set differences.

**Theorem 3.2.65.** *If  $A$  and  $B$  are sets, then:*

$$A = (A \cap B) \cup (A \setminus B) \quad (3.2.18)$$

*Proof.* For let  $\Omega = A \cup B$ . Then  $A \subseteq \Omega$  and  $B \subseteq \Omega$ , and thus:

$$(A \cap B) \cup (A \setminus B) = (A \cap B) \cup (A \cap B^C) \quad (3.2.19a)$$

$$= A \cap (B \cup B^C) \quad (3.2.19b)$$

$$= A \cap \Omega \quad (3.2.19c)$$

But by Thm. 3.2.32,  $A \cap \Omega = A$ . Therefore, etc.  $\square$

**Theorem 3.2.66.** *If  $A$ ,  $B$ , and  $C$  are sets, then:*

$$A \cap (B \setminus C) = (A \cap B) \cap (A \setminus C) \quad (3.2.20)$$

*Proof.* For:

$$A \cap (B \setminus C) = A \cap (B \cap C^C) \quad (3.2.21a)$$

$$= (A \cap A) \cap (B \cap C^C) \quad (3.2.21b)$$

$$= (A \cap B) \cap (A \cap C^C) \quad (3.2.21c)$$

$$= (A \cap B) \cap (A \setminus C) \quad (3.2.21d)$$

$\square$

Intersections do distribute over set differences.

**Theorem 3.2.67.** *If  $A$ ,  $B$ , and  $C$  are sets, then:*

$$A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C) \quad (3.2.22)$$

*Proof.* For:

$$(A \cap B) \setminus (A \cap C) = (A \cap B) \cap (A \cap C)^C \quad (3.2.23a)$$

$$= (A \cap B) \cap (A^C \cup C^C) \quad (3.2.23b)$$

$$= [(A \cap B) \cap A^C] \cup [(A \cap B) \cap C^C] \quad (3.2.23c)$$

$$= [(A \cap A^C) \cap B] \cup [(A \cap B) \cap C^C] \quad (3.2.23d)$$

$$= \emptyset \cup [(A \cap B) \cap C^C] \quad (3.2.23e)$$

$$= (A \cap B) \cap C^C \quad (3.2.23f)$$

$$= A \cap (B \cap C^C) \quad (3.2.23g)$$

$$= A \cap (B \setminus C) \quad (3.2.23h)$$

Therefore, etc.  $\square$

Unions do not, however. For let  $A$  be non-empty and let  $A = B = C$ . Then  $A \cup (B \setminus C) = A$ , but  $(A \cup B) \setminus (A \cup C) = \emptyset$ .

**Theorem 3.2.68.** *If  $A$  and  $B$  are sets and  $A \subset B$ , then  $B \setminus (B \setminus A) = A$ .*

*Proof.* For:

$$Y_o \quad (3.2.24)$$

$$\begin{aligned} [x \in B \setminus (B \setminus A)] &\Rightarrow [x \in B \wedge x \notin \{x \in B : x \notin A\}] \Rightarrow [x \in A \subset B]. \\ [x \in A] &\Rightarrow [x \notin B \setminus A] \Rightarrow [x \in B \setminus (B \setminus A)]. \end{aligned} \quad \square$$

The previous theorem shows that  $(A^C)^C = A$ .

**Theorem 3.2.69.** *If  $A$ ,  $B$ , and  $C$  are sets, and if  $A \subseteq C$  and  $B \subseteq C$ , then:*

$$B \setminus A = B \cap (C \setminus A) \quad (3.2.25)$$

*Proof.* For if  $x \in B \setminus A$ , then  $x \in B$  and  $x \notin A$ . But  $B \subseteq C$ , and thus if  $x \in B$ , then  $x \in C$ . But if  $x \notin A$ , then  $x \in C \setminus A$ . Therefore  $B \setminus A \subseteq B \cap (C \setminus A)$ . Similarly,  $B \cap (C \setminus A) \subseteq B \setminus A$ , and therefore  $B \setminus A = B \cap (C \setminus A)$ .  $\square$

While set difference appears similar to subtraction that one finds in basic arithmetic, the two have their differences. For any two real numbers  $a$  and  $b$ ,  $b = a - (a - b)$ . For sets this is not true. For let  $A$  be the empty set, and let  $B$  be non-empty. Then  $A \setminus (A \setminus B) = \emptyset$ , which is not  $B$ . Also, while it may seem convincing that  $A \setminus (B \setminus A) = A \setminus B$ , this is not true. For let  $A$  be a non-empty set and let  $B = A$ . Then  $A \setminus (B \setminus A) = A$ , but  $A \setminus B = \emptyset$ . The concept of set difference can then be used to define the concept of complement.

### 3.3 Relations

#### Definition 3.3.1: Relation on a Set

A **relation** on a **set**  $A$  is a **subset**  $R$  of the **Cartesian product**  $A \times A$ .

We use a special notation for relations on a set.

#### Notation 3.3.1: Relation Notation

If  $A$  is a set, if  $R$  is a relation on  $A$ , and if  $(a, b) \in R$ , we write  $aRb$ .

$$\forall_x \forall_y (aRb) \Leftrightarrow ((a, b) \in R)$$

For a relation  $R$  it is not necessary true that  $aRb$  implies  $bRa$ , nor is it necessarily true that  $aRa$ . These are called symmetric and reflexive relations, respectively.

#### Example 3.3.1: Examples of Relations

Let  $A = \mathbb{R}$  and consider the relation of equality. That is, let  $R_=\subseteq \mathbb{R}^2$  be defined by:

$$R_-=\{(x, y) \in \mathbb{R}^2 \mid x = y\} \quad (3.3.1)$$

Then  $R_-$  is a relation on  $\mathbb{R}^2$ . Rather than writing  $(x, y) \in R_-$  or  $xR_-y$  we commonly write  $x = y$ . Note that this relation is defined entirely by the *diagonal* of the Cartesian product  $\mathbb{R} \times \mathbb{R}$ . Another simple relation is that of ordering. Let  $R_<$  be defined as follows:

$$R_<=\{(x, y) \in \mathbb{R}^2 \mid x < y\} \quad (3.3.2)$$

This is also a relation since it is a subset of the Cartesian product, but it is a slightly more complicated one. There are many *off-diagonal* elements of this relation. ■

**Theorem 3.3.1.** *If  $B$  is a set, if  $A \subseteq B$ , and if  $R$  is a relation on  $B$ , then there is a relation  $R_A$  such that  $R_A$  is a relation on  $A$  and  $R_A \subseteq R$ .*

*Proof.* For let  $P$  be the proposition *True if  $(x, y) \in A^2$ , false otherwise*. By the axiom schema of specification (Ax. 3.1.3) there is a set:

$$R_A = \{(x, y) \in R \mid P((x, y))\} \quad (3.3.3)$$

But then  $(x, y) \in R_A$  if and only if  $(x, y) \in R$  and  $(x, y) \in A^2$ . □

This set is called the *restriction* of  $R$  to the subset  $A$ .

### Definition 3.3.2: Restriction of a Relation

The restriction of a relation  $R$  on a set  $B$  to a subset  $A$  is the set  $R_A$  defined by:

$$R_A = \{ (x, y) \in R \mid (x, y) \in A^2 \}$$

There are many basic properties that relations have, and we prove them now.

**Theorem 3.3.2.** *If  $A$  is a set, then  $A \times A$  is a relation on  $A$ .*

*Proof.* For if  $A$  is a set, then  $A \times A \subseteq A \times A$ . Therefore, etc. □

**Theorem 3.3.3.** *If  $A$  is a set, and then  $\emptyset$  is a relation on  $A$ .*

*Proof.* For if  $A$  is a set, then  $\emptyset \subseteq A \times A$ . Therefore, etc. □

**Theorem 3.3.4.** *Set inclusion  $\subseteq$  is a relation. Proper set inclusion  $\subsetneq$  is a relation. These define partial orderings.*

### Definition 3.3.3: Domain of a Relation

The **domain** of a **relation**  $R$  on a **set**  $A$  is the set:

$$\text{dom}(R) = \{a \in A \mid \exists b \in A \text{ such that } aRb\}$$

### Definition 3.3.4: Range of a Relation

The **range** of a **relation**  $R$  on a **set**  $A$  is the set:

$$\text{ran}(R) = \{b \in A \mid \exists a \in A \text{ such that } aRb\}$$

**Definition 3.3.5: Field of a Relation**

The **field** of a **relation**  $R$  on a set  $A$  is the set:

$$\text{field}(R) = \text{dom}(R) \cup \text{ran}(R)$$

Where  $\text{dom}(R)$  is the **domain** of  $R$  and  $\text{ran}(R)$  is the **range** of  $R$ .

These provide the two most basic examples of relations on a set. The empty set is the relation that says no two elements are related. Indeed, even single elements are unrelated to themselves. The second, the entire Cartesian product  $A \times A$ , says that everything is related. These are the two extreme cases, but provide useful examples and counterexamples in various contexts. More useful is that the union and intersection of relations is also a relation. We prove this now.

**Theorem 3.3.5.** *If  $A$  is a set and if  $R_1$  and  $R_2$  are relations on  $A$ , then  $R_1 \cap R_2$  is a relation on  $A$ .*

*Proof.* For let  $R = R_1 \cap R_2$  and suppose  $R$  is not a relation on  $A$ . Then there is an  $x \in R$  such that  $x \notin A \times A$ . But if  $x \in R$  then  $x \in R_1$  and  $x \in R_2$ . But for all  $x \in R_1$ ,  $x \in A \times A$ , since  $R_1$  is a relation on  $A$ , a contradiction as  $x \notin A \times A$ . Therefore,  $R$  is a relation on  $A$ .  $\square$

**Theorem 3.3.6.** *If  $A$  is a set and if  $R_1$  and  $R_2$  are relations on  $A$ , then  $R_1 \cup R_2$  is a relation on  $A$ .*

*Proof.* For let  $R = R_1 \cup R_2$  and suppose  $R$  is not a relation on  $A$ . Then there is an  $x \in R$  such that  $x \notin A \times A$ . But if  $x \in R$  then  $x \in R_1$  or  $x \in R_2$ . But for all  $x \in R_1$  and for all  $x \in R_2$ ,  $x \in A \times A$ , since  $R_1$  and  $R_2$  are relations on  $A$ , a contradiction. Therefore, etc.  $\square$

**Theorem 3.3.7.** *If  $A$  is a set and  $R$  is a relation on  $A$ , then there is a relation  $\mathcal{U}$  on  $A$  such that  $R \cap \mathcal{U} = R$ .*

*Proof.* For let  $\mathcal{U} = A \times A$ . Then by Thm. 3.3.2,  $A \times A$  is a relation on  $A$ . But since  $R$  is a relation,  $R \subseteq A \times A$ . But then  $R \cap \mathcal{U} = R$ . Therefore, etc.  $\square$

**Theorem 3.3.8.** *If  $A$  is a set and  $R$  is a relation on  $A$ , then there is a relation  $\mathcal{U}$  on  $A$  such that  $R \cup \mathcal{U} = R$*

*Proof.* For let  $\mathcal{U} = \emptyset$ . Then by Thm. 3.3.3,  $\mathcal{U}$  is a relation. But if  $R$  is a set, then  $R \cup \emptyset = R$ . Thus,  $R \cup \mathcal{U} = R$ . Therefore, etc.  $\square$

Since a general relation is simply a subset of  $A \times A$ , there's not much structure on them, and thus there's not a lot that can be said about them. We can add

more constraints to certain relations to get the more familiar properties we're used to.

### Definition 3.3.6: Reflexive Relations

A reflexive relation on a set  $A$  is a relation  $R$  on  $A$  such that for all  $a \in A$  it is true that  $aRa$ .

A reflexive relation on  $A$  is simply any subset of  $A \times A$  that contains the entire *diagonal*. That, all of the pairs  $(a, a)$ . A reflexive relation can contain more than this, however. The only strict requirement is that  $aRa$  for all  $a \in A$ .

**Theorem 3.3.9.** *If  $A$  is a set, and if  $R_1$  and  $R_2$  are reflexive relations on  $A$ , then  $R_1 \cap R_2$  is a reflexive relation on  $A$ .*

*Proof.* For let  $R = R_1 \cap R_2$ . Then by Thm. 3.3.5,  $R$  is a relation. Suppose  $R$  is not reflexive. Then there is an  $a \in A$  such that  $(a, a) \notin R$ . But if  $a \in A$ , then  $(a, a) \in R_1$ , since  $R_1$  is reflexive. Similarly,  $(a, a) \in R_2$  since  $R_2$  is reflexive. But if  $(a, a) \in R_1$  and  $(a, a) \in R_2$ , then  $(a, a) \in R$  since  $R = R_1 \cap R_2$ , a contradiction. Therefore,  $R$  is reflexive.  $\square$

**Theorem 3.3.10.** *If  $A$  is a set, if  $R_1$  is a reflexive relation on  $A$ , and if  $R_2$  is a relation on  $A$ , then  $R_1 \cup R_2$  is a reflexive relation on  $A$ .*

*Proof.* For let  $R = R_1 \cup R_2$ . Since  $R_1$  and  $R_2$  are relations, by Thm. 3.3.6,  $R$  is a relation. Suppose it is not reflexive. Then there is an  $a \in A$  such that  $(a, a) \notin R$ . But if  $a \in A$  then  $(a, a) \in R_1$  since  $R_1$  is reflexive. But if  $(a, a) \in R_1$  then  $(a, a) \in R_1 \cup R_2$ , a contradiction. Therefore, etc.  $\square$

Given an arbitrary relation  $R$  on a set  $A$ , it may not be true that  $R$  is reflexive. It may often be useful to add in only the necessary points of  $A$  that will make  $R$  reflexive. This is called the reflexive closure of  $R$ .

### Definition 3.3.7: Reflexive Closure of a Relation

The reflexive closure of a relation  $R$  on a set  $A$  is the set:

$$S = R \cup \{(a, a) : a \in A\} \quad (3.3.4)$$

**Theorem 3.3.11.** *If  $A$  is a set,  $R$  is a relation on  $A$ , and if  $S$  is the reflexive closure of  $R$ , then  $S$  is a reflexive relation on  $A$ .*

**Theorem 3.3.12.** *If  $A$  is a set, if  $R$  is a relation on  $A$ , if  $S$  is the reflexive closure of  $R$ , and if  $T$  is a reflexive relation on  $A$  such that  $R \subseteq T$ , then  $S \subseteq T$ .*

*Proof.* For if  $x \in S$ , then either  $x \in R$  or there is an  $a \in A$  such that  $x = (a, a)$ . But if  $x \in R$ , then  $x \in T$  since  $R \subseteq T$ . If  $x \notin R$  then there is an  $a \in A$  such that  $x = (a, a)$ . But  $T$  is reflexive, and therefore  $(a, a) \in T$ . But then  $x \in T$ . Therefore,  $S \subseteq T$ .  $\square$

Thm. 3.3.12 says that the reflexive closure of a relation  $R$  is, in a sense, the *smallest* relation that is reflexive and contains  $R$  as a subset.

**Theorem 3.3.13.** *If  $A$  is a set,  $R_1$  and  $R_2$  are relations on  $A$ , and if  $S_1$  and  $S_2$  are the reflexive closures of  $R_1$  and  $R_2$ , respectively, then the reflexive closure of  $R_1 \cap R_2$  is:*

$$S = S_1 \cap S_2 \quad (3.3.5)$$

*Proof.* By the definition of reflexive closure, we have:

$$S_1 = R_1 \cup \{(a, a) : a \in A\} \quad (\text{Def. 3.3.7})$$

$$S_2 = R_2 \cup \{(a, a) : a \in A\} \quad (\text{Def. 3.3.7})$$

$$\begin{aligned} S_1 \cap S_2 &= (R_1 \cup \{(a, a) : a \in A\}) \cap (R_2 \cup \{(a, a) : a \in A\}) \\ &= (R_1 \cap R_2) \cup \{(a, a) : a \in A\} \quad (\text{Distributive Law}) \end{aligned}$$

But by the definition of the transitive closure of  $R_1 \cap R_2$ :

$$S = (R_1 \cap R_2) \cup \{(a, a) : a \in A\} \quad (\text{Def. 3.3.7})$$

Therefore, etc.  $\square$

### Definition 3.3.8: Symmetric Relation

A symmetric relation on a set  $A$  is a relation  $R$  on  $A$  such that for all  $a, b \in A$  such that  $aRb$ , it is true that  $bRa$ .

**Theorem 3.3.14.** *If  $A$  is a set, if  $S_1$  and  $S_2$  are symmetric relations on  $A$ , then  $S_1 \cap S_2$  is a symmetric relation on  $A$ .*

*Proof.* For since  $S_1$  and  $S_2$  are relations,  $S_1 \cap S_2$  is a relation (Thm. 3.3.5). Suppose it is not symmetric. Then there is an  $(x, y) \in S_1 \cap S_2$  such that  $(y, x) \notin S_1 \cap S_2$ . But if  $(x, y) \in S_1 \cap S_2$ , then  $(x, y) \in S_1$  and  $(x, y) \in S_2$  (Def. 3.1.8). But  $S_1$  is symmetric and if  $(x, y) \in S_1$ , then  $(y, x) \in S_1$  (Def. 3.3.8). Similarly  $(y, x) \in S_2$ , and therefore  $(y, x) \in S_1 \cap S_2$  (Def. 3.1.8)), a contradiction. Therefore,  $S_1 \cap S_2$  is symmetric.  $\square$

### Definition 3.3.9: Transitive Relation

A transitive relation on a set  $A$  is a relation  $R$  on  $A$  such that for all  $a, b, c \in A$  such that  $aRb$  and  $bRc$ , is it true that  $aRc$ .

**Theorem 3.3.15.** *If  $A$  is a set, then  $A \times A$  is a transitive relation on  $A$ .*

*Proof.* For suppose not. Then there exists  $a, b, c \in A$  such that  $(a, b) \in A \times A$  and  $(b, c) \in A \times A$ , yet  $(a, c) \in A \times A$ . But if  $a \in A$  and  $c \in A$ , then  $(a, c) \in A \times A$  (Def. 3.1.13), a contradiction. Therefore  $A \times A$  is a transitive relation on  $A$ .  $\square$

Using the Cartesian product definition of a relation, we can visualize the requirement imposed on transitive relations in the diagram below (Fig. 3.22).

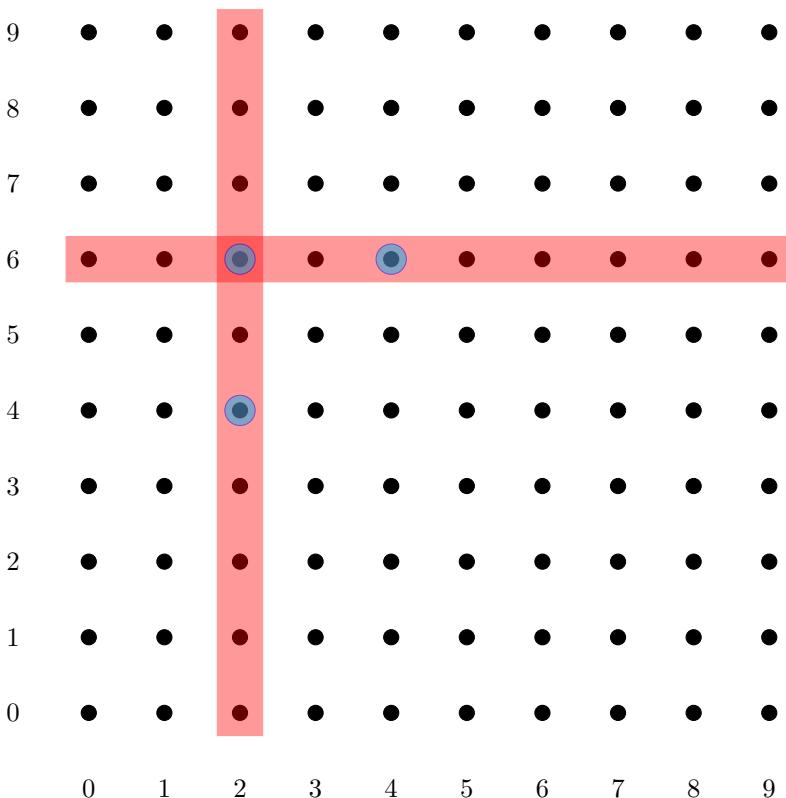


Fig. 3.22: Diagram for a Transitive Relation

Given a point  $(a, b)$  that is in the relation and another point  $(b, c)$ , for the

relation to be transitive requires  $(a, c)$  to be contained in it. That is, if we take the first coordinate from the first element and the second coordinate from the second element and then combine them to form a new ordered pair, this element must also be in the relation.

**Theorem 3.3.16.** *If  $A$  is a set, if  $T_1$  and  $T_2$  are transitive relations on  $A$ , and if  $R = T_1 \cap T_2$ , then  $R$  is a transitive relation.*

*Proof.* For since  $T_1$  and  $T_2$  are relations,  $T_1 \cap T_2$  is a relation (Thm. 3.3.5). Suppose it is not transitive. Then there are  $(x, y), (y, z) \in R$  such that  $(x, z) \notin R$  (Def. 3.3.9). But if  $(x, y), (y, z) \in R$ , then by the definition of intersection,  $(x, y), (y, z) \in T_1$  and  $(x, y), (y, z) \in T_2$  (Def. 3.1.8). But  $T_1$  is transitive, and thus if  $xT_1y$  and  $yT_1z$ , then  $xT_1z$ . But similarly  $T_2$  is transitive, and therefore  $xT_2z$ . But then  $(x, z) \in T_1$  and  $(x, z) \in T_2$ , and thus  $(x, z) \in T_1 \cap T_2$ , a contradiction. Therefore,  $R$  is transitive.  $\square$

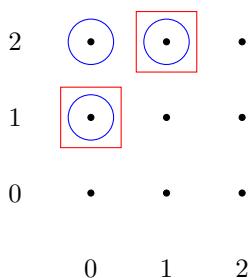
**Example 3.3.2** The requirement that both relations  $T_1$  and  $T_2$  are transitive cannot be weakened. For consider the relations  $S$  and  $T$  on  $\mathbb{Z}_3$  defined by:

$$S = \{(0, 1), (1, 2)\} \quad (3.3.6a) \quad T = \{(0, 1), (1, 2), (0, 1)\} \quad (3.3.6b)$$

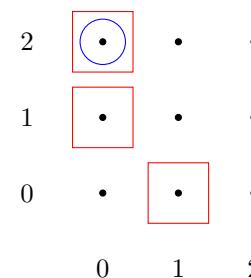
Then  $T$  is transitive and  $S$  is not. Moreover  $S \subseteq T$ , and hence  $S \cap T = S$  (Thm. 3.2.37), and therefore the intersection is not transitive. This example is demonstrated in Fig. 3.23.1. The opposite is possible, and to construct an example we need only find a transitive relation  $T$  and a non-transitive relation  $S$  such that  $T \subseteq S$ . Define:

$$T = \{(0, 0)\} \quad (3.3.7a) \quad S = \{(0, 0), (0, 1), (1, 2)\} \quad (3.3.7b)$$

Then  $T$  is transitive,  $S$  is not, and  $S \cap T = T$  (See 3.23.2).



3.23.1: The intersection is not transitive



3.23.2: The intersection is transitive

Fig. 3.23: The Intersection of Transitive and Non-Transitive Relations

We can strengthen our claim that the intersection of two transitive relations is again transitive and show that any arbitrary intersection will again be transitive.

**Theorem 3.3.17.** *If  $A$  is a set, if  $\mathcal{P}(A \times A)$  denotes the power set of  $A \times A$ , if  $\mathcal{O} \subseteq \mathcal{P}(A \times A)$  is such that for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $\mathcal{U}$  is a transitive relation on  $A$ , if  $\mathcal{T}$  is defined by:*

$$\mathcal{T} = \bigcap_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \quad (3.3.8)$$

*Then  $\mathcal{T}$  is a transitive relation on  $A$ .*

*Proof.* For suppose not. Then there exists  $a, b, c \in A$  such that  $(a, b) \in \mathcal{T}$  and  $(b, c) \in \mathcal{T}$ , yet  $(a, c) \notin \mathcal{T}$ . But if  $(a, b) \in \mathcal{T}$ , then for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $(a, b) \in \mathcal{U}$  (Def. 3.1.10). Similarly, for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $(b, c) \in \mathcal{U}$ . But by hypothesis, for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $\mathcal{U}$  is a transitive relation and thus if  $(a, b) \in \mathcal{U}$  and  $(b, c) \in \mathcal{U}$ , then it is true that  $(a, c) \in \mathcal{U}$  (Def. 3.3.9). But then for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $(a, c) \in \mathcal{U}$ , and therefore  $(a, c) \in \mathcal{T}$  (Def. 3.1.10), a contradiction. Therefore,  $\mathcal{T}$  is a transitive relation on  $A$ .  $\square$

This allows us to define the transitive closure of any relation  $R$  on a set  $A$ . It is, in a sense, the *smallest* transitive relation that contains  $R$ .

**Theorem 3.3.18.** *If  $A$  is a set and if  $R$  is a relation on  $A$ , then there exists a transitive relation  $\mathcal{T}$  on  $A$  such that  $R \subseteq \mathcal{T}$  and such that for transitive relations  $T$  on  $A$  such that  $R \subseteq T$  it is true that  $\mathcal{T} \subseteq T$ .*

*Proof.* For let  $P$  be the proposition *True if  $S$  is a transitive relation on  $A$  such that  $R \subseteq S$ , false otherwise*. Then by the axiom schema of specification (Ax. 3.1.3) there exists a set:

$$\mathcal{O} = \{ S \in \mathcal{P}(A \times A) \mid P(S) \} \quad (3.3.9)$$

But then for all  $S \in \mathcal{O}$ ,  $P(S)$  is true and therefore  $R \subseteq S$  and  $S$  is transitive. Moreover,  $\mathcal{O}$  is non-empty since by Thm. 3.3.15,  $A \times A$  is a transitive relation. Define  $\mathcal{T}$  by:

$$\mathcal{T} = \bigcap_{S \in \mathcal{O}} S \quad (3.3.10)$$

Then by Thm. 3.3.17,  $\mathcal{T}$  is a transitive relation. Moreover, suppose  $S$  is a transitive relation such that  $R \subseteq S$ . But if  $S$  is a relation on  $A$ , then  $S \subseteq A \times A$  (Def. 3.3.1) and therefore  $S \in \mathcal{P}(A \times A)$  (Def. 3.1.12). But if  $S$  is a transitive relation and if  $R \subseteq S$ , then  $P(S)$  is true, and therefore  $S \in \mathcal{P}$ . Thus,  $\mathcal{T} \subseteq S$ .  $\square$

**Definition 3.3.10: Transitive Closure**

The transitive closure of a relation  $R$  on a set  $A$  is the the set  $R^t \subseteq A \times A$  defined by:

$$R^t \quad (3.3.11)$$

**Definition 3.3.11: Asymmetric Relation**

An asymmetric relation on a set  $A$  is a relation  $R$  on  $A$  such that for all  $a, b \in A$  such that  $aRb$  it is true that  $(b, a) \notin R$ .

**Definition 3.3.12: Total Relation**

A total relation on a set  $A$  is a relation  $R$  on  $A$  such that for all  $a, b \in A$  it is true that either  $aRb$  or  $bRa$ , or both.

The notion of equality can be defined as a relation with the following properties:

1. Equality is Reflexive:  $a = a$  for all  $a \in A$ .
2. Equality is Symmetric:  $a = b$  if and only if  $b = a$ .
3. Equality is Transitive: If  $a = b$  and  $b = c$ , then  $a = c$ .
4. The relation is uniquely defined by the set  $\{(a, a) \in A \times A : a \in A\}$ .

That is, equality can be seen as the *diagonal* in the Cartesian product  $A \times A$ .

**Definition 3.3.13: Antisymmetric Relation**

n antisymmetric relation on a set  $A$  is a relation  $R$  on  $A$  such that for all  $a, b \in A$  such that  $aRb$  and  $bRa$ , it is true that  $a = b$ .

**Definition 3.3.14: Equivalence Relation**

An equivalence relation on a set  $A$  is a relation  $R$  on  $A$  such that  $R$  is reflexive, symmetric, and transitive.

Equivalence relations attempt to model equality. They are fundamental in mathematics as they allow us to define *equivalence classes*, which are used to define quotients. There are many examples such as quotient topologies, quotient groups, quotient rings, and quotient modules, all of which will be discussed later.

### Definition 3.3.15: Equivalence Class

The equivalence class of an element  $x$  in a set  $A$  by an equivalence relation  $R$  is the set:

$$[x] = \{ y \in A \mid xRy \}$$

It's important to note that the term class here is different from the notion of a collection of sets. And equivalence class of an element  $x$  in a set  $A$  under an equivalence relation  $R$  will indeed be a set in *ZFC*.

**Theorem 3.3.19.** *If  $A$  is a set, if  $R$  is an equivalence relation on  $A$ , and if  $x, y \in A$ , then either  $[x] = [y]$  or  $[x] \cap [y] = \emptyset$ .*

*Proof.* For suppose not and suppose  $[x] \neq [y]$  and that  $[x] \cap [y] \neq \emptyset$ . That is, suppose:

$$\neg([x] = [y]) \wedge \neg([x] \cap [y] = \emptyset)$$

If  $[x] \cap [y]$  is non-empty then there is a  $z \in A$  such that  $z \in [x]$  and  $z \in [y]$  (Def. 3.1.1). But if  $z \in [x]$ , then  $xRz$  (Def. 3.3.15). But also  $z \in [y]$  and therefore  $yRz$ . But  $R$  is an equivalence relation and is therefore symmetric (Def. 3.3.14) and thus if  $yRz$  then  $zRy$  (Def. 3.3.8). But an equivalence relation is also transitive, and thus if  $xRz$  and  $zRy$ , then  $xRy$  (Def. 3.3.9). But if  $[x] \neq [y]$  then either  $[x] \not\subseteq [y]$  or  $[y] \not\subseteq [x]$ . Suppose  $[x] \not\subseteq [y]$  and let  $a \in [x]$  be such that  $a \notin [y]$ . But if  $a \in [x]$  then  $xRa$  (Def. 3.3.15). But since equivalence relations are symmetric, if  $xRa$ , then  $aRx$ . But it was proven that  $xRy$  and since equivalence relations are transitive, if  $aRx$  and  $xRy$ , then  $aRy$ . But again if  $aRy$ , then  $yRa$  and therefore  $a \in [y]$ , a contradiction. Therefore  $[x] \subseteq [y]$ . Similarly,  $[y] \subseteq [x]$  and therefore  $[x] = [y]$ , a contradiction. By the law of the excluded middle, the negation is true:

$$\neg(\neg([x] = [y]) \wedge \neg([x] \cap [y] = \emptyset)) = ([x] = [y]) \vee ([x] \cap [y] = \emptyset)$$

Thus, either  $[x] = [y]$  or  $[x] \cap [y] = \emptyset$ . □

**Definition 3.3.16: Quotient Set**

The quotient set of a set  $A$  by an equivalence relation  $R$  on  $A$  is the set:

$$A/R = \{ [x] \in \mathcal{P}(A) \mid x \in A \}$$

Where  $[x]$  is the equivalence class of  $x$  under  $R$ .

**Example 3.3.3** The definition of the quotient set comes naturally when one considers functions between sets. Suppose  $A$  and  $B$  are sets, and suppose  $f : A \rightarrow B$  is a function. In general, it may not be true that  $f(a_1) = f(a_2)$  implies that  $a_1 = a_2$ , and so we wish to find a subset of  $A$  with this property. The quotient set does this. Let  $R$  be the relation:

$$R = \{ (a, b) \in A^2 \mid f(a) = f(b) \} \quad (3.3.12)$$

If we form the quotient set  $A/R$  and consider the projective mapping  $\pi : A \rightarrow A/R$  that sends  $a \in A$  to its equivalence class. That is,  $\pi(a) = [a]$ . We then seek a function  $\tilde{f} : A/R \rightarrow B$  such that  $\tilde{f} \circ \pi = f$ . That is, we wish to make the diagram below *commute*.

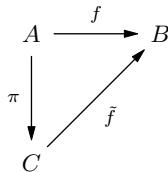


Fig. 3.24: Commutative Diagram for the Quotient Set

So we need to map  $[x]$  to  $f(x)$ . That is,  $\tilde{f}([x]) = f(x)$ . For this problem to be well posed requires that the equivalence class that make up the elements of  $A/R$  come from equivalence relations. That is, that the relation  $R$  is transitive, symmetric, and reflexive.

**Theorem 3.3.20.** *If  $A$  is a set and if  $R$  is an equivalence relation on  $A$ , then  $A/R$  is a partition of  $A$ .*

*Proof.* For by Thm. 3.3.19, if  $\mathcal{U}, \mathcal{V} \in A/R$ , then either  $\mathcal{U} = \mathcal{V}$  or  $\mathcal{U} \cap \mathcal{V} = \emptyset$ . But also, for all  $x \in A$ , there is a  $\mathcal{U} \in A/R$  such that  $x \in \mathcal{U}$  since  $x \in [x]$  and  $[x] \in R/A$ . Therefore,  $A/R$  is a partition of  $A$ .  $\square$

**Theorem 3.3.21.** *If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, and if  $R$  is the relation on  $A$  defined by:*

$$R = \{ (a_0, a_1) \in A \times A \mid f(a_0) = f(a_1) \} \quad (3.3.13)$$

Then  $R$  is an equivalence relation on  $A$ .

**Theorem 3.3.22.** If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, if  $R$  is an equivalence relation on  $A$ , if  $A/R$  is the quotient set, and if  $\tilde{f} \subseteq (A/R) \times B$  defined by:

$$\tilde{f} = \{ (\tilde{a}, b) \in (A/R) \times B \mid \exists_{a \in A} \exists_{b \in B} (\tilde{a} = [a]) \text{ and } (f(a) = b) \} \quad (3.3.14)$$

is such that for all  $a \in A$  and  $b \in B$  such that  $f(a) = b$ ,  $f^{-1}[\{b\}]$  is the union of elements of the equivalence class  $[a]$ , then  $\tilde{f} : A/R \rightarrow B$  is a function.

## CHAPTER 4

---

### The Real Numbers

---



# CHAPTER 5

---

## Function Theory

---

Functions serve as a basic tool for studying mathematics, so much so that it is often taken as fundamental and no definition is given. We've adopted the definition that a function from a set  $A$  to a set  $B$ ,  $f : A \rightarrow B$ , is a subset of the Cartesian product  $A \times B$  with a few properties (see Def. 3.1.14). We now take the time to examine the implications of this definition.

### 5.1 Basic Definitions and Theorems

We start by presenting some basic, but crucial, facts about the image and pre-images of subsets under a function.

#### Definition 5.1.1: Identity Function

The identity function on a set  $A$  is the subset  $\text{id}_A \subseteq A \times A$  defined by:

$$\text{id}_A = \{ \mathbf{x} \in A \times A \mid \exists_{x \in A} (\mathbf{x} = (x, x)) \}$$

It would be poor to prove by definition, and thus we show that the identity function is indeed a function.

**Theorem 5.1.1.** *If  $A$  is a set, and if  $\text{id}_A$  is the identity function on  $A$ , this  $\text{id}_A$  is a function from  $A$  to  $A$ .*

*Proof.* For suppose not. But if  $x \in A$ , then  $(x, x) \in A \times A$  (Def. 3.1.13), and hence  $(x, x) \in \text{id}_A$  (Def. 5.1.1). Thus, for all  $x \in A$  there is a  $y \in A$  such that  $(x, y) \in \text{id}_A$ . Thus, if  $\text{id}_A$  is not a function, then there is an  $x \in A$  such that there are at least two distinct  $y_1, y_2 \in A$  such that  $(x, y_1) \in \text{id}_A$  and  $(x, y_2) \in \text{id}_A$  (Def. 3.1.14). But by the definition of the identity function,

$(x, y_1) \in \text{id}_A$  if and only if there is a  $z \in A$  such that  $(x, y_1) = (z, z)$  (Def. 5.1.1). But then  $x = z$  and  $y_1 = z$ , and similarly  $y_2 = z$ , and thus by the transitivity of equality (Thm. 3.2.11),  $y_1 = y_2$ , a contradiction since  $y_1$  and  $y_2$  are distinct. Thus,  $\text{id}_A$  is a function.  $\square$

We can write the identity in a nicer way by defining the image of an element  $x \in A$ .

$$\text{id}_A(x) = x \quad (5.1.1)$$

**Theorem 5.1.2.** *If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, then  $f(\emptyset) = \emptyset$ .*

*Proof.* For suppose not and suppose  $y \in f(\emptyset)$ . But then by the definition of the image of a subset (Def. 3.1.16) there is an  $x \in \emptyset$  such that  $f(x) = y$ , a contradiction since for all  $x$  it is true that  $x \notin \emptyset$  (Def. 3.2.1). Therefore,  $f(\emptyset) = \emptyset$ .  $\square$

**Theorem 5.1.3.** *If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, then  $f^{-1}(\emptyset) = \emptyset$ .*

*Proof.* Suppose not and suppose  $x \in f^{-1}(\emptyset)$ . Then by the definition of the pre-image (Def. 3.1.17)  $f(x) \in \emptyset$ , a contradiction since for all  $y$  it is true that  $y \notin \emptyset$  (Def. 3.2.1). Therefore,  $f^{-1}(\emptyset) = \emptyset$ .  $\square$

**Theorem 5.1.4.** *If  $A$  and  $B$  are sets,  $\mathcal{U} \subseteq B$ , and  $f : A \rightarrow B$  is a function, then:*

$$f(f^{-1}(\mathcal{U})) \subseteq \mathcal{U}$$

*Proof.* For if  $y \in f(f^{-1}(\mathcal{U}))$ , then there is an  $x \in f^{-1}(\mathcal{U})$  such that  $y = f(x)$  (Def. 3.1.16). But if  $x \in f^{-1}(\mathcal{U})$ , then  $f(x) \in \mathcal{U}$  (Def. 3.1.17). Thus,  $y \in \mathcal{U}$ .  $\square$

We may not be able to attain equality.

**Example 5.1.1** If  $A$  and  $B$  are non-empty sets and if there exists  $y_1, y_2 \in B$  such that  $y_1 \neq y_2$ , then there is a function  $f : A \rightarrow B$  and a  $\mathcal{U} \subseteq B$  such that:

$$f(f^{-1}(\mathcal{U})) \neq \mathcal{U} \quad (5.1.2)$$

For if  $A$  and  $B$  are non-empty, let  $f : A \rightarrow B$  be defined by:

$$f = \{(x, y_1) \in A \times B \mid x \in A\} \quad (5.1.3)$$

That is,  $f(x) = y_1$  for all  $x \in A$ . Then  $f$  is a function since  $f \subseteq A \times B$  and for all  $x \in A$  there is a unique  $y \in B$  such that  $(x, y) \in f$  (Def. 3.1.14). However since for all  $x \in A$  it is true that  $f(x) = y_1$ , we have:

$$f^{-1}(\{y_2\}) = \emptyset \quad (5.1.4)$$

For suppose not. If  $x \in f^{-1}(\{y_2\})$ , then  $f(x) = y_2$  (Def. 3.1.17), but for all  $x \in A$  it is true that  $f(x) = y_1$ , and  $y_1 \neq y_2$ , a contradiction. Thus  $f^{-1}(\{y_2\}) = \emptyset$  (Def. 3.2.1). But by Thm. 5.1.2, the image of the empty set is the empty set and thus:

$$f(f^{-1}(\{y_2\})) = \emptyset \quad (5.1.5)$$

But  $\{y_2\} \neq \emptyset$  and  $\{y_2\} \subseteq B$ . This shows the converse of Thm. 5.1.4 may fail.

**Theorem 5.1.5.** *If  $A$  and  $B$  are sets,  $\mathcal{U} \subseteq A$ , and  $f : A \rightarrow B$  is a function, then:*

$$\mathcal{U} \subseteq f^{-1}(f(\mathcal{U}))$$

*Proof.* For if  $x \in \mathcal{U}$ , then  $f(x) \in f(\mathcal{U})$  (Def. 3.1.16). But if  $f(x) \in \mathcal{U}$ , then  $x \in f^{-1}(f(\mathcal{U}))$  (Def. 3.1.17), completing the proof.  $\square$

Again, the converse need not be true.

**Example 5.1.2** Let  $A = B = \mathbb{N}$ , and define  $f(n) = 0$ . Then  $f^{-1}(\{0\}) = \mathbb{N}$ , but  $\{0\} \neq \mathbb{N}$ . Thus, the converse of Thm. 5.1.5 may fail.

**Theorem 5.1.6.** *If  $A$  and  $B$  are sets, if  $\mathcal{U} \subseteq A$  is a non-empty subset of  $A$ , and if  $f : A \rightarrow B$  is a function, then  $f(\mathcal{U})$  is non-empty.*

*Proof.* For suppose not. If  $f(\mathcal{U})$  is empty, then  $f^{-1}(f(\mathcal{U}))$  is empty (Thm. 5.1.3). But by Thm. 5.1.5,  $\mathcal{U} \subseteq f^{-1}(f(\mathcal{U}))$ , and thus  $\mathcal{U} \subseteq \emptyset$ . But  $\emptyset \subseteq \mathcal{U}$  (Thm. 3.2.1), and thus  $\mathcal{U} = \emptyset$  (Def. 3.1.3), a contradiction as  $\mathcal{U}$  is non-empty.  $\square$

A nice rewording of this theorem comes from the contrapositive.

**Theorem 5.1.7.** *If  $A$  and  $B$  are sets, if  $\mathcal{U} \subseteq A$  is a non-empty set, if  $f : A \rightarrow B$  is a function, and if  $f(\mathcal{U}) = \emptyset$ , then  $\mathcal{U} = \emptyset$ .*

*Proof.* Suppose not. Then  $f(\mathcal{U})$  is non-empty (Thm. 5.1.6), a contradiction.  $\square$

We cannot reverse this theorem for the pre-image. That is, the inverse image of a non-empty set may indeed be empty.

**Example 5.1.3** Consider the case when  $A = B = \mathbb{R}$  and define the function  $f : \mathbb{R} \rightarrow \mathbb{R}$  by  $f(x) = 0$  for all  $x \in \mathbb{R}$ . Formally, this is subset of  $\mathbb{R}^2$  defined by:

$$f = \{ (x, 0) \in \mathbb{R}^2 \mid x \in \mathbb{R} \} \quad (5.1.6)$$

The pre-image of any subset  $\mathcal{U} \subseteq \mathbb{R}$  that does not contain 0 will be empty, even though the set itself need not be empty. A surjective function is a function with the property that the pre-image of a non-empty set is non-empty (see Def. 5.1.2).

**Theorem 5.1.8.** If  $A$  and  $B$  are sets, if  $\mathcal{U}, \mathcal{V} \subseteq A$ , if  $\mathcal{U} \subseteq \mathcal{V}$ , and if  $f : A \rightarrow B$  is a function, then:

$$f(\mathcal{U}) \subseteq f(\mathcal{V})$$

*Proof.* For suppose not. Then by the definition of subset, there is a  $y \in f(\mathcal{U})$  such that  $y \notin f(\mathcal{V})$  (Def. 3.1.2). But if  $y \in f(\mathcal{U})$ , then there is an  $x \in \mathcal{U}$  such that  $f(x) = y$  (Def. 3.1.16). But  $\mathcal{U} \subseteq \mathcal{V}$ , and therefore  $x \in \mathcal{V}$  (Def. 3.1.2). But if  $x \in \mathcal{V}$ , then  $f(x) \in f(\mathcal{V})$  (Def. 3.1.16). Thus,  $y \in f(\mathcal{V})$ , a contradiction.  $\square$

**Theorem 5.1.9.** If  $A$  and  $B$  are sets, if  $\mathcal{U}, \mathcal{V} \subseteq B$ , if  $\mathcal{U} \subseteq \mathcal{V}$ , and if  $f : A \rightarrow B$  is a function, then:

$$f^{-1}(\mathcal{U}) \subseteq f^{-1}(\mathcal{V})$$

*Proof.* For suppose not. Then by the definition of subsets there is an  $x \in f^{-1}(\mathcal{U})$  such that  $x \notin f^{-1}(\mathcal{V})$  (Def. 3.1.2). But by the definition of pre-image, if  $x \in f^{-1}(\mathcal{U})$ , then there is a  $y \in \mathcal{U}$  such that  $f(x) = y$  (Def. 3.1.17). But  $\mathcal{U} \subseteq \mathcal{V}$ , and thus  $y \in \mathcal{V}$  (Def. 3.1.2). But then  $x \in f^{-1}(\mathcal{V})$ , a contradiction.  $\square$

### Theorem 5.1.10: Preservation of Intersection by Pre-Images

If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, if  $\mathcal{P}(B)$  is the power set of  $B$ , and if  $\mathcal{O} \subseteq \mathcal{P}(B)$ , then:

$$f^{-1}\left(\bigcap_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) = \bigcap_{\mathcal{U} \in \mathcal{O}} f^{-1}(\mathcal{U})$$

*Proof.* For if  $x \in f^{-1}(\bigcap \mathcal{U})$ , then  $f(x) \in \bigcap \mathcal{U}$  (Def. 3.1.17). But then for all  $\mathcal{U} \in \mathcal{O}$ ,  $f(x) \in \mathcal{U}$  (Def. 3.1.10). But then for all  $\mathcal{U} \in \mathcal{O}$ ,  $x \in f^{-1}(\mathcal{U})$  (Def. 3.1.17) and therefore  $x \in \bigcap f^{-1}(\mathcal{U})$  (Def. 3.1.10). Thus, by the definition of subsets:

$$f^{-1}\left(\bigcap_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) \subseteq \bigcap_{\mathcal{U} \in \mathcal{O}} f^{-1}(\mathcal{U}) \quad (5.1.7)$$

Moreover, if  $x \in \bigcap f^{-1}(\mathcal{U})$  then for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $x \in f^{-1}(\mathcal{U})$  (Def. 3.1.10). But then for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $f(x) \in \mathcal{U}$  (Def. 3.1.17) and therefore  $f(x) \in \bigcap \mathcal{U}$ . But then  $x \in f^{-1}(\bigcap \mathcal{U})$  (Def. 3.1.17) and therefore:

$$\bigcap_{\mathcal{U} \in \mathcal{O}} f^{-1}(\mathcal{U}) \subseteq f^{-1}\left(\bigcap_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) \quad (5.1.8)$$

From the definition of equality (Def. 3.1.3) we can conclude the proof.  $\square$

While the proof of this theorem is very straight forward, we have highlighted it to emphasize that it is a very important and useful theorem that will be used

frequently once we delve into measure theory and topology. This theorem has a counterpart for unions that is equally important. The forward image of a function lacks this preservation property, and we can only show that one side is a subset (perhaps proper) of the other.

**Theorem 5.1.11.** *If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, if  $\mathcal{P}(A)$  is the power set of  $A$ , and if  $\mathcal{O} \subseteq \mathcal{P}(A)$ , then:*

$$f\left(\bigcap_{U \in \mathcal{O}} U\right) \subseteq \bigcap_{U \in \mathcal{O}} f(U) \quad (5.1.9)$$

*Proof.* For if not, then by the definition of subset there is a  $y \in f(\bigcap \mathcal{U})$  such that  $y \notin \bigcap f(\mathcal{U})$  (Def. 3.1.2). But then by the definition of image, there is an  $x \in \bigcap \mathcal{U}$  such that  $f(x) = y$  (Def. 3.1.16). But by the definition of intersection, for all  $U \in \mathcal{O}$  it is true that  $x \in U$  (Def. 3.1.10), and thus for all  $U \in \mathcal{O}$ ,  $f(x) \in f(U)$  (Def. 3.1.16). Thus,  $y \in \bigcap f(\mathcal{U})$  (Def. 3.1.10), a contradiction.  $\square$

**Example 5.1.4** Equality need not be attained for a general function. Indeed, let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be defined by  $f(n) = 0$ . Note the  $\mathbb{N}_e$  and  $\mathbb{N}_o$ , the set of even and odd integers, respectively, are disjoint and thus  $\mathbb{N}_e \cap \mathbb{N}_o = \emptyset$ . Applying Thm. 5.1.2, we have:

$$f(\mathbb{N}_e \cap \mathbb{N}_o) = \emptyset \quad (5.1.10)$$

However,  $f(\mathbb{N}_e) = \{0\}$ , and similarly for  $\mathbb{N}_o$ . Thus:

$$f(\mathbb{N}_e) \cap f(\mathbb{N}_o) = \{0\} \quad (5.1.11)$$

and thus equality is not attained. Functions that attain equality are *injective*.

Returning to pre-images, we have preservation once again.

### Theorem 5.1.12: Preservation of Union by Pre-Images

If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, if  $\mathcal{P}(B)$  is the power set of  $B$ , and if  $\mathcal{O} \subseteq \mathcal{P}(B)$ , then:

$$f^{-1}\left(\bigcup_{U \in \mathcal{O}} U\right) = \bigcup_{U \in \mathcal{O}} f^{-1}(U)$$

*Proof.* For if  $x \in f^{-1}(\bigcup \mathcal{U})$ , then by the definition of pre-image,  $f(x) \in \bigcup \mathcal{U}$  (Def. 3.1.17). But then by the definition of union, there is a  $U \in \mathcal{O}$  such that  $f(x) \in U$ . But then again by the definition of pre-image,  $x \in f^{-1}(U)$ . And

since  $f^{-1}(\mathcal{U}) \subseteq \bigcup f^{-1}(\mathcal{U})$  we have:

$$f^{-1}\left(\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) \subseteq \bigcup_{\mathcal{U} \in \mathcal{O}} f^{-1}(\mathcal{U}) \quad (5.1.12)$$

Similarly, if  $x \in \bigcup f^{-1}(\mathcal{U})$ , then there is a  $\mathcal{U} \in \mathcal{O}$  such that  $x \in f^{-1}(\mathcal{U})$  (Def. 3.1.6). But then  $f(x) \in \mathcal{U}$ , and therefore  $f(x) \in \bigcup \mathcal{U}$ . But then  $x \in f^{-1}(\bigcup \mathcal{U})$  (Def. 3.1.17), and therefore:

$$\bigcup_{\mathcal{U} \in \mathcal{O}} f^{-1}(\mathcal{U}) \subseteq f^{-1}\left(\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) \quad (5.1.13)$$

By the definition of equality (Def. 3.1.3), we are done.  $\square$

While the forward image does not preserve intersections, it does preserve unions. Since the pre-image preserves both intersection and union, it is often used as the basic ingredient of many theories of mathematics. For example, in topology one defines continuous functions by looking at the pre-image of open sets. In measure theory one defines measurable functions by looking at the pre-image of measurable sets. In doing so we can very easily prove many useful theorems about such functions by resorting to Thm. 5.1.10 and Thm. 5.1.12. We now show that the forward image is at least half useful.

### Theorem 5.1.13: Preservation of Union by Images

If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, if  $\mathcal{P}(A)$  is the power set of  $A$ , and if  $\mathcal{O} \subseteq \mathcal{P}(A)$ , then:

$$f\left(\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) = \bigcup_{\mathcal{U} \in \mathcal{O}} f(\mathcal{U})$$

*Proof.* For if  $y \in f(\bigcup \mathcal{U})$ , then there is an  $x \in \bigcup \mathcal{U}$  such that  $y = f(x)$  (Def. 3.1.16). But by the definition of union there is a  $\mathcal{U} \in \mathcal{O}$  such that  $x \in \mathcal{U}$  (Def. 3.1.6). But then  $f(x) \in f(\mathcal{U})$ , and therefore  $y \in f(\mathcal{U})$ . But  $f(\mathcal{U}) \subseteq \bigcup f(\mathcal{U})$ , and thus:

$$f\left(\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) \subseteq \bigcup_{\mathcal{U} \in \mathcal{O}} f(\mathcal{U}) \quad (5.1.14)$$

Similarly, if  $y \in \bigcup f(\mathcal{U})$ , then there is a  $\mathcal{U} \in \mathcal{O}$  such that  $y \in f(\mathcal{U})$  (Def. 3.1.6). But then there is an  $x \in \mathcal{U}$  such that  $y = f(x)$  (Def. 3.1.16). But if  $x \in \mathcal{U}$ ,

then  $x \in \bigcup \mathcal{U}$ , and thus  $f(x) \in \bigcup \mathcal{U}$ . That is:

$$\bigcup_{\mathcal{U} \in \mathcal{O}} f(\mathcal{U}) \subseteq f\left(\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) \quad (5.1.15)$$

From the definition of equality (Def. 3.1.3) we obtain the result.  $\square$

These results are by no means *deep* theorems, but their applications are spread across a myriad of subjects. So much so that many authors forget to thank them! Equally useful theorems relate the pre-image of set differences. Again, the image fails to preserve this.

**Theorem 5.1.14.** *If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, and if  $\mathcal{U}, \mathcal{V} \subseteq A$ , then:*

$$f(\mathcal{U}) \setminus f(\mathcal{V}) \subseteq f(\mathcal{U} \setminus \mathcal{V})$$

*Proof.* For suppose not. Then by the definition of subset, there is a  $y \in f(\mathcal{U} \setminus \mathcal{V})$  such that  $y \notin f(\mathcal{U}) \setminus f(\mathcal{V})$  (Def. 3.1.2). But by the definition of image, there is an  $x \in \mathcal{U} \setminus \mathcal{V}$  such that  $f(x) = y$  (Def. 3.1.16). But if  $x \in \mathcal{U} \setminus \mathcal{V}$ , then it is true that  $x \in \mathcal{U}$  and  $x \notin \mathcal{V}$  (Def. 3.2.2). But if  $x \in \mathcal{U}$ , then  $f(x) \in f(\mathcal{U})$  (Def. 3.1.16). Similarly, if  $x \notin \mathcal{V}$  then  $f(x) \notin f(\mathcal{V})$ . But then  $f(x) \in f(\mathcal{U})$  and  $f(x) \notin f(\mathcal{V})$ , and therefore  $f(x) \in f(\mathcal{U}) \setminus f(\mathcal{V})$  (Def. 3.2.2), a contradiction.  $\square$

**Example 5.1.5** The converse of Thm. 5.1.14 may not be true in general. For let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be defined by  $f(x) = 0$ . Let  $\mathcal{U} = \mathbb{N}_e$  and  $\mathcal{V} = \mathbb{N}_o$  be the even and odd non-negative integers, respectively. Then  $f(\mathcal{U}) = f(\mathcal{V}) = \{0\}$ , and therefore  $f(\mathcal{U}) \setminus f(\mathcal{V}) = \emptyset$ . But  $\mathcal{U} \setminus \mathcal{V} = \mathcal{U}$  since the even and odd numbers are disjoint, and thus  $f(\mathcal{U} \setminus \mathcal{V}) = f(\mathcal{U}) = \{0\}$ , and so equality is not attained.

We now prove that pre-images preserve set differences.

### Theorem 5.1.15: Preservation of Set Difference by Pre-Image

If  $A$  and  $B$  are sets, if  $f : A \rightarrow B$  is a function, and if  $\mathcal{U}, \mathcal{V} \subseteq B$ , then:

$$f^{-1}(\mathcal{U} \setminus \mathcal{V}) = f^{-1}(\mathcal{U}) \setminus f^{-1}(\mathcal{V})$$

*Proof.* For if  $x \in f^{-1}(\mathcal{U} \setminus \mathcal{V})$ , then  $f(x) \in \mathcal{U} \setminus \mathcal{V}$  (Def. 3.1.17). But then  $f(x) \in \mathcal{U}$  and  $f(x) \notin \mathcal{V}$  (Def. 3.2.2). But then  $x \in f^{-1}(\mathcal{U})$  and  $x \notin f^{-1}(\mathcal{V})$  (Def. 3.1.17), and therefore  $x \in f^{-1}(\mathcal{U}) \setminus f^{-1}(\mathcal{V})$  (Def. 3.2.2). Thus:

$$f^{-1}(\mathcal{U} \setminus \mathcal{V}) \subseteq f^{-1}(\mathcal{U}) \setminus f^{-1}(\mathcal{V}) \quad (5.1.16)$$

Similarly, if  $x \in f^{-1}(\mathcal{U}) \setminus f^{-1}(\mathcal{V})$ , then  $x \in f^{-1}(\mathcal{U})$  and  $x \notin f^{-1}(\mathcal{V})$  (Def. 3.2.2). But then  $f(x) \in \mathcal{U}$  and  $f(x) \notin \mathcal{V}$  (Def. 3.1.17). Thus,  $f(x) \in \mathcal{U} \setminus \mathcal{V}$  (Def. 3.2.2).

But then  $x \in f^{-1}(\mathcal{U} \setminus \mathcal{V})$  (Def. 3.2.2). Therefore:

$$f^{-1}(\mathcal{U}) \setminus f^{-1}(\mathcal{V}) \subseteq f^{-1}(\mathcal{U} \setminus \mathcal{V}) \quad (5.1.17)$$

Thus, by the definition of equality (Def. 3.1.3), we are done.  $\square$

The theorems presented can be summarized by saying that pre-images preserve the notions of inclusion, intersections, unions, and set differences, whereas images only preserve unions and inclusion. When we add more structure to function, say by adding *surjectivity* or *injectivity*, then we can show that forward images have more preservation properties.

### 5.1.1 Surjections

#### Definition 5.1.2: Surjective Functions

A **surjective function** from a **set**  $A$  to a set  $B$  is a **function**  $f : A \rightarrow B$  such that  $f(A) = B$ . That is, for all  $y \in B$  there is an  $x \in A$  such that  $f(x) = y$ .

That is, every point  $y \in B$  gets mapped to by at least one point in  $A$ . Surjective functions are also called *onto*. It may also be true that many points in  $A$  map to the same point in  $B$ . Excluding this possibility gives rise to the definition of an *injective* function.

**Example 5.1.6** Given a set  $A$  we can define the *identity function* on  $A$  as the function  $\text{id}_A : A \rightarrow A$  defined by  $\text{id}_A(x) = x$  for all  $x \in A$ . This function will be surjective since for all  $x \in A$ ,  $x$  is the image of itself under  $\text{id}_A$ . Thus every point in  $A$  is mapped to by some point in  $A$ . We can think of less trivial examples if we ponder functions in  $\mathbb{R}$ . Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $f(x) = x^3$ . That is, the cubing function. Given any number  $y \in \mathbb{R}$ , we choose  $x = \sqrt[3]{y}$  (the cubed root of  $y$ ), and from this we obtain  $f(x) = y$ . One might recall that the square root of a negative number is not defined on  $\mathbb{R}$ , and thus if we define  $g : \mathbb{R} \rightarrow \mathbb{R}$  by  $g(x) = x^2$ , then  $g$  is not surjective since  $-1$  is not mapped to by any point.

**Example 5.1.7** The notion of a surjective function depends on both the domain and range of a function. For example, consider the arctan function from trigonometry  $\tan^{-1}$ . If we define this as a function from  $\mathbb{R}$  into the interval  $(-\frac{\pi}{2}, \frac{\pi}{2})$ , then it is indeed a surjection. However, if instead we write  $\tan^{-1} : \mathbb{R} \rightarrow \mathbb{R}$ , then it is no longer a surjection since  $4 \in \mathbb{R}$ , but there is no real number  $x$  such that  $\tan^{-1}(x) = 4$ .

**Theorem 5.1.16.** *If  $A, B, C$  are sets, and if  $f : A \rightarrow B$  and  $g : B \rightarrow C$  are surjective functions, then  $g \circ f : A \rightarrow C$  is a surjective function.*

*Proof.* For suppose not. Then there is a  $z \in C$  such that for all  $x \in X$  it is true that  $f(x) \neq z$ . But since  $g$  is a surjection, there is a  $y \in B$  such that  $g(y) = z$  (Def. 5.1.2). But since  $f$  is a surjection, if  $y \in B$ , then there is an  $x \in A$  such that  $f(x) = y$ . But then  $(g \circ f)(x) = g(f(x)) = g(y) = z$ , a contradiction. Thus,  $g \circ f$  is surjective.  $\square$

The converse of this theorem is partially true.

**Theorem 5.1.17.** *If  $A, B$ , and  $C$  are sets, if  $f : A \rightarrow B$  and  $g : B \rightarrow C$  are functions, and if  $g \circ f : A \rightarrow C$  is surjective, then  $g$  is surjective.*

*Proof.* For suppose not. Then there is a  $z \in C$  such that for all  $y \in B$  it is true that  $g(y) \neq z$ . But  $g \circ f$  is surjective, and thus there is an  $x \in A$  such that  $(g \circ f)(x) = z$ . But  $f : A \rightarrow B$  is a function, and thus there is a  $y \in B$  such that  $f(x) = y$ . But then  $(g \circ f)(x) = g(f(x)) = g(y) = z$ , a contradiction. Thus,  $g$  is surjective.  $\square$

**Example 5.1.8** The full converse of Thm. 5.1.16 is not true, in general. For let  $A = \mathbb{R}$ ,  $B = \mathbb{R}^2$ , and  $C = \mathbb{R}$ . Define  $f$  and  $g$  by:

$$f(x) = (x, 0) \quad (5.1.18a) \quad g(x, y) = x \quad (5.1.18b)$$

Then  $g \circ f$  is simply the identity function. That is:

$$(g \circ f)(x) = g(f(x)) = g(x, 0) = x \quad (5.1.19)$$

and this is certainly surjective. However,  $f$  is not surjective. To see this, note that the point  $(0, 1)$  is never mapped to. Thus, the converse of Thm. 5.1.16 is not true.

With surjectivity we can recover the converse of Thm. 5.1.4.

**Theorem 5.1.18.** *If  $A$  and  $B$  are sets, if  $\mathcal{U} \subseteq B$ , and if  $f : A \rightarrow B$  is a surjective function, then:*

$$f(f^{-1}(\mathcal{U})) = \mathcal{U}$$

*Proof.* For suppose not. By Thm. 5.1.4,  $f(f^{-1}(\mathcal{U})) \subseteq \mathcal{U}$ , and thus by the definition of equality (Def. 3.1.3),  $\mathcal{U} \not\subseteq f(f^{-1}(\mathcal{U}))$ . But then there is a  $y \in \mathcal{U}$  such that  $y \notin f(f^{-1}(\mathcal{U}))$ . But  $f$  is surjective, and thus there is an  $x \in A$  such that  $f(x) = y$  (Def. 5.1.2). But if  $y \in \mathcal{U}$ , then  $x \in f^{-1}(\mathcal{U})$  (Def. 3.1.17). But if  $x \in f^{-1}(\mathcal{U})$ , then  $f(x) \in f(f^{-1}(\mathcal{U}))$  (Def. 3.1.16), and thus  $y \in f(f^{-1}(\mathcal{U}))$ , a contradiction.  $\square$

### 5.1.2 Injections

#### Definition 5.1.3: Injective Function

An **injective function** is a function  $f : A \rightarrow B$  such that for all distinct  $x, y \in A$  it is true that  $f(x) \neq f(y)$ .

That is, an injective function is a function  $f : A \rightarrow B$  such that  $f(x_1) = f(x_2)$  if and only if  $x_1 = x_2$ . Such functions are also called *one-to-one*.

**Example 5.1.9** Consider the natural logarithm  $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ . This is an injective function. For let  $x, y \in \mathbb{R}^+$  be such that  $x \neq y$ . Suppose  $\ln(x) = \ln(y)$ . But then:

$$\ln(x) - \ln(y) = \ln\left(\frac{x}{y}\right) = 0 \quad (5.1.20)$$

Recall the definition of the natural logarithm:

$$\ln(t) = \int_1^t \frac{1}{x} dx \quad (5.1.21)$$

But then  $\ln(t) = 0$  if and only if  $t = 1$ . Thus  $x = y$ , a contradiction. Therefore  $\ln$  is an injective function. Not every function is injective, for define  $f : \mathbb{R} \rightarrow \mathbb{R}$  by  $f(x) = x^2$ . Then, for all  $x \in \mathbb{R}^+$ ,  $f(-x) = f(x)$ , and thus  $f$  cannot be an injective function.

One might think that most functions are not injective, and indeed for the *finite* case, this is true. For let  $A$  and  $B$  be finite sets with  $n$  and  $m$  elements, respectively. If  $m < n$ , there can't be any injective function. Consider the case when  $n = m$ . Then we are simply counting the number of ways to permute the elements of  $A$ . This is  $n!$ . On the other hand, the total number of functions is  $n^n$ . Thus, the ratio of the number of injective functions to the number of functions is  $n!/n^n$ , and this decays to zero rapidly as  $n$  gets large. Finally, if  $m > n$ , then the total number of injective functions is  $n! \binom{m}{n}$ , where  $\binom{m}{n}$  is the binomial coefficient. The total number of functions is  $n^m$ . The ratio is thus:

$$\frac{n! \binom{m}{n}}{n^m} = \frac{n! \frac{m!}{n!(m-n)!}}{n^m} = \frac{m!}{(m-n)!n^m} \quad (5.1.22)$$

And again, this decays rapidly to zero and  $n$  and  $m$  get large. Later, when we define infinite sets and the notion of Cardinality, we'll show that this trend continues. That is, in a sense, *most* functions from a set  $A$  to a sufficiently large set  $B$  are not injective.

### 5.1.3 Bijections

#### Definition 5.1.4: Bijective Functions

A **bijective function** is a function that is both injective and surjective.

**Example 5.1.10** The perspectivity mapping shown in Fig. 3.14 is a geometric example of a bijective mapping from one line to another. To prove this requires the axioms of Euclid and shall be saved for later during our discussion of geometry. It is hoped that the figure makes this claim clear.

#### Definition 5.1.5: Permutation

A **permutation** on a **set**  $A$  is a **bijective function**  $f : A \rightarrow A$ .

**Theorem 5.1.19.** If  $A$  and  $B$  are sets and if  $f : A \rightarrow B$  is a bijective function, then  $f^{-1} : B \rightarrow A$  is bijective.

*Proof.* For if  $f^{-1}(y_1) = f^{-1}(y_2)$ , then there exists  $x \in A$  such that  $f(x) = y_1$  and  $f(x) = y_2$ . But  $f$  is bijective, and therefore  $y_1 = y_2$ . Thus,  $f^{-1}$  is injective. But by definition,  $f^{-1}$  is surjective, and therefore it is bijective.  $\square$

**Theorem 5.1.20.** If  $A, B$ , and  $C$  are sets, if  $f : A \rightarrow B$  and  $g : B \rightarrow C$  are bijective functions, and if  $\mathcal{V} \subset C$ , then  $(g \circ g)^{-1}(\mathcal{V}) = f^{-1}(g^{-1}(\mathcal{V}))$ .

**Theorem 5.1.21.** If  $A, B$ , and  $C$  are sets, and if  $f : A \rightarrow B$  and  $g : B \rightarrow C$  are bijective functions, then  $g \circ f : A \rightarrow C$  is bijective.

*Proof.* For if  $f$  and  $g$  are bijective, then they are surjective (Def. 5.1.4). But if  $f$  and  $g$  are surjective, then  $g \circ f$  is surjective (Thm. 5.1.16). But if  $f$  and  $g$  are bijective, then they are injective (Def. 5.1.4). But if  $f$  and  $g$  are injective, then  $g \circ f$  is injective. Thus  $g \circ f$  is surjective and injective, and is therefore bijective (Def. 5.1.4).  $\square$

**Theorem 5.1.22.** If  $f : A \rightarrow B$  is bijective,  $A_1 \subset A$ , and  $f(A_1) = B$ , then  $A_1 = A$ .

*Proof.* For if  $A \setminus A_1$  is non-empty, then the image  $f(A \setminus A_1)$  is also non-empty (Thm. 5.1.6). But  $f$  is bijective and therefore  $f(A_1) \cap f(A \setminus A_1) = \emptyset$ . But then there exists  $y \in B$  such that  $y \notin f(A_1)$ , a contradiction.  $\square$

### 5.1.4 Compositions

The composition of two functions is an important concept that allows one to construct new functions from old ones. Given a function  $f : A \rightarrow B$  and a function  $g : B \rightarrow C$ , we seek a function  $h : A \rightarrow C$  that arises in a *natural* way.

One way to describe this is by using *commutative diagrams*. Consider the diagram in Fig. 5.1.

The function  $h$  takes an element of  $A$  and maps it to  $C$ . We say that this diagram commutes if following the arrows  $A \rightarrow B \rightarrow C$  leads to the same result as  $A \rightarrow C$ . That is, take  $x \in A$  and map it to some point  $y \in B$  under the function  $f$ . Then take  $y$  map this to  $C$  under  $g$ . The resulting element in  $C$  should be the same value that  $x$  is mapped to under  $h$ . We define  $h$  to be the *function composition* of  $f$  and  $g$  and we denote this by  $g \circ f$ . The formula for composition then becomes clear:

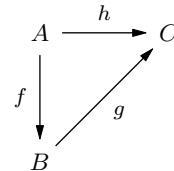


Fig. 5.1: A Commutative Diagram

$$(g \circ f)(x) = g(f(x)) \quad (5.1.23)$$

Note that  $g \circ f : A \rightarrow C$  is a function from  $A$  into  $C$ .

#### Definition 5.1.6: Function Composition

The composition of a function  $f : A \rightarrow B$  with a function  $g : B \rightarrow C$  is the function  $g \circ f : A \rightarrow B$  defined by:

$$(g \circ f)(x) = g(f(x)) \quad x \in A$$

**Example 5.1.11** Let  $A = [0, \infty)$  denote all of the non-negative real numbers and let  $f : A \rightarrow \mathbb{R}$  be defined by  $f(x) = \sqrt{x}$  and  $g : \mathbb{R} \rightarrow A$  by  $g(x) = x^2$ . Since the target of  $f$  is equal to the domain of  $g$ , the composition  $g \circ f$  is well defined. We have:

$$(g \circ f)(x) = (\sqrt{x})^2 = x \quad (5.1.24)$$

So  $g \circ f : A \rightarrow A$  is simply the identity function,  $g \circ f = \text{id}_A$ .

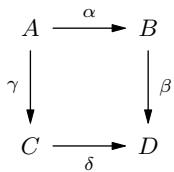


Fig. 5.2: A Slightly Complicated Commutative Diagram

When we have several sets and several functions between them, it is often convenient to specify function compositions by drawing the commutative diagram that corresponds to the given configuration. Suppose we have four sets  $A, B, C$ , and  $D$  and four functions  $\alpha : A \rightarrow B$ ,  $\beta : B \rightarrow D$ ,  $\gamma : A \rightarrow C$ , and  $\delta : C \rightarrow D$ . We can write out explicitly that  $\beta \circ \alpha = \delta \circ \gamma$ , or we can say that Fig. 5.2 commutes. For more complicated scenarios, we'd very much prefer to use diagrams. In homological algebra one comes across the Five Lemma, which involves the following commutative diagram shown in Fig. 5.3.

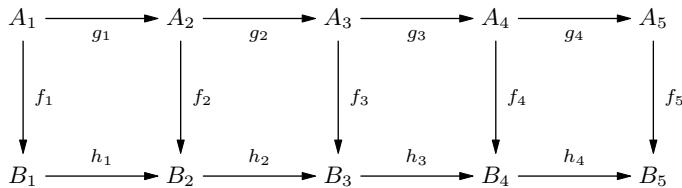


Fig. 5.3: A Very Complicated Commutative Diagram

It would be very tedious to directly spell out all of the compositions involved in this theorem, and so usually one uses commutative diagrams to ease this effort. One of the most important properties of function composition is that it is associative. That is, if  $f : A \rightarrow B$ ,  $g : B \rightarrow C$ , and  $h : C \rightarrow D$  are functions, then the functions  $h \circ (g \circ f) : A \rightarrow D$  and  $(h \circ g) \circ f : A \rightarrow D$  are the same. This can be expressed in terms of the commutative diagram given below (Fig. 5.4). As a side note, the diagram shown in Fig. 5.4 has an edge connecting every vertex to every other vertex. That is,  $A$  is connected to  $B, C$ , and  $D$ , and similarly for all of the other combinations. It turns out that four vertices is the largest number possible for such a configuration to be drawn without edges crossing. If one tries to connect all of the points on a pentagon to each other, one must have two of the edges cross. This is the topic of planarity, and is discussed in Chapt. 9

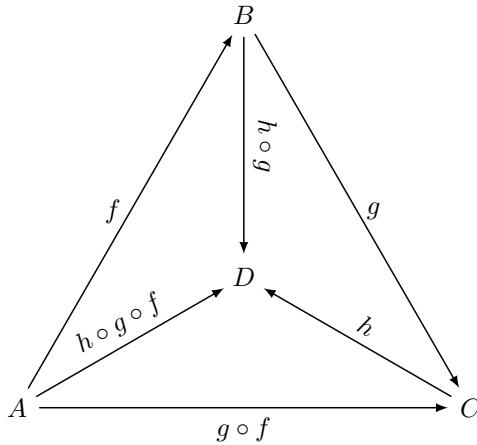


Fig. 5.4: Associativity of Function Composition

### Theorem 5.1.23: Associativity of Composition

If  $A$ ,  $B$ ,  $C$ , and  $D$  are sets, if  $f : A \rightarrow B$ ,  $g : B \rightarrow C$ , and  $h : C \rightarrow D$  are functions, if  $h \circ g : B \rightarrow D$  is the composition of  $h$  with  $g$ , and if  $g \circ f : A \rightarrow C$  is the composition of  $g$  with  $f$ , then:

$$h \circ (g \circ f) = (h \circ g) \circ f$$

*Proof.* For if not, then there is an  $x \in A$  such that  $(h \circ (g \circ f))(x) \neq ((h \circ g) \circ f)(x)$ . But by the definition of function composition (Def. 5.1.6), we have:

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) \quad (5.1.25)$$

But also:

$$((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x))) \quad (5.1.26)$$

By the transitivity of equality (Thm. 3.2.11),  $(h \circ (g \circ f))(x) = ((h \circ g) \circ f)(x)$ , a contradiction.  $\square$

One of the most important aspects of functions is the notion of *inverses*.

**Definition 5.1.7: Right Invertible Function**

A right invertible function from a set  $A$  to a set  $B$  is function  $f : A \rightarrow B$  such that there exists a function  $g : B \rightarrow A$  such that  $f \circ g = \text{id}_B$ .

**Example 5.1.12** If we consider  $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , where  $\mathbb{R}^+$  is the set of positive real numbers, given by  $f(x) = x^2$ , then  $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  defined by  $g(x) = \sqrt{x}$  is a right inverse since  $(f \circ g)(x) = (\sqrt{x})^2 = x$ , and thus  $f \circ g = \text{id}_{\mathbb{R}^+}$ .

**Example 5.1.13** If we let  $f : \mathbb{Z} \rightarrow \mathbb{Z}$  be defined by  $f(n) = n + 1$ , then  $g : \mathbb{Z} \rightarrow \mathbb{Z}$  defined by  $g(n) = n - 1$  is a right inverse since  $(f \circ g)(n) = f(n - 1) = (n - 1) + 1 = n$ , and hence  $f \circ g = \text{id}_{\mathbb{Z}}$ .

A common feature of both of these examples is that the function  $f$  is surjective. This is both a necessary and sufficient condition for a function to have a right inverse.

**Theorem 5.1.24.** *If  $A$  and  $B$  are sets, and if  $f : A \rightarrow B$  is a function, then there exists a right inverse  $g : B \rightarrow A$  if and only if  $f$  is surjective.*

*Proof.* For suppose  $f$  is surjective, and let  $G : B \rightarrow \mathcal{P}(A)$  be defined by:

$$G(y) = f^{-1}[\{y\}] \quad (5.1.27)$$

and let  $\mathcal{O} = G[B]$ . Then  $\emptyset \notin \mathcal{O}$ . For suppose not. But  $\emptyset \in \mathcal{O}$  if and only if there is a  $y \in B$  such that  $f^{-1}[\{y\}] = \emptyset$ . But  $f$  is surjective, and thus if  $y \in B$  then there is an  $x \in A$  such that  $f(x) = y$  (Def. 5.1.2). But if  $f(x) = y$ , then  $x \in f^{-1}[\{y\}]$  (Def. 3.1.18), a contradiction since for all  $x$  it is true that  $x \notin \emptyset$  (Def. 3.2.1). Hence,  $\emptyset \notin \mathcal{O}$ . But then  $\mathcal{O}$  is a collection of non-empty sets, and hence by the axiom of choice there is a function  $F : \mathcal{O} \rightarrow \bigcup \mathcal{O}$  such that for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $F(\mathcal{U}) \in \mathcal{U}$  (Ax. 3.1.8). But  $f$  is a function, and hence for all  $x \in A$  there is a  $y \in B$  such that  $f(x) = y$  (Def. 3.1.14). But then for all  $x \in A$ , there is a  $\mathcal{U} \in \mathcal{O}$  such that  $x \in \mathcal{U}$ , and hence  $x \in \bigcup \mathcal{O}$  (Def. 3.1.6). But then  $X \subseteq \bigcup \mathcal{O}$  (Def. 3.1.2), and hence  $X = \bigcup \mathcal{O}$ . Thus,  $F$  is a function from  $\mathcal{O}$  to  $X$ . Let  $g : B \rightarrow A$  be defined by  $F \circ G$ . But then for all  $y \in B$ ,  $g(x) = (F \circ G)(y) = F(G(y))$ , and thus by the definition of  $F$  and  $G$ ,  $g(y) \in f^{-1}[\{y\}]$ . But then  $f(g(x)) \in \{y\}$  (Def. 3.1.18), and therefore  $(f \circ g)(y) = y$ , and therefore  $f \circ g$  is the identity function. Therefore,  $f$  is right invertible (Def. 5.1.7). In the other direction, if  $f$  is right invertible, then there is a function  $g : B \rightarrow A$  such that  $f \circ g = \text{id}_B$  (Def. 5.1.7). Suppose  $f$  is not surjective. Then there is a  $y \in B$  such that for all  $x \in A$  it is true that  $f(x) \neq y$  (Def. 5.1.2). But  $f \circ g = \text{id}_B$ , and thus  $(f \circ g)(y) = y$  (Def. 5.1.1). But then  $g(y) \in A$  is such that  $f(g(y)) = y$ , a contradiction. Therefore,  $f$  is surjective.  $\square$

**Definition 5.1.8: Left Invertible Function**

A left invertible function from a set  $A$  to a set  $B$  is a function  $f : A \rightarrow B$  such that there exists a function  $g : B \rightarrow A$  such that  $g \circ f = \text{id}_A$ .

**Theorem 5.1.25.** *If  $A$  and  $B$  are non-empty sets, and if  $f : A \rightarrow B$  is a function, then  $f$  is left invertible if and only if  $f$  is injective.*

*Proof.* For suppose  $f$  is injective. Since  $A$  is non-empty, there is an element  $a \in A$  (Def. 3.1.1). Let  $g \subseteq B \times A$  be defined by:

$$g = \{ (y, x) \in B \times A \mid \exists_{x \in A} (y = f(x)) \text{ or } (y \notin f[A] \text{ and } x = a) \} \quad (5.1.28)$$

Then  $g$  is a function from  $B$  to  $A$ . for if  $y \in B$ , by the law of the excluded middle either  $y \in f[A]$  or  $y \notin f[A]$ . But if  $y \in f[A]$ , then there is an  $x \in A$  such that  $y = f(x)$  (Def. 3.1.16). But then  $(y, x) \in g$ . If  $y \notin f[A]$ , then  $(y, a) \in g$ . Thus, if  $g$  is not a function, then there is a  $y \in B$  such that there are two distinct elements  $x_1, x_2 \in A$  such that  $(y, x_1) \in g$  and  $(y, x_2) \in g$ . But if  $y \notin f[A]$ , then  $x_1 = a$  and  $x_2 = a$ , and hence by the transitivity of equality,  $x_1 = x_2$  (Thm. 3.2.11), a contradiction. Hence  $y \in f[A]$ . But if  $(y, x_1) \in g$  and  $(y, x_2) \in g$ , then  $f(x_1) = y$  and  $f(x_2) = y$ . Thus, by the transitivity of equality,  $f(x_1) = f(x_2)$  (Thm. 3.2.11). But  $f$  is injective, and therefore if  $f(x_1) = f(x_2)$ , then  $x_1 = x_2$  (Def. 5.1.3), a contradiction. Therefore,  $g : B \rightarrow A$  is a function. Moreover, for all  $x \in A$ ,  $(g \circ f)(x) = g(f(x)) = x$  by definition of  $g$ , and thus  $f$  is left invertible (Def. 5.1.8). In the other direction, if  $f$  is left invertible, then there is a function  $g : B \rightarrow A$  such that  $g \circ f = \text{id}_A$ . Suppose  $f$  is not injective. Then there exists  $x_1, x_2 \in A$  such that  $x_1 \neq x_2$  and  $f(x_1) = f(x_2)$  (Def. 5.1.3). But then:

$$x_1 = g(f(x_1)) = g(f(x_2)) = x_2 \quad (5.1.29)$$

a contradiction. Therefore,  $f$  is injective.  $\square$

**Definition 5.1.9: Invertible Function**

An invertible function from a set  $A$  to a set  $B$  is a function  $f : A \rightarrow B$  such that  $f$  is right invertible and left invertible.

**Theorem 5.1.26.** *If  $A$  and  $B$  are non-empty sets, and if  $f : A \rightarrow B$  is a function, then  $f$  is invertible if and only if it is a bijection.*

*Proof.* For if  $f$  is invertible, then it is right invertible and left invertible. But if  $f$  is right invertible, then it is surjective (Thm. 5.1.24). And if  $f$  is left invertible,

then it is injective (Thm. 5.1.25). But then  $f$  is injective and surjective, and hence it is bijective (Def. 5.1.4). In the other direction, if  $f$  is bijection, then it is injective and surjective.  $\square$

## 5.2 Binary Operations

Binary operations are the standard tools used to develop arithmetic. As such, the most familiar examples of binary operations are addition, multiplication, and subtraction with real numbers. Division is *not* a binary operation on the real numbers since division by zero is undefined. To make this explicit we need to give a rigorous definition to binary operations. We can do this with the language of functions and Cartesian products.

### Definition 5.2.1: Binary Operation

A **binary operation** on a **set**  $A$  is a function  $* : A \times A \rightarrow A$ .

**Example 5.2.1** Let  $\mathbb{R}$  be the set of real numbers and  $+$  denote the addition of two real numbers. Then  $+$  is a binary operation on  $\mathbb{R}$ . Similarly, if  $\cdot$  denotes the multiplication of two real numbers, than it too is a binary operation on  $\mathbb{R}$ . For division we are lacking the requirement that *for all*  $(a, b) \in \mathbb{R}^2$  there is a unique  $c \in \mathbb{R}$  such that  $a \div b = c$ , since if  $b = 0$  our expression is undefined. That is, this is not a function from  $\mathbb{R}^2$  to  $\mathbb{R}$ . If we consider all of the non-zero elements, then division is a binary operation. That is, division is a binary operation on  $\mathbb{R} \setminus \{0\}$ .

### Example 5.2.2: Binary Operation on the Set of Functions

If  $A$  is a set, and if  $\mathcal{F}(A, A)$  denotes the set of all functions  $f : A \rightarrow A$ , and if  $\circ$  denotes function composition, then  $\circ$  is a binary operation on  $\mathcal{F}(A, A)$ . That is, for any two functions  $f, g \in \mathcal{F}(A, A)$ , the composition  $g \circ f : A \rightarrow A$  is again an element of  $\mathcal{F}(A, A)$   $\blacksquare$

Just like functions, there are three important conditions that a binary operation must satisfy. Given any ordered pair  $(a, b) \in A \times A$ , it must be true that  $*(a, b)$  is defined. This comes from the definition of a function on a set (Def. 3.1.14). Next, the image of  $(a, b)$  must be unique. That is, if  $*(a, b) = c$  and  $*(a, b) = d$ , then  $c = d$ . Note that this is not the same as requiring that  $*(a, b) = *(b, a)$ , and in general this is not true. Such binary operations are called *commutative*. Lastly, for any  $(a, b) \in A \times A$ ,  $*(a, b)$  must be an element of  $A$ . That is,

$*(a, b) \in A$ . All of these requirements come from the definition of a function, so in a sense it is redundant to repeat these. In practice one defines a binary operation by a formula  $\varphi$ , and it then becomes necessary to show that this formula satisfies these properties before we can rightly call it a binary operation.

**Example 5.2.3** Let  $A = \mathbb{Z}_2$  and consider all of the binary operations on  $\mathbb{Z}_2$ . We can count these by constructing tables:

$(x, y)$	$*(x, y)$
$(0, 0)$	0
$(0, 1)$	0
$(1, 0)$	1
$(1, 1)$	0

Table 5.1: Simple Binary Operation on  $\mathbb{Z}_2$

This is one such binary operation, there are 15 others. To see this, recall that the number of functions from a set  $A$  to a set  $B$ , where both  $A$  and  $B$  are finite sets with  $m$  and  $n$  elements, respectively, is  $n^m$ . Since  $\mathbb{Z}_2$  has 2 elements, and since a binary operation is a function  $*: \mathbb{Z}_2 \times \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$ , the total number of binary operations is  $2^{(2^2)} = 2^4 = 16$ . In general, if  $A$  has  $n$  elements, and if  $B$  is the set of all binary operations on  $A$ , then:

$$\text{Card}(B) = n^{(n^2)} \quad (5.2.1)$$

**Example 5.2.4** Let's consider formulas that take in numbers and return numbers and see if they can define operations on various sets. Suppose we have:

$$a * b = \{ r \in \mathbb{R} \mid r^2 = |ab| \} \quad (5.2.2)$$

Where  $|ab|$  denotes the absolute value of  $a$  times  $b$ . If we take the positive square root we can write this as  $a * b = \sqrt{|ab|}$ . If we consider this formula on the rational numbers  $\mathbb{Q}$ , does it define a binary operation? One might recall that  $\sqrt{2}$  is not a rational number, and thus  $1 * 2$  is not well-defined. Hence,  $*$  is not a binary operation on  $\mathbb{Q}$ . It is a binary operation on  $\mathbb{R}$ , however. Suppose we change the formula to state:

$$a * b = \{ r \in \mathbb{R} \mid r^2 - a^2 b^2 = 0 \} \quad (5.2.3)$$

and where we consider this formula to take inputs from  $\mathbb{R}$ . This is not a binary operation since it is poorly defined. That is, should  $1 * 1 = 1$ , or should  $1 * 1 = -1$ ? The formula is ambiguous and thus  $*$  is not a binary operation.

**Example 5.2.5** If we consider subtraction on the integers  $\mathbb{Z}$ , this is a binary operation. The operation is well defined and returns an integer for all integer

inputs. If instead we consider subtraction on  $\mathbb{N}$ , this is *not* a binary operation since it may take in non-negative integers and return a negative integer. For example,  $1 - 2 = -1$ , and  $-1 \notin \mathbb{N}$ . A simple fix for this is considering again the absolute value function. If we define  $n * m = |n - m|$ , then  $*$  is indeed a binary operation on  $\mathbb{N}$ .

### Notation 5.2.1: Binary Operation

If  $A$  is a set and if  $* : A \times A \rightarrow A$  is a binary operation on  $A$ , for any ordered pair  $(a, b) \in A^2$ , the image of  $*(a, b)$  is denoted  $a * b$ .

Given a binary operation  $*$  on a set  $A$ , and given three distinct elements  $a, b, c \in A$ , the expression  $a * b * c$  is ambiguous. A binary operation takes in two elements at a time, and thus the question arises as to whether this should denote  $(a * b) * c$  or  $a * (b * c)$ . To rid ourselves of such problems, we consider *associative* operations.

### Definition 5.2.2: Associative Operation

A *associative operation* on a *set*  $A$  is a *binary operation*  $*$  such that, for all  $a, b, c \in A$  it is true that  $a * (b * c) = (a * b) * c$ .

**Example 5.2.6** The usual arithmetic operations addition and multiplication are associative. Subtraction is not, for  $a - (b - c) = a + (-b) + c$ , whereas  $(a - b) - c = a + (-b) + (-c)$ , and these are only equal if  $c = 0$ . Similarly, division is not associative for  $(1/2)/3 = 1/6$ , whereas  $1/(2/3) = 3/2$ , and these are not equal.

**Example 5.2.7** We can place an arithmetic structure on the set of  $n \times n$  matrices over a set  $A$  if  $A$  has the structure of a *ring*. This is the standard arithmetic of matrices that one may be familiar with when one considers real valued entries. Both matrix addition and multiplication are associative.

**Example 5.2.8** Consider a finite set  $A$  and consider the set of all functions from  $\mathbb{Z}_n$  to  $A$ . That is,  $\mathcal{F}_n(\mathbb{Z}_n, A)$ . Define  $A[x]$  by:

$$\mathcal{F} = \bigcup_{n \in \mathbb{N}} \mathcal{F}_n \tag{5.2.4}$$

That is, the set of all finite sequences in  $A$ . We can form an associative operation on this set by defining the concatenation operation. Given  $f, g \in \mathcal{F}$ ,

suppose  $f \in \mathcal{F}(\mathbb{Z}_m, A)$  and  $g \in \mathcal{F}(\mathbb{Z}_n, A)$ . We define  $f * g \in \mathcal{F}(\mathbb{Z}_{m+n}, A)$  as follows:

$$(f * g)(k) = \begin{cases} f(k), & k \in \mathbb{Z}_m \\ g(k-m), & k \in \mathbb{Z}_{m+n} \text{ and } k \geq m \end{cases} \quad (5.2.5)$$

That is, given two sequences  $f_0, f_1, \dots, f_{m-1}$  and  $g_0, g_1, \dots, g_{n-1}$ , we concatenate them to form the sequence  $f_0, \dots, f_{m-1}, g_0, \dots, g_{n-1}$ . This operation is associative since if  $f, g, h \in \mathcal{F}$ , then:

$$f * (g * h) = (f_0, f_1, \dots, f_{m-1}) * (g_0, g_1, \dots, g_{m-1}, h_0, h_1, \dots, h_{r-1}) \quad (5.2.6a)$$

$$= f_0, f_1, \dots, f_{m-1}, g_0, g_1, \dots, g_{m-1}, h_0, h_1, \dots, h_{r-1} \quad (5.2.6b)$$

$$= (f_0, f_1, \dots, f_{m-1}, g_0, g_1, \dots, g_{m-1}) * (h_0, h_1, \dots, h_{r-1}) \quad (5.2.6c)$$

$$= (f * g) * h \quad (5.2.6d)$$

If  $A$  has more than one point than  $*$  is not commutative. For let  $f, g : \mathbb{Z}_1 \rightarrow A$  be defined by  $f(0) = a$  and  $g(0) = b$ , respectively. Then  $f * g = a, b$  but  $g * f = b, a$ , and thus  $f * g \neq g * f$ . There is, however, an identity. Consider  $\mathbb{Z}_0$ , which is the empty set. Any function from  $\mathbb{Z}_0$  to  $A$  is therefore the *empty sequence*. If we concatenate  $f$  with the empty sequence we get back  $f$ , and this then acts as our unital element.

### Definition 5.2.3: Idempotent

An idempotent element of a set  $A$  under a binary operation  $*$  is an element  $a \in A$  such that  $a * a = a$ .

It is occasionally useful to think of binary operations purely as functions, and so we will use function notation at these times. For the most part we will stick with notation defined in Not. 5.2.1. There are several types of binary operations worth studying, and several key properties that these operations can have. One of the most fundamental is the existence of a *unital element*, also known as an identity.

### Definition 5.2.4: Left Unital Element

A left unital element in a set  $A$  under a binary operation  $*$  on  $A$  is an element  $e_L \in A$  such that, for all  $a \in A$  it is true that  $e_L * a = a$ .

**Theorem 5.2.1.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , and if  $e_L$  is a left unital element of  $A$ , then  $e_L$  is idempotent.*

*Proof.* For  $e_L = e_L * e_L$  (Def. 5.2.4), and thus  $e_L$  is idempotent (Def. 5.2.3).  $\square$

**Example 5.2.9** From the definition of a left unital element (Def. 5.2.4) it would seem natural to define a right unital element. The importance is to note that the existence of a left unital element does not imply the existence of a right. Indeed, if  $A$  is a set and  $*$  is a binary operation, given a left identity  $e_L$  and a right identity  $e_R$  it will be true that  $e_R = e_L$  and thus all left and right unital elements will be the same (see Thm. 5.2.2). Thus to find counterexamples to the claim that the existence of a left unital element implies the existence of a right unital element we need to think of strange operations. Let  $A = \mathbb{R}$  and let  $*$  be defined by  $a * b = b$  for all  $a, b \in \mathbb{R}$ . Then every element of  $\mathbb{R}$  is a left unital element. Moreover, none of the elements of  $\mathbb{R}$  are right unitals.

### Definition 5.2.5: Right Unital Element

A right unital element of a set  $A$  under a binary operation  $*$  is an element  $e_R$  such that for all  $a \in A$  it is true that  $a * e_R = a$ .

**Example 5.2.10** Consider  $\mathbb{R}$  with the operation  $*$  defined by  $a * b = a + b + 1$ . This operation has a right unital element,  $-1$ . For:

$$a * (-1) = a + (-1) + 1 = a + 0 = a \quad (5.2.7)$$

And this is true for all  $a \in \mathbb{R}$ , so  $-1$  is a right unital element. It turns out this is also a left unital element, and hence a unital element, and this can be proven if addition is known to be an *associative* operation.

**Theorem 5.2.2.** If  $A$  is a set, if  $*$  is a binary operation on  $A$ , if  $e_L$  is a left unital element of  $A$ , and if  $e_R$  is a right unital element of  $A$ , then  $e_L = e_R$ .

*Proof.* For  $e_L = e_L * e_R = e_R$  (Defs. 5.2.4 and 5.2.5) and thus  $e_L = e_R$ .  $\square$

**Example 5.2.11** Consider a non-empty set  $A$  and the set of all functions from  $A$  to itself,  $\mathcal{F}(A, A)$ . Let  $\circ$  denote the binary operation of function composition. Then  $\mathcal{F}(A, A)$  has a right identity under  $\circ$ , and a left identity. For the identity function  $\text{id}_A$  acts as a right identity:

$$(f \circ \text{id}_A)(x) = f(\text{id}_A(x)) = f(x) \quad (5.2.8)$$

And thus  $\text{id}_A$  is a right identity. By Thm. 5.2.2, any left identity must also be a right identity, and so the likely candidate to check is  $\text{id}_A$ . And indeed we have:

$$(\text{id}_A \circ f)(x) = \text{id}_A(f(x)) = f(x) \quad (5.2.9)$$

And thus  $\text{id}_A$  is a left identity as well.

**Theorem 5.2.3.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , if  $e_L$  and  $e'_L$  are left unital elements, and if  $e_R$  is a right unital element, then  $e_L = e'_L$*

*Proof.* For if  $e_L$  and  $e'_L$  are left unitals and  $e_R$  is a right unital, then  $e_L = e_R$  and  $e'_L = e_R$  (Thm. 5.2.2). By the transitivity of equality,  $e_L = e'_L$ .  $\square$

The same is true of right identities and thus if we find a left identity and a right identity, then they're the same and they are *the* identity.

### Definition 5.2.6: Unital Element

A **unital element** of a **set**  $A$  under a **binary operation**  $*$  is an element  $e \in A$  that is both a right unital element and a left unital element.

**Example 5.2.12** Let  $\mathbb{R}$  be the set of real numbers and let  $+$  be the usual notion of addition. Then  $0$  is a unital element of  $\mathbb{R}$  with respect to this operation. That is, for any real number  $x$  we have  $x + 0 = 0 + x = x$ . For multiplication the unital element is  $1$ . This is because  $1 \cdot x = x \cdot 1 = x$ . Subtraction has a right unital element, and again it is  $0$  since  $x - 0 = x$ , but no left identity. To see this, suppose  $e - x = x$  for all  $x$ . Applying some algebra we have that  $e = 2x$ , meaning there is no constant  $e \in \mathbb{R}$  such that for all  $x$ ,  $e - x = x$ . Since subtraction has no left unital element, it has no unital element either.

**Theorem 5.2.4.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , and if  $e$  and  $e'$  are unital elements of  $A$ , then  $e = e'$*

*Proof.* For then  $e$  is a right unital element and  $e'$  is a left unital element (Def. 5.2.6). But then  $e = e'$  (Thm. 5.2.2).  $\square$

**Theorem 5.2.5.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , if  $e$  is a unital element, and  $e_R$  is a right unital element, then  $e = e_R$ .*

*Proof.* For if  $e$  is a unital element, then it is a left unital element (Def. 5.2.6). But if  $e$  is a left unital element and  $e_R$  is a right unital element, then  $e = e_R$  (Thm. 5.2.2).  $\square$

Thus, if one has an operation that contains a unital element  $e$ , then there is only one right unital element, and only one left unital element, and that is  $e$ . The next thing to discuss is the notion of inverses.

### Definition 5.2.7: Weakly Right Invertible

A weakly right invertible element in a **set**  $A$  under a **binary operation**  $*$  on  $A$  is an element  $a \in A$  such that there is a  $b \in A$  such that  $a * b$  is a right unital element.

That is to say, an element  $a \in A$  is weakly right invertible if there is a  $b \in A$  such that, for all  $r \in A$  the following is true:

$$r * (a * b) = r \quad (5.2.10)$$

The reason for the word weakly is because we do not require  $r * (a * b) = r$ . The justification for *right* is because we also do not require  $r * (b * a) = r$ . That is,  $a$  is weakly right invertible if there is some other element  $b$  that we can multiply on the right of  $a$  such that the product acts as a right identity. One would think that such a restriction would be useless, but it turns out the notions allow one to determine if something is a *group*.

**Theorem 5.2.6.** *If  $A$  is a set, if  $*$  is an associative operation on  $A$ , if  $a \in A$  is weakly right invertible, and if  $b$  is such that  $a * b$  is a right unital element, then  $b * a$  is idempotent.*

*Proof.* For:

$$\begin{aligned} b * a &= (b * (a * b)) * a && (a * b \text{ is a right unital element (Def. 5.2.5)}) \\ &= ((b * a) * b) * a && (\text{Associativity}) \\ &= (b * a) * (b * a) && (\text{Associativity}) \end{aligned}$$

And therefore  $b * a$  is idempotent (Def. 5.2.3).  $\square$

**Theorem 5.2.7.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $a, b \in A$  are weakly right invertible, and if  $a * b$  is a right unital element, then  $b * a$  is a right unital element.*

*Proof.* For if  $b$  is weakly right invertible, then there is a  $c \in A$  such that  $b * c$  is a right unital element (Def. 5.2.7). But then:

$$\begin{aligned} b * a &= b * (a * (b * c)) && (b * c \text{ is a right unital element (Def. 5.2.5)}) \\ &= b * ((a * b) * c) && (\text{Associativity}) \\ &= (b * (a * b)) * c && (\text{Associativity}) \\ &= b * c && (a * b \text{ is a right unital element (Def. 5.2.5)}) \end{aligned}$$

And thus by transitivity,  $b * a = b * c$ . But  $b * c$  is a right unital element and therefore  $b * a$  is a right unital element.  $\square$

**Theorem 5.2.8.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $e \in A$  is a left unital element, and if  $a \in A$  is weakly right invertible, then  $e$  is a unital element.*

*Proof.* For if  $a$  is weakly right invertible, then there is a  $b \in A$  such that  $a * b$  is a right unital element (Def. 5.2.7). But if  $e$  is a left unital element and  $a * b$  is a right unital element, then  $e = a * b$  (Thm. 5.2.2). Therefore,  $e$  is a unital element.  $\square$

**Theorem 5.2.9.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $e$  is a right unital element, and if  $a, b \in A$  are weakly right invertible elements such that  $a * b = e$ , then  $e * a = a$ .*

*Proof.* For:

$$\begin{aligned} e * a &= (a * b) * a && (\text{Hypothesis}) \\ &= a * (b * a) && (\text{Associativity}) \\ &= a && (\text{Thm. 5.2.7}) \end{aligned}$$

And therefore  $e * a = a$ . □

**Theorem 5.2.10.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $e \in A$  is a unique right unital element, and if for all  $a \in A$  it is true that  $a$  is weakly right invertible, then  $e$  is a unital element.*

*Proof.* For suppose not. Then there is an  $a \in A$  such that  $e * a \neq a$ . But since  $a \in A$ ,  $a$  is weakly right invertible and thus there is a  $b \in A$  such that  $a * b$  is a right unital element (Def. 5.2.7). But by hypothesis  $e$  is the unique right unital element, and thus  $a * b = e$ . But since  $b \in A$ , by hypothesis  $b$  is weakly right invertible. But then  $a$  and  $b$  are weakly right invertible elements such that  $a * b = e$ , and thus by Thm. 5.2.9,  $e * a = a$ , a contradiction. Therefore,  $e$  is a unital element. □

We've almost set up the definition of a group. We have that a set with an associative binary operation that contains a unique right unital element and such that all elements are weakly right invertible will necessarily have a unique unital element. Next we need to show that every element will also be weakly left invertible, and we'll have a group. It is crucial to note the requirement that the right unital element in Thm. 5.2.10 is unique. If it is not, we may not have a unital element at all. Consider again the operation  $a * b = a$  on some set with at least two element. Then every element is a right unital element, and similarly every element is weakly right invertible since, for any  $a, b$ , we have:

$$a * (b * b) = a * b = a \quad (5.2.11)$$

and thus  $b$  is weakly right invertible, and is it's own weak inverse. This structure has no left unital element since  $e * a = e$  for any  $e \in A$ , and thus  $e$  cannot be a left unital. The uniqueness of  $e$  in Thm. 5.2.10 is what prevents such pathological examples from appearing.

### Definition 5.2.8: Right Invertible

A right invertible element of a **set**  $A$  under a **binary operation** is an element  $a \in A$  such that there exists  $b \in A$  such that  $a * b$  is a **unital element**.

**Theorem 5.2.11.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , and if  $a \in A$  is right invertible, then it is weakly right invertible.*

*Proof.* For then there is a  $b \in A$  such that  $a * b$  is a unital element (Def. 5.2.8). But then unital elements are right unital elements (Def. 5.2.6), and thus  $a * b$  is a right unital element. Therefore,  $a$  is weakly right invertible.  $\square$

Nothing to deep here and we've simply strengthened the requirement that  $a * b$  not only be a right unital element, but also a left unital element as well.

**Theorem 5.2.12.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $a$  is right invertible and idempotent, then  $a$  is a unital element.*

*Proof.* For if  $a$  is right invertible then there is a  $b \in A$  such that  $a * b$  is a unital element.

$$\begin{aligned} a &= a * (a * b) && (a * b \text{ is a unital element (Def. 5.2.6)}) \\ &= (a * a) * b && (\text{Associativity}) \\ &= a * b && (a \text{ is idempotent (Def. 5.2.3)}) \end{aligned}$$

And thus  $a = a * b$ . But  $a * b$  is a unital element, and thus so is  $a$ .  $\square$

### Definition 5.2.9: Weakly Left Invertible

A weakly left invertible element of a set  $A$  under a binary operation  $*$  is an element  $a \in A$  such that there exists a  $b \in A$  such that  $b * a$  is a left unital element.

All of the theorems about weakly right invertible elements apply to weakly left invertible elements, but we need to swap all of the orders of multiplication.

**Theorem 5.2.13.** *If  $A$  is a set, if  $*$  is an associative operation on  $A$ , if  $a \in A$  is weakly left invertible, and if  $b \in A$  is such that  $b * a$  is a left unital element, then  $a * b$  is idempotent.*

*Proof.* For:

$$\begin{aligned} a * b &= a * ((b * a) * b) && (b * a \text{ is a left unital element (Def. 5.2.4)}) \\ &= a * (b * (a * b)) && (\text{Associativity}) \\ &= (a * b) * (a * b) && (\text{Associativity}) \end{aligned}$$

And therefore  $a * b$  is idempotent (Def. 5.2.3).  $\square$

**Theorem 5.2.14.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $a, b \in A$  are weakly left invertible, and if  $b * a$  is a left unital element, then  $a * b$  is a left unital element.*

*Proof.* For if  $b$  is weakly left invertible, then there is a  $c \in A$  such that  $c * b$  is a left unital element (Def. 5.2.9). But then:

$$\begin{aligned} a * b &= ((c * b) * a) * b && (c * b \text{ is a left unital element (Def. 5.2.4)}) \\ &= (c * (b * a)) * b && (\text{Associativity}) \\ &= c * ((b * a) * b) && (\text{Associativity}) \\ &= c * b && (b * a \text{ is a left unital element (Def. 5.2.4)}) \end{aligned}$$

And thus by transitivity,  $a * b = c * b$ . But  $c * b$  is a left unital element and therefore  $a * b$  is a left unital element.  $\square$

**Theorem 5.2.15.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $e \in A$  is a left unital element, and if  $a \in A$  is weakly right invertible, then  $e$  is a unital element.*

*Proof.* For if  $a$  is weakly left invertible, then there is a  $b \in A$  such that  $b * a$  is a left unital element (Def. 5.2.7). But if  $e$  is a right unital element and  $b * a$  is a left unital element, then  $e = a * b$  (Thm. 5.2.2). Therefore,  $e$  is a unital element.  $\square$

**Theorem 5.2.16.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $e$  is a left unital element, and if  $a, b \in A$  are weakly left invertible elements such that  $b * a = e$ , then  $a * e = a$ .*

*Proof.* For:

$$\begin{aligned} a * e &= a * (b * a) && (\text{Hypothesis}) \\ &= (a * b) * a && (\text{Associativity}) \\ &= a && (\text{Thm. 5.2.14}) \end{aligned}$$

And therefore  $e * a = a$ .  $\square$

**Theorem 5.2.17.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $e \in A$  is a unique left unital element, and if for all  $a \in A$  it is true that  $a$  is weakly left invertible, then  $e$  is a unital element.*

*Proof.* For suppose not. Then there is an  $a \in A$  such that  $a * e \neq a$ . But since  $a \in A$ ,  $a$  is weakly left invertible and thus there is a  $b \in A$  such that  $b * a$  is a left unital element (Def. 5.2.9). But by hypothesis  $e$  is the unique right unital element, and thus  $b * a = e$ . But since  $b \in A$ , by hypothesis  $b$  is weakly left invertible. But then  $a$  and  $b$  are weakly left invertible elements such that  $b * a = e$ , and thus by Thm. 5.2.16,  $e * a = a$ , a contradiction. Therefore,  $e$  is a unital element.  $\square$

**Theorem 5.2.18.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $a \in A$  is a weakly left invertible element, and if  $b \in A$  is a weakly right invertible element, then there exists a unique  $e \in A$  such that  $e$  is a unital element.*

*Proof.* For if  $a$  is weakly left invertible, then there exists  $r_L \in A$  such that  $a_L * a$  is a left unital element (Def. 5.2.4). But if  $b$  is weakly right invertible then there is a  $b_R$  such that  $b * b_R$  is a right unital element. But then there exists a left unital element  $b_L * b$  and a right unital element  $a * a_R$ , and thus  $b_L * b = a * a_R$  (Thm. 5.2.2), and thus there exists a unital element (Def. 5.2.6). And unital elements are unique (Thm. 5.2.4), completing the proof.  $\square$

**Theorem 5.2.19.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , if there is a unique right unital element  $e \in A$ , and if for all  $a \in A$  it is true that  $a$  is weakly right invertible, then for all  $a \in A$  it is true that  $a$  is weakly left invertible.*

*Proof.* For suppose not and suppose there is an  $a \in A$  such that  $a$  is not weakly left invertible. But  $a$  is weakly right invertible and thus there is a  $b \in A$  such that  $a * b$  is a right unital element. But  $e$  is the unique right unital element, and thus  $a * b = e$ . But if  $a * b$  is a right unital element, then  $b * a$  is a right unital element (Thm. 5.2.9), and thus  $b * a = e$ . But since every element of  $A$  is weakly right invertible, and since  $e$  is the unique right unital element, it is true that  $e$  is a unital element (Thm. 5.2.10). But then  $b * a$  is a unital element, and is therefore a left unital element (Def. 5.2.6), a contradiction. Thus,  $a$  is weakly left invertible.  $\square$

### Definition 5.2.10: Left Invertible

A left invertible element of a set  $A$  under a binary operation is an element  $a \in A$  such that there exists a  $b \in A$  such that  $b * a$  is a unital element.

**Theorem 5.2.20.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , and if  $a \in A$  is left invertible, then it is weakly left invertible.*

*Proof.* For then there is a  $b \in A$  such that  $b * a$  is a unital element (Def. 5.2.10). But then unital elements are right unital elements (Def. 5.2.6), and thus  $b * a$  is a left unital element. Therefore,  $a$  is weakly left invertible.  $\square$

Again, a rather obvious theorem that comes straight from the definition. We use the notions of right and left invertible to define invertible.

**Theorem 5.2.21.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , and if  $a$  is left invertible and idempotent, then  $a$  is a unital element.*

*Proof.* For if  $a$  is left invertible, there is a  $b \in A$  such that  $b * a$  is unital. But then:

$$\begin{aligned} a &= (b * a) * a && (b * a \text{ is a unital element (Def. 5.2.6)}) \\ &= b * (a * a()) && (\text{Associativity}) \\ &= b * a && (a \text{ is idempotent (Def. 5.2.3)}) \end{aligned}$$

And thus  $a = b * a$ . But  $b * a$  is unital, and thus so is  $a$ .  $\square$

### Definition 5.2.11: Invertible Element

An invertible element of a set  $A$  under a binary operation is an element  $a \in A$  that is both left invertible and right invertible.

**Theorem 5.2.22.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , and if  $e$  is a unital element of  $A$ , then  $e$  is an invertible element.*

*Proof.* For since  $e$  is a unital element, it is true that  $e = e * e$  (Def. 5.2.6). Therefore  $e$  is invertible element (Def. 5.2.11).  $\square$

**Theorem 5.2.23.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , and if  $a \in A$  is weakly left invertible and weakly right invertible, then  $a$  is invertible.*

*Proof.* For if  $a$  is weakly right invertible and weakly left invertible, then there is a unique unital element  $e \in A$  (Thm. 5.2.18). But since  $a$  is weakly right invertible, there is a  $b$  such that  $a * b$  is a right unital element (Def. 5.2.7). And since  $e$  is the unique unital element, it is the unique right unital element and thus  $a * b = e$ . But since  $a$  is weakly left invertible there is a  $c \in A$  such that  $c * a$  is a left unital element (Def. 5.2.9), and again since  $e$  is unique we have that  $c * a = e$ . But then  $a$  has a right inverse and a left inverse (Defs. 5.2.8 and 5.2.10), and is therefore invertible (Def. 5.2.11).  $\square$

**Theorem 5.2.24.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $e$  is a unique right unital element, and if for all  $a \in A$  it is true that  $a$  is weakly right invertible, then for all  $a \in A$  it is true that  $a$  is invertible.*

*Proof.* For by Thm. 5.2.19, for all  $a \in A$  it is true that  $a$  is weakly left invertible. But then for all  $a \in A$ ,  $a$  is weakly left invertible and weakly right invertible, and is therefore invertible (Thm. 5.2.23).  $\square$

**Theorem 5.2.25.** *If  $A$  is a set, if  $*$  is a associative binary operation on  $A$ , if  $a$  is an invertible element, and if  $b, c \in A$  are such that  $b * a$  and  $a * c$  are unital elements, then  $b = c$ .*

*Proof.* For:

$$\begin{aligned} b &= b * (a * c) && (a * b \text{ is a unital element}) \\ &= (b * a) * c && (\text{Associativity}) \\ &= c && (b * a \text{ is a unital element}) \end{aligned}$$

And therefore  $b = c$ .  $\square$

**Theorem 5.2.26.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $a \in A$  is invertible, and if  $b, c \in A$  are such that  $a * b$  and  $a * c$  are unital elements, then  $b = c$ .*

*Proof.* For if  $a$  is invertible, there is an  $l \in A$  such that  $l * a$  is a unital element (Def. 5.2.6). But then  $l * a$  is a unital element and  $a * b$  is a unital element, and thus  $l = b$  (Thm. 5.2.25). Similarly  $l = c$ , and thus by the transitivity of equality,  $b = c$ .  $\square$

**Theorem 5.2.27.** *If  $A$  is a set, if  $*$  is an associative binary operation on  $A$ , if  $a \in A$  is invertible, and if  $b, c \in A$  are such that  $b * a$  and  $c * a$  are unital elements, then  $b = c$ .*

*Proof.* For if  $a$  is invertible, there is an  $r \in A$  such that  $r * a$  is a unital element (Def. 5.2.6). But then  $a * r$  is a unital element and  $b * a$  is a unital element, and thus  $b = r$  (Thm. 5.2.25). Similarly  $r = c$ , and thus by the transitivity of equality,  $b = c$ .  $\square$

Combining Thms. 5.2.26-5.2.27 we have that invertible elements have a unique *inverse* element. Depending on the operation under discussing, the inverse of  $a$  is denoted as either  $-a$  or  $a^{-1}$ .

**Theorem 5.2.28.** *If  $A$  is a set, if  $*$  is a binary operation on  $A$ , if  $a, b \in A$  are invertible elements, then  $a * b$  is invertible and  $(a * b)^{-1} = b^{-1} * a^{-1}$ .*

*Proof.* For if  $a$  is invertible, then there is an  $a^{-1}$  such that  $a * a^{-1}$  is a unital element (Def. 5.2.11). But unital elements are unique (Thm. 5.2.4). Let  $e \in A$  be the unital element. But  $b$  is also invertible and thus there is a  $b^{-1} \in A$  such that  $b * b^{-1} = e$ . Thus:

$$\begin{aligned} e &= a * a^{-1} && (\text{Inverse Property}) \\ &= (a * e) * a^{-1} && (\text{Identity}) \\ &= (a * (b * b^{-1})) * a^{-1} && (\text{Inverse Property}) \\ &= ((a * b) * b^{-1}) * a^{-1} && (\text{Associativity}) \\ &= (a * b) * (b * b^{-1}) && (\text{Associativity}) \end{aligned}$$

And therefore  $(a * b) * (b^{-1} * a^{-1})$  is a unital element, and thus  $a * b$  is right invertible (Def. 5.2.8). And similarly,  $(b^{-1} * a^{-1}) * (a * b) = e$ , and thus  $a * b$  is left invertible (Def. 5.2.10). Therefore,  $\square$

**Definition 5.2.12: Commuting Elements**

Commuting elements in a set  $A$  with respect to a binary operation  $*$  on  $A$  are elements  $a, b \in A$  such that  $a * b = b * a$ .

**Definition 5.2.13: Commutative Operation**

A commutative operation on a set  $A$  is a binary operation  $*$  such that for all  $(a, b) \in A^2$  it is true that  $a * b = b * a$ .

**Definition 5.2.14: Distributive Operation**

A distributive operation over a binary operation  $+$  on a set  $A$  is a binary operation  $*$  on  $A$  such that, for all  $a, b, c \in A$  the following is true:

$$a * (b + c) = (a * b) + (a * c)$$

## 5.3 Boolean Algebras

We now attempt to make set theory more algebraic. We wish to model as an object the triple  $(\mathcal{P}(X), \cup, \cap)$ , where  $\mathcal{P}(X)$  is the power set of some set, and  $\cup$  and  $\cap$  and union and intersection, respectively. These can be seen as binary operations on  $\mathcal{P}(X)$ . We take a few of the properties of this structure and state them as the definition for our new object: *Boolean Algebras*.

**Definition 5.3.1: Complement in a Boolean Algebra**

A complement of a set  $A$  with respect to two binary operations  $*$  and  $\circ$  is an element  $a^{-1} \in A$  such that:

$$a * a^{-1} = a^{-1} * a = e_{\circ} \quad a \circ a^{-1} = a^{-1} \circ a = e_{*}$$

Where  $e_{\circ}$  and  $e_{*}$  are the unital elements of  $\circ$  and  $*$ , respectively.

### Definition 5.3.2: Boolean Algebras

A Boolean algebra is a set  $A$  with two **commutative binary operations**  $\circ$  and  $*$  on  $A$  with **unital elements**  $e_*$  and  $e_\circ$ , respectively, such that:

- 1.)  $\circ$  **distributes** over  $*$  and  $*$  distributes over  $\circ$ .
- 2.) For all  $a \in A$  there is a complement of  $a$ .

The second property is known as the complement property and it is very different from the notion of inverses. An inverse of an element  $a$  with respect to an operation  $\cdot$  is such that  $a \cdot b$  is a unital element with respect to the operation  $\cdot$ . A complement produces a unital element with respect to the other operation. That is,  $a * a^{-1}$  is a unital element of  $\circ$ , and not  $*$ . Similarly,  $a \circ a^{-1}$  is a unital element of  $*$  and not  $\circ$ .

**Theorem 5.3.1.** *If  $(A, \circ, *)$  is a Boolean algebra and if  $b \in X$  is a unital element of  $\circ$ , then  $b = e_\circ$ .*

*Proof.* For unital elements are unique (Thm. 5.2.4), and therefore  $b = e_\circ$ .  $\square$

**Theorem 5.3.2.** *If  $(X, \circ, *)$  is a Boolean algebra and if  $b \in X$  is a unital element of  $*$ , then  $b = e_*$ .*

*Proof.* For unital elements are unique (Thm. 5.2.4), and therefore  $b = e_*$ .  $\square$

**Theorem 5.3.3.** *If  $(X, \circ, *)$  is a Boolean algebra, if  $e_*$  is the unital element of  $*$ , and if  $a \in X$ , then  $a \circ e_* = e_*$ .*

*Proof.* For if  $a \in X$  then there is an  $a^{-1} \in X$  such that  $a \circ a^{-1} = e_*$  (Def. 5.3.2). But then:

$$\begin{aligned} e_* &= a \circ a^{-1} && (\text{Complement}) && = e_* * (a \circ e_*) && (\text{Complement}) \\ &= a \circ (a^{-1} * e_*) && (\text{Identity}) && = a \circ e_* && (\text{Identity}) \\ &= (a \circ a^{-1}) * (a \circ e_*) && && && \\ &&& (\text{Distributivity}) && && \end{aligned}$$

And therefore  $e_* = a \circ e_*$ .  $\square$

This theorem is equivalent to the notion that a Boolean algebra is a bounded lattice and the  $e_*$  is a boundary. The theorem holds for  $\circ$  as well.

**Theorem 5.3.4.** *If  $(X, \circ, *)$  is a Boolean algebra, if  $e_*$  is the unital element of  $*$ , and if  $a \in X$ , then  $a * e_\circ = e_\circ$ .*

*Proof.* For if  $a \in X$  then there is an  $a^{-1} \in X$  such that  $a * a^{-1} = e_o$  (Def. 5.3.2). But then:

$$\begin{aligned} e_o &= a * a^{-1} && \text{(Complement)} &= e_o \circ (a * e_o) && \text{(Complement)} \\ &= a * (a^{-1} \circ e_o) && \text{(Identity)} &= a * e_o && \text{(Identity)} \\ &= (a * a^{-1}) \circ (a * e_*) && \text{(Distributivity)} \end{aligned}$$

And therefore  $e_o = a * e_o$ .  $\square$

**Theorem 5.3.5.** If  $(A, \circ, *)$  is a Boolean algebra, if  $e_o$  and  $e_*$  are the unital elements of  $\circ$  and  $*$ , respectively, then  $e_o$  is the complement of  $e_*$  and  $e_*$  is the complement of  $e_o$ .

*Proof.* From identity:

$$e_o \circ e_* = e_* \circ e_o = e_* \quad \text{(Identity)} \qquad e_o * e_* = e_* * e_o = e_o \quad \text{(Identity)}$$

Thus,  $e_*$  is a complement of  $e_o$  and  $e_o$  is a complement of  $e_*$  (Def. 5.3.1).  $\square$

Every element of a Boolean algebra is idempotent with respect to both operations.

**Theorem 5.3.6.** If  $(A, \circ, *)$  is a Boolean algebra and if  $a \in A$ , then  $a * a = a$ .

*Proof.* For:

$$\begin{aligned} a &= a * e_* && \text{(Identity)} &= (a * a) \circ e_o && \text{(Complement)} \\ &= a * (a \circ a^{-1}) && \text{(Complement)} &= a * a && \text{(Identity)} \\ &= (a * a) \circ (a * a^{-1}) && \text{(Distributivity)} \end{aligned}$$

And therefore  $a = a * a$ .  $\square$

**Theorem 5.3.7.** If  $(A, \circ, *)$  is a Boolean algebra and if  $a \in A$ , then  $a \circ a = a$ .

*Proof.* For:

$$\begin{aligned} a &= a \circ e_o && \text{(Identity)} &= (a \circ a) * e_* && \text{(Complement)} \\ &= a \circ (a * a^{-1}) && \text{(Complement)} &= a \circ a && \text{(Identity)} \\ &= (a \circ a) * (a \circ a^{-1}) && \text{(Distributivity)} \end{aligned}$$

And therefore  $a = a \circ a$ .  $\square$

**Theorem 5.3.8.** If  $(A, \circ, *)$  is a Boolean algebra, if  $a, b \in A$ , if  $a \circ b = a$ , and if  $a * b = a$ , then  $b = a$ .

*Proof.* For:

$$\begin{aligned}
 b &= b * e_* && \text{(Identity)} \\
 &= b * (a \circ a^{-1}) && \text{(Complement)} \\
 &= (b * a) \circ (b * a^{-1}) && \text{(Distributivity)} \\
 &&& \quad \text{(Distributivity)} \\
 &= a \circ (b * a^{-1}) && \text{(Hypothesis)} \\
 &= (a \circ b) * (a \circ a^{-1}) && \text{(Distributivity)} \\
 &= (a \circ b) * e_* && \text{(Complement)} \\
 &= a \circ b && \text{(Identity)} \\
 &= a && \text{(Hypothesis)}
 \end{aligned}$$

And therefore  $a = b$ . □

**Theorem 5.3.9.** If  $(A, \circ, *)$  is a Boolean algebra, if  $a \in A$  is such that  $a = a^{-1}$ , then  $a = e_\circ = e_*$ .

*Proof.* For let  $a \in A$  and let  $a = a^{-1}$ . Then by Thm. 5.3.6:

$$a = a * a = a * a^{-1} = e_\circ \quad (5.3.4)$$

Similarly,  $a = e_*$ . □

**Theorem 5.3.10.** If  $(A, \circ, *)$  is a Boolean algebra, if  $a \in A$ , and if  $b, c \in A$  are complements of  $a$ , then  $b = c$ .

*Proof.* For:

$$\begin{aligned}
 b &= b * e && \text{(Identity)} \\
 &= b * (a \circ c) && \text{(Complement)} \\
 &= (b * a) \circ (b * c) && \text{(Distributivity)} \\
 &= e_\circ \circ (b * c) && \text{(Complement)} \\
 &= (c * a) \circ (b * c) && \text{(Complement)}
 \end{aligned}$$

$$\begin{aligned}
 &= (c * a) \circ (c * b) && \text{(Commutativity)} \\
 &= c \circ (a * b) && \text{(Distributivity)} \\
 &= c \circ e_\circ && \text{(Complement)} \\
 &= c && \text{(Identity)}
 \end{aligned}$$

Therefore,  $b = c$ . □

**Theorem 5.3.11.** If  $(A, \circ, *)$  is a Boolean algebra and if  $a \in A$ , then  $(a^{-1})^{-1} = a$ .

*Proof.* For:

$$\begin{aligned}
 a &= a * e_* && \text{(Identity)} \\
 &= a * (a^{-1} \circ (a^{-1})^{-1}) && \text{(Complement)} \\
 &= (a \circ a^{-1}) * (a \circ (a^{-1})^{-1}) && \text{(Distributivity)} \\
 &= e_* * (a \circ (a^{-1})^{-1}) && \text{(Complement)} \\
 &= a \circ (a^{-1})^{-1} && \text{(Identity)}
 \end{aligned}$$

And similarly  $a * (a^{-1})^{-1} = a$ . But if  $a * b = a$  and  $a \circ b = a$ , then  $a = b$  (Thm. 5.3.8). Therefore,  $a = (a^{-1})^{-1}$ .  $\square$

### Theorem 5.3.12: Absorption Laws

If  $(A, \circ, *)$  is a Boolean algebra, if  $a \in A$  and if  $b \in A$ , then  $a * (a \circ b) = a$  and  $a \circ (a * b) = a$ .  $\blacksquare$

*Proof.* For:

$$\begin{aligned}
 a * (a \circ b) &= (a * e_*) \circ (a * b) && \text{(Identity)} \\
 &= a * (e_* \circ b) && \text{(Distributivity)} \\
 &= a * e_* && \text{(Thm. 5.3.3)} \\
 &= a && \text{(Identity)}
 \end{aligned}$$

And therefore  $a * (a \circ b) = a$ . Similarly:

$$\begin{aligned}
 a \circ (a * b) &= (a \circ e_\circ) * (a * b) && \text{(Identity)} \\
 &= a \circ (e_\circ * b) && \text{(Distributivity)} \\
 &= a \circ e_\circ && \text{(Thm. 5.3.4)} \\
 &= a && \text{(Identity)}
 \end{aligned}$$

And therefore  $a \circ (a * b) = a$ .  $\square$

We can weaken the hypothesis of Thm. 5.3.8 to obtain a more general result.

**Theorem 5.3.13.** *If  $(A, \circ, *)$  is a Boolean algebra, if  $a, b \in A$ , and if  $a * b = a \circ b$ , then  $a = b$ .*

*Proof.* For:

$$\begin{aligned}
 a &= a * e_* && \text{(Identity)} \\
 &= a * (b \circ b^{-1}) && \text{(Complement)} \\
 &= (a * b) \circ (a * b^{-1}) && \text{(Distributivity)} \\
 &= (a \circ b) \circ (a * b^{-1}) && \text{(Hypothesis)} \\
 &= ((a \circ b) \circ a) * ((a \circ b) \circ b^{-1}) && \text{(Distributivity)} \\
 &= ((a \circ b) \circ a) * (a \circ (b \circ b^{-1})) && \text{(Associativity)} \\
 &= ((a \circ b) \circ a) * (a \circ e_*) && \text{(Complement)} \\
 &= ((a \circ b) \circ a) * e_* && \text{(Thm. 5.3.3)} \\
 &= (a \circ b) \circ a && \text{(Identity)} \\
 &= (a \circ a) \circ b && \text{(Associativity and Commutativity)} \\
 &= a \circ b && \text{(Thm 5.3.7)}
 \end{aligned}$$

Thus  $a = a \circ b$ . But  $a \circ b = a * b$ , and so  $a = a * b$ . By Thm. 5.3.8,  $a = b$ .  $\square$

**Definition 5.3.3** For  $a \in S$ , an inverse, or normal inverse, of the First Operation is an element  $b \in S$  such that  $a \circ b = e_\circ$ . An inverse of the Second Operation is similarly defined. The normal inverses are denoted  $a^*$  and  $a^\circ$ .

**Theorem 5.3.14.** *If  $a \in S$  has a normal inverse for either operation, than it is unique.*

*Proof.* For suppose not. Let  $a \in S$  have a normal inverse for the First Operation. That is, there is an  $a^\circ \in S$  such that  $a \circ a^\circ = e_\circ$  and let  $a'^\circ$  be a second normal inverse not equal to the first. But then  $a^\circ = a^\circ \circ e_\circ = a^\circ \circ (a \circ a'^\circ)$  and from associativity we have  $a^\circ = (a^\circ \circ a) \circ a'^\circ = a'^\circ$ . Thus, the normal inverse is unique. Similarly if there is an inverse for the Second Operation  $\square$

**Theorem 5.3.15.** *If  $a \in S$  has a normal inverse, say  $a'$ , for one operation, then  $a^{-1} = a'^{-1}$ .*

*Proof.* For let  $a \in S$  have a normal inverse  $a'$  for the First Operation. That is,  $a \circ a' = e_\circ$ . But  $a' \circ a'^{-1} = e_*$ , and from theorem 1.3  $a \circ e_* = e_*$ . So  $a \circ (a' \circ a'^{-1}) = e_*$ . And from theorem 1.4,  $a \circ a = a$ , so we have  $(a \circ a) \circ (a' \circ a'^{-1}) = a \circ (a \circ a') \circ a'^{-1} = a \circ a'^{-1} = e_*$ . But  $a \circ a^{-1} = e_\circ$ . And pseudo-inverses are unique. Thus,  $a^{-1} = a'^{-1}$ .  $\square$

**Theorem 5.3.16.** *The identities have normal inverses for their respective operations.*

*Proof.* As normal inverses are unique, it suffices to find inverses for both identities. But  $e_\circ \circ e_\circ = e_\circ$ , so  $e_\circ$  is its own inverse for the First Operation. Similarly,  $e_* \circ e_* = e_*$ .  $\square$

**Theorem 5.3.17.** (*The Not-A-Field Theorem*) Only the identities have normal inverses.

*Proof.* For suppose not. Suppose  $a \in S$ ,  $a \neq e_o$ ,  $a \neq e_*$  and  $a$  has an inverse for the First Operation. That is  $\exists a^\circ \in S | a \circ a^\circ = e_o$ . But by theorem 1.4,  $a \circ a^\circ = (a \circ a) \circ a^\circ$ . By associativity, we have  $e_o = a \circ a^\circ = a \circ (a \circ a^\circ) = a \circ e_o = a$ . Thus,  $a = e_o$ . But by hypothesis,  $a \neq e_o$ . Thus, there is no inverse for  $a$ . Similarly,  $a$  has no inverse for the Second Operation.  $\square$

**Theorem 5.3.18.** There exist pseudo-fields with only one element.

*Proof.* For let  $e_o = e_*$ , and let no other elements be in the set.  $\square$

**Theorem 5.3.19.** A pseud-field has one element if and only if  $e_o = e_*$ .

*Proof.* For suppose there is another element  $a \neq e_o$ . But then  $a \circ e_o = a$ , but also  $a \circ e_o = a \circ e_* = e_*$ . So  $a = e_*$ . If there is only one element, then clearly  $e_o = e_*$  as otherwise there would be two elements.  $\square$

**Definition 5.3.4** A generating set on a pseudo-field is a subset  $g_S \subset S$  such that every element of  $S$  can be written as a finite combination of elements in  $g_S$  using  $\circ$  or  $*$ .

**Theorem 5.3.20.** The number of elements in a finite pseudo-field is a power of 2.

*Proof.* Consider the set of all generators  $g_S$  on  $S$ . Clearly for all such generators,  $1 \leq |g_S| \leq |S|$ . Let  $G$  be the smallest generator, such that  $|G| \leq |g_S|$  for any other given generator.  $\square$

## 5.4 Sequences and Matrices

Matrices are the fundamental object studied in linear algebra, and are used in the study of general algebra as well. To discuss the more interesting properties requires some notion of arithmetic that we do not yet posses. In particular, matrices are most interesting when there is an underlying *ring* structure. For now we simply introduce the set theoretic definition of a matrix, relate this to the familiar *grid of numbers* definition, and provide examples.

### Definition 5.4.1: Matrix

An  $n \times m$ ,  $n, m \in \mathbb{N}$ , matrix on a set  $X$  is a function  $A : \mathbb{Z}_n \times \mathbb{Z}_m \rightarrow X$ .

## CHAPTER 6

---

### Arithmetic

---



## CHAPTER 7

---

### Cardinality

---



**Part II**

# **Lattice Theory**



## CHAPTER 8

---

### Relations of Order

---



## CHAPTER 9

---

### Graph Theory

---



## **Part III**

# **Categories and Models**



## CHAPTER 10

---

### Category Theory

---

Book Two

Algebra





**Part IV**

**Group Theory**



# CHAPTER 11

---

## Elementary Group Theory

---

One of the most fundamental structures studied in mathematics is a [group](#). Recalling some notions from Book One we define a group to be a [set](#)  $G$  together with a [binary operation](#)  $*$  that is [associative](#), has an [identity](#), and such that all elements are [invertible](#).<sup>1</sup> Groups then seem to be simple objects, and indeed the standard arithmetic that one is familiar with in  $\mathbb{R}$  has far more structure.<sup>2</sup> In another sense perhaps groups have too much structure. We can certainly strip away invertibility, leaving associativity and identity intact, and this results in a [monoid](#). We can further rid of *the* identity, maintaining only associativity, and it may be reasonable that such objects are useful. Furthermore we can take away the requirement that we have a binary operation and replace this with a partial function. This latter object gives rise to the notion of a [groupoid](#), which has applications in geometry and analysis. We'll start with ordinary binary operations, but only require associativity. These are called [semigroups](#).

### 11.1 Group-Like Structures

One motivating reason for studying semigroups is that many foundational theorems of groups do not require all of the structure they are endowed with. Indeed almost nothing was needed to prove the identity is unique (Thm. 5.2.4). Given an element  $e$  such that  $a*e = e*a = a$  for all  $a \in G$ , if  $e'$  is also an identity, then  $e = e * e' = e'$ . Requiring associativity and the existence of inverses is superfluous. However, associativity is important for otherwise there is an ambiguity in how to combine three elements. Associativity is not the only resolution to this problem but is common and appears in our familiar arithmetic.<sup>3</sup>

---

<sup>1</sup> All of these notions are developed in Chapt. 5.

<sup>2</sup>  $\mathbb{R}$  forms a [field](#).

<sup>3</sup> The Jacobi identity is another means. It appears in the *cross product* of vectors.

### 11.1.1 Semigroups and Monoids

Let's briefly recall some vocabulary from Chapt. 5. Given a set  $G$  with a binary operation  $*$ , a right unital element is an  $e_R \in G$  such that  $a * e_R = a$  for all  $a \in G$ . Left unitals are similarly defined,  $e_L * a = a$ . Unital elements are those which are both left and right unital elements. Left and right identities need not be unique, but two-sided ones do. If there exists both left and right unital elements  $e_L$  and  $e_R$ , then they coincide  $e_L = e_R$  giving us an identity (Thm. 5.2.2). A right invertible element is an  $a \in G$  such that there is a  $b \in G$  with  $a * b$  a unital element. Left invertible is similarly defined,  $b * a$  being a unital element, and invertible elements are those that are both left and right invertible. If the binary operation is associative, then inverses are unique and we may write  $a^{-1}$ . We relax these conditions to *weakly* left and right invertible, requiring for  $a \in G$  that there is a  $b \in G$  with  $b * a$  equal to a left unital element or  $a * b$  equal to a right unital element, respectively. This is *usually* a strictly weaker notion. Without further ado we give the definition of a semigroup.

#### Definition 11.1.1: Semigroup

A **semigroup** is an **ordered pair**  $(G, *)$  where  $G$  is a **set** and  $*$  is an **associative operation**  $*$  on  $G$ . That is, for all  $a, b, c \in G$  we have:

$$a * (b * c) = (a * b) * c \quad (1)$$

Associativity was formally defined in Chapt. 5 Def. 5.2.2. This states there is no ambiguity in combining three elements. Indeed, continuing inductively, there is no ambiguity in combining an ordered list of  $n$  elements for any  $n \in \mathbb{N}$  with  $n \geq 2$ . It is important to note we may not be able to swap the order of our operation. That is, for distinct  $a, b \in G$  it may not be true that  $a * b = b * a$ . The binary operation  $*$  is not required to be **commutative**.

Defining semigroups as **ordered pairs** allows us to distinguish a semigroup from the underlying set. Recalling Kuratowski's definition (Def. 3.1.5) we have:

$$(G, *) = \{ \{G\}, \{G, *\} \} \quad (11.1.1)$$

By regularity, since  $\{G\} \in (G, *)$  and  $G \notin G$  (Thm. 3.1.11), we have a set-theoretic way of distinguishing between the semigroup structure on  $G$  and  $G$  itself. In category theory one speaks of the *forgetful functor*. The language is often presented very loosely: We map the semigroup  $G$  to the underlying set  $G$ . That is, we *forget* about the binary operation  $*$ . A more satisfying way of presenting this is to express  $\mathcal{F} : \mathbf{Semigroups} \rightarrow \mathbf{Sets}$  by:

$$\mathcal{F}((G, *)) = G \quad (11.1.2)$$

We did not adopt any axioms about proper classes in Book One and hence such things are of little concern to us. Without further delay we present examples.

**Example 11.1.1** If  $G = \emptyset$  and  $*$  is the empty function, then  $(G, *)$  is the *empty semigroup*. This is vacuously true since there are no elements that do not satisfy Def. 11.1.1, and thus  $(\emptyset, \emptyset)$  is a semigroup.

**Example 11.1.2** Let  $G = \mathbb{Z}_1 = \{0\}$ , and let  $*$  be the only possible binary operation on  $G$ :  $0 * 0 = 0$ . Then  $(G, *)$  is a semigroup trivially. Since there is only one element, we compute:

$$0 * (0 * 0) = 0 * 0 = (0 * 0) * 0 \quad (11.1.3)$$

Thus  $(G, *)$  is a set with an associative binary operation, making it a semigroup (Def. 11.1.1).  $(G, *)$  has additional structure: there is a unital element 0 and every element is invertible.<sup>4</sup> This is the structure of a group.

**Example 11.1.3** The defining example of a semigroup is the positive integers  $\mathbb{N}^+$  with the usual notion of addition  $+$ . More precisely, we take the standard additive binary operation  $+: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  and restrict it to  $\mathbb{N}^+$  (but rather than writing  $+|_{\mathbb{N}^+}$  we stick to  $+$ ). This makes  $(\mathbb{N}^+, +)$  a semigroup. Since the restriction of  $+$  to  $\mathbb{N}^+$  is closed<sup>5</sup> we have that  $+$  is a binary operation on  $\mathbb{N}^+$ . And since the restriction of an associative operation is again associative,  $+$  is associative on  $\mathbb{N}^+$ , and hence the criteria of Def. 11.1.1 are satisfied.

Note the semigroup  $(\mathbb{N}^+, +)$  does not have any of the additional structure mentioned in the forward to this chapter. There is no unital element, nor are there inverse elements for any  $n \in \mathbb{N}^+$ . To see this, note  $(\mathbb{Z}, +)$  has a unital element 0 and for all  $n \in \mathbb{Z}$  there is an inverse  $-n$ . Since unital elements are unique (Thm. 5.2.4) and since inverses are unique for associative operations (Thm. 5.2.25), we have that 0 is the unique unital element and  $-n$  is the unique inverse of  $n \in \mathbb{Z}$ . But  $0 \notin \mathbb{N}^+$  and for all  $n \in \mathbb{N}^+$  we know that  $-n \notin \mathbb{N}^+$ . Suppose  $(\mathbb{N}^+, +)$  has a unital element  $e$ . That is, for all  $n \in \mathbb{N}^+$  we have  $e + n = n$ . We must show  $e = 0$ , contradicting our claim that  $e \in \mathbb{N}^+$ . But then, since  $\mathbb{N}^+ \subseteq \mathbb{Z}$ , for all  $n \in \mathbb{N}^+$  there is a  $-n \in \mathbb{Z}$  such that  $n + (-n) = 0$ . Applying associativity, we obtain:

$$0 = n + (-n) = (e + n) + (-n) = e + (n + (-n)) = e + 0 = e \quad (11.1.4)$$

So  $e = 0$ , a contradiction, and thus  $\mathbb{N}^+$  has no unital element. Similarly, there are no invertible elements. From this we see that  $(\mathbb{N}^+, +)$  is purely a semigroup, and is neither a monoid nor a group.<sup>6</sup> This is what was meant by the claim that  $(\mathbb{N}^+, +)$  is the defining example of a semigroup.<sup>7</sup>

<sup>4</sup> There is only one element, and unital elements are invertible (Thm. 5.2.22).

<sup>5</sup> That is, the sum of positive integers is again a positive integer.

<sup>6</sup> See the start of this chapter for a rough definition or Defs. 11.1.2 and 11.1.3, respectively.

<sup>7</sup> It's slightly more, since it is a *commutative* semigroup.

In proving  $\mathbb{N}^+$  has no unital element we used the fact that it embeds naturally as a subset of a group, the integers  $(\mathbb{Z}, +)$ , and the argument presented required the use of invertible elements. If we have a semigroup  $(G, *)$  with a unital element  $e \in G$ , it is possible that  $G \setminus \{e\}$  has a *new* unital element  $e'$ . Since identities are unique,  $e'$  is not a unital element in  $G$ , but rather a unital element in  $G \setminus \{e\}$ . We'll see examples once Thm. 11.1.1 has been proven.

**Example 11.1.4** We can put another associative binary operation on  $\mathbb{N}^+$ , and may define  $*$  by  $m * n = n + m + nm$ . Then  $(\mathbb{N}, *)$  is a semigroup since  $*$  is associative (Def. 11.1.1). To see this, we compute:

$$a * (b * c) = a * (b + c + bc) \quad (11.1.5)$$

$$= a + (b + c + bc) + a(b + c + bc) \quad (11.1.6)$$

$$= a + b + c + bc + ab + ac + abc \quad (11.1.7)$$

$$= a + b + ab + c + (a + b + ab)c \quad (11.1.8)$$

$$= (a + b + ab) * c \quad (11.1.9)$$

$$= (a * b) * c \quad (11.1.10)$$

and hence  $*$  is indeed associative (Def. 5.2.2).

**Example 11.1.5** Given a set  $A$ ,  $(\mathcal{P}(A), \cup)$  is a semigroup since union is a well defined binary operation and moreover is associative (Thm. 3.2.30). Similarly,  $(\mathcal{P}(A), \cap)$  is a semigroup since intersection is associative (Thm. 3.2.49).

Semigroups can be truly reckless. We demonstrate this with an example.

**Example 11.1.6** Let  $X$  be any non-empty set and define  $*$  by  $a * b = b$ . That is, we simply choose the right element. Then  $*$  is associative since:

$$a * (b * c) = a * c = c \quad (11.1.11a) \qquad (a * b) * c = b * c = c \quad (11.1.11b)$$

In this structure every element is weakly left invertible and also a left unital element but there are no right identities or weakly right invertible elements.

**Example 11.1.7** Similar to Ex. 11.1.6 we can define  $a * b = a$  which again gives us a semigroup. Here every element is weakly right invertible and everything is a right identity. That is, if  $a \in G$ , let  $b \in G$  be arbitrary. Then  $a * b$  is a right unital element since every element is right unital. Hence,  $a$  is weakly right invertible and  $b$  is a weak right inverse. There are no left unital or weakly left invertible elements.

There is solace in the fact that any semigroup can be extended to a monoid by adding a single element. That is, any semigroup can be given an identity. This is done in Thm. 11.1.1. First, we formally define monoids.

**Definition 11.1.2:** Monoid

A **monoid** is a **semigroup**  $(G, *)$  such that there exists a **unital element**  $e \in G$ . That is, there is an  $e \in G$  and for all  $a, b, c \in G$  we have:

$$a * (b * c) = a * (b * c) \quad (1) \qquad a * e = a \quad e * a = a \quad (2)$$

From the uniqueness of unital elements with respect to any binary operation (Thm. 5.2.4), the identity in a monoid  $(G, *)$  is unique.

**Example 11.1.8** Unlike semigroups, there is no *empty* monoid since by definition we assume the existence of a unital element. Hence  $G$  must be nonempty.

**Example 11.1.9** Any monoid is necessarily a semigroup, but the converse need not hold. Ex. 11.1.3 showed that the positive integers with the usual notion of addition  $+$  form a pure semigroup that is neither a monoid nor a group. That is, since  $(\mathbb{N}^+, +)$  does not have a unital element it is not a monoid.

**Example 11.1.10** The set  $\mathbb{Z}_1$  with the operation  $*$  defined by  $0 * 0 = 0$  is a monoid. As stated before it is also a group.

**Example 11.1.11** The quintessential example of a monoid is the natural numbers  $\mathbb{N}$  with the usual additive binary operation  $+$ . This makes  $(\mathbb{N}, +)$  a monoid with 0 acting as the identity. Much like the semigroup  $(\mathbb{N}^+, +)$ , the monoid  $(\mathbb{N}, +)$  lacks inverses and is not a group, but rather a pure monoid.<sup>8</sup>

We now prove any semigroup can be extended to a monoid by adding one point.

**Theorem 11.1.1: Embedding Theorem for Semigroups**

If  $(G, *)$  is a semigroup, then there is a monoid  $(\tilde{G}, \times)$  such that  $G \subseteq \tilde{G}$  and such that  $\times|_G = *$ . That is, the restriction of  $\times$  to  $G$  is  $*$ . ■

*Proof.* If  $G$  is a set, then  $\{G\}$  is a set (Thm. 3.1.4). But  $G \neq \{G\}$  (Thm. 3.1.13) and  $\{G\} \notin G$  (Thm. 3.2.10). Let  $\tilde{G} = G \cup \{G\}$ . Then  $G \subseteq \tilde{G}$  (Thm. 3.2.15). Define  $\times$  as follows:

$$a \times b = \begin{cases} a * b, & a \in G, b \in G \\ a, & a \in G, b = G \\ b, & a = G, b \in G \\ G, & a = G, b = G \end{cases} \quad (11.1.12)$$

Then since  $\{G\} \notin G$ , this is well defined, and by definition  $\times|_G = *$ . But  $\times$  is an associative operation with a unital element, since by definition  $G$  is the

---

<sup>8</sup> It also has the property that addition is commutative making it a **commutative monoid**.

unital element of  $\tilde{G}$ . It is associative since there are only a few cases to check: If  $a, b, c \in G$ , then since  $a * b \in G$  and  $b * c \in G$ , and since  $G \not\subseteq \tilde{G}$  we have:

$$a \times (b \times c) = a \times (b * c) = a * (b * c) = (a * b) * c = (a * b) \times c = (a \times b) \times c \quad (11.1.13)$$

If  $a, b \in G$  and  $c = G$ , we have:

$$a \times (b \times G) = a \times b = (a \times b) \times G \quad (11.1.14)$$

And similarly if  $a = G$  or  $b = G$ . Hence,  $(\tilde{G}, \times)$  is a monoid (Def. 11.1.2).  $\square$

Let's see what happens if we perform the construction in Thm. 11.1.1 on a semigroup that is already a monoid. Taking  $\mathbb{Z}$  as our monoid (It's actually a group), we pick up a new element  $\{\mathbb{Z}\}$  which we may as well label  $\infty$  with the strange operation:

$$n + \infty = n \quad n \in \mathbb{Z} \quad (11.1.15)$$

This new monoid contains  $\mathbb{Z}$  embedded inside, although  $\mathbb{Z}$  already contains an identity, namely zero. So what is the identity for our appended space  $\tilde{\mathbb{Z}}$ ? It's  $\infty$  since we have  $0 + \infty = 0$  and hence 0 is no longer an identity. This leads to the paradoxical conclusion: There are monoids  $(G, *)$  with proper subsets  $A \subsetneq G$  that do **not** contain the identity element, but such that  $(A, *)$  is still a monoid (contains an identity).

In this example we see that 0 still really wants to be an identity, and the only thing stopping it from being so is  $\infty$ . Removing this element allows zero to be what its always wanted to be: The additive identity. This problem is resolved by groups. While monoids are structurally simpler than groups, the notion of a group is one of the most basic and fundamental algebraic structures that one can consider. The existence of inverse elements removes this unwanted phenomenon and if semigroup  $(G, *)$  embeds into a group  $(\tilde{G}, \times)$ , then either the unital element in  $G$  is the same as the one in  $\tilde{G}$ , or  $G$  has no identity.

### Definition 11.1.3: Group

A group is a monoid  $(G, *)$  such that for all  $g \in G$ ,  $g$  is an invertible element with respect to  $*$ . That is:

The binary operation  $*$  is associative (1)

There exists a unital element  $e \in G$  (2)

For every element  $a \in A$  there is an inverse element  $a^{-1}$  (3)

Note that it is not necessarily true that  $a * b = b * a$ . Such groups are called Abelian. There are several examples of groups that one is likely familiar with.

**Example 11.1.12** If  $+$  denotes the usual addition on  $\mathbb{Z}$ , then  $(\mathbb{Z}, +)$  is a group. The unital element is 0, and for all  $n \in \mathbb{Z}$ ,  $-n$  is an inverse element of  $n$ . That is,  $n + (-n) = 0$ , which is a unital element. Moreover, addition is associative and therefore  $(\mathbb{Z}, +)$  is a group.

**Example 11.1.13** If  $\mathbb{R}^+$  denotes the positive real numbers, and if  $\cdot$  denotes the usual multiplication of real numbers, then  $(\mathbb{R}, \cdot)$  is a group. Here 1 is the unital element and for all  $r \in \mathbb{R}^+$   $1/r$  is the inverse element. Lastly, the operation is indeed associative.

**Example 11.1.14** Consider  $\mathbb{Z}_n$  with modular addition  $+$ . Then  $(\mathbb{Z}_n, +)$  is a group. The unital element is 0 since for all  $k \in \mathbb{Z}_n$ ,  $k + 0 = k$ . The inverse of  $k \in \mathbb{Z}_n$  is the unique integer  $m \in \mathbb{Z}_n$  such that  $k + m = n$ . That is,  $m = n - k$ . For then  $k + m = n$ , and  $n$  is equivalent to zero in  $\mathbb{Z}_n$ . Since modular addition is commutative, it follows that 0 is a left and right unital element and that each element is left and right invertible, and thus  $(\mathbb{Z}_n, +)$  is a group.

Wishing to be as lazy as possible, we present the following.

**Theorem 11.1.2.** *If  $G$  is a set, if  $*$  is an associative binary operation on  $G$ , if  $e \in G$  is a unique right unital element, and if for all  $g \in G$  it is true that  $g$  is weakly right invertible, then  $G$  is a group.*

*Proof.* For by Thm. 5.2.18,  $e$  is a unital element. By Thm. 5.2.24, for all  $g \in G$  it is true that  $g$  is invertible. Therefore,  $(G, *)$  is a group (Def. 11.1.3).  $\square$

One should consult the proofs of the referenced theorems in Chapt. 5 if there is doubt. This eases the task of proving something is a group, we need only check that there is a right identity  $g * e_R = g$  and for all  $a \in G$  there is a  $b \in G$  with  $g * (a * b) = g$ . In other words,  $a * b = e_R$ .

Some consequences of Chapt. 5. The identity of a group is unique (Thm. 5.2.4), as are the inverses of elements (Thm. 5.2.25). Moreover, the inverse of a unital element is itself (Thm. 5.2.22), and the inverse of a product is the product of inverses *with the order swapped*:  $(a * b)^{-1} = b^{-1} * a^{-1}$  (Thm. 5.2.28). Lastly, inverting an inverse gives you the original element:  $(a^{-1})^{-1} = a$ .

One of the ways of representing semigroups (and groups and monoids) is via their *Cayley Table*, named after the English mathematician Arthur Cayley (1821-1895 C.E.) who was one of the early pioneers of group theory. In 1854 he published *On the Theory of Groups* which contains the first occurrence of these tables. The concept is quite simple, given a semigroup with  $n$  elements we present an  $n \times n$  table and fill in the  $(i, j)$  spot by multiplying the  $i^{th}$  element with the  $j^{th}$  element, left to right. Of course, we need to put some order on the elements, but this is immaterial.

**Example 11.1.15** Take  $\mathbb{Z}_2 = \{0, 1\}$  and bequeath it with modulo arithmetic:  $0 + 0 = 0$ ,  $0 + 1 = 1$ ,  $1 + 0 = 1$ ,  $1 + 1 = 0$ . This produces the structure of a group. We use this to compute the Cayley table (Tab. 11.1)

+	0	1
0	0	1
1	1	0

Table 11.1: Cayley Table for  $\mathbb{Z}_2$

We could equivalently have chosen 1 in the left column and 0 in the right, the order being irrelevant. All of the information about a group or monoid or semigroup can be encoded in its Cayley table.<sup>9</sup>

**Example 11.1.16** Consider  $\mathbb{Z}_4$  with its modular addition  $+$ . Then  $(\mathbb{Z}_4, +)$  is a group. We can represent the operation  $+$  with the following table:

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Table 11.2: The Group Structure of  $\mathbb{Z}_4$

**Example 11.1.17** Rotations about a point  $(x, y) \in \mathbb{R}^2$  defined by:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (11.1.16)$$

form a group under rotation by an angle  $\theta$ .  $\theta = 0$  is the identity, and given a rotation  $\theta$ , the rotation  $-\theta$  in the reverse direction serves as inverse.

**Example 11.1.18** Let  $X$  denote the set of all  $\mathbf{x} \in \mathbb{R}^2$  such that  $x_0 \neq 0$ . That is, the Cartesian plane with the  $x$  axis removed. Define  $*$  on  $X$  by:

$$(a, b) * (c, d) = (ac, bc + d) \quad (11.1.17)$$

This is a group. The element  $(1, 0)$  serves as identity since:

$$(1, 0) * (c, d) = (1 \cdot c, 0 \cdot c + d) = (c, d) \quad (11.1.18)$$

$$(a, b) * (1, 0) = (a \cdot 1, b \cdot 1 + 0) = (a, b) \quad (11.1.19)$$

---

<sup>9</sup> When the underlying set is infinite this is no longer useful.

Lastly, given  $(a, b)$ , the inverse is  $(1/a, -b/a)$  since:

$$(a, b) * \left(\frac{1}{a}, -\frac{b}{a}\right) = \left(a \cdot \frac{1}{a}, \frac{b}{a} - \frac{b}{a}\right) = (1, 0) \quad (11.1.20)$$

and similarly for left multiplication. Hence, this is a group.

With Cayley tables presented, it is convenient to briefly discuss Latin squares. These are studied in combinatorics but have their uses in groups theory. A Latin square on a collection of  $n$  letters is an  $n \times n$  grid of these letters such that each symbol appears exactly once in each row and each column.

A	B	C
B	C	A
C	A	B

Table 11.3: Example of a Latin Square

We can write this out algebraically. Consider a group with its Cayley table. Fix some row element  $a \in G$  and look at  $a * x$  for all  $x \in G$ . The Cayley table is then a Latin square if this is bijective for all  $a \in G$ , and if  $x * b$  is bijective for all  $b \in G$ . We can write this more succinctly, requiring for all  $a, b \in G$  there is a unique  $x \in G$  such that  $a * x = b$ . This property has been well studied and given a name.

#### Definition 11.1.4: Quasigroup

A quasigroup is an ordered pair  $(G, *)$  where  $G$  is a set and  $*$  is a binary operation on  $G$  where for all  $a, b \in G$  there exists unique  $x, y \in G$  with:

$$a * x = b \quad (1) \qquad y * a = b \quad (2)$$

Note associativity is **not** assumed in the definition. The condition being imposed is known as the *Latin square property* and allows one to define division unambiguously. This is quite clever since no unital elements are assumed to exist. In most settings we define division by inverse elements. That is, given  $a$  we define  $1/a = a^{-1}$  where  $a^{-1}$  is the inverse of  $a$ . Such a definition presumes the existence of an identity since then  $a * a^{-1}$  is a unital element by definition. Quasigroups allow for division without identities since we require the solution to  $a * x = b$  to be unique allowing us to define  $a/b = x$ .

**Example 11.1.19** The integers with subtraction serve as the modeling example of quasigroups. Subtraction satisfies the Latin square property since if  $a - b = a - c$ , then  $b = c$ , and similarly if  $b - a = c - a$ . There are no other features associated with this structure. Subtraction is *not* associative since

$a - (b - c) = a - b + c \neq (a - b) - c$  and it is also not commutative since  $a - b = b - a$  if and only if  $a = b$ . There is also no unital element. Zero is a right unital, but not a left. That is,  $a - 0 = a$  but  $0 - a = -a$

**Example 11.1.20** The non-zero rationals  $\mathbb{Q} \setminus \{0\}$  form a quasigroup with division in a similar manner to subtraction with the integers. This is also non-commutative, non-associative, and lacks a left identity.

**Example 11.1.21** In group theory one speaks of the *quaternion group*, which is a group consisting of 8 elements that has found its way into geometry, topology, and physics. In quasigroup theory we have the *hyperbolic quaternion quasigroup* which similarly has applications in physics. Again we have a set with 8 elements  $\{\pm 1, \pm i, \pm j, \pm k\}$  and define  $i^2 = j^2 = k^2 = 1$ . We then fill out the rest of the Cayley table the way one does the quaternions.<sup>10</sup>

.	1	$i$	$j$	$k$	-1	$-i$	$-j$	$-k$
1	1	$i$	$j$	$k$	-1	$-i$	$-j$	$-k$
$i$	$i$	1	$k$	$-j$	$-i$	-1	$-k$	$j$
$j$	$j$	$-k$	1	$i$	$-j$	$k$	-1	$-i$
$k$	$k$	$j$	$-i$	1	$-k$	$-j$	$i$	-1
-1	-1	$-i$	$-j$	$-k$	1	$i$	$j$	$k$
$-i$	$-i$	-1	$-k$	$j$	$i$	1	$k$	$-j$
$-j$	$-j$	$k$	-1	$-i$	$j$	$-k$	1	$i$
$-k$	$-k$	$-j$	$i$	-1	$k$	$j$	$-i$	1

Table 11.4: Cayley Table for the Hyperbolic Quaternion Quasigroup

There is an equivalent condition one can impose that gives rise to this structure. A binary operation is called cancellative if  $a * b = a * c$  can be simplified to  $b = c$ . Hence addition in  $\mathbb{Z}$  is cancellative but multiplication is not since  $0 \cdot 2 = 0 \cdot 3$  but  $2 \neq 3$ . Quasigroups necessarily have a cancellative binary operation.

**Theorem 11.1.3.** *If  $(G, *)$  is a quasigroup, then  $*$  is a cancellative binary operation. That is, for all  $a, b, c \in G$ , if  $a * b = a * c$ , then  $b = c$ , and if  $b * a = c * a$ , then  $b = c$ .*

*Proof.* For suppose not. Then there exists  $a, b, c \in G$  such that  $a * b = a * c$ , but  $b \neq c$ . Let  $y = a * b$ . But  $a * b = a * c$ , and hence by the transitivity of equality  $y = a * c$ . Since  $a \in G$  and  $y \in G$ , and since  $(G, *)$  is a quasigroup, there is a unique  $x \in G$  with  $a * x = y$  (Def. 11.1.4 Eqn. 1). But  $a * b = y$  and hence  $b = x$ . But also  $a * c = y$  and hence  $c = x$ . By the transitivity of equality,  $b = c$ . Similarly, if  $b * a = c * a$  apply Def. 11.1.4 Eqn. 2.  $\square$

---

<sup>10</sup> We will return to the quaternion group and hyperbolic quaternion quasigroup later.

The converse, that cancellative binary operations give rise to quasigroups, holds for finite sets. This is not a deep theorem but the method of proof is vital. We will mimic this with Cayley's theorem.

**Theorem 11.1.4.** *If  $G$  is a finite set and if  $*$  is a cancellative binary operation on  $G$ , then  $(G, *)$  is a quasigroup.*

*Proof.* For suppose not. Then there exists  $a, b \in G$  such that there is not a unique  $x \in G$  with  $a * x = b$  or a unique  $y \in G$  such that  $y * a = b$ . Suppose the former and let  $f : G \rightarrow G$  be defined by  $f(x) = a * x$ . Then  $f$  is injective. For if not, then there exists  $x_1, x_2 \in G$  such that  $f(x_1) = f(x_2)$  but  $x_1 \neq x_2$  (Def. 5.1.3). But  $*$  is cancellative, and hence if  $a * x_1 = a * x_2$ , then  $x_1 = x_2$ , a contradiction. But if  $f$  is an injective function from  $G$  to itself, and if  $G$  is finite, then  $f$  is surjective. Hence there is an  $x \in G$  with  $f(x) = b$ . But then  $a * x = b$ , a contradiction.  $\square$

**Example 11.1.22** We cannot extend Thm. 11.1.4 to the infinite case. Let  $G = \mathbb{Z} \setminus \{0\}$  and define  $a * b = 2ab$ . If  $a * b = a * c$ , then  $2ab = 2ac$  and since  $a \neq 0$  we conclude that  $b = c$ . This is not a quasigroup, since  $2, 3 \in G$  but there is no element  $x \in G$  with  $2 * x = 3$  since  $2 * x = 4x$  which must be even.

**Example 11.1.23** Quasigroups are not quite *non-associative groups* since there not be an identity. Moreover, even if  $(G, *)$  has a unital element and inverses, without associativity this may not be a quasigroup. Take the three element set  $\{a, b, e\}$  and let  $e$  be the identity. Define  $a * a = a$  and  $b * b = b$ . Furthermore, let  $a$  and  $b$  be inverses of each other so  $a * b = b * a = e$ . This operation is not a group since it is non-associative:

$$a * (a * b) = a * e = a \quad (11.1.21a) \quad (a * a) * b = a * b = e \quad (11.1.21b)$$

It is also not a quasigroup since it lacks the uniqueness requirement. We have  $a * a = a * e$  but  $a \neq e$ .

Quasigroups are allowed to be empty since there is no hypothesized element to exist. It may seem, given the bizarre examples, that quasigroups are quite removed from groups, but this is not so. If  $(G, *)$  is a non-empty quasigroup and if  $*$  is associative, then the structure is actually a group. That is, unital elements and inverses come for free.

**Theorem 11.1.5.** *If  $(G, *)$  is a quasigroup, if  $G$  is non-empty, and if  $*$  is associative, then there is a unital element  $e \in G$ .*

*Proof.* For if  $G$  is non-empty, then there is an  $a \in G$  (Def. 3.1.1). But  $(G, *)$  is a quasigroup and hence if  $a \in G$  then there is an  $e_R \in G$  such that  $a * e_R = a$  (Def. 11.1.4 Eqn. 1). Suppose  $e_R$  is not a right identity. Then there is a  $b \in G$  such that  $b * e_R \neq b$  (Def. 5.2.5). But  $(G, *)$  is a quasigroup, and hence there

is a unique  $y \in G$  such that  $y * a = b$  (Def. 11.1.4 Eqn. 2). But by hypothesis  $*$  is associative, and hence:

$$b * e_R = (y * a) * e_R = y * (a * e_R) = y * a = b \quad (11.1.22)$$

a contradiction. Hence,  $e_R$  is a right unital element. Similarly there exists a left unital element  $e_L$ . But if there exists a left unital element  $e_L$  and a right unital element  $e_R$ , then  $e_L = e_R$  (Def. 5.2.2). But then  $e_L$  is both a left and right unital element, and is therefore a unital element (Def. 5.2.6).  $\square$

**Theorem 11.1.6.** *If  $(G, *)$  is a quasigroup, if  $G$  is non-empty, if  $*$  is associative, and if  $a \in G$ , then  $a$  is invertible.*

*Proof.* For if  $G$  is non-empty and  $*$  is associative, then there is a unital element  $e \in G$  (Thm. 11.1.5). But if  $a \in G$ , then since  $G$  is a quasigroup there is a unique  $x \in G$  such that  $a * x = e$  (Def. 11.1.4 Eqn. 1) and similarly a unique  $y \in G$  such that  $y * a = e$  (Def. 11.1.4 Eqn. 2). But then  $a$  is both left and right invertible (Defs. 5.2.10, 5.2.8), and is therefore invertible (Def. 5.2.11).  $\square$

**Theorem 11.1.7.** *If  $(G, *)$  is a semigroup that is also a quasigroup, and if  $G$  is non-empty, then  $(G, *)$  is a group.*

*Proof.* For if  $(G, *)$  is a semigroup, then  $*$  is associative (Def. 11.1.1). Then by Thm. 11.1.5 there exists an identity, and by Thm. 11.1.6 there exist inverses. Thus,  $(G, *)$  is a group (Def. 11.1.3).  $\square$

**Theorem 11.1.8.** *If  $(G, *)$  is a cancellative semigroup, and if  $*$  is associative, then  $(G, *)$  is a group.*

*Proof.* For if  $G$  is finite and  $*$  is cancellative, then  $(G, *)$  is a quasigroup. But if  $G$  is non-empty and  $(G, *)$  is both a semigroup and a quasigroup, then it is a group.  $\square$

## 11.1.2 Subsemigroups

Now that we've shown that semigroups can be embedded into monoids, we return to the study of pure semigroups and see if there are any gems to be found. We introduce the notion of subsemigroup, a generalization of *subgroup* (see Def. 11.2.6). First we need to develop the idea of the product of subsets of a semigroup.

**Definition 11.1.5: Semigroup Set Product**

The product of a subset  $A \subseteq G$  against a subset  $B \in G$  in a semigroup  $(G, *)$  is the set  $A * B$  defined by:

$$A * B = \{ x \in G \mid \text{There exists } a \in A \text{ and } b \in B \text{ such that } x = a * b \}$$

We can write the semigroup product of  $A, B \subseteq G$  more suggestively as follows:

$$A * B = \{ a * b \mid a \in A \text{ and } b \in B \} \quad (11.1.23)$$

The definition we've used is simply to stay consistent with our notation used for the axiom schema of separation (Ax. 3.1.3) which we need to prove that  $A * B$  exists in the first place.

We've now doubly used the symbol  $*$  which is perhaps poor notation, but hopefully will not cause confusion. If  $A$  and  $B$  are *subsets* of  $G$ , then  $A * B$  is another subset of  $G$ . If  $a$  and  $b$  are elements of  $G$ , then  $a * b$  is again an element of  $G$ . Therein lies the distinction between the notations.

**Example 11.1.24** Let  $(\mathbb{N}^+, +)$  be the usual additive semigroup on the positive integers, and let  $A = \{1\}$  and  $B = \mathbb{N}_o$ , the set of all positive odd integers. The set  $AB_{\mathbb{N}^+}$  is then the set:

$$A + B = \{ a + b \mid a \in A \text{ and } b \in B \} \quad (11.1.24)$$

But  $A = \{1\}$  and  $B = \mathbb{N}_o$ , the set of all odd positive integers, and so for all  $b \in B$  we can write  $b = 2n + 1$  for some  $n \in \mathbb{N}$  (possibly  $n = 0$ ). But then:

$$A + B = \{ 1 + (2n + 1) \mid n \in \mathbb{N} \} \quad (11.1.25a)$$

$$= \{ 2n + 2 \mid n \in \mathbb{N} \} \quad (11.1.25b)$$

$$= \{ 2(n + 1) \mid n \in \mathbb{N} \} \quad (11.1.25c)$$

$$= \{ 2n \mid n \in \mathbb{N}^+ \} \quad (11.1.25d)$$

This is precisely the even positive natural integers. That is, we have found that:

$$A + B = \mathbb{N}_e \setminus \{0\} \quad (11.1.26)$$

**Example 11.1.25** Let  $(\mathbb{N}, \cdot)$  be the semigroup of the natural numbers with the usual notion of multiplication. If we let  $A = \{2\}$  and  $B = \mathbb{N}$ , then we obtain:

$$A \cdot B = \{ 2 \cdot n \mid n \in \mathbb{N}^+ \} \mathbb{N}_e \setminus \{0\} \quad (11.1.27)$$

same as the previous example. It is occasionally convenient to simply write  $2\mathbb{N}$ , much the way we write  $n\mathbb{Z}$  to denote the multiples of  $\mathbb{Z}$  by  $n$ .

To define a subsemigroup of a semigroup  $(G, *)$  we want a subset  $A \subseteq G$  that is closed under the binary operation  $*$ . That is, for all  $a, b \in A$  we want it to be true that  $a * b \in A$ . We can phrase this in terms of the semigroup product by requiring that  $AA_G \subseteq A$ . This need not always be the case, one need only consider  $(\mathbb{N}^+, +)$  with  $A = \{1\}$ . Then  $AA = \{2\}$ , and thus  $A$  is not closed under  $+$ . With this, we now define subsemigroups.

#### Definition 11.1.6: Subsemigroup

A subsemigroup of a semigroup  $(G, *)$  is a subset  $A \subseteq G$  such that for all  $a, b \in A$  it is true that  $a * b \in A$

**Theorem 11.1.9.** *If  $(G, *)$  is a semigroup, then  $A \subseteq G$  is a subsemigroup if and only if  $AA_G \subseteq A$ , where  $AA_G$  is the semigroup product of  $A$  with itself in  $G$ .*

*Proof.* For suppose  $A$  is a subsemigroup and  $AA_G \not\subseteq A$ . Then there exists  $x \in AA_G$  such that  $a \notin A$  (Def. 3.1.2). But by the definition of semigroup product,  $x \in AA_G$  if and only if there exists  $a, b \in A$  such that  $a * b = x$  (Def. 11.1.5). But  $A$  is a subsemigroup and thus for all  $a, b \in A$  it is true that  $a * b \in A$ , a contradiction. Hence,  $AA_G \subseteq A$ . Now suppose  $AA_G \subseteq A$  and suppose  $A$  is not a subsemigroup. Then there exists  $a, b \in A$  such that  $a * b \notin A$  (Def. 11.1.6). But if  $a, b \in A$ , then  $a * b \in AA_G$  (Def. 11.1.5) and by hypothesis  $AA_G \subseteq A$ , and thus  $a * b \in A$  (Def. 3.1.2), a contradiction. Thus,  $A$  is a subsemigroup if and only if  $AA_G \subseteq A$ .  $\square$

**Theorem 11.1.10.** *If  $(G, *)$  is a semigroup, then  $G$  is a subsemigroup of  $(G, *)$ .*

*Proof.* For if  $(G, *)$  is a semigroup, then  $*$  is an associative binary operation (Def. 11.1.1), and hence  $*$  is a binary operation. But then if  $a, b \in G$ , then  $a * b \in G$  (Def. 5.2.1). Thus,  $G$  is a subsemigroup.  $\square$

**Theorem 11.1.11.** *If  $(G, *)$  is a semigroup, then  $\emptyset$  is a subsemigroup of  $G$ .*

*Proof.* For suppose not. Then there exists  $a, b \in \emptyset$  such that  $a * b \notin \emptyset$ , a contradiction since for all  $x$  it is true that  $x \notin \emptyset$  (Def. 3.2.1). Hence,  $\emptyset$  is a subsemigroup.  $\square$

**Theorem 11.1.12.** *If  $(G, *)$  is a semigroup, if  $\mathcal{S} \subseteq \mathcal{P}(G)$  is such that for all  $A \in \mathcal{S}$  it is true that  $A$  is a subsemigroup of  $G$ , then  $\bigcap \mathcal{S}$  is a subsemigroup of  $G$ .*

*Proof.* For if not then there exists  $a, b \in \bigcap \mathcal{S}$  such that  $a * b \notin \bigcap \mathcal{S}$  (Def. 11.1.6). But if  $a, b \in \bigcap \mathcal{S}$  then for all  $\mathcal{U} \in \mathcal{S}$  it is true that  $a, b \in \mathcal{U}$  (Def. 3.1.10). But by hypothesis, for all  $\mathcal{U} \in \mathcal{S}$  it is true that  $\mathcal{U}$  is a subsemigroup of  $G$ , and thus  $a * b \in \mathcal{U}$  (Def. 11.1.6). But if  $a * b \in \mathcal{U}$  for all  $\mathcal{U} \in \mathcal{S}$ , then  $a * b \in \bigcap \mathcal{S}$  (Def. 3.1.10), a contradiction. Thus,  $\bigcap \mathcal{S}$  is a subsemigroup.  $\square$

We can immediately apply this to any finite collection. In particular:

**Theorem 11.1.13.** *If  $(G, *)$  is a semigroup, if  $A, B \subseteq G$  are subsemigroups, then  $A \cap B$  is a subsemigroup.*

*Proof.* For by Thm. 3.1.3, the set  $\{A, B\}$  exists. Thus, by Thm. 11.1.12  $\bigcap \{A, B\}$  is a subsemigroup. But  $\bigcap \{A, B\} = A \cap B$ , hence  $A \cap B$  is a subsemigroup.  $\square$

Thus the structure of subsemigroups of a semigroup forms a complete lattice.

### Theorem 11.1.14: Lattice Theorem of Semigroups

---

If  $(G, *)$  is a semigroup, if  $\mathcal{S} \subseteq \mathcal{P}(G)$  is the set of all subsemigroups of  $G$ :

$$\mathcal{S} = \{ A \in \mathcal{P}(G) \mid A \text{ is a subsemigroup of } G \} \quad (11.1.28)$$

If  $(\mathcal{P}(G), \subseteq)$  is the usual partial ordering by inclusion, then  $(\mathcal{S}, \subseteq |_{\mathcal{S}})$  is a complete lattice.  $\blacksquare$

*Proof.* For let  $A, B \in \mathcal{S}$ . By Thm. 11.1.13,  $A \cap B$  is a subsemigroup. But  $A \cap B \subseteq A$  (Thm. 3.2.32) and  $A \cap B \subseteq B$  (Thm. 3.2.33), and hence  $A \cap B$  is a lower bound for  $A$  and  $B$ . Suppose it is not the greatest lower bound. Then there is a set  $C \in \mathcal{S}$  such that  $C \subseteq A$ ,  $C \subseteq B$ , and  $A \cap B \subsetneq C$ . But then there exists  $c \in C$  such that  $c \notin A \cap B$  (Thm. 3.2.12). But  $C \subseteq A$  and thus  $c \in A$  (Def. 3.1.2). Similarly,  $C \subseteq B$  and hence  $c \in B$ . But then  $c \in A \cap B$  (Def. 3.1.8), a contradiction. Thus,  $A \cap B$  is the greatest lower bound of  $A$  and  $B$ . By the axiom schema of specification (Ax. 3.1.3), there exists the set:

$$\mathcal{U}_{AB} = \{ \mathcal{U} \in \mathcal{S} \mid A \in \mathcal{U} \text{ and } B \in \mathcal{U} \} \quad (11.1.29)$$

But  $A, B \subseteq G$  and  $G$  is a subsemigroup of  $G$  (Thm. 11.1.10), and thus  $\mathcal{U}_{AB} \neq \emptyset$ . But by Thm. 11.1.12,  $\bigcap \mathcal{U}_{AB}$  is a subsemigroup. But by construction, for all  $\mathcal{U} \in \mathcal{U}_{AB}$ ,  $A \in \mathcal{U}$  and  $B \in \mathcal{U}$ , and hence  $A, B \in \bigcap \mathcal{U}_{AB}$ . Thus  $\bigcap \mathcal{U}_{AB}$  is an upper bound for  $A$  and  $B$ . Suppose it is not the least upper bound. Then there exists  $C \in \mathcal{S}$  such that  $A \subseteq C$ ,  $B \subseteq C$ , and  $C \subsetneq \bigcap \mathcal{U}_{AB}$ . But then  $C$  is a subsemigroup such that  $A \in C$  and  $B \in C$ , and thus  $C \in \mathcal{U}_{AB}$ . But then  $\bigcap \mathcal{U}_{AB} \subseteq C$ , a contradiction. Thus,  $\bigcap \mathcal{U}_{AB}$  is the least upper bound of  $A$  and  $B$ . Thus,  $(\mathcal{S}, \subseteq)$  is a complete lattice.  $\square$

**Example 11.1.26** Let  $X$  be a set with several distinct elements and let  $\mathcal{F}$  be the set of all constant mappings  $f : X \rightarrow X$ . That is, there is some  $c \in X$  such that for all  $x \in X$  it is true that  $f(x) = c$ , so  $f[X] = \{c\}$ . Let  $\circ$  denote function composition. We know that function composition is associative (Thm. 5.1.23). Moreover,  $\circ$  takes elements of  $\mathcal{F}$  to  $\mathcal{F}$ . For if  $f, g \in \mathcal{F}$  then there are  $c_f, c_g \in X$  such that  $f[X] = \{c_f\}$  and  $g[X] = \{c_g\}$ . But then, for all  $x \in X$  we have:

$$(g \circ f)(x) = g(f(x)) = g(c_f) = c_g \quad (11.1.30)$$

and thus  $g \circ f$  is a constant mapping as well. Therefore  $\circ$  is an associative binary operation on  $\mathcal{F}$  and  $(\mathcal{F}, \circ)$  is a semigroup (Def. 11.1.1).

The example shown in Ex. 11.1.26 is missing most algebraic properties. Notably, there is no identity. Since we chose  $X$  to have at least several distinct elements, for any distinct functions  $f, g \in \mathcal{F}$  we have  $g \circ f \neq f \circ g$  and hence there can be no unital element. There can also be no left unital element, and as such there can be no left invertible elements. Every element is a right unital element since  $g \circ f = g$ , and as such every element is *weakly* right invertible. Moreover, as this previous expression shows, the operation is not commutative. Thus  $(\mathcal{F}, \circ)$  is a semigroup but can't possibly be any of the nicer objects like monoids or groups. It motivates an important object in mathematics, the transformation semigroup. Avoiding proof by naming, we first show this is a semigroup.

**Theorem 11.1.15.** *If  $X$  is a set, if  $\mathcal{F}(X, X)$  is the set of all function from  $X$  to  $X$ , and if  $\circ$  denotes function composition, then  $(\mathcal{F}(X, X), \circ)$  is a semigroup.*

*Proof.* Since function composition is a binary operation on  $\mathcal{F}(X, X)$ , and since  $\circ$  is associative (Thm. 5.1.23), this forms a semigroup (Def. 11.1.1).  $\square$

### Definition 11.1.7: Semigroup of Transformations

The semigroup of transformations on a set  $X$  is the semigroup  $(\mathcal{F}(X, X), \circ)$  of the set of all functions  $f : X \rightarrow X$   $\mathcal{F}(X, X)$  with function composition.

This will be much more important later when we consider certain subsets of transformation semigroups which actually form groups. It turns out *every* group is of this form.<sup>11</sup> We now prove an analogous result for semigroups, that every semigroup is equivalent to a subset of a semigroup of transformations. To do this we need a precise notion of equivalence and we need to introduce the idea of subsemigroups, which are the semigroup analog of subsets in set theory.

<sup>11</sup> Up to isomorphism. This is Cayley's Theorem.

## 11.2 Groups

**Theorem 11.2.1.** If  $n \in \mathbb{N}^+$ , if  $\tilde{+}$  is addition modulo  $n$ , then  $(\mathbb{Z}/n\mathbb{Z}, \tilde{+})$  is a group.

*Proof.* For  $\tilde{+}$  is associative (Thm. ??),  $[0]$  is a unital element (Thm. ??), and for all  $x \in \mathbb{Z}/n\mathbb{Z}$  there is a  $y \in \mathbb{Z}/n\mathbb{Z}$  such that  $x\tilde{y}$  is a unital element (Thm. ??). Therefore,  $(\mathbb{Z}/n\mathbb{Z}, \tilde{+})$  is a group.  $\square$

**Theorem 11.2.2.** If  $p \in \mathbb{N}^+$  is prime, and if  $\tilde{\cdot}$  is the restriction of multiplication modulo  $n$  to  $\mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}$ , then  $(\mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}, \tilde{\cdot})$  is a group.

*Proof.* For  $[1]$  is a unital element (Thm. ??),  $\tilde{\cdot}$  is associative (Thm. ??). But  $p$  is prime, and hence for all non-zero  $x \in \mathbb{Z}/p\mathbb{Z}$  there is a multiplicative inverse (Thm. ??). Hence,  $(\mathbb{Z}/p\mathbb{Z} \setminus \{[0]\}, \tilde{\cdot})$  is a group.  $\square$

**Example 11.2.1** If  $G \subseteq \mathbb{C}$  is defined by:

$$G = \{ z \in \mathbb{C} \mid \text{There exists } n \in \mathbb{N}^+ \text{ such that } z^n = 1 \} \quad (11.2.1)$$

then  $G$  is a group under multiplication. For these are just the *roots of unity* and using the polar representation of a complex number we have  $z = \exp(2\pi im/n)$  for some  $m, n \in \mathbb{N}$ . Multiplying  $z \cdot w$  yields:

$$\exp\left(\frac{2\pi im}{n}\right) \exp\left(\frac{2\pi ij}{k}\right) = \exp\left(\frac{2\pi i(mk + nj)}{nk}\right) \quad (11.2.2)$$

showing that  $G$  is closed to multiplication. The identity element is contained in here since  $1 = 1^1$ , hence  $1 \in G$ . Lastly, inverses: If  $z = \exp(2\pi im/n)$ , let  $z^{-1} = \exp(-2\pi im/n)$ . However,  $G$  is not a group under addition since  $1 + 1$  is not contained in  $G$ . That is,  $1 + 1 = 2 \exp(2\pi i)$  in polar form and hence any power of this will not result in one since:

$$\|(2 \exp(2\pi i))^n\| = 2^n \|\exp(2\pi in)\| = 2^n \quad (11.2.3)$$

and this is not equal to 1 for all  $n \in \mathbb{N}^+$ .

**Example 11.2.2** Define  $G \subseteq \mathbb{R}$  by:

$$G = \{ a + b\sqrt{2} \mid a, b \in \mathbb{Q} \} \quad (11.2.4)$$

Then  $G$  is a group under addition. Given  $x, y \in G$  we have:

$$x + y = (a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2} \quad (11.2.5)$$

and hence  $G$  is closed under addition. The identity element is contained in  $G$  since  $0 = 0 + 0\sqrt{2}$ , and moreover so are additive inverses: Let  $-x = -a - b\sqrt{2}$ .

Since addition is also associative, we have that  $G$  is a group under addition. If we consider  $G \setminus \{0\}$  with multiplication, then this too is a group. Give  $x, y \in G$  we have:

$$x \cdot y = (a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bd)\sqrt{2} \quad (11.2.6)$$

since  $ad + bd \in \mathbb{Q}$  and  $ac + 2bd \in \mathbb{Q}$ , we have that  $x \cdot y$  is again an element of  $\mathbb{Q}$ . The identity is in  $G$ , setting  $a = 1$  and  $b = 0$ . Lastly, if  $x \in G$  then by definition  $x$  is not zero, and hence if  $x = a + b\sqrt{2}$  then either  $a$  is non-zero or  $b$  is non-zero. But then  $a^2 - 2b^2$  is non-zero since  $\sqrt{2}$  is irrational. Let  $x^{-1} = (a - b\sqrt{2})/(a^2 - 2b^2)$ . Then:

$$x \cdot x^{-1} = (a + b\sqrt{2}) \left( \frac{a - b\sqrt{2}}{a^2 - 2b^2} \right) = \frac{(a + b\sqrt{2})(a - b\sqrt{2})}{a^2 - 2b^2} = \frac{a^2 - 2b^2}{a^2 - 2b^2} = 1 \quad (11.2.7)$$

**Theorem 11.2.3.** *If  $(G, *)$  is a group, and if  $g \in G$  has order  $n \in \mathbb{N}$ , then  $g^{-1} = g^{n-1}$ .*

*Proof.* For  $e = g^n = g \cdot g^{n-1}$ , so the uniqueness of inverses  $g^{n-1} = g^{-1}$ .  $\square$

**Theorem 11.2.4.** *If  $(G, *)$  is a group, if  $x, y \in G$ , and if  $x$  and  $y$  commute, then  $y^{-1} * x * y = x$  and  $x^{-1} * y^{-1} * x * y = e$ .*

*Proof.* Since  $x$  and  $y$  commute,  $x * y = y * x$ . Apply the cancellation laws.  $\square$

**Theorem 11.2.5** (Power Laws). *If  $(G, *)$  is a group,  $a \in G$ ,  $n, m \in \mathbb{Z}$ , then:*

$$a^n * a^m = a^{n+m} \quad (11.2.8)$$

$$(a^{-1})^n = (a^n)^{-1} \quad (11.2.9)$$

**Theorem 11.2.6.** *If  $x \in G$  has odd order, there is a  $n \in \mathbb{Z}$  such that  $x = (x^2)^{n+1}$ .*

*Proof.* For  $e = x^{2n+1} = x^{2n}x$  and hence by cancellation  $x = x^{2n+2} = (x^2)^{n+1}$ .  $\square$

**Theorem 11.2.7.** *If  $(G, *)$  is a group,  $a \in G$  has order  $n$ ,  $g \in G$ , then  $g^{-1}ag$  has order  $n$ . Also, the order of  $a * b$  is the order of  $b * a$ .*

*Proof.* We show  $(g^{-1}ag)^n = g^{-1}a^n g$  by induction:

$$(g^{-1}ag)^{n+1} = (g^{-1}ag)(g^{-1}ag)^n = (g^{-1}ag)(g^{-1}a^n g) = g^{-1}a^{n+1}g \quad (11.2.10)$$

hence if  $a$  has order  $n$ , applying this shows that  $g^{-1}ag$  has order  $n$ .  $\square$

**Theorem 11.2.8.** *If  $a, b$  commute, then  $(ab)^n = a^n b^n$ .*

**Theorem 11.2.9** (Cyclic Subgroup). *If  $(G, *)$  is a group, if  $x \in G$ , and if  $H \subseteq G$  is defined by:*

$$H = \{x^n \mid n \in \mathbb{Z}\} \quad (11.2.11)$$

*Then  $(H, *)$  is a group.*

*Proof.* It contains the identity since  $x^0 = e$  by definition. It contains inverses since  $(x^n)^{-1} = x^{-n}$ . Lastly, it is closed under  $*$  since  $x^n * x^m = x^{n+m}$ . Hence, it is a group.  $\square$

**Theorem 11.2.10.** *If  $(A, *_A)$  and  $(B, *_B)$  are groups, if  $(A \times B, *)$  is the direct product of  $A$  and  $B$ , then  $(A \times B, *)$  is Abelian if and only if  $(A, *_A)$  and  $(B, *_B)$  are Abelian.*

*Proof.* For if  $(A, *_A)$  and  $(B, *_B)$  are Abelian,  $(a, b), (c, d) \in A \times B$ , then:

$$(a, b) * (c, d) = (a *_A c, b *_B d) = (c *_A a, d *_B b) = (c, d) * (a, b) \quad (11.2.12)$$

and hence  $(A \times B, *)$  is Abelian. The other direction is similar.  $\square$

**Theorem 11.2.11.** *If  $(A, *_A)$  and  $(B, *_B)$  are groups, if  $e_A \in A$  is the unital element of  $A$ , if  $e_B \in B$  is the unital element of  $B$ , if  $a \in A$ , if  $b \in B$ , and if  $(A \times B, *)$  is the direct product of  $A$  and  $B$ , then  $(a, e_B)$  and  $(e_A, b)$  commute in  $A \times B$ .*

*Proof.* For

$$(a, e_B) * (e_A, b) = (a *_A e_A, e_B *_B b) = (e_A *_A a, b *_B e_B) = (e_A, b) * (a, e_B) \quad (11.2.13)$$

$\square$

**Theorem 11.2.12.** *If  $(A, *_A)$  and  $(B, *_B)$  are groups, if  $a \in A$  has order  $n \in \mathbb{N}^+$ , if  $b \in B$  has order  $m \in \mathbb{N}^+$ , and if  $(A \times B, *)$  is the direct product of  $A$  and  $B$ , then  $(a, b)$  has order  $\text{LCM}(n, m)$  in  $A \times B$ .*

*Proof.* For:

$$(a, b)^N = ((a, e_B) * (e_A, b))^N = (a, e_B)^N * (e_A, b)^N = (a^N, e_B) * (e_A, b^N) \quad (11.2.14)$$

and hence the least  $N$  that makes  $a^N = e$  and  $b^N = e$  is the least  $N$  such that  $n|N$  and  $m|N$ . But this is just  $\text{LCM}(n, m)$ .  $\square$

DF 1.1.32

**Theorem 11.2.13.** *If  $(G, *)$  is a group,  $x \in G$  has order  $n \in \mathbb{N}$ , and if  $f : \mathbb{Z}_n \rightarrow G$  is defined by  $f(k) = x^k$ , then  $f$  is injective.*

*Proof.* For if not then there are distinct  $k_1, k_2 \in \mathbb{Z}_n$  such that  $f(k_1) = f(k_2)$ . Suppose  $k_1 < k_2$ . Then  $x^{k_2} = x^{k_1}$  and hence  $x^{k_2 - k_1} = e$ , a contradiction since  $k_2 - k_1 < n$  and  $n$  is the least such element in  $\mathbb{N}^+$  with  $x^n = e$ .  $\square$

**Theorem 11.2.14.** *If  $(G, *)$  is a group, if  $x \in G$  has infinite order, and if  $f : \mathbb{Z} \rightarrow G$  is defined by  $f(n) = x^n$ , then  $f$  is injective.*

*Proof.* For suppose not. Then  $x^m = x^n$  which implies  $x^{m-n} = e$ , a contradiction since  $x$  has infinite order. Hence,  $m - n = 0$  and thus  $m = n$ .  $\square$

**Theorem 11.2.15.** *If  $(G, *)$  is a group, if  $n \in \mathbb{N}^+$ , if  $x \in G$  has order  $n$ , and if  $N \in \mathbb{N}$ , then there is a  $K \in \mathbb{Z}_n$  such that  $x^K = x^N$ .*

*Proof.* For either  $N = 0$  or  $N \neq 0$ . But if  $N = 0$ , then since  $0 \in \mathbb{Z}_n$  we have that  $x^N = x^0$  and we're done. Suppose  $N \neq 0$ . But then  $N, n \in \mathbb{Z} \setminus \{0\}$  and hence by Euclid's division algorithm there exists  $q \in \mathbb{Z}$  and  $r \in \mathbb{N}$  such that  $r < n$  and  $N = qn + r$ . But then:

$$x^N = x^{qn+r} = x^{qn}x^r = (x^n)^qx^r = e^qx^r = x^r \quad (11.2.15)$$

proving the claim.  $\square$

**Theorem 11.2.16.** *If  $(G, *)$  is a group and  $a, b \in G$ , then:*

$$(a * b)^{-1} = b^{-1} * a^{-1} \quad (11.2.16)$$

*Proof.* For if  $(G, *)$  is a group, and if  $a, b \in G$ , then  $*$  is associative and  $a$  and  $b$  are invertible (Def. 11.1.3). But then by Thm. 5.2.28,  $a * b = b^{-1} * a^{-1}$ .  $\square$

**Theorem 11.2.17.** *If  $(G, *)$  is a group and  $a \in G$ , then:*

$$(a^{-1})^{-1} = a \quad (11.2.17)$$

*Proof.* For if  $e$  is the unital element of  $G$ , then:

$$\begin{aligned} a^{-1} * (a^{-1})^{-1} &= (a^{-1} * a)^{-1} && \text{(Thm. 11.2.16)} \\ &= e && \text{(Inverse Property)} \end{aligned}$$

From the uniqueness of inverses (Thm. 5.2.25),  $(a^{-1})^{-1} = a$ .  $\square$

One of the fundamental aspects of elementary algebra is adding and subtracting like expressions from two sides of an equation. That is, if we are given  $4x + 1 = 9$ , we solve this by adding  $-1$  to both sides, and then dividing both sides by  $4$  yielding  $x = 2$ . Such procedures are valid because of the *cancellation laws* of real number arithmetic. These laws hold for groups, as well.

**Theorem 11.2.18: Left Cancellation Law**

If  $(G, *)$  is a group, if  $a, b, c \in G$ , and if  $a * b = a * c$ , then  $b = c$ . ■

*Proof.* For let  $e$  be the unital element of  $G$ . Then:

$$\begin{aligned} b &= e * b && \text{(Identity)} &= (a^{-1} * a) * c && \text{(Associativity)} \\ &= (a^{-1} * a) * b && \text{(Inverse)} &= c * e && \text{(Inverse)} \\ &= a^{-1} * (a * b) && \text{(Associativity)} &= c && \text{(Identity)} \\ &= a^{-1} * (a * c) && \text{(Hypothesis)} \end{aligned}$$

Thus by the transitivity of equality (Thm. 3.2.11),  $b = c$ . □

**Theorem 11.2.19: Right Cancellation Law**

If  $(G, *)$  is a group, if  $a, b, c \in G$ , and if  $b * a = c * a$ , then  $b = c$ . ■

*Proof.* For let  $e$  be the unital element of  $G$ . Then:

$$\begin{aligned} b &= b * e && \text{(Identity)} &= c * (a * a^{-1}) && \text{(Associativity)} \\ &= b * (a * a^{-1}) && \text{(Inverse)} &= e * c && \text{(Inverse)} \\ &= (b * a) * a^{-1} && \text{(Associativity)} &= c && \text{(Identity)} \\ &= (c * a) * a^{-1} && \text{(Hypothesis)} \end{aligned}$$

And therefore by transitivity (Thm. 3.2.11),  $b = c$ . □

With this we can perform the basic substitution operations from elementary algebra.

**Example 11.2.3** Let  $(G, *)$  be a group with unital element  $e \in G$ , and suppose  $x \in G$  and  $a \in G$  are elements satisfying the following:

$$x^2 = b \quad (11.2.18a) \qquad x^7 = e \quad (11.2.18b)$$

where  $x^n = x^{n-1} * x$ . We can use the cancellation laws to solve for  $x$ . We have:

$$b^3 * x = (x^2)^3 * x = x^6 * x = x^7 = e \quad (11.2.19)$$

And thus we have  $b^3 * x = e$ . But  $b^3 * (b^3)^{-1} = e$ , and from the left cancellation law (Thm. 11.2.18) we obtain  $x = (b^3)^{-1}$ .

**Example 11.2.4** As a more complicated example, let  $(G, *)$  be a group with unital element  $e \in G$ , and  $a, b, x \in G$  satisfying the following:

$$a * x^2 = b \quad (11.2.20a) \qquad x^3 = e \quad (11.2.20b)$$

We can solve for  $x$  in terms of the other variables. Note that if we multiply both sides of the first equation on the right by  $x$  we obtain:

$$a * x^3 = b * x \quad (11.2.21)$$

But  $x^3 = e$ , and thus we can simplify this to  $a = b * x$ . One solution to this is  $b^{-1} * a$ , and applying the right cancellation law (Thm. 11.2.19) we see that this is the only solution. Thus,  $x = b^{-1} * a$ .

**Theorem 11.2.20.** *If  $(G, *)$  is a group, if  $a, b, c \in G$ , and if  $a * b = c$ , then  $b = a^{-1} * c$ .*

*Proof.* For  $a * (a^{-1} * c) = (a * a^{-1}) * c = e * c = c$ , by associativity and identity. But  $a * b = c$ , and thus by the left cancellation law (Thm. 11.2.18),  $b = a^{-1} * c$ .  $\square$

**Theorem 11.2.21.** *If  $(G, *)$  is a group, if  $a, b, c \in G$ , and if  $a * b = c$ , then  $a = c * b^{-1}$ .*

*Proof.* For  $(c * b^{-1}) * b = c * (b^{-1} * b) = c * e = c$ , by associativity and identity. But  $a * b = c$ , and thus by the right cancellation law (Thm. 11.2.19),  $a = c * b^{-1}$ .  $\square$

**Theorem 11.2.22.** *If  $(G, *)$  is a group and if  $a \in G$  is idempotent, then  $a$  is the unital element.*

*Proof.* If  $(G, *)$  is a group, then there is a unital element  $e \in G$  (Def. 11.1.3). But if  $a$  is idempotent, then  $a * a = a$  (Def. 5.2.3). But  $a * e = a$ , and thus by the left cancellation law (Thm. 11.2.18)  $a = e$ .  $\square$

Note that this does not say that  $a^2 = e$  implies that  $a = e$ , for consider the group  $(\mathbb{Z}_2, +)$ . Then  $1 + 1 = 0$ , which is the identity element, but  $1 \neq 0$ . Real valued arithmetic does not have this property. If  $a$  is a real number such that  $a + a = 0$ , then  $a = 0$ . For multiplication, if  $a \cdot a = 1$ , then we can conclude that either  $a = 1$  or  $a = -1$ . In a general group, however, there may be infinitely many elements such that  $a * a = e$ , yet  $a \neq e$ . Groups such that every element has the property that  $a^2 = e$  are called *Boolean groups*.

Groups may also lack *square roots*. That is, given a group  $(G, *)$  and an element  $a \in G$ , there may not be an element  $b \in G$  such that  $a = b * b$ . One need only think of the group  $(\mathbb{Q}^+, \cdot)$ , where  $\cdot$  denotes usual multiplication. This is a group, and  $2 \in \mathbb{Q}^+$ , but there is no element  $x \in \mathbb{Q}^+$  such that  $x^2 = 2$ . Groups do, however, possess the Latin square property.

### Theorem 11.2.23: Latin Square Property of Groups

If  $(G, *)$  is a group,  $a, b \in G$ , then there is a unique  $x \in G$  such that  $a * x = b$ .



*Proof.* For since  $(G, *)$  is a group, there is a unital element  $e \in G$  and an element  $a^{-1} \in G$  such that  $a * a^{-1} = e$  (Def. 11.1.3). Let  $x = a^{-1} * b$ . Then:

$$a * x = a * (a^{-1} * b) = (a * a^{-1}) * b = e * b = b \quad (11.2.22)$$

Moreover, if  $a * y = b$ , then by the left cancellation law 11.2.18,  $y = x$ . Therefore,  $x$  is unique.  $\square$

**Theorem 11.2.24.** *If  $(G, *)$  is a group, if  $a, b \in G$  are commuting elements, and if  $a^{-1}$  and  $b^{-1}$  are the inverses of  $a$  and  $b$ , respectively, then  $a^{-1}$  and  $b^{-1}$  are commuting elements.*

*Proof.* For:

$$\begin{aligned} b^{-1} * a^{-1} &= (a * b)^{-1} && (\text{Thm. 11.2.16}) \\ &= (b * a)^{-1} && (\text{Hypothesis}) \\ &= a^{-1} * b^{-1} && (\text{Thm. 11.2.16}) \end{aligned}$$

By transitivity,  $a^{-1} * b^{-1} = b^{-1} * a^{-1}$ . Thus,  $a^{-1}$  and  $b^{-1}$  commute (Def. 5.2.12).  $\square$

We will later define the center of a group to be the subset of all commuting elements. The above theorem will be used to show that the center is a subgroup of the original group.

**Theorem 11.2.25.** *If  $(G, *)$  is a group, if  $a, b \in G$  commute, and if  $n \in \mathbb{N}$ , then  $a^n$  and  $b$  commute.*

*Proof.* For by induction, suppose not. Then by the well ordering of  $\mathbb{N}$  there is a least  $n \in \mathbb{N}$  such that  $a^n$  and  $b$  do not commute. But  $a$  and  $b$  commute, and thus  $n > 1$ . But then  $a^{n-1}$  commutes with  $b$ . Thus:

$$\begin{aligned} a^n * b &= (a^{n-1} * a) * b && = (b * a^{n-1}) * a \quad (\text{Commutativity}) \\ && \quad (\text{Definition of } a^n) &= b * (a^{n-1} * a) && (\text{Associativity}) \\ &= a^{n-1} * (a * b) && \quad (\text{Associativity}) &= b * a^n && (\text{Definition of } a^n) \\ &= a^{n-1} * (b * a) && \quad (\text{Commutativity}) \\ && \quad (\text{Commutativity}) \\ &= (a^{n-1} * b) * a && \quad (\text{Associativity}) \end{aligned}$$

By transitivity,  $a^n * b = b * a^n$ , a contradiction. Thus, for all  $n \in \mathbb{N}$  it is true that  $a^n$  and  $b$  commute.  $\square$

**Theorem 11.2.26.** *If  $(G, *)$  is a group, if  $a, b \in G$  commute, and if  $n \in \mathbb{N}$ , then  $a^n$  and  $b^n$  commute.*

*Proof.* For by Thm. 11.2.25,  $b$  commutes with  $a^n$ . But again, if  $a^n$  commutes with  $b$ , then  $a^n$  commutes with  $b^n$  (Thm. 11.2.25), completing the proof.  $\square$

**Theorem 11.2.27.** *If  $(G, *)$  is a group, if  $a, b \in G$  commute, and if  $b^{-1}$  is the inverse element of  $b$ , then  $a = b * a * b^{-1}$*

*Proof.* For:

$$a = a * (b * b^{-1}) = (a * b) * b^{-1} = (b * a) * b^{-1} \quad (11.2.23)$$

By transitivity,  $a = b * a * b^{-1}$ .  $\square$

**Theorem 11.2.28.** *If  $(G, *)$  is a group, if  $a, b \in G$  are such that  $a = b * a * b^{-1}$ , then  $a$  and  $b$  commute.*

*Proof.* For:

$$a * b = (b * a * b^{-1}) * b = b * (a * b * b^{-1}) = b * a \quad (11.2.24)$$

By transitivity,  $a * b = b * a$ , and thus  $a$  and  $b$  commute (Def. 5.2.12).  $\square$

**Theorem 11.2.29.** *If  $(G, *)$  is a group, if  $a, b \in G$  commute, and if  $b^{-1}$  is the inverse of  $b$ , then  $a$  and  $b^{-1}$  commute.*

*Proof.* For let  $e \in G$  be the unital element. Then:

$$\begin{aligned} b^{-1} * a &= b^{-1} * (b * a * b^{-1}) \quad (\text{Thm. 11.2.27}) &= e * (a * b^{-1}) && (\text{Inverse}) \\ &= (b^{-1} * b) * (a * b^{-1}) &= a * b^{-1} && (\text{Identity}) \\ && && (\text{Associativity}) \end{aligned}$$

By transitivity,  $a * b^{-1} = b^{-1} * a$ , and thus  $a$  and  $b^{-1}$  commute (Def. 5.2.12).  $\square$

**Theorem 11.2.30.** *If  $(G, *)$  is a group, if  $e \in G$  is the unital element, and if  $a, b \in G$  commute, then  $a * b * a^{-1} * b^{-1} = e$ .*

*Proof.* For  $(a * b)^{-1} = b^{-1} * a^{-1}$  (Thm. 11.2.16). But if  $a$  and  $b$  commute, and  $a^{-1}$  and  $b^{-1}$  commute (Thm. 11.2.24), and thus  $b^{-1} * a^{-1} = a^{-1} * b^{-1}$ . But then:

$$e = (a * b) * (b^{-1} * a^{-1}) = (a * b) * (a^{-1} * b^{-1}) \quad (11.2.25)$$

Completing the proof.  $\square$

**Theorem 11.2.31.** *If  $(G, *)$  is a group, if  $a, g \in G$ , and if  $n \in \mathbb{N}$ , then"*

$$(g * a * g^{-1})^n = g * a^n * g^{-1} \quad (11.2.26)$$

*Proof.* For by induction, suppose not and let  $n \in \mathbb{N}$  be the least  $n$  such that the proposition fails. But by the definition of  $(g * a * g^{-1})^n$ , we have that:

$$(g * a * g^{-1})^1 = g * a * g^{-1} = g * a^1 * g^{-1} \quad (11.2.27)$$

and therefore  $n > 1$ . But then the proposition holds for  $n - 1$ . But then:

$$\begin{aligned} (g * a * g^{-1})^n &= (g * a * g^{-1})^{n-1} * (g * a * g^{-1}) && (\text{Definition of } x^n) \\ &= (g * a^{n-1} * g^{-1}) * (g * a * g^{-1}) && (\text{Inductive Hypothesis}) \\ &= (g * a^{n-1}) * (g^{-1} * g) * (a * g^{-1}) && (\text{Associativity}) \\ &= (g * a^{n-1}) * e * (a * g^{-1}) && (\text{Inverse}) \\ &= (g * a^{n-1}) * (a * g^{-1}) && (\text{Identity}) \\ &= g * (a^{n-1} * a) * g^{-1} && (\text{Associativity}) \\ &= g * a^n * g^{-1} && (\text{Definition of } x^n) \end{aligned}$$

A contradiction. Therefore, the proposition is true for all  $n \in \mathbb{N}$ .  $\square$

**Theorem 11.2.32.** If  $(G, *)$  is a group, if  $a, b \in G$  are commuting elements, and if  $n \in \mathbb{N}$ , then  $(a * b)^n = a^n * b^n$ .

*Proof.* For by induction, suppose not and let  $n \in \mathbb{N}$  be the least element such that the proposition fails. But  $(a * b)^1 = a * b = a^1 * b^1$  and therefore  $n > 1$ . But then the proposition is true for  $n - 1$ . But then:

$$\begin{aligned} (a * b)^n &= (a * b)^{n-1} * (a * b) && (\text{Definition of } x^n) \\ &= (a^{n-1} * b^{n-1}) * (a * b) && (\text{Inductive Hypothesis}) \\ &= ((a^{n-1} * b^{n-1}) * a) * b && (\text{Associativity}) \\ &= (a^{n-1} * (b^{n-1} * a)) * b && (\text{Associativity}) \\ &= (a^{n-1} * (a * b^{n-1})) * b && (\text{Thm. 11.2.25}) \\ &= ((a^{n-1} * a) * b^{n-1}) * b && (\text{Associativity}) \\ &= (a^n * b^{n-1}) * b && (\text{Definition of } x^n) \\ &= a^n * (b^{n-1} * b) && (\text{Associativity}) \\ &= a^n * b^n && (\text{Definition of } x^n) \end{aligned}$$

A contradiction. Thus, the proposition holds for all  $n \in \mathbb{N}$ .  $\square$

**Theorem 11.2.33.** If  $(G, *)$  is a group,  $a, b \in G$ , and  $a * b = b^{-1}$ , then  $(a * b)^2 = a$ .

*Proof.* For:

$$(a * b)^2 = (a * b) * (a * b) = a * (b * (a * b)) = a * (b * b^{-1}) = a * e = a \quad (11.2.28)$$

Completing the proof.  $\square$

**Theorem 11.2.34.** If  $(G, *)$  is a group, if  $a, b \in G$ , if  $a * b = b^{-1}$ , and if  $n \in \mathbb{N}$ , then  $(a * b)^{2n} = a^n$ .

*Proof.* For by induction, suppose not. Let  $n \in \mathbb{N}$  be the least integer such that the proposition fails. But by Thm. 11.2.33,  $n > 1$ . Thus the proposition is true for  $n - 1$ . But:

$$\begin{aligned}(a * b)^{2n} &= (a * b)^{2(n-1)} * (a * b)^2 && (\text{Definition of } x^n) \\&= a^{n-1} * (a * b)^2 && (\text{Inductive Hypothesis}) \\&= a^{n-1} * a && (\text{Thm. 11.2.33}) \\&= a^n && (\text{Definition of } x^n)\end{aligned}$$

A contradiction. Thus, the proposition is true for all  $n \in \mathbb{N}$ .  $\square$

### Definition 11.2.1: Square Root

A square root of an element  $a$  in a group  $(G, *)$  is an element  $b \in G$  such that  $b * b = a$ . We write  $b = \sqrt{a}$ .

This is the same as the usual notion of squares roots with real numbers, but we've now generalized the concept to abstract groups.

**Example 11.2.5** If we consider  $(\mathbb{R}^+, \cdot)$ , then every element  $r \in \mathbb{R}^+$  has a square root. This comes from the *completeness* of  $\mathbb{R}$ .

**Example 11.2.6** If we consider  $(\mathbb{C} \setminus \{0\}, \cdot)$ , then again, every element  $z \in \mathbb{C} \setminus \{0\}$  has a square root. To see this, let  $z = r \exp(i\theta)$ , where  $r \in \mathbb{R}^+$  and  $\theta \in [0, 2\pi)$ . This is the polar representation of a complex number. Since  $r > 0$ , by the previous example we know that it has a square root,  $\sqrt{r}$ . We then see that  $\sqrt{r} \exp(i\theta/2)$  is a square root of  $z$ . Thus, every element of  $\mathbb{C} \setminus \{0\}$  has a square root. Zero has a square root as well (itself), but  $(\mathbb{C}, \cdot)$  is not a group since 0 has no inverse element, and hence we excluded it from this example.

**Example 11.2.7** The multiplicative group  $(G, \cdot)$  does contain square roots for all of its elements. One need only consider  $2 \in \mathbb{Q}^+$ , where it is well known that  $\sqrt{2}$  is irrational, and thus  $\sqrt{2} \notin \mathbb{Q}^+$ .

**Theorem 11.2.35.** If  $(G, *)$  is a group, if  $a \in G$ , if  $e \in G$  is the unital element, and if  $a^3 = e$ , then  $a$  has a square root in  $G$ .

*Proof.* For if  $a^3 = e$ , then  $a^2 * a = e$ , and thus by the left cancellation law (Thm. 11.2.18),  $a^2 = a^{-1}$ . But then again by the left cancellation law,  $a = (a^{-1})^2$ . Thus,  $\sqrt{a} = a^{-1}$  (Def. 11.2.1).  $\square$

**Theorem 11.2.36.** If  $(G, *)$  is a group, if  $a, b, g \in G$ , and if  $g * a * g = b$ , then  $a * b$  has a square root in  $G$ .

*Proof.* For  $a * b = a * (g * a * g) = (a * g) * (a * g) = (a * g)^2$ , and thus  $\sqrt{a * b} = a * g$   $\square$

**Theorem 11.2.37.** If  $(G, *)$  is a group, if  $a \in G$ , and if  $a^{-1}$  has a square root in  $G$ , then  $a$  has a square root in  $G$ .

*Proof.* For if  $a^{-1}$  has a square root, there is a  $b \in G$  such that  $b^2 = a^{-1}$ . But  $(a^{-1})^{-1} = a$  (Thm. 11.2.17) and thus  $a = (b^2)^{-1}$ . But  $(b^2)^{-1} = (b^{-1})^2$  (Thm. 11.2.16). Thus,  $\sqrt{a} = b^{-1}$ .  $\square$

Similarly, we can define cube roots.

### Definition 11.2.2: Cube Root

A cube root of an element  $a$  in a group  $(G, *)$  is an element  $b \in G$  such that  $b^3 = a$ . We write  $b = \sqrt[3]{a}$ .

**Theorem 11.2.38.** If  $(G, *)$  is a group, if  $a \in G$ , and if  $a^{-1}$  has a cube root, then  $a$  has a cube root.

*Proof.* For if  $a^{-1}$  has a cube root, there exists a  $b \in G$  such that  $b^3 = a^{-1}$ . But  $(a^{-1})^{-1} = a$  (Thm. 11.2.17), and thus  $a = (b^3)^{-1}$ . But  $(b^3)^{-1} = (b^{-1})^3$  (Thm. 11.2.16), and thus  $a = (b^{-1})^3$ . Thus,  $\sqrt[3]{a} = b^{-1}$  (Def. 11.2.2).  $\square$

**Theorem 11.2.39.** If  $(G, *)$  is a group, if  $a, g \in G$ , and if  $g^2 * a * g = a^{-1}$ , then  $g * a * g^2 = a^{-1}$ .

*Proof.* For:

$$g^2 * a = a^{-1} * g^{-1} \quad (\text{Thm. 11.2.21})$$

$$\Rightarrow g^2 = a^{-1} * g^{-1} * a^{-1} \quad (\text{Thm. 11.2.21})$$

But then:

$$\begin{aligned} g * a * g^2 &= g * a * (a^{-1} * g^{-1} * a^{-1}) \\ &= g * (a * a^{-1}) * (g^{-1} * a^{-1}) && (\text{Associativity}) \\ &= (g * e) * (g^{-1} * a^{-1}) && (\text{Inverse}) \\ &= g * (g^{-1} * a^{-1}) && (\text{Identity}) \\ &= (g * g^{-1}) * a^{-1} && (\text{Associativity}) \\ &= e * a^{-1} && (\text{Inverse}) \\ &= a^{-1} && (\text{Identity}) \end{aligned}$$

Completing the proof. □

**Theorem 11.2.40.** *If  $(G, *)$  is a group, if  $a, g \in G$ , and if  $g^2 * a * g = a^{-1}$ , then  $a$  has a cube root.*

*Proof.* For if  $g^2 * a * g = a^{-1}$ , then:

$$\begin{aligned}
 (g * a * g)^3 &= (g * a * g) * (g * a * g) * (g * a * g) && (\text{Definition of } x^3) \\
 &= (g * a * g^2) * (a * (g^2 * a * g)) && (\text{Associativity}) \\
 &= (g * a * g^2) * (a * a^{-1}) && (\text{Hypothesis}) \\
 &= (g * a * g^2) * e && (\text{Inverse}) \\
 &= g * a * g^2 && (\text{Identity})
 \end{aligned}$$

But if  $g^2 * a * g = a^{-1}$ , then  $g * a * g^2 = a^{-1}$  (Thm. 11.2.39). Therefore,  $a^{-1}$  has a cube root (Def. 11.2.2). But if  $a^{-1}$  has a cube root, then  $a$  has a cube root (Thm. 11.2.38). Thus,  $a$  has a cube root. □

### Definition 11.2.3: Abelian Group

An **Abelian group** is a **group**  $(G, *)$  such that  $*$  is a **commutative operation**.

### Example 11.2.8: The Dihedral Group $D_6$

Not every group is Abelian, and a classic non-Abelian group is the group of symmetries on an equilateral triangle. This is the dihedral group  $D_6$ . It is formed by considering all of the distinct ways one can rearrange the three points on an equilateral triangle by means of rotation by  $60^\circ$  and by reflection across the  $y$  axis, as well as any combination of these two (see Fig. 11.1).

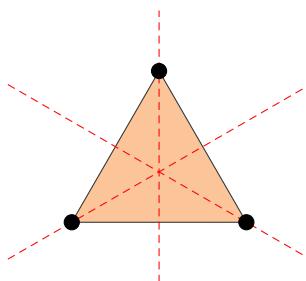


Fig. 11.1: The Dihedral Group  $D_6$

It turns out there are 6 distinct such moves, but by considering  $*$  to be the *successor* operation,  $(D_6, *)$  forms a group. The successor operation means that if  $r$  denotes rotation and  $a$  denotes reflection, then  $r * a$  denotes *rotate and then reflect*. By studying the triangle we get the *Cayley table* (Tab. 11.5) of this operation. A few things to note, the identity of our group is the *do nothing* symmetry. That is, we neither rotate nor reflect. Also note that reflecting twice in a row or rotating three times in a row is equivalent to doing nothing. The last thing to note is that reflection, rotation, then reflecting again is the same as rotating *backwards*. In other words, we have the following constraints:

$$r^3 = e \quad a^2 = e \quad (a * r) * (a * r) = e \quad (11.2.29)$$

We can verify this last statement via pictures. Fig. 11.2 shows that  $a * r * a * r$  is equivalent to doing nothing, as claimed.

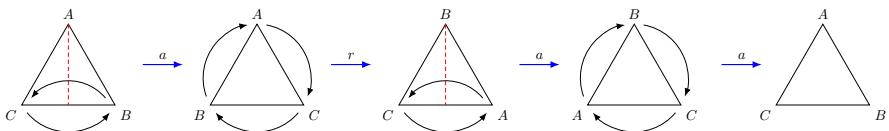
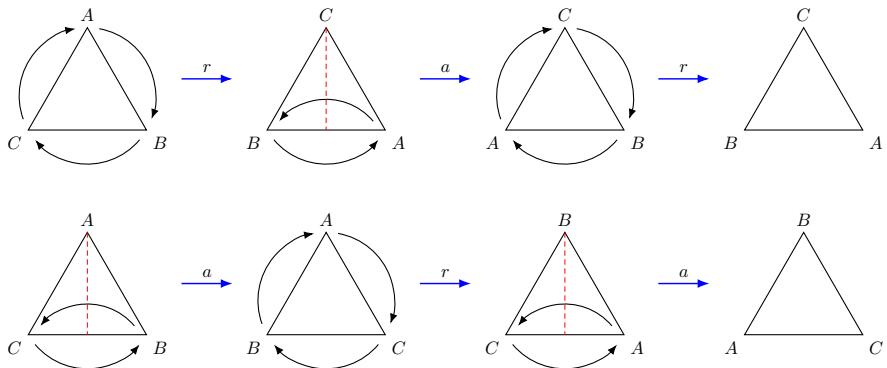


Fig. 11.2: Restraints on the Dihedral Group  $D_6$

This tells us that rotation and reflection has an inverse notion (rotate backwards and reflect again, respectively). Since the group is determined by these two operations, we need only check that  $r * (ar) = (r * a) * r$  and  $a * (r * a) = (a * r) * a$  to determine the associative of the rest of the group. We can do this by examining the triangle (see Fig. 11.3). That is, if we rotate, and then follow by reflecting and then rotating, it's the same thing as rotating and then reflecting, following by rotating again. Similarly, if we reflect, and then follow by rotating and then reflecting, this is equivalence to reflecting and then rotating, and then following with another reflection.

Fig. 11.3: Associativity of the Dihedral Group  $D_6$ 

With this we can compute the table (Tab. 11.5). We now see that  $(D_6, *)$  is not an Abelian group since the successor operation is not commutative. That is,  $r * a = a * r^2 \neq a * r$ , and thus  $a * r \neq r * a$ . ■

*	$e$	$r$	$r^2$	$a$	$a * r$	$a * r^2$
$e$	$e$	$r$	$r^2$	$a$	$a * r$	$a * r^2$
$r$	$r$	$r^2$	$e$	$a * r^2$	$a$	$a * r$
$r^2$	$r^2$	$e$	$r$	$a * r$	$a * r^2$	$a$
$a$	$a$	$a * r$	$a * r^2$	$e$	$r$	$r^2$
$a * r$	$a * r$	$a * r^2$	$a$	$r^2$	$e$	$r$
$a * r^2$	$a * r^2$	$a$	$a * r$	$r$	$r^2$	$e$

Table 11.5: Cayley Table of  $D_6$ 

**Example 11.2.9** The trivial group is the set  $G = \{\emptyset\}$ , although we usually label  $G = \{e\}$  or  $G = \{0\}$  when considering the trivial group, and we combine this with the operation  $* : G \times G \rightarrow G$  defined by  $e * e = e$ . This makes  $(G, *)$  a group, but moreover it is an Abelian group since  $*$  is trivially commutative.

**Example 11.2.10** Addition in  $\mathbb{Z}$  and multiplication in  $\mathbb{Q}^+$  are commutative operations, and thus  $(\mathbb{Z}, +)$  and  $(\mathbb{Q}^+, \cdot)$  are Abelian groups.

**Example 11.2.11** Addition of complex numbers  $\mathbb{C}$  is commutative and associative, as is the multiplication of non-zero complex numbers. Thus  $(\mathbb{C}, +)$  and  $(\mathbb{C} \setminus \{0\}, \cdot)$  are Abelian groups. The identities are  $0 = 0 + i0$  and  $1 = 1 + i0$ , respectively. Recall that we developed the arithmetic of  $\mathbb{C}$  based on the structure of  $\mathbb{R}^2$  together with the arithmetic of  $\mathbb{R}$ . We define:

$$a + ib = (a, b) \quad a, b \in \mathbb{R} \quad (11.2.30)$$

And defined addition and multiplication as follows:

$$(a + ib) + (c + id) = (a + c) + i(b + d) \quad (11.2.31a)$$

$$(a + ib) \cdot (c + id) = (ac - bd) + i(bc + ad) \quad (11.2.31b)$$

The commutativity and associativity of these operations stems from the arithmetic of  $\mathbb{R}$ , as does the fact that 0 and 1 are identities for these operations.

**Example 11.2.12** There are other Abelian group structures we can place on  $\mathbb{R}^2$ . Much the way the arithmetic of  $\mathbb{Q}$  was developed by building from the arithmetic of  $\mathbb{Z}$  and putting a structure on  $\mathbb{Z}^2$ , we can do the same of  $\mathbb{R}^2$ . Let  $*$  be the binary operation on  $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$  defined as follows:

$$(a, b) * (c, d) = (ad + bc, bd) \quad (11.2.32)$$

where  $ad$  denotes the usual real valued multiplication of  $a$  and  $b$ , and where  $+$  is the usual addition in  $\mathbb{R}$ . This operation will be clearer if we write it out as:

$$\frac{a}{b} * \frac{c}{d} = \frac{ad + bc}{bd} \quad (11.2.33)$$

Now we see why we required the second entry to be non-zero, this is the usually additive operation for fractions of real numbers. As such it is associative and commutative. Moreover, there is an identity  $(0, 1)$ , and an inverse  $(-a, b)$ . Thus, we have an Abelian group.

### Example 11.2.13: Group Operation on Power Set

---

When studying Boolean algebras we saw that on a set  $A$  the structure  $(\mathcal{P}(A), \cup, \cap)$  does not yield inverse elements. That is, given  $B, C \subseteq A$ , if  $A \cup C = \emptyset$  then  $A = C = \emptyset$ , and if  $B \cap C = A$ , then  $A = B = C$ . So Boolean algebras do not have an underlying group structure. That is, neither  $(\mathcal{P}(A), \cup)$  nor  $(\mathcal{P}(A), \cap)$  are groups. We can place a group structure on  $\mathcal{P}(A)$  by considering another familiar operation: The symmetric difference,  $\ominus$ . Recall that this is defined as:

$$B \ominus C = (B \cup C) \setminus (B \cap C) \quad (11.2.34)$$

The symmetric difference is both associative and commutative, and moreover there is a unital element since  $B \ominus \emptyset = B$ . Lastly, there is an inverse element, since  $B \ominus B = \emptyset$ . Thus,  $(\mathcal{P}(A), \ominus)$  forms an Abelian group. ■

### Example 11.2.14: Reflections on a Square

---

We now consider the reflections on a square, but do not consider rotations. We can visualize this by consider a point on one of the four vertices of a square and allowing it to move diagonally, horizontally, or vertically. Our operation is again the *successor* operation. Let's label  $h$  as horizontal, and similarly  $d$  and  $v$  for diagonal and vertical. Let  $e$  denote the *do nothing* reflection. Note that each of these *generators* is it's own inverse. Reflecting across the diagonal twice is equivalent to doing nothing, as are reflecting twice vertically or horizontally. We thus obtain the following restrictions:

$$h * h = e \quad v * v = e \quad d * d = e \quad (11.2.35)$$

Using these constraints, and the diagram below (Fig. 11.4), we can once again compute the Cayley table for this operation. The table is given by:

$e$	$e$	$h$	$v$	$d$
$e$	$e$	$h$	$v$	$d$
$h$	$h$	$e$	$d$	$v$
$v$	$v$	$d$	$e$	$h$
$d$	$d$	$v$	$h$	$e$

Table 11.6: Cayley Table for Reflection on a Square

This group is equivalent to the group structure that can be placed on  $\mathbb{Z}_2 \times \mathbb{Z}_2$  by considering pointwise modular addition. For example, we have:

$$(0, 1) + (1, 1) = (0 + 1, 1 + 1) = (1, 0) \quad (11.2.36)$$

and simillary for the other elements. This group structure, called the *direct product* of  $(\mathbb{Z}_2, +)$  with itself, is the same as the group structure we've described here. By looking at the Cayley table we see that the group of reflections on the square is an Abelian group. This is contrary to  $D_6$  where we allowed both rotations and reflections and saw that the group is not Abelian. ■

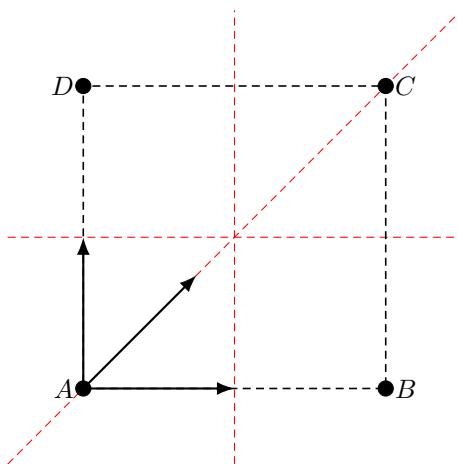


Fig. 11.4: Reflections on a Square

### Example 11.2.15: Group of Two Coins

Consider two coins located at points  $A$  and  $B$ . We group generated by the following operations:

- $e$ : Do nothing.
- $f$ : Flip the coin at  $A$ .
- $s$ : Swap the coins  $A$  and  $B$ .

From this we get that there are 8 moves total that can be generated from the *successor* operation, which generated by the following constraints:

$$f^2 = e \quad s^2 = e \quad (11.2.37)$$

That is, flipping the coin at  $A$  twice does nothing, nor does swapping the two coins twice. We also need one more restraint, and this tells us how to flip both coins without swapping them. We can achieve this by doing flip-swap-flip-swap, or equivalently by swap-flip-swap-flip. Thus, we have one more restraint:

$$fsfs = sfsf \quad (11.2.38)$$

This is equivalent to requiring  $(fsfs)^2 = e$ . We can now compute the table (Tab. 11.7) finding that this group is non-Abelian. ■

*	$e$	$f$	$s$	$fs$	$sf$	$fsf$	$sfs$	$sfsf$
$e$	$e$	$f$	$s$	$fs$	$sf$	$fsf$	$sfs$	$sfsf$
$f$	$f$	$e$	$fs$	$s$	$fsf$	$sf$	$sfsf$	$sfs$
$s$	$s$	$sf$	$e$	$sfs$	$f$	$sfsf$	$fs$	$sfs$
$fs$	$fs$	$fsf$	$f$	$sfsf$	$e$	$sfs$	$s$	$sf$
$sf$	$sf$	$s$	$sfs$	$e$	$sfsf$	$f$	$sfs$	$fs$
$fsf$	$fsf$	$fs$	$sfsf$	$f$	$sfs$	$e$	$sf$	$s$
$sfs$	$sfs$	$sfsf$	$sf$	$fsf$	$s$	$fs$	$e$	$f$
$sfsf$	$sfsf$	$sfs$	$fsf$	$sf$	$fs$	$s$	$f$	$e$

Table 11.7: Cayley Table of Group Formed on Two Coins

**Theorem 11.2.41.** If  $(G, *)$  is a group, if  $a, b \in G$  commute, and if  $g \in G$ , then  $g * a * g^{-1}$  and  $g * b * g^{-1}$  commute.

*Proof.* For:

$$\begin{aligned}
 (g * a * g^{-1}) * (g * b * g^{-1}) &= (g * a) * (g^{-1} * g) * (b * g^{-1}) && \text{(Associativity)} \\
 &= (g * a) * e * (b * g^{-1}) && \text{(Inverse)} \\
 &= (g * a) * (b * g^{-1}) && \text{(Identity)} \\
 &= g * (a * b) * g^{-1} && \text{(Associativity)} \\
 &= g * (b * a) * g^{-1} && \text{(Commutativity)} \\
 &= g * ((b * e) * a) * g^{-1} && \text{(Identity)} \\
 &= g * ((b * (g^{-1} * g)) * a) * g^{-1} && \text{(Inverse)} \\
 &= g * (((b * g^{-1}) * g) * a) * g^{-1} && \text{(Associativity)} \\
 &= g * (b * g^{-1}) * (g * a) * g^{-1} && \text{(Associativity)} \\
 &= (g * b * g^{-1}) * (g * a * g^{-1}) && \text{(Associativity)}
 \end{aligned}$$

And thus  $g * a * g^{-1}$  and  $g * b * g^{-1}$  commute (Def. 5.2.12).  $\square$

**Theorem 11.2.42.** If  $(G, *)$  is a group, if  $e$  is the unital element of  $G$ , and if  $a, b \in G$  are such that  $a * b = a$ , then  $b = e$ .

*Proof.* For since  $e$  is the unital element,  $a * e = a$  (Def. 5.2.6). But then by the left cancellation law (Thm. 11.2.18),  $b = e$ .  $\square$

**Theorem 11.2.43.** If  $(G, *)$  is a group, if  $e$  is the unital element of  $G$ , and if  $a, b \in G$  are such that  $a * b = b$ , then  $a = e$ .

*Proof.* For since  $e$  is the unital element,  $e * b = b$  (Def. 5.2.6). But then by the right cancellation law (Thm. 11.2.19),  $a = e$ .  $\square$

These two theorems state some slightly stronger than the uniqueness of the unital element, and show that to check that an element  $e \in G$  is the unital element it suffices to see that  $a * e = a$  for at least one  $a \in G$ .

**Theorem 11.2.44.** *If  $(G, *)$  is a group, if  $a \in G$ , and if  $f : G \rightarrow G$  is defined by  $f(g) = a * g$  for all  $g \in G$ , then  $f$  is surjective.*

*Proof.* For suppose not. Then there is a  $y \in G$  such that, for all  $x \in G$  it is true that  $f(x) \neq y$ . But by the Latin square property (Thm. 11.2.23), if  $a, y \in G$  then there is a unique  $x \in G$  such that  $a * x = y$ . But then  $f(x) = y$ , a contradiction. Therefore,  $f$  is surjective.  $\square$

**Theorem 11.2.45.** *If  $(G, *)$  is a group, if  $a \in G$ , and if  $f : G \rightarrow G$  is defined by  $f(g) = a * g$  for all  $g \in G$ , then  $f$  is injective.*

*Proof.* For suppose not. Then there exists  $x, y \in G$  such that  $x \neq y$  and  $f(x) = f(y)$ . But then  $a * x = a * y$ , and thus by the left cancellation law (Thm. 11.2.18),  $x = y$ , a contradiction. Therefore,  $f$  is injective.  $\square$

**Theorem 11.2.46.** *If  $(G, *)$  is a group, if  $a \in G$ , and if  $f : G \rightarrow G$  is defined by  $f(g) = a * g$  for all  $g \in G$ , then  $f$  is a permutation.*

*Proof.* For by Thm. 11.2.44,  $f$  is surjective. And by Thm. 11.2.45,  $f$  is injective. But then  $f$  is a bijection (Def. 5.1.4), and therefore  $f$  is a bijective function from  $G$  to itself, and is therefore a permutation (Def. 5.1.5).  $\square$

The function presented in these three theorems is the main tool used in proving Cayley's Theorem, which is one of the classic results of group theory.

#### Definition 11.2.4: Boolean Group

A Boolean group is a group  $(G, *)$  such that for all  $a \in G$  it is true that  $a^2 = e$ , where  $e$  is the unital element of  $G$ .

**Theorem 11.2.47.** *If  $(G, *)$  is a Boolean group, then it is Abelian.*

*Proof.* For suppose not. Then there are  $a, b \in G$  such that  $a * b \neq b * a$ . But:

$$\begin{aligned}
 a * b &= (a * b) * e && (\text{Identity}) \\
 &= (a * b) * (b * a)^2 && (\text{Boolean Property}) \\
 &= ((a * b) * (b * a)) * (b * a) && (\text{Associativity}) \\
 &= (a * (b * b) * a) * (b * a) && (\text{Associativity}) \\
 &= (a * e * a) * (b * a) && (\text{Boolean Property}) \\
 &= (a * a) * (b * a) && (\text{Identity}) \\
 &= e * (b * a) && (\text{Boolean Property}) \\
 &= b * a && (\text{Identity})
 \end{aligned}$$

A contradiction. Thus,  $(G, *)$  is Abelian.  $\square$

### 11.2.1 Direct Product

#### Definition 11.2.5: Group Direct Product

The direct product of two groups  $(G, *)$  and  $(G', \circ)$  is the Cartesian product  $G \times G'$  together with the binary operation  $\star : (G \times G') \times (G \times G') \rightarrow G \times G'$  defined by:

$$(a, a') \star (b, b') = (a * b, a' \circ b')$$

**Theorem 11.2.48.** *If  $(G, *)$  and  $(G', \circ)$  are groups, and if  $(G \times G', \star)$  is their direct product, then it is a group.*

**Theorem 11.2.49.** *If  $(G, *)$  and  $(G', \circ)$  are Abelian groups, and if  $(G \times G', \star)$  is their direct product, then it is an Abelian group.*

**Theorem 11.2.50.** *The direct product of Boolean groups is Boolean.*

**Example 11.2.16** As we will soon see, taking the direct product of two groups may not produce anything new and exciting. For example, the direct product of  $\mathbb{Z}_2$  with  $\mathbb{Z}_3$  is equivalent (isomorphic) to  $\mathbb{Z}_6$ . That is,  $\mathbb{Z}_2 \times \mathbb{Z}_3 \simeq \mathbb{Z}_6$ . We've already seen examples where this is not the case, and indeed  $\mathbb{Z}_4$  and  $\mathbb{Z}_2 \times \mathbb{Z}_2$  are different groups.

### 11.2.2 Subgroups

Subgroups are the group analog of subsets and subspaces. Given a group  $(G, *)$ , we consider a subset  $H \subseteq G$  such that  $H$  is *closed* under the group operation  $*$ , and such that it is closed to inverses. We can phrase this precisely.

**Definition 11.2.6:** Subgroup

A subgroup of a group  $(G, *)$  is a non-empty subset  $H \subseteq G$  such that for all  $a \in H$  it is true that  $a^{-1} \in G$ , and for all  $a, b \in H$  it is true that  $a * b \in G$ .

**Example 11.2.17** If we consider  $(\mathbb{Z}, +)$  and let  $\mathbb{Z}_e$  be the even integers, then  $\mathbb{Z}_e$  is a subgroup of  $\mathbb{Z}$ . For the sum of two even integers is again even, and the negative of an even integer is even, and thus  $\mathbb{Z}_e$  is closed under addition and under negation.

**Example 11.2.18** Let  $(\mathbb{R}, \cdot)$  denote the multiplicative group of positive real numbers. Then  $\mathbb{Q}^+$  is a subgroup. That is, the product of rational numbers is rational numbers, and the inverse of  $p/q$  (with  $p, q > 0$ ) is  $q/p$ , which is rational. Thus  $\mathbb{Q}^+$  is closed under multiplicative and inverses and is thus a subgroup of  $\mathbb{R}^+$ .

**Example 11.2.19** Let  $(\mathbb{Z}_4, +)$  denote the group of 4 elements under modulo arithmetic. There is a  $\mathbb{Z}_2$  subgroup hiding in here, for let  $H = \{0, 2\}$ . We can check case by case that this is a subgroup:

$$0 + 0 = 0 \quad 0 + 2 = 2 \quad 2 + 0 = 2 \quad 2 + 2 = 0 \quad (11.2.39)$$

and thus  $H$  is closed under modular addition. Finally, 0 is its own inverse, and since  $2 + 2 = 0$  we see that 2 is its own inverse as well, and thus  $H$  is closed under inversion.  $H$  is a subgroup of  $\mathbb{Z}_4$ . Our claim that this is a  $\mathbb{Z}_2$  in disguised can be realized by the function  $f : H \rightarrow \mathbb{Z}_2$  defined by  $f(0) = 2$  and  $f(2) = 1$ . This shows that  $H$  and  $\mathbb{Z}_2$  are essentially relabellings of the same structure.

**Example 11.2.20** If  $\mathcal{F}(\mathbb{R}, \mathbb{R})$  is the set of all functions  $f : \mathbb{R} \rightarrow \mathbb{R}$ , and if  $+$  denotes function addition:

$$(f + g)(x) = f(x) + g(x) \quad x \in \mathbb{R} \quad (11.2.40)$$

Then  $(\mathcal{F}(\mathbb{R}, \mathbb{R}), +)$  is a group. The identity is the zero function,  $\mathcal{O}(x) = 0$  for all  $x \in \mathbb{R}$ , and associativity stems from the associativity of addition in  $\mathbb{R}$ . To see that there are inverses, let  $-f$  denote the function:

$$(-f)(x) = -f(x) \quad x \in \mathbb{R} \quad (11.2.41)$$

If we let  $\mathcal{C}(\mathbb{R}, \mathbb{R})$  denote the set of all *continuous* functions, then this is a subgroup. Similarly, if  $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$  denotes the set of all differentiable functions, then this too is a subgroup. In general, if  $k \in \mathbb{N}$  and if  $\mathcal{C}^k(\mathbb{R}, \mathbb{R})$  denotes the set of all  $k$  times differentiable function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , then this is a subgroup of  $(\mathcal{F}(\mathbb{R}, \mathbb{R}))$ .

**Example 11.2.21** Again letting  $\mathcal{F}(\mathbb{R}, \mathbb{R})$ , the subset of all periodic functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  with period  $T > 0$  is again a subgroup  $(\mathcal{F}(\mathbb{R}, \mathbb{R}), +)$ .

**Example 11.2.22** If  $A$  and  $B$  are sets, and if  $A \subseteq B$ , then  $\mathcal{P}(A)$  is a subgroup of  $(\mathcal{P}(B), \ominus)$ , where  $\ominus$  denotes the symmetric difference operation, and  $\mathcal{P}(B)$  is the power set of  $B$ .

**Theorem 11.2.51.** *If  $(G, *)$  is a group, and if  $e \in G$  is the unital element, then  $\{e\}$  is a subgroup of  $(G, *)$ .*

*Proof.* For since  $e * e = e$ ,  $\{e\}$  is closed under  $*$  and to inverses. Thus,  $\{e\}$  is a subgroup (Def. 11.2.6).  $\square$

This is called the trivial subgroup of a group. At the other extreme, the entire group is a subgroup of itself.

**Theorem 11.2.52.** *If  $(G, *)$  is a group, then  $G$  is a subgroup of  $(G, *)$ .*

*Proof.* For since  $(G, *)$  is a group,  $G$  is closed to  $*$  and inverses (Def. 11.1.3). Thus,  $G$  is a subgroup (Def. 11.2.6).  $\square$

**Theorem 11.2.53.** *If  $(G, *)$  is a group, if  $H \subseteq G$  is a subgroup, and if  $e \in G$  is the unital element, then  $e \in H$ .*

*Proof.* For if  $H$  is a subgroup of  $(G, *)$  then it is non-empty (Def. 11.2.6). But if  $H$  is non-empty, then there is an  $a \in H$  (Def. 3.1.1). But  $H$  is a subgroup, and thus it is true that  $a^{-1} \in H$  (Def. 11.2.6) and since subgroups are closed under  $*$  we have that  $a * a^{-1} \in H$ . But  $a * a^{-1}$  is a unital element, and unital elements are unique (Thm. 5.2.4). Thus,  $e \in H$ .  $\square$

**Theorem 11.2.54.** \* *If  $(G, *)$  is a group, if  $H \subseteq G$  is a subgroup, and if  $*|_H$  is the restriction of  $*$  to  $H$ , then  $*|_H$  is a binary operation on  $H$ .*

*Proof.* Since the restriction of a function is a function, it suffices to show that the range of  $*|_H$  is  $H$ . For suppose not. Then there exists  $a, b \in H$  such that  $a * b \notin H$ . But  $H$  is a subgroup and thus if  $a, b \in H$ , then  $a * b \in H$  (Def. 11.2.6), a contradiction. Thus,  $*|_H$  is a binary operation on  $H$  (Def. 5.2.1).  $\square$

**Theorem 11.2.55.** *If  $(G, *)$  is a group, if  $H \subseteq G$  is a subgroup, and if  $*|_H$  is the restriction of  $*$  to  $H$ , then  $(H, *|_H)$  is a group.*

*Proof.* For  $*|_H$  is a binary operation on  $H$  (Thm. 11.2.54). Moreover since  $H$  is a subgroup, if  $a \in H$  then  $a^{-1} \in H$  (Def. 11.2.6) and thus for all  $a \in H$  it is true that  $a$  is invertible. Lastly, there is a unital element of  $(H, *|_H)$  (Thm. 11.2.53). Therefore,  $(H, *|_H)$  is a group (Def. 11.1.3).  $\square$

**Theorem 11.2.56.** If  $(G, *)$  is an Abelian group, and if  $H \subseteq G$  is defined by:

$$H = \{ a \in G \mid a^2 = e \} \quad (11.2.42)$$

then  $H$  is a subgroup of  $G$ .

*Proof.* Since  $(G, *)$  is Abelian,  $*$  is commutative (Def. 11.2.3). Thus, if  $a, b \in G$ , then:

$$\begin{aligned} (a * b)^2 &= (a * b) * (a * b) && \text{(Definition of } x^n\text{)} \\ &= (a * (b * a)) * b && \text{(Associativity)} \\ &= (a * (a * b)) * b && \text{(Commutativity)} \\ &= ((a * a) * b) * b && \text{(Associativity)} \\ &= (e * b) * b && \text{(Hypothesis)} \\ &= b * b && \text{(Identity)} \\ &= e && \text{(Hypothesis)} \end{aligned}$$

and thus  $a * b \in H$ . Moreover, if  $a \in H$  then:

$$\begin{aligned} (a^{-1})^2 &= a^{-1} * a^{-1} && \text{(Definition of } x^n\text{)} \\ &= (a^{-1} * e) * a^{-1} && \text{(Identity)} \\ &= (a^{-1} * (a * a)) * a^{-1} && \text{(Hypothesis)} \\ &= ((a^{-1} * a) * a) * a^{-1} && \text{(Associativity)} \\ &= (e * a) * a^{-1} && \text{(Inverse)} \\ &= a * a^{-1} && \text{(Identity)} \\ &= e && \text{(Inverse)} \end{aligned}$$

□

**Theorem 11.2.57.** If  $(G, *)$  is an Abelian group, if  $H \subseteq G$  is defined by:

$$H = \{ a \in G \mid \text{There exists } b \in G \text{ such that } b^2 = a \} \quad (11.2.43)$$

Then  $H$  is a subgroup of  $(G, *)$ .

*Proof.* For if  $a, b \in H$ , then there exists  $c, d \in G$  such that  $c^2 = a$  and  $d^2 = b$ . But  $G$  is Abelian, and thus  $*$  is commutative (Def. 11.2.3). Thus:

$$a * b = c^2 * d^2 = (c * d) * (c * d) = (c * d)^2 \quad (11.2.44)$$

and thus  $a * b \in H$ . Moreover  $a^{-1} \in H$ .

□

**Theorem 11.2.58.** If  $(G, *)$  is a group, if  $H, K \subseteq G$  are subgroups of  $(G, *)$ , and  $H \subseteq K$ , then  $H$  is a subgroup of  $(K, *|_K)$ , where  $*|_K$  is the restriction of  $*$  to  $K$ .

*Proof.* For if  $H \subseteq K$ , then for all  $a \in H$  it is true that  $a \in K$  (Def. 3.1.2). But since  $H$  is a subgroup of  $(G, *)$ , for all  $a, b \in H$  it is true that  $a * b \in H$  (Def. 11.2.6). But then  $a * b \in K$  and thus  $a * b = a * |_K b$ , and thus  $H$  is closed under  $*|_K$ . Moreover, if  $H$  is a subgroup of  $G$ , then for all  $a \in H$  it is true that  $a^{-1} \in H$ . But if  $a^{-1} \in H$ , then  $a^{-1} \in K$  since  $H \subseteq K$ . Thus,  $H$  is a subgroup of  $(K, *|_K)$ .  $\square$

### Definition 11.2.7: Center of a Group

The center of a group  $(G, *)$  is the set  $Z(G, *)$  defined by:

$$Z(G, *) = \{ a \in G \mid a * b = b * a \text{ for all } b \in G \}$$

Note that the center of a group is non-empty since the unital element of a group commutes with everything. Moreover, the center of a group is a subgroup.

**Theorem 11.2.59.** If  $(G, *)$  is a group and if  $Z(G, *)$  is the center of  $(G, *)$ , then  $Z(G, *)$  is a subgroup of  $(G, *)$ .

**Theorem 11.2.60.** If  $(G, *)$  is a finite group, and if  $S \subseteq G$  is closed under  $*$ , then  $S$  is a subgroup of  $G$ . That is,  $S$  is closed under inverses as well.

The set all periods of a function  $f : G \rightarrow G$  from a group  $(G, *)$  to itself forms a subgroup of  $G$ . That is, the set of all elements  $g \in G$  such that, for all  $x \in G$  it is true that  $f(g * x) = f(x)$ .

**Theorem 11.2.61.** If  $(G, *)$  is a group, if  $f : G \rightarrow G$ , and if  $H \subseteq G$  is defined by:

$$H = \{ g \in G \mid \text{For all } x \in G, f(g * x) = f(g) \} \quad (11.2.45)$$

then  $H$  is a subgroup of  $(G, *)$ .

**Theorem 11.2.62.** If  $(G, *)$  is a group, if  $(G', \circ)$  is a group, if  $e \in G$  is the unital element, and if  $H$  is defined by:

$$H = \{(e, g') \in G \times G' \mid g' \in G'\} \quad (11.2.46)$$

Then  $H$  is a subgroup of the direct product  $G \times G'$ .

**Theorem 11.2.63.** If  $(G, *)$  is a group, if  $(G', \circ)$  is a group, if  $e' \in G'$  is the unital element, and if  $H$  is defined by:

$$H = \{(g, e') \in G \times G' \mid g \in G\} \quad (11.2.47)$$

Then  $H$  is a subgroup of the direct product  $G \times G'$ .

**Theorem 11.2.64.** If  $(G, *)$  is a group, if  $\mathcal{P}(G)$  is the power set of  $G$ , if  $\mathcal{G} \subseteq \mathcal{P}(G)$  is such that for all  $H \in \mathcal{G}$  it is true that  $H$  is a subgroup of  $(G, *)$ , then the set  $N$  defined by:

$$N = \bigcap_{H \in \mathcal{G}} H \quad (11.2.48)$$

is a subgroup of  $(G, *)$ .

*Proof.* For  $N$  is not empty since for all  $H \in \mathcal{G}$ ,  $e \in H$  (Thm. 11.2.53). Suppose there exists  $a, b \in N$  such that  $a * b \notin N$ . But if  $a, b \in N$ , then for all  $H \in \mathcal{G}$  it is true that  $a, b \in H$  (Def. 3.1.10). But since  $H$  is a subgroup of  $(G, *)$ ,  $a * b \in H$ . But if  $a * b \in H$  for all  $H \in \mathcal{G}$ , then  $a * b \in N$ , a contradiction. Thus  $N$  is closed under  $*$ . Suppose there is an  $a \in N$  such that  $a^{-1} \notin N$ . But then  $a \in H$  for all  $H \in \mathcal{G}$  and thus  $a^{-1} \in H$  for all  $H$ , since  $H$  is a subgroup (Def. 11.2.6). But then  $a^{-1} \in a$  (Def. 3.1.10), a contradiction. Therefore,  $N$  is closed to inverses. Thus,  $N$  is a subgroup.  $\square$

This theorem allows us to define the subgroup of a group  $(G, *)$  generated from some set  $S \subseteq G$ .

**Theorem 11.2.65.** If  $(G, *)$  is a group, and if  $S \subseteq G$ , then there exists a set  $\mathcal{G}$  such that  $\mathcal{G}$  is non-empty and such that for all  $H \in \mathcal{G}$  it is true that  $S \subseteq H$  and  $H$  is a subgroup of  $(G, *)$ .

*Proof.* Apply the axiom schema of specification (Ax. 3.1.3) to the proposition  $P$  where  $P(H)$  is true if and only if  $S \subseteq H$  and  $H$  is a subgroup of  $G$  to the power set  $\mathcal{P}(G)$ . That is:

$$\mathcal{G} = \{ H \in \mathcal{P}(G) \mid P(H) \} \quad (11.2.49)$$

This completes the proof.  $\square$

This theorems means the following is well-defined.

### Definition 11.2.8: Generated Subgroup

The subgroup of a group  $(G, *)$  generated by a subset  $S \subseteq G$  is the set:

$$\langle S \rangle = \bigcap_{H \in \mathcal{G}} H$$

where  $\mathcal{G}$  is the collection of all subgroups of  $(G, *)$  that contain  $S$ .

**Example 11.2.23** Consider  $\mathbb{Z}_6$  with usual modular addition. The entire group can be considered as the group generated by the element 1. That is,  $2 = 1 + 1$ ,

$3 = 1 + 1 + 1$ , and so on, up until  $0 = 1 + 1 + 1 + 1 + 1 + 1$ . Thus we can write  $(\mathbb{Z}_6, +) = \langle 1 \rangle$ . Such groups are called finitely generated.

**Example 11.2.24** The Dihedral group  $D_6$  discussed before can be generated from two elements  $r, f$  with the requirements that  $r^3 = f^2 = e$  and that  $(rf)^2 = e$ . We can write this as:

$$D_6 = \langle r, f \mid r^3, f^2, (rf)^2 \rangle \quad (11.2.50)$$

This is called a *presentation* of  $D_6$ . Similarly, the dihedral group on 4 points,  $D_8$ , has the presentation:

$$D_8 = \langle r, f \mid r^4, f^2, (rf)^2 \rangle \quad (11.2.51)$$

The quaternion group  $Q$ , which is another non-Abelian group with 8 elements, has the presentation:

$$Q = \langle -1, i, j, k \mid (-1)^2 = 1, i^2 = j^2 = k^2 = ijk = -1 \rangle \quad (11.2.52)$$

### 11.2.3 Cayley Diagrams

Cayley diagrams can be used to represent finite groups and are formed from a directed graph. The points on the graph correspond to the elements of the group  $(G, *)$ , and the director arrows correspond to multiplying by generators of the group.

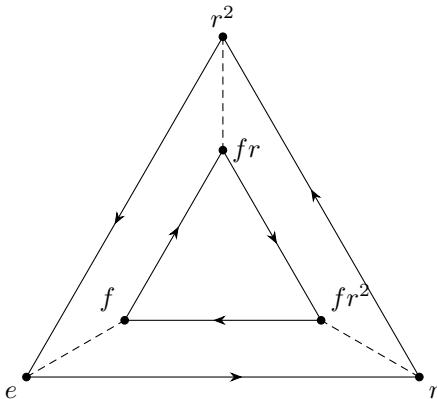
---

#### Example 11.2.25: Cayley Diagram of the Dihedral Group $D_6$

Consider the Dihedral group  $D_6$ , defined as the group on six points generated by  $r, f \in G$  such that  $r^3 = f^2 = (rf)^2 = e$ . That is, the group of rotational and reflectional symmetries on a triangle. It is thus not surprising that the Cayley diagram of  $D_6$  consists of triangle. The 6 points are:

$$G = \{ e, r, f, r^2, rf, r^2f \} \quad (11.2.53)$$

We'll draw a dotted arrow from  $g$  to  $h$  meaning that  $g * f = h$ , and a solid arrow for  $g * r = h$ .

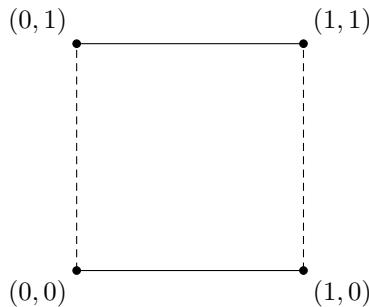
Fig. 11.5: Cayley Diagram of  $D_6$ 

Note that since  $f^2 = e$ , multiplying by  $f$  or by  $f^{-1}$  results in the same move, and thus arrows are redundant. Hence in the Cayley diagram, the dotted lines which represent reflections have no arrows. The Cayley diagram contains all of the information about the group. If we wish to compute  $(fr) * (fr^2)$ , we need only follow the arrows. That is, start at  $fr$  and then apply associativity to unwrap this expression. Applying  $f$  to  $fr$ , we move to  $r^2$ . We then apply  $r$  and end up at  $e$ . Applying  $r$  again, we arrive at  $r$ , and thus  $(fr) * (fr^2) = r$ . We can also compute inverses. The inverse of  $fr^2$  is the path that leads back to  $e$ . We see that  $(fr^2) * (r) * (f) = e$  and thus the inverse is  $rf = fr^2$ . ■

We can form the Cayley diagram of any finite group, and indeed we can form this for infinite groups as well, though the resulting graph is infinite as well. In particular, the *free group* on 2 generators can be expressed nicely via an infinite Cayley graph. Note that a directed graph is a Cayley graph for some group with generators  $a_k$  if and only if for all points  $g$  on the graph there is an arrow for  $a_k$  which starts at  $g$  and an arrow that ends on  $g$ .

### Example 11.2.26: Cayley Diagram on $\mathbb{Z}_2 \times \mathbb{Z}_2$

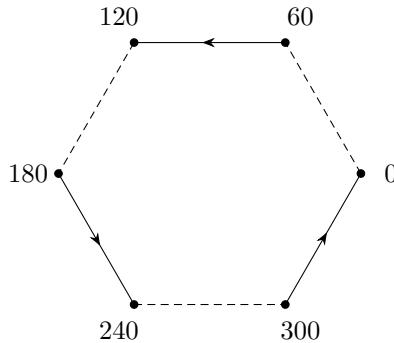
Consider the Cayley diagram on the group  $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ , where  $+$  is pointwise modular addition. That is, the arithmetic that stems from the direct product of  $\mathbb{Z}_2$  with itself. We arrive at Fig. 11.6. The dashed lines denote  $+(0, 1)$  and the solid lines denote  $+(1, 0)$ . Since  $(1, 0) + (1, 0) = (1+1, 0+0) = (0, 0)$ , and similarly for  $(0, 1)$ , there's no need to draw arrows on the diagram. ■

Fig. 11.6: Cayley Diagram of  $\mathbb{Z}_2 \times \mathbb{Z}_2$ **Example 11.2.27: Cayley Diagram of  $S_3$** 

Let  $n \in \mathbb{N}$  and  $G = \mathbb{Z}_n$ . Consider the set of all *permutations* on  $\mathbb{Z}_n$ . That is, the set of all bijective functions from  $\mathbb{Z}_n$  to itself. This is a group under function composition and is called the symmetric group  $S_n$ . In particular,  $S_n$  has  $n!$  points. Now consider  $S_3$ . We thus have that  $S_3$  has  $3! = 6$  points, and it can be seen that it is non-Abelian. As it turns out,  $S_3$  is equivalent to  $D_6$ , and indeed there are only two unique groups on 6 points ( $\mathbb{Z}_6$  and  $D_6$ ), this result will be proven later. But for now we can compute the Cayley diagram of  $S_3$ .  $S_3$  has the following presentation:

$$S_3 = \langle a, b \mid a^2 = b^2 = (ab)^3 = e \rangle \quad (11.2.54)$$

To see that this is the same thing as the dihedral group  $D_6$ , let  $a = f$  and let  $b = fr$ . Thus we have a different presentation that produces the same group, however this will produce a different Cayley diagram.

Fig. 11.7: Cayley Diagram for  $S_3$ 

While this Cayley diagram certainly looks different from the one presented in

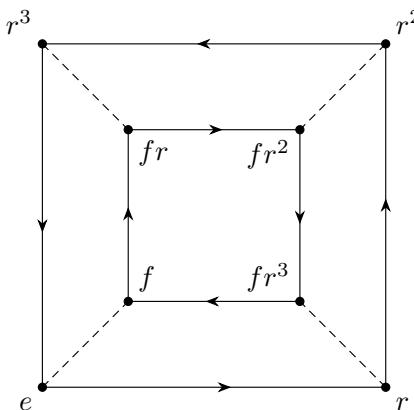
Ex. 11.2.25, they both represent the same group  $D_6$ . ■

### Example 11.2.28: The Dihedral Group $D_8$

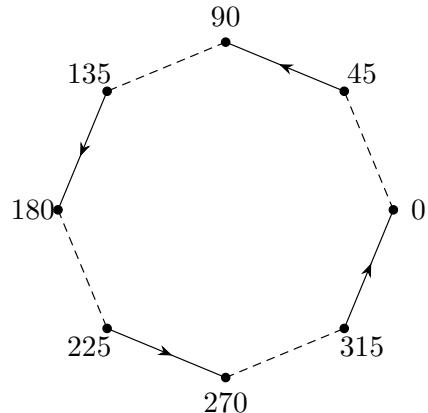
The dihedral group on 4 points,  $D_8$ , is the group of symmetries on the square. If we use as generators rotation and reflection, with the constraints  $r^4 = f^2 = e$ , and that  $(rf)^2 = e$ , we arrive at the Cayley Diagram shown in Fig. 11.8.1. Rather than representing the symmetries of a square by rotation and reflection, we can represent it by two different reflections. That is, let  $d$  denote diagonal reflection, and  $h$  represent horizontal reflection. We have the following presentation:

$$h^2 = d^2 = (hd)^4 = e \quad (11.2.55)$$

While this is equivalent to  $D_8$ , it presents a different, but equivalent, Cayley diagram (Fig. 11.8.2). ■



11.8.1: Standard Diagram for  $D_8$



11.8.2: Alternate Diagram for  $D_8$

Fig. 11.8: Cayley Diagrams for  $D_8$

### 11.2.4 Comments from Meeting with David

Abelianization (UMP, quotient out commutator)  $D_\infty$ ,  $D_{2n} = \mathbb{Z}_n \rtimes \mathbb{Z}_2$ . Chinese remainder theorem for proving Euler totient is multiplicative.  $D_\infty = \mathbb{Z} \rtimes \mathbb{Z}_2$ . We can write:

$$D_\infty = \langle r, s \mid s^2 = e, rsr^{-1} = s^{-1} \rangle = \langle a, b \mid a^2 = e = b^2 \rangle \quad (11.2.56)$$

Set  $b = s$  and  $a = rs$ .

## 11.3 Group Morphisms

### 11.3.1 Homomorphisms

#### Definition 11.3.1: Group Homomorphism

A group homomorphism from a group  $(G, *)$  to a group  $(G', \circ)$  is a bijective function  $\varphi : G \rightarrow G'$  such that for all  $a, b \in G$  it is true that:

$$\varphi(a * b) = \varphi(a) \circ \varphi(b)$$

#### Definition 11.3.2: Group Isomorphism

A group isomorphism from a group  $(G, *)$  to a group  $(G', \circ)$  is a bijective group homomorphism  $\varphi : G \rightarrow G'$ .

**Theorem 11.3.1.** If  $(G, *)$  and  $(G', \circ)$  are isomorphic with identities  $e_*$  and  $e_\circ$  are the identities, then  $f(e_*) = e_\circ$ .

*Proof.*  $\forall a \in G$ ,  $f(a) = f(a * e_*) = f(a) \circ f(e_*)$  as  $f$  is an isomorphism. As identities are unique,  $f(e_*) = e_\circ$ .  $\square$

**Theorem 11.3.2.** If  $(G, *)$  and  $(G', \circ)$  are isomorphic, with isomorphism  $f$ , and if  $a \in G$ , then  $f(a^{-1}) = f(a)^{-1}$ .

*Proof.* For:

$$e_\circ = f(e_*) = f(a * a^{-1}) = f(a^{-1} * a) = f(a) \circ f(a^{-1}) = f(a^{-1}) \circ f(a) \quad (11.3.1)$$

As inverses are unique,  $f(a^{-1}) = f(a)^{-1}$ .  $\square$

#### Definition 11.3.3: Group Epimorphism

An epimorphism from a group  $(G, *)$  to a group  $(H, \circ)$  is a homomorphism  $h : G \rightarrow H$  such that  $h$  is surjective.

**Definition 11.3.4: Group Monomorphism**

A monomorphism from a group  $(G, *)$  to a group  $(H, \circ)$  is a homomorphism  $h : G \rightarrow H$  such that  $h$  is injective.



# CHAPTER 12

---

## Finite Groups

---

Finite groups are of fundamental interest not only to mathematicians, but throughout many of the other sciences. Indeed, chemists and physicists make regular use of the theory of finite groups, and its application can be found in general relativity, quantum mechanics, and studying the lattice structure of organic molecules. A finite group is exactly what it sounds like: A group  $(G, *)$  where  $G$  is a finite set.

### Definition 12.0.1: Finite Group

A finite group is a [group](#)  $(G, *)$  such that  $G$  is a finite set.

One of the fundamental problems of group theory is a combinatorial one. Given an integer  $n \in \mathbb{N}$ , how many groups with  $n$  elements are there (up to isomorphism)? This challenging problem can be aided by the theorems of Cayley, Cauchy, Lagrange, and Sylow, and it is our aim to develop this theory.

## 12.1 Permutation Groups

Recall that a permutation on a set  $A$  is a bijective function  $f : A \rightarrow A$ . That is,  $f$  is a rearranging of  $A$ . Under the operation of function composition, given a non-empty set  $A$ , the set of all permutations on  $A$  together with function composition  $\circ$  have a group structure.

**Theorem 12.1.1.** *If  $A$  is a non-empty set, if  $S_A$  is the set of all permutations of  $A$ , and if  $\circ$  denotes function composition, then  $(S_A, \circ)$  is a group.*

*Proof.* For  $\circ$  is indeed a binary operation on  $S_A$ . There also exists a unital element, since  $\text{Id}_A$  is a permutation on  $A$ . Lastly, if  $f \in S_A$ , then it is a

bijection and thus there exists an inverse function  $g : A \rightarrow A$ . But the inverse of a permutation is a permutation, and thus  $g \in S_A$ . Thus,  $(S_A, \circ)$  is closed to inverses and is therefore a group (Def. 11.1.3).  $\square$

We will be most interested in case when  $A = \mathbb{Z}_n$  for some  $n \in \mathbb{N}$ . The set of all permutations on a set  $A$  is called the *symmetric group* of  $A$ .

### Definition 12.1.1: Symmetric Group

The symmetric group of a set  $A$  is the group  $(S_A \circ)$  of all permutations of  $A$  under function composition  $\circ$ .

By Thm. 12.1.1, the symmetric group is a group. The reason we required the underlying set to be non-empty is because the set of permutations of the empty set is empty, and thus  $S_\emptyset$  cannot be a group since groups are required to have a unital element, and thus cannot be empty.

### Example 12.1.1: Symmetric Group $S_3$

We've seen the symmetric group for  $\mathbb{Z}_3$  before and noted that it is isomorphic to  $D_6$ . Given a permutation  $f \in S_3$ , we can describe  $f$  via the following matrix:

$$f = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix} \quad (12.1.1)$$

The first row of the matrix is the input, and the second row is the output. This matrix tells us that  $f$  can be defined as follows:

$$f(n) = \begin{cases} 1, & n = 0 \\ 0, & n = 1 \\ 2, & n = 2 \end{cases} \quad (12.1.2)$$

We have that there are  $3! = 6$  permutations on  $\mathbb{Z}_3$ , and we can list them as follows:

$$\text{Id}_{\mathbb{Z}_3} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \quad (12.1.3a) \qquad \gamma = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix} \quad (12.1.3d)$$

$$\alpha = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix} \quad (12.1.3b) \qquad \delta = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix} \quad (12.1.3e)$$

$$\beta = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix} \quad (12.1.3c) \qquad \epsilon = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix} \quad (12.1.3f)$$

We can use this to compute compositions of permutations.

$$\beta \circ \delta = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} = \text{Id}_{\mathbb{Z}_3} \quad (12.1.4)$$

This should be read as *0 goes to 1 and 1 goes to 0, so 0 goes to 0*. That is, we read the  $\delta$  matrix first, and then feed this result to the  $\beta$  matrix. Similarly, 1 goes to 2 and 2 goes to 1, so 1 goes to 1. Lastly, 2 goes to 0 and 0 goes to 2, so 2 goes to 0. The resulting permutation is the identity permutation. Note that we are **not** performing matrix multiplication. On the one hand, we've yet to define matrix multiplication at this point, and on the other matrix multiplication is *undefined* for matrices of these sizes. That is, we cannot multiply a  $2 \times 3$  matrix by a  $2 \times 3$  matrix in the usual fashion. This representation is simply to aid in ones understanding of groups of permutations. The symmetric group is non-Abelian, as is  $D_6$ . We can see that it is non-Abelian by considering  $\alpha \circ \beta$  and  $\beta \circ \alpha$ . In  $\alpha \circ \beta$  we have that 0 goes to 2 and 2 goes to 1, so 0 goes to 1. However in  $\beta \circ \alpha$  we see that 0 goes to 0, and then 0 goes to 2, so 0 goes to 2. Thus  $\alpha \circ \beta \neq \beta \circ \alpha$ , so  $(S_3, \circ)$  is not Abelian. We can form a new representation of  $S_3$  to show that it is isomorphic to  $D_6$ . In fact, there are only two groups with 6 elements (up to isomorphism). ■

We do not have to consider *all* permutations on a given group, and the more general *group of permutations* is formed by considering subgroups of a symmetric group. What's remarkable is that *every* finite group is a subgroup group of one of the symmetric groups  $S_n$ . This result is known as Cayley's Theorem and will be proved shortly.

### 12.1.1 Finite Permutations

When dealing with permutations on a finite set, it is convenient to break up a given permutation into disjoint *cycles*. For example, suppose we have the following permutation on  $\mathbb{Z}_8$ :

$$f = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 5 & 6 & 4 & 0 & 7 & 3 & 1 \end{pmatrix} \quad (12.1.5)$$

We can draw this as two disjoint cycles as follows:

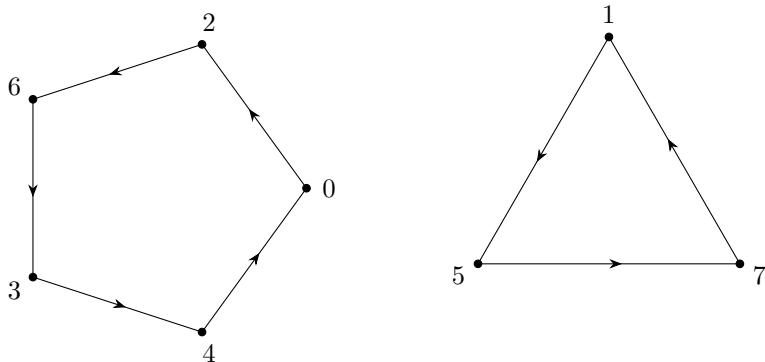


Fig. 12.1: Cycle Diagram for a Permutation

We can write this simply as the product of two cycle of the permutation:

$$f = (02634)(157) \quad (12.1.6)$$

First we need to rigorously define a cycle.

### Definition 12.1.2: Cycle Permutation

A cycle permutation on a finite set  $A$  is a permutation  $f \in S_A$  such that there exists two disjoint subsets  $M, N \subseteq A$  such that  $M \cup N = A$ ,  $f|_N = \text{Id}_A|_N$ , and such that there

**Theorem 12.1.2.** *Disjoint cycles commute.*

**Theorem 12.1.3.** *If  $A$  is a set and if  $f$  is a permutation on  $A$ , then  $f$  is the product of finitely many disjoint cycles. The product is unique.*

### Definition 12.1.3: Transposition

A transposition is a cycle of length 2.

**Theorem 12.1.4.** *Every cycle is the product of transpositions.*

**Theorem 12.1.5.** *If  $\alpha$  is a cycle of length  $s$ , and if  $\alpha^2$  is a cycle, then  $s$  is odd.*

**Theorem 12.1.6.** *If  $\alpha$  is a permutation, then  $\alpha^2$  is an even permutation.*

The decomposition of a cycle into transpositions need not be unique, and even the number of transpositions in such a decomposition can vary.

**Theorem 12.1.7.** *The identity is always the product of an even number of transpositions.*

**Theorem 12.1.8.** *The number of transpositions in the decomposition of a permutation is either always odd or always even.*

This always us to define the alternating group.

#### Definition 12.1.4: Alternating Group

The alternating group on a set  $B$  is the subgroup of  $S_B$  consisting of all even permutations. It is denoted  $A_B$ .

#### Notation 12.1.1: Set of Permutations

The set of all permutations from a set  $X$  to itself is denoted  $S_X$ .

**Theorem 12.1.9.** *If  $X$  is a set, if  $S_X$  is the set of all permutations of  $X$ , and if  $\circ$  is function composition, then  $(S_X, \circ)$  is a group.*

*Proof.* Since  $\text{id}_X$  is a permutation, there exists a unital element. Moreover,  $\circ$  is associative. Lastly, if  $f \in S_X$ , then since  $f$  is a permutation it is a bijection and hence there is an inverse function  $f^{-1}$ . Therefore,  $(S_X, \circ)$  is a group.  $\square$

The particularly important case is when the set  $X$  is finite. Indeed, historically these were the first groups to be studied. The important fact about  $S_n$ , as we will soon prove, is that every finite group is just a subgroup of  $S_n$  for some appropriate  $n$ .

**Theorem 12.1.10.** *If  $n \in \mathbb{N}^+$ , and if  $S_n$  is the set of all permutations of  $\mathbb{Z}_n$ , then  $\text{Card}(S_n) = n!$ .*

*Proof.* By induction. The base case of  $n = 1$  is true since the only permutation of  $\mathbb{Z}_1$  is the identity, and hence  $\text{Card}(S_1) = 1 = 1!$ . Suppose it is true for  $n \in \mathbb{N}^+$ . If  $f$  is a permutation of  $\mathbb{Z}_{n+1}$ , then there are  $n+1$  options for  $n$  to map to. Restricting  $f$  to  $\mathbb{Z}_n$  and removing  $f(n)$  from the image gives us a permutation of  $\mathbb{Z}_n$ , and there are  $n!$  of those. Thus, there are  $(n+1) \cdot n!$  total permutations, and this is just  $(n+1)!$ .  $\square$

## 12.2 Dihedral Groups

Pretty pictures, presentation, etc. If  $k \in \mathbb{Z}_k$  then  $fr^k = r^{-k}f$ . Every element has unique representation  $f^j r^k$  with  $j = 0$  or  $j = 1$  and  $k \in \mathbb{Z}_n$ .

### Definition 12.2.1: Generator of a Group

A generator of a group  $(G, *)$  is a subset  $S \subseteq G$  such that for all  $g \in G$  there exists an  $n \in \mathbb{N}$  and a sequence  $a : \mathbb{Z}_n \rightarrow S$  such that:

$$g = \prod_{k \in \mathbb{Z}_n} a_k$$

Note that sequences allow for repetition and have a notion of order, and so this respects the potential non-commutativity of a group. Another way of interpreting this definition is that every element of  $G$  can be written as the finite product of elements in  $S$ .

**Example 12.2.1** For any dihedral group  $(D_{2n}, \circ)$ , the set  $S = \{r, f\}$  containing the rotation and the reflection element is a generator for  $D_{2n}$ . To tell two such dihedral groups apart we introduce *relations*. For example,  $r^n = e$  and  $f^2 = e$ , and these are the least such positive integers with these properties. Moreover,  $rf = fr^{-1}$ .

**Example 12.2.2** Consider  $G$  with the presentation:

$$G = \langle a, b \mid a^n = e, b^2 = e, ab = ba^2 \rangle \quad (12.2.1)$$

Let's see if we can determine what this group is. We have:

$$a = ae \tag{12.2.2a} \qquad \qquad \qquad = (ba)(ba^2) \tag{12.2.2f}$$

$$= ab^2 \tag{12.2.2b} \qquad \qquad \qquad = (bab)a^2 \tag{12.2.2g}$$

$$= (ab)b \tag{12.2.2c} \qquad \qquad \qquad = b(ba^2)a^2 \tag{12.2.2h}$$

$$= (ba^2)b \tag{12.2.2d} \qquad \qquad \qquad = b^2a^4 \tag{12.2.2i}$$

$$= (ba)(ab) \tag{12.2.2e} \qquad \qquad \qquad = a^4 \tag{12.2.2j}$$

and hence by the cancellation law we conclude that  $a^3 = e$ . Hence, we may take  $n$  to be either 1, 2, or 3. If  $n = 1$  we are left with the group presented by a single variable  $b$  such that  $b^2 = e$ , and this is just  $\mathbb{Z}/2\mathbb{Z}$ . If  $n = 2$  then from  $ab = ba^2$  we conclude  $ab = b$ , and hence  $a = e$ , again leading us to  $\mathbb{Z}/2\mathbb{Z}$ . Finally, with  $n = 3$  we have  $ab = ba^2$  implying that  $ab = ba^{-1}$ , and this is precisely the presentation for the dihedral group  $D_6$ .

**Example 12.2.3** Consider  $G$  defined by:

$$G = \langle a, b \mid a^4 = e, b^3 = e, ab = b^2a^2 \rangle \quad (12.2.3)$$

Let's show that  $G$  is just the trivial group. We have:

$$bab = b(ab) = b(b^2a^2) = b^3a^2 = ea^2 = a^2 \quad (12.2.4)$$

and hence:

$$e = a^4 = (bab)^2 = (bab)(bab) = bab^2ab = bab^2b^2a^2 = bab^3ba^2 = baba^2 \quad (12.2.5)$$

and since  $b^3 = e$ , we apply the cancellation law to obtain  $b^2 = aba^2$ . But then:

$$a = ae = ab^3 = (ab)b^2 = b^2a^2(aba^2) = b^2a^3ba^2 = b^{-1}a^{-1}ba^2 \quad (12.2.6)$$

and hence  $aba = ba^2$ , and thus  $ab = ba$ . So the operation commutes. But then:

$$ab = b^2a^2 = a^2b^2 \quad (12.2.7)$$

and so applying cancellation we have  $ab = e$  and hence  $a = b^{-1}$ . But  $b^{-1} = b^2$  since  $b^3 = e$  and thus  $a = b^2$ . But then:

$$a^3 = (b^2)^3 = (b^3)^2 = e \quad (12.2.8)$$

But  $a^4 = e$  and therefore by the cancellation law  $a = e$ . And since  $b = a^{-1}$  we have that  $b = e$ . Hence,  $G$  is the trivial group.

The presentation of the general dihedral group  $D_{2n}$  is:

$$D_{2n} = \langle r, f \mid r^n = e, f^2 = e, rf = fr^{-1} \rangle \quad (12.2.9)$$

**Theorem 12.2.1.** If  $D_{2n}$  is the dihedral group with rotational generator  $r$  and reflectional generator  $f$ , and if  $x \in D_{2n}$  is not a power of  $r$ , then  $rx = xr^{-1}$ .

*Proof.* By induction. Suppose  $a : \mathbb{Z}_k \rightarrow \{r, f\}$  is a least sequence such that the product is  $x$ . In the base case of  $k = 1$ , since  $x$  is not a power of  $r$  we simply have that  $x = f$ . But  $rf = fr^{-1}$ , so we are done. Suppose it is true of  $k \in \mathbb{N}$  and let  $a : \mathbb{Z}_{k+1} \rightarrow \{r, f\}$  be a sequence whose product equals  $x$ . Then either  $a_k = r$  or  $a_k = f$ . Suppose  $a_k = r$ . Since  $x$  is not a power of  $r$ , there is and  $i \in \mathbb{Z}_k$  such that  $a_i \neq r$ . Define  $y$  by:

$$y = \prod_{j \in \mathbb{Z}_n} a_j \quad (12.2.10)$$

Then since  $y$  is not a power of  $r$ ,  $ry = yr^{-1}$ . But then:

$$rx = r(yr) = (ry)r = (yr^{-1})r = y(r^{-1}r) = ye = y \quad (12.2.11)$$

but  $x = yr$ , and hence  $y = xr^{-1}$ . Thus  $rx = x^{-1}$ . If  $a_k = f$ , let  $y$  be defined similarly. If  $y = r^k$ , then:

$$rx = ryf = rr^kf = r^krf = r^kfr^{-1} = xr^{-1} \quad (12.2.12)$$

if  $y$  is not a power of  $r$ , then by the induction hypothesis we have that  $ry = yr^{-1}$ . But then:

$$rx = r(ry) = (ry)r = (yr^{-1})r = y(r^{-1}r) = ye = y \quad (12.2.13)$$

But  $x = yr$  and hence  $y = xr^{-1}$ , so  $rx = xr^{-1}$ , as claimed.  $\square$

**Theorem 12.2.2.** *In  $D_{2n}$  every element that is not a power of  $r$  has order 2.*

*Proof.* For every such element can be written  $fr^k$  and hence  $x^2 = (fr^k)(fr^k) = (fr^k)(r^{-k}f) = f^2 = e$ .  $\square$

**Theorem 12.2.3.** *If  $n \in \mathbb{N}^+$ , if  $D_{2n}$  is the dihedral group with rotational element  $r$ , then the order of  $r$  is  $n$ .*

*Proof.* By definition of  $D_{2n}$ ,  $r^n$  is the identity element.  $\square$

**Example 12.2.4** We can also consider the dihedral group on two points. We have the presentation:

$$D_4 = \langle r, f \mid r^2 = e, f^2 = e, rf = fr^{-1} \rangle \quad (12.2.14)$$

This is isomorphic to  $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ . To see this, note that the presentation allows us to show that  $D_4$  is commutative. Since  $r^2 = e$  we know that  $r^{-1} = r$  and thus the last relation shows that  $rf = fr$ , so  $D_4$  is commutative. Let  $\phi : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \rightarrow D_4$  be defined by:

$$\phi(x) = \begin{cases} e, & x = (0, 0) \\ r, & x = (1, 0) \\ f, & x = (0, 1) \\ rf, & x = (1, 1) \end{cases} \quad (12.2.15)$$

Then  $\phi$  is injective. Moreover, it is surjective. For if  $y \in D_4$  then since  $D_4$  is commutative we know that  $y = r^j f^k$ . But we only care about  $j$  and  $k$  mod 2, and so we have four possibilities, each of which is mapped onto by  $\phi$ . Moreover, it is an isomorphism.

## 12.3 Polyhedra Groups

DF 1.2 Problems 9-13.

**Part V**

# **Ring Theory**



# CHAPTER 13

---

## Rings

---

We now add more structure by considering a set with two operations. Everything we've studied so far (semigroups, quasigroups, monoids, groups) has had only one operation associated to it, but in the most fundamental forms of arithmetic there are two. The only structure we've encountered with two operations so far has been Boolean algebras (see Book One), but as we will see when we study topology, there is essentially only one type of Boolean algebra and thus this study is, in a sense, complete. If we are going to axiomatize some algebraic structure it is then wise to avoid recreating Boolean algebras and so instead we try to model the arithmetic of the real numbers. The most fundamental properties can be stated quite succinctly:  $(\mathbb{R}, +)$  is an Abelian group and  $(\mathbb{R}, \cdot)$  is a monoid. We cannot just leave it there, however, since we've no way of knowing how  $+$  and  $\cdot$  play together. As presented, we have two potentially unrelated binary operations and thus cannot proceed any further. To complete our structure, we add [distributivity](#).

### 13.1 Definitions

The notion described in the preceding paragraph is that of a [ring](#). We can weaken this slightly and describe what is called a *rng*.

#### 13.1.1 Rngs

##### Definition 13.1.1: Rng

A [rng](#) is an [Abelian group](#)  $(R, +)$  and a [semigroup](#)  $(R, \cdot)$  such that  $\cdot$  is a [distributive operation](#) over  $+$ . A rng is denoted  $(R, +, \cdot)$ .

The unital element of the Abelian group  $(R, +)$  is often denoted 0 and is called the zero element. Zero has the property that multiplication by zero returns zero for any element.

**Theorem 13.1.1.** *If  $(R, +, \cdot)$  is a rng, if 0 is the unital element of  $(R, +)$ , and if  $r \in R$ , then  $r \cdot 0 = 0$ .*

*Proof.* For since  $(R, +)$  is an Abelian group, if  $r \cdot 0 \in R$  then there is an inverse element  $-r \cdot 0$  (Def. 11.1.3). But then:

$$\begin{aligned} 0 &= r \cdot 0 - r \cdot 0 && \text{(Inverse Property of Groups)} \\ &= r \cdot (0 - 0) && \text{(Distributive Property)} \\ &= r \cdot 0 && \text{(Identity Property)} \end{aligned}$$

And therefore  $r \cdot 0 = 0$ . □

The converse of this theorem fails in a general rng, but is true in a ring (see Thm. 13.1.3). To see this we will show that we can construct a rng from any Abelian group in a trivial manner such that the converse fails.

**Theorem 13.1.2.** *If  $(R, +)$  is an Abelian group, then there is an associative binary operation  $\cdot$  on  $R$  such that  $(R, +, \cdot)$  is a rng.*

*Proof.* For let  $\cdot : R \times R \rightarrow R$  be defined by  $r \cdot s = 0$  for all  $r, s \in R$ . Then  $\cdot$  is a binary operation since for all  $(r, s) \in R \times R$  there is a unique  $x \in R$  such that  $r \cdot s = x$ , and thus  $\cdot$  is a function (Def. 3.1.14). Moreover, it is associative. For if not then there are  $r, s, t \in R$  such that  $r \cdot (s \cdot t) \neq (r \cdot s) \cdot t$ . But by definition  $s \cdot t = 0$ , and thus  $r \cdot 0 = 0$ . Similarly,  $r \cdot s = 0$  and  $0 \cdot t = 0$ , a contradiction. Thus  $\cdot$  is associative. Thus,  $(R, \cdot)$  is a semigroup (Def. 11.1.1). Finally,  $\cdot$  is associative over  $+$ . For let  $a, b, c \in R$ . Then:

$$a \cdot (b + c) = 0 = 0 + 0 = (a \cdot b) + (a \cdot c) \tag{13.1.1}$$

and thus by the transitivity of equality (Thm. 3.2.11),  $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ . Thus,  $(R, +, \cdot)$  is a rng (Def. 13.1.1). □

A rng with the property that  $a \cdot b = 0$  for all  $a, b \in R$  is occasionally called a rng of square zero. If we take an Abelian group with at least two elements, the rng of square zero generated from this will have the property that there exists a  $b \in R$  such that  $b \neq 0$  and for all  $a \in R$ ,  $a \cdot b = b \cdot a = 0$ .

### 13.1.2 Rings

#### Definition 13.1.2: Ring

A **ring** is a **rng**  $(R, +, \cdot)$  such that  $(R, \cdot)$  is a **monoid**. That is, an **Abelian group**  $(R, +)$  and a monoid  $(R, \cdot)$  such that  $\cdot$  is a **distributive operation** over  $+$ .

Once  $(R, \cdot)$  has a unital element we can prove that zero is the only element such that  $0 \cdot r = 0$  for all  $r \in R$ . That is, the converse of Thm. 13.1.1 is true in a ring.

**Theorem 13.1.3.** *If  $(R, +, \cdot)$  is a ring, if 0 is the unital element of  $(R, +)$ , and if  $r$  is such that for all  $s \in R$  it is true that  $r \cdot s = 0$ , then  $r = 0$ .*

*Proof.* For if  $(R, +, \cdot)$  is a ring, then  $(R, \cdot)$  is a monoid (Def. 16.3.1) and thus there is a unital element, 1, of  $(R, \cdot)$  (Def. 11.1.2). But then for all  $s \in R$  we have:

$$\begin{aligned} r + s &= (r + s) \cdot 1 && \text{(Identity Property of 1)} \\ &= (r \cdot 1) + (s \cdot 1) && \text{(Distributive Property)} \\ &= 0 + s \cdot 1 && \text{(Hypothesis)} \\ &= s \cdot 1 && \text{(Identity Property of 0)} \\ &= s && \text{(Identity Property of 1)} \end{aligned}$$

And thus, for all  $s \in R$ ,  $r + s = s$ . But  $(R, +)$  is an Abelian group, and thus if  $r + s = s$ , then  $s + r = s$  (Def. 11.2.3) and therefore  $r$  is a unital element of  $(R, +)$  (Def. 5.2.6). But the unital element of a group is unique, and therefore  $r = 0$ .  $\square$

Given any Abelian group  $(R, +)$ , there is an associative binary operation  $\cdot$  such that  $(R, +, \cdot)$  is a rng (Thm. 13.1.2). Rings do not have this property, as we will now demonstrate.

**Theorem 13.1.4.** *If  $(R, +)$  is an Abelian group such that for all  $a \in R$  there is an  $n \in \mathbb{N}^+$  such that  $na = 0$ , if  $E$  is the set:*

$$E = \{ n \in \mathbb{N}^+ \mid \exists_{a \in R} : na = 0 \wedge \forall_{k \in \mathbb{Z}_n \setminus \{0\}} (ka \neq 0) \} \quad (13.1.2)$$

*and if  $E$  is infinite, then there is no binary operation  $\cdot$  on  $R$  such that  $(R, +, \cdot)$  is a ring.*

*Proof.* For suppose not and let  $\cdot$  be such a binary operation. But if  $(R, +, \cdot)$  is a ring, then  $(R, \cdot)$  is a monoid (Def. 16.3.1) and thus there is a unital element

1 of  $(R, \cdot)$  (Def. 11.1.2). But by hypothesis there is an  $n \in \mathbb{N}$  such that  $n1 = 0$ . But then for all  $a \in R$ , by associativity and Thm. 13.1.1 we have:

$$ka = k(1 \cdot a) = (k1) \cdot a = 0 \cdot a = 0 \quad (13.1.3)$$

Thus the set  $E$  is bounded by  $k$ . But then  $E \subseteq \mathbb{Z}_k$ , and is therefore finite. A contradiction as  $E$  is infinite. Thus,  $(R, +, \cdot)$  is not a ring.  $\square$

**Example 13.1.1** Let  $(\mathbb{Q}, +)$  and  $(\mathbb{Z}, +)$  be the usual additive groups and consider the quotient group on  $\mathbb{Q}/\mathbb{Z}$ . This is Abelian since  $\mathbb{Z}$  is a normal subgroup of  $(\mathbb{Q}, +)$ , but for every element  $a \in \mathbb{Q}/\mathbb{Z}$  there is an  $n \in \mathbb{N}^+$  such that  $na = 0$ . That is, given an equivalence class  $a \in \mathbb{Q}/\mathbb{Z}$ , let  $p/q \in [0, 1)$  be a representative. Then  $qa = [p] = [0]$ , where  $[p]$  is the equivalence class of  $[p]$  in the quotient group. Thus every element is a torsion element. Moreover, for all  $n \in \mathbb{N}^+$  there is a  $a \in \mathbb{Q}/\mathbb{Z}$  such  $na = 0$  and for all  $k \in \mathbb{Z}_n^+$ ,  $ka \neq 0$ . To see this, let  $a = [1/n]$ . So by Thm. 13.1.4 there is no ring structure on  $\mathbb{Q}/\mathbb{Z}$ .

## 13.2 Ring Morphisms

### Definition 13.2.1: Ring Homomorphism

A **ring homomorphism** from a **ring**  $(R_1, +, \cdot)$  to a **ring**  $(R_2, +', *)$  is a **function**  $f : R_1 \rightarrow R_2$  such that, for all  $x, y \in R_1$ , the following are true:

$$f(x + y) = f(x) +' f(y) \quad (\text{Preservation of Addition})$$

$$f(x \cdot y) = f(x) * f(y) \quad (\text{Preservation of Multiplication})$$

$$f(1_{R_1}) = 1_{R_2} \quad (\text{Preservation of Identities})$$

Where  $1_{R_1}$  is the unital element of  $R_1$  and  $1_{R_2}$  is the unital element of  $R_2$ .

There's a special name for a homomorphism from a ring  $(R, +, \cdot)$  to itself.

### Definition 13.2.2: Ring Endomorphisms

A **ring endomorphism** on a **ring**  $(R, +, \cdot)$  is a **ring homomorphism** from  $(R, +, \cdot)$  to itself. That is, a ring homomorphism  $f : R \rightarrow R$ .

**Notation 13.2.1: Set of Ring Endomorphisms**

The set of ring endomorphisms on a ring  $\mathcal{R} = (R, +, \cdot)$  is denoted  $\text{End}(\mathcal{R})$ .



# CHAPTER 14

---

## Fields

---

### 14.1 Definitions

#### Definition 14.1.1: Fields

A field is a commutative ring  $(\mathbb{F}, +, \cdot)$  such that, for all  $a \in \mathbb{F}$  such that  $a$  is not the unital element of  $(\mathbb{F}, +)$ , it is true that  $a$  is an invertible element of  $(\mathbb{F}, \cdot)$ .

#### Definition 14.1.2: Subfield

A subfield of a field  $(F, +, \cdot)$  is a set  $K \subset F$ , such that  $(K, +, \cdot)$  is a field.

Given an element  $a \in \mathbb{F}$ , if  $b$  is such that  $a + b = 0$  then we write  $b = -a$ . Subtraction of two elements  $a$  and  $c$ , denoted  $a - c$ , is defined as  $a + (-c)$ . The structure  $(\mathbb{F}, +)$  forms an Abelian group. From this we have that the identity is unique, as are additive inverses. It is common in the definition of a field to require that  $0 \neq 1$ . This is because if  $0 = 1$  then we have  $\mathbb{F} = \{0\}$ . This comes from the following.

#### Theorem 14.1.1: Multiplication by Zero

---

If  $(\mathbb{F}, +, \cdot)$  is a field, and if  $a \in \mathbb{F}$ , then  $a \cdot 0 = 0$ . ■

*Proof.* For we have:

$$0 = a \cdot (0) - a \cdot (0) = a \cdot (0 - 0) = a \cdot 0 \quad (14.1.1)$$

This simply combines the distributive law with the additive property of zero, completing the proof.  $\square$

**Theorem 14.1.2.** *If  $(\mathbb{F}, +, \cdot)$  is a field, and if  $0 = 1$ , then  $\mathbb{F} = \{0\}$ .*

*Proof.* For suppose not, and let  $a \in \mathbb{F}$  be such that  $a \neq 0$ . But then, by Thm. 14.1.1:

$$a = a \cdot 1 = a \cdot 0 = 0 \quad (14.1.2)$$

And thus  $a = 0$ , a contradiction. Therefore,  $\mathbb{F}$  is trivial.  $\square$

It is thus common to either call such a field a trivial field, or to require that  $0 \neq 1$ .

### Example 14.1.1: Examples of Fields

There are several fields that should be familiar to the reader. If we let  $\mathbb{R}$  denote the real numbers and  $+$  and  $\cdot$  be the usual notations of addition and multiplication, then  $(\mathbb{R}, +, \cdot)$  is a field. Similarly, letting  $\mathbb{Q}$  denote the rational numbers and  $\mathbb{C}$  denote the complex numbers,  $(\mathbb{Q}, +, \cdot)$  is a field, as is  $(\mathbb{C}, +, \cdot)$ . There are finite fields as well. Let  $\mathbb{F}_2 = \{0, 1\}$  and define multiplication and addition as follows:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Table 14.1: The Arithmetic of  $\mathbb{F}_2$

$(\mathbb{F}_2, +, \cdot)$  forms a field. Finally, if  $p \in \mathbb{N}$  is prime, and if  $+$  and  $\cdot$  are addition and multiplication mod  $p$ , respectively, then  $(\mathbb{Z}_p, +, \cdot)$  is a field.  $\blacksquare$

**Definition 14.1.3: Vector Space**

A vector space over a field  $(\mathbb{F}, +, \cdot)$  is a set  $V$  and a function  $\cdot : \mathbb{F} \times V \rightarrow V$  and a binary operation  $+$  on  $V$ , usually called scalar multiplication and vector addition, respectively, such that for all  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ , and all  $a, b \in \mathbb{F}$ , the following is true:

1.  $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$  [Associative of Vector Addition]
2.  $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$  [Commutativity of Vector Addition]
3. There is a  $\mathbf{0} \in V$  such that  $\mathbf{0} + \mathbf{x} = \mathbf{x}$  [Existence of Zero Vector]
4. For all  $\mathbf{x}$  there is a  $\mathbf{y}$  such that  $\mathbf{x} + \mathbf{y} = \mathbf{0}$  [Additive Inverses]
5.  $(a \cdot b) \cdot \mathbf{x} = a \cdot (b \cdot \mathbf{x})$  [Compatibility of Multiplication]
6.  $(a + b) \cdot \mathbf{x} = (a \cdot \mathbf{x}) + (b \cdot \mathbf{x})$  [Distributive Law for Field Addition]
7.  $a \cdot (\mathbf{x} + \mathbf{y}) = (a \cdot \mathbf{x}) + (a \cdot \mathbf{y})$  [Distributive Law for Vector Addition]

It is quite common not to distinguish between scalar multiplication  $\cdot$  and field multiplication  $\cdot$ , which may cause confusion. It is also common to drop the use of a symbol altogether and simply represent multiplication by concatenation of the two variables, for example  $a\mathbf{x}$  or  $ab$ , which represents scalar multiplication and field multiplication, respectively.

**Example 14.1.2** If we let  $\mathbb{F} = \mathbb{R}$  and let  $V = \mathbb{R}^n$ , where addition, multiplication, scalar multiplication, and vector addition are defined in their usual manner, then this forms a vector space. Similarly, the space  $C([a, b])$  of continuous functions forms a vector space over  $\mathbb{R}$ , as does  $L^2(\mathbb{R})$ , the space of square integrable functions.

**Definition 14.1.4: Bilinear Operations**

A bilinear operation on a vector space  $(V, +, \cdot)$  over a field  $(\mathbf{F}, +, \cdot)$  is a function  $[] : V \times V \rightarrow V$  such that, for all  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ , and for all  $a, b \in \mathbf{F}$ , the following is true:

1.  $[\mathbf{x} + \mathbf{y}, \mathbf{z}] = [\mathbf{x}, \mathbf{z}] + [\mathbf{y}, \mathbf{z}]$  [Right Distributive Law]
2.  $[\mathbf{x}, \mathbf{y} + \mathbf{z}] = [\mathbf{x}, \mathbf{y}] + [\mathbf{x}, \mathbf{z}]$  [Left Distributive Law]
3.  $[a \cdot \mathbf{x}, b \cdot \mathbf{y}] = (a \cdot b) \cdot [\mathbf{x}, \mathbf{y}]$  [Compatibility with Scalars]

**Example 14.1.3: Examples of Bilinear Operations**

The quintessential example of a bilinear operation is the cross product that one encounters in a multivariable calculus course. That is, for any three vectors  $\mathbf{x}, \mathbf{y}, \mathbf{z}$ , we have:

$$\mathbf{x} \times (\mathbf{y} + \mathbf{z}) = \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z} \quad (14.1.3)$$

Similarly for right sided multiplication. The compatibility of the cross product with scalar multiplication is also true:

$$(ax) \times (by) = ab(\mathbf{x} \times \mathbf{y}) \quad (14.1.4)$$

This serves somewhat as a motivating example for bilinear operations. If we think of the field of invertible matrices, then multiplication forms a bilinear operation as well, with scalar multiplication being the usual entry wise operation that is done on matrices. Lastly, if  $\langle \rangle$  is an inner product on  $\mathbb{R}$  or  $\mathbb{C}$ , then this is a bilinear operation, the vector space being the underlying field itself.

**Definition 14.1.5: Algebra over a Field**

An algebra of a field  $(\mathbf{F}, +, \cdot)$  is a vector space  $(\mathbf{V}, +, \cdot)$  and a bilinear operation  $[] : V \times V \rightarrow V$ .

**Definition 14.1.6: Associative Algebra over a Field**

An associative algebra over a field  $(\mathbb{F}, +, \cdot)$  is an algebra  $(V, [] )$  over  $\mathbb{F}$  such that, for all  $r \in \mathbb{F}$  and for all  $\mathbf{x}, \mathbf{y} \in V$ , the following is true:

$$r[\mathbf{x}, \mathbf{y}] = [r\mathbf{x}, \mathbf{y}] = [\mathbf{x}, r\mathbf{y}] \quad (14.1.5)$$

**Definition 14.1.7: Derivation on an Algebra**

A derivation on an algebra  $(V, [] )$  is a function  $D : V \rightarrow V$  such that for all  $\mathbf{x}, \mathbf{y} \in V$ , the following (Liebniz's Rule) is true:

$$D([\mathbf{x}, \mathbf{y}]) = [\mathbf{x}, D(\mathbf{y})] + [D(\mathbf{x}), \mathbf{y}] \quad (14.1.6)$$

**Theorem 14.1.3.** *In a field, 0 and 1 are unique.*

*Proof.* For suppose not, and let  $0'$  and  $1'$  be other identities. Then  $1' = 1' \cdot 1 = 1$  and  $0' = 0' + 0 = 0$ .  $\square$

**Theorem 14.1.4.** *For any field  $\langle F, +, \cdot \rangle$  and  $a \in F$ ,  $a \cdot 0 = 0$ .*

*Proof.* For:

$$0 = a \cdot 0 + (-a \cdot 0) = a \cdot (0 + 0) + (-a \cdot 0) = a \cdot 0 + a \cdot 0 + (-a \cdot 0) = a \cdot 0 \quad (14.1.7)$$

Thus,  $a \cdot 0 = 0$ .  $\square$

If  $1 = 0$ , then  $a = a \cdot 1 = a \cdot 0 = 0$ , and thus every element is zero. A very boring field.

**Theorem 14.1.5.** *In a field  $\langle F, +, \cdot \rangle$ , if  $0 \neq 1$ , then 0 has no inverse.*

*Proof.* For let  $a$  be such an inverse. Then  $a \cdot 0 = 1$ . But for any element of  $F$ ,  $a \cdot 0 = 0$ . But  $0 \neq 1$ , a contradiction.  $\square$

**Theorem 14.1.6.** *If  $a + b = 0$ , then  $b = (-1) \cdot a$  where  $(-1)$  is the solution to  $1 + (-1) = 0$ .*

*Proof.*  $a + (-1)a = a(1 + (-1)) = a \cdot 0 = 0$ . From uniqueness,  $b = (-1)a$ . We may thus write additive inverses as  $-a$ .  $\square$

**Definition 14.1.8** Given two fields  $(F, +, \cdot)$  and  $(F', +', \times)$ , a bijection function  $f : F \rightarrow F'$  is said to be a field isomorphism if and only if for all elements  $a, b \in F$ ,  $f(a + b) = f(a) +' f(b)$ , and  $f(a \cdot b) = f(a) \times f(b)$

**Definition 14.1.9**  $(F, +, \cdot)$  and  $(F', +', \times)$ , are said to be isomorphic if and only if they have an isomorphism.

**Theorem 14.1.7.** *If  $(F, +_F, \cdot_F)$  is a field, if  $(K, +_K, \cdot_K)$  is a field, and if  $\phi : F \rightarrow K$  is a field isomorphism, then  $\phi(1_F) = 1_K$  and  $\phi(0_F) = \phi(0_K)$ .*

*Proof.* For suppose not. If  $\phi(1_F) \neq 1_K$ , then  $\phi(1_F)$  is not the unital element of  $K$  and thus there exists  $y \in K$  such that  $\phi(1_F) \cdot_K y \neq y$ . But since  $\phi$  is an isomorphism, it is bijective and thus there is an  $x \in F$  such that  $\phi(x) = y$ . But since  $\phi$  is an isomorphism, we have:

$$y = \phi(x) = \phi(x \cdot_F 1_F) = \phi(x) \cdot_K \phi(1_F) \quad (14.1.8)$$

A contradiction. Thus,  $\phi(1_F) = 1_K$ . Similarly for  $0_F$ .  $\square$

**Theorem 14.1.8.** *In a field  $(F, +, \cdot)$ ,  $(a + b)^2 = a^2 + 2ab + b^2$  (2 being the solution to  $1 + 1$ ).*

*Proof.* For:

$$(a + b)^2 = (a + b)(a + b) \quad (14.1.9)$$

$$= a(a + b) + b(a + b) \quad (14.1.10)$$

$$= a^2 + ab + ba + b^2 \quad (14.1.11)$$

$$= a^2 + ab(1 + 1) + b^2 \quad (14.1.12)$$

$$= a^2 + 2ab + b^2 \quad (14.1.13)$$

$\square$

# **Part VI**

# **Modules**



# CHAPTER 15

## Elementary Module Theory

Modules generalize the notion of a vector space that one might find in a linear algebra or multivariable calculus course. More than being a generalization for the sake of generalizing, they are found abundantly in the wild and are used in both geometry and algebra. This chapter is devoted to developing the basics of modules, providing definitions, theorems, and plenty of examples.

### 15.1 Definitions

The first thing to do is to define a module over a given ring.

#### Definition 15.1.1: Modules

A **module** on a **ring**  $(R, +_R, \cdot_R)$  is an **Abelian group**  $(M, +)$  with a **function**  $\star : R \times M \rightarrow M$  such that for all  $r_1, r_2 \in R$  and for all  $m_1, m_2 \in M$ , the following are true:

$$r_1 \star (m_1 + m_2) = (r_1 \star m_1) + (r_1 \star m_2) \quad (\text{Scalar Distributivity})$$

$$(r_1 +_R r_2) \star m_1 = (r_1 \star m_1) +_M (r_2 \star m_1) \quad (\text{Module Distributivity})$$

$$r_1 \star (r_2 \star m_1) = (r_1 \cdot_R r_2) \star m_1 \quad (\text{Associativity})$$

$$1 \star m_1 = m_1 \quad (\text{Identity})$$

Where  $1$  is the unital element of  $R$ . We denote a module by  $(M, +, \star)$ .

Note that, while we've called these properties *associative* and *distributive*, we are not using these words as defined in Def. 5.2.2 and Def. 5.2.14, respectively. Associativity is a property of a binary operation, and  $\star$  is not a binary opera-

tion. Recall that a binary operation on a set  $A$  is a function  $* : A \times A \rightarrow A$ . Here, we've defined a function  $\star : R \times M \rightarrow M$ . Since, in general, it may not be true that  $R = M$ , it is not generally true that  $\star$  is a binary operation. These equations and definitions seem to *mimic* binary operations, and hence it seems appropriate to attribute these notions with the same words. Thus, we reuse the terms associativity and distributivity and hope that no confusion arises from this.

**Example 15.1.1** If  $(R, +_R, \cdot_R)$  is a *field*, and if  $(M, +, \star)$  is a module over this field, then it will also be a *vector space*. That is, modules are the ring-analog of vector spaces. Vectors spaces will be discussed in detail later.

**Example 15.1.2** If  $(R, +, \cdot)$  is a ring, then it can be thought of as a module over itself. That is, define  $(R, +, \star)$  trivially as follows:

$$r_1 + r_2 = r_1 + r_2 \quad (15.1.1a) \qquad r_1 \star r_2 = r_1 \cdot r_2 \quad (15.1.1b)$$

From this we have that  $R$  is a module over itself. This follows since by the definition of a *ring*,  $(R, +)$  is an Abelian group and  $(R, \cdot)$  is a *monoid* such that  $\cdot$  is a distributive operation over  $+$  (Here we are using the word distributive correctly. That is, as defined by Def. 5.2.14). Hence we obtain scalar and module distributivity, as well as associativity (monoids are associative). The identity property is simply a restatement of the fact that 1 is a unital element of the monoid  $(R, \cdot)$ .

For the sake of using Ex. 15.1.2 in later theorems, we take the time to prove it.

**Theorem 15.1.1.** *If  $(R, +, \cdot)$  is a ring, then  $(R, +, \star)$  is a module over  $R$ .*

*Proof.* For if  $(R, +, \cdot)$  is a ring, then  $(R, +)$  is an Abelian group and  $(R, \cdot)$  is a monoid such that  $\cdot$  distributed over  $+$  (Def. 16.3.1). But then  $\cdot$  is an associative binary operation (Def. 11.1.2), and hence for all  $r_1, r_2 \in R$  and for all  $m_1, m_2 \in R$ , we have:

$$\begin{aligned} r_1 \cdot (m_1 + m_2) &= r_1 \cdot m_1 + r_1 \cdot m_2 && (\text{Associativity}) \\ (r_1 + r_2) \cdot m_1 &= r_1 \cdot m_1 + r_2 \cdot m_1 && (\text{Distributivity}) \\ r_1 \cdot (r_2 \cdot m_1) &= (r_1 \cdot r_2) \cdot m_1 && (\text{Associativity}) \\ 1 \cdot m_1 &= m_1 && (\text{Identity}) \end{aligned}$$

Thus,  $(R, +, \star)$  is a module over  $(R, +, \cdot)$  (Def. 15.1.1).  $\square$

**Example 15.1.3** Another example arises from the polynomial ring  $R[x]$  over a commutative ring  $(R, +_R, \cdot_R)$ . Define  $\star : R \times R[x] \rightarrow R[x]$  by:

$$r \star m(x) = r \star \left( \sum_{k=1}^n a_k x^k \right) = \sum_{k=1}^n (r \cdot_R a_k) x^k \quad (15.1.2)$$

where  $\cdot_R$  is ring multiplication and the  $a_k$  are elements of  $R$ . Stated another way, given a finitely supported sequence  $a : \mathbb{N} \rightarrow R$ , we define  $r \star a : \mathbb{N} \rightarrow R$  by:

$$(r \star a)_n = r \cdot_R a_n \quad n \in \mathbb{N} \quad (15.1.3)$$

Module addition is then just sequence addition. Given two finitely supported sequences  $a, b : \mathbb{N} \rightarrow R$ , we define  $(a + b) : \mathbb{N} \rightarrow R$  by:

$$(a + b)_n = a_n +_R b_n \quad (15.1.4)$$

Since  $a$  and  $b$  are finitely supported there exists  $N_1, N_2$  such that for all  $n \in \mathbb{N}$  with  $n > N_1$  it is true that  $a_n = 0$ , and similarly for all  $n > N_2$  it is true that  $b_n = 0$ . Thus, for all  $n \in \mathbb{N}$  such that  $n > \max\{N_1, N_2\}$  we have that  $a_n +_R b_n = 0 + 0 = 0$ , hence  $a + b$  is finitely supported. Hence  $+$  is a well defined binary operation on  $R[x]$ , and since  $(R, +, \cdot)$  is a commutative ring,  $+_R$  is a commutative operation, which in turn implies that  $+$  is a commutative operation, and thus  $(R[x], +)$  an Abelian group. Then  $(R, +, \star)$  is a module over  $(R, +_R, \cdot_R)$ .

**Example 15.1.4** We can also think about modules on  $R[x]$  itself, since  $R[x]$  is a ring. If we consider the case where  $R$  is itself a field, then any module over  $R[x]$  will be equivalent to a vector space over  $R$  that is equipped with a *linear transformation* from the vector space to itself.

**Example 15.1.5** If  $(G, *)$  is an Abelian group, then it may be thought of as a module over the ring of integers  $(\mathbb{Z}, +, \cdot)$ . Define  $\star : \mathbb{N} \times G \rightarrow G$  by:

$$1 \star g = g \quad (n+1) \star g = (n \star g) * g \quad (15.1.5)$$

For negative elements, define the following:

$$(-n) \star g = (n \star g)^{-1} \quad (15.1.6)$$

where  $(n \star g)^{-1}$  is the group inverse element of  $n \star g$ . This is well defined since  $n \star g$  is an element of  $G$  and  $G$  is a group. Trivially, define:

$$0 \star g = e \quad (15.1.7)$$

where  $e$  is the unital element of the group  $(G, *)$ . With this we  $(G, *, \star)$  is a module over  $(\mathbb{Z}, +, \cdot)$ .

It is worthwhile to prove the claim that  $(G, *, \star)$  is a module over  $(\mathbb{Z}, +, \cdot)$  since this result is used a lot. The proof is straight forward, but laborious.

### Theorem 15.1.2: Abelian Groups as $\mathbb{Z}$ -Modules

If  $(G, *)$  is an Abelian group, then there is a function  $\star : \mathbb{Z} \times G \rightarrow G$  such that  $(G, *, \star)$  is a module over  $(\mathbb{Z}, +, \cdot)$ .

$$(n + m) \star g = (n + m - 1 + 1) \star g \quad (15.1.8a)$$

$$= ((n + m - 1) \star g) * g \quad (15.1.8b)$$

$$= ((n \star g) * ((m - 1) \star g)) * g \quad (15.1.8c)$$

$$= (n \star g) * (((m - 1) \star g) * g) \quad (15.1.8d)$$

$$= (n \star g) * ((m - 1 + 1) \star g) \quad (15.1.8e)$$

$$= (n \star g) * (m \star g) \quad (15.1.8f)$$

For negatives we have:

$$((-n) + (-m)) \star g = ((n + m) \star g)^{-1} \quad (15.1.9a)$$

$$= ((n \star g) * (m \star g))^{-1} \quad (15.1.9b)$$

$$= (n \star g)^{-1} * (m \star g)^{-1} \quad (15.1.9c)$$

This last equality comes from the fact that  $(G, *)$  is an Abelian group, and therefore  $(a * b)^{-1} = a^{-1} * b^{-1}$ . Continuing, we obtain:

$$((-n) + (-m)) \star g = (n \star g)^{-1} * (m \star g)^{-1} \quad (15.1.10a)$$

$$= ((-n) \star g) * ((-m) \star g) \quad (15.1.10b)$$

And thus we have the distributive law holds again. Next we need to check for when we have one positive and one negative. We get:

$$((n + 1) + (-m)) \star g = ((n + (-m)) + 1) \star g \quad (15.1.11a)$$

$$= ((n + (-m)) \star g) * g \quad (15.1.11b)$$

$$= ((n \star g) * ((-m) \star g)) * g \quad (15.1.11c)$$

Since  $(G, *)$  is an Abelian group, we can simplify this further to get:

$$((n + 1) + (-m)) \star g = ((n \star g) * ((-m) \star g)) * g \quad (15.1.12a)$$

$$= ((n \star g) * g) * ((-m) \star g) \quad (15.1.12b)$$

$$= ((n + 1) \star g) * ((-m) \star g) \quad (15.1.12c)$$

If either  $n$  or  $m$  is zero, the identity holds trivially. Thus we have the preservation of module distributivity. For the scalar distributive law, we have by

induction (and since  $(G, *)$  is Abelian):

$$(n+1) \star (g * h) = (n \star (g * h)) * (g * h) \quad (15.1.13a)$$

$$= (n \star g) * (n \star g) * (g * h) \quad (15.1.13b)$$

$$= ((n \star g) * g) * ((n \star h) * h) \quad (15.1.13c)$$

$$= ((n+1) \star g) * ((n+1) \star h) \quad (15.1.13d)$$

And therefore the distributive law over modules holds for positive integers. For negative integers we have:

$$(-n) \star (g * h) = (n \star (g * h))^{-1} \quad (15.1.14a)$$

$$= ((n \star g) * (n \star h))^{-1} \quad (15.1.14b)$$

$$= (n \star g)^{-1} * (n \star h)^{-1} \quad (15.1.14c)$$

$$= ((-n) \star g) * ((-n) \star h) \quad (15.1.14d)$$

For  $n = 0$  this is trivial since the product is simply the identity element of  $G$ . The last thing to check is the compatibility of ring multiplication with  $\star$ . If  $n, m \in \mathbb{N}$ , we have:

$$n \star ((m+1) \star g) = n \star ((m \star g) * g) \quad (15.1.15a)$$

$$= (n \star (m \star g)) * (n \star g) \quad (15.1.15b)$$

$$= ((n \cdot m) \star g) * (n \star g) \quad (15.1.15c)$$

$$= (n \cdot m + n) \star g \quad (15.1.15d)$$

$$= (n \cdot (m+1)) \star g \quad (15.1.15e)$$

The structure of a  $\mathbb{Z}$  module is different from a vector space. In  $\mathbb{Z}$ ,  $\{2\}$  is a linearly independent set that cannot be extended to a basis. This is because for all  $n \in \mathbb{Z}$  such that  $n \neq 2$  we have that  $0 = x \cdot 2 + (-2) \cdot x$ , and thus the set  $\{2, x\}$  is linearly dependent. This is contrary to the usual notions of vector spaces where a linearly independent set can always be extended to a basis. Moreover,  $\{2, 3\}$  is a generated set of  $\mathbb{Z}$  but no subset of  $\{2, 3\}$  is a basis. To see that it generates  $\mathbb{Z}$ , note that  $1 = 1 \cdot 3 - 1 \cdot 2$ , and 1 generated  $\mathbb{Z}$ . We've already seen that  $\{2\}$  is not a basis and cannot be extended to form one, and neither can 3.

**Theorem 15.1.3.** *If  $(M, +, \star)$  is a module over a ring  $(R, +, \cdot)$  and if  $r \in R$ , then the function  $f : M \rightarrow M$  defined by:*

$$f(m) = r \star m \quad (15.1.16)$$

*is a group endomorphism on the group  $(M, +)$ .*

*Proof.* For let  $m_1, m_2 \in M$ . But since  $(M, +, \star)$  is a module,  $\star$  left-distributes over  $+$  (Def 15.1.1). Therefore:

$$f(m_1 + m_2) = r \star (m_1 + m_2) = (r \star m_1) + (r \star m_2) = f(m_1) + f(m_2) \quad (15.1.17)$$

Thus,  $f$  is a group endomorphism on  $(M, +)$ .  $\square$

**Theorem 15.1.4.** *If  $(M, +, \star)$  is a module over a ring  $(R, +, \cdot)$ , if  $\text{End}(M)$  is the set of all group endomorphisms on  $(M, +)$ , if  $+$ ' denotes function addition, and if  $\circ$  denotes function composition, then there exists a ring homomorphism  $\varphi : R \rightarrow (\text{End}(M), +, \circ)$ .*

*Proof.* For let  $\varphi : R \rightarrow \text{End}(M, +, \star)$  be defined by the function that maps  $r \in R$  to the function  $\varphi_r$  by:

$$\varphi_r(m) = r \star m \quad m \in M \quad (15.1.18)$$

By Thm. 15.1.3, for all  $r \in R$ ,  $\varphi_r \in \text{End}(M, +, \star)$ . Moreover,  $\varphi$  is a ring homomorphism. For if  $r_1, r_2 \in R$ , then for all  $m \in M$ :

$$\begin{aligned} \varphi_{r_1+r_2}(m) &= (r_1 + r_2) \star m && \text{(Definition of } \varphi\text{)} \\ &= (r_1 \star m) + (r_2 \star m) && \text{(Module Distributivity)} \\ &= \varphi_{r_1}(m) + \varphi_{r_2}(m) && \text{(Definition of } \varphi\text{)} \\ &= (\varphi_{r_1} +' \varphi_{r_2})(m) && \text{(Definition of } +'\text{)} \end{aligned}$$

And thus  $\varphi$  preserves addition. Moreover:

$$\begin{aligned} \varphi_{r_1 \cdot r_2}(m) &= (r_1 \cdot r_2) \star m && \text{(Definition of } \varphi\text{)} \\ &= r_1 \star (r_2 \star m) && \text{(Associativity)} \\ &= r_1 \star (\varphi_{r_2}(m)) && \text{(Definition of } \varphi\text{)} \\ &= \varphi_{r_1}(\varphi_{r_2}(m)) && \text{(Definition of } \varphi\text{)} \\ &= (\varphi_{r_1} \circ \varphi_{r_2})(m) && \text{(Definition of } \circ\text{)} \end{aligned}$$

And thus  $\varphi$  preserves multiplication. Lastly:

$$\begin{aligned} \varphi_{1_R}(m) &= 1_R \star m && \text{(Definition of } \varphi\text{)} \\ &= m && \text{(Identity)} \end{aligned}$$

And thus  $\varphi_1 = \text{id}_M$ , and  $\text{id}_M$  is the unital element of  $M$ . Thus  $\varphi$  preserves identity, and therefore  $\varphi$  is a ring homomorphism (Def. 13.2.1).  $\square$

The converse of this theorem is true as well. That is, if  $(M, +)$  is an Abelian group and if  $\varphi : (R, +, \cdot) \rightarrow (\text{End}(M), +, \circ)$  is a ring homomorphism, then there exists an operation  $\star : R \times M \rightarrow M$  that makes  $(M, +, \circ)$  a module over  $(R, +, \cdot)$ .

**Theorem 15.1.5.** *If  $(R, +, \cdot)$  is a ring, if  $(M, +)$  is an Abelian group, and if  $\varphi : (R, +, \cdot) \rightarrow (\text{End}(M), +, \circ)$  is a ring homomorphism, then there is a function  $\star : R \times M \rightarrow M$  such that  $(M, +, \star)$  is a module over  $(R, +, \cdot)$ .*

*Proof.* For let  $\star : R \times M \rightarrow M$  be defined by:

$$r \star m = \varphi_r(m) \quad (15.1.19)$$

Where  $\varphi_r$  is the endomorphism that  $r \in R$  gets mapped to by  $\varphi$ . Then  $(M, +, \star)$  is a module over  $(R, +, \cdot)$ . For if 1 is the unital element of  $R$ , then  $\varphi_1 = \text{id}_M$ , since  $\varphi$  is a ring homomorphism. But then:

$$1 \star m = \varphi_1(m) = \text{id}_M(m) = m \quad (15.1.20)$$

If  $r_1, r_2 \in R$  and if  $m \in M$ , then:

$$(r_1 + r_2) \star m = \varphi_{r_1+r_2}(m) \quad (15.1.21)$$

But  $\varphi$  is a ring homomorphism and therefore  $\varphi_{r_1+r_2} = \varphi_{r_1} +' \varphi_{r_2}$ , and thus:

$$(r_1 + r_2) \star m = (\varphi_{r_1} +' \varphi_{r_2})(m) = \varphi_{r_1}(m) + \varphi_{r_2}(m) = (r_1 \star m) + (r_2 \star m) \quad (15.1.22)$$

And thus we have module distributivity. For scalars, if  $r \in R$  and if  $m_1, m_2 \in M$ , then since for all  $r \in R$  it is true that  $\varphi_r$  is an endomorphism on  $(M, +)$ , we have:

$$r \star (m_1 + m_2) = \varphi_r(m_1 + m_2) = \varphi_r(m_1) + \varphi_r(m_2) = (r \star m_1) + (r \star m_2) \quad (15.1.23)$$

And thus the distributive law for scalars is upheld. Lastly, to check for the compatibility of multiplication. If  $r_1, r_2 \in R$  and if  $m \in M$ , then:

$$r_1 \star (r_2 \star m) = r_1 \star (\varphi_{r_2}(m)) = \varphi_{r_1}(\varphi_{r_2}(m)) = (\varphi_{r_1} \circ \varphi_{r_2})(m) \quad (15.1.24)$$

But  $\varphi$  is a ring homomorphism, and therefore  $\varphi_{r_1 \cdot r_2} = \varphi_{r_1} \circ \varphi_{r_2}$ . Thus, we obtain:

$$r \star (m_1 + m_2) = (r \star m_1) + (r \star m_2) = \varphi_{r_1 \cdot r_2}(m) = (r_1 \cdot r_2) \star m \quad (15.1.25)$$

Thus,  $(M, +, \star)$  is a module over  $(R, +, \cdot)$ . □

### Example 15.1.6: Another Example of a Module

Let  $(V, +, \cdot)$  be a finite dimensional vector space over the field  $(\mathbb{F}, +, \cdot)$  and let  $T : V \rightarrow V$  be a linear operator and let  $\mathbb{F}[x]$  be the polynomial ring with coefficients in  $\mathbb{F}$ . We can define a module structure over  $V$  by letting  $\star : \mathbb{F}[x] \times V \rightarrow V$  be defined by:

$$f \star v = \sum_{k=0}^n a_k T^k(v) \quad (15.1.26)$$

where  $a_k$  are the coefficients of the polynomial  $f \in \mathbb{F}[x]$  and where  $T^k$  is the  $k^{th}$  composition of  $T$  with itself. That is,  $T^2 = T \circ T$ ,  $T^{n+1} = T^n \circ T$ . Letting  $+$ ' denote polynomial addition, we have that  $(V, +', \star)$  is a module over  $\mathbb{F}[x]$  with its usual ring structure. ■

### Definition 15.1.2: Module Homomorphism

A **module homomorphism** from a **module**  $(M_1, +, \star)$  to another module  $(M_2, +', \diamond)$  over a **ring**  $(R, +, \cdot)$  is a **function**  $f : M_1 \rightarrow M_2$  such that, for all  $x, y \in M_1$ , and for all  $r \in R$ , the following are true:

$$f(x + y) = f(\mathbf{x}) +' f(\mathbf{y}) \quad (\text{Preservation of Addition})$$

$$f(r \star x) = r \diamond f(x) \quad (\text{Preservation of Scalar Multiplication})$$

**Example 15.1.7** If  $(M_1, +, \star)$  and  $(M_2, +', \diamond)$  are modules over  $(R, +, \cdot)$ , if if  $f : M_1 \rightarrow M_2$  is the zero map:  $f(m) = 0$  for all  $m \in M_1$ , then  $f$  is a module homomorphism. As a non-trivial example, given a field  $(\mathbb{F}, +, \cdot)$  and two vector spaces  $(V_1, +, \cdot)$ ,  $(V_2, +', \cdot')$  any linear transformation  $T : V_1 \rightarrow V_2$  is a module homomorphism. Preservation of addition comes from the fact that  $T$  is linear. That is:

$$T(\mathbf{x} + \mathbf{y}) = T(x) + T(y) \quad (15.1.27)$$

The preservation of scalar multiplication also holds since for linear transformations we have:

$$T(a \cdot \mathbf{x}) = a \cdot T(\mathbf{x}) \quad (15.1.28)$$

Thus  $T$  is a module homomorphism. Remember that any vector space over a field can be thought of as a module over the underlying field, since any field is also a ring. It is in this sense that  $T$  may be thought of as a module homomorphism.

**Example 15.1.8** We've seen that any ring  $(R, +, \cdot)$  can be seen as a module over itself. We've also seen that the ring of endomorphisms over  $R$  can be given a module structure as well by letting  $+$ ' denote the sum of two endomorphisms and  $\star$  denoting the product by scalar elements of  $R$ . That is,  $r \star f$  is the map defined by  $(r \star f)(s) = r \cdot f(s)$ , for all  $s \in R$ . There is bijective homomorphism between these two modules. For let  $\varphi : R \rightarrow \text{End}(R)$  be defined by:

$$(\varphi(r))(s) = r \cdot s \quad (15.1.29)$$

Then  $\varphi$  is a module homomorphism. For we have:

$$(\varphi(r + s))(t) = (r + s) \cdot t = (r \cdot t) + (s \cdot t) = (\varphi(r))(t) + (\varphi(s))(t) \quad (15.1.30)$$

And thus addition is preserved. Moreover, if  $r, s \in S$ , then:

$$(\varphi(r \cdot s))(t) = (r \cdot s) \cdot t = r \cdot (s \cdot t) = r \cdot (\varphi(s))(t) \quad (15.1.31)$$

And thus  $\varphi$  is a module homomorphism. Moreover, the identity is mapped to the identity, for:

$$(\varphi(1))(s) = 1 \cdot s = \text{id}_R(s) \quad (15.1.32)$$

We will later show that this homomorphism is bijective.

And from this we can define a module isomorphism.

### Definition 15.1.3: Module Isomorphism

A **module isomorphism** from a **module**  $(M_1, +, \star)$  to a **module**  $(M_2, +', \diamond)$  over a **ring**  $(R, +, \cdot)$  is a **module homomorphism**  $f$  such that  $f$  is a **bijective function**.

**Example 15.1.9** Using the previous example we see with a ring  $(R, +, \cdot)$  acting as a module over itself, and the ring of endomorphisms  $(\text{End}(R), +', \diamond)$  given the module structure  $(r \star f)(s) = r \cdot f(s)$ , are homomorphic. The homomorphism given, mapping  $r$  to the endomorphism  $f(s) = r \cdot s$ , is bijective. It is injective for:

$$(\varphi(r))(t) = (\varphi(s))(t) \Leftrightarrow r \cdot t = s \cdot t \Leftrightarrow (r - s) \cdot t = 0 \quad (15.1.33a)$$

But if  $(r - s) \cdot t = 0$  for all  $t \in R$ , then  $r - s = 0$  (Thm. 13.1.3) and therefore  $r = s$ . Moreover every element of  $\text{End}(R)$  is mapped to. For let  $f \in \text{End}(R)$ . Then since  $f$  is a homomorphism,  $f(1) = 1$ , so we have:

$$f(s) = f(s \cdot 1) = s \cdot f(1) = s \cdot r \quad (15.1.34)$$

### Definition 15.1.4: Submodule

A **submodule** of a **module**  $(M, +, \star)$  over a **ring**  $(R, +, \cdot)$  is a **subgroup**  $(N, +)$  of  $(M, +)$  such that, for all  $n \in N$  and for all  $r \in R$  it is true that  $r \star n \in N$ .

**Definition 15.1.5: Linear Transformation**

A linear transformation from a module  $(M_1, +, \star)$  to a module  $(M_2, +', \diamond)$  over a ring  $(R, +, \cdot)$  is a function  $T : M_1 \rightarrow M_2$  such that, for all  $m_1, m_2 \in M_1$  and for all  $r_1, r_2 \in R$ , the following is true:

$$f((r_1 \star m_1) + (r_2 \star m_2)) = (r_1 \diamond f(m_1)) +' (r_2 \diamond f(m_2))$$

**Definition 15.1.6: Invariant Subspace**

And invariant subspace of a vector space  $(V, +, \cdot)$  over a field  $(\mathbb{F}, +, \cdot)$  under a linear transformation  $T : V \rightarrow V$  is a subspace  $W \subseteq V$  such that the image of  $W$  under  $T$  is a subset of  $W$ . That is,  $T(W) \subseteq W$ .

**Theorem 15.1.6.** *If  $(V, +, \cdot)$  is a vector space over a field  $(\mathbb{F}, +, \cdot)$ , if  $(\mathbb{F}[x], +', *)$  is the ring of polynomials over  $\mathbb{F}$ , if  $T : V \rightarrow V$  is a linear transformation, if  $\star : \mathbb{F}[x] \times V \rightarrow V$  is defined by:*

$$f \star v = \sum_{k=1}^n a_k T^k(v) \quad (15.1.35)$$

where  $a_k$  are the coefficients of  $f \in \mathbb{F}[x]$  and  $T^k$  is the  $k^{th}$  composition of  $T$  with itself, then a subset  $W \subseteq V$  is a submodule of  $V$  over  $\mathbb{F}[x]$  if and only if  $W$  is an invariant subspace of  $V$  under  $T$ .

*Proof.* For let  $W$  be a submodule of  $V$ , let  $w \in W$ , and let  $f$  be the polynomial  $f(x) = x$ . But then  $f \star w = T(w)$ , and since  $W$  is a submodule, it is true that  $f \star w \in W$ . But then  $T(w) \in W$ , and therefore  $T(W) \subseteq W$ . Thus,  $W$  is an invariant subspace under  $T$  (Def. 15.1.6). Now suppose  $W$  is an invariant subspace of  $V$  under  $T$ . Then for all  $n \in \mathbb{N}$ ,  $T^n(W) \subseteq W$  since by induction  $T^n(W) = T(T^{n-1}(W)) \subseteq T(W) \subseteq W$ . Let  $f \in \mathbb{F}[x]$  with coefficients  $a_k$  and let  $w \in W$ . Since, for all  $n \in \mathbb{N}$  it is true that  $T^n(w) \in W$ , and since  $W$  is a subspace of  $V$ , for all  $k$  it is true that  $a_k \cdot T^n(w) \in W$ . Therefore  $f \star w \in W$ . That is,  $W$  is a submodule of  $V$  over  $\mathbb{F}[x]$ .  $\square$

If we have a finite dimensional vector space  $V$  and a subspace  $W$  with basis  $\mathcal{B}'$ , we can extend this to a basis  $\mathcal{B}$ . The representiting matrix then has the form:

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} [T|_W]_{\mathcal{B}'}^{\mathcal{B}'} & * \\ 0 & * \end{bmatrix} \quad (15.1.36)$$

**Theorem 15.1.7.** If  $(R_1, +, \cdot)$  and  $(R_2, +', *)$  are rings, and if  $(M, +, \star)$  is a module over  $(R_2, +', *)$ , and if  $\varphi : R_1 \rightarrow R_2$  is a ring homomorphism, then there is a function  $\diamond : R_1 \times M \rightarrow M$  such that  $(M, +, \diamond)$  is a module over  $(R, +, \cdot)$ .

*Proof.* For let  $\diamond : R_1 \times M \rightarrow M$  be defined by:

$$r \diamond m = \varphi(r) \star m \quad (15.1.37)$$

The  $(M, +, \diamond)$  is a module over  $(R, +, \cdot)$ . For since  $(M, +, \diamond)$  is a module over  $(R, +', *)$ ,  $(M, +)$  is an Abelian group (Def. 15.1.1). Moreover, if  $r_1, r_2 \in R_1$ , and if  $m \in M$ , then:

$$\begin{aligned} (r_1 + r_2) \diamond m &= \varphi(r_1 + r_2) \star m && \text{(Definition of } \diamond\text{)} \\ &= (\varphi(r_1) +' \varphi(r_2)) \star m && (\varphi \text{ is a Homomorphism}) \\ &= (\varphi(r_1) \star m) + (\varphi(r_2) \star m) && \text{(Distributive Property)} \\ &= (r_1 \diamond m) + (r_2 \diamond m) && \text{(Definition of } \diamond\text{)} \end{aligned}$$

And thus  $\diamond$  is distributive over module elements. If  $r \in R_1$  and if  $m_1, m_2 \in M$ , then:

$$\begin{aligned} r \diamond (m_1 + m_2) &= \varphi(r) \star (m_1 + m_2) && \text{(Definition of } \diamond\text{)} \\ &= (\varphi(r) \star m_1) + (\varphi(r) \star m_2) && \text{(Distributive Property)} \\ &= (r \diamond m_1) + (r \diamond m_2) && \text{(Definition of } \diamond\text{)} \end{aligned}$$

Also, the identity is preserved. For let  $\mathbb{I}_1$  and  $\mathbb{I}_2$  be the unital elements of  $(R_1, \cdot)$  and  $(R_2, *)$ , respectively. Since  $\varphi$  is a ring homomorphism,  $\varphi(\mathbb{I}_1) = \mathbb{I}_2$ , and therefore:

$$\begin{aligned} \mathbb{I}_1 \diamond m &= \varphi(\mathbb{I}_1) \star m && \text{(Definition of } \diamond\text{)} \\ &= \mathbb{I}_2 \star m && \text{(Identity Property of Homomorphisms)} \\ &= m && \text{(Identity Property of Modules)} \end{aligned}$$

And thus  $\diamond$  preserves the identity. Lastly,  $\diamond$  is compatible with scalar multiplication.

$$\begin{aligned} r_1 \diamond (r_2 \diamond m) &= \varphi(r_1) \star (\varphi(r_2) \star m) && \text{(Definition of } \diamond\text{)} \\ &= (\varphi(r_1) * \varphi(r_2)) \star m && \text{(Module Compatibility with Multiplication)} \\ &= (r_1 \cdot r_2) \diamond m && \text{(Definition of } \diamond\text{)} \end{aligned}$$

Therefore,  $(M, +, \diamond)$  is a module over  $(R_1, +, \cdot)$  (Def. 15.1.1).  $\square$

## 15.2 Old

An attempt has been made to preserve the differences between the various operations in a left module.  $+$  and  $\cdot$  are binary operations that act on elements of  $R$ . That is, for  $a, b \in R$ ,  $a + b$  gives another element of  $R$ , as does  $a \cdot b$ . However,  $+_M$  is a binary operation over  $M$ . If  $a, b \in R$ ,  $a +_M b$  has no meaning. For  $a, b \in M$ ,  $a +_M b$  is well defined, and returns another element of  $M$ . The “function,”  $*$  takes an ordered pair  $(r, a)$ , where  $r \in R$  and  $a \in M$ , and returns another element in  $M$ . For convenience we write  $r * a$ . If  $a, b \in R$ , then  $a * b$  has no meaning, and if  $a, b \in M$  then  $a * b$  also has no meaning. Usually this is very unimportant, and  $+$  and  $+_M$  are given the same symbol, as are  $\cdot$  and  $*$ . We can then more loosely rewrite the definition as, for all  $r, s \in R$ , and all  $a, b \in M$ :

1.  $r(a + b) = ra + rb$
2.  $(rs)(a) = r(sa)$
3.  $(r + s)(a) = (ra) + (rs)$
4.  $1a = a$

This is the more natural notation one finds when defining vector spaces. A module is analogous to a vector space: In a vector space one has a set  $V$  and a *field*  $K$ , whereas in a module one has a set  $M$  and a *ring with identity*  $R$ .

**Theorem 15.2.1.** *If  $(G, +)$  is an Abelian group, then there is a function  $* : \mathbb{Z} \times G \rightarrow G$  such that  $(G, +_G)$  is a left module of  $(\mathbb{Z}, +, \cdot)$ , where  $+$  and  $\cdot$  are the standard arithmetic operations over  $\mathbb{Z}$ .*

*Proof.* For define  $0 * a = e$  and  $1 * a = a$  for all  $a \in G$ , and for all  $n \in \mathbb{Z}$ ,  $n > 1$  inductively define  $(n+1)*a = n*a+_G a$ . For  $n < 0$  define  $n*a = ((-n)*a)^{-1}$ , where the inverse is taken with respect to the group  $G$ . Then  $*$  is a function  $* : \mathbb{Z} \times G \rightarrow G$ . If  $n > 0$ , we have the following:

$$\begin{aligned} (n * a) +_G (n * b) &= \underbrace{(a +_G \cdots +_G a)}_n +_G \underbrace{(b +_G \cdots +_G b)}_n \\ &= \underbrace{(a + b) +_G \cdots +_G (a + b)}_n \\ &= n * (a + b) \end{aligned}$$

If  $n, m > 0$ , we have:

$$n * a +_G m * a = \underbrace{a +_G \cdots +_G a}_{n+m} = (n + m) * a$$

And finally:

$$(n \cdot m) * a = \underbrace{a +_G \cdots +_G a}_{n \cdot m} = n * \underbrace{(a +_G \cdots +_G a)}_m = n * (m * a)$$

Similarly for when  $n, m < 0$ ,  $n < 0 < m$ , or  $m < 0 < n$ .  $\square$

Thus, every Abelian group  $(G, +_G)$  can be seen as a left module over  $(R, +, \cdot)$ . Moreover the function  $*$  is unique, so this correspondence is unique as well. As another example, every vector space  $V$  over a field  $K$  is a left module over  $K$ , since any field  $K$  is also a ring with identity. A **Left Ideal** of a ring  $(R, +, *)$  is a subset  $I \subseteq R$  such that:

1.  $\forall_{a,b \in I}, a + b \in I$
2.  $\forall_{r \in R} \forall_{a \in I}, r \cdot a \in I$

This can be rephrased by saying that  $(I, +)$  is a subgroup of  $(R, +)$ , and  $I$  absorbs left-multiplication. A **Right Ideal** replaces  $r \cdot a$  with  $a \cdot r$ . An **Ideal** or **Two-Sided Ideal** is a subset that is both a left and a right ideal.

**Theorem 15.2.2.** *If  $(R, +, \cdot)$  is a ring with identity and  $(I, +)$  is a left ideal of  $R$ , then there is a function  $* : R \times I \rightarrow I$  such that  $(I, +)$  is a left module over  $R$ .*

*Proof.* For let  $*$  be the restriction of  $\cdot$  to  $R \times I$ . Then, for all  $a \in I$ ,  $1 \cdot a = a$ . If  $a, b \in I$  and  $r \in R$ , then:

$$r * (a + b) = r \cdot (a + b) = (r \cdot a) + (r \cdot b)$$

If  $r, s \in R$  and  $a \in I$ , then:

$$\begin{aligned} (r \cdot s) * a &= (r \cdot s) \cdot a = r \cdot (s \cdot a) \\ (r + s) * (a) &= (r + s) \cdot a = (r \cdot a) + (s \cdot a) \end{aligned}$$

$\square$

**Definition 15.2.1** The Annihilator of a Left Module  $(M, +_M)$  over a ring with identity  $(R, +, \cdot)$  is the set:

$$I = \{r \in R : \forall_{m \in M}, r * m = 0\}$$

**Theorem 15.2.3.** *If  $(M, +_M)$  is a Left Module over a ring with identity  $(R, +, \cdot)$ , and if  $I$  is the annihilator of  $M$ , then  $I$  is a two-sided ideal of  $R$ .*

*Proof.* For if  $r, s \in I$ , then for all  $m \in M$ ,  $r * m = 0$  and  $s * m = 0$ . But  $(r + s) * m = (r * m) +_M (s * m) = 0$ . Therefore  $r + s \in I$ . If  $r \in R$  and  $s \in I$ , then  $(r \cdot s) * m = r * (s * m) = r * 0 = 0$ , and therefore  $r \cdot s \in I$ . Furthermore,  $(s \cdot r) * m = s * (r * m) = 0$ , and thus  $s \cdot r \in I$ . Therefore  $I$  is a two-sided ideal.  $\square$



**Part VII**

# Galois Theory



# CHAPTER 16

---

## Elementary Field Theory

---

Fields are logically the next step in adding structure to algebra. Most of the problems that arise in ring theory stem from the fact that non-zero elements may not be invertible, and also that zero divisors may exist. Field theory rids us of these problems and establishes the 9 familiar properties of arithmetic that one is familiar with from elementary school. While the structure of a field is certainly nice, there can still be some problems that may appear when one lets their intuition get in the way (for example, in the field  $(\mathbb{Z}_2, +, \cdot)$  we have  $1 + 1 = 0$ ).

### 16.1 What's the Point?

From the historical perspective, we can start with numbers. There's the standard inclusion:

$$\mathbb{N}^+ \subseteq \mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C} \quad (16.1.1)$$

Suppose we are given the equation  $2 + x = 1$ . If we only know about the positive integers, then we cannot solve this equation. We thus need to introduce negative integers. Next we could write  $2x = 1$ , and we are now forced to introduce the rational numbers. In ancient Greece, the solution to  $x^2 = 2$  was proved to be irrational, and thus we must go beyond  $\mathbb{Q}$  and develop the real numbers (or at the very least, the algebraic numbers  $\mathbb{A}$ ). Pushing beyond this, polynomial equations such as  $x^2 + 1 = 0$  were studied in Italy during the Renaissance era. It was known that there are no real solutions to this, as one can see from the graph of  $x^2 + 1$  (it never crosses the  $x$  axis). To solve such equations one must invent  $\mathbb{C}$ . The complex numbers are the set  $\mathbb{C}$  of the form:

$$\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\} \quad (16.1.2)$$

where  $i^2 = -1$ , by definition, which is the solution to the equation  $z^2 + 1 = 0$ . This equation has no solutions in  $\mathbb{R}$  and so  $i$  is not a real number, and hence is called the *imaginary* unit. We can picture complex numbers by use of the plane  $\mathbb{R}^2$ . But there's nothing too special about the equation  $z^2 + 1 = 0$ , and we can consider  $z^2 + z + 1 = 0$  and again we can ask if this has real solutions. Unlike the first equation, it's not so obvious that this has no real solution. We can look at the quadratic formula, and in particular the discriminant, obtaining:

$$\Delta = b^2 - 4ac = 1 - 4 = -2 \quad (16.1.3)$$

Since this is negative, there are no real solutions, and hence  $z^2 + z + 1$  has no solution in  $\mathbb{R}$ . It does have roots in  $\mathbb{C}$ :

$$\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \quad (16.1.4a) \qquad \bar{\omega} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i \quad (16.1.4b)$$

We can further consider the set  $\mathbb{R}[\omega]$  defined by:

$$\mathbb{R}[\omega] = \{x + y\omega \mid x, y \in \mathbb{R}\} \quad (16.1.5)$$

This has a nice field structure, like  $\mathbb{C}$ , and indeed this is equal to  $\mathbb{C}$ . That is,  $\mathbb{R}[\omega] = \mathbb{C}$ . We can see this since  $\mathbb{R}[\omega]$  is a subspace of  $\mathbb{C}$  with a basis consisting of two elements:  $\{1, \omega\}$ , and thus has the same dimension as  $\mathbb{C}$ . Hence, it is equal to the whole thing. We can be even more explicit:

$$x + y\omega = x + y\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = \left(x - \frac{1}{2}y\right) + \left(\frac{\sqrt{3}}{2}y\right)i \quad (16.1.6)$$

And this is of the form  $x' + y'i$ , where:

$$x' = x - \frac{1}{2}y \quad (16.1.7a) \qquad y' = \frac{\sqrt{3}}{2}y \quad (16.1.7b)$$

Since this is always solvable for both  $(x, y)$  and  $(x', y')$ , the two spaces are the same. And indeed, we can generalize. If  $f(x) = ax^2 + bx + c$ , with  $a, b, c \in \mathbb{R}$  such that  $b^2 - 4ac < 0$ , then defining:

$$\alpha = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad (16.1.8)$$

which is a complex root of  $f$ , then  $\mathbb{R}[\alpha] = \mathbb{C}$ . This shows there's nothing too special about  $i$ : extending  $\mathbb{R}$  with any complex root of a quadratic gives the entirety of  $\mathbb{C}$ , we need not only choose  $z^2 + 1 = 0$ . Even if we were to stick with this polynomial, we could still choose  $-i$ , since this too is a solution. Choosing  $i$  over  $-i$  seems to purely be an accident of history. Going from one choice to another is an  $\mathbb{C}$  automorphism:  $x + iy \mapsto x - iy$ . An  $\mathbb{R}$  automorphism is a

bijective ring homomorphism  $f : \mathbb{C} \rightarrow \mathbb{C}$ . That is, an isomorphism from  $\mathbb{C}$  to itself:

$$f(z_1 + z_2) = f(z_1) + f(z_2) \quad (16.1.9a) \qquad f(z_1 z_2) = f(z_1)f(z_2) \quad (16.1.9b)$$

And also requiring:

$$f(1) = 1 \quad (16.1.10)$$

The automorphism  $x + iy \mapsto x - iy$  is called complex conjugation. If we don't like  $i$ , and have a complex number such as  $\omega$ , we can still take as an  $\mathbb{R}$  automorphism the function  $\sigma : \mathbb{C} \rightarrow \mathbb{C}$  where  $x + y\omega \mapsto x + y\bar{\omega}$ . As it turns out, this is the same as the automorphism  $x + iy \mapsto x - iy$  since we can write:

$$i = \frac{1 + 2\omega}{\sqrt{3}} \quad (16.1.11)$$

This is the object we wish to stress as the important part of the theory of complex numbers. Neither  $i$  nor  $\omega$  are too important, but rather the notion of complex conjugation is. The group of  $\mathbb{R}$  automorphisms of  $\mathbb{C}$  is equal to:

$$\text{Aut}_{\mathbb{R}}(\mathbb{C}) = \{\text{id}_{\mathbb{R}}, \sigma\} \quad (16.1.12)$$

Where  $\sigma$  is complex conjugation. That is,  $\sigma$  is the unique non-trivial  $\mathbb{C}$  automorphism that has the property that it exchanges the roots of any  $f(x) = ax^2 + bx + c$  with  $b^2 - 4ac < 0$ . The group structure comes from function composition. Since function composition is associative, the identity map is an automorphism, and since bijections have inverse functions, this is indeed a group. We can summarize all of this as follows: The roots of any real polynomial are either real or come in complex conjugate pairs.

Looking at the numerology of the problem, there seems to be something special about the number two. This is the size of the automorphism group  $\text{Aut}_{\mathbb{R}}(\mathbb{C})$ , and this is also the dimension of  $\mathbb{C}$ , and lastly it is the degree of  $\mathbb{C}$  over  $\mathbb{R}$ :  $[\mathbb{C} : \mathbb{R}]$ . More generally, consider any field  $\mathbb{F}$  with characteristic not equal to 2 (that is,  $1 + 1 \neq 0$ ), and any function  $f(x) = ax^2 + bx + c$ ,  $a, b, c \in \mathbb{F}$  such that  $f(x) = 0$  has no solutions in  $\mathbb{F}$ . For example,  $\mathbb{R}$  with  $f(x) = x^2 + 1$ , or  $\mathbb{Q}$  with  $f(x) = x^2 - 2$ . If we have such conditions, then there is a field  $\mathbb{K}$  and an inclusion  $\mathbb{F} \subseteq \mathbb{K}$  making  $\mathbb{F}$  a subfield, such that  $f(x) = a(x - \alpha)(x - \beta)$ , where  $\alpha, \beta \in \mathbb{K}$ . Moreover,  $\mathbb{K} = \mathbb{F}[\alpha]$ . That is:

$$\mathbb{K} = \{x + ya \mid x, y \in \mathbb{F}\} \quad (16.1.13)$$

Similarly,  $\mathbb{K} = \mathbb{F}[\beta]$ . Lastly, the automorphism group is

$$\text{Aut}_{\mathbb{F}}(\mathbb{K}) = \{\text{id}_{\mathbb{F}}, \sigma\} \quad (16.1.14)$$

where  $\sigma$  is the unique automorphism such that  $\sigma(\alpha = \beta)$ . The proof is simply an application of the quadratic formula, where we invoke the fact that  $2 \neq 0$  in a field whose characteristic is not 2.

### 16.1.1 Cubic Equations and Higher

In the 16<sup>th</sup> century the Italians were able to solve the cubic equation:

$$x^3 + px - q = 0 \quad (16.1.15)$$

This may not look like the general cubic, but since we are interested in roots we may always divide off by the leading coefficient of  $x^3$ , and the quadratic term may be absorbed by completing the square, and thus any cubic can be written in such a form. The solution is much less elegant than the quadratic formula:

$$x = \quad (16.1.16)$$

By the 18<sup>th</sup> century the Italians were able to solve the general quartic equation. The next natural question is the solution to the quintic, but this was shown not to exist. The Abel-Ruffini theorem shows that the general quintic equation can not be solved using nested radicals. Galois went to prove that a polynomial has a root that can be written in terms of nested radicals if and only if  $K/F$ , the splitting field, has an automorphism group  $\text{Aut}_F(K)$  that is solveable.

### 16.1.2 Some Reminders

**Definition 16.1.1** A field  $(\mathbb{F}, +, \cdot)$  is an Abelian group  $(\mathbb{F}, +)$  such that  $(\mathbb{F}^* \setminus \{0\}, \cdot)$  is an Abelian group as well. This is the group of *units*.

**Example 16.1.1** The classic examples are  $\mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$ , as well as the finite fields  $\mathbb{F}_p$ , also commonly denoted  $\mathbb{Z}/p\mathbb{Z}$  or simply  $\mathbb{Z}_p$ .

**Definition 16.1.2** A field extension of a field  $F$  is a field  $K$  such that  $F \subseteq K$ . We may also say that  $F$  is a subfield of  $K$ .

We often denote that  $K$  is a field extension of  $F$  by writing  $K/F$ . This is not to denote a quotient or anything of that manner and is simply to denote that  $F$  is a subfield of  $K$ .

**Example 16.1.2**  $\mathbb{C}$  is a field extension of  $\mathbb{R}$  since both are fields and  $\mathbb{R} \subseteq \mathbb{C}$ . We can go backwards, thinking of  $\mathbb{R}$  as a field extension  $\mathbb{R}$ .

Also important, if  $K$  is a field extension of  $F$ ,  $K/F$ , then  $K$  has the structure of an  $F$  vector space. That is,  $K$  can be seen as a vector space over  $F$ . One thing that we write is this bracket notation  $[K : F]$ , which again is not to be confused with the notation found in groups about the cardinality of certain things.  $[K : F]$  is the simply the dimension of the vector space  $K$  over  $F$ :

$$[K : F] = \dim_K(F) \quad (16.1.17)$$

This is also called the degree of the extension  $K/F$ . If the dimension is finite,  $[K : F] < \infty$ , we say that  $K/F$  is a finite extension.

**Example 16.1.3**  $\mathbb{C}$  is a two dimensional vector space over  $\mathbb{R}$  and thus  $[\mathbb{C}, \mathbb{R}] = 2$ . To see this, use  $\{1, i\}$  as a basis.

**Theorem 16.1.1.** *Any countable dimensional vector space over a countable field is also countable.*

**Example 16.1.4** Using this theorem shows that  $\mathbb{R}$ , as a vector space over  $\mathbb{Q}$ , is not only an infinite dimensional vector space, but also has an uncountably infinite basis. Thus,  $[\mathbb{R} : \mathbb{Q}]$  is uncountably infinite.

**Example 16.1.5** Consider  $\mathbb{Q}[\sqrt{2}]$ , defined by:

$$\mathbb{Q}[\sqrt{2}] = \{x + y\sqrt{2} \mid x, y \in \mathbb{Q}\} \quad (16.1.18)$$

This is a subfield of  $\mathbb{R}$ ,  $\mathbb{Q}[\sqrt{2}] \subseteq \mathbb{R}$ . Addition and multiplication are easy enough to see, and 0 and 1 are contained in there, we need only check multiplicative inverses. But:

$$(x + \sqrt{2}y)^{-1} = \frac{x - \sqrt{2}y}{x^2 - 2y^2} \quad (16.1.19)$$

And  $x^2 - 2y^2$  is only zero when  $x = y = 0$ , since if  $x^2 - 2y^2 = 0$ , then rearrange this to obtain  $y^2 = 2$ . But by the arguments of the ancient Greeks, there is no rational number whose square is 2, and thus the denominator is never zero for non-zero rational ordered pairs.

**Example 16.1.6**  $\mathbb{R}/\mathbb{Q}[\sqrt{2}]$  is uncountably infinite, but  $\mathbb{Q}[\sqrt{2}]/\mathbb{Q}$  has degree 2 with a basis  $\{1, \sqrt{2}\}$ .

### 16.1.3 Polynomials

We use  $F[x]$  to denote the ring of polynomials with coefficients in  $F$ . For example:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + x_0 \quad a_k \in F \quad (16.1.20)$$

Then  $f \in F[x]$ . The degree of a polynomial is the largest power of the polynomial with non-zero coefficient. Some things can be said about the degree of polynomials:

$$\deg(f + g) \leq \max\{\deg(f), \deg(g)\} \quad (16.1.21)$$

$$\deg(fg) = \deg(f) + \deg(g) \quad f, g \neq 0 \quad (16.1.22)$$

The degree of a polynomial is zero if and only if the polynomial is constant. Since  $F[x]$  has a ring structure,  $F[x]^*$  can be seen as the set of all non-zero constant polynomials.

**Theorem 16.1.2.**  *$F[x]$  is a Euclidean domain. That is, for any polynomial  $f \in F[x]$  and for any non-zero  $g \in F[x]$ , there exist unique polynomials  $r, q \in F[x]$  such that  $f = qg + r$  where either  $r = 0$  or  $\deg(r) < \deg(g)$ .*

**Theorem 16.1.3.** *The polynomial ring  $F[x]$  is a principal ideal domain. That is, every ideal  $I \subseteq F[x]$  is principal. That is, every ideal is generated by a single element.*

**Theorem 16.1.4.** *Every Euclidean domain is a principle ideal.*

Thus, there is a bijection between ideals  $I \subseteq F[x]$  and monic polynomials in  $F[x]$ . Recall that if  $R$  is a commutative ring with unity, then  $r \in R$  is called irreducible if  $r \neq 0$ ,  $r$  not a unit, and if  $r = ab$ , then either  $a$  or  $b$  is a unit. We take this definition to exclude some trivialities. For example, in  $\mathbb{Z}$ , 3 is irreducible, however  $3 = (-1) \cdot (-3)$ . We don't care about this product, since -1 is a unit. Moreover, an element  $r \in R$  is prime if  $(r) \subseteq R$  is a prime ideal. That is if  $r$  divides  $ab$ , then either  $r$  divides  $a$  or  $r$  divides  $b$ . By divides,  $r|a$ , we mean that  $a = r \cdot s$  for some  $s \in R$ .

**Example 16.1.7** If  $F$  is a field,  $f \in F[x]$ , then  $f$  is irreducible if and only if  $f$  is not the product of two polynomials with degrees strictly less than  $f$ . That is, if  $f = gh$ , then one of these must be a constant.

**Example 16.1.8** In  $\mathbb{Z}$ , prime if and only if irreducible.

**Theorem 16.1.5.** *If  $R$  is a integral domain, and if  $r$  is prime, then it is irreducible. That is, if there are no zero divisors then prime implies irreducible.*

**Theorem 16.1.6.** *If  $R$  is a principal ideal domain and if  $r$  is irreducible, then the ideal generated by  $r$  is maximal.*

**Theorem 16.1.7.** *A maximal ideal is a prime ideal.*

Recall that an ideal is called prime if  $R/I$  is a domain. That is, if  $ab \in I$ , then either  $a \in I$  or  $b \in I$ . A maximal ideal is an ideal that has no proper ideals between it and the entire ring. Another way to say this is that  $R/I$  is a field. In other words, if  $I \subseteq J \subseteq R$ , then  $I = J$ . Using this we see that a maximal ideal is prime since  $R/I$  will be a field, which is certainly an integral domain.

**Theorem 16.1.8.** *The fourth isomorphism theorem says that if  $I \subseteq R$  is an ideal, then there is a bijection between ideals containing  $I$ ,  $I \subseteq J \subseteq R$ , and ideals of  $R/I$ .*

**Theorem 16.1.9.**  *$f \in F[x]$  is irreducible if and only if  $F[x]/(f)$  is a field.*

Note that  $F$  can be seen as a subfield of  $F[x]/(f)$  since  $F$  can be identified with all constant polynomials, which can further be seen to live inside of  $F[x]/(f)$ .

**Theorem 16.1.10.** If  $\bar{g} \in F[x]/(f)$  then there exists a unique  $g_0 \in F[x]$  such that  $\deg(g_0) < \deg(f)$ , with  $\overline{g_0} = \bar{g}$ .

If  $n$  is the degree of  $f$ , then the set  $\{\bar{1}, \bar{x}, \dots, \bar{x^{n-1}}\}$  is a basis for  $F[x]/(f)$  over  $F$ .

**Theorem 16.1.11.** If  $f$  is irreducible of degree  $n$ , then  $F[x]/(f)$  is a field extension of  $F$  of degree  $n$ .

**Example 16.1.9** In  $\mathbb{R}$ , the polynomial  $f(x) = x^2 + 1$  is irreducible since it cannot be factored any further. Thus  $\mathbb{R}[x]/(x^2 + 1)$  is a field extension of  $\mathbb{R}$  of degree 2.

**Theorem 16.1.12.**  $\mathbb{R}[x]/(x^2 + 1)$  is isomorphic to  $\mathbb{C}$ .

*Proof.* For since  $\{\bar{1}, \bar{x}\}$  is a basis, we have:

$$\mathbb{R}[x]/(x^2 + 1) = \{a\bar{1} + b\bar{x} \mid a, b \in \mathbb{R}\} \quad (16.1.23)$$

So we trivial map  $a\bar{1} + b\bar{x}$  to  $a + bi$ . □

**Example 16.1.10** Consider now  $\mathbb{Q}[x]$  with  $x^2 - 2$ . This is irreducible since it cannot be factored ( $\sqrt{2}$  is irrational). Then  $\mathbb{Q}[x]/(x^2 - 2)$  is isomorphic to  $\mathbb{Q}[\sqrt{2}]$ .

#### 16.1.4 Review of Previous Lecture

If  $F$  is a field, and if  $f \in F[x]$  is irreducible, then  $F[x]/(f)$  is a field extension of  $F$  of degree  $\deg(f)$ . Also,  $\bar{x} = x + (f) \in F[x]/(f)$  is a root of  $f(x)$  in this field  $F[x]/(f)$ . That is,  $f(\bar{x}) = \bar{f}(x)$ , and this maps to zero. The fact that  $f(\bar{x}) = \bar{f}(x)$  is simply the statement that the quotient ring is well defined.

## 16.2 Stuff

**Problem 16.2.1** Let  $f(x) = x^4 - 1$  and  $g(x) = 3x^2 + 3x$ . Find the quotient and remainder after dividing  $f$  by  $g$ , the GCD and  $f$  and  $g$ , and the expression of the GCD in terms of  $af + bg$  with  $a, b \in \mathbb{Q}[x]$ . Use the Euclidean algorithm and Bézout's identity.

*Solution* We first try to find the quotient  $q$ . Noting that  $f$  is monic, and that the leading coefficient of  $g$  is 3, we try  $q_0 = \frac{1}{3}x^2$ . This gives:

$$q_0(x)g(x) = \frac{1}{3}x^2(3x^2 + 3x) = x^4 + x^3 = (x^4 - 1) + (x^3 + 1) = f(x) + (x^3 + 1) \quad (16.2.1)$$

so we need to subtract off  $x^3 + 1$ . We try  $q_1 = -\frac{1}{3}x$  and obtain:

$$q_1(x)g(x) = -\frac{1}{3}x(3x^2 + 3x) = -x^3 - x^2 = (-x^3 - 1) + (-x^2 + 1) \quad (16.2.2)$$

so the remainder is now  $-x^2 + 1$ . We try  $\frac{1}{3}$ , and get:

$$q_2(x)g(x) = \frac{1}{3}(3x^2 + 3x) = x^2 + x = (x^2 - 1) + (x + 1) \quad (16.2.3)$$

and so we are left with remainder  $-x - 1$ . That is:

$$x^4 - 1 = \left(\frac{1}{3}x^2 - \frac{1}{3}x + \frac{1}{3}\right)(3x^2 + 3x) + (-x - 1) \quad (16.2.4)$$

We double check just to be safe:

$$\begin{aligned} \left(\frac{1}{3}x^2 - \frac{1}{3}x + \frac{1}{3}\right)(3x^2 + 3x) + (-x - 1) &= x^4 - x^3 + x^2 + x^3 - x^2 + x + (-x - 1) \\ &= x^4 - 1 \end{aligned}$$

and we may rejoice. To find the GCD, we apply the Euclidean algorithm and perform repeated division with remainder. We've done step one, now we need to divide  $g$  by the remainder  $-x - 1$ . We have:

$$g(x) = 3x^2 + 3x = -3(-x - 1) \quad (16.2.5)$$

so the remainder is zero, and hence the GCD is  $-x - 1$ . The coefficients for B'ezout identity are thus trivially computed, with  $a = 1$  and  $b = -g$ . That is:

$$(x^4 - 1) + \left(-\frac{1}{3}x^2 + \frac{x}{3} - \frac{1}{3}\right)(3x^2 + 3x) = -x - 1 \quad (16.2.6)$$

**Problem 16.2.2** Show that various polynomials are reducible or irreducible.

*Solution* Since the only irreducible polynomials over  $\mathbb{R}$  are linear and quadratic ones,  $x^4 + 1$  is irreducible even though it has no roots. We look for complex roots, and then multiply conjugate pairs. The roots are  $\pm\sqrt{\pm i}$ , which corresponds to  $(\pm 1 \pm i)/\sqrt{2}$ . Multiplying in conjugate pairs, we have:

$$x^4 + 1 = \left(x - \frac{1+i}{\sqrt{2}}\right)\left(x - \frac{1-i}{\sqrt{2}}\right)\left(x - \frac{-1+i}{\sqrt{2}}\right)\left(x - \frac{-1-i}{\sqrt{2}}\right) \quad (16.2.7)$$

this simplifies to:

$$x^4 + 1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1) \quad (16.2.8)$$

$x^4 + 1$  is irreducible over  $\mathbb{Q}$ . We prove this via the sliding technique and applying Eisenstein's criterion. We have:

$$f(x+1) = (x+1)^4 + 1 = x^4 + 4x^3 + 6x^2 + 4x + 2 \quad (16.2.9)$$

so 2 is a prime that divides the  $a_k$  except for the leading one, and such that  $2^2$  does not divide the last coefficient. Hence, by Eisenstein's criterion,  $f(x+1)$

is irreducible. But if  $f(x+1)$  is irreducible, then so is  $f(x)$ . Next up is  $x^7 + 11x^3 - 33x + 22$ . Reducing modulo 2 is instantly noticed to be reducible since we can factor our an  $x$ . Three and five are trickier, but can be done. Finally in modulo 7 it is irreducible, and therefore it is irreducible in  $\mathbb{Q}$ . For  $1+x+x^2+x^3+x^4$  we invoke the partial sum rule for the geometric series and shift this polynomial by 1. We have:

$$f(x+1) = \sum_{k=0}^4 (x+1)^k = \frac{1-(x+1)^5}{-x} = x^4 + 5x^3 + 10x^2 + 10x + 5 \quad (16.2.10)$$

which is irreducible by Eisenstein setting  $p = 5$ . For the last one,  $x^7 - 7x^2 + 3x + 3$  we note that  $x = 1$  is a root since  $1 - 7 + 3 + 3 = -6 + 6 = 0$ . Hence  $x - 1$  is a factor. Performing division we have:

$$x^3 - 7x^2 + 3x + 3 = (x-1)(x^2 - 6x + 1) \quad (16.2.11)$$

By the quadratic formula, the roots of  $x^2 - 6x + 1$  are irrational and hence this is the reduced form of  $x^3 - 7x^2 + 3x + 1$ .

**Problem 16.2.3** Find all monic irreducible polynomials of degree  $\leq 3$  in  $\mathbb{Z}_3/3\mathbb{Z}$ . Determine the number of irreducible polynomials of degree 4 as well.

*Solution* Since we do not consider constants as irreducible, we move on to degree 1. In this case, since  $f$  is required to be monic, we have that  $f(x) = x+a$  is monic for any  $a \in \mathbb{Z}/3\mathbb{Z}$ . Hence  $x$ ,  $x+1$ , and  $x+2$  are all irreducible. For degree two, if  $f$  is reducible then it must be the product of two degree polynomials of degree one:  $f(x) = (x+a)(x+b) = x^2 + (a+b)x + ab$  and so we must avoid such quadratics. In other words,  $f$  must have no roots. If we write  $f(x) = x^2 + ax + b$ , we require  $b \neq 0$  for otherwise 0 is a root of  $f$ . We also require  $1+a+b \neq 0$  and  $1-a+b \neq 0$ . If  $b = 1$ , then  $2+a \neq 0$  and  $2-a \neq 0$ . The first equation means  $a \neq 1$  and the second implies  $a \neq 2$ . Hence,  $a = 0$ . That is, if  $b = 1$  we are forced to choose  $f(x) = x^2 + 1$ . If  $b = 2$ , we have  $1+a+2 \neq 0$ , which implies  $a \neq 0$ . The second equation gives us  $-a \neq 0$ , but this is redundant, and hence the only constraint is the  $a \neq 0$ . So  $x^2 + x + 2$  and  $x^2 + 2x + 2$  are both irreducible polynomials.

	a=0	a=1	a=2
b=0	X	X	X
b=1	✓	X	X
b=2	X	✓	✓

Table 16.1: Monic Irreducible Quadratics in  $\mathbb{Z}/3\mathbb{Z}$

For degree three, if  $f$  is reducible then either it splits into linear terms, or factors into a linear times an irreducible quadratic. In either scenario, if  $f$  is

reducible, then it has a root. So we begin by trying to avoid monic cubics with roots. If  $f(x) = x^3 + ax^2 + bx + c$ , then automatically we require  $c \neq 0$  or else 0 would be a root of  $f$ . We also require that  $1 + a + b + c \neq 0$  and  $-1 + a - b + c \neq 0$ . If  $a = 0$  then this reduces to  $b + c \neq -1$  and  $-b + c \neq 1$ . If we try  $b = 0$ , we have  $c \neq 0$ ,  $c \neq 1$ , and  $c \neq -1$ , which is impossible. So  $a = 0$  and  $b = 0$  are off the table. If we try  $b = 1$  we again have no legal choices for  $c$ , and so this column is crossed out as well. Finally, if  $b = 2$  we have  $c \neq 0$  as the only constraint, and  $x^3 + 2x + 1$  and  $x^2 + 2x + 2$  are irreducible. If  $a = 1$ , we are left with  $b + c \neq 1$  and  $-b + c \neq 0$ , in addition to the constraint that  $c \neq 0$ . The second relation says  $b \neq c$  and so we can cross out the diagonal in our table. But  $b = 0$  we simply have  $c \neq 1$ , so  $x^3 + x^2 + 2$  does the trick. When  $b = 1$  we have  $c \neq 1$ , leaving us with  $x^3 + x^2 + x + 2$ . The only spot left on our table is  $c = 1$  and  $b = 2$ , and this is indeed irreducible. In the final case with  $a = 2$ , our equations are  $1 + 2 + b + c \neq 0$  and  $-1 + 2 - b + c \neq 0$ , which reduces to  $b + c \neq 0$  and  $-b + c \neq -1$ . If  $b = 0$ , we have  $c \neq 0$  and  $c \neq -1$ , and so we are stuck with  $c = 1$ . Taking  $b = 1$  we have  $c \neq 0$  and  $c \neq -1$ , so  $c = 2$ . Finally we conclude with  $b = 2$ , which means  $c \neq 1$ , and so  $c = 2$ .

$a=0$	$b=0$	$b=1$	$b=2$	$a=1$	0	1	2	$a=2$	0	1	2
$c=0$	X	X	X		X	X	X		X	X	X
$c=1$	X	X	✓		X	X	✓		✓	X	X
$c=2$	X	X	✓		✓	✓	X		X	✓	✓

Table 16.2: Monic Irreducible Cubics in  $\mathbb{Z}/3\mathbb{Z}$ 

Hence we have a total of 8 irreducible monic cubic polynomials over  $\mathbb{Z}/3\mathbb{Z}$ . So in total there are 3 irreducible linear polynomials, 3 irreducible quadratics, and 8 irreducible cubics. To compute the number of irreducible quartics it suffices to compute the number of reducible quartics. There are a total of  $3^4$  monic quartic polynomials  $x^4 + ax^3 + bx^2 + cx + d$  since there are 3 choices for each of the  $a, b, c$ , and  $d$ . If such a polynomial is reducible then it factors either as a product of linear terms, a product of an irreducible cubic and a linear term, a product of two linear terms and an irreducible quadratic, or the product of two irreducible quadratics.

## 16.3 Rings

### Definition 16.3.1: Ring

A ring is a set  $R$  with two binary operations  $+$  and  $\cdot$ , denoted  $(R, +, \cdot)$ , such that  $(R, +)$  is a group,  $(R, \cdot)$  is a monoid, and such that  $\cdot$  distributes over  $+$ . That is, for all  $a, b, c \in R$  it is true that:

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$

The unital element of  $(R, +)$  is denoted  $0$  and the unital element of  $(R, \cdot)$  is written as  $1$ .  $+$  is called addition and  $\cdot$  is called multiplication.

If we relax the requirement that  $(R, \cdot)$  is a monoid and merely require it to be a semigroup (that is, there need not exist a multiplicative identity), then the resulting structure is called a *rng*. Some authors reserve the word ring for the more general case of what we're calling a rng, and use the phrasing *ring with identity* for what we're calling a ring. Commutative rings are rings  $(R, +, \cdot)$  where  $(R, \cdot)$  is an Abelian monoid. Rings (with identity) are automatically commutative in  $+$ . That is,  $(R, +)$  is an Abelian group.

**Theorem 16.3.1.** *If  $(R, +, \cdot)$  is a ring, then  $(R, +)$  is an Abelian group.*

*Proof.* For suppose not. Then there exist  $a, b \in R$  such that  $a + b \neq b + a$ . But  $(R, +, \cdot)$  is a ring, and hence  $(R, \cdot)$  is a monoid and thus there is an identity element  $1 \in R$  with respect to multiplication. But since  $R$  is a ring, multiplication distributes over addition and therefore:

$$\begin{aligned} (1 + 1) \cdot (a + b) &= ((1 + 1) \cdot a) + ((1 + 1) \cdot b) && \text{(Left Distributive Law)} \\ &= ((1 \cdot a) + (1 \cdot a)) + ((1 \cdot b) + (1 \cdot b)) && \text{(Right Distributive Law)} \\ &= (a + a) + (b + b) && \text{(Multiplicative Identity)} \\ &= a + (a + (b + b)) && \text{(Associative Law)} \end{aligned}$$

But again by distributivity, we obtain:

$$\begin{aligned} (1 + 1) \cdot (a + b) &= (1 \cdot (a + b)) + (1 \cdot (a + b)) && \text{(Right Distributive Law)} \\ &= (a + b) + (a + b) && \text{(Multiplicative Identity)} \\ &= a + (b + (a + b)) && \text{(Associative Law)} \end{aligned}$$

By the transitivity of equality, we have:

$$a + (a + (b + b)) = a + (b + (a + b)) \tag{16.3.1}$$

But  $(R, +)$  is a group, and thus by the left cancellation law  $a + (b + b) = b + (a + b)$ . By associativity  $a + (b + b) = (a + b) + b$  and  $b + (a + b) = (b + a) + b$ , and hence  $(a + b) + b = (b + a) + b$ . Thus by the right cancellation law,  $a + b = b + a$ , which is a contradiction. Therefore,  $(R, +)$  is Abelian.  $\square$

**Example 16.3.1** If  $R = \{0\}$ , and if  $+$  and  $\cdot$  are the only binary operations one can define on  $R$ :  $0 \cdot 0 = 0$  and  $0 + 0 = 0$ , then  $(R, +, \cdot)$  is a ring. Since  $(R, +)$  and  $(R, \cdot)$  are simply the Abelian group  $\mathbb{Z}_1$ , we need only check the distributive law, but this holds trivially since  $0 \cdot (0 + 0) = 0 \cdot 0 = 0$  and  $0 \cdot 0 + 0 \cdot 0 = 0$ . This is often called the *zero ring*, but also the trivial ring.

**Example 16.3.2** There are other types of different trivial rng structures on any Abelian group  $(G, *)$ . Defined  $\cdot : G \times G \rightarrow G$  by  $a \cdot b = e$  for all  $a, b \in G$ , where  $e \in G$  is the unital element. This is a rng since  $\cdot$  is associative, and hence  $(R, \cdot)$  is a semigroup. Multiplication also distributes over  $*$  trivially since:

$$a \cdot (b * c) = e \quad (16.3.2a) \quad (a \cdot b) * (a \cdot c) = e * e = e \quad (16.3.2b)$$

Thus,  $(R, +, \cdot)$  is a rng. If  $R$  has at least two elements then this cannot be a proper ring, since there can be no multiplicative identity.

**Example 16.3.3** The rational, real, and complex numbers, together with their usual arithmetic, form commutative rings (moreover, they form fields). That is, letting  $+$  and  $\cdot$  denote the familiar forms of addition and multiplication, respectively,  $(\mathbb{R}, +, \cdot)$  is a ring, and similarly for the other two.

**Example 16.3.4** The quintessential example of a ring is  $\mathbb{Z}$  quipped with its usual arithmetic. That is,  $(\mathbb{Z}, +, \cdot)$  is a ring.  $(\mathbb{Z}, +)$  is an Abelian group and  $(\mathbb{Z}, \cdot)$  forms an Abelian monoid. Moreover, multiplication distributes over addition, and hence  $(\mathbb{Z}, +, \cdot)$  is a commutative ring.

**Example 16.3.5** Endowing  $\mathbb{Z}_n$  with its modulo arithmetic structure,  $(\mathbb{Z}_n, +, \cdot)$  also forms a ring.

An important class of rings comes from studying function spaces.

**Theorem 16.3.2.** *If  $X$  is a set, if  $(R, +_R, \cdot_R)$  is a ring, if  $\mathcal{F}(X, R)$  is the set of all functions  $f : X \rightarrow R$ , if  $+$  is the binary operation on  $\mathcal{F}(X, R)$  defined by:*

$$(f + g)(x) = f(x) +_R g(x) \quad (16.3.3)$$

*and if  $\cdot$  is the binary operation defined by:*

$$(f \cdot g)(x) = f(x) \cdot_R g(x) \quad (16.3.4)$$

*then  $(\mathcal{F}(X, R), +, \cdot)$  is a ring.*

**Example 16.3.6** The set of even integers  $\mathbb{Z}_e$  with addition and multiplication forms a rng, but not a ring. The identity element of  $\mathbb{Z}$  is 1, and 1 is not even.

**Theorem 16.3.3.** *If  $(R, +, \cdot)$  is a rng, if 0 is the unital element of  $(R, +)$ , and if  $a \in R$ , then  $a \cdot 0 = 0$  and  $0 \cdot a = 0$ .*

*Proof.* For since  $(R, +, \cdot)$  is a rng,  $\cdot$  distributes over addition, and hence both left and right distributes. And since 0 is the additive identity, we obtain:

$$a \cdot 0 = a \cdot (0 + 0) = (a \cdot 0) + (a \cdot 0) \quad (16.3.5)$$

by the left cancellation law we have that  $a \cdot 0 = 0$ . Similary:

$$0 \cdot a = (0 + 0) \cdot a = (0 \cdot a) + (0 \cdot a) \quad (16.3.6)$$

Again by the cancellation law,  $0 \cdot a = 0$ .  $\square$

**Theorem 16.3.4.** *If  $(R, +, \cdot)$  is a ring, if 0 is the multiplicative identity, if 1 is the additive identity, and if  $0 = 1$ , then  $R = \{0\}$ .*

*Proof.* For suppose not. Then there is an element  $x \in R$  such that  $x \neq 0$ . But 1 is the multiplicative identity, and hence  $a = a \cdot 1$ . But  $1 = 0$ , and thus  $a = a \cdot 0$ . But  $a \cdot 0 = 0$  (Thm. 16.3.3), and thus by the transitivity of equality we have  $a = 0$ , a contradiction.  $\square$

**Theorem 16.3.5.** *If  $(R, +, \cdot)$  is a rng, if  $a, b \in R$ , and if  $-a \in R$  is the additive inverse of  $a$ , then:*

$$(-a) \cdot b = -(a \cdot b) \quad (16.3.7)$$

*Proof.* For:

$$a \cdot b + (-a) \cdot b = (a + (-a)) \cdot b = 0 \cdot b = 0 \quad (16.3.8)$$

and thus  $(-a) \cdot b = -(a \cdot b)$ .  $\square$

**Theorem 16.3.6.** *If  $(R, +, \cdot)$  is a rng, if  $a, b \in R$ , and if  $-a, -b \in R$  are their additive inverses, respectively, then:*

$$(-a) \cdot (-b) = a \cdot b \quad (16.3.9)$$

*Proof.* For by Thm. 16.3.5,  $(-a) \cdot (-b) = -(a \cdot (-b))$ . But then:

$$a \cdot b - ((-a) \cdot (-b)) = a \cdot b + (a \cdot (-b)) = a \cdot (b + (-b)) = a \cdot 0 = 0 \quad (16.3.10)$$

and hence  $a \cdot b = (-a) \cdot (-b)$ .  $\square$

**Theorem 16.3.7.** *If  $(R, +, \cdot)$  is a ring, if  $1 \in R$  is the multiplicative identity, and if  $-1$  is the additive inverse of 1, then  $(-1)^2 = 1$ .*

*Proof.* For  $(-1)^2 = (-1) \cdot (-1)$  and by Thm. 16.3.6  $(-1) \cdot (-1) = 1 \cdot 1$ . But 1 is the multiplicative identity and thus  $1 \cdot 1 = 1$ .  $\square$

**Theorem 16.3.8.** *If  $(R, +, \cdot)$  is a ring, if  $r \in R$  has a multiplicative inverse, and if  $-r$  is the additive inverse of  $r$ , then  $-r$  has a multiplicative inverse.*

*Proof.* For if  $r$  has a multiplicative inverse, then there is an element  $r^{-1} \in R$  such that  $r \cdot r^{-1} = 1$ . But then:

$$\begin{aligned} (-r) \cdot (-r^{-1}) &= r \cdot r^{-1} && \text{(Thm. 16.3.6)} \\ &= 1 && \text{(Multiplicative Inverse)} \end{aligned}$$

and thus  $-r$  is invertible.  $\square$

### Definition 16.3.2: Zero Divisor

A zero divisor of a rng  $(R, +, \cdot)$  is an element  $a \in R$  such that  $a \neq 0$  and there exists a  $b \in R$  such that  $b \neq 0$  and either  $a \cdot b = 0$  or  $b \cdot a = 0$ .

**Example 16.3.7** The ring of integers contains no zero divisors. This follows from the work of Euclid. If  $n \cdot m = 0$ , then either  $n = 0$  or  $m = 0$  and hence there can be no zero divisors.

**Example 16.3.8** Let  $n \in \mathbb{N}^+$  be positive, and suppose  $a \in \mathbb{Z}_n$  is such that  $a$  divides  $n$ . That is, there is an  $m \in \mathbb{N}$  such that  $a \cdot m = n$ . Then in the ring of integers modulo  $n$ ,  $(\mathbb{Z}/n\mathbb{Z}, \tilde{+}, \tilde{\cdot})$ , we have that both  $a$  and  $m$  are zero divisors. That is, since  $a \cdot m = n$ , they are congruent to zero modulo  $n$  and hence  $a \cdot m \equiv 0 \pmod{n}$ . We can also see that if  $p$  is prime then there are no zero-divisors. This is just a consequence of Thm. ???. That is, if  $p$  is prime, then  $(\mathbb{Z}/p\mathbb{Z}, \tilde{+}, \tilde{\cdot})$  is a ring without zero divisors.

**Theorem 16.3.9.** *If  $(R, +, \cdot)$  is a ring, and if  $R^\times$  is the set:*

$$R^\times = \{ r \in R \mid \exists_{r^{-1} \in R} (r \cdot r^{-1} = 1) \} \quad (16.3.11)$$

*then  $(R^\times, \cdot_{R^\times})$  is a group.*

*Proof.* For if  $a, b \in R$  are invertible, then  $a \cdot b$  is. Moreover, the identity is invertible, and  $\cdot$  is associative. Hence, we have a group.  $\square$

**Theorem 16.3.10.** *If  $(R, +, \cdot)$  is a non-zero ring, and if  $a \in R$  is a zero divisor, then  $a$  is not invertible.*

*Proof.* For suppose not. Then there is a  $b \in R$  such that  $a \cdot b = 1$ . But since  $a$  is a zero divisor,  $a \neq 0$  and there is a  $c \in R$  such that  $c \neq 0$  and either  $a \cdot c = 0$  or  $c \cdot a = 0$  (Def. 16.3.2)  $c \cdot a = 0$ . But then:

$$\begin{array}{llll} c = c \cdot 1 & \text{(Identity)} & = 0 \cdot b & \text{(Hypothesis)} \\ = c \cdot (a \cdot b) & \text{(Inverse)} & = 0 & \text{(Thm. 16.3.3)} \\ = (c \cdot a) \cdot b & \text{(Associativity)} & & \end{array}$$

a contradiction. Similarly if  $a \cdot c = 0$ .  $\square$

### Definition 16.3.3: Integral Domain

An integral domain is a ring  $(R, +, \cdot)$  such that for all  $r \in R$  it is true that  $r$  is not a zero divisor.

**Example 16.3.9** By our previous discussion,  $(\mathbb{Z}, +, \cdot)$  is an integral domain since it contains no zero divisors.

**Theorem 16.3.11.** *If  $(R, +, \cdot)$  is an integral domain, if  $a, b, c \in R$ , and if  $a \cdot b = a \cdot c$ , then either  $b = c$  or  $a = 0$ .*

*Proof.* For suppose not. Then:

$$a \cdot (b - c) = a \cdot b - a \cdot c = a \cdot c - a \cdot c = 0 \quad (16.3.12)$$

But  $b \neq c$  and thus  $b - c \neq 0$ . But then  $a$  is a non-zero element of  $R$  such that there exists a non-zero element  $b - c$  with  $a \cdot (b - c) = 0$ , and hence  $a$  is a zero divisor (Def. 16.3.2). But  $(R, +, \cdot)$  is an integral domain and hence there are no zero divisors (Def. 16.3.3), a contradiction.  $\square$

**Theorem 16.3.12.** *If  $(R, +, \cdot)$  is an integral domain, if  $a, b \in R$ , and if  $a \cdot b = 0$ , then either  $a = 0$  or  $b = 0$ .*

*Proof.* For if  $(R, +, \cdot)$  is a ring, then  $a \cdot 0 = 0$  (Thm. 16.3.3). But  $a \cdot b = 0$  by hypothesis, and hence by the transitivity of equality  $a \cdot b = a \cdot 0$ . But if  $(R, +, \cdot)$  is an integral domain, then by the cancellation law either  $a = 0$  or  $b = 0$  (Thm. 16.3.11).  $\square$

### Definition 16.3.4: Field

A field is a commutative ring  $(R, +, \cdot)$  such that for all  $r \in R \setminus \{0\}$  there exists a multiplicative inverse  $r^{-1} \in R \setminus \{0\}$  and such that  $0 \neq 1$ .

**Example 16.3.10** Since by definition a division ring is a non-zero ring, fields are required to have at least two elements. The smallest field is thus  $(\mathbb{Z}_2, +, \cdot)$  with modulo arithmetic.

**Theorem 16.3.13.** *If  $(F, +, \cdot)$  is a field, then  $(F, +, \cdot)$  is an integral domain.*

*Proof.* For suppose not. But if  $(F, +, \cdot)$  is not an integral domain, then there is a zero divisor  $r \in F$  (Def. 16.3.3). But if  $r$  is a zero divisor, then  $r \neq 0$ , and hence  $r \in R \setminus \{0\}$ . But  $F$  is a field, and therefore  $r$  is invertible (Def. 16.3.4). But zero divisors are not invertible (Thm. 16.3.10), a contradiction.  $\square$

**Theorem 16.3.14.** *If  $(R, +, \cdot)$  is an integral domain, if  $a \in R$  is non-zero, and if  $f : R \rightarrow R$  is defined by  $f(x) = a \cdot x$ , then  $f$  is injective.*

*Proof.* For suppose not. Then there are elements  $x, y \in R$  such that  $x \neq y$  and  $f(x) = f(y)$ . But then  $a \cdot x = a \cdot y$ . But since  $R$  is an integral domain, if  $a \cdot x = a \cdot y$ , then either  $a = 0$  or  $x = y$  (Thm. 16.3.11). But  $a \neq 0$  and  $x \neq y$ , a contradiction.  $\square$

**Theorem 16.3.15.** *If  $(R, +, \cdot)$  is an integral domain, if  $R$  is finite, and if  $0 \neq 1$ , then  $(R, +, \cdot)$  is a field.*

*Proof.* For suppose not. Then there is an  $r \in R \setminus \{0\}$  such that  $r$  has no multiplicative inverse. Let  $f : R \rightarrow R$  be defined by  $f(x) = r \cdot x$ . But then  $f$  is injective (Thm. 16.3.14). And since  $R$  is finite, it is therefore surjective. But then there is an  $r^{-1} \in R$  such that  $f(r^{-1}) = 1$ . But then  $r \cdot r^{-1} = 1$ , a contradiction.  $\square$

### Definition 16.3.5: Subring

A subring of a ring  $(R, +, \cdot)$  is a subset  $S \subseteq R$  such that  $(S, +_S)$  is a subgroup of  $(R, +)$  and  $(S, \cdot_S)$  is a submonoid of  $(R, \cdot)$ .

**Example 16.3.11** If we take  $(\mathbb{C}, +, \cdot)$  to be usual field structure on the complex numbers (which is therefore a ring), there are several familiar subrings.  $\mathbb{R}$ ,  $\mathbb{Q}$ , and  $\mathbb{Z}$  are all subrings.

**Theorem 16.3.16.** *If  $(F, +, \cdot)$  is a field, and if  $R \subseteq F$  is a subring of  $F$ , then  $(R, +_R, \cdot_R)$  is an integral domain.*

*Proof.* For suppose not. Then there is a zero divisor  $a \in R$  (Def. 16.3.3), and hence there is a non-zero element  $b \in R$  such that either  $a \cdot b = 0$  or  $b \cdot a = 0$ . But  $R \subseteq F$ , and hence  $a, b \in F$ . But if  $(F, +, \cdot)$  is a field, then it is an integral domain (Thm. 16.3.13). But if  $(F, +, \cdot)$  is an integral domain, and if  $a \cdot b = 0$ , then either  $a = 0$  or  $b = 0$  (Thm. 16.3.12), a contradiction.  $\square$

**Example 16.3.12** If  $n \subseteq \mathbb{N}$  is square free (that is, there is no  $m \in \mathbb{N}$  such that  $m^2 = n$ ), then we can adjoin  $\sqrt{n}$  to  $\mathbb{Z}$  by considering the set:

$$\mathbb{Z}[\sqrt{n}] = \{ a + b\sqrt{n} \mid a, b \in \mathbb{Z} \} \quad (16.3.13)$$

This is a subring of  $\mathbb{R}$ . More precisely, it is a subring of  $\mathbb{Q}(\sqrt{n})$ . If  $a, b, c, d \in \mathbb{Z}$ , then:

$$(a + b\sqrt{n}) \cdot (c + d\sqrt{n}) = ac + bdn + (ad + bc)\sqrt{n} \quad (16.3.14)$$

And hence we have that  $\mathbb{Z}[\sqrt{n}]$  is closed under multiplication.

**Example 16.3.13** If  $r \in \mathbb{Q}^+$  is not a square of another rational number, we can define  $\mathbb{Q}(r)$  as follows:

$$\mathbb{Q}(r) = \{ a + b\sqrt{r} \mid a, b \in \mathbb{Q} \} \quad (16.3.15)$$

Since  $\mathbb{Q}$  is a field, it may be reasonable to suspect that  $\mathbb{Q}(\sqrt{r})$  is a field as well. First, we show that it is closed under addition:

$$(a + b\sqrt{r}) + (c + d\sqrt{r}) = (a + c) + (b + d)\sqrt{r} \quad (16.3.16)$$

and similarly for multiplication:

$$(a + b\sqrt{r}) \cdot (c + d\sqrt{r}) = (ac + bdr) + (ad + bc)\sqrt{r} \quad (16.3.17)$$

and hence  $\mathbb{Q}(\sqrt{r})$  is a subring of  $\mathbb{R}$ . It is also a field. We need only show that multiplicative inverses for non-zero elements exist. Give  $a + b\sqrt{r}$ , define:

$$(a + b\sqrt{r})^{-1} = \frac{a}{a^2 - rb^2} - \frac{b}{a^2 - rb^2}\sqrt{r} \quad (16.3.18)$$

this is well defined since  $a^2 \neq rb^2$ , otherwise  $r = a^2/b^2$  which contradicts the fact that  $r$  is square free. Then  $(a + b\sqrt{r})^{-1}$  is an element of  $\mathbb{Q}(\sqrt{r})$  and moreover:

$$(a + b\sqrt{r})^{-1} \cdot (a + b\sqrt{r}) = \left( \frac{a}{a^2 - rb^2} - \frac{b}{a^2 - rb^2}\sqrt{r} \right) \cdot (a + b\sqrt{r}) \quad (16.3.19a)$$

$$= \frac{1}{a^2 - rb^2} ((a - b\sqrt{r}) \cdot (a + b\sqrt{r})) \quad (16.3.19b)$$

$$= \frac{a^2 - rb^2}{a^2 - rb^2} \quad (16.3.19c)$$

$$= 1 \quad (16.3.19d)$$

which shows that  $a + b\sqrt{r}$  is invertible. Thus,  $\mathbb{Q}(\sqrt{r})$  is a subfield of  $\mathbb{R}$ .

**Example 16.3.14** The Gaussian integers are another common example of a subring of the complex numbers  $\mathbb{Q}$ . Consider the subring  $\mathbb{Z}[i]$  defined as follows:

$$\mathbb{Z}[i] = \{ a + ib \mid a, b \in \mathbb{Z} \} \quad (16.3.20)$$

this is a subring of  $\mathbb{C}$  since:

$$(a + ib) + (c + id) = (a + c) + i(b + d) \quad (16.3.21)$$

and:

$$(a + ib) \cdot (c + id) = ac - bd + i(ad + bc) \quad (16.3.22)$$

hence,  $\mathbb{Z}[i] \subseteq \mathbb{C}$  is a subring.

## 16.4 Ideals

### Definition 16.4.1: Left Ideal

An ideal in a ring  $(R, +, \cdot)$  is a subset  $I \subseteq R$  such that for all  $a, b \in I$  it is true that  $a + b \in I$  and for all  $r \in R$  and  $a \in I$  it is true that  $r \cdot a \in I$ .

**Example 16.4.1** The motivating example of an ideal is the even integers in  $(\mathbb{Z}, +, \cdot)$ . The sum of two even integers is again an even integer, and for any even element  $2n$  and any  $m \in \mathbb{Z}$  we have  $2n \cdot m = 2 \cdot (n \cdot m)$ , which is again even.

**Theorem 16.4.1.** *If  $(R, +, \cdot)$  is a ring, and if  $I \subseteq R$  is a non-empty left ideal, then  $I$  is a subgroup of  $(R, +)$ .*

*Proof.* For since  $I$  is a left ideal,  $I$  is closed under addition (Def. 16.4.1). Moreover, since  $I$  is non-empty there is an  $x \in I$ . But since  $I$  is a left ideal,  $0 \cdot x \in I$  (Def. 16.4.1). But  $0 \cdot x = 0$  (Thm. 16.3.3) and hence  $0 \in I$ . Lastly, if  $x \in I$ , thus  $(-1) \cdot x = -x \in I$  and hence  $I$  is closed to additive inverses. Hence,  $I$  is a subgroup of  $(R, +)$ .  $\square$

We similarly define what it means to be a right ideal.

### Definition 16.4.2: Right Ideal

A ring ideal of a ring  $(R, +, \cdot)$  is a subset  $I \subseteq R$  such that for all  $a, b \in I$  it is true that  $a + b \in I$  and for all  $r \in R$  and  $a \in I$  it is true that  $a \cdot r \in I$ .

That we have two different definitions may indicate that these are different objects (Left ideals vs. right ideals). This stems from the fact that ring multiplication need not be commutative.

**Example 16.4.2** The set of all  $n \times n$  matrices where the entire last row is all zeroes forms a right ideal in the ring of  $n \times n$  matrices, but not a left ideal. Similarly, the set of all  $n \times n$  matrices with the last column set to zero is a left ideal, but not a right ideal.

To rid ourselves of such things, we often speak plainly of *ideals*. We define this now.

### Definition 16.4.3: Two Sided Ideal

A two sided ideal in a ring  $(R, +, \cdot)$  is a subset  $I \subseteq R$  such that  $I$  is a left ideal and a right ideal in  $R$ .

**Example 16.4.3** In  $(\mathbb{Z}, +, \cdot)$ , the even integers form a two sided ideal. As discussed previously, they form a left sided ideal, and from the commutativity of integer multiplication they then necessarily form a right ideal.

**Theorem 16.4.2.** *If  $(R, +, \cdot)$  is a commutative ring, and if  $I \subseteq R$  is either a left ideal or a right ideal, then it is a two sided ideal.*

*Proof.* For suppose not. Then either  $I$  is a left ideal and not a two sided ideal, or  $I$  is a right ideal and not two sided. But if  $I$  is a left ideal and not two sided, then it is not a right ideal (Def. 16.4.3) and hence there is an  $a \in I$  and an  $r \in R$  such that  $a \cdot r \notin I$ . But  $R$  is commutative, so  $a \cdot r = r \cdot a$ . But  $I$  is a left ideal and thus  $a \cdot r \in I$  (Def. 16.4.1), a contradiction. Similarly if  $I$  is a right ideal.  $\square$

**Theorem 16.4.3.** *If  $(F, +, \cdot)$  is a commutative ring, then it is a field if and only if there only ideals are the zero ideal and the entire ring.*

*Proof.* For if  $(F, +, \cdot)$  is a field then for all  $a \in F$  such that  $a \neq 0$  there exists a multiplicative inverse  $a^{-1} \in F$  such that  $a \cdot a^{-1} = 1$ . But then if  $I$  is a non-zero subring of  $F$ , there is a non-zero element  $a \in I$ , and thus  $a \cdot a^{-1} \in I$ , and therefore  $1 \in I$ . But then for all  $r \in F$ ,  $r \cdot 1 \in I$ , but  $r \cdot 1 = r$ . Hence,  $I = F$ . In the other direction, if the only ideals are the zero ideal and the entire ring, then for any non-zero  $a \in F$  the ideal generated by  $a$  must be the entire ring. Hence, there is a  $b \in I$  such that  $a \cdot b = 1$ , and so  $F$  is a field.  $\square$

**Example 16.4.4** Fields:  $\mathbb{Q}$ ,  $\mathbb{R}$ ,  $\mathbb{C}$ ,  $\mathbb{Z}_p$  with  $p$  a prime.

#### Definition 16.4.4: Prime Ideal

A prime ideal of a commutative ring  $(R, +, \cdot)$  is a proper subset  $I \subsetneq R$  such that  $I$  is a two sided ideal in  $R$  and for all  $a, b \in R$  with  $a \cdot b \in I$ , it is true that either  $a \in I$  or  $b \in I$ .

This is a generalization of Euclid's prime number lemma (Thm. ??). More precisely, it is a generalization of the corollary that follows (Thm. ??). We can use this connection to prove the following theorem.

**Theorem 16.4.4.** *If  $(\mathbb{Z}, +, \cdot)$  is the usual ring structure on  $\mathbb{Z}$ , and if  $p \in \mathbb{Z}$ , then  $p\mathbb{Z}$  is a prime ideal of  $\mathbb{Z}$  if and only if  $p$  is prime.*

*Proof.* For if  $p \in \mathbb{N}$  is prime, and if  $a, b \in \mathbb{Z}$  are such that  $a \cdot b \in p\mathbb{Z}$ , then there is an  $m \in \mathbb{Z}$  such that  $a \cdot b = m \cdot p$ . But then  $p$  divides  $a \cdot b$ , and thus either  $p$  divides  $a$  or  $p$  divides  $b$  (Thm. ??). But if  $p$  divides  $a$ , then  $a \in p\mathbb{Z}$ , and similarly if  $p$  divides  $b$ . Hence,  $p\mathbb{Z}$  is a prime ideal (Def. 16.4.4). In the other direction, if  $p\mathbb{Z}$  is a prime ideal, suppose  $p$  is not prime. Then there are integers  $a, b \in \mathbb{Z}$ , neither of which are units, such that  $p = a \cdot b$ . But by hypothesis  $p\mathbb{Z}$  is a prime ideal, and hence if  $a \cdot b \in p\mathbb{Z}$ , then either  $a \in p\mathbb{Z}$  or  $b \in p\mathbb{Z}$ . But if  $a \in p\mathbb{Z}$ , then there is an  $m \in \mathbb{Z}$  such that  $a = m \cdot p$ . But  $a \cdot b = p$ , and hence  $m \cdot p \cdot b = p$ . But then  $m \cdot b = 1$ , and thus  $b$  is a unit, which is a contradiction. Hence,  $p$  is prime.  $\square$

We can use this to generate examples of prime ideals.

**Example 16.4.5** The even integers  $\mathbb{Z}_e$  form a prime ideal of  $\mathbb{Z}$ . This can be seen immediately once one notes that  $\mathbb{Z}_e = 2\mathbb{Z}$ , and since 2 is a prime number, by Thm. 16.4.4 we see that  $2\mathbb{Z}$  is a prime ideal.

A useful equivalent definition of prime ideals goes as follows:

**Theorem 16.4.5.** *If  $(R, +, \cdot)$  is a ring, and if  $I \subsetneq R$  is a two sided ideal, then  $I$  is a prime ideal if and only if for all  $a, b \in R \setminus \{I\}$  it is true that  $a \cdot b \in R \setminus \{I\}$ .*

*Proof.* For if  $I$  is a prime ideal, then for all  $a, b \in R$  such that  $a \cdot b \in I$ , it is true that either  $a \in I$  or  $b \in I$ . Suppose there exists  $a, b \in R \setminus \{I\}$  such that  $a \cdot b \notin R \setminus \{I\}$ . But since  $\cdot$  is a binary operation it is true that  $a \cdot b \in R$ , and hence if  $a \cdot b \notin R \setminus I$ , then  $a \cdot b \in I$ . But  $I$  is a prime ideal, and thus if  $a \cdot b \in I$  then either  $a \in I$  or  $b \in I$ , a contradiction since by hypothesis  $a, b \notin I$ . Therefore,  $a \cdot b \in R \setminus \{I\}$ . In the other direction, if  $R \setminus \{I\}$  is multiplicatively closed, suppose  $a, b \in R$  are such that  $a \cdot b \in I$ . But  $R \setminus \{I\}$  is multiplicatively closed, and thus if  $a, b \in R \setminus \{I\}$ , then  $a \cdot b \in R \setminus \{I\}$ , a contradiction, and thus either  $a \in I$  or  $b \in I$ . Thus,  $I$  is a prime ideal.  $\square$

**Theorem 16.4.6.** *If  $(R, +_R, \cdot_R)$  and  $(S, +_S, \cdot_S)$  are rings, if  $\phi : R \rightarrow S$  is a ring homomorphism, and if  $I \subseteq S$  is a two sided ideal, then  $\phi^{-1}[I]$  is a two sided ideal.*

*Proof.* Since  $I$  is a two sided ideal, it is a left ideal (Def. 16.4.3) and hence  $I$  is a subgroup of  $(S, +_S)$  (Thm. 16.4.1) and therefore  $0_S \in I$ . But  $\phi$  is a ring homomorphism and therefore  $\phi(0_R) = 0_S$ . Hence,  $\phi^{-1}[I]$  is non-empty. But if  $a, b \in \phi^{-1}[I]$ , then  $\phi(a), \phi(b) \in I$ . But again since  $\phi$  is a ring homomorphism we have  $\phi(a +_R b) = \phi(a) +_S \phi(b)$ , and since  $I$  is an ideal it is true that  $\phi(a) +_S \phi(b) \in I$  (Def. 16.4.3). Therefore  $\phi(a +_R b) \in I$ , and thus  $a +_R b \in \phi^{-1}[I]$ . Hence  $\phi^{-1}[I]$  is closed to addition. But if  $r \in R$  and  $a \in \phi^{-1}[I]$ , then since  $\phi$  is a ring homomorphism it is true that  $\phi(r \cdot_R a) = \phi(r) \cdot_S \phi(a)$ . But since  $a \in \phi^{-1}[I]$  it is true that  $\phi(a) \in I$ , and since  $I$  is an ideal in  $S$  it is therefore true that  $\phi(r) \cdot_S \phi(a) \in I$ . Thus,  $\phi(r \cdot_R a) \in I$ , and hence  $r \cdot_R a \in \phi^{-1}[I]$ . Therefore,  $\phi^{-1}[I]$  is an ideal in  $R$ .  $\square$

**Theorem 16.4.7.** *If  $(R, +_R, \cdot_R)$  and  $(S, +_S, \cdot_S)$  are commutative rings, if  $I \subsetneq S$  is a prime ideal, and if  $\phi : R \rightarrow S$  is a ring homomorphism, then  $\phi^{-1}[I]$  is a prime ideal in  $R$ .*

*Proof.* For if  $I$  is an ideal in  $S$  and  $\phi : R \rightarrow S$  is a ring homomorphism, then  $\phi^{-1}[I]$  is an ideal in  $R$  (Thm. 16.4.6). Suppose  $a, b \in R$  are such that  $a \cdot_R b \in \phi^{-1}[I]$ . But then  $\phi(a \cdot_R b) \in I$ , and since  $\phi$  is a ring homomorphism it is therefore true that  $\phi(a \cdot_R b) = \phi(a) \cdot_S \phi(b)$ . But  $I$  is a prime ideal, and if  $\phi(a) \cdot_S \phi(b) \in I$ , then either  $\phi(a) \in I$  or  $\phi(b) \in I$ . But then either  $a \in \phi^{-1}[I]$  or  $b \in \phi^{-1}[I]$ , and hence  $\phi^{-1}[I]$  is a prime ideal.  $\square$

Given a ring  $(R, +, \cdot)$ , the existence of prime ideals follows directly from the existence of *maximal* ideals. We first must define such things, and then prove that every maximal ideal is prime.

#### Definition 16.4.5: Maximal Ideal

A maximal ideal in a ring  $(R, +, \cdot)$  is an ideal  $I \subsetneq R$  such that for every ideal  $J \subsetneq R$  such that  $I \subseteq J$ , it is true that  $I = J$ .

**Theorem 16.4.8.** *If  $(R, +, \cdot)$  is an ideal, if  $I \subseteq R$  is a proper ideal, and if  $1 \in R$  is the multiplicative identity, then  $1 \notin I$ .*

*Proof.* For suppose not. But if  $1 \in I$ , then since  $I$  is an ideal for all  $r \in R$  it is true that  $r \cdot 1 \in I$ . But 1 is the multiplicative identity, and hence  $r \cdot 1 \in I$ . But then for all  $r \in R$  it is true that  $r \in I$ , and hence  $r = I$ , a contradiction since  $I$  is a proper ideal. Thus,  $1 \notin I$ .  $\square$

**Theorem 16.4.9: Krull's Theorem**

If  $(R, +, \cdot)$  is a non-zero commutative ring, then there is a maximal ideal  $I \subseteq R$ .

*Proof.* For let  $J \subseteq \mathcal{P}(R)$  be defined as follows:

$$J = \{ I \in \mathcal{P}(R) \mid I \text{ is a proper ideal of } R \} \quad (16.4.1)$$

But  $R$  is non-zero, and hence  $\{0\}$  is a proper ideal of  $R$ , and therefore  $J$  is non-empty. Then if  $\subseteq$  is the inclusion relation, then  $(J, \subseteq)$  is a partially ordered set. Suppose  $\Lambda \subseteq J$  is a chain. But then  $\bigcup \Lambda \in J$ . For if  $x, y \in \bigcup \Lambda$ , then there is a  $I_x \in \Lambda$  and an  $I_y \in \Lambda$  such that  $x \in I_x$  and  $y \in I_y$ . But since  $\Lambda$  is a chain, either  $I_x \subseteq I_y$  or  $I_y \subseteq I_x$ . Suppose  $I_x \subseteq I_y$ . But then  $x, y \in I_y$ , and since  $I_y \in J$  it is an ideal. But then  $x + y \in I_y$ , and similarly if  $I_y \subseteq I_x$ . Hence,  $x + y \in \bigcup \Lambda$ . Moreover, if  $r \in R$  and  $x \in \bigcup \Lambda$ , then there is an  $I \in \Lambda$  such that  $x \in I$ . But  $I$  is an ideal, and therefore  $r \cdot x \in I$ . But then  $r \cdot x \in \bigcup \Lambda$ . Therefore,  $\bigcup \Lambda$  is an ideal. Moreover, since for all  $I \in \Lambda$  it is true that  $I \in J$ ,  $I$  is thus a proper ideal. But if  $I$  is a proper ideal, then  $1 \notin I$  (Thm. 16.4.8). But this is true of all  $I \in \Lambda$ , and hence  $1 \notin \bigcup \Lambda$ , and therefore  $\bigcup \Lambda$  is a proper ideal. Therefore, every chain of  $(J, \subseteq)$  is bounded, and therefore there exists a maximal element  $I \in J$ . But then  $I$  is a proper ideal of  $R$  that is not contained in any other proper ideals, and is therefore a maximal ideal (Def. 16.4.5).  $\square$

Zorn's lemma is indeed required for this proof. Krull's original proof invoked transfinite induction, a consequence of the well-ordering theorem, and indeed Krull's theorem implies the axiom of choice. Since the axiom of choice and Zorn's lemma are equivalent, any axiomatic system capable of proving Krull's theorem must contain, or be able to prove, the axiom of choice. Nevertheless, we can now prove that every non-zero ring contains a prime ideal.

**Theorem 16.4.10: Maximal Ideals are Prime**

If  $(R, +, \cdot)$  is a ring, if  $I \subseteq R$  is a maximal ideal, then  $I$  is a prime ideal.  $\blacksquare$

*Proof.* For suppose not. Then there exists  $a, b \in R$  such that  $a, b \notin I$ , but  $a \cdot b \in I$  (Def. 16.4.4). But then neither  $a$  nor  $b$  are units, for otherwise since  $I$  is an ideal, since  $a \cdot b \in I$ , it is then true that  $a^{-1} \cdot a \cdot b \in I$ , and thus  $b \in I$ , a contradiction. But then the ideal generated by  $I \cup \{a\}$  is a proper ideal since it does not contain 1. But then  $I \subsetneq (I \cup \{a\})$  since  $a \notin I$ , a contradiction since  $I$  is maximal. Thus,  $I$  is prime.  $\square$

**Theorem 16.4.11.** *If  $(R, +, \cdot)$  is a non-zero ring, then there exists a prime ideal  $I \subseteq R$ .*

*Proof.* For by Krull's theorem, there exists a maximal ideal  $I \subseteq R$  (Thm. 16.4.9). But maximal ideals are prime ideals (Thm. 16.4.10). Hence,  $I$  is a prime ideal.  $\square$

We now return to fields for a moment to develop their structure.

**Theorem 16.4.12.** *If  $(F, +_F, \cdot_F)$  and  $(K, +_K, \cdot_K)$  are fields, and if  $\phi : F \rightarrow K$  is a field homomorphism, then it is injective.*

*Proof.* For  $\phi^{-1}[\{0_K\}]$  is an ideal in  $F$  (Thm. 16.4.6), and since it doesn't contain 1 by the definition of a field homomorphism, it must be a proper ideal. But the only proper ideal of a field is the zero ideal (Thm. 16.4.3), and thus  $\phi^{-1}[\{0_K\}] = \{0_F\}$ . But then if  $\phi(a) = \phi(b)$ , then  $\phi(a - b) = 0_K$ , and thus  $a - b = 0_F$ . That is,  $a = b$ .  $\square$

**Theorem 16.4.13.** *If  $(F, +_F, \cdot_F)$  is a field, if  $(\mathbb{Z}, +, \cdot)$  is the standard ring of integers, if  $(\mathbb{Q}, +, \cdot)$  is the field of rational numbers, and if  $f : \mathbb{Z} \rightarrow F$  is an injective ring homomorphism, then there exists a field homomorphism  $\tilde{f} : \mathbb{Q} \rightarrow F$ .*

*Proof.* For any  $n \in \mathbb{N}^+$ , we have:

$$f(n) = f\left(\sum_{k \in \mathbb{Z}_n} 1\right) = \sum_{k \in \mathbb{Z}_n} f(1) \quad (16.4.2)$$

and similarly for any negative integer. But  $f$  is injective, and hence  $f(1) \neq 0_R$ . Thus, since  $(F, +_F, \cdot_F)$  is a field, for all  $m \in \mathbb{Z} \setminus \{0\}$  there is a unique  $x \in F$  such that  $f(m) \cdot_F x = 1_F$ . Label  $x$  as  $f(m)^{-1}$  and define  $\tilde{f}$  as follows:

$$\tilde{f}((n, m)) = \begin{cases} 0, & n = 0 \\ f(n) \cdot_F f(m)^{-1}, & \text{otherwise} \end{cases} \quad (16.4.3)$$

$\square$

**Definition 16.4.6** A field of characteristic zero is a field  $(F, +, \cdot)$  such that there exists a field homomorphism  $f : \mathbb{Q} \rightarrow F$ .

Equivalently one could say that adding  $1_F$  to itself  $n$  times never results in zero.

**Example 16.4.6**  $\mathbb{Q}$ ,  $\mathbb{R}$ , and  $\mathbb{C}$  are fields of characteristic zero.

**Definition 16.4.7** The characteristic of a field  $(F, +, \cdot)$  is the smallest non-zero  $n \in \mathbb{N}$  such that:

$$\sum_{k \in \mathbb{Z}_n} 1_F = 0_F \quad (16.4.4)$$

**Theorem 16.4.14.** If  $n \in \mathbb{N}^+$  is a non-negative integer, and if  $(F, +, \cdot)$  is a field of characteristic  $n$ , then  $n$  is a prime number.

*Proof.* For suppose not. Then there are integers  $p, q \in \mathbb{N}^+$  such that  $p, q < n$  and  $p \cdot q = n$ . Let  $f : \mathbb{Z} \rightarrow F$  be defined by:

$$f(n) = \begin{cases} 0_F, & n = 0 \\ \sum_{k \in \mathbb{Z}_n} 1_F, & n > 0 \\ -\sum_{k \in \mathbb{Z}_{|n|}} 1_F, & n < 0 \end{cases} \quad (16.4.5)$$

This is a ring homomorphism, and hence  $f(p \cdot q) = f(p) \cdot f(q)$ . But  $f(n) = 0$ , and since all fields are integral domains, either  $f(p) = 0$  or  $f(q) = 0$ . But then there exists a positive integer smaller than  $n$  such that  $f(p) = 0$ , a contradiction as  $n$  is the characteristic of  $F$ . Hence,  $n$  is a prime.  $\square$

### Theorem 16.4.15: Binomial Theorem

If  $(R, +, \cdot)$  is a commutative ring, if  $a, b \in R$ , and if  $n \in \mathbb{N}$ , then:

$$(a + b)^n = \sum_{k \in \mathbb{Z}_n} \binom{n}{k} a^k b^{n-k}$$

where  $\binom{n}{k}x$  denotes the sum of  $x$  with itself  $\binom{n}{k}$  times.

**Theorem 16.4.16.** If  $p \in \mathbb{N}$  is a prime number, if  $r \in \mathbb{N}$  is such that  $1 \leq r$  and  $r \leq p^n - 1$ , then  $p$  divides  $\binom{p^n}{r}$ .

**Theorem 16.4.17.** If  $(F, +, \cdot)$  is a field of characteristic  $p \in \mathbb{N}^+$ , if  $n \in \mathbb{N}$ , and if  $a, b \in F$ , then:

$$(a + b)^{p^n} = a^{p^n} + b^{p^n} \quad (16.4.6)$$

*Proof.* Apply the binomial in combination with the previous theorem.  $\square$

## 16.5 Polynomial Rings

### Definition 16.5.1: Polynomial Ring

The ring of polynomials over a ring  $(R, +_R, \cdot_R)$  is the set  $R[x]$  of all finitely supported sequences  $a : \mathbb{N} \rightarrow F$  with the following addition and multiplication:

$$(a + b)_n = a_n +_F b_n \quad (16.5.1)$$

$$(ab)_n = \sum_{k=0}^n a_k \cdot_F b_{n-k} \quad (16.5.2)$$

This is very much mimicing polynomials. The sum rule says we simply add the coefficients of two polynomials, and the product is the Cauchy product of two polynomials. That is, we multiply  $(a_0 + a_1x + \cdots + a_nx^n)$  by  $(b_0 + b_1x + \cdots + b_nx^n)$  and collect the coefficients of all terms with order  $x^k$  and group them. The resulting coefficient is precisely this sum. We now show that, given a ring  $(R, +_R, \cdot_R)$ , the polynomial ring  $R[x]$  is indeed a ring. That is,  $(R, +)$  is a group,  $(R, \cdot)$  is a monoid, and  $\cdot$  distributes over  $+$ .

**Theorem 16.5.1.** *If  $(R, +_R, \cdot_R)$  is a ring, then  $(R[x], +, \cdot)$  is a ring.*

*Proof.* For let  $e : \mathbb{N} \rightarrow R$  be the sequence defined by  $e_k = 0_R$  for all  $k \in \mathbb{N}$ . Then  $e$  is finitely supported, and hence  $e \in R[x]$ , and moreover  $e$  is a unital element with respect to  $+$ . For if  $f \in R[x]$ , then for all  $n \in \mathbb{N}$  we have:

$$(e + f)(n) = e_n +_R f_n = 0_R + f_n = f_n \quad (16.5.3)$$

and hence  $e + f = f$ . Similarly,  $f + e = f$ , and therefore  $e$  is a unital element. If  $f \in R[x]$ , let  $-f$  be the sequence defined by  $(-f)(n) = -1_R \cdot (f(n))$ . Since  $f$  is finitely supported,  $-f$  is also finitely supported, and hence an element of  $R[x]$ , and moreover for all  $n \in \mathbb{N}$  we have:

$$(f + (-f))(n) = f(n) +_R (-f)(n) \quad (16.5.4a)$$

$$= f(n) +_R (-1_R) \cdot f(n) \quad (16.5.4b)$$

$$= f(n) \cdot_R (1 + (-1)) \quad (16.5.4c)$$

$$= f(n) \cdot_R 0 \quad (16.5.4d)$$

$$= 0 \quad (\text{Thm. 16.3.3})$$

That is,  $f + (-f) = e$ , and hence  $f$  has an additive inverse. Lastly,  $+$  is associative, for if  $f, g, h \in R[x]$ , if  $A = (f + g) + h$ , and if  $B = f + (g + h)$ , then

for all  $n \in \mathbb{N}$  we obtain:

$$A(n) = ((f + g) + h)(n) \quad (16.5.5a)$$

$$= (f + g)(n) +_r h(n) \quad (16.5.5b)$$

$$= (f(n) +_R g(n)) +_R h(n) \quad (16.5.5c)$$

$$= f(n) +_R (g(n) +_R h(n)) \quad (16.5.5d)$$

$$= f(n) +_R (g + h)(n) \quad (16.5.5e)$$

$$= (f + (g + h))(n) \quad (16.5.5f)$$

$$= B(n) \quad (16.5.5g)$$

and therefore  $+$  is associative. Next,  $(R, \cdot)$  is a monoid. For let  $I : \mathbb{N} \rightarrow R$  be defined by  $I_0 = 1$  and  $I_k = 0$  for all  $k \in \mathbb{N}^+$ . Then for all  $f \in R[x]$ , we have:

$$(f \cdot I)(n) = \sum_{k=0}^n f(k) \cdot_R I(n - k) = f(n) \quad (16.5.6)$$

since  $I(n - k) = 0$  for all  $k \neq 0$ . Thus,  $f \cdot I = f$ , and similarly  $I \cdot f = f$ . Therefore  $I$  is a multiplicative identity. Multiplication is also associative, for if  $f, g, h \in R[x]$ , and if  $A = (f \cdot g) \cdot h$  and  $B = f \cdot (g \cdot h)$ , then:

$$A(n) = ((f \cdot g) \cdot h)(n) = \sum_{k=0}^n (f \cdot g)(k) \cdot_R h(n - k) \quad (16.5.7a)$$

$$= \sum_{k=0}^n \left( \sum_{j=0}^n f(j) \cdot_R g(n - j) \right) \cdot_R h(n - k) \quad (16.5.7b)$$

□

**Theorem 16.5.2.** *If  $(F, +, \cdot)$  is a field, then  $(F[x], +, \cdot)$  is a commutative algebra over  $F$ .*

There is a natural embedding of  $F$  into  $F[x]$  by looking at the subspace of all sequences  $a : \mathbb{N} \rightarrow F$  such that  $a_k = 0$  for all  $k > 0$ . That is, the only possible non-zero term is  $a_0$ .

**Theorem 16.5.3.** *If  $(F, +_F, \cdot_F)$  is a field, if  $(R, +_R, \cdot_R)$  is a ring, if  $F \subseteq R$  is a subring, and if  $r \in R$ , then there is a unique homomorphism  $f : F[x] \rightarrow R$  such that for all  $a \in F$ ,  $f(a) = r$ .*

**Definition 16.5.2** The degree of a non-zero polynomial  $a \in F[x]$  is the largest  $n \in \mathbb{N}$  such that  $a_n \neq 0$ . The degree of the zero polynomial is zero.

Since  $F[x]$  is the space of all finitely supported sequences, for any such  $a \in F[x]$  there will eventually be an  $N \in \mathbb{N}$  such that for all  $n > N$  it is true that  $a_n = 0$ .

By the well-ordering all of the integers, there will be a least such element, and hence the above definition is well defined. There's a natural way of looking at  $F[x]$  as a subset of  $\mathcal{F}(F, F)$ , the set of all functions from  $F$  to itself. Given  $a \in F[x]$  of degree  $n \in \mathbb{N}$  we consider the function  $f \in \mathcal{F}(F, F)$  defined by:

$$f(x) = \sum_{k \in \mathbb{Z}_{n+1}} a_k \cdot_F x^k = a_0 +_F a_1 \cdot_k x +_F a_2 \cdot_F x^2 +_F \cdots +_F a_n \cdot_F x^n \quad (16.5.8)$$

In more familiar language (dropping the subscripts and using concatenation to denote multiplication), we have:

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \quad (16.5.9)$$

The sequence definition is good for rigor and solving theorems, since the Cauchy product allows one to easily manipulate expressions without worrying about the non-existent dummy variable  $x$ , whereas the function definition (as a subset of  $\mathcal{F}(F, F)$ ) is good for intuition.

**Definition 16.5.3** A monic polynomial of degree  $n$  in a field  $(F, +, \cdot)$  is a polynomial  $a \in F[x]$  of degree  $n \in \mathbb{N}$  such that  $a_n = 1$ .

**Theorem 16.5.4.** *If  $(F, +, \cdot)$  is a field, then  $F[x]$  is a Euclidean domain.*

**Theorem 16.5.5.** *If  $(F, +, \cdot)$  is a field, if  $(I, +_I, \cdot_I)$  is an ideal in  $F[x]$ , and if  $a \in I$  is of least degree, then  $I = (a)$ , where  $(a)$  is the ideal generated by  $a$ .*

*Proof.* For given  $b \in (a)$ , there is an  $r \in F[x]$  such that  $b = r \cdot a$ . But  $I$  is an ideal, and  $a \in I$ , and hence  $b \in I$ . Thus,  $(a) \subseteq I$ . If  $b \in I$ , then by the division algorithm there are polynomials  $p, q \in F[x]$  such that  $b = aq + r$  where the degree of  $r$  is strictly less than the degree of  $a$ . But then  $r = b - aq$ . And since  $b \in I$  and  $q \in F[x]$ , it is true that  $bq \in I$  since  $I$  is an ideal. But if  $a \in I$  and  $bq \in I$ , then  $b - aq \in I$  since  $I$  is an ideal. Thus,  $r \in I$ . But  $r$  is a polynomial of degree strictly less than  $a$ , and  $a$  is a non-zero polynomial of least degree in  $I$ . Therefore,  $r = 0$ . But then  $b = aq$ , and hence  $b \in (a)$ . Thus,  $I \subseteq (a)$ . From the definition of equality,  $I = (a)$ .  $\square$

**Theorem 16.5.6.** *There exists a bijection between monic polynomials in  $F[x]$  and the ideals of  $F[x]$ .*

**Definition 16.5.4** A root of a polynomial  $a \in F[x]$  of degree  $n \in \mathbb{N}$  over a field  $(F, +, \cdot)$  is an element  $r \in F$  such that:

$$\sum_{k \in \mathbb{Z}_{n+1}} a_k \cdot_F x^k = 0_F \quad (16.5.10)$$

**Theorem 16.5.7.** *If  $(\mathbb{Q}, +, \cdot)$  is the field of rational numbers, if  $a \in \mathbb{Q}[x]$  is a polynomial of degree  $n \in \mathbb{N}$ , if  $r \in \mathbb{Q}$  is a root of  $a$ , and if  $p, q \in \mathbb{Q}$  are such that  $r = p/q$  and  $\text{GCD}(p, q) = 1$ , then  $p$  divides  $a_0$  and  $q$  divides  $a_n$ .*

*Proof.* For if  $r$  is a root, then:

$$\sum_{k \in \mathbb{Z}_{n+1}} a_k \cdot r^k = 0 \quad (16.5.11)$$

But  $r = p/q$ , and thus:

$$\sum_{k \in \mathbb{Z}_{n+1}} a_k \cdot \left(\frac{p}{q}\right)^k = 0 \quad (16.5.12)$$

Simplifying, and multiplying both sides by  $q^k$ , we have:

$$\sum_{k \in \mathbb{Z}_{n+1}} a_k \cdot p^k q^{n-k} = 0 \quad (16.5.13)$$

From this,  $q$  divides  $a_n p^n$ . But  $\text{GCD}(p, q) = 1$ , and hence  $q$  does not divide  $p^n$ . Thus,  $q$  divides  $a_n$ . Similarly,  $p$  divides  $a_0$ .  $\square$

**Example 16.5.1** Sticking with  $\mathbb{Q}[x]$ , the polynomial  $f(x) = x^3 - 3x - 1$  is irreducible over  $\mathbb{Q}$ . By the previous theorem, the only possible roots  $p/q$  must be such that  $p$  divides  $-1$  and  $q$  divides  $1$ . Hence,  $p/q = \pm 1$ . But  $f(1) = -3$  and  $f(-1) = 1$ , neither of which are zero. Hence,  $f$  has no roots over  $\mathbb{Q}$ . Since it is a cubic, it must be irreducible.

### Theorem 16.5.8: Gauss's Lemma for Polynomials

If  $(\mathbb{Z}, +, \cdot)$  is the ring of integers, if  $(\mathbb{Q}, +, \cdot)$  is the field of rational numbers, if  $a \in \mathbb{Z}[x]$  is such that  $a$  factors non-trivially in  $\mathbb{Q}[x]$ , then  $a$  factors non-trivially in  $\mathbb{Z}[x]$ .

*Proof.* For if  $b, c \in \mathbb{Q}[x]$  are such that  $a = b \cdot c$ , if  $M, N \in \mathbb{N}$  are the products of the denominators of  $b$  and  $c$ , respectively, then  $Ma, Nb \in \mathbb{Z}[x]$ . But then  $NMa \in \mathbb{Z}[x]$ . By the fundamental theorem of arithmetic, there exists a prime  $p \in \mathbb{N}$  such that  $p$  divides  $MN$ . But then  $Ma \cdot Nb$  is the zero polynomial in  $\mathbb{F}_p[x]$  and since  $\mathbb{F}_p$  is a field, this means that  $p$  divides the coefficients of every element of either  $Ma$  or  $Nb$ . Continuing we remove all of the prime factor of  $MNa$  and obtain a factorization of  $a$  in  $\mathbb{Z}[x]$ .  $\square$

**Theorem 16.5.9.** If  $(\mathbb{Z}, +, \cdot)$  is the ring of integers, if  $a \in \mathbb{Z}[x]$  is a monic polynomial of degree  $n \in \mathbb{N}$ , and if  $b \in \mathbb{Q}[x]$  is a monic factor of  $a$ , then  $b \in \mathbb{Z}[x]$ .

*Proof.* For if  $b, c \in \mathbb{Q}[x]$  are such that  $a = bc$ , with  $b$  a monic polynomial, then by the Cauchy product, since  $a$  is monic,  $c$  must also be monic. Let  $m$  and  $n$  by the least integers such that  $mb, nc \in \mathbb{Z}[x]$ . If  $p \in \mathbb{N}$  is a prime that divides

$mn$ , then it divides all of the coefficients of either  $mb$  or  $nc$  and hence either  $(m/p)b \in \mathbb{Z}[x]$  or  $(n/p)c \in \mathbb{Z}[x]$ , a contradiction since  $m$  and  $n$  are the least such integers with this property. Hence,  $m = 1$  and  $n = 1$ .  $\square$

**Definition 16.5.5** An algebraic integer in  $\mathbb{C}$  is a complex number  $z \in \mathbb{C}$  such that  $z$  is the root of a monic polynomial  $a \in \mathbb{Z}[x]$ .

### Theorem 16.5.10: Eisenstein's Criterion

If  $(\mathbb{Z}, +, \cdot)$  is the ring of integers, if  $a \in \mathbb{Z}[x]$  is a polynomial in  $\mathbb{Z}$  of degree  $n \in \mathbb{N}$ , if  $p \in \mathbb{N}$  is a prime number such that  $p$  does not divide  $a_n$ ,  $p^2$  does not divide  $a_0$ , and such that  $p$  divides  $a_k$  for all  $k \in \mathbb{Z}_n$ , then  $a$  is irreducible in  $\mathbb{Q}[x]$ .

*Proof.* For suppose  $a$  factors in  $\mathbb{Q}[x]$ . But then it factors in  $\mathbb{Z}[x]$ . Suppose  $b, c \in \mathbb{Z}[x]$  are non-trivial factors. By the Cauchy product,  $a_0 = b_0c_0$ , and thus  $p$  divides either  $b_0$  or  $c_0$ . Suppose it divides  $b_0$ . But since  $p^2$  does not divide  $a_0$ ,  $p$  does not divide  $c_0$ . But again by the Cauchy product,  $a_1 = b_0c_1 + b_1c_0$ . But  $p$  divides  $a_1$ , and hence  $p$  divides  $b_1$ . Continuing by induction on the Cauchy product,  $p$  divides all of the  $b_k$ , contradicting that  $p$  does not divide  $a_n$ .  $\square$

It is important again to note that  $\mathbb{Z}_n = \{0, 1, \dots, n - 1\}$ , hence  $\mathbb{Z}_n$  does not contain  $n$ . So we require  $p$  to divide  $a_0, \dots, a_{n-1}$ , but not  $a_n$ , and we require  $p^2$  not to divide  $a_0$ . Eisenstein's criterion holds for any unique factorization domain.

**Theorem 16.5.11.** If  $(\mathbb{Z}, +, \cdot)$  is the ring of integers, if  $a \in \mathbb{Z}[x]$  is a reducible polynomial in  $\mathbb{Z}$  of degree  $n \in \mathbb{N}$ , if  $p \in \mathbb{N}$  is a prime, and if  $p$  does not divide  $a_n$ , then  $\bar{a} \in \mathbb{F}_p[x]$  is reducible.

*Proof.* For since  $a \in \mathbb{Z}[x]$  is reducible, there exists  $b, c \in \mathbb{Z}[x]$  such that  $a = b \cdot c$ . But since  $p$  does not divide  $a_n$ , it does not divide  $a_0b_n + a_nb_0$ , and thus  $\bar{a} = \bar{b} \cdot \bar{c}$  is a non-trivial factorization. Hence,  $\bar{a}$  is reducible in  $\mathbb{F}_p$ .  $\square$

A more useful result is the contrapositive.

**Theorem 16.5.12.** If  $(\mathbb{Z}, +, \cdot)$  is the ring of integers, if  $p \in \mathbb{N}$  is a prime, if  $(\mathbb{F}_p, +_p, \cdot_p)$  is the field of integers modulo  $p$ , if  $a \in \mathbb{Z}[x]$  is a polynomial of degree  $n \in \mathbb{N}$  such that  $p$  does not divide  $a_n$ , and if  $\bar{a} \in \mathbb{F}_p[x]$  is irreducible, then  $a$  is irreducible in  $\mathbb{Z}[x]$ .

The converse of this theorem does not hold. Indeed, there exist polynomials  $a \in \mathbb{Z}[x]$  that are reducible in  $\mathbb{F}_p[x]$  for every prime integer  $p \in \mathbb{N}$ , yet  $a$  is not reducible in  $\mathbb{Z}[x]$ .

**Theorem 16.5.13.** If  $(F, +, \cdot)$  is a field, if  $a \in F$ , and if  $f \in F[x]$ , then there exists a polynomial  $q \in F[x]$  such that, for all  $x \in F$  it is true that:

$$f(x) = q(x) \cdot (x - a) + f(a) \quad (16.5.14)$$

*Proof.* For by the division algorithm, there exists polynomials  $q, r \in F[x]$  such that  $f(x) = q(x) \cdot (x - a) + r(x)$  and  $r$  has degree less than  $x - a$ . But the degree of  $x - a$  is 1, and hence  $r$  is a constant (has degree zero). Moreover:

$$f(a) = q(a) \cdot (a - a) + r(a) = q(a) \cdot 0 + r(a) = 0 + r(a) = r(a) \quad (16.5.15)$$

Therefore  $r(x) = f(a)$ . Thus, for all  $x \in F$ ,  $f(x) = q(x) \cdot (x - a) + f(a)$ .  $\square$

**Theorem 16.5.14.** If  $(F, +, \cdot)$  is a field, if  $f \in F[x]$  is a polynomial, if  $a \in F$ , and if  $x - a$  divides  $f(x)$ , then  $f(a) = 0$ .

*Proof.* For by the previous theorem, there is a polynomial  $q \in F[x]$  such that  $f(x) = q(x) \cdot (x - a) + f(a)$ . But if  $f(a) = 0$ , then  $f(x) = q(x) \cdot (x - a)$ , and hence  $x - a$  divides  $f$ . In the other direction, if  $x - a$  divides  $f(x)$ , then there is a  $q \in F[x]$  such that  $f(x) = q(x) \cdot (x - a)$ . But then  $f(a) = q(a) \cdot (a - a)$ , and thus  $f(a) = 0$ .  $\square$

**Theorem 16.5.15.** If  $(F, +, \cdot)$  is a field, if  $f \in F[x]$  is a non-zero polynomial in  $F$  of degree  $n \in \mathbb{N}$ , the there are at most  $n$  roots.

*Proof.* For by induction. If  $f \in F[x]$  is a polynomial of degree 1, then  $f(x) = ax + b$  for some  $a, b \in F$ . But if  $\alpha$  is a root, then  $f(x) = q(x) \cdot (x - \alpha)$ . But since  $f$  is degree 1, and since  $x - \alpha$  is degree 1,  $q$  must be of degree 0, and hence a constant. But  $f$  is non-zero, and hence  $\alpha \neq 0$ . Thus  $f(x) = 0$  if and only if  $x = \alpha$ . Suppose the proposition is true for  $n \in \mathbb{N}$ , and let  $f \in F[x]$  be a non-zero polynomial of degree  $n + 1$ . Suppose  $f$  has more than  $n + 1$  roots. But if  $\alpha$  is a root, then  $x - \alpha$  divides  $f$ . Hence  $f(x) = q(x) \cdot (x - \alpha)$ , where  $q$  is a polynomial of degree less than  $f$ . But by hypothesis  $q$  then has at most  $n$  roots. And any root of  $q$  is a root of  $f$ , and hence  $f$  has at most  $n + 1$  roots, a contradiction. Thus,  $f$  has at most  $n + 1$  roots.  $\square$

**Theorem 16.5.16.** If  $(G, *)$  is a finite Abelian group, if  $m \in \mathbb{N}$  is such that  $m$  divides  $\text{Card}(G)$ , then there are at most  $m$  elements in  $G$  whose order divides  $m$  if and only if  $G$  is cyclic.

*Proof.* For if  $G$  is cycle, and if  $d$  divides  $\text{Card}(G)$ , let  $G_d = \{x \in G \mid x^d = 1\}$ . If  $G_d$  is empty, then  $\text{Card}(G_d) < m$ . If not then there is a  $y \in G_d$ . But then  $\langle y \rangle \subseteq G_d$ . For if  $x \in \langle y \rangle$ , then  $x = y^k$  for some  $k \in \mathbb{Z}_n$ . But then  $x^d = (y^k)^d = y^{kd} = 1$ , and hence  $x \in G_d$ . Therefore,  $\langle y \rangle \subseteq G_d$ . Now let  $x \in G_d$ . Since  $(G, *)$  is cyclic there exists  $z \in G$  such that  $\langle z \rangle = G$ . But then there exists  $k_1, k_2 \in \mathbb{Z}_n$  such that  $z^{k_1} = y$  and  $z^{k_2} = x$ . But then  $y^{k_2 - k_1} = x$ ,

and therefore  $x \in \langle y \rangle$ . Thus,  $\langle y \rangle = G_d$ . But  $\langle y \rangle$  has  $d$  elements, and hence there are at most  $d$  elements of order  $d$ . Going the other way, suppose that if  $d$  divides  $n$  then there are at most  $d$  elements of order  $d$ . Since  $(G, *)$  is finite and Abelian, it is the product of finite cyclic groups  $(\mathbb{Z}_{p_k^{n_k}}, +)$ . Suppose two of the primes  $p_k$  are equal. But then there are at least  $p_k^2$  elements of order  $p_k$ , and  $p_k$  divides the order of the group, a contradiction. Thus, all of the primes are distinct. By the direct product of coprime cyclic groups is a cyclic group. Thus,  $(G, *)$  is cyclic.  $\square$

**Theorem 16.5.17.** *If  $(F, +, \cdot)$  is a field, if  $(F \setminus \{0\}, \cdot)$  is the multiplicative group of  $F$ , then it is cyclic.*

*Proof.* Since  $(F, +, \cdot)$  is a field,  $(F \setminus \{0\}, \cdot)$  is Abelian. Suppose  $d \in \mathbb{N}$  divides the cardinality of  $F \setminus \{0\}$  and let  $f \in F[x]$  be the polynomial  $f(x) = x^d - 1$ . Then by the previous theorem there are at most  $d$  roots. But then there are at most  $d$  elements of  $F \setminus \{0\}$  such that  $x^d = 1$ , and thus by the previous theorem  $(F \setminus \{0\}, \cdot)$  is cyclic.  $\square$

With this we now return to the claim that there are irreducible polynomials in  $\mathbb{Q}[x]$  that are reducible if  $\mathbb{Z}_p[x]$  for every prime  $p \in \mathbb{N}$ .

**Theorem 16.5.18.** *If  $(G, *)$  is a cyclic group, if  $a, b \in G$  do not have square roots, then  $a * b$  has a square root.*

*Proof.* For since  $(G, *)$  is cyclic there is an  $x \in G$  such that  $\langle x \rangle = G$ . But then if  $a, b \in G$ , there exists  $n, m \in \mathbb{N}$  such that  $x^n = a$  and  $x^m = b$ . But  $a$  and  $b$  do not have square roots, and hence  $n$  and  $m$  are odd. But then  $a * b = x^n * x^m = x^{n+m}$ , and  $n + m$  is even. Hence,  $x^{n+m}$  has a square root, and thus  $a * b$  has a square root.  $\square$

**Example 16.5.2** The polynomial  $f(x) = x^4 - 10x^2 + 1$  is irreducible in  $\mathbb{Z}[x]$ , yet it is reducible in  $\mathbb{F}_p[x]$  for every prime  $p \in \mathbb{N}$ . If  $p \in \mathbb{N}$  is such that 2 has a square root in  $\mathbb{Z}_p$ , then we can factor this to obtain:

$$(x^2 - 2\sqrt{2}x - 1)(x^2 + 2\sqrt{2}x - 1) = x^4 - 10x^2 + 1 \quad (16.5.16)$$

And hence  $f$  is reducible over all such  $\mathbb{Z}_p[x]$ . If  $p$  is such that 3 has a square root, then:

$$(x^2 - 2\sqrt{3}x - 1)(x^2 + 2\sqrt{2}x + 1) = x^4 - 10x^2 + 1 \quad (16.5.17)$$

For all other such  $p$ , neither 2 nor 3 are square roots, and hence by the previous theorem their product  $2 \cdot 3 = 6$  does have a square root. But then:

$$x^4 - 10x^2 + 1 = (x^2 - (5 + 2\sqrt{6}))(x^2 - (5 - 2\sqrt{6})) \quad (16.5.18)$$

There is a stronger result that for every non-prime  $n \in \mathbb{N}^+$  there is a polynomial  $f \in \mathbb{Z}[x]$  of degree  $n$  such that  $f$  is irreducible over  $\mathbb{Z}[x]$  but reducible over  $\mathbb{Z}_p[x]$  for every prime  $p$ . (See Brandl, R. American Mathematical Monthly, 1986).

**Definition 16.5.6** An extension field of a field  $(F, +_F, \cdot_F)$  is a field  $(K, +_K, \cdot_K)$  such that  $(F, +_F, \cdot_F)$  is a subfield of  $(K, +_K, \cdot_K)$ .

Thus, a field extension is to fields as supersets are to sets. Unlike set theory where one considers subsets most of the time, most of field theory is obsessed with field extensions.

**Theorem 16.5.19.** *If  $(F, +_F, \cdot_F)$  and  $(K, +_K, \cdot_K)$  are fields, and if  $K$  is an extension field of  $F$ , then  $(K, +_K, \cdot_K)$  is a vector field over  $(F, +_F, \cdot_F)$ .*

**Definition 16.5.7** The degree of a field extension  $(K, +_K, \cdot_K)$  of a given field  $(F, +_F, \cdot_F)$  is the dimension of the vector space  $(K, +_K, \cdot_K)$  over  $(F, +_F, \cdot_F)$ . This is denoted  $[K : F]$ .

**Definition 16.5.8** A finite extension of a field  $(F, +_F, \cdot_F)$  is an extension field  $(K, +_K, \cdot_K)$  of  $F$  such that  $[K : F] \in \mathbb{N}$ .

**Example 16.5.3** The field of complex numbers  $\mathbb{C}$  is a field extension over  $\mathbb{R}$ . Moreover, it is a finite extension since  $\mathbb{C}$  has the bases  $\{1, i\}$ . Thus,  $[\mathbb{C} : \mathbb{R}] = 2$

**Example 16.5.4** The field  $\mathbb{R}$  is also a field extension over the rational numbers  $\mathbb{Q}$ . However, unlike  $[\mathbb{C} : \mathbb{R}]$ , which is finite,  $[\mathbb{R} : \mathbb{Q}] = \text{Card}(\mathbb{R})$ . This can be shown since any finite extension of  $\mathbb{Q}$  must be countable, and any countable extension must have cardinality  $\text{Card}(\mathbb{Q} \times \mathbb{Q})$ , but this is equal to  $\text{Card}(\mathbb{N})$ , and  $\mathbb{R}$  is uncountable. Hence, at the very least,  $[\mathbb{Q} : \mathbb{R}]$  is uncountable, and if we assume the continuum hypothesis then the cardinality must then be equal to the cardinality of  $\mathbb{R}$  (for it is certinaly not greater).

**Example 16.5.5** The Gaussian numbers are a subfield of  $\mathbb{C}$  defined as follows:

$$\mathbb{Q}(i) = \{a + ib \in \mathbb{C} \mid a, b \in \mathbb{Q}\} \quad (16.5.19)$$

We can then see that  $[\mathbb{Q} : \mathbb{Q}(i)] = 2$  since this vector field has  $\{1, i\}$  as a basis.

**Example 16.5.6** While  $[\mathbb{Q} : \mathbb{R}]$  is uncountable, there are countably infinite extension fields. Given any field  $(F, +, \cdot)$ , the field of rational functions  $F(x)$  has the countable basis  $x^n$  for all  $n \in \mathbb{N}$ . The subspace of polynomials  $F[x]$ , while not a field, is still a vector space over  $F$  and has the same basis.

**Theorem 16.5.20.** *If  $(F, +_F, \cdot_F)$  is a field, if  $(K, +_K, \cdot_K)$  is a finite field extension of  $F$ , and  $(L, +_L, \cdot_L)$  is a finite field extension of  $K$ , then  $L$  is a finite field extension of  $F$  and:*

$$[L : F] = [L : K][K : F] \quad (16.5.20)$$

*Proof.* For if  $(K, +_K, \cdot_K)$  is a finite field extension of  $(F, +_F, \cdot_F)$ , then it is a finite dimensional vector space over  $F$  and thus there is an  $m \in \mathbb{N}$  and a finite sequence  $a : \mathbb{Z}_m \rightarrow K$  such that  $a[\mathbb{Z}_m]$  is a basis for  $K$  over  $F$ . And similarly if  $(L, +_L, \cdot_L)$  is a finite field extension over  $(K, +_K, \cdot_K)$  then there is an  $n \in \mathbb{N}$  and a finite sequence  $b : \mathbb{Z}_n \rightarrow L$  such that  $b[\mathbb{Z}_n]$  is a basis for  $L$  over  $K$ . Define  $A : \mathbb{Z}_m \times \mathbb{Z}_n \rightarrow L$  by  $A(i, j) = a_i \cdot b_j$  for all  $(i, j) \in \mathbb{Z}_m \times \mathbb{Z}_n$ . Let  $f : \mathbb{Z}_{m \cdot n} \rightarrow \mathbb{Z}_m \times \mathbb{Z}_n$  be a bijection and define  $e : \mathbb{Z}_{n \cdot m} \rightarrow L$  by  $e = A \circ f$ . Suppose  $e[\mathbb{Z}_{n \cdot m}]$  does not span all of  $L$  over  $F$ . Then there exists  $x \in L$  such that for every sequence  $c : \mathbb{Z}_{n \cdot m} \rightarrow F$ ,  $\alpha$  is not the sum over  $c_i \cdot e_i$ . But  $b[\mathbb{Z}_n]$  forms a basis of  $L$  over  $K$ , and thus there is a sequence  $\beta : \mathbb{Z}_n \rightarrow K$  such that:

$$\alpha = \sum_{k \in \mathbb{Z}_n} \beta_k \cdot_L b_k \quad (16.5.21)$$

But for each  $k$  it is true that  $\beta_k \in K$ , and since  $a[\mathbb{Z}_m]$  is a basis for  $K$  over  $F$ , there is a sequence  $\gamma_k : \mathbb{Z}_m \rightarrow F$  such that:

$$\beta_k = \sum_{i \in \mathbb{Z}_m} \gamma_i \cdot_K a_i \quad (16.5.22)$$

Let  $c : \mathbb{Z}_m \times \mathbb{Z}_n \rightarrow F$  be defined by  $c(i, j) = \gamma_i$  and let  $d : \mathbb{Z}_{m \cdot n} \rightarrow F$  be defined by  $d = c \circ f$ . But then:

$$\alpha = \sum_{j \in \mathbb{Z}_n} \beta_j \cdot_L b_j \quad (16.5.23)$$

$$= \sum_{j \in \mathbb{Z}_n} \left( \sum_{i \in \mathbb{Z}_m} \gamma_i \cdot_K a_i \right) \cdot_L b_j \quad (16.5.24)$$

$$= \sum_{j \in \mathbb{Z}_n} \sum_{i \in \mathbb{Z}_m} (\gamma_k \cdot_K (a_i \cdot_L b_j)) \quad (16.5.25)$$

$$= \sum_{j \in \mathbb{Z}_n} \sum_{i \in \mathbb{Z}_m} c(i, j) \cdot_K A(i, j) \quad (16.5.26)$$

$$= \sum_{k \in \mathbb{Z}_{n \cdot m}} ((c \cdot_K A) \circ f)(k) \quad (16.5.27)$$

$$= \sum_{k \in \mathbb{Z}_{n \cdot m}} d_k \cdot_K e_k \quad (16.5.28)$$

A contradiction. Hence,  $e[\mathbb{Z}_{n \cdot m}]$  does span  $L$ . Moreover, it is linearly independent. For if  $d : \mathbb{Z}_{m \cdot n} \rightarrow F$  is a sequence such that:

$$\sum_{k \in \mathbb{Z}_{m \cdot n}} d_k \cdot_K e_k = 0 \quad (16.5.29)$$

The composing with  $f^{-1}$  such that:

$$\sum_{j \in \mathbb{Z}_n} \sum_{i \in \mathbb{Z}_m} d_{ij} a_i b_j = 0 \quad (16.5.30)$$

From the linear independence of the  $a_i$  and the  $b_j$ , all of the  $d_{ij}$  are zero. Hence, by the basis theorem,  $[L : K] = n \cdot m$ .  $\square$

Given a field  $F$  and a monic irreducible polynomial of positive degree  $f \in F[x]$ , we can always obtain a field extension for  $F$  by considering the quotient  $F[x]/(f)$ , where  $(f)$  is the ideal generated by  $f$ .

**Example 16.5.7** Consider the field of real numbers  $(\mathbb{R}, +, \cdot)$  and let  $f(x) = x^2 + 1$ . This is irreducible over  $\mathbb{R}$  since it is a quadratic with no roots, as one can determine via the quadratic formula. If we consider  $\mathbb{R}[x]/(x^2 + 1)$ , we obtain the following arithmetic:

$$(a + bx) + (c + dx) = (a + c) + (b + d)x \quad (16.5.31)$$

Multiplication is carried out as follows:

$$(a + bx)(c + dx) = ac + (bc + ad)x + bdx^2 \quad (16.5.32)$$

But in the quotient we have  $x^2 + 1 = 0$ , and hence  $x^2 = -1$ . Thus, we can simplify this to:

$$(a + bx)(c + dx) = (ac - bd) + (ad + bc)x \quad (16.5.33)$$

And this is precisely the multiplicative structure on  $\mathbb{C}$ . Thus the field  $\mathbb{R}[x]/(x^2 + 1)$  is isomorphic to the complex numbers  $\mathbb{C}$  under the mapping  $\phi(1) = 1$  and  $\phi(x) = i$ .

Such fields are called stem fields. That is, stem fields are fields of the form  $F[x]/(f)$ . The intersection of subrings is a subring. The subring generated by a subset  $S \subseteq R$  in a ring  $(R, +, \cdot)$  is the intersection of all subrings containing  $S$ . Given a ring  $(R, +, \cdot)$  and a subring  $(S, +, \cdot)$ , and given any subset  $A \subseteq R$ , the subring generated by  $S$  over  $A$  is the intersection of all subrings of  $R$  that contains  $A \cup S$ .

**Example 16.5.8** Letting  $\mathbb{C}$  and  $\mathbb{R}$  have their usual field structures, and taking  $A \subseteq \mathbb{C}$  to be  $A = \{i\}$ , the subring generated by  $\mathbb{R}$  over  $A$  is simply the entire complex plane  $\mathbb{C}$ . We can write this as  $\mathbb{C} = \mathbb{R}[i]$  or  $\mathbb{C} = \mathbb{R}[\sqrt{-1}]$ .

The subring generated by a subset is equal to the set of all possible linear combinations of elements in the subset. Rigorously, we look at all finite sequence  $a : \mathbb{Z}_n \rightarrow F$  and  $b : \mathbb{Z}_n \rightarrow \mathcal{P}(S)$  such that  $b_k$  is finite for all  $k \in \mathbb{Z}_n$ , and we form the sums:

$$z = \sum_{k \in \mathbb{Z}_n} \left( a_k \prod_{\alpha \in b_k} \alpha \right) \quad (16.5.34)$$

**Theorem 16.5.21.** *If  $(R, +, \cdot)$  is an integral domain, if  $(F, +_F, \cdot_F)$  is a subring of  $R$  such that  $(F, +_F, \cdot_F)$  is also a field, and if  $R$  is a finite dimensional vector space over  $F$ , then  $(R, +, \cdot)$  is a field.*

*Proof.* For if not then there is a non-zero element  $a \in R$  with no inverse element. But the function  $f : R \rightarrow R$  defined by  $f(x) = a \cdot x$  is injective. For if  $f(x) = f(y)$ , then:

$$a \cdot x - a \cdot y = a \cdot (x - y) = 0 \quad (16.5.35)$$

But  $(R, +, \cdot)$  is an integral domain and thus either  $a = 0$  or  $x - y = 0$ . But by hypothesis  $a$  is non-zero, and hence  $x - y = 0$ . Thus,  $f$  is injective. But it is also linear since:

$$f(x + y) = a \cdot (x + y) = a \cdot x + a \cdot y = f(x) + f(y) \quad (16.5.36)$$

And thus  $f$  is a linear map from a finite dimensional vector space to itself and is therefore surjective. But then there exists  $x \in R$  such that  $f(x) = 1$ . But then  $a \cdot x = 1$ , a contradiction. Hence,  $(R, +, \cdot)$  is a field.  $\square$

**Theorem 16.5.22.** *If  $(F, +, \cdot)$  is a field, if  $(K, +, \cdot)$  is a finite extension field, and if  $(S, +, \cdot)$  is a subring of  $K$  that contains  $F$ , then it is a field.*

**Theorem 16.5.23.** *For such a ring will be an integral domain, and will also be a finite dimensional vector space over  $F$ , and hence will be a field by the previous theorem.*

The intersection of subfields is a field. We can similarly define field generated by subset. Given the ring generated by  $S$ ,  $F[S]$ , the field generated by  $S$   $F(S)$  is also called the field of fractions of  $F[S]$ . This is because it is obtained by considering all fractions of elements in  $F[S]$ . For example, in the polynomial ring  $F[x]$ , the field that this generates is the field of rational functions  $F(x)$ .

**Example 16.5.9** The ring  $\mathbb{Q}[\pi]$  is a subset of  $\mathbb{R}$  that is all linear combinations of powers of  $\pi$  with rational coefficients. The field generated by this  $\mathbb{Q}(\pi)$  is the set of all  $f(\pi)/g(\pi)$  where  $f, g$  are rational polynomials evaluated at  $\pi$ , and  $g$  is non-zero.

**Definition 16.5.9** A simple extension of a field  $(F, +, \cdot)$  is an extension field  $(K, +, \cdot)$  such that there exists  $\alpha \in K$  such that  $K = F(\alpha)$ .

**Definition 16.5.10** The composite of two subfields  $(F_1, +, \cdot)$  and  $(F_2, +, \cdot)$  of a field  $(F, +, \cdot)$  is the field generated by  $F_1 \cup F_2$ .

**Theorem 16.5.24.** *If  $(F, +, \cdot)$  is a field, if  $(F_1, +, \cdot)$  and  $(F_2, +, \cdot)$  are subfields, and if  $K$  is the composite of  $F_1$  and  $F_2$ , then  $K$  is the field generated by  $F_1$  over  $F_2$ , and  $K$  is the field generated by  $F_2$  over  $F_1$ .*

Given a field  $(F, +, \cdot)$  and an extension field  $(K, +, \cdot)$ , for any element  $\alpha \in E$  we can define the homomorphism  $\phi : F[x] \rightarrow E$  be  $\phi(f) = f(\alpha)$ . If the kernel of this is zero, then we say  $\alpha$  is transcendental. It then turns out that  $\phi : F[x] \rightarrow F[\alpha]$  is an isomorphism and this extends to an isomorphism  $\tilde{\phi} : F(x) \rightarrow F(\alpha)$ . In the latter case, where the kernel is non-zero, we call  $\alpha$  algebraic.

**Definition 16.5.11** An algebraic element of a field extension  $(K, +, \cdot)$  over a field  $(F, +, \cdot)$  is an element  $\alpha \in K$  such that there exists a polynomial  $f \in F[x]$  such that  $f(\alpha) = 0$ .

**Definition 16.5.12** A minimal polynomial of an algebraic element  $\alpha$  of a field extension  $(K, +, \cdot)$  over a field  $(F, +, \cdot)$  is a monic irreducible polynomial such that  $f(\alpha) = 0$ .

Minimal polynomials are unique. We can define  $\phi : F[x]/(f) \rightarrow F[\alpha]$  by  $\phi(\bar{f}) = f(\alpha)$ . Since  $F[x]/(f)$  is a field,  $F[\alpha]$  is as well. Therefore  $F[\alpha] = F(\alpha)$ , and  $F[\alpha]$  is a stem field.

**Definition 16.5.13** An algebraic extension field of a field  $(F, +, \cdot)$  is an extension field  $(K, +, \cdot)$  such that for all  $\alpha \in K$  it is true that  $\alpha$  is algebraic in  $F$ .

**Example 16.5.10** The set of algebraic numbers in  $\mathbb{R}$  are an algebraic extension over  $\mathbb{Q}$ . The complex numbers are an algebraic extension over  $\mathbb{R}$ . The real numbers are not an algebraic extension over  $\mathbb{Q}$  however, since  $\pi$  is transcendental. Indeed, the set of all algebraic numbers in  $\mathbb{R}$  form a countable subset, and so most real numbers are transcendental.

**Theorem 16.5.25.** *If  $(F, +, \cdot)$  is a field, and if  $(K, +, \cdot)$  is an extension field, then  $K$  is a finite extension if and only if it is algebraic and finitely generated.*

### Theorem 16.5.26: Liouville's Transcendental Theorem

If  $a \in \mathbb{N}$ , if  $a \geq 2$ , and if  $s \in \mathbb{R}$  is defined by:

$$s = \sum_{k \in \mathbb{N}} \frac{1}{a^{n!}} \quad (16.5.37)$$

then  $s$  is transcendental.

*Proof.* For suppose not. Then  $s$  is algebraic and thus there is a  $d \in \mathbb{N}$  and a minimal polyomial  $f \in \mathbb{Q}[x]$  of degree  $d$  such that  $f(s) = 0$ . But then  $\mathbb{Q}[s]$  is a  $d$  dimensional vector space over  $\mathbb{Q}$ . Let  $N \in \mathbb{N}$  be such that  $N \cdot f \in \mathbb{Z}[x]$  and let  $S_n$  be the  $n^{\text{th}}$  partial sums:

$$S_n = \sum_{k \in \mathbb{Z}_n} \frac{1}{a^{k!}} \quad (16.5.38)$$

Thus,  $|S_n - s| \rightarrow 0$ . If  $s$  is rational, then the minimal polynomial has degree one and hence  $f(x) = x - s$ . Otherwise, since  $f$  is irreducible and of degree

greater than one,  $f$  has no rational roots. Moreover, for all  $n \in \mathbb{N}$ ,  $S_n \neq s$  and hence  $f(S_n) \neq 0$ . Moreover, by induction  $S_n \in \mathbb{Q}$  for all  $n \in \mathbb{N}$ , and  $(a^{n!})^d DS_n$  is an integer, and therefore:

$$|(a^{n!})^d DS_n| \geq 1 \quad (16.5.39)$$

From the fundamental theorem of algebra,  $f$  splits into its roots and so:

$$f(x) = \prod_{k \in \mathbb{Z}_d} (x - \alpha_k) \quad (16.5.40)$$

Where the  $\alpha_k$  are the complex roots of  $f$ . Let  $M_1 = \max\{|\alpha_k| : k \in \mathbb{Z}_d \setminus \{0\}\}$ . That is,  $M$  is the largest modulus of all the roots, neglecting the zeroth term. Let  $M = \max\{M_1, 1\}$ . Then:

$$|f(S_n)| = \prod_{k \in \mathbb{Z}_d} |S_n - \alpha_k| \leq |S_n - \alpha_0| (S_n + M)^{d-1} \quad (16.5.41)$$

But we can simplify this further, since:

$$|S_n - \alpha_1| = \sum_{k=n+1}^{\infty} \frac{1}{a^{k!}} \leq \frac{1}{a^{(k+1)!}} \sum_{k \in \mathbb{N}} \frac{1}{a^k} = \frac{a}{a-1} \frac{1}{a^{(n+1)!}} \quad (16.5.42)$$

Piecing this together, we obtain:

$$|f(S_n)| \leq \frac{2}{2(n+1)!} (S_n + M)^{d-1} \quad (16.5.43)$$

And this converges to zero. Therefore:

$$|(a^{n!})^d DS_n| \leq \frac{a}{a-1} \frac{a^{d \cdot n!}}{a^{(n+1)!}} (S_n + M)^{d-1} \quad (16.5.44)$$

Which thus converges to zero, contradicting that this is always greater than one.  $\square$

**Definition 16.5.14** A splitting polynomial in a field  $(F, +, \cdot)$  is a polynomial  $f \in F[x]$  of degree  $n \in \mathbb{N}$  such that there exists finite sequences  $a, b : \mathbb{Z}_n \rightarrow F$  such that:

$$f(x) = \prod_{k \in \mathbb{Z}_n} (ax - b_k) \quad (16.5.45)$$

**Theorem 16.5.27.** *If  $(F, +, \cdot)$  is a field, then every non-constant polynomial  $f \in F[x]$  is a splitting polynomial if and only if every non-constant polynomial has at least one root in  $F$ .*

*Proof.* If every non-constant polynomial  $f \in F[x]$  splits, then given such an  $f$

of degree  $n \in \mathbb{N}$  there are sequences  $a, b : \mathbb{Z}_n \rightarrow F$  such that:

$$f(x) = \prod_{k \in \mathbb{Z}_n} (a_k x - b_k) \quad (16.5.46)$$

But then  $b^k/a_k$  is a root for all  $k$ , and hence there is at least one root. Going the other way, let  $f \in F[x]$  be a non-constant polynomial. Then there is a root  $\alpha$  and hence  $f(x) = (x - \alpha)g(x)$ . Then either  $g$  is a constant or a non-constant polynomial. In the latter case it has a root  $\beta$  and so  $f(x) = (x - \alpha)g(x) = (x - \alpha)(x - \beta)h(x)$ . Continuing inductively we find that  $f$  is a splitting polynomial.  $\square$

**Theorem 16.5.28.** *If  $(F, +, \cdot)$  is a field, then every non-constant polynomial  $f \in F[x]$  has at least one root in  $F$  if and only if every irreducible polynomial in  $F[x]$  has degree 1.*

*Proof.* For if  $f \in F[x]$  is a non-constant polynomial, then either the degree of  $f$  is 1 or it is greater than 1. But if the degree of  $f$  is 1, then  $f(x) = ax + b$  and thus  $f$  has a root. If the degree of  $f$  is greater than 1, then it is reducible and hence there are polynomials  $g, h \in F[x]$  of positive degree such that  $f = gh$ . Continuing inductively we eventually factor  $f$  down to the product of degree 1 polynomials, and thus  $f$  has a root. Going the other way, if  $f \in F[x]$  is non-constant and irreducible then it has a root, and thus  $f(x) = ax + b$ . Thus, the only irreducible polynomials have degree 1.  $\square$

**Theorem 16.5.29.** *If  $(F, +, \cdot)$  is a field, then every irreducible polynomial  $f \in F[x]$  has degree 1 if and only if for every finite field extension  $(K, +, \cdot)$  over  $F$ ,  $K = F$ .*

*Proof.* For if  $(K, +, \cdot)$  is a finite field extension over  $F$ , then the minimal polynomial  $f \in F[x]$  for any element  $\alpha \in K$  has degree 1 and hence  $f(x) = x - \alpha$ . But then  $\alpha \in F$ , and thus  $F = K$ . In the reverse direction, if  $f \in F[x]$  is an irreducible polynomial, then  $F[x]/(f)$  is a finite extension field of  $F$ . But by hypothesis,  $F[x]/(f) = F$ , and since the degree of  $f$  is equal to  $[F[x]/(f) : F] = 1$ ,  $f$  is a degree 1 polynomial.  $\square$

**Definition 16.5.15** An algebraically closed field is a field  $(F, +, \cdot)$  such that for all non-constant polynomials  $f \in F[x]$ , there exists a root  $\alpha \in F$  of  $f$ .

**Definition 16.5.16** An algebraic closure of a field  $(F, +, \cdot)$  is an extension field  $(K, +, \cdot)$  of  $F$  such that  $K$  is algebraically closed.

**Example 16.5.11** By the fundamental theorem of algebra,  $\mathbb{C}$  is algebraically closed. It is therefore an algebraic closure of  $\mathbb{R}$ . To note that  $\mathbb{R}$  is not algebraically closed one need only consider the polynomial  $f(x) = x^2 + 1$ .

**Theorem 16.5.30.** *If  $(F, +, \cdot)$  is a field, if  $(K, +, \cdot)$  is an algebraic extension field of  $F$ , and if for every polynomial  $f \in F[x]$  it is true that  $f$  is a splitting polynomial in  $K[x]$ , then  $(K, +, \cdot)$  is algebraically closed.*

*Proof.* For if  $f \in K[x]$  is a non-constant polynomial then there is a finite extension field  $(L, +, \cdot)$  of  $K$  such that  $f \in L[x]$  has a root  $\alpha \in L$ . But then by the tower law,  $L$  is a finite extension field over  $F$ . Thus  $\alpha$  is algebraic over  $F$  and so there is a polynomial  $g \in F[x]$  such that  $g(\alpha) = 0$ . But by hypothesis,  $g$  splits in  $K$  and so the roots of  $g$  lie in  $K$ . Thus,  $\alpha \in K$ . Hence,  $K$  is algebraically closed.  $\square$

**Theorem 16.5.31.** *If  $(K, +, \cdot)$  is an extension field of a field  $(F, +, \cdot)$  and if  $\mathbb{A}_F$  is the set:*

$$\mathbb{A}_F = \{ \alpha \in \Omega \mid \alpha \text{ is algebraic over } F \} \quad (16.5.47)$$

*Then  $(\mathbb{A}_F, +, \cdot)$  is a field.*

*Proof.* For if  $\alpha, \beta \in K$  are algebraic over  $F$ , then  $F[\alpha, \beta]$  is a field of finite degree over  $F$ . Thus every element of  $F[\alpha, \beta]$  is algebraic over  $F$ , and hence  $\alpha + \beta$ ,  $\alpha - \beta$ ,  $\alpha \cdot \beta$ , and  $\alpha/\beta$  are all algebraic. Therefore,  $\mathbb{A}_F$  is a field.  $\square$

**Definition 16.5.17** The algebraic closure of a field  $(F, +, \cdot)$  with respect to an extension field  $(K, +, \cdot)$  is the subfield  $(\mathbb{A}_F, +, \cdot)$  defined by:

$$\mathbb{A}_F = \{ \alpha \in \Omega \mid \alpha \text{ is algebraic over } F \} \quad (16.5.48)$$

By the previous theorem, the algebraic of a field  $(F, +, \cdot)$  with respect to any field extension  $(K, +, \cdot)$  is a subfield, and hence this is well defined.

**Theorem 16.5.32.** *If  $(F, +, \cdot)$  is a field, if  $(K, +, \cdot)$  is an algebraically closed field extension of  $F$ , and if  $\mathbb{A}_F$  is the algebraic closure of  $F$  with respect to  $K$ , then  $\mathbb{A}_F$  is an algebraic closure of  $F$ .*

*Proof.* Since  $\mathbb{A}_F$  is algebraic over  $F$ , and since every polynomial  $f \in F[x]$  splits in  $\mathbb{A}_F[x]$ , we thus have that, since  $K$  is algebraically closed, the  $f$  has a root in  $\mathbb{A}_F$ . Hence,  $\mathbb{A}_F$  is an algebraic closure of  $F$ .  $\square$

Combining this with the fundamental theorem of algebra we see that every subfield of  $\mathbb{C}$  has an algebraic closure.

**Example 16.5.12** Let  $\mathbb{Q}$  denote the standard field of rational numbers, and let  $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$  be the field extension obtained by appending  $\sqrt{2}$  and  $\sqrt{3}$ . The degree of  $[\mathbb{Q}[\sqrt{2}, \sqrt{3}] : \mathbb{Q}]$  is four, and to prove this we use the tower law. Firstly, note that  $[\mathbb{Q}[\sqrt{2}] : \mathbb{Q}] = 2$  for  $\{1, \sqrt{2}\}$  form a linear independent basis of  $\mathbb{Q}[\sqrt{2}]$

over  $\mathbb{Q}$ . Next,  $[\mathbb{Q}[\sqrt{2}, \sqrt{3}] : \mathbb{Q}[\sqrt{2}]] = 2$ . For  $\{1, \sqrt{3}\}$  certainly forms a basis, but moreover it is linearly independent. For suppose not and suppose we have:

$$a + b\sqrt{3} = 0 \quad a, b \in \mathbb{Q}[\sqrt{2}] \quad (16.5.49)$$

Then we have that  $\sqrt{3} = -a/b$  with  $a, b \in \mathbb{Q}[\sqrt{2}]$ . With this we may write:

$$\sqrt{3} = -\frac{a_1 + a_2\sqrt{2}}{b_1 + b_2\sqrt{2}} \quad (16.5.50)$$

Squaring both sides yields:

$$a_1^2 + 2a_1a_2\sqrt{2} + 2a_2^2 = 3b_1^2 + 6b_1b_2\sqrt{2} + 6b_2^2 \quad (16.5.51)$$

If we collect the  $\sqrt{2}$  terms, we obtain:

$$\sqrt{2}(2a_1a_2 - 6b_1b_2) = 3b_1^2 - a_1^2 + 6b_2^2 - 2a_2^2 \quad (16.5.52)$$

Now if  $2a_1a_2 - 6b_1b_2 \neq 0$ , then we may divide through by this showing that  $\sqrt{2}$  is a rational number, which is a contradiction. Thus,  $2a_1a_2 - 6b_1b_2 = 0$  and hence  $a_1a_2 = 3b_1b_2$ . But then:

$$a_1 + a_2\sqrt{2} + b_1\sqrt{3} + b_2\sqrt{6} = 0 \quad (16.5.53)$$

$$\Rightarrow a_1^2 + 3b_1b_2\sqrt{2} + a_1b_1\sqrt{3} + a_1b_2\sqrt{6} = 0 \quad (16.5.54)$$

$$\Rightarrow \quad (16.5.55)$$

**Definition 16.5.18** An  $F$  homomorphism from an extension field  $(E, +, \cdot)$  over a field  $(F, +, \cdot)$  to an extension field  $(E', +, \cdot)$  over  $F$  is a field homomorphism  $\phi : E \rightarrow E'$  such that  $\phi|_F = \text{id}_F$ .

**Theorem 16.5.33.** *If  $(F, +, \cdot)$  is a field, if  $(E, +, \cdot)$  is a simply field extension of  $F$ , if  $(K, +, \cdot)$  is a field extension of  $F$ , and if  $\varphi : E \rightarrow K$  is an  $F$  homomorphism, if  $\alpha \in E$  is such that  $E = F(\alpha)$ , if  $\alpha$  is transcendental over  $F$ , then  $\varphi(\alpha)$  is transcendental over  $F$  and the function  $f : \text{Hom}(E, K) \rightarrow E \setminus \mathbb{A}_F$  defined by  $f(\varphi) = \varphi(\alpha)$  is a bijection.*

**Example 16.5.13** Take the polynomial ring  $\mathbb{C}[x]$ . The prime ideals on this space are  $(z-a)$  for some  $a \in \mathbb{C}$ . The topology is then  $\{(x-a) \mid a \in \mathbb{C}\} \cup \{(0)\}$ . The weirdness about this is that  $(0)$  is dense in this space.

**Theorem 16.5.34.** *If  $(F, +_F, \cdot_F)$  is a field, if  $(K, +_K, \cdot_K)$  is a field extension, and if  $\alpha \in K$ , then  $F[\alpha]$  is a field if and only if  $\alpha$  is algebraic.*

**Theorem 16.5.35.**

**Definition 16.5.19: Irreducible Elements**

An irreducible element of a ring  $(R, +, \cdot)$  is an element  $x \in R$  such that for all  $a, b \in R$  such that  $a \cdot b = x$ , it is true that either  $a$  is a unit or  $b$  is a unit.

**Definition 16.5.20: Prime Element**

A prime element of a ring  $(R, +, \cdot)$  is an element  $p \in R$  such that for all  $a, b \in R$  such that  $p$  divides  $a \cdot b$ , then either  $p$  divides  $a$  or  $p$  divides  $b$ .

**Definition 16.5.21: Unique Factorization Domain**

An integral domain is a ring  $(R, +, \cdot)$  such that for all  $r \in R$  there exists finitely many irreducible elements  $a_i$  such that  $\prod a_i = r$  and such that for any other sequence of irreducible elements  $b_k$  such that  $\prod b_k = r$ , there exists units  $u_i$  such that  $a_i = u_i \cdot b_i$ .

**Example 16.5.14** The fundamental theorem of arithmetic states the  $\mathbb{Z}$  is a UFD.

**Theorem 16.5.36.** *If  $F$  is a field, then  $F[x]$  is a UFD.*

*Proof.* Any principal ideal domain is a unique factorization domain.  $\square$

**Example 16.5.15** There are non-unique factorization domains. For example,  $\mathbb{Z}[\sqrt{-5}]$  since  $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ .

The converse of the previous example is not true.

**Theorem 16.5.37.**  $\mathbb{Z}[x]$  is a UFD.

However,  $\mathbb{Z}[x]$  is not a PID since  $(2, x)$  is not principal.

### 16.5.1 Roots and Irreducibility

**Definition 16.5.22** An element  $\alpha \in F$  is a root of a polynomial  $f(x)$  if  $f(\alpha) = 0$ .

**Theorem 16.5.38.**  $\alpha \in F$  is a root of  $f \in F[x]$  if and only if  $(x - \alpha)$  divides  $f$ .

*Proof.* One was is clear, if  $(x - \alpha) | f$ , then  $f(x) = (x - \alpha)r(x)$  for some  $r \in F[x]$ . But then  $f(\alpha) = 0 \cdot r(\alpha) = 0$ , and thus  $\alpha$  is a root. In the other direction, use the division algorithm. Write  $f(x) = (x - \alpha)q(x) + r(x)$ . But  $r(x)$  must have degree less than  $x - \alpha$ , and is therefore a constant. But then  $f(\alpha) = r = 0$ , so  $r = 0$ . Hence,  $f(x) = (x - \alpha)q(x)$ , so  $x - \alpha$  divides  $f$ .  $\square$

**Theorem 16.5.39.** *If  $f \in F[x]$  is a polynomial of degree  $n$ , then there are at most  $n$  roots.*

*Proof.* By induction. If there are no roots, we are done. If not, write  $f(x) = (x - \alpha)q(x)$ . Then  $q$  is a polynomial of degree  $n - 1$ , and by the induction hypothesis has at most  $n - 1$  roots. Thus, there are at most  $n$  roots.  $\square$

**Example 16.5.16**  $x^2 - 1 \in \mathbb{Z}_8[x]$  has 4 roots. That is, 1, 3, 5, 7 are all roots. This does not contradict the previous theorem since  $\mathbb{Z}_8$  is not a field.

### Theorem 16.5.40: Fundamental Theorem of Algebra

If  $f : \mathbb{C} \rightarrow \mathbb{C}$  is a non-constant polynomial, then there exists an  $\alpha \in \mathbb{C}$  such that  $f(\alpha) = 0$ .

**Theorem 16.5.41.** *Any linear polynomial  $ax + b \in F[x]$  is irreducible.*

**Theorem 16.5.42.** *If  $f \in F[x]$  is irreducible and  $\deg(f) \geq 2$ , then  $f$  has no roots. Moreover, the converse holds if  $\deg(f) = 2$  or 3.*

**Example 16.5.17** The previous theorem is very special for degree 2 and 3. Let  $f(x) = x^2 + x + 1$  where  $f \in \mathbb{F}_p[x]$ ,  $p$  a prime. For the case of  $p = 2$  have seen that this is a irreducible. For  $p = 3$  we have that 1 is a root:  $1^2 + 1 + 1 = 3 \cong 0$  in  $\mathbb{F}_3$ . For  $p = 5$ , this is once again irreducible, but not in  $\mathbb{F}_7$ .

**Example 16.5.18** Consider  $x^4 + x^2 + 1 \in \mathbb{F}_2[x]$ . This has no roots, since 0 and 1 both map to 1, but it is reducible since  $x^4 + x^2 + 1 = (x^2 + x + 1)^2$  in  $\mathbb{F}_2[x]$ . That is, the so called *freshman's dream* is true in this case, and we can distribute the power. Thus we have a reducible polynomial with no roots. Note, however, that the degree is not 2 or 3.

Next, we present Gauss's Lemma. Note that being irreducible over  $\mathbb{Z}$  is stronger than being irreducible over  $\mathbb{Z}$ . Consider  $f(x) = 2x$ . While this is linear in  $\mathbb{Q}$ , and since  $\mathbb{Q}$  is a field,  $f$  is irreducible in  $\mathbb{Q}[x]$ . However when viewed in  $\mathbb{Z}[x]$ , we have  $2x = 2 \cdot x$ , neither of which are units, and hence  $f$  is reducible in  $\mathbb{Z}[x]$ .

**Theorem 16.5.43.** *If  $f \in \mathbb{Z}[x]$  is irreducible, then it is irreducible in  $\mathbb{Q}[x]$ .*

### 16.5.2 Reduction Modulo a Prime

Consider the projection map  $\pi : \mathbb{Z}[x] \rightarrow \mathbf{F}_p[x]$ , reduction mod  $p$  of the coefficients:

$$f(x) = \sum a_k x^k \longrightarrow \bar{f}(x) \sum \bar{a}_k x^k \quad (16.5.56)$$

The mapping  $\pi$  is a ring homomorphism.

**Example 16.5.19**  $x^3 + 21x + 31 \mapsto x^3 + x + 1$  in  $\mathbb{F}_2[x]$ .

**Theorem 16.5.44.** *If  $f \in \mathbb{Z}[x]$ , if  $p$  is prime, and if  $p \nmid a_n$ , and if  $\pi(f) \in \mathbb{F}_p[x]$  is irreducible, then  $f \in \mathbb{Z}[x]$  is irreducible.*

*Proof.* We prove by the contrapositive. If  $f(x) = g(x)h(x)$ , then  $\bar{f}(x) = \bar{g}(x)\bar{h}(x)$  since  $\pi$  is a homomorphism. But  $\bar{g}$  and  $\bar{h}$  are not units since  $p$  does not divide  $a_n$ , and thus the degree of the reduction is equal to the degree of the original polynomial. That is,  $\deg(f) = \deg(\bar{f})$ . But also  $\deg(g) = \deg(\bar{g})$  and similarly for  $h$  since:

$$\deg(\bar{g}) + \deg(\bar{h}) = \deg(g) + \deg(h) = \deg(f) \quad (16.5.57)$$

Since  $p$  does not divide the leading coefficients of either  $g$  or  $h$ , the degrees of  $\bar{g}$  and  $\bar{h}$  remain the same, and hence are not units. Thus,  $\bar{f}$  is reducible.  $\square$

**Example 16.5.20** Consider  $x^3 + 21x + 31 \in \mathbb{Z}[x]$ . In  $\mathbb{F}_3[x]$  we have  $x^3 + x + 1$ , which is indeed irreducible, and hence the original polynomial  $x^3 + 21x + 31$  is irreducible.

**Example 16.5.21** The converse is not true. There are polynomials that are reducible for all  $p$ , yet in  $\mathbb{Z}[x]$  it is irreducible. Consider  $(2x + 1)(x + 1) = 2x^2 + 3x + 1$ . This is reducible, since we have the factorization, however in  $\mathbb{F}_2[x]$  this is simply  $x + 1$ , which is irreducible. However, 2 divides 2 so the theorem does not apply.

**Example 16.5.22** In  $\mathbb{Z}[x]$ , consider  $x^2 + x + 1$  which we know is irreducible since it is irreducible in  $\mathbb{R}[x]$  (the roots are complex). In  $\mathbb{F}_2[x]$  we can check and see that there are no roots, and thus  $x^2 + x + 1$  is irreducible by Gauss' lemma. However, in  $\mathbb{F}_3[x]$  it is reducible since  $x^2 + x + 1 = (x - 1)^2$ . That is, 1 is a root of  $x^2 + x + 1$  in  $\mathbb{F}_2[x]$  with multiplicity 2.

#### Theorem 16.5.45: Eisenstein's Criterion

If  $p$  is prime, if  $f \in \mathbb{Z}[x]$ , if  $p$  does not divide  $a - n$ , if  $p$  divides  $a_k$  for all  $k < n$ , and if  $p^2$  does not divide  $a_0$ , then  $f$  is irreducible.

*Proof.* We prove by the contradiction. Suppose  $f$  is reducible. If  $p$  divides all of the  $a_i$ , then in the reduction map  $\pi : \mathbb{Z}[x] \rightarrow \mathbb{F}_p[x]$ , all of the  $a_i$  map to zero. So we have:

$$\bar{a}_n x^n = \bar{f}(x) = \bar{g}(x) \bar{h}(x) \quad (16.5.58)$$

But since  $\mathbb{Z}[x]$  is a UFD, both  $\bar{g}$  and  $\bar{h}$  are monomials,  $\bar{g}(x) = g_0 x^i$  and  $\bar{h}(x) = h_0 x^j$ . If  $i, j < n$ , then  $i, j > 0$ , and so both constant terms must be divisible by  $p$ . Thus  $g(x)h(x)$  has a constant term divisible by  $p^2$ , a contradiction.  $\square$

**Example 16.5.23** Let  $f(x) = x^4 + 22x^2 + 33x + 44$ . By Eisenstein, with  $p = 11$ , we have that this is irreducible.

### 16.5.3 Review of Previous Lecture

If  $\mathbf{F}$  is a field, then  $\zeta \in \mathbf{F}$  is called a root of unity if there is some  $n \geq 1$  such that  $\zeta^n = 1$ .

**Example 16.5.24** In the field  $\mathbb{R}$ , the roots of unity of 1 and -1. The number 1 is a first root of unity, whereas -1 is a second root of unity.

**Example 16.5.25** In the complex numbers  $\mathbb{C}$ , there is an  $n^{th}$  root of unity for all  $n \in \mathbb{N}^+$ . Let  $\zeta = \exp(2\pi i/n)$ . The roots of unity are scattered along the unit circle.

A root of unity is some element  $\zeta \in \mathbb{F}$  such that  $\zeta$  is a root of the polynomial  $f(x) = x^n - 1$ .

**Example 16.5.26** Let  $p$  be a prime integer, and consider the roots of  $f(x) = x^p - 1$ . There is always a root since 1 satisfies this criterion, and hence we can factor this and obtain:

$$f(x) = x^p - 1 = (x - 1)(x^{p-1} + x^{p-2} + \cdots + x + 1) \quad (16.5.59)$$

This latter polynomial  $x^{p-1} + x^{p-2} + \cdots + x + 1$  is irreducible over  $\mathbb{Q}$ .

**Theorem 16.5.46.** *If  $p$  is prime and  $f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$ ,  $f \in \mathbb{R}[x]$ , then  $f$  is irreducible over  $\mathbb{Q}$ .*

*Proof.* For:

$$xf(x+1) = (x+1)^p - 1 \quad (16.5.60)$$

by the binomial theorem we have:

$$xf(x+1) = (x+1)^p - 1 = -1 + \sum_{k=0}^p \binom{p}{k} x^k = \sum_{k=1}^p \binom{p}{k} x^k \quad (16.5.61)$$

Thus, simplifying, we have:

$$f(x+1) = \sum_{k=1}^p \binom{p}{k} x^{k-1} \quad (16.5.62)$$

Thus by the Eisenstein criterion,  $f(x+1)$  is irreducible over  $\mathbb{Q}$ . But if  $f(x+1)$  is irreducible over  $\mathbb{Q}$ , then  $f(x)$  is as well.  $\square$

The converse of the statement before is not true. If the translation of a polynomial is reducible, it need not mean the original polynomial was reducible. For let  $f(x) \in \mathbb{F}_2[x]$  be defined by  $f(x) = x^2 + x + 1$ . Plugging in  $x^2$  we get  $f(x)^2 = x^4 + x^2 + 1 = (x^2 + x + 1)^2$ , which is reducible.

## 16.6 Gauss's Lemma

**Definition 16.6.1** A polynomial  $f \in \mathbb{Z}[x]$  is called primitive if  $\text{GCD}(a_0, \dots, a_n) = 1$ . Equivalently, for all primes  $p$  there is an  $i$  such that  $p \nmid a_i$ . That is,  $p$  does not divide  $a_i$ .

**Example 16.6.1** Let  $f(x) = 10x^3 + 5x^2 + 2x + 23$ . This is a primitive polynomial. If  $p$  divides all of the  $a_i$ , it must divide 2, but 2 is the only prime that divides 2, and hence  $p = 2$ . But 5 and 23 are odd, and hence 2 does not divide them. So,  $f$  is primitive.

### Theorem 16.6.1: Gauss's Lemma

For any  $f \in \mathbb{Q}[x]$  there is a unique  $c \in \mathbb{Q}$  and a unique primitive polynomial  $g \in \mathbb{Z}[x]$  such that  $f(x) = c \cdot g(x)$ .  $\blacksquare$

*Proof.* Existence is straight forward. Clear out the denominators of  $f$ , let  $c$  be the common factor, and we're done. Suppose  $cg(x) = c'g'(x)$ . If  $c$  and  $c'$  are not integers, multiply by their denominators and thus we may assume  $c$  and  $c'$  are integers. Then we have:

$$cg(x) = c'(b_0 + b_1x + \dots + b_nx^n) \quad (16.6.1)$$

And hence:

$$c = \text{GCD}(c'b_0, \dots, c'b_n) = c\text{GCD}((b_0, \dots, b_n)) \quad (16.6.2)$$

But  $g'$  is trivial, so the greatest common denominator is 1. Hence,  $c = c'$ . Also,  $g = g'$ .  $\blacksquare$

### Theorem 16.6.2: Gauss' Lemma Version 2

If  $g \in \mathbb{Z}[x]$  is primitive and  $f \in \mathbb{Z}[x]$ , and if  $g|f$  in  $\mathbb{Q}[x]$ , then  $g|f$  in  $\mathbb{Z}[x]$ .  $\blacksquare$

*Proof.* For if  $g|f$ , then  $f = gh$  with  $h \in \mathbb{Q}[x]$ . But by Gauss' lemma there is a unique  $c \in \mathbb{Q}$  such that  $h(x) = ch_0(x)$ , where  $h_0 \in \mathbb{Z}[x]$  is primitive. So  $f(x) = cg(x)h_0(x)$ . But then  $g(x)h_0(x)$  is primitive, and hence  $c$  is an integer. Thus,  $h \in \mathbb{Z}[x]$ . That is, since the product of primitive is primitive, the numerator of  $c$  is equal to the denominator of  $c$  times the GCD of the coefficients of  $f$ . But this GCD is a positive integer, and hence the numerator divides it, and hence  $c$  is an integer itself.  $\square$

### Theorem 16.6.3: Gauss' Lemma V3

If  $f \in \mathbb{Z}[x]$  and if  $g, h \in \mathbb{Q}[x]$  are such that  $f = gh$ , then  $f = g_0h_0$ . ■

## 16.7 Field Extensions

### Theorem 16.7.1: Tower Law

If  $\mathbb{F}$ ,  $\mathbb{K}$ , and  $\mathbb{L}$  are fields, if  $\mathbb{K}$  is a field extension of  $\mathbb{F}$ , if  $\mathbb{L}$  is a field extension over  $\mathbb{K}$ , then  $\mathbb{L}/\mathbb{F}$  is a finite dimensional vector space if and only if  $\mathbb{L}/\mathbb{K}$  and  $\mathbb{K}/\mathbb{F}$  is finite. Moreover:

$$[\mathbb{L} : \mathbb{F}] = [\mathbb{L} : \mathbb{K}][\mathbb{K} : \mathbb{F}]$$

Suppose  $K/F$  is a field extension. That is,  $F \subseteq K$  and  $F$  is a subfield of  $K$ . Let  $X \subseteq K$  be any non-empty subset. We now define the adjointment of  $X$  to  $F$ .

**Definition 16.7.1** The field extension generated by a subfield  $F$  of a field  $K$  by a subset  $X \subseteq K$  is the subfield:

$$F(X) = \bigcap_{\substack{F \subseteq L \subseteq K \\ X \subseteq L}} L \tag{16.7.1}$$

Where  $L$  is a subfield.

$F(X)$  is non-empty since  $K$  is in the intersection, and the intersection of subfields is again a subfield, so  $F(X)$  is a field. If  $X = \{a_0, a_1, \dots\}$ , we often write  $F(X) = F(a_0, a_1, \dots)$  and if  $X$  is finite we call  $F(X)/F$  finitely generated. If  $X$  is a single element,  $F(\alpha)$  is called a simple extension. A field extension is simple if it can be written as a simple extension.

**Example 16.7.1**  $\mathbb{R}/\mathbb{Q}$  is not finitely generated. A simple cardinality argument works here, for if it were countably generated then since  $\mathbb{Q}$  is countable,  $\mathbb{R}$  would again be countable, which is false.

**Example 16.7.2**  $\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}, \dots)$  is not finitely generated. We can show this by building a chain of subfields. First consider  $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ . This is a field extension of degree two. One can see this since  $\sqrt{2}$  is not rational. The next field extension  $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}(\sqrt{2})$  is also an extension of degree 2. For suppose not and suppose:

$$\sqrt{3} = a + b\sqrt{2} \quad (16.7.2)$$

Squaring, we obtain:

$$3 = a^2 + 2b^2 + 2ab\sqrt{2} \quad (16.7.3)$$

Now since  $\sqrt{2}$  is irrational we must conclude that  $ab = 0$ . Thus either  $a = 0$  or  $b = 0$ . If  $b = 0$ , then  $\sqrt{3}$  is an integer, which is false. If  $a = 0$  then  $3 = 2b^2$ , but 3 is prime, a contradiction. Thus  $\sqrt{3}$  is not contained in  $\mathbb{Q}(\sqrt{3})$ .

**Theorem 16.7.2.** *If  $F$  is a field, if  $K$  is a field extension, if  $\alpha \in K$ , then:*

$$F(\alpha) = \left\{ \frac{f(\alpha)}{g(\alpha)} \mid f, g \in F[x], g(\alpha) \neq 0 \right\} \quad (16.7.4)$$

**Theorem 16.7.3.** *A field extension of a field  $F$  is a field extension  $K$  of  $F$  is an injective homomorphism  $\iota : F \rightarrow K$ .*

**Definition 16.7.2** A morphism of field extensions  $K/F$  and  $K'/F$  is a homomorphism  $\varphi : K \rightarrow K'$  such that the injective homomorphisms  $\iota$  and  $\iota'$  commute with  $\varphi$ . That is,  $\iota' = \varphi \circ \iota$ .

**Example 16.7.3**  $\mathbb{C}$  and  $\mathbb{R}[x]/(x^2 + 1)$  are isomorphic field extensions of  $\mathbb{R}$ .  $\varphi : \mathbb{R}[x] \rightarrow \mathbb{R}$  mapping  $f \mapsto f(i)$  is surjective since  $a + ib = a(i) + b(i)^4$ . Thus this is a surjective  $\mathbb{R}$  algebra homomorphism. The kernel is the ideal generated by  $x^2 + 1$ . Then by the first isomorphism theorem,  $\tilde{\varphi} : \mathbb{R}[x]/(x^2 + 1) \rightarrow \mathbb{C}$  is an isomorphism. Hence, these are isomorphic.

## 16.8 Minimal Polynomial

If  $F \subseteq K$  is a subfield, if  $\alpha \in K$ , then there exists an  $F$  algebra homomorphism  $\varphi_\alpha : F[x] \rightarrow K$ . Then  $\text{Ker}(\varphi_\alpha) \subseteq F[x]$ . If  $\varphi_\alpha$  is injective, then the kernel is 0, and this is true if and only if  $f(\alpha) \neq 0$  for all  $f \in F[x]$ . That is,  $\alpha$  is *transcendental* over  $F$ . On the other hand, if  $\varphi_\alpha$  is not injective then  $\text{Ker}(\varphi_\alpha) = (m_\alpha/F(x))$  where  $m_\alpha/F(x)$  is a monic polynomial in  $F[x]$ , and this is called the minimal polynomial of  $\alpha$  over  $F$ . In this case we say that  $\alpha$  is *algebraic* over  $F$ .

**Theorem 16.8.1.** *If  $K/F$  is a field extension,  $\alpha \in K$  is algebraic over  $F$ , if  $f \in F[x]$  with  $f(\alpha) = 0$ , then  $m_{\alpha/F}(x) | f$ , hence  $m_{\alpha/F}(x)$  is the unique monic polynomial over  $F$  of minimal degree with  $\alpha$  as a root. Moreover,  $m_{\alpha/F}(x)$  is irreducible over  $F$  and is the unique monic irreducible polynomial over  $F$  with  $\alpha$  as a root.*

**Example 16.8.1** Let  $d \in \mathbb{Z}$  be a non-square. Then the minimal polynomial of  $d$  is  $m_{\sqrt{d}}/\mathbb{Q}(x) = x^2 - d$ .

**Example 16.8.2** Let  $\alpha = \sqrt{2} + \sqrt{3}$ . What is the minimal polynomial of  $\alpha$  over  $\mathbb{Q}$ ? Well we first try to find a polynomial that has  $\alpha$  as a root. Squaring, we have:

$$\alpha^2 = 2 + 2\sqrt{6} + 3 = 5 + 2\sqrt{6} \quad (16.8.1)$$

Bringing the 5 over and squaring:

$$(\alpha^2 - 5)^2 = 24 = \alpha^4 - 10\alpha^2 + 25 \quad (16.8.2)$$

Thus letting  $f(x) = x^4 - 10x + 1$ , we obtain a polynomial with  $\alpha$  as a root.

**Theorem 16.8.2.** *If  $\alpha \in K$  is algebraic over  $F$ , then  $F[x]/(m_\alpha/F(x))$  is isomorphic to  $F[\alpha]$ , and  $F[\alpha] = F(\alpha)$ .*

Since  $F(\alpha) \subseteq F[\alpha]$  because the first part,  $F[x]$  is a field since  $m_\alpha/F(x)$  is irreducible, so  $F[x]/(m_\alpha/F(x))$  is a field. Hence  $F(\alpha) \subseteq F[\alpha]$  since  $F(\alpha)$  is minimal field containing  $\alpha$ . Given  $\alpha$  algebraic over  $F$ ,  $\alpha^{-1}$  is a polynomial in  $\alpha$ . There exists an algorithm demonstrating this using Bezout's identity.

**Definition 16.8.1** A field extension  $K/F$  is algebraic if for all  $\alpha \in K$ ,  $\alpha$  is algebraic over  $F$ .

**Theorem 16.8.3.**  *$K/F$  finite if and only if  $K/F$  is algebraic and finitely generated.*

## 16.9 Review of Previous Lectures

A field extension  $K/F$  is algebraic if every element  $\alpha \in K$  is algebraic over  $F$ .

**Theorem 16.9.1.** *A field extension  $K/F$  is finite if and only if  $K/F$  is algebraic and finitely generated.*

*Proof.* For if  $K/F$  is finite, then  $K$  is a finite dimensional vector space over  $F$ . Thus, if  $\alpha \in K$ , then the set  $\{1, \alpha, \alpha^2, \dots\}$  is linearly dependent. That is, there exists  $a_0, \dots, a_n \in F$  such that  $a_0 + \dots + a_n \alpha^n = 0$ . Let  $\alpha_1, \dots, \alpha_m$  be an  $F$  basis for  $K$ . Then  $K = F(\alpha_1, \dots, \alpha_n)$ , and so  $K$  is finitely generated. The converse is trickier. Since  $K$  is finitely generated,  $K = F(\alpha_1, \dots, \alpha_n)$ . But since  $K$  is algebraic over  $F$ ,  $\alpha_i$  is algebraic in  $F$ . Thus we build a tower:

$$F \rightarrow F(\alpha_1) \rightarrow F(\alpha_1, \alpha_2) \rightarrow \dots \rightarrow F(\alpha_1, \dots, \alpha_n) = K \quad (16.9.1)$$

By the generalized tower law,  $K$  is finite over  $F$ .  $\square$

$K/F$  is transcendental (that is, not algebraic) implies that  $K/F$  is of infinite degree. Given  $\alpha \in K$  transcendental over  $F$ , then  $\varphi_\alpha : F[x] \rightarrow F[\alpha] \subseteq F(\alpha)$ .

**Theorem 16.9.2.** *If  $K/L$  is algebraic and  $L/F$  is algebraic, then  $K/F$  is algebraic.*

*Proof.* For let  $\alpha \in K$ . We want to show that  $\alpha$  is algebraic over  $F$ . This is equivalent to the claim that  $F(\alpha)$  is finite over  $F$ . But  $\alpha$  is algebraic over  $L$ , and hence there is a minimal polynomial  $m_\alpha/L(x) = a_0 + \dots + a_{n-1}x^{n-1} + x^n \in L[x]$ . Since  $L/F$  is algebraic,  $a_i$  is algebraic over  $F$ .  $\square$

The converse is true as well.

## 16.10 Compass and Straight Edge

Consider some subset  $S \subseteq \mathbb{C}$ , equipped with some rules:

- Given two points  $P, Q$  there is a line through  $P$  and  $Q$ .
- Given  $P, Q$ , there is a circle  $C(P, |Q - P|)$  centering at  $P$  with radius  $|P - Q|$ .

Any point that is the intersection of any of the lines and circles is said to be constructible by compass and straightedge.

**Example 16.10.1** Bisecting a line is possible. Bisecting an angle is possible. Can you trisect an angle? What about double a cube?

We'll define inductively some sets  $P_n$ ,  $L_n$ , and  $C_n$ .  $P_0 = \{0, 1\} \subseteq \mathbb{C}$ ,  $L_0, C_0 = \emptyset$ . If  $L_n$  has been constructed, let  $L_{n+1}$  be the set of all lines through all points in  $P_n$ . If  $C_n$  has been constructed, let  $C_{n+1}$  be the set of all circles about all points in  $P_n$  with radii all of the distances  $|z_1 - z_2|$  for points  $z_1, z_2 \in P_n$ . Then  $P_n$  is finite for all  $n \in \mathbb{N}$  and thus the union over all  $P_n$  is at most countable. This cardinality argument shows that there are inconstructible numbers. Moreover,  $P = \bigcup P_n$  is a subfield of  $\mathbb{C}$ . To see this we must show that  $P$  is closed to addition, multiplication, and inverses.

**Theorem 16.10.1.** *The set of constructible numbers is a subfield of  $\mathbb{C}$ .*

*Proof.* It suffices to show that  $P \cap \mathbb{R}$  is a subfield. For let  $a, b \in P \cap \mathbb{R}$ . We need to show that  $a + b$ ,  $a \cdot b$ , and  $a/b$  are constructible. Given the length  $a$  and the length  $b$ , we can translate the length  $b$  to start at  $a$ , given us the point  $a + b$ , which will have length  $a + b$ . For  $ab$  we construct two triangles, one with length 1 and the other with length  $a$ . We build a triangle on the first with a length  $b$ , and a similar triangle on the second length which will have length  $ab$  by similarity. Lastly, do the same triangle with  $b$  on the inside to get  $a/b$ . Draw some pictures later.  $\square$

**Theorem 16.10.2.**  $\mathbb{Q} \subseteq P$ .

*Proof.* Since  $1 \in P$ , every integer multiple of 1 is also contained in  $P$  since we can add 1 to itself  $n$  times. Thus  $n/m$  is contained in  $P$ , so  $\mathbb{Q} \subseteq P$ .  $\square$

**Theorem 16.10.3.**  $P$  is closed to square roots.

*Proof.* Draw a circle of diameter  $a + 1$ . Do that fancy circle.  $\square$

Let  $Q^{py}$  be the intersection of all subfields  $K \subseteq \mathbb{C}$  such that for all  $z \in K$  it is true that  $\sqrt{z} \in K$ . This is the smallest subfield of  $\mathbb{C}$  in which one can always take square roots. It's called the Pythagorean closure of  $\mathbb{Q}$ . From the previous theorem,  $Q^{py} \subseteq P$  is a subfield.

**Theorem 16.10.4.**  $Q^{py} = P$ .

*Proof.* For let  $z \in P$ . It suffices to show that  $P_n \subseteq Q^n$  for all  $n$ , since the  $\bigcup P_n = P \subseteq Q^{py}$ . We prove this by induction. The base case is true since  $P_0 = \{0, 1\}$  and this is a subset of  $Q^{py}$ . Suppose  $P_n \subseteq Q^{py}$  and recall that  $P_{n+1}$  is defined as all  $z \in \mathbb{C}$  such that  $z \in L \cap L'$ , or  $z \in L \cap C$ , or  $z \in C \cap C'$ , where  $L$ ,  $L'$ ,  $C$ , and  $C'$  are lines and circles through points in  $P_n$ . Case 1,  $z \in L \cap L'$  Then there are four points  $z_1, z_2, z_3, z_4$  such that:

$$z = z_1\alpha + (1 - \alpha)z_2 = z_3\beta + (1 - \beta)z_4 \quad (16.10.1)$$

We can solve this and note that  $P_n$  is closed under complex conjugation. Thus, in case 1 we have that  $z$  is contained in  $Q^{py}$ . Here we have:

$$z = z_1\alpha + (1 - \alpha)z_2 = z_3 + r \exp(i\theta) \quad (16.10.2)$$

This be true as well. The final case is two circles:

$$z = z_1 + r_1 \exp(i\theta_1) = z_2 + r_2 \exp(i\theta_2) \quad (16.10.3)$$

Which also be like it is.  $\square$

**Theorem 16.10.5.**  $Q^{py}$  is algebraic over  $\mathbb{Q}$ .

We can build the Pythagorean closure from  $\mathbb{Q}$  by considering  $\sqrt{\mathbb{Q}}$ , all numbers of the for  $a + b\sqrt{c}$  with  $a, b, c \in \mathbb{Q}$ , and then  $\sqrt{\sqrt{\mathbb{Q}}}$ , and so on.

**Theorem 16.10.6.**  $z \in Q^{py}$  if and only if there is a tower of extensions  $K_0, \dots, K_n$  such that  $\mathbb{Q} = K_0$  and  $K_n = \mathbb{Q}[z]$  where  $K_{j+1}$  has degree 2 over  $K_j$ .

**Theorem 16.10.7.** If  $\alpha \in Q^{py}$ , then  $[\mathbb{Q}[\alpha] : \mathbb{Q}] = 2^n$  for some  $n \in \mathbb{N}$ .

*Proof.* Apply the tower law to the previous theorem.  $\square$

From this we can do all of the impossibility proofs of various constructions.

**Example 16.10.2** It is impossible to double the volume of a cube with lengths 1. The minimal polynomial of  $\sqrt[3]{2}$  is  $x^3 - 2$  by Eisenstein. Thus the degree of the field extension  $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$  which is not a power of 2. It is also impossible to trisect angles. For let  $\theta = 2\pi/3$ . Trisecting this is equivalent to constructing  $2\pi/9$ . The minimal polynomial of this is  $x^3 - \frac{3}{4}x + \frac{1}{8}$ , and thus the degree of the extension is again 3, which is not a power of 2.

## 16.11 Review

If  $f \in F[x]$  is irreducible, then  $f$  has no roots. The converse is false. If  $f \in F[x]$  has no roots, it may still be reducible. This holds for any field if  $f$  has degree 2 or 3.

**Example 16.11.1** If  $f(x) = x^4 + 1$ ,  $f \in \mathbb{R}[x]$ , then  $f$  is reducible. For we have:

$$x^4 + 1 = (x^2 - \sqrt{-x}x + 1)(x^2 + \sqrt{2}x + 1) \quad (16.11.1)$$

We can now complete this since  $x^2 - \sqrt{2}x + 1$  and  $x^2 + \sqrt{2}x + 1$  have no real roots by the quadratic formula, and hence these are irreducible. Thus,  $x^4 + 1$  is reducible and factors as above.

**Example 16.11.2** There is only one irreducible quadratic polynomial over  $\mathbb{F}_2[x]$  and that is  $x^2 + x + 1$ . Using this, is the polynomial  $x^4 + x^2 + x + 1$  irreducible in  $\mathbb{F}_2[x]$ ? Since this has no roots, we know that it has no linear factors, and thus if it is reducible it must be the product of irreducible quadratics. But there is only one irreducible quadratic, so we can use this to check. We have:

$$(x^2 + x + 1)^2 = x^4 + 2x^3 + x^2 + 2x + 1 = x^4 + x^2 + 1 \quad (16.11.2)$$

Since in  $\mathbb{F}_2$  we have that  $2 = 0$ . But this is not equal to the original polynomial  $x^4 + x^3 + x^2 + x + 1$ , and hence this is indeed irreducible.

If  $\alpha \in \mathbb{C}$  is constructible, and if  $K_i$  is a tower of field extensions each of degree 2 over the previous one, then  $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 2^n$ .

**Theorem 16.11.1.** *A regular  $n$  gon is constructible if and only if the complex number  $\exp(2\pi i/n)$  is constructible.*

If  $p$  is a prime number, then the minimal polynomial of  $\exp(i2\pi i/n)$  is just  $x^{p-1} + \dots + 1$ . Thus  $[\mathbb{Q}(\zeta_p) : \mathbb{Q}] = p - 1$ . Thus we need  $p - 1$  to be a power of 2.

**Theorem 16.11.2.** *If  $p - 1$  is not a power of 2 then we can not construct a regular  $p$  gone.*

**Theorem 16.11.3.** *If  $2^k + 1$  is a prime, then  $k = 2^n$  for some  $n$ .*

There are only 5 Fermat primes,  $p = 3, 5, 17, 257, 65537$ . An unsolved problem (as of 2020) is whether or not there are more such primes, or are there infinitely many primes?

## 16.12 Splitting Fields

If  $f \in F[x]$ , we have given a construction of  $K/F$  in which  $f$  has a root  $\alpha \in K$ . That is,  $(x - \alpha)$  divides  $f$  in  $K[x]$ . For if  $f$  splits completely (factors into a product of linear polynomials over  $F$ ), then there is nothing to prove. If not, let  $g(x)$  be an irreducible factor of  $f$  and define  $K = F[x]/(g)$ , where  $(g)$  is the ideal generated by  $g$  in  $F[x]$ . Then  $g(x)$  has a root  $\bar{x}$ ,  $K = F(\bar{x})$ , so  $(x - \bar{x})$  divides  $g$ , which divides  $f$ , so we're done.

**Example 16.12.1** Let  $f(x) = x^3 - 2$ ,  $f \in \mathbb{Q}[x]$ . This has three roots,  $\sqrt[3]{2}$ ,  $\sqrt[3]{2}\omega$ , and  $\sqrt[3]{2}\bar{\omega}$ , where  $\omega$  is a complex cubed root of unity, and  $\bar{\omega}$  is its complex conjugate. Thus  $\mathbb{Q}[x]/(x^3 - 2)$  is isomorphic to  $\mathbb{Q}(\sqrt[3]{2})$ .

**Definition 16.12.1** An extension  $K/F$  is a splitting field of  $f \in F[x]$  if  $f$  splits completely in  $K[x]$  and does not split in any proper subfield.

**Theorem 16.12.1.** *Splitting fields exist.*

*Proof.* Proof by induction on the degree of  $n$ . If it is true for  $n$ , write  $f = (x - \alpha)g$  for some  $\alpha$  in  $F[x]/(h)$ . Then  $g$  is of degree  $n$  or less, and hence has a splitting field  $K$ . Adjoining  $\alpha$  to  $K$ . Let  $E$  be the intersection of all subfields of  $E$  containing this  $K$  adjoing  $\alpha$ .  $\square$

### 16.12.1 Review from Previous Lecture

If  $F$  is a field, and if  $f \in F[x]$ , then an extension field  $K$  over  $F$  is a splitting field if  $f \in K[x]$  is a splitting polynomial. That is, there are linear factors  $a_kx + b_k$ , with  $a_k, b_k \in K$ , such that:

$$f(x) = \prod_{k \in \mathbb{Z}_n} (a_kx + b_k) \tag{16.12.1}$$

Splitting fields exists since given  $f$  we can find an irreducible factor  $g$ , and thus  $F[x]/(g)$  will be a field extension of  $F$ . Letting  $\alpha_1 = \bar{x}$ , the equivalence class of  $x$ , this will be a root of  $g$  in  $F[x]/(g)$ , and hence a root of  $f$  in  $G$ . The extension field will have degree less than or equal to  $n$ . Continuing on inductively, obtaining  $\alpha_k$ , we find that  $f \in F[\alpha_1, \dots, \alpha_n]$  splits completely and that this field extension has at most degree  $n!$ . If  $f$  was irreducible to begin with then  $f$  is a minimal polynomial up to scalar multiplication, and hence the

degree of  $F[\alpha_1]$  over  $F$  will be  $n$ . From the tower law,  $n$  will divide the degree of  $F[\alpha_1, \dots, \alpha_n]$ .

**Example 16.12.2** Let  $f(x) = x^4 + 4$ . This factors as:

$$x^4 + 4(x^2 - 2x + 2)(x^2 + 2x + 2) \quad (16.12.2)$$

We can invoke the quadratic formula to find the roots of these two over  $\mathbb{C}$ :

$$\alpha_1 = \frac{2 \pm \sqrt{-4}}{2} = 1 \pm i \quad (16.12.3a) \quad \alpha_2 = \frac{-2 \pm \sqrt{-4}}{2} = -1 \pm i \quad (16.12.3b)$$

So the splitting field is  $K = \mathbb{Q}(\pm 1 \pm i)$ , and this is just  $\mathbb{Q}(i)$ . Thus,  $[K : \mathbb{Q}] = 2$ .

**Theorem 16.12.2.** *If  $F, F'$  are fields, if  $\varphi : F \rightarrow F'$  is a field isomorphism, and if  $f \in F[x]$ , if  $f'(x)(\varphi \circ f)(x) \in F'[x]$ , if  $K$  is a splitting field for  $f$ , and if  $K'$  is a splitting field for  $f'$ , then there is an isomorphism  $\sigma : K \rightarrow K'$  such that for all  $c \in F$ ,  $\sigma(c) = \varphi(c)$ .*

*Proof.* We prove by induction on the degree of  $f$ . Since  $\varphi$  is an isomorphism,  $f'$  and  $f$  have the same degree. Thus if  $f \in F[x]$  and  $f' \in F'[x]$  are degree one polynomials, then they already split and hence if we let  $\sigma = \varphi$ , then we have shown that the splitting fields are isomorphic. Assume the claim is true for  $n \in \mathbb{N}$ . We can assume that  $f$  has an irreducible factor  $g$  of degree greater than 1. Then  $g$  has a root in  $K$ , label it  $\alpha$ , and let  $g' = \varphi \circ g$ . Then  $g'|f'$  and  $g'$  is irreducible over  $F'$ . So let  $\alpha'$  be a root of  $g'$  in  $K'$ . Then  $F[x]/(g) \simeq F[\alpha] \simeq F'[x]/(g')$  and thus we can extend this isomorphism one step up. So now we have:

$$f(x) = (x - \alpha)f_1(x) \quad (16.12.4a) \quad f'(x) = (x - \alpha')f'_1(x) \quad (16.12.4b)$$

$f_1$  splits in  $K$  and thus we need to show that  $K$  is minimal. Suppose  $K/L/F_1$ . Then since  $f_1$  splits in  $L$ ,  $F$  splits in  $L$  as well, contradicting the minimality of  $K$  over  $F$ , and hence  $K$  is minimal over  $F_1$ . Thus,  $K$  is a splitting field for  $f_1$  over  $F_1$ . Thus by induction there is an isomorphism  $\sigma : K \rightarrow K'$  with the desired property over  $F_1$  and  $F'_1$ . But  $F_1$  and  $F'_1$  are extensions over  $F$  and  $F'$ , so we're done.  $\square$

## 16.13 Separability

Let  $F$  be a field,  $f \in F[x]$  be a polynomial of degree  $n$ , and  $K$  the splitting field of  $f$  over  $F$ . Then:

$$f(x) = a_n \prod_{k \in \mathbb{Z}_r} (x - \alpha_k)^{m_k} \quad (16.13.1)$$

where the  $\alpha_k$  are distinct roots, and the  $m_k$  are all positive integers.

**Definition 16.13.1** A simply root of a polynomial  $f$  over a field  $F$  is an element  $\alpha \in F$  such that  $\alpha$  is a root of  $f$  and  $(x - \alpha)^2$  does not divide  $f$ . Otherwise  $\alpha$  is called a multiple root.

**Example 16.13.1** In  $\mathbb{Q}[x]$  we can consider  $x^2 - 1$ . This splits as  $(x - 1)(x + 1)$  and so we see that the roots are  $\pm 1$ . However neither  $(x - 1)^2$  nor  $(x + 1)^2$  divides  $f$ , and hence both roots are simple. If we consider  $x^2 + 2x + 1$  then we know this factors as  $(x + 1)^2$  and so the only root is  $-1$  and it is a multiple root with multiplicity 2.

**Definition 16.13.2** A separable polynomial is one such that every root is simple. Otherwise,  $f$  is inseparable.

Over  $\mathbb{C}$  we can use calculus to determine separability. Let  $\alpha \in \mathbb{C}$  be a root of  $f(x) \in \mathbb{C}[x]$  and suppose:

$$f(x) = (x - \alpha)^m g(x) \quad (16.13.2)$$

with  $m \geq 1$ . If we look at the derivative of  $f$ , we obtain:

$$\dot{f}(x) = m(x - \alpha)^{m-1}g(x) + (x - \alpha)^m\dot{g}(x) \quad (16.13.3)$$

where  $g(\alpha) \neq 0$ . Then there are two possibilities:

$$\dot{f}(\alpha) = \begin{cases} 0, & m \geq 2 \\ g(\alpha), & m = 1 \end{cases} \quad (16.13.4)$$

Since  $g(\alpha)$  is non-zero we see that  $\alpha$  is a simply root if and only if the derivative of  $f$  evaluated at  $\alpha$  is non-zero. The idea is to generalize such notions to general fields where we may not have the structure to perform calculus. We thus define a derivation on  $F[x]$ .

**Definition 16.13.3** A derivation on an algebra  $\mathcal{A}$  over a field  $F$  is a function  $D : \mathcal{A} \rightarrow \mathcal{A}$  such that:

$$D(af + bg) = aD(f) + bD(g) \quad (16.13.5)$$

$$D(x^n) = nx^{n-1} \quad (16.13.6)$$

$$D(f \circ g) = Df(g(x)) \cdot Dg(x) \quad (16.13.7)$$

$$D(fg) = D(f)g + fD(g) \quad (16.13.8)$$

**Theorem 16.13.1.** *Linearity, Liebnizean, and  $D(x) = 1$  imply the other properties.*

**Example 16.13.2** Let  $p \in \mathbb{N}$  be a prime, and consider the field  $(\mathbb{Z}_p, +, \cdot)$  with the usual arithmetic modulo  $p$ . Let  $t \in \mathbb{Z}_p$  be some fixed constant, and consider the polynomial  $f \in \mathbb{Z}_p[x]$  defined by  $f(x) = x^p - t$ . Then  $f$  is irreducible, regardless of choice of  $t$  by applying Eisenstein's criterion to the field  $\mathbb{Z}_p[t] \subseteq \mathbb{Z}_p$ . Let  $\mathbb{Z}_p(\sqrt[p]{t}) = \mathbb{Z}_p[x]/(x^p - t)$ . Then:

$$f(x) = x^p - t = x^p - (\sqrt[p]{t})^p - (x - \sqrt[p]{t})^p \quad (16.13.9)$$

and thus  $f$  has root  $\sqrt[p]{t}$  with multiplicity  $p$ . Note that we can distribute the power only because  $\mathbb{Z}_p$  has characteristic  $p$ , and hence  $(a + b)^p = a^p + b^p$ , a consequence of the binomial theorem.

**Theorem 16.13.2.** *If  $(F, +, \cdot)$  is a field,  $f \in F[x]$  is a nonconstant, then  $f$  is separable if and only if  $f$  and  $Df$  are relatively prime.*

*Proof.* First note that the greatest common denominator is independent of the base field for two polynomials  $f$  and  $g$ . Thus, suppose  $K$  is a splitting field for  $f$  and  $\alpha$  is a root. Then  $f$  being separable implies that:

$$f(x) = (x - \alpha)g(x) \quad (16.13.10)$$

$$\Rightarrow Df(x) = g(x) + (x - \alpha)Dg(x) \quad (16.13.11)$$

$$\Rightarrow (Df)(\alpha) = g(\alpha) \quad (16.13.12)$$

But  $g(\alpha) \neq 0$ , and hence  $Df(\alpha) \neq 0$ . In the other direction, let  $\alpha \in K$  be a multiple root of  $f$  so that  $f(x) = (x - \alpha)^2 h(x)$  with  $h \in K[x]$ . But then:

$$f(x) = (x - \alpha)^2 h(x) \quad (16.13.13)$$

$$\Rightarrow Df(x) = 2(x - \alpha)h(x) + (x - \alpha)^2 Dh(x) \Rightarrow Df(\alpha) = 0 \quad (16.13.14)$$

So  $\alpha$  is also a root of  $Df$  in  $K$ . If  $f \in F[x]$  and  $K/F$  is an extension field with  $\alpha \in K$  such that  $f(\alpha) = 0$ , then  $m_{\alpha/F}(x)$  divides  $f(x)$ . So  $m_{\alpha/F}(x)$  divides the greatest common denominator of  $f$  and  $Df$ , and thus  $f$  and  $Df$  are relatively prime.  $\square$

**Example 16.13.3** Let  $F$  be a field, and let  $f(x) = x^n - a$  for some  $a \in F$ . Using the previous theorem we can determine when  $f$  is separable over  $F$ . Suppose  $a \neq 0$ . Then we have  $Df(x) = nx^{n-1}$ , and thus we have two cases:  $n = 0$  and  $n \neq 0$ . In the first case we have that the characteristic of  $F$  then divides  $n$ , and hence  $f$  is not separable. In the latter,  $n \neq 0$  and thus the characteristic does not divide  $n$ , and since the GCD of  $nx^{n-1}$  and  $x^n - a$  is 1, for otherwise 0 would be a root of  $f$  but it is not, and thus  $f$  is separable.

**Theorem 16.13.3.** *If  $f \in F[x]$  is irreducible, then  $f$  is separable if and only if  $Df$  is non-zero.*

**Theorem 16.13.4.** *If  $F$  is a field with characteristic zero, then any irreducible polynomial is separable. Hence  $f \in F[x]$  is separable if and only if  $f$  is the product of distinct irreducibles.*

**Theorem 16.13.5.** *If  $F$  is a field with characteristic  $p > 0$  then any irreducible  $f(x) = g(x^{p^r})$  for  $g \in F[x]$  is irreducible, separable, and uniquely determined for some  $r \geq 1$ .*

**Definition 16.13.4** The Frobenius map on a field of characteristic  $p \in \mathbb{N}$  is the function  $\phi : F \rightarrow F$  defined by  $\phi(x) = x^p$ .

**Definition 16.13.5** A perfect field is a field  $F$  such that every irreducible polynomial is separable.

**Example 16.13.4** Every field of characteristic zero is perfect by the previous theorems.

Do there exist fields of characteristic  $p > 0$  that are perfect? The answer is yes, and this can be categorized by the following theorem.

**Theorem 16.13.6.** *If  $(F, +, \cdot)$  is a field, then  $F$  is perfect if and only if either  $F$  has characteristic zero, or if the Frobenius map  $\phi : F \rightarrow F$  is surjective.*

## 16.14 Subfields and Automorphisms

If  $K/F$  is a finite field extension,  $A = \text{Aut}_F(K)$  the automorphism group,  $H \subseteq A$  a subgroup, the fixed field of  $H$  is the set  $K^H$  defined by all  $\alpha \in K$  such that for all  $\sigma \in H$  it is true that  $\sigma(\alpha) = \alpha$ .

**Theorem 16.14.1.** *If  $(F, +, \cdot)$  is a field, if  $(K, +, \cdot)$  is a field extension of  $F$ , if  $\text{Aut}_F(K)$  is the automorphism group of  $K$  over  $F$ , and if  $H \subseteq \text{Aut}_F(K)$  is a subgroup, then the fixed field  $K^H$  is a subfield of  $K$ .*

*Proof.* For if  $\alpha, \beta \in K^H$ , then:

$$\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta) = \alpha + \beta \quad (16.14.1)$$

and thus  $\alpha + \beta \in K^H$ . Similarly:

$$\sigma(\alpha \cdot \beta) = \sigma(\alpha) \cdot \sigma(\beta) = \alpha \cdot \beta \quad (16.14.2)$$

and therefore  $\alpha \cdot \beta \in K^H$ . Lastly:

$$\sigma(\alpha^{-1}) = \sigma(\alpha)^{-1} = \alpha^{-1} \quad (16.14.3)$$

and hence  $\alpha^{-1} \in K^H$ . □

Note that since  $\sigma \in \text{Aut}_F(K)$ , for all  $x \in F$  it is true that  $\sigma(x) = x$ . Hence,  $F \subseteq K^H$ .

**Theorem 16.14.2.** *If  $(F, +, \cdot)$  is a field, if  $(L, +, \cdot)$  is a field extension of  $F$ , and if  $(K, +, \cdot)$  is a field extension of  $L$ , then  $\text{Aut}_L(K)$  is a subgroup of  $\text{Aut}_F(K)$ .*

**Theorem 16.14.3.** *If  $(F, +_F, \cdot_F)$  is a field, if  $(K, +_K, \cdot_K)$  is a field extension of  $F$ , if  $H$  is a subgroup of the automorphism group  $\text{Aut}_F(K)$ , and if  $K^H$  is the fixed field of  $H$ , then  $H$  is a subgroup of  $\text{Aut}_{K^H}(K)$ .*

**Theorem 16.14.4.** *If  $(F, +_F, \cdot_F)$  is a field, if  $(L, +_L, \cdot_L)$  is an extension field of  $F$ , if  $(K, +_K, \cdot_K)$  is an extension field of  $L$ , if  $\text{Aut}_L(K)$  is the automorphism group, and if  $K^{\text{Aut}_L(K)}$  is the fixed field, then  $L \subseteq K^{\text{Aut}_L(K)}$ .*

**Theorem 16.14.5.**  $K/L_1/L_2/F$  implies that  $\text{Aut}_{L_2}(K)$  is a subgroup of  $\text{Aut}_{L_1}(K)$ .

**Theorem 16.14.6.** *If  $H_1 \subseteq H_2 \subseteq \text{Aut}_F(K)$ , then  $K^{H_2} \subseteq K^{H_1}$ .*

**Definition 16.14.1** A Galois extension of a field  $(F, +_F, \cdot_F)$  is a field  $(K, +_K, \cdot_K)$  such that  $\text{Card}(\text{Aut}_F(K)) = [K : F]$ .

## 16.15 More on Symmetric Polynomials

Given the symmetric group  $S_n$ , this acts on  $F[x_1, \dots, x_n]$  by permuting the variables. The symmetric polynomials are those that are invariant under permutations of variables. This is a ring. If  $K/F$  is a finite extension field, then the number of automorphisms is bounded by the degree. That is,  $\text{Card}(\text{Aut}_F(K)) \leq [K : F]$ . Let  $K/F$  and  $L/F$  are extensions, where  $K/F$  is finite (there is no constraint on  $L/F$ ). We look at  $\text{Hom}_F(K, L)$  which is the set of all homomorphisms  $\varphi : K \rightarrow L$  that fix  $F$ . Then  $\text{Hom}_F(K, L)$  is finite. For if  $K/F$  is finite, then it is finitely generated:  $K = F(\alpha_1, \dots, \alpha_n)$ , and thus any  $\varphi$  is determined by where it maps the  $\alpha_i$ , and the  $\varphi(\alpha_i)$  are roots of the minimal polynomial  $m_{\alpha_i/F}$ . So there are finitely many such choices for  $\varphi$ . If we consider the vector space homomorphisms  $\text{Hom}_{F(VS)}(K, L)$  of all linear transformations  $\varphi : K \rightarrow L$  that fix  $F$ , then there are infinitely many elements. If  $L/F$  is also finite, then the dimension of this space is also finite since it's the space of  $m \times n$  matrices where  $m = [K : F]$  and  $n = [L : F]$ .

**Theorem 16.15.1.**  $\text{Hom}_F(K, L) \subseteq \text{Hom}_{F(VS)}(K, L)$ .

*Proof.* Assume  $\text{Hom}_F(K, L)$  is dependent. So there exists elements  $a_j \in L$ , not all of which are zero, and function  $\varphi_j \in \text{Hom}_F(K, L)$  such that:

$$\sum_{i=1}^m a_j \varphi_j = 0 \tag{16.15.1}$$

□

**Theorem 16.15.2.** *Letting  $L = K$ , we have  $\text{Aut}_F(K)$  has less than  $[K : F]$  elements.*

Heading back to the before time:

**Theorem 16.15.3.** *If  $(K, +, \cdot)$  is a field, if  $H \subseteq \text{Aut}(K)$  is a finite subgroup of the automorphism group of  $K$ , and if  $K^H$  is the fixed field of  $H$ , then  $K/K^H$  is a finite field extension and  $[K : K^H] = \text{Card}(H)$ .*

Let  $K/F$  be a finite Galois extension, and let  $G \in \text{Aut}_F(G)$ . There is a bijection between the subgroups of  $H$  and the subextensions of  $L$  such that  $K/L/F$ . We map  $H \mapsto K^H$ , the fixed field of  $H$ , and we map  $L$  to  $\text{Aut}_L(K)$ . Moreover,  $H = \text{Aut}_{K^H}(K)$ . Next we need to prove that for any extension field  $L$  of  $F$  such that  $K$  is an extension of  $L$ , we must show that  $L = K^{\text{Aut}_L(K)}$ . That is,  $L$  is equal to the fixed field generated by the automorphism group  $\text{Aut}_L(K)$ .

**Theorem 16.15.4.** *If  $(F, +_F, \cdot_F)$  is a field, if  $(L, +_L, \cdot_L)$  is a field extension of  $F$ , is  $(K, +_K, \cdot_K)$  is a field extension of  $L$ , and if  $K$  is a Galois extension of  $F$ , then  $K$  is a Galois extension of  $L$ .*

*Proof.* For let  $X = \text{Hom}_F(L, K)$ . Then for  $\iota \in X$ ,  $\iota : L \rightarrow K$  is an inclusion mapping. Note that  $G$  acts on  $\text{Aut}_F(K)$  by  $g \cdot \varphi = g \circ \varphi$  for all  $g \in \text{Aut}_F(K)$  and  $\varphi \in X$ . Then the stabilizer  $G_\iota$  of  $\iota$ , which is the set of all  $g \in G$  such that  $g \cdot \iota = \iota$ . But then  $G_\iota = \text{Aut}_L(K)$  since  $\iota$  fixes  $L$ . But since  $K$  is a Galois extension of  $F$ ,  $[K : F] = \text{Card}(\text{Aut}_F(G))$ . But by the orbit stabilizer theorem:

$$\text{Card}(\text{Aut}_F(G)) = \text{Card}(\text{Aut}_F(K)_\iota) \cdot \text{Card}(\text{Aut}_F(K) \cdot \iota) \quad (16.15.2)$$

where  $G \cdot \iota = \{g \cdot \iota \mid g \in \text{Aut}_F(K)\}$ . But  $\text{Aut}_F(K)_\iota = \text{Aut}_L(K)$ , and hence:

$$[K : F] = \text{Card}(\text{Aut}_L(K)) \cdot \text{Card}(\text{Aut}_F(K) \cdot \iota) \quad (16.15.3)$$

But  $G \cdot \iota \subseteq X$ , and hence:

$$[K : F] = \text{Card}(\text{Aut}_L(K)) \cdot \text{Card}(X) \leq [K : L] \cdot [L : F] = [K : F] \quad (16.15.4)$$

and hence these inequalities are actually equalities. Thus,  $\text{Aut}_L(K) = [K : L]$ .  $\square$

**Theorem 16.15.5.** *If  $(F, +_F, \cdot_F)$  is a field, and if  $(K, +_K, \cdot_K)$  is a Galois extension of  $F$ , then  $F = K^{\text{Aut}_F(K)}$ , where  $K^{\text{Aut}_F(G)}$  is the fixed field of  $\text{Aut}_F(G)$ .*

*Proof.* For let  $E = K^{\text{Aut}_F(G)}$ . We know that  $F \subseteq E$ . But since  $K$  is a Galois extension of  $F$ , it is true that  $[K : F] = \text{Card}(\text{Aut}_F(K))$ . But any  $\varphi : K \rightarrow K$  that fixed  $F$  also fixed  $E$ , and hence  $\text{Aut}_F(G) = \text{Aut}_E(G)$ , and therefore:

$$[K : F] = \text{Card}(\text{Aut}_F(K)) = \text{Card}(\text{Aut}_E(K)) \leq [K : E] = [K : F]/[E : F] \quad (16.15.5)$$

and thus  $[E : F] = 1$ . But if  $E$  has degree 1 over  $F$ , then  $E$  is equal to  $F$ . Thus,  $F = K^{\text{Aut}_F(K)}$ .  $\square$

Some known properties of the bijection given by the fundamental theorem of Galois theory: If  $H_1 \subseteq H_2$ , then  $K^{H_2} \subseteq K^{H_1}$ . That is, the bijection is inclusion reversing. If  $H$  is a subgroup of  $\text{Aut}_F(K)$ , then the index  $[G : H] = [K^H : F]$ . In other words, if  $L$  is the subextension corresponding to  $H$ , then the index of  $G$  over  $H$  is equal to the degree of  $L$  over  $F$ .

**Theorem 16.15.6.** *If  $K/L/F$  is a subextension, if  $K/F$  is a finite Galois extension, if  $G = \text{Aut}_F(K)$ , and if  $H = \text{Aut}_L(K)$ , then  $L/F$  is Galois if and only if  $H$  is a normal subgroup of  $G$ .*

*Proof.* Let  $X = \text{Hom}_F(L, K)$ ,  $\iota \in X$ .  $G$  acts on  $X$  and  $G_\iota = \text{Aut}_L(K) = H$ . If  $L/F$  is Galois, then  $\text{Card}(\text{Aut}_F(L)) = [L : F]$  and this is equal to  $\text{Card}(X)$ . The mapping  $\text{Aut}_F(L) \rightarrow X$  defined by  $\tau \mapsto \iota \circ \tau$  is injective since  $\iota$  is injective, and hence this is also surjective since  $\text{Card}(\text{Aut}_F(L)) = \text{Card}(X)$ , and these are finite sets. So given any  $\sigma \in \text{Aut}_F(K)$ ,  $\sigma|_L : L \rightarrow K$ , so  $\sigma_L \in X$  and hence there is a  $\tau \in \text{Aut}_F(L)$  such that  $\sigma|_L = \iota \circ \tau$ . Hence  $\sigma(L) = \iota(\tau(L)) = L$ .  $\square$

The fundamental theorem of Galois theory states that if  $K/F$  is a finite Galois extension, then there is a bijection between the subgroups of  $\text{Aut}_F(K)$  and subextensions  $K/L/F$ . We map  $H \mapsto K^H$ , the fixed field of  $H$ , and we map  $L \mapsto \text{Aut}_L(K)$ .

**Theorem 16.15.7.** *If  $K/F$  is a finite Galois extension,  $K/L/F$  a subextension, and if  $H = \text{Aut}_L(K)$ , then the following are equivalent:*

- $L/F$  is Galois.
- $\sigma(L) = L$  for all  $\sigma \in \text{Aut}_F(K)$
- $H$  is a normal subgroup of  $\text{Aut}_F(K)$ .

**Theorem 16.15.8.** *If  $L/F$  and  $K/F$  are field extensions, and if  $K/F$  is finite, then  $\text{Card}(\text{Hom}_F(K, L)) \leq [L : F]$ .*

**Theorem 16.15.9.** *Given a Galois correspondence, if the subgroup  $H$  corresponds to the subextension  $L$ , and if  $\sigma \in \text{Aut}_F(K)$ , then  $\sigma H \sigma^{-1}$  corresponds to  $\sigma(L)$ .*

**Example 16.15.1**  $\mathbb{Q}(\sqrt[3]{2}, \omega)$ .

**Example 16.15.2**  $\mathbb{Q}(\exp(2\pi i/7))$ . The minimal polynomial is  $1 + x + \dots + x^5 + x^6$ . Roots of minimal polynomial are

## 16.16 Normal Extension

**Definition 16.16.1** A normal extension of a field  $(F, +_F, \cdot_F)$  is a field extension  $(K, +_K, \cdot_K)$  of  $F$  such that for all  $f \in F[x]$  such that  $f$  has a root in  $K$ , it is true that  $f$  splits over  $K$ .

**Theorem 16.16.1.** *The minimal polynomial  $m_{\alpha/F}(x)$  of any  $\alpha \in K$  splits in  $K$ . Equivalently, given any irreducible polynomial  $f(x)$  with coefficients in  $F$ , either  $f$  is irreducible over  $K$  or splits over  $K$ .*

**Theorem 16.16.2.** *If  $K/F$  is finite, and normal, then  $K$  is the splitting field of a polynomial  $f \in F[x]$ . The converse is true as well.*

Show that if  $a_1, \dots, a_n \in F$ , with  $F$  a field of characteristic not equal to 2 such that no product  $a_i \cdot a_j$  is a square. Prove  $K = F(\sqrt{a_1}, \dots, \sqrt{a_n})$  has degree  $2^n$  over  $F$ . We prove by induction on  $n$ . The base case of  $n = 1$  is true. We use the tower law and the hypothesis that  $F(\sqrt{a_1}, \dots, \sqrt{a_{n-1}})$  has degree  $2^{n-1}$  and show that  $F(\sqrt{a_1}, \dots, \sqrt{a_n})$  has degree 2 over  $F(\sqrt{a_1}, \dots, \sqrt{a_{n-1}})$ . We just need to show that  $\sqrt{a_n}$  is not in  $F(\sqrt{a_1}, \dots, \sqrt{a_{n-1}})$ . An algebraic extension  $K/F$  is normal if for every  $f \in F[x]$  with a root  $\alpha \in K$ , then  $f$  splits over  $K$ .

**Theorem 16.16.3.** *If  $K/F$  is a finite, then  $K$  is normal if and only if  $K$  is the splitting field of some  $f \in F[x]$ .*

*Proof.* For if  $K = F(\alpha_1, \dots, \alpha_n)$ , and if  $f$  is the minimum polynomial  $f = \prod_i m_{\alpha_i/F}$ , then  $K$  is the splitting field of  $f$ . Now, suppose  $K$  is the splitting field of  $f \in F[x]$  where  $f$  has degree  $n$ . Let  $\alpha \in K$  and  $g$  the minimal polynomial of  $\alpha$  over  $F$ . We want to prove that  $g$  splits in  $K$ . Let  $M$  be the splitting field of  $g$  over  $K$  and let  $\beta \in M$  be any root of  $g$ .  $\square$

**Definition 16.16.2** The normal closure of an algebraic field extension  $K/F$  is an extension  $N/K$  such that  $N/F$  is normal and minimal with this property.

**Theorem 16.16.4.** *If  $K/F$  is finite, then a normal closure  $N/K$  exists and is unique up to  $F$  isomorphism.*

*Proof.* Let  $K = F(\alpha_1, \dots, \alpha_n)$  and  $f$  the product over the minimal polynomials. Let  $N$  be the splitting field of  $f$  over  $F$ .  $\square$

**Definition 16.16.3** A separable extension of a field  $(F, +_F, \cdot_F)$  is an algebraic field extension  $(K, +_K, \cdot_K)$  of  $F$  such that for all  $\alpha \in K$  and  $m_{\alpha/F}(x) \in F[x]$  it is true that  $m_{\alpha/F}$  is separable.

### Theorem 16.16.5: Separable Field Extensions Theorem

If  $K/F$  is finite, then it is Galois if and only if it is separable and normal.

## 16.17 Radical Stuff

**Theorem 16.17.1.** *If  $K/F$  is a radical Galois extension, then  $\text{Aut}(F)K$  is solvable.*

*Proof.* If  $K$  is a radical field extension, then  $K = F(\alpha_1, \dots, \alpha_n)$  with  $\alpha_i^{n_i} \in F(\alpha_i, \dots, \alpha_{i-1})$ . We may assume that all of the  $n_i$  are prime or equal to 1. For let  $p$  be a prime that divides  $n_m$ . So then  $n_m = p^r k$ , where  $p$  does not divide  $k$ . But then:

$$F(\alpha_1, \dots, \alpha_m) = F(\alpha_1, \dots, \alpha_m^k, \alpha_m^{p^r}) \quad (16.17.1)$$

Since  $\alpha_m^k$  and  $\alpha_m^{p^r}$  can be generated by  $\alpha_m$ , the right side is a subset of the left. By Bezout's identity there exists  $a, b$  such that  $ak + bp^r = 1$ , since  $k$  and  $p^r$  are coprime. But then:

$$\alpha_m = (\alpha_m^k)^a (\alpha_m^{p^r})^b \quad (16.17.2)$$

and hence we have equality. Some stuff, and now we prove the theorem by induction on the length  $m$  of elements  $\alpha_i$ . Since  $K$  is a Galois extension, it is normal, and hence  $\alpha_1$  has another conjugate  $\beta$ , that is, a root of the minimal polynomial  $m_{\alpha_1/F}(x)$ , in  $K$ , call it  $\beta$ . But then  $\alpha_1^{n_1} \in F$  implies that  $\beta^{n_1} \in F$ , and then  $(\alpha_1/\beta)^{n_1} = 1$ .  $\square$

## 16.18 Quartic Polynomials

Suppose  $G \subseteq S_4$  is a transitive subgroup. The possibilities are the alternating group  $A_4$ , the dihedral group  $D_8 = \langle (12)(34), (1234) \rangle$ , the Klein 4 group, denoted  $V_4 = \langle 1, (12)(34), (13)(24), (14)(23) \rangle$ , and the cyclic group  $\mathbb{Z}_4 = \langle (1234) \rangle$ . The only normal subgroups of  $S_4$  are  $A_4$ ,  $D_8$ ,  $V_4$ , and  $S_4$  itself. Moreover,  $V_4$  is a normal subgroup of  $A_4$ . If the discriminant is a square in  $F$ , then  $G$  is a subgroup of  $A_4$  and hence either  $G = V_4$  or  $G = A_4$ . If  $f(x) = x^4 + ax^2 + b$ , then  $G$  is a subgroup of  $D_8$  and hence  $G$  is isomorphic to the trivial group,  $\mathbb{Z}/2\mathbb{Z}$ ,  $\mathbb{Z}/4\mathbb{Z}$ ,  $\mathbb{Z}_2 \times \mathbb{Z}_2$ , or  $D_8$ . Since  $f$  is irreducible we can conclude that  $G$  is isomorphic to either  $\mathbb{Z}/4\mathbb{Z}$ ,  $\mathbb{Z}_2 \times \mathbb{Z}_2$ , or  $D_8$ . Let  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  be roots of  $f(x)$  and defined  $\beta_i$  as follows:

$$\beta_1 = \alpha_1\alpha_4 + \alpha_2\alpha_3 \quad (16.18.1a)$$

$$\beta_2 = \alpha_1\alpha_3 + \alpha_2\alpha_4 \quad (16.18.1b)$$

$$\beta_3 = \alpha_1\alpha_2 + \alpha_3\alpha_4 \quad (16.18.1c)$$

Then the Galois group  $G$  of  $f$  acts on  $\{\beta_1, \beta_2, \beta_3\}$ , so we get the following diagram:

$$\begin{array}{ccc} G & \longrightarrow & S_3 \\ \downarrow & & \downarrow \\ S_4 & \longrightarrow & S_3 \end{array}$$

Fig. 16.1: Lagrange Discriminant

the kernel of the induced homomorphism  $\varphi : S_3 \rightarrow S_4$  is the Klein 4 group  $V_4$ . If  $G = S_4$  then  $g$  is irreducible. So, in summary, if  $f \in F[x]$  is an irreducible quartic, if  $g \in F[x]$  is its cubic resolvent:

$$g(x) = (x - \beta_1)(x - \beta_2)(x - \beta_3) \quad (16.18.2)$$

then we have the following table of facts:

$\Delta(f)$	$g(x)$	$\text{Aut}_F(K)$
Not a square	Irreducible	$S_4$
Square	Irreducible	$A_4$
Not a square	Has one root	$\mathbb{Z}_4$ or $D_8$
Square	Splits	$\mathbb{Z}_2 \times \mathbb{Z}_2$

Table 16.3: Classifying Quartics

## 16.19 Stuff

If  $\langle \cdot | \cdot \rangle$  is a symmetric bilinear form, it is represented by its Gram matrix relative to a basis  $\mathcal{B}$  of the finite dimensional vector space  $\mathcal{B}$ . We have:

$$\langle v | w \rangle = [V]_{\mathcal{B}}^T G_{\mathcal{B}} [W]_{\mathcal{B}} \quad (16.19.1)$$

Let  $\langle \cdot | \cdot \rangle : V \times V \rightarrow k$  be bilinear, let  $\mathcal{B}$  and  $\mathcal{C}$  be bases of  $V$ , and let  $G_{\mathcal{B}}$  and  $G_{\mathcal{C}}$  be the corresponding Gram matrices. Let  $x, y \in V$ . Then:

$$[x]_{\mathcal{B}}^T G_{\mathcal{B}} [y]_{\mathcal{B}} = \langle x | y \rangle \quad (16.19.2a)$$

$$= [x]_{\mathcal{C}}^T G_{\mathcal{C}} [Y]_{\mathcal{C}} \quad (16.19.2b)$$

$$= [\text{Id}(x)]_{\mathcal{C}}^T G_{\mathcal{C}} [\text{Id}(y)]_{\mathcal{C}} \quad (16.19.2c)$$

$$= \left( [\text{Id}]_{\mathcal{C}}^{\mathcal{B}} [x]_{\mathcal{B}} \right)^T G_{\mathcal{C}} \left( [\text{Id}]_{\mathcal{C}}^{\mathcal{B}} [y]_{\mathcal{B}} \right) \quad (16.19.2d)$$

From this, we obtain the formula for the Gram matrix:

$$G_{\mathcal{B}} = P^T G_{\mathcal{C}} P \quad (16.19.3)$$

**Definition 16.19.1: Congruent Matrices**

Congruent matrices are matrices  $A, B \in M_n(k)$  such that there is an invertible matrix  $P \in GL_n(k)$  such that:

$$B = P^T AP \quad (16.19.4)$$

## 16.20 Orthogonal Transformations

**Theorem 16.20.1: Sylvester's Theorem**

If  $V$  is a finite dimensional vector space over  $\mathbb{R}$  and if  $\langle \cdot | \cdot \rangle$  is a symmetric bilinear form, then there exists a basis  $\mathcal{B}$  of  $V$  such that:

$$G_{\mathcal{B}} = \begin{bmatrix} 1 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & -1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} \quad (16.20.1)$$

Where there are  $r$  1's,  $s$  negative 1's, and  $t$  zeros, and  $r, s$ , and  $t$  are uniquely determined.

Let  $(V, \langle \cdot | \cdot \rangle)$  be a nondegenerate bilinear space. Then a linear map  $T : V \rightarrow V$  is orthogonal if or a linear isometry if:

$$\langle T(v) | T(w) \rangle = \langle v | w \rangle \quad (16.20.2)$$

Let  $\mathcal{B}$  be a basis of  $V$  and let  $G_{\mathcal{B}}$  be the Gram matrix of  $\langle \cdot | \cdot \rangle$ . Let  $A$  be the representing matrix of  $T$  over the basis  $\mathcal{B}$ . Then:

$$\langle v | w \rangle = \langle T(v) | T(w) \rangle = [T(v)]_{\mathcal{B}}^T G_{\mathcal{B}} [T(w)]_{\mathcal{B}} = [v]_{\mathcal{B}}^T A^T G_{\mathcal{B}} A [w]_{\mathcal{B}} \quad (16.20.3)$$

From this we can compute what the Gram matrix is:

$$G_{\mathcal{B}} = A^T G_{\mathcal{B}} A \quad (16.20.4)$$

If  $A$  represents an orthogonal transformation, then  $A$  must satisfy this equation. In the special case of when  $\mathcal{B}$  is orthonormal, then  $G_{\mathcal{B}}$  is simply the identity matrix and thus we have that  $A^T A = I$ , or  $A^T = A^{-1}$ .

A nondegenerate skew symmetric bilinear form on a  $2n$  dimensional real vector space is called a symplectic form. The Gram matrix is:

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \quad (16.20.5)$$

A transformation  $A$  such that  $A^TJA = J$  is called a symplectic or a canonical transformation.

If  $\langle \cdot | \cdot \rangle$  is the Lorentz form on  $\mathbb{R}^4$ , then an orthogonal transformation is called a Lorentz transformation.

## 16.21 Sesquilinear Geometry

Let  $V$  and  $W$  be  $\mathbb{R}$  inner product spaces. That is,  $V$  and  $W$  are equipped with a symmetric bilinear form  $\langle \cdot | \cdot \rangle_V$  and  $\langle \cdot | \cdot \rangle_W$  that are positive-definite:

$$\langle v | v \rangle \geq 0 \quad (16.21.1)$$

With equality if and only if  $v = 0$ . Note that positive definite implies non-degenerate since  $\langle v | v \rangle > 0$  for nonzero  $v$ . Let  $T : V \rightarrow W$  be a linear map. Then  $T$  induces  $T^* : W^* \rightarrow V^*$  by something. Let  $V$  be a finite dimension  $\mathbb{R}$  vector space. An inner product on  $V$  is a bilinear form that is symmetric and positive-definite. Euclidean geometry is derived from the inner product. This induces a norm and the notion of angle:

$$\|v\| = \sqrt{\langle v | v \rangle} \quad \theta = \cos^{-1} \left( \frac{\langle v | w \rangle}{\|v\| \|w\|} \right) \quad (16.21.2)$$

Cauchy-Schwartz comes out of this. Let  $V$  be a vector space over  $\mathbb{C}$ , say  $V = \mathbb{C}^n$ . If we define:

$$\langle v | w \rangle = \sum_{k=1}^n v_k w_k \quad (16.21.3)$$

We lose positive-definiteness since  $(1, 2i) \cdot (1, 2i) = -3$ , in  $\mathbb{C}^2$ , for example. We define the dot product on  $\mathbb{C}^n$  by:

$$\langle z | w \rangle = \sum_{k=1}^n \bar{z}_k \bar{w}_k \quad (16.21.4)$$

That is,  $\langle v | w \rangle = \bar{v} \cdot w$ . From this we have that  $\langle z | \cdot \rangle$  is linear on  $\mathbb{C}$ . However, looking at  $\langle \cdot | w \rangle$ , we have that it is  $\mathbb{R}$  linear but only conjugate linear over  $\mathbb{C}$ . This gives rise to the notion of a sesquilinear product.

**Definition 16.21.1: Sesquilinear Product**

A sesquilinear product on a vector space  $V$  over  $\mathbb{C}$  is a function  $\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{C}$  such that  $\langle z | \cdot \rangle$  is  $\mathbb{C}$  linear and  $\langle \cdot | w \rangle$  is  $\mathbb{R}$  linear and  $\mathbb{C}$  conjugate linear.

With this we can define a Hermitian form on a  $\mathbb{C}$  vector space  $V$ . Note that a sesquilinear form is conjugate symmetric. That is,  $\langle w | z \rangle = \overline{\langle z | w \rangle}$ .

**Definition 16.21.2: Hermitian Form**

A Hermitian form on a  $\mathbb{C}$  vector space  $V$  is a function  $\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{C}$  that is sesquilinear and conjugate symmetric.

**Definition 16.21.3: Hermitian Inner Product Space**

A Hermitian inner product space is a vector space  $V$  over  $\mathbb{C}$  with a Hermitian inner product.

Let  $\mathcal{B}$  be a basis of a Hermitian inner product space  $V$  and let  $G_{\mathcal{B}} = [g_{ij}]_{ij}$  where  $g_{ij} = \langle e_i | e_j \rangle$ . Then for  $v, w \in V$ , we have:

$$v = \sum_{k=1}^n a_k e_k \quad w = \sum_{k=1}^n b_k e_k \quad (16.21.5)$$

And moreover:

$$\langle v | w \rangle = \left\langle \sum_{j=1}^n a_j e_j \mid \sum_{k=1}^n b_k e_k \right\rangle = \sum_{k=1}^n \sum_{j=1}^n \bar{a}_j b_k \langle e_j | e_k \rangle = \sum_{j=1}^n \sum_{k=1}^n \bar{a}_j g_{jk} b_k \quad (16.21.6)$$

So we have:

$$\langle v | w \rangle = \overline{[v]_{\mathcal{B}}^T} G_{\mathcal{B}} [w]_{\mathcal{B}} \quad (16.21.7)$$

This gives rise to the definition of a Hermitian transpose.

**Definition 16.21.4: Hermitian Transpose**

The Hermitian transpose of a matrix  $A$  is the matrix:

$$A^H = \overline{A^T}$$

## 16.22 Unitary Transformations

### Definition 16.22.1: Unitary Transformations

A unitary transformation on a vector space  $V$  over  $\mathbb{C}$  is a linear function  $T : V \rightarrow V$  such that:

$$\langle T(v)|T(w)\rangle = \langle v|w\rangle$$

Let  $V$  and  $W$  be Hermitian inner product spaces and let  $T : V \rightarrow W$  be  $\mathbb{C}$  linear.  $T$  induces  $T^* : W^* \rightarrow V^*$  by  $T^*(\psi) = \psi \circ T$ .

### Definition 16.22.2: Self-Adjoint Hermitian Operator

A self-adjoint operator on a Hermitian inner product space  $V$  is a linear function  $T : \mathbb{C} \rightarrow \mathbb{C}$  such that  $T = T^H$ .

### Definition 16.22.3: Normal Hermitian Operator

A function  $T : \mathbb{C} \rightarrow \mathbb{C}$  such that  $TT^H = T^HT$

## 16.23 Spectral Theorem

**Theorem 16.23.1.** *If  $V$  is a finite dimensional vector space over  $\mathbb{C}$ , if  $T : V \rightarrow V$  is a linear map, then  $T$  has an eigenvalue.*

*Proof.* For the characteristic polynomial is non-constant and thus by the fundamental theorem of algebra there exists a root.  $\square$

**Theorem 16.23.2.** *If  $V$  is a finite dimensional Hermitian inner product space and if  $T : V \rightarrow V$  is a Hermitian operator, then the eigenvalues of  $T$  are real and if  $v$  is an eigenvector of  $\lambda$  and if  $w$  is a  $\mu$  eigenvector for two different eigenvalues  $\lambda \neq \mu$ , then  $v$  and  $w$  are orthogonal.*

*Proof.* For let  $\lambda$  be an eigenvalue and let  $v$  be a non-zero eigenvector for  $\lambda$ . Then  $T(v) = \lambda v$ . But  $T$  is Hermitian, and therefore:

$$\langle T^H(v)|v\rangle = \langle v|T(v)\rangle = \langle v|\lambda v\rangle = \lambda \langle v|v\rangle \quad (16.23.1)$$

But also:

$$\langle T^H(v)|v\rangle = \langle T(v)|v\rangle = \langle \lambda v|v\rangle = \bar{\lambda}\langle v|v\rangle \quad (16.23.2)$$

And therefore, since  $\langle v|v\rangle \neq 0$ , we have  $\lambda = \bar{\lambda}$ , and therefore  $\lambda$  is real. For the second part, we have:

$$\langle v|T(w)\rangle = \langle v|\mu w\rangle = \mu\langle v|w\rangle = \langle T^H(v)|w\rangle = \langle T(v)|w\rangle = \langle \lambda v|w\rangle = \bar{\lambda}\langle v|w\rangle \quad (16.23.3)$$

But we just proved that  $\lambda$  is real, and thus  $\bar{\lambda} = \lambda$ . Moreover  $\lambda \neq \mu$ , and thus for equality to occur we must have  $\langle v|w\rangle = 0$ .  $\square$

**Theorem 16.23.3.** *If  $V$  is a finite dimensional Hermitian inner product space, if  $T : V \rightarrow V$  is linear, and if  $W \subseteq V$  is a  $T$  invariant subspace (that is,  $T(W) \subseteq W$ ), then  $W^\perp$  is  $T^H$  invariant.*

*Proof.* For let  $x \in W^\perp$  and let  $w \in W$ . Then:

$$\langle T^H(x)|w\rangle = \langle x|T(w)\rangle = 0 \quad (16.23.4)$$

And thus  $T(x) \in W^\perp$ . Therefore,  $T(W^\perp) \subseteq W^\perp$ .  $\square$

### Theorem 16.23.4: Unitary Triangulation Theorem

If  $V$  is a finite dimensional Hermitian inner product space over  $V$  and if  $T : V \rightarrow V$  is a linear operator, then there exists an orthonormal basis  $\mathcal{B}$  of  $V$  such that  $[T]_{\mathcal{B}}^{\mathcal{B}}$  is upper triangular.  $\blacksquare$

*Proof.* For consider  $T^H : V \rightarrow V$ . It has an eigenvalue  $\lambda$ . Let  $v$  be a  $\lambda$  eigenvector of  $T^H$  and let  $W = \mathbb{C} \cdot v = \text{Span}\{zv : z \in \mathbb{C}\}$ . Since  $v$  is an eigenvector of  $T^H$  we have that  $W$  is a  $T^H$  invariant subspace so therefore  $W^\perp$  is invariant under  $(T^H)^H = T$ . That is,  $W^\perp$  is a  $T$  invariant subspace. Since  $W$  is non-zero,  $W^\perp$  has dimension less than  $V$  and thus by induction there is an orthonormal basis of  $W^\perp$  such that  $[T_{W^\perp}]_{\mathcal{B}'}^{\mathcal{B}'}$  is upper triangular. Extending  $\mathcal{B}$  from  $\mathcal{B}'$  gives an orthonormal basis such that  $[T]_{\mathcal{B}}^{\mathcal{B}}$  is upper triangular.  $\square$

**Theorem 16.23.5.** *If  $A \in M_n(\mathbb{C})$  then there is a unitary matrix  $P \in U(n)$  such that  $PAP^{-1}$  is upper triangular.*

### Theorem 16.23.6: Spectral Theorem for Hermitian Operators

If  $V$  is a finite dimensional Hermitian inner product space and if  $T : V \rightarrow V$  is a Hermitian operator then there exists an orthonormal basis  $\mathcal{B}$  of  $V$  such that  $[T]_{\mathcal{B}}^{\mathcal{B}}$  is diagonal.

*Proof.* For the representing matrix of  $T$  is upper triangular, and thus the representing matrix for  $T^H$  is lower triangular. But  $T = T^H$  and therefore the representing matrix of  $T$  is both upper and lower triangular, and therefore it is diagonal.  $\square$

The theorem holds for normal operators as well. For Hermitian we see that the eigenvalues are real. For skew-Hermitian ( $T = -T^H$ ) the eigenvalues are purely imaginary, and lastly for a unitary  $T$  the eigenvalues are unit modulus.

### Theorem 16.23.7: Real Spectral Theorem

If  $V$  is a finite dimensional inner product space, if  $T : V \rightarrow V$  is a self-adjoint operator then there is an orthonormal basis of  $V$  such that the representing matrix is diagonal.

*Proof.* Let  $A$  be the representing matrix of  $T$  relative to the standard basis of  $\mathbb{R}^n$ . Then  $A = A^T$  and thus  $A$  defines a linear map  $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$  by mapping  $A(v) = Av$  as a matrix operation. Then  $A$  is a Hermitian operator and therefore  $A$  has real eigenvalues. Let  $v \in \mathbb{C}^n$  be a complex eigenvector for  $\lambda$ . Let  $v = x + iy$  for  $x, y \in \mathbb{R}^n$ . Then since  $Av = \lambda v$  we have:

$$Ax + iAy = \lambda x + i\lambda y \quad (16.23.5)$$

But  $A$  is real, as are  $x$  and  $y$ , and thus we have:

$$Ax = \lambda x \quad Ay = \lambda y \quad (16.23.6)$$

But since  $v$  is non-zero, at least one of  $x$  or  $y$  is non-zero. Thus  $A$  has a real eigenvector,  $x$  or  $y$ . Let  $u$  be a unit real eigenvector and let  $W$  be the real span of  $u$ . Then  $W$  is  $T$  invariant and therefore  $W^\perp$  is  $T^H$  invariant, but  $T = T^H$ . We complete the proof by induction.  $\square$

## 16.24 Tensor Products

Let  $R$  be a ring and let  $M$  be a right  $R$  module and let  $N$  be a left  $R$  module. We seek a universal product  $M \times N$  into  $M \otimes_R N$ . Examples of products: Dot products, cross products, bilinear forms, matrix multiplication. All of these are Biadditive. Matrices also have the *balanced* property:

$$A(cB) = (Ac)B \quad (16.24.1)$$

If the underlying ring is not commutative, we may not have perfect bilinearity and so we replace this requirement with the balanced property. So we seek a

function  $\beta : M \times N \rightarrow M \otimes_R N$  such that:

$$\beta(a + b, c) = \beta(a, c) + \beta(b, c) \quad (\text{Left Additivity})$$

$$\beta(a, b + c) = \beta(a, b) + \beta(a, c) \quad (\text{Right Additivity})$$

$$\beta(ar, b) = \beta(a, rb) \quad (\text{Balanced})$$

Universal means that given any Abelian group  $A$  and any biadditive  $R$  balanced map  $\mu : M \times N \rightarrow A$  there is a unique  $\mathbb{Z}$  module homomorphism  $\tilde{\mu} : M \otimes_R N \rightarrow A$  such that the diagram commutes. Since if it exists it is unique up to unique isomorphism, we need only construct such a thing. Long complicated construction to follow. By construction,  $M \otimes_R N$  is generated by  $m \otimes n$  for all  $m \in M, n \in N$ . As a warning, not every element of  $M \otimes_R N$  need be decomposable. It's folly to try to define a map  $M \times N \rightarrow A$  by  $m \otimes n \mapsto f(m, n)$  unless  $f(m, n)$  is biadditive and  $R$  balanced.

**Theorem 16.24.1.** *If  $N$  is a left  $R$  module, then  $R \otimes_R N$  is isomorphic to  $N$  as a  $\mathbb{Z}$  module.*

*Proof.* For let  $\mu : R \times N \rightarrow N$  by defined by  $\mu(r, n) = r \cdot n$ . Then  $\mu$  is biaddiative and balanced.  $\square$

The idea is to use scalar multiplication  $R \times N \rightarrow N$  to construct such a module. The by the universal mapping property there exists a unique  $\mathbb{Z}$  module homomorphism  $\tilde{\mu} : R \otimes_R N \rightarrow N$  such that the diagram commutes. Define  $\tilde{\mu}$  by:

$$\tilde{\mu}(r \otimes n) = r \cdot n \quad (16.24.2)$$

Define  $j : N \rightarrow R \otimes_R N$  by  $j(n) = 1 \otimes n$ . This is a  $\mathbb{Z}$  module since the tensor product is biaddiative and balanced. Moreover,  $\tilde{\mu}$  and  $j$  are inverses of each other.

**Theorem 16.24.2.** *If  $R$  is a ring, if  $I$  is an ideal of  $R$ , and if  $N$  is a left  $R$  module, then:*

$$R/I \otimes_R N \simeq N/IN \quad (16.24.3)$$

*Proof.* For define  $\mu : R/I \times N \rightarrow N/IN$  by:

$$\mu(\bar{r}, n) = \mu(r + I, n) = rn + IN = \bar{rn} \quad (16.24.4)$$

For  $n \in N$  let  $f_N : R \rightarrow N/IN$  be defined by  $f_N(r) = rn + IN = \bar{rn}$ . Then  $f_N$  is a map of left  $R$  modules and therefore induces a map  $\bar{f}_N$  from  $R/I$  to  $N/IN$  by  $\bar{r} \mapsto \bar{rn}$ . By the universal mapping property there is a  $\mathbb{Z}$  linear map  $\tilde{\mu}$  such that the diagram commutes. So we have that  $\tilde{\mu}(\bar{r} \otimes n) = \bar{rn}$ . We now want a map  $N/IN \rightarrow R/I \otimes_R N$ . Let  $j$  be defined by  $j(n) = \bar{1} \otimes n$ . Then  $j$  vanishes on generators of  $IN$  and therefore induces  $\bar{j}$  such that  $\bar{n} \mapsto \bar{1} \otimes n$ . Thus we now have two maps that are well defined and now we need only check that they are inverses of each other. And it is so. So we are done.  $\square$

**Theorem 16.24.3.** *Given a sequence of modules over  $R$ ,  $N' \rightarrow N \rightarrow N'' \rightarrow 0$ , and suppose that for all left  $R$  modules  $Y$  the sequence:*

$$0 \rightarrow \text{Hom}_R(N'', Y) \rightarrow \text{Hom}_R(N, Y) \rightarrow \text{Hom}_R(N', Y) \quad (16.24.5)$$

*is exact, then the original sequence is exact.*

*Proof.* For let  $Y = \text{coker}(v) = N''/v(N)$ . Then since:

$$0 \rightarrow \text{Hom}_R(N'', Y) \rightarrow \text{Hom}_R(N, Y) \rightarrow \text{Hom}_R(N', Y) \quad (16.24.6)$$

is exact, we have that the canonical projection  $\pi$  gets mapped to  $v^*(\pi)$ . But  $v^*(\pi) = \pi \circ v$ , and this is zero since  $n \mapsto v(n)$  which maps to 0 by  $\pi$ . Thus  $v^*$  is surjective and injective. On the other side, let  $Y = N''$ . Then  $\text{id}_{N''}$  maps to  $v$  under  $v^*$ , and thus  $v$  maps to 0 under  $u^*$  since the sequence is exact, and thus  $u \circ v = 0$ . Thus  $\text{im}(u) \subseteq \ker(v)$ .  $\square$

An algebra over a commutative ring  $k$  is a left  $k$  module with a ring structure such that the ring multiplication is compatible with scalar multiplication:

$$(\lambda \star a) \cdot b = \lambda \star (a \cdot b) = a \cdot (\lambda \star b) \quad (16.24.7)$$

An equivalent definition is a ring  $A$  with a ring homomorphism  $\varphi : k \rightarrow Z(A)$ . We can also define an algebra in terms of the tensor product. An algebra over  $k$  is a  $k$  module  $A$  with a homomorphism  $\mu : A \otimes A \rightarrow A$  and a module homomorphism  $\eta : k \rightarrow A$  such that some diagram commutes.

## 16.25 Bonus Stuff

Given a ring homomorphism  $\varphi : R \rightarrow S$ , there is a map  $\varphi^\# : \text{Spec}(S) \rightarrow \text{Spec}(R)$ , where  $\text{Spec}(X)$  is the set of all prime ideals on  $X$ . Zariski topology. Define, for all  $r \in R$ :

$$X_r = \{p \in \text{Spec}R \mid r \in p\} \quad (16.25.1)$$

The topology is  $\tau = \{X_r \mid r \in R\}$ .

## **Part VIII**

# **Unsorted Stuff**



# CHAPTER 17

---

## Combinatorics

---

### 17.1 Introduction

We use the following notation:

$$[n] = \mathbb{Z}_n = \{1, 2, \dots, n\} \quad (17.1.1)$$

In combinatorics we study functions of the form  $f : [n] \rightarrow [m]$ . These functions can be one-to-one, onto, or both. A permutation of the elements of  $[n]$  is simply a bijection  $f : [n] \rightarrow [n]$ . A partial permutation is a permutation of length  $k \leq n$  of the set  $[n]$ . That is, a permutation on some subset of  $\mathbb{Z}_n$ . We all study sets. In particular, the power set of  $[n]$  and subsets from  $k$  chosen elements of  $[n]$ . Another topic of study is that of lattice paths on  $\mathbb{Z} \times \mathbb{Z}$ . That it, paths from  $(i, j)$  to  $(n, m)$  using a prescribed set of rules. For example, how many ways can you get from  $(0, 0)$  to  $(21, 7)$  if you're only allowed to move North and East. Restricted paths impose more rules, for example the number of moves east must be greater than the number of moves north. There are also things called Catalan paths and Motzkin paths.

Another topic in combinatorics is that of words. An alphabet is a set  $[n]$ , and we wish to study the number of words of length  $k$  in  $[n]$ . Binary words are words when the alphabet is  $\{0, 1\}$ . There are also words with a prescribed number for each letter. There are also circular arrangements of the elements in  $[n]$ , and the idea of multi-sets. Multi-sets are sets that allow for repetition. From elementary set theory, we have:

$$\{a, b, c\} = \{a, a, b, c\} \quad (17.1.2)$$

Sets are uniquely defined by the elements they contain. Multi-sets allow for repetition, and this distinguishes two different sets. We write:

$$A = \{\{1, 1, 1, 2, 3, 3\}\} \quad (17.1.3)$$

Note that  $A \neq \{\{1, 2, 3\}\}$ . The multiplicity of an elements  $a \in A$  is the number of times the element  $a$  occurs in the multi-set  $A$ . To know how many multi-sets chosen from  $[n]$  with  $k$  elements, we wish to study the following equation:

$$\sum_{i=1}^n m_i = k \quad (17.1.4)$$

Where  $m_i$  is the multiplicity of the  $i^{th}$  element of  $[n]$ . Where wish to find integer solutions to this equation, in particular solutions with non-negative integers. We can also place restrictions on the multi-sets, for example by requiring that each element occurs at least once. Thus we'd have:

$$\sum_{i=1}^n m_i = k \quad m_i \geq 1 \quad (17.1.5)$$

A partition of  $[n]$  is a collection of sets  $B_1, b_2, \dots, B_k$  such that:

$$\cup_{i=1}^k B_i = [n] \quad B_i \cap B_j = \emptyset \quad i \neq j \quad (17.1.6)$$

The  $B_i$  are called blocks. We also study partitions of numbers. Given  $n \geq 0$ , we want  $\lambda = (\lambda_1, \dots, \lambda_\ell)$  such that:

$$\lambda_{i+1} \leq \lambda_i, \quad i = 1, 2, \dots, \ell - 1 \quad (17.1.7)$$

And such that:

$$\sum_{i=1}^{\ell} \lambda_i = n \quad (17.1.8)$$

Another commonly studied object is a graph. Labelled trees, colorings of graphs, and spanning trees. Finally we study tableaux's. These are fillings of arrays of boxes with objects, such as numbers, sets, and multi-sets.

## 17.2 Counting Techniques

### 17.2.1 Basic Numbers

$$n^k = |\{f : [k] \rightarrow [n]\}| = \text{The number of words over the alphabet } [n] \quad (17.2.1)$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \text{The number of ways to choose } k \text{ objects from } [n] \quad (17.2.2)$$

Stirling's number of the second kind, denoted  $S(k, n)$ , is the number of ways to partition  $[n]$  into  $n$  blocks.  $P(n)$  is the number of partitions of the integer  $n$ . And lastly,  $n!$  is the number of bijections from  $[n]$  into  $[n]$ . That is,  $n!$  is the number of permutations on  $[n]$ .

### 17.2.2 Basic Counting Principles

Sum Rule (Divide and Conquer): If you cannot count the set, divide it into pieces and count the pieces.

$$S = \bigcup_{i=1}^k S_i \quad S_i \cap S_j = \emptyset, \quad i \neq j \quad (17.2.3)$$

$$|S| = \sum_{i=1}^k |S_i| \quad (17.2.4)$$

Application: Classify or partition the elements in  $S$  according to a set of mutually disjoint properties  $p_1, \dots, p_k$  and let:

$$S_k = \{x \subseteq S : p_k(x)\} \quad (17.2.5)$$

We often use such a scheme to prove recurrences. For example Pascal's triangle:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \quad (17.2.6)$$

We can prove this by plugging in the formula and simplifying, but we wish to give a combinatorial proof. Let  $S \subseteq [n]$ , where  $\text{curl}(S) = k \leq n$ . Define:

$$S_1 = \{A \subseteq S : n \in A\} \quad (17.2.7)$$

$$S_2 = \{A \subseteq S : n \notin A\} \quad (17.2.8)$$

Then  $S_1$  and  $S_2$  partition  $S$ , and thus  $\text{curl}(S) = \text{curl}(S_1) + \text{curl}(S_2)$ . But:

$$\text{curl}(S_1) = \binom{n-1}{k-1} \quad (17.2.9)$$

$$\text{curl}(S_2) = \binom{n-1}{k} \quad (17.2.10)$$

This completes the proof. Next is the difference rule (Count the opposite):

$$S \subseteq \mathcal{U} \quad (17.2.11)$$

$$\bar{S} = \mathcal{U} \setminus S \quad (17.2.12)$$

$$\text{curl}(S) = \text{curl}(\mathcal{U}) - \text{curl}(\bar{S}) \quad (17.2.13)$$

For example, how many permutations on  $[n]$  are there so that 1 and 2 are not next to each other? We count 1 and 2 being together as 12 or 21. It's thus easier to think of this as one element. So we're counting the number of permutation on  $[n-1]$  by consider 12 as one element, and again by consider 21 as one element. This gives  $2(n-1)!$ . By taking the difference, we have:

$$\text{curl}(S) = n! - 2(n-1)! = (n-1)!(n-2) \quad (17.2.14)$$

## 17.3 Posets

Let  $n \in \mathbb{N}$ ,  $X = 2^{[n]} = \mathcal{P}(\mathbb{Z}_n)$ , and let consider  $(X, \subseteq)$ , where  $\subseteq$  is the partial ordering of inclusion. If  $\text{curl}(A) = \text{curl}(B)$ , then either  $A = B$ , or  $A$  and  $B$  are not comparable, and hence  $\subseteq$  is a partial order.

**Theorem 17.3.1.** *For any  $n \geq 1$ ,  $2^{[n]}$  has a SCD.*

*Proof.* Let  $x \in 2^{[n]}$ . Then  $x$  has a binary representation. We want to create 1-0 pairs as if they were parentheses. For example, suppose  $x = \{1, 5, 7\}$ . Then  $x = 0110001$ . We match this up to  $)((\leftrightarrow 0((0)01$ . To get the chain containing  $x$ , we need to describe how to go up and how to go down in the chain. To go up, take the right-most unpaired zero and change it to a one. To go down, take the left-most one and change it to a zero. For the example of  $x = \{1, 5, 6\}$ , we have:

$$0110000 \rightarrow 0110001 \rightarrow 0110011 \rightarrow 1110011$$

The size of the smallest set in the chain the the number of parenthesizations. If this number is  $i$ , then the largest set has size  $(n-2i)+i = n-i$ . By construction, the chain is saturated. At every step we only add one element. Thus, the constructions produce symmetric chains. Notice that we never produce new 1-0 pairings in the algorithm. Thus, all of the sets in the chain have the same pairings. So two sets  $X$  and  $Y$  produce either the same chain or disjoint chains. For example, consider  $2^{[4]}$ . The chain is:

$$\begin{aligned} 0000 &\rightarrow 0001 \rightarrow 0011 \rightarrow 0111 \rightarrow 1111 \\ 0010 &\rightarrow 0110 \rightarrow 1110 \\ 0100 &\rightarrow 0101 \rightarrow 1101 \\ 1000 &\rightarrow 1001 \rightarrow 1011 \end{aligned}$$

□

### Theorem 17.3.2: Sperner's Theorem

Any anti-chain of  $2^{[n]}$  elements has at most  $\binom{n}{\lfloor n/2 \rfloor}$  subsets. ▀

*Proof.* Notice that in the diagram of  $2^{[n]}$  each level contains  $\binom{n}{k}$  elements. An SCD partitions  $2^{[n]}$ . Let  $G$  be an SCD. with  $m$  chains. Then the maximum number of incomparable elements is  $m$ . otherwise, two imcomparable elements are in the same chain. Thus,  $m \geq \binom{n}{\lfloor n/2 \rfloor}$ . Also, each chain will intersect a level in the graph of the poset at most once. Therefore,  $m \leq \binom{n}{\lfloor n/2 \rfloor}$ . Thus, etc. □

The existence of a symmetric chain decomposition gives an elegant combinatorial proof that the sequence  $\binom{n}{k}$ ,  $k = 0, 1, \dots, n$ , is unimodal. To prove unimodality, it would suffice to show that, for  $k \leq n/2$ , there exists an injection from  $k$  subsets to  $k + 1$  subsets. If we have a SCD, map a  $k$  subset to its successor in the chain. This gives the injection.

## 17.4 Binomial Coefficients and Multi-Sets

Recall that a multi-set is similar to a set, except that repetitions are allowed. For example, if we consider [3], then a multi-set of size 5 could be:

$$\{\{1, 1, 2, 2, 3\}\} \quad (17.4.1)$$

This has 5 elements, and is a multi-set of size 5.

**Theorem 17.4.1.** *The number of  $k$  multisets of an  $n$  element set is:*

$$\frac{n^k}{k!} = \frac{n(n+1)\cdots(n+k-1)}{k!} = \binom{n+k-1}{k} \quad (17.4.2)$$

*Proof.* For let  $S$  be the set of multi-sets of size  $k$  of elements of an  $n$  element set, and let  $T$  be subsets of size  $k$  in  $[n+k-1]$ . We need to produce a map  $f : S \rightarrow [T]$ . Let:

$$M = \{\{1 \leq a_1 \leq a_2 \leq \cdots \leq a_k \leq n\}\} \quad (17.4.3)$$

This maps to the set:

$$A = \{1 \leq a_1 < a_2 + 1 < a_3 + 2 < \cdots < a_k + k - 1 \leq n + k - 1\} \quad (17.4.4)$$

This mapping is reversible. Therefore, etc.  $\square$

**Example 17.4.1** Let  $n = 3$ , and  $k = 5$ . Also, define:

$$M = \{\{1, 1, 2, 2, 3\}\} \quad (17.4.5)$$

Then:

$$A = \{1, 2, 4, 5, 7\} \subseteq [7] \quad (17.4.6)$$

Multi-sets can be seen as binary sequences (Stars and bars). For example, let  $M = \{3, 3, 4, 7, 7\}$ . We can write this as  $|| * | * ||| * *$ . This helps count out the repetitions of various elements in the multi-set.  $\binom{n}{k}$  can be seen as the number of functions that map  $k$  elements to 0 and  $n - k$  elements to 1. We can generalize to functions  $[n] \rightarrow [m]$ . Let  $k_1, k_2, \dots, k_m$  be such that:

$$\sum_{i=1}^m k_i = n \quad (17.4.7)$$

And such that, for all  $i, k_i \geq 0$ . Then  $\binom{n}{k_1, \dots, k_m}$  is the number of ways to map  $[n] \rightarrow [m]$  such that  $k_i$  elements map to  $i$ , where:

$$\binom{n}{k_1, k_2, \dots, k_m} = \binom{n}{k_1} \binom{n-k_1}{k_2} \binom{n-k_1-k_2}{k_3} \cdots \binom{k_m}{k_m} \quad (17.4.8)$$

$$= \frac{n!}{k_1! k_2! \cdots k_m!} \quad (17.4.9)$$

### Theorem 17.4.2: Multinomial Theorem

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{k_1 + \cdots + k_m = n} \binom{n}{k_1, \dots, k_m} x_1^{k_1} \cdots x_m^{k_m} \quad (17.4.10)$$



#### 17.4.1 Lattice Paths

Let  $\mathbb{Z}^d$  be an integer lattice of dimension  $d$ , where  $d \in \mathbb{N}$  and  $d \geq 1$ .

##### Definition 17.4.1: Lattice Path

A lattice path in  $\mathbb{Z}^d$  with  $k$  steps in  $S \subseteq \mathbb{Z}^d$  is a subset  $L \subseteq \mathbb{Z}^d$  such that  $L = \{v_1, \dots, v_k\}$  such that, for all  $i = 1, 2, \dots, k-1$ ,  $v_{i+1} - v_i \in S$ . ■

**Example 17.4.2** If  $d = 2$ ,  $S = \{(0,1), (1,0)\}$ , then there are 6 paths from  $(0,0)$  to  $(2,2)$ .

**Theorem 17.4.3.** if  $v = (a_1, \dots, a_d) \in \mathbb{Z}^d$  and if  $e_i$  is the  $i^{th}$  unit vector in  $\mathbb{Z}^d$ , then the number of lattice paths in  $\mathbb{Z}^d$  from the origin to  $v$  with steps in  $\{e_i : i \in \}$  is given by the multinomial coefficient  $\binom{\|v\|_1}{a_1, \dots, a_d}$ .

*Proof.* For let  $v_0, \dots, v_k$  be a path. Then  $v_1 - v_0, v_2 - v_1, \dots, v_k - v_{k-1}$  consist of the  $e_i$ . Thus there are  $a_1$   $e_1$ 's,  $a_2$   $e_2$ 's, and so on. The total number is thus the multinomial coefficient. □

**Theorem 17.4.4.** The number of lattice paths from  $(0,0)$  to  $(n,m)$  with steps in  $\{(0,1), (1,0)\}$  is  $\binom{n+m}{n}$ .

#### 17.4.2 The Involution Principle

**Theorem 17.4.5.** The number of lattice paths from  $(i,j)$  to  $(m,n)$  using steps  $(1,0)$  and  $(0,1)$  is  $\binom{m-i+n-j}{m-i}$ .

Given a set  $S$  and a partition  $S = S^+ \cup S^-$  into a negative part  $S^-$  and a positive part  $S^+$ , then  $S$  is called a signed set. We are interested in computing  $\text{curl}(S^+) - \text{curl}(S^-)$ .

---

**Definition 17.4.2: Sign Reversing Involution**


---

A sign reversing involution is an involution  $\psi : S \rightarrow S$  such that for all  $x \in S$  such that  $\psi(x) \neq x$ , then  $\psi(x) \in S^+$  for all  $x \in S^-$  and  $\psi(x) \in S^-$  for all  $x \in S^+$ . ■

**Theorem 17.4.6.** *If  $\psi$  is a sign reversing involution, if  $F^+$  are the fixed points of  $\psi$  in  $S^+$ , and  $F^-$  are the fixed points are  $\psi$  in  $S^-$ , then:*

$$\text{curl}(S^+ \setminus F^+) = \text{curl}(S^- \setminus F^-) \quad (17.4.11)$$

$$\text{curl}(S^+) - \text{curl}(S^-) = \text{curl}(F^+) - \text{curl}(F^-) \quad (17.4.12)$$

Suppose we are given a set  $X$  and we want to compute  $\text{curl}(X)$ . Embed  $X$  into a signed set  $S = S^+ \cup S^-$  such that for all  $x \in X$  there is a corresponding  $s \in S^+$ .

---

**Definition 17.4.3: Catalan Path**


---

A Catalan path is a lattice path from  $(0, 0)$  to  $(n, n)$  using steps  $(0, 1)$  and  $(1, 0)$  such that the path never crosses the line  $x = y$ . ■

We are interested in counting the number of Catalan paths for a given  $n \in \mathbb{N}$ . The first few numbers are  $1, 2, 5, 14, 42, \dots$  and occur frequently in mathematics. Let  $S^+$  be the set of paths from  $(1, 0)$  to  $(n + 1, n)$  and  $S^-$  be the set of paths from  $(0, 1)$  to  $(n + 1, n)$ . Using the previous theorem:

$$\text{curl}(S^+) = \binom{2n}{n} \quad (17.4.13)$$

$$\text{curl}(S^-) = \binom{2n}{n-1} \quad (17.4.14)$$

Now we need to embed the Catalan paths into  $S^+$ . The embedding comes from shifting the graphs 1 unit to the right. Note that the image never touches the line  $x = y$ . Define a sign reversing involution  $\psi : S \rightarrow S$  by letting  $P$  be any path in  $S$  that does not touch  $x = y$ , and defining  $\psi(P) = P$ . If  $P$  touches or crosses  $x = y$ , let  $p_0$  be the first such crossing. Let  $\psi(P)$  be the path from  $(1, 0)$  (Respectively, from  $(0, 1)$ ), such that the points from  $(0, 0)$  to  $p_0$  are reflected, and the points from  $p$  to  $(n + 1, n)$  stay the same. Now  $F^-$  is empty, since given any path from  $(1, 0)$  to  $(n + 1, n)$ , it must cross the line  $y = x$ . Thus there are no fixed points in  $S^-$ . But then:

$$\text{curl}(F^+) = \text{curl}(S^+) - \text{curl}(S^-) \quad (17.4.15)$$

But the Catalan number  $C_n$  is equal to the size of  $F^+$ , and thus we have:

$$C_n = \binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1} \binom{2n}{n} \quad (17.4.16)$$

### 17.4.3 Diagonal Lattice Paths

#### Definition 17.4.4: Diagonal Lattice Paths

A diagonal lattice path is a lattice path with steps  $(1, 1)$  and  $(1, -1)$ . ■

**Example 17.4.3** Consider all diagonal paths from  $(0, 0)$  to  $(4, 0)$ . Since any step increasing the  $x$  coordinate by 1, there must be 4 steps in the lattice path. But since the path must end at 0, there must be an equal number of steps that go up as there are steps that go down. So, we must have two up steps and two down steps. The total number of diagonal lattice paths is thus  $\binom{4}{2} = 6$ . In general, the total number of lattice paths from  $(0, 0)$  to  $(2n, 0)$  is  $\binom{2n}{n}$ .

These diagonal lattice paths can be seen as binary words with  $d = (1, -1)$  and  $(u = 1, 1)$  such that the number of occurrences of  $d$  is equal to the number of occurrences of  $u$ . We can establish a correspondence between diagonal lattice paths and Catalan paths by considering as the bijection a reflection about the  $x = y$  axis, and then a rotation by  $45^\circ$ .

#### Definition 17.4.5: Dyck Paths

A Dyck is a diagonal lattice path that never goes below its starting point. ■

## 17.5 q-Analogues

In combinatorics, a  $q$  analogue of a counting function, such as  $n!$ , is typically a polynomial in  $q$  which evaluates to the function if we set  $q = 1$ , and if not a polynomial we take the limit as  $q \rightarrow 1$ . We want the  $q$ -analogue to preserve the same recurrence properties that the counting function has.

**Example 17.5.1** A  $q$ -Analogue of a real number  $x \in \mathbb{R}$  could be:

$$[x]_q = \frac{1 - q^x}{1 - q} \quad (17.5.1)$$

Taking the limit as  $q \rightarrow 1$ , we see that this expression evaluates to  $x$  by using L'Hôpital's Rule. If  $x = n \in \mathbb{N}$ , then:

$$\frac{1 - q^n}{1 - q} = 1 + q + \cdots + q^{n-1} \quad (17.5.2)$$

This allows us to construct a  $q$ -Analogue of  $n!$ :

$$[n]_q! = [1]_q [2]_q \cdots [n]_q \quad (17.5.3)$$

This can be used to put statistics on sets.

**Definition 17.5.1: Statistic on a Finite Set**

A statistic on a finite set  $S$  is a function  $f : S \rightarrow \mathbb{N}_0$  ■

Let  $S_n$  denote the symmetric group, which is the set of permutations of  $1, 2, \dots, n$  under the operation of composition. Then:

$$\text{curl}(S_n) = n! \quad (17.5.4)$$

**Definition 17.5.2: Inversion of a Word**

An inversion of a word  $\sigma$  is a pair  $(i, j)$ , where  $1 \leq i < j \leq n$ , where  $\sigma_i > \sigma_j$ . ■

**Example 17.5.2** Let  $\sigma = (132)(45)(6)(7)$ . Then  $(1, 3)$  is an inversion, since  $\sigma_1 = 3 > \sigma_3 = 2$ .

**Definition 17.5.3: Inversion Statistic**

The inversion statistic on  $S_n$  is the number of inversions of  $\sigma \in S_n$ . ■

**Theorem 17.5.1.** *If  $S_n$  is the permutation group, then:*

$$\sum_{\sigma \in S_n} q^{\text{inv}(\sigma)} = [n]_q! \quad (17.5.5)$$

## 17.6 Lecture 6

Last week we introduced q-Analogs, and proved the following identities:

$$\sum_{\sigma \in S_n} q^{\text{inv}(\sigma)} = \sum_{\sigma \in S_n} q^{\text{maj}(\sigma)} = [n]_q! \quad (17.6.1)$$

Where  $\text{inv}(\sigma)$  is the number of inversions, and  $\text{maj}(\sigma)$  is the number of descents. We now want a q-Analog of  $\binom{n}{k}$ . Recall the inversion tables:

$$\mathcal{I}_n = \{(a_1, \dots, a_n) : 0 \leq a_i \leq i\} \quad (17.6.2)$$

We can write this as:

$$\mathcal{I}_n = \{0\} \times \{0, 1\} \times \{0, 1, 2\} \times \cdots \times \{0, 1, 2, \dots, n-1\} \quad (17.6.3)$$

From this we obtain:

$$\text{curl}(\mathcal{I}_n) = n! \quad (17.6.4)$$

We define the function  $\Psi_1 : \mathcal{I}_n \rightarrow S_n$  by mapping:

$$\Psi_1(a_1, \dots, a_n) = \sigma \quad (17.6.5)$$

Where  $\sigma$  is the permutation with inversions  $a_1, \dots, a_n$ , and  $a_i$  is the number inversions created by  $i$ . We also define  $\Psi_2 : \mathcal{I}_n \rightarrow S_n$  and re-interpreted  $a_i$  to be the contribution of  $i$  to the major index. Then  $\Psi = \Psi_2 \circ \Psi_1^{-1}$  is a bijection from  $S_n$  to itself. We now want to find a good q-Analog for  $\binom{n}{k}$  that would satisfy similar properties as the binomial coefficient. One nice property is Pascal's Identity:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \quad (17.6.6)$$

Perhaps the obvious choice is to choose:

$$\binom{n}{k}_q = \frac{[n]_q!}{[k]_q![n-k]_q!} \quad (17.6.7)$$

These are called the Gaussian polynomials, and it seems surprising that these are polynomials in the first place, since it appears to be a rational function. However, we can see just by plugging in that:

$$\binom{n+1}{k}_q \neq \binom{n}{k}_q + \binom{n}{k-1}_q \quad (17.6.8)$$

And thus this is not a good q-Analog for the binomial coefficients. Let:

$$\mathcal{R}(1^k 0^{n-k}) = \{\text{Set of binary words of length } n \text{ with } k \text{ 1's}\} \quad (17.6.9)$$

Then:

$$\text{curl}(\mathcal{R}(1^k 0^{n-k})) = \binom{n}{k} \quad (17.6.10)$$

This is equivalent to saying:

$$k!(n-k)! \text{curl}(\mathcal{R}(1^k 0^{n-k})) = n! \quad (17.6.11)$$

But the left-hand side of this equation is the cardinality of  $S_k \times S_{n-k} \times \mathcal{R}(1^k 0^{n-k})$ , and the right-hand side is the cardinality of  $S_n$ . We need to define a function:

$$f : S_k \times S_{n-k} \times \mathcal{R}(1^k 0^{n-k}) \quad (17.6.12)$$

Add  $n - k$  to the numbers in the permutation  $S_k$ . For example consider:

$$(132, 14523, 10011000) \mapsto (687, 14523, 10011000) \quad (17.6.13)$$

Send the left-most number in order from left to right to the 1's in the binary word (The third entry). Send the second entry to the 0's in the binary word. So, finally we have:

$$(132, 14523, 10011000) \mapsto (61487523) \quad (17.6.14)$$

We now want to show that:

$$\sum_{r \in \mathcal{R}(1^k 0^{n-k})} q^{inv(r)} = \binom{n}{k}_q = \frac{[n]_q!}{[k]_q! [n-k]_q!} \quad (17.6.15)$$

We can do this in a similar manner as before. We need a function  $f$  from  $S_{n-k} \times S_k \times \mathcal{R}(1^k 0^{n-k})$  that is bijective. Let's use the one defined previously. We now need to show that  $f$  preserves inversions.

**Theorem 17.6.1.**

$$\binom{n+1}{k}_q = q^k \binom{n}{k}_q + \binom{n}{k-1}_q \quad (17.6.16)$$

**Theorem 17.6.2: Foata's Theorem**

---

$$\sum_{r \in \mathcal{R}(1^k 0^{n-k})} q^{maj(r)} = \binom{n}{k}_q \quad (17.6.17)$$



### 17.6.1 Lattice Paths and Gaussian Polynomials

#### Definition 17.6.1: Partitions of Integers

---

A partition of  $\mathbb{Z}_n$ ,  $n \in \mathbb{N}$ , is a weakly decreasing sequence  $\lambda = (\lambda_1, \dots, \lambda_\ell)$ , such that:

$$|\lambda| = \sum_{k=1}^{\ell} \lambda_k = n \quad (17.6.18)$$

$|\lambda|$  is called the weight of  $\lambda$  and  $\lambda_i$  are called the parts of  $\lambda$ . The length of  $\lambda$  is the number of non-zero parts.  $\blacksquare$

A young diagram is a graphical representation of a partition  $\lambda = (\lambda_1, \dots, \lambda_\ell)$ . The conjugate of a partition  $\lambda$  is obtained by transposing the Young diagram of  $\lambda$ . For example:

$$(3, 3, 1) \mapsto (3, 2, 2) \quad (17.6.19)$$

We denote the conjugate by  $\lambda'$ . Recall that  $\binom{n+m}{n}$  is the number of lattice paths from  $(0, 0)$  to  $(n, m)$  using steps  $(1, 0)$  and  $(0, 1)$ .

**Theorem 17.6.3.** *There exists a bijection between the set of lattice paths from  $(0, 0)$  to  $(m, n)$  and the set of partitions of such that  $\lambda_1 \leq m$  and  $\ell(\lambda) \leq n$ .*

If  $p(m, n)$  is the number of partitions that fit in the  $m \times n$  rectangle, that is  $\ell(\lambda) \leq N$  and  $\lambda_1 \leq m$ , then  $p(m, n) = \binom{n+m}{n}$ . A statistic on partitions is given by the weight of  $\lambda$ ,  $|\lambda|$ .

**Theorem 17.6.4.** *For  $m, n \in \mathbb{N}$ :*

$$\binom{n}{m}_q = \sum_{\lambda \subseteq [m^n]} q^{|\lambda|} \quad (17.6.20)$$

*Proof.* We can show this by proving that the sum satisfies the same recurrence relation and initial conditions as the q binomial.  $\square$

## 17.7 Lecture 7 (I Think)

We're currently discussing q-analogues. We want to extend the q-analogue defined for the factorial function to the binomial coefficient. The Gaussian polynomials are one such attempt at this:

$$\binom{n}{k}_q = \frac{[n]_q!}{[k]_q![n-k]_q!} \quad (17.7.1)$$

Another such attempt was to sum over all binary words of length  $n$  with  $k$  one's, and obtain:

$$\binom{n}{k}_q = \sum_{r \in \mathcal{R}(1^k, 0^{n-k})} q^{inv(r)} \quad (17.7.2)$$

Using this definition, we obtained the following equation:

$$\binom{n+1}{k}_q = q^k \binom{n}{k}_q + \binom{n}{k-1}_q \quad (17.7.3)$$

Then we discussed lattice paths  $L(m, n)$ , which are paths from  $(0, 0)$  to  $(m, n)$  using steps in  $(1, 0)$  and  $(0, 1)$ . Next we discussed partitions of numbers.

$n$	Partitions
0	$\emptyset$
1	$(1)$
2	$(2), (1, 1)$
3	$(3), (2, 1), (1, 1, 1)$

Table 17.1: Caption

Given such a partition, we assign the weight to be:

$$|\lambda| = \sum_{k=1}^{\ell} \lambda_k \quad (17.7.4)$$

Then we define:

$$P(m, n) = \{(\lambda_1, \dots, \lambda_\ell : \lambda_1 \leq m, \ell(\lambda) \leq n\} \quad (17.7.5)$$

We showed that there is a bijection between  $L(m, n)$  and  $P(m, n)$ . Thus, we have:

$$\text{curl } (P(n, m)) = \binom{m+n}{m} \quad (17.7.6)$$

**Theorem 17.7.1.** *If  $m, n \in \mathbb{N}$ , then:*

$$\binom{m+n}{m}_q = \sum_{\lambda \in P(m, n)} q^{|\lambda|} \quad (17.7.7)$$

*Proof.* The strategy of the proof is to show that this sum satisfies the same initial conditions and the same recurrence as the original definition. We have that  $p(m, 0) = 1 = q^0$  since there is only the empty partition in the rectangle

$m \times 0$ , and similarly  $p(0, n) = 1 = q^0$  since there is only the empty partition in the rectangle  $0 \times n$ . Moreover,  $p(m, m) = 1 = q^0$  since:

$$\binom{m}{m}_q = \frac{[m]_q}{[0]_q [m]_q} = \frac{[m]_q}{[m]_q} = 1 \quad (17.7.8)$$

We now must show that the recurrence relation is satisfied. We want:

$$p(m, n) = q^m p(n-1, m) + p(n, m-1) \quad (17.7.9)$$

We have that:

$$p(m, n) = \sum_{\lambda \in P(m, n)} q^{|\lambda|} \quad (17.7.10)$$

$$= \sum_{\lambda_1=m} q^{|\lambda|} + \sum_{\lambda_1 < m} q^{|\lambda|} \quad (17.7.11)$$

$$= q^m \sum_{\lambda \in P(m, n-1)} q^{|\lambda|} + \sum_{\lambda \in P(m-1, n)} q^{|\lambda|} \quad (17.7.12)$$

This completes the proof.  $\square$

### Definition 17.7.1: $x$ Factorization

Let  $w \in X^*$  be a word in the alphabet  $X$ . Let  $x \in X$  and suppose  $w = vy$ , where  $v$  is a word and  $y$  is a letter in  $X$ . That is,  $y$  is the last letter of  $w$ . Then the factorization is:

$$w = v_1 y_1 \dots v_k y_k \quad (17.7.13)$$

Where, if  $y > x$  ( $X$  is totally ordered):

$$y_i > x, \quad y_i \in X \quad (17.7.14)$$

$$v_i \in L_x^* \quad (17.7.15)$$

Where:

$$L_x = \{a : a \leq x\} \quad (17.7.16)$$

If  $y \leq x$ , then:

$$y_i \leq x \quad y_i \in X \quad (17.7.17)$$

and:

$$v_i \in R_x^* \quad (17.7.18)$$

Where:

$$R_x = \{a : a > x\} \quad (17.7.19)$$



**Example 17.7.1** Let  $x = 3$  and let:

$$w = 125312641237 \quad (17.7.20)$$

Then  $y = 7$ , and thus  $y > x$ . Then we can write:

$$w = |12|5|312|6||4|123|7 \quad (17.7.21)$$

We allow for empty words. As another example, consider:

$$w = 135712136412 \quad (17.7.22)$$

Then  $y = 2$ , and thus  $y < 2$ . We obtain:

$$w = 1|3|57|2|1|3|641|2 \quad (17.7.23)$$

**Theorem 17.7.2.** *The following is true:*

$$\sum_{r \in \mathcal{R}(1^k, 0^{n-k})} q^{maj(r)} = \binom{n}{k}_q \quad (17.7.24)$$

*Proof.* We want to define a bijection  $\varphi$  from  $\mathcal{R}(1^k, 0^{n-k})$  to itself such that, for any  $r$ , we have  $maj(r) = inv(\varphi(r))$ . Let  $X \subseteq \mathbb{N}$ . Let  $X^*$  be the set of all words over  $X$ . For example if  $X = \{0, 1\}$ , then  $X^*$  is the set of all binary words. Define  $\varphi : X^* \rightarrow X^*$  be such that  $maj(w) = inv(\varphi(w))$  for any  $w \in X^*$ . Note that inversions and descents are defined in the same way as for permutations. This is why we required the set to be totally ordered,  $\mathbb{N}$  in our case. Define  $\gamma_x : X^* \rightarrow X^*$  by:

$$\gamma_x(w) = \begin{cases} \emptyset, & w = \emptyset \\ y_1 v_1 \dots y_k v_k, & w = v_1 y_1 \dots v_k y_k \end{cases} \quad (17.7.25)$$

We define  $\varphi$  as follows:

$$\varphi(w) = \begin{cases} \emptyset, & w = \emptyset \\ w, & w \in X \\ \gamma_x(\varphi(v)), & w = vx, x \in X \end{cases} \quad (17.7.26)$$

This is a recursive definition. For example, let:

$$w = 121314 \quad (17.7.27)$$

Then  $\varphi(1) = 1$ , and thus  $\varphi(12) = \gamma_2(\varphi(1))2 = 12$ . The first interesting case is with three elements. We have:

$$\varphi(121) = \gamma_1(\varphi(12))1 = \gamma_1(12)1 = 211 \quad (17.7.28)$$

Note that  $\text{inv}(211) = 2$  and  $\text{maj}(121) = 2$ . This function works since, if  $w \in X^*$ , then let  $r_x$  be the number of letters in  $w$  that are greater than  $x$ , and let  $\ell_x$  be the number of letters in  $w$  that are less than or equal to  $x$ . Let  $v \in X^*$  and  $x \in X$ . Then:

$$\text{inv}(vx) = \text{inv}(v) + r_x(v) \quad (17.7.29)$$

Also, when the last letter of  $v$  is less than or equal to  $x$ , we have:

$$\text{inv}(\gamma_x(v)) = \text{inv}(v) - r_x(v) \quad (17.7.30)$$

And otherwise we have:

$$\text{inv}(\gamma_x(v)) = \text{inv}(v) + \ell_x(v) \quad (17.7.31)$$

Moreover, if the last letter  $v$  is less than or equal to  $x$ , then:

$$\text{maj}(vx) = \text{maj}(v) \quad (17.7.32)$$

And otherwise:

$$\text{maj}(vx) + |v| \quad (17.7.33)$$

□

## 17.8 Lecture 9

As a summary, we were studying q-Analogues. We have shown:

$$\sum_{\sigma \in S_n} q^{\text{inv}(\sigma)} = \sum_{\sigma \in S_n} q^{\text{maj}(\sigma)} = [n]_q! \quad (17.8.1)$$

Also:

$$\sum_{w \in \mathcal{R}(1^k, 0^{n-k})} q^{\text{inv}(w)} = \sum_{w \in \mathcal{R}(1^k, 0^{n-k})} q^{\text{maj}(w)} = \binom{n}{k}_q \quad (17.8.2)$$

### Theorem 17.8.1: q-Binomial Theorem

The following is true:

$$\prod_{i=1}^n (x + q^i y) = \sum q^{\binom{n-k+1}{2}} \binom{n}{k}_q x^k y^{n-k} \quad (17.8.3)$$

■

For all  $n, k \in \mathbb{N}$ , we have:

$$\binom{n+k}{k}_q = \sum_{\lambda \in P(n,k)} q^{|\lambda|} \quad (17.8.4)$$

We can use rising factorials to define  $\binom{x}{k}_q$  for all  $x \in \mathbb{R}$  and  $k \in \mathbb{N}$ . That is:

$$\binom{x}{k}_q = \frac{(1-q^{x-k+1})(1-q^{x-k+1})\dots}{(1-q)(1-q^2)\dots} \quad (17.8.5)$$

### 17.8.1 q-Catalan Analogue

Recall that  $C_n$  is the number of lattice paths from  $(0,0)$  to  $(n,n)$  that do not go below or above the line  $x = y$  in the plane. We showed earlier that:

$$C_n = \frac{1}{n+1} \binom{2n}{n} \quad (17.8.6)$$

Recall that  $L(m,n)$  is the number of lattice paths from  $(0,0)$  to  $(m,n)$ . We define  $L^+(m,n)$  to be the set of lattice paths from  $(0,0)$  to  $(m,n)$  that do not go below the line  $y = \frac{n}{m}x$ . In particular,  $L^+(n,n) = C_n$ . The Catalan numbers satisfy the following recurrence:

$$C_n = \sum_{k=1}^n C_{k-1} C_{n-k} \quad (17.8.7)$$

Recall that  $\omega : L(m,n) \rightarrow \mathcal{R}(0^n, 1^m)$ , where  $\omega$  maps  $N \rightarrow 0$  and  $E$  to 1, north and east. Let  $\mathcal{R}^+(0^n 1^n)$  be the elements of  $\mathcal{R}(1^n 0^n)$  that correspond to Catalan paths. From a previous observation, the words corresponding to the Catalan paths are characterized by always having more 0's than 1's for any initial word.

### Theorem 17.8.2: MacMahon's Theorem

---

The following is true:

$$\sum_{p \in L^+(n,n)} q^{\text{maj}(w(p))} = \frac{1}{[n+1]_q} \binom{2n}{n}_q \quad (17.8.8)$$



*Proof.* Define the following:

$$\mathcal{R}^-(0^n 1^n) = \mathcal{R}(0^n 1^n) - \mathcal{R}^+(0^n 1^n) \quad (17.8.9)$$

$$L^-(n,n) = L(n,n) - L^+(n,n) \quad (17.8.10)$$

Given a path  $P$  in  $L^-(n, n)$ , let  $A$  be the lattice point with the smallest  $x$  coordinate among all the lattice points  $(i, j)$  with  $i - j$  maximized, whose distance from the  $x = y$  line in the south east direction is maximized. Let  $B$  be the lattice point just before  $A$ . Notice that the step  $B \rightarrow A$  must be an east step. Create a new path as follows. Change the east step to a north step, and then take the remaining path from  $A$  to  $(n, n)$  and shift it up one and two the left one. This path ends on  $(n - 1, n + 1)$ . Let  $\varphi$  denote the new path. Then:

$$\text{maj}(w(\varphi(p))) = \text{maj}(w(p)) - 1 \quad (17.8.11)$$

For suppose  $B \neq (0, 0)$ . The the step that goes to  $B$  must be an east step. For if not, then  $A$  does not have the smallest  $x$  coordinate with maximal distance to the line  $y = x$ . If  $B = (0, 0)$ , then the first position goes away. Moreover, the algorithm is reversible. For let  $P'$  be a lattice path from  $(0, 0)$  to  $(n, m)$ , and let  $A'$  be the point with maximal  $x$  coordinate such that  $i - j$  is maximized. This point corresponds to the  $B$  in the previous path. Therefore:

$$\sum_{w \in \mathcal{R}(1^n 0^n)} q^{\text{maj}(w)} = \sum_{w \in \mathcal{R}(1^{n+1} 0^{n-1})} q^{\text{maj}(w)+1} = q \binom{2n}{n+1} \quad (17.8.12)$$

□

There is another q-analogue for  $C_n$  due to Carlitz and Riordan. Let  $p \in L^+(n, n)$  and define  $a_i(p)$  to be the number of complete squares between the path and the  $x = y$  line in row  $i$ . The number  $a_i(p)$  is called the length of the  $i^{th}$  row of  $p$  and the sequence  $(a_1(p), \dots, a_n(p))$  is called the co-area vector of  $p$ . The co-area statistic on  $p$  is defined as:

$$\text{Coarea}(p) = \sum_{i=1}^n a_i(p) \quad (17.8.13)$$

### Theorem 17.8.3: Carlitz-Riordan Theorem

The following is true:

$$C_n(q) = \sum_{p \in L^+(n, n)} q^{\text{Coarea}(p)} \quad (17.8.14)$$

■

Using the Carlitz-Riordan theorem, we can show the following result.

**Theorem 17.8.4.** *The following is true:*

$$C_n(q) = \sum_{k=1}^n q^{k-1} C_{k-1}(q) C_{n-k}(q) \quad (17.8.15)$$

If we set  $x \mapsto q^i x$ , and  $y \mapsto 1$  in the q-binomial theorem, then we obtain:

$$(-x; q) = \prod_{k=0}^{n-1} (q^i x + 1) = \sum_{k=0}^n q^{\binom{k}{2}} \binom{n}{k}_q x^k \quad (17.8.16)$$

Using this q-binomial, we get the following.

**Theorem 17.8.5.** *If  $h, n, m \in \mathbb{N}$ , then:*

$$\sum_{k=0}^n q^{(n-k)(h-k)} \binom{n}{k}_q \binom{m}{n-k}_q = \binom{m+n}{h}_q \quad (17.8.17)$$

## 17.9 Generating Functions

Given a q-analogue and a set  $S$ , we can then define a statistic,  $\lambda$ . We then have:

$$\sum_{s \in S} q^{\lambda(s)} = \sum_{i=0}^n a_i q^i \quad (17.9.1)$$

Where  $a_i$  is the number of elements in  $S$  with statistic value  $i$ . Thus we can think of a statistic :  $S \rightarrow \mathbb{N}$ , called the value function.

**Example 17.9.1** Let  $n \in \mathbb{N}$  and consider  $S = \mathcal{P}(\mathbb{Z}_n)$ . One easy statistic we can place on  $S$  is the cardinality function. That is, we define  $f : S \rightarrow \mathbb{N}$  by:

$$f(\omega) = \text{curl}(\omega) \quad \omega \in \mathcal{P}(\mathbb{Z}_n) \quad (17.9.2)$$

Let's compute this a different way. Given a set  $A \subseteq \mathcal{P}(\mathbb{Z}_n)$ , either  $1 \in A$  or  $1 \notin A$ . Similarly, either  $2 \in A$  or  $2 \notin A$ . For all  $k \in \mathbb{Z}$ , either  $k \in A$  or  $k \notin A$ . Thus, we have:

$$(q^1 + q^0) \cdots (q^1 + q^0) = \prod_{k=1}^n (q^1 + q^0) = [2]_q^n = \sum_{k=0}^n \binom{n}{k} q^k \quad (17.9.3)$$

Next we want to consider  $S$  being infinite. To get a generating function we require that  $a_i$  is equal to the number of elements in  $S$  with value  $i$  being finite. We obtain the following power series:

$$a_0 q^0 + a_1 q^1 + \cdots = \sum_{i=0}^{\infty} a_i q^i \quad (17.9.4)$$

Let  $\mathbb{C}[[q]]$  denote the ring of formal power series. This is a ring. For let:

$$A(q) = \sum_{i=0}^{\infty} a_i q^i \quad (17.9.5a)$$

$$B(q) = \sum_{i=0}^{\infty} b_i q^i \quad (17.9.5b)$$

Then the sum is well defined, and we have:

$$A(q) + B(q) = \sum_{i=0}^{\infty} (a_i + b_i)q^i \quad (17.9.6)$$

We can also define the product by using the convolution product, or Cauchy sums:

$$A(q)B(q) = \sum_{i=0}^{\infty} c_i q^i \quad (17.9.7)$$

Where:

$$c_k = \sum_{i=0}^k a_i b_{k-i} \quad (17.9.8)$$

Some properties of  $\mathbb{C}[[q]]$  is that it is a commutative ring. This is because we considered the coefficients to be over  $\mathbb{C}$ . Moreover, it is an integral domain. That is,  $\mathbb{C}[[q]]$  has no zero divisors. The units, or invertible elements, are formal power series such that  $a_0 \neq 0$ . Indeed, this is a necessary and sufficient condition for an element to be invertible. For example, consider:

$$A(q) = \sum_{i=0}^{\infty} q^i \quad (17.9.9)$$

This is a geometric sum, and we can show that for  $|q| < 1$ , this formal sum is a convergent sum and evaluates to  $(1 - q)^{-1}$ . However, for all formal sums, this formal power series has an inverse, and the inverse is indeed  $(1 - q)^{-1}$ . Ivan Niven has a nice article on  $\mathbb{C}[q]$  in the American Mathematical Monthly, 1969. This has applications in counting partitions of numbers. See Andrews Theory of Partitions. Let  $p(n)$  be the number of partitions on  $n$ . Then:

$$\sum_{n=0}^{\infty} p(n)q^n = 1 + q + 2q^2 + 3q^3 + 5q^4 + 7q^5 + 11q^6 + 15q^7 + \dots \quad (17.9.10)$$

The value function is thus the weight of the partition  $|\lambda| = \lambda_1 + \dots + \lambda_m$ .

### Theorem 17.9.1: Euler's Partition Theorem

If  $P$  is the set of partitions, then:

$$\sum_{\lambda \in P} q^{|\lambda|} = \prod_{k=1}^{\infty} \frac{1}{1 - q^k} \quad (17.9.11)$$

Give an arbitrary partition  $\lambda$ , consider the parts of size one.  $\lambda$  does not have a part of size 1 or  $\lambda$  has a part of size one, or  $\lambda$  has a part of size two, and so on. Now do the same for each of the parts of size  $k$ , in general.

---

**Theorem 17.9.2: Euler's Second Partition Theorem**


---

The number of partitions with  $n$  distinct parts is equal to the number of partitions of  $n$  with only odd parts. █

*Proof.* For:

$$\sum_{n=0}^{\infty} a_n q^n = \prod_{i=1}^{\infty} (1 + q^i) = \prod_{i=1}^{\infty} \frac{(q + q^i)(1 - q^i)}{1 - q^i} \quad (17.9.12)$$

We can then simplify: □

---

**Definition 17.9.1: Durfee Square**


---

The largest square that fits into a partition  $\lambda$  is called the Durfee square of  $\lambda$ . █

---

**Definition 17.9.2: Self-Conjugate Partition**


---

A self-conjugate partition is a partition  $\lambda$  such that  $\lambda = \lambda'$ , where  $\lambda'$  is the conjugate of  $\lambda$ . █

---

**Theorem 17.9.3: Euler's Other-Other Theorem**


---

The following is true:

$$\sum_{\lambda=\lambda'} q^{|\lambda|} = \prod_{n=0}^{\infty} (1 + q^{2n+1}) \quad (17.9.13)$$

Where  $\lambda = \lambda'$  are all of the self-conjugate partitions. █

The product is the generating function for partitions with distinct odd parts. Euler's theorem then says that the generating function for this set is the equal to the sum over all of the self-conjugate partitions.

## 17.10 Euler's Theorem

$$\sum_{n \in \mathbb{N}} p(n) q^n = \prod_{i=1}^{\infty} \frac{1}{1 - q^i} \quad (17.10.1)$$

Where  $p(n)$  is the number of partitions of  $n$ . We also define the Euler function, not to be confused with the Euler totient function, as:

$$\phi(q) = \prod_{i=1}^{\infty} (1 - q^i) \quad (17.10.2)$$

Let's try to simplify this:

$$\phi(q) = \prod_{i=1}^{\infty} (1 - q^i) = \sum_{k=0}^{\infty} b_k q^k \quad (17.10.3)$$

We want to find the  $b_k$ . Multiplying through by the original series from Euler's theorem, we get:

$$\left( \sum_{i=1}^{\infty} b_i q^i \right) \left( \sum_{j=0}^{\infty} p(j) q^j \right) = 1 \quad (17.10.4)$$

Using the convolution product, we have for all  $k \geq 1$ :

$$\sum_{j=0}^k b_j p(k-j) = 0 \quad (17.10.5)$$

This gives a recursion for  $p(k)$ . This gives us Euler's Pentagonal Number Theorem.

### Theorem 17.10.1: Euler's Pentagonal Number Theorem

---

$$\phi(q) = \prod_{i=1}^{\infty} (1 - q^i) \quad (17.10.6a)$$

$$= 1 + \sum_{m=1}^{\infty} (-1)^m \left( q^{\frac{m(3m-2)}{2}} + q^{\frac{m(3m+1)}{2}} \right) \quad (17.10.6b)$$

$$= \sum_{m=-\infty}^{\infty} (-1)^m q^{\frac{m(3m-1)}{2}} \quad (17.10.6c)$$



Let  $p_e(d, n)$  denote the number of partitions  $n$  with distinct parts and even length. Similarly, define  $p_o(d, x)$  for odd length.

### Theorem 17.10.2.

$$p_e(d, n) - p_o(d, n) = \begin{cases} (-1)^n, & n = \frac{m(3m\pm 1)}{2} \\ 0, & \text{Otherwise} \end{cases} \quad (17.10.7)$$

After some reflection, it should be easy to see that the first case is the inverse of the second case. Moreover, cases 1, 2, and 3 cover all partitions with distinct parts. For any partition in case three only one of  $a$  or  $b$  is true. The bijection

constructed proves Euler's Pentagonal Theorem. Now we can compute  $p$ :

$$p(0) = 1 \quad (17.10.8a)$$

$$p(1) = 1 \quad (17.10.8b)$$

$$p(2) = 2 \quad (17.10.8c)$$

$$p(3) = 3 \quad (17.10.8d)$$

$$p(4) = p(3) + p(2) = 5 \quad (17.10.8e)$$

$$p(5) = p(4) + p(3) - p(0) = 7 \quad (17.10.8f)$$

$$p(6) = p(5) + p(4) - p(1) = 11 \quad (17.10.8g)$$

$$p(7) = p(6) + p(5) - p(2) - p(0) = 15 \quad (17.10.8h)$$

And in general:

$$p(n) = p(n-1) + p(n-2) - p(n-5) - p(n-7) + \dots \quad (17.10.9)$$

Where we add and subtract over the pentagonal numbers. Gauss then turned to the question of computing powers of  $\phi(q)$ .

### Theorem 17.10.3: Gauss's Pentagonal Theorem

$$\phi(q)^3 = \prod_{i=1}^{\infty} (1 - q^i)^3 = \sum_{r=0}^{\infty} (-1)^r (2r+1) q^{\frac{r(r+1)}{2}} \quad (17.10.10)$$



This identity occurs in many different areas of mathematics, such as homological algebra, complex analysis, and hyperbolic geometry. The proof comes from Jacobi's Triple Product Identity.

### Theorem 17.10.4: Jacobi's Triple Product Identity

$$\sum_{n=-\infty}^{\infty} z^n q^{n^2} = \prod_{n=0}^{\infty} (1 - q^{2n+2})(1 + zq^{2n+1})(1 + z^{-1}q^{2n+1}) \quad (17.10.11)$$



The proof of Gauss' identity then uses Sylvester's bijection. To get this from Jacobi, do a shift of index starting from  $n = 0$  to  $n = 1$ . Differentiate both sides with respect to  $q$ , and then put  $z = -q$ . Finally, map  $q^2$  to  $q$ .

Felix Klein computed  $\phi(q)^8$ . In the theory of modular forms there is something called the  $\tau$  function, due to Ramanujan. This has the property that:

$$\sum_{n=1}^{\infty} \tau(n) q^{n-1} = \phi(q)^{24} = \prod_{m=1}^{\infty} (1 - q^m)^{24} \quad (17.10.12)$$

Freeman Dyson also had some contributions to this subject and came up with nice formula for  $\phi(q)^d$  when  $d = 3, 8, 10, 14, 15, 21, 24, 26, 28, 35, 36, \dots$ . With the exception of 26, these are the dimensions of the Lie algebras. Ian McDonald, working on the same problem, saw this as well. He came up with the following:

$$\phi(q)^{n^2-1} = \sum \varepsilon(k_1, \dots, k_n) \prod_{i=1}^n \binom{k_i}{n-i} q^{k_n} \quad (17.10.13)$$

Where this is summed over all tuples  $(k_1, \dots, k_n)$  of non-negative integers such that:

$$\sum_{i=1}^n k_i^2 = \sum_{i=1}^k k_i + \sum_{i=1}^{n-1} k_i k_{i+1} + k_n k_1 \quad (17.10.14)$$

And where  $\varepsilon(k_1, \dots, k_m) = \pm 1$ .

## 17.11 Generating Function for Multisets

Recall that a multiset is a collection with repetition. For example:

$$A = \{\{1, 1, 1, 3, 3, 4\}\} \quad (17.11.1)$$

This is different from the set  $B = \{1, 1, 1, 3, 3, 4\}$ , since sets cannot account for repetition. That is,  $B$  can be reduced down to  $B = \{1, 3, 4\}$ . Note that partitions are multisets and we have shown that  $\binom{k+(n-1)}{k}$  is the number of multi-sets of size  $k$  chosen from the set  $[n]$ . We want to compute the generating function for the number of multi-sets of size  $k$ :

$$f(M) = \sum_M q^{|M|} \quad (17.11.2)$$

Where  $M$  is a multi-set of elements in  $[n]$ , and  $|M|$  denotes the number of elements in  $M$ , with repetitions included. If  $M$  is an arbitrary multi-set, then either  $1 \neq M$ , or  $1 \in M$ , or  $1, 1 \in M$ , and so on. So, in general, we get:

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \quad (17.11.3)$$

In general:

$$\sum_M q^{|M|} = \frac{1}{(1-q)^n} = \sum_{k=0}^{\infty} \binom{n-1+k}{k} q^k \quad (17.11.4)$$

From the binomial theorem, we get:

$$(1+q)^n = \sum_{k=0}^n \binom{n}{k} q^k \quad (17.11.5)$$

And thus, we can define:

$$\binom{-n}{k} = \binom{n-1+k}{k} (-1)^k \quad (17.11.6)$$

Extending the binomial coefficient to all  $n \in \mathbb{Z}$ . Next, recall the q-binomial theorem.

$$\sum_{k=0}^n \binom{n}{k}_q q^{\binom{k}{2}} x^k = \prod_{k=1}^{n-1} (1 + q^k x) \quad (17.11.7)$$

And also:

$$\sum_{k=0}^{\infty} \binom{n+k}{k}_q x^k = \prod_{i=0}^n (1 - q^i x)^{-1} \quad (17.11.8)$$

See Stanley and Macdonald.

## 17.12 Symmetric Functions

### 17.12.1 Symmetric Polynomials

Let  $x_1, \dots, x_N$  commute, and let  $f \in \mathbb{C}[x_1, \dots, x_N]$ . Then  $f$  is symmetric if, for all  $\sigma \in S_N$ , then:

$$f(x_1, \dots, x_N) = f(x_{\sigma(1)}, \dots, x_{\sigma(N)}) \quad (17.12.1)$$

Where  $S_N$  is the symmetric group, and  $\sigma$  is any permutation. It is convenient to work with infinitely many variables. We impose the requirements that there are countable many variables, so that we may list them, and that the commute. In this case we have power series instead of polynomials. We thus get sums of the form:

$$f = \sum_{\alpha} C_{\alpha} x^{\alpha} \quad (17.12.2)$$

Where  $\alpha$  is a sequence of non-negative integers. We require that the sum over  $\alpha$  be finite, and thus this implies that all of the monomials are of finite degree.  $f$  is said to be homogeneous if all of the monomials have the same degree. We require that  $f$  be invariant under any permutation  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ .

#### Definition 17.12.1: Monomial Symmetric Basis

---

A monomial symmetric basis is:

$$m_{\lambda} = m_{\lambda}(x) + \sum_{\alpha} x^{\alpha} \quad (17.12.3)$$

Where  $\alpha$  is a rearrangement of  $\lambda$ . ■

That is,  $m_\lambda$  is the sum of all monomials in  $x_i$  whose exponents are the parts of  $\lambda$ .

### Example 17.12.1

$$m_{1,1} = \sum_{i < j} x_i x_j = x_1 x_2 + x_1 x_3 + \cdots + x_2 x_3 + \dots \quad (17.12.4a)$$

$$m_{2,1,1}(x_1, x_2, x_3) = x_1^2 x_2 x_2 + x_1 x_2^2 x_3 + x_1 x_2 x_3^2 \quad (17.12.4b)$$

$$m_2(x) = \sum_{i=1}^{\infty} x_i^2 \quad (17.12.4c)$$

---

### Definition 17.12.2: Elementary Symmetric Function

---

The elementary symmetric function is defined as:

$$e_k = m_{1^k} = \sum_{i_1 < i_2 < \dots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k} \quad (17.12.5)$$

For  $k \in \mathbb{N}$ . ■

Note that:

$$\prod_{i=1}^{\infty} (1 + zx_i) = \sum_{n=0}^{\infty} e_n z^n \quad (17.12.6)$$

---

### Definition 17.12.3: Power Symmetric Function

---

The power symmetric function is defined as:

$$P_k = m_k = \sum_{i=1}^{\infty} x_i^k \quad (17.12.7)$$

For  $k \in \mathbb{N}$ . That is,  $k \geq 1$ . ■

---

### Definition 17.12.4: Complete Homogeneous Symmetric Function

---

The complete homogeneous symmetric function is defined as:

$$h_k = \sum_{\lambda+k} m_\lambda \quad (17.12.8)$$

That is, the sum of all monomials of degree  $k$ . ■

**Theorem 17.12.1.** *The generating function for the homogeneous symmetric function is:*

$$\prod_{i=1}^{\infty} \frac{1}{1-zx_i} = \sum_{n=0}^{\infty} h_n z^n \quad (17.12.9)$$

Where  $h_0$  is defined as  $h_0 = 1$ .

An endomorphism is a function such that  $w(fg) = w(f)w(g)$ ,  $w(f + g) = w(f) + w(g)$ , and  $w(cf) = cw(f)$ .

#### **Definition 17.12.5: $\omega$ Involution**

The  $\omega$  involution is the endomorphism defined by:

$$\omega(p_k) = (-1)^{k=1} P_k \quad (17.12.10)$$

And:

$$\omega(p_\lambda) = (-1)^{n-\ell(\lambda)} P_\lambda \quad (17.12.11)$$



### 17.12.2 Misc Problems

**Problem 17.12.1** Let  $S = \{1, 2, 3, 4, 5, 6, 7, 8\}$ .

1. How many subsets of  $S$  are there with exactly three elements?
2. How many subsets of  $S$  contain exactly one even number and two odd numbers?
3. How many subsets of  $S$  contain exactly three elements, at most one of which is even?
4. How many subsets of  $S$  are there with exactly three elements satisfying the condition that the subset contains the numbers 1 or 8 (Or both)?
5. How many subsets of  $S$  are there with exactly three elements which satisfy the property that two of its elements sum to 9?

*Solution.*    1.  $\binom{8}{3} = \frac{8!}{3!(8-3)!} = 56$

2.  $\binom{4}{1} \binom{4}{2} = 24$

3.  $\binom{4}{1} \binom{4}{2} + \binom{4}{0} \binom{4}{3} = 28$ .

4.  $\binom{7}{2}$  contain 1,  $\binom{7}{2}$  contain 8, and  $\binom{6}{1}$  contain both.  $\binom{7}{2} + \binom{7}{2} - \binom{6}{1} = 36$

5.  $8 + 1 = 7 + 2 = 6 + 3 = 5 + 4 = 9$ . 4 pairs with  $\binom{6}{1}$  subsets per pair. We have  $4 \cdot 6 = 24$ .

□

**Problem 17.12.2** A club has 20 members.

- In how many different ways can the club select a president, vice-president, and secretary?
- In how many different ways can a social committee with four members be elected?
- The club contains 12 men and 8 women. In how many different ways can a committee of four people be selected if the committee must have two men and two women?
- In how many different ways can a four person committee be selected from the club members if one person is designated as the leader of the committee?

*Solution*    1.  $P(20, 3) = \frac{20!}{(20-3)!} = 6840$

2.  $\binom{20}{4} = \frac{20!}{4!(20-4)!} = 4845$

3.  $\binom{12}{2} \binom{8}{2} = 1848$

4.  $\binom{19}{3} = 969$

**Problem 17.12.3** A woman has eight friends. Answer the following:

- In how many different ways can she invite four of her friends to dinner?
- Two of her friends dislike each other. If she invites one friend, she can't invite the other. How many ways can she invite her friends?
- Five are her friends are men and three are women. How many ways can she invite four of her friends if she wants two men and two women.
- The eight friends consist of two married couples and four single people. How many ways can she invite four friends under the condition that if she invites one spouse she must invite the other?

*Solution.*

1.  $\binom{8}{4} = 70$

2.  $2\binom{6}{3} + \binom{6}{4} = 55$

3.  $\binom{5}{2} \binom{3}{2} = 30$

4.  $\binom{4}{0} + \binom{4}{1} + \binom{4}{2} + \binom{4}{3} = 14$

□

# CHAPTER 18

---

## Discrete Structure

---

### 18.1 Combinatorics

If one event can happen  $n$  ways, and another independent event can happen  $m$  ways, then the total number of possibilities is  $nm$ . If there are 5 entrees and 3 sides at a restaurant, then there are 15 total possible meals.

**Definition 18.1.1** The factorial of a positive integer  $n$ , denoted  $n!$ , is  $n! = n \cdot (n - 1) \cdots 2 \cdot 1$ . We define  $0! = 1$ .

The factorial of a number  $n$  is the number of ways to permute  $n$  objects. The number of ways to permute  $k$  objects from a set of  $n$  is  $P(n, k) = n!/(n - k)!$ , and the number of ways to choose  $k$  objects from a set of  $n$  objects (Taking the order into account) is the binomial coefficient  $\binom{n}{k}$ . The number of permutations of  $n$  objects into groups  $n_1, n_2, \dots, n_N$ , where  $n_1 + n_2 + \dots + n_N = n$ , is  $n!/(n_1!n_2!\dots n_N!)$ . Stirling's Approximation says that, for large  $n$ , the factorial can be approximated as follows:

$$n! \approx \sqrt{2\pi n} n^n e^{-n} \quad (18.1.1)$$

**Example 18.1.1** What is the probability of rolling four 4's out six tosses of a six sided dice? The number of ways to choose two numbers that are not 4 is  $\binom{6}{2} = 15$ . The probability of an event with four 4's and two numbers that aren't 4 is  $(\frac{1}{6})^4(\frac{5}{6})^2$ . So the probability of rolling four 4's is  $15(\frac{1}{6})^4(\frac{5}{6})^2 = 0.008$ .

### 18.2 Exams

**Problem 18.2.1** Calculate the following:

1.  $\sum_{i=0}^2 \sum_{j=1}^3 (3i - j)$

$$2. \prod_{i=1}^n \frac{2i}{i+1}$$

$$3. \cup_{i=1}^n \{i, i+1\}$$

*Solution.* 1.  $\sum_{i=0}^2 \sum_{j=1}^3 (3i - j) = \sum_{i=0}^2 ((3i - 1) + (3i - 2) + (3i - 3)) = \sum_{i=0}^2 (9i - 6) = -6 + 3 + 12 = 9.$

2.  $\prod_{i=1}^n \frac{2i}{i+1} = \frac{2^n}{n+1}$ . We prove by induction. The base case is trivial.

Suppose it is true for  $n \in \mathbb{N}$ . Then  $\prod_{i=1}^{n+1} \frac{2i}{i+1} = \frac{2(n+1)}{(n+1)+1} \prod_{i=1}^n \frac{2i}{i+1} = \frac{2n}{(n+1)+1} \frac{2^n}{n+1} = \frac{2^{n+1}}{(n+1)+n}$ . Therefore  $\prod_{i=1}^{n+1} \frac{2i}{i+1} = \frac{2^{n+1}}{(n+1)+1}$ .

3.  $\cup_{i=1}^n A_i = \mathbb{Z}_{n+1}$ . We prove by induction. The base case is trivial. Suppose it is true for  $n \in \mathbb{N}$ . Then  $\cup_{i=1}^{n+1} A_i = (\cup_{i=1}^n A_i) \cup A_{n+1} \mathbb{Z}_{n+1} \cup \{n+1, n+2\} = \mathbb{Z}_{n+2}$ . Thus  $\cup_{i=1}^{n+1} A_i = \mathbb{Z}_{n+2}$ .

□

**Problem 18.2.2** 1. Compute the binary representation of 75.

2. Convert  $1100101_2$  to decimal.

*Solution.* 1.  $75 = 2^6 + 2^3 + 2^1 + 2^0 = 1001011_2$

2.  $1100101_2 = 2^0 + 2^2 + 2^5 + 2^6 = 1 + 4 + 32 + 64 = 101$

□

**Problem 18.2.3** A bit string is a sequence of numbers consisting of 0's and 1's. Answer the following:

1. How many strings of length 5 are there?
2. How many strings of length 6 have an even number of 1's?
3. How many strings of length 5 begin with 0 or end with 1 (Or both)?
4. How many strings of length 6 contain exactly three 1's?
5. How many strings of length 6 contain at least three 1's?
6. How many strings of length 6 are palindromic (Same from left to right as from right to left)?

*Solution.*

$$1. 2^5 = 32.$$

$$2. \binom{6}{0} + \binom{6}{2} + \binom{6}{4} + \binom{6}{6} = 32$$

$$3. 2^3 + 2^3 + 2^3 = 24$$

4.  $\binom{6}{3} = 20$

5.  $\binom{6}{3} + \binom{6}{4} + \binom{6}{5} + \binom{6}{6} = 42$

6.  $2^3 = 8$

□

**Problem 18.2.4** Let  $A$  and  $B$  be sets,  $|A \cup B| = 50$ ,  $|A| = 37$ , and  $|A \cap B| = 20$ . Calculate:

1.  $|B|$

2.  $|A \setminus B|$

3.  $|B \setminus A|$

4.  $|A \oplus B|$

*Solution.* 1.  $|B| = |A \cup B| - |A| + |A \cap B| = 33$ .

2.  $|A \setminus B| = |A| - |A \cap B| = 17$ .

3.  $|B \setminus A| = |B| - |A \cap B| = 13$ .

4.  $|A \oplus B| = |A \cup B| - |A \cap B| = 30$ .

□

**Problem 18.2.5** Let  $A, B$  be finite sets such that  $|A \cap B| = 10$ ,  $|A| = 22$ , and  $|B| = 15$ . Calculate:

1.  $|A \cup B|$

2.  $|A \setminus B|$

3.  $|A \oplus B|$

4.  $|A \times B|$

*Solution.* 1.  $|A \cup B| = |A| + |B| - |A \cap B| = 27$

2.  $|A \setminus B| = |A| - |A \cap B| = 12$

3.  $|A \oplus B| = |A \cup B| - |A \cap B| = 17$

4.  $|A \times B| = |A||B| = 330$ .

□

**Problem 18.2.6** Find a formula for  $|A \cup B \cup C|$ , where  $A, B, C$  are finite sets.

*Solution.*  $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$  □

### 18.2.1 Exam I

**Problem 18.2.7** A deck of cards contains 28 cards. Each card is either red, yellow, green, or blue, and there are seven cards for each color labelled with integers 1 to 7. Answer the following:

- How many ways can a hand of four cards be selected such that each card is a different color.
- How many ways can four cards be selected so that no cards are the same color or number.
- How many ways can a hand of four cards be selected so there is one red and three yellow cards.
- How many ways can four cards be drawn such that all cards are the same color?

*Solution.* 1.  $\frac{28 \cdot 21 \cdot 14 \cdot 7}{4!} = 2401$

2.  $\frac{28 \cdot 18 \cdot 10 \cdot 4}{4!} = 840$

3.  $\binom{7}{1} \binom{7}{3} = 245$

4.  $4 \binom{7}{4} = 140$

□

**Problem 18.2.8** A true false quiz has 6 questions. Answer the following:

- How many ways are there to fill out the quiz so that there are exactly four true answers?
- How many different ways are there to fill out the quiz so that there are at most two false answers?

*Solution.* 1.  $\binom{6}{2} = \frac{6!}{2!(6-4)!} = 15$

2.  $\binom{6}{4} + \binom{6}{5} + \binom{6}{6} = 22$

□

**Example 18.2.1** We can use the binomial theorem to easily compute the coefficients of a power of the form  $(x+a)^n$ . For example, let's find the coefficient of  $x^3y^6$  in the polynomial  $(x - 10y)^9$ . We have:

$$(x - 10y)^9 = \sum_{k=0}^9 \binom{9}{k} x^{n-k} (-10y)^k \quad (18.2.1)$$

Now we just need to compute the case when  $k = 6$ . We obtain:

$$\binom{9}{6} (-10)^6 = 84,000,000 \quad (18.2.2)$$

**Example 18.2.2** As another example, find the coefficient of  $x^2y^3$  in  $(7x - 10y)^5$ . We get:

$$(7x - 10y)^5 = \sum_{k=0}^5 \binom{5}{k} (7x)^{5-k} (-10y)^k \quad (18.2.3)$$

Computing the binomial coefficient for the case of  $k = 3$ , we have:

$$\binom{5}{3} 7^2 (-10)^3 = -490,000 \quad (18.2.4)$$

**Example 18.2.3** Expand  $(2x - 5y)^4$  using the binomial theorem.

$$(2x - 5y)^4 = \sum_{k=0}^4 \binom{n}{k} (2x)^{4-k} (-5y)^k \quad (18.2.5a)$$

$$= 16x^4 - 160x^3y + 600x^2y^2 - 1000xy^3 + 625y^4 \quad (18.2.5b)$$

**Problem 18.2.9** Expand and simplify  $\sum_{i=0}^3 \sum_{j=1}^2 (2x^i - x^j)$

*Solution.*  $\sum_{i=0}^3 \sum_{j=1}^2 (2x^i - x^j) = \sum_{i=0}^3 (4x^i - x^1 - x^2) = 4 + 4x^1 + 4x^2 + 4x^3 - 4x^1 - 4x^2 = 4 + 4x^3 \quad \square$

## 18.2.2 Practice Exam II

**Problem 18.2.10** Make a true table for  $p, q, p \wedge q, p \vee q, (p \wedge q) \vee (\neg p \wedge q)$

*Solution.*

$p$	$q$	$p \wedge q$	$p \vee q$	$(p \wedge q) \vee (\neg p \wedge q)$
0	0	0	0	0
0	1	0	1	1
1	0	0	1	0
1	1	1	1	1

Table 18.1: Truth Table for Problem 18.2.10

$\square$

**Problem 18.2.11** Prove that  $(p \wedge q) \vee (\neg p \wedge q)$  is equivalent to  $q$ .

*Solution.* By the distributive law,  $(p \wedge q) \vee (\neg p \wedge q) \Leftrightarrow (p \vee \neg p) \wedge q \Leftrightarrow 1 \wedge q \Leftrightarrow q$ .  $\square$

**Problem 18.2.12** Prove that if  $n$  is an integer, then  $n^2 + 6n$  is even if and only if  $n$  is even.

*Solution.* If  $n$  is even, then  $n = 2k$  for some integer  $k$ . Thus  $n^2 + 6n = 4k^2 + 12k = 2(2k^2 + 6k)$ , so  $n^2 + 6n$  is even. If  $n^2 + 6n$  is even  $n$  is odd, then  $n = 2k + 1$  for some integer  $k$ . Thus  $n^2 + 6n = 2(k^2 + 8k) + 1$ , which is odd. A contradiction. Therefore  $n$  is even.  $\square$

**Problem 18.2.13** Prove by contradiction that if  $7n^2 + 2n + 1$  is even, then  $n$  odd.

*Proof.* If  $n$  is even, then  $n = 2k$  for some integer  $k$ . Thus  $7n^2 + 2n + 1 = 28k^2 + 14k + 1 = 2(14k^2 + 7k) + 1$ , which is odd, a contradiction. Therefore  $n$  is odd.  $\square$

**Problem 18.2.14** Prove the following:  $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$

*Solution* We prove by induction. The base case is true. Suppose it is true for some  $n \in \mathbb{N}$ . Then:

$$\sum_{k=1}^{n+1} k^2 = (n+1)^2 + \sum_{k=1}^n k^2 = (n+1)^2 + \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(n+2)(2n+3)}{6} = \frac{(n+1)(n+2)(2n+3)}{6} \quad (18.2.6)$$

**Problem 18.2.15** Does  $\neg p \wedge (p \rightarrow q)$  imply  $\neg q$ ?

*Solution* No, it does not.

*Solution.* 1. There exists an integer  $x$  such that  $x$  is even and prime. This is true for 2 is such an integer.

2. For all positive integers  $n$ ,  $n^2 > 0$ . This is true of all real numbers.
3. For all positive integers  $x$  there is a rational number  $y$  such that  $xy = 1$ . This is true for let  $y = \frac{1}{x}$ .
4. There exists a rational number  $y$  such that for all positive integers  $x$ ,  $xy = 1$ . This is false. For if  $x_1y = 1$  and  $x_2y = 1$ , we have  $(x_1 - x_2)y = 0$ . But  $x_1 \neq x_2$ , so  $y = 0$ . A contradiction as  $x_1y = 1$ .

$\square$

**Problem 18.2.16** Let  $U = \mathcal{P}(\{1, 2, 3, 4, 5\})$ , and consider the following propositions over  $U$ :

$$p(A) : A \cap \{1, 2, 3\} = \emptyset \quad q(A) : A \subset \{1, 2, 5\} \quad r(A) : |A| = 2$$

Find the truth sets of the following  $p, r, q \wedge r, q \vee p$ .

*Solution.*

1.  $T_p = \mathcal{P}(\{4, 5\})$
2.  $T_r = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{2, 3\}, \{2, 4\}, \{2, 5\}, \{3, 4\}, \{3, 5\}, \{4, 5\}\}$ .
3.  $T_{q \wedge r} = \{\{1, 3\}, \{1, 5\}, \{3, 5\}\}$ .
4.  $T_{q \vee p} = \{\emptyset, \{1\}, \{3\}, \{5\}, \{1, 3\}, \{1, 5\}, \{3, 5\}, \{4\}, \{4, 5\}, \{1, 3, 5\}\}$ .

□

**Problem 18.2.17** Consider the following propositions:

$$p(n) : n^2 = 100 \quad q(n) : 2n - 20 = 0 \quad r(n) : n^4 = 10,000$$

1. Suppose  $p, q, r$  are over  $\mathbb{R}$ . Which are equivalent? Which propositions imply another?
2. Suppose  $p, q, r$  are over  $\mathbb{C}$ . Which are equivalent? Which propositions imply another?

*Solution.*

1. Over  $\mathbb{R}$ , we have  $T_p = \{-10, 10\}$ ,  $T_q = \{10\}$ , and  $T_r = \{-10, 10\}$ . So  $p$  and  $r$  are equivalent since  $T_p = T_r$ . Also  $q \Rightarrow p$  and  $q \Rightarrow r$  since  $T_q \subset T_p$  and  $T_q \subset T_r$ .
2. Over  $\mathbb{C}$ , we have  $T_p = \{-10, 10\}$ ,  $T_q = \{10\}$ ,  $T_r = \{-10i, 10i, -10, 10\}$ . So none of the propositions are equivalent, however  $q \Rightarrow p \Rightarrow r$ .

□

**Problem 18.2.18** Make a truth table for  $p \wedge 1$ ,  $p \rightarrow p \vee q$ ,  $p \vee q \rightarrow q$ ,  $p \vee 0 \leftrightarrow \neg p$ . Determine which are tautologies, contradictions, or neither.

*Solution.*  $p \rightarrow p \vee q$  is a tautology,  $p \vee 0 \leftrightarrow \neg p$  is a contradiction, both  $p \wedge 1$  and  $p \vee q \rightarrow q$  are neither.

$p$	$q$	$\neg p$	$\neg q$	0	1	$p \vee q$	$p \vee 0$	$p \wedge 1$	$p \rightarrow p \vee q$	$p \vee q \rightarrow q$	$p \vee 0 \leftrightarrow \neg p$
0	0	1	1	0	1	0	0	0	1	1	0
1	0	0	1	0	1	1	1	1	1	0	0
0	1	1	0	0	1	1	0	0	1	1	0
1	1	0	0	0	1	1	1	1	1	1	0

Table 18.2: Truth Table for Problem 18.2.18

□

### 18.2.3 Exam II

**Problem 18.2.19** Let  $n$  be an integer. Prove the  $n$  is odd if and only if  $n^2 + 2n + 5$  is even.

*Solution.* Suppose  $n$  is odd. Then there is a  $k \in \mathbb{Z}$  such that  $n = 2k + 1$ . So  $n^2 + 2n + 5 = (2k + 1)^2 + 2(2k + 1) + 5 = 4k^2 + 4k + 1 + 4k + 2 + 5 = 4k^2 + 8k + 8 = 2(2k^2 + 4k + 4)$ , which is even. Thus if  $n$  is odd, then  $n^2 + 2n + 5$  is even. Let  $n^2 + 2n + 5$  be even and suppose  $n$  is even. Then  $n = 2k$ . But then  $n^2 + 2n + 5 = 4k^2 + 4k + 5 = 2(2k^2 + 2k + 2) + 1$ . But this is an odd number, a contradiction. Therefore,  $n$  is not an even number. □

**Problem 18.2.20** Prove that for all  $n \in \mathbb{N}$ ,  $\sum_{k=1}^n 3k(k+1) = n(n+1)(n+1)$ .

*Solution.* By induction. The base case is trivial. Suppose it is true for  $n \in \mathbb{N}$ . Then  $\sum_{k=1}^{n+1} 3k(k+1) = 3(n+1)(n+2) + \sum_{k=1}^n 3k(k+1) = 3(n+1)(n+2) + n(n+1)(n+2) = (n+3)(n+1)(n+2) = (n+1)((n+1)+1)((n+1)+2)$ .  $\square$

**Problem 18.2.21** Consider  $(p \wedge q) \leftrightarrow (\neg q \vee \neg p)$ .

1. Write down the truth table for  $(p \wedge q) \leftrightarrow (\neg q \vee \neg p)$  and  $\neg(q \vee p)$
2. Determine whether this proposition is a tautology, contradiction, or neither.
3. Is  $\neg(q \vee p)$  equivalent to  $\neg q \vee \neg p$

*Solution.* The Proposition is a tautology. Also,  $\neg q \vee \neg p$  is not equivalent to  $\neg(q \vee p)$ .

$p$	$q$	$p \wedge q$	$p \vee q$	$\neg q$	$\neg p$	$\neg q \vee \neg p$	$\neg(q \vee p)$	$(p \wedge q) \leftrightarrow (\neg q \vee \neg p)$
0	0	0	0	1	1	1	1	0
0	1	0	1	0	1	1	0	0
1	0	0	1	1	0	1	0	0
1	1	1	1	0	0	0	0	0

Table 18.3: Truth Table for Problem 18.2.21

$\square$

**Problem 18.2.22** Let  $n$  be an integer.

1. Write the contrapositive of : If  $n^2 \geq 1$ , then  $n \geq 1$ . Is this true?
2. Write the converse of: If  $n^2 \geq 1$ , then  $n \geq 1$ . Is this true?

*Solution.*

1. If  $n < 1$ , then  $n^2 < 1$ . This is false. If  $n = -5$ , then  $n < 1$ , yet  $n^2 = 25 > 1$ .
2. If  $n \geq 1$ , then  $n^2 \geq 1$ . This is true, for if  $n \geq 1$ , then squaring both sides we get  $n^2 \geq 1^2 = 1$ .

$\square$

**Problem 18.2.23** Let  $U$  be a set with  $n$  elements. How many different propositions of  $U$  can you list without listing two that are equivalent?

*Solution.* Two propositions  $p$  and  $q$  are equivalent if  $T_p = T_q$ . There are  $|\mathcal{P}(U)| = 2^n$  possible truth sets. So there are  $2^n$  non-equivalent propositions over  $U$ .  $\square$

**Problem 18.2.24** Consider the following propositions of  $\mathbb{N}$ .

$$p(n) : 0 \leq n \leq 5 \quad q(n) : n \text{ is a prime greater than } 7 \quad r(n) : n^2 \leq 9 \quad s(m,$$

1. Translate the following statement into English:  $\forall_{n \in \mathbb{Z}} \exists_{m \in \mathbb{Z}} (s(m, n))$
2. Translate the following symbolically: There exists an integer  $n$  so that  $n^2 > 9$  or  $0 \leq n \leq 5$ .
3. Determine the truth sets for  $p(n), q(n), r(n)$ .
4. Does one of the propositions imply another?
5. Is the statement in part 1 true?

*Solution.*

1. For all integers  $n$  there exists an integer  $m$  so that  $2m + n = 4$ .
2.  $\exists_{n \in \mathbb{Z}} (\neg r(n) \vee p(n))$
3. (a)  $T_p = \{0, 1, 2, 3, 4, 5\}$   
 (b)  $T_q = \{2, 3, 5\}$   
 (c)  $\{\}, 3, 2, 1, 0, 1, 2, 3\}$
4. Yes,  $q(n) \Rightarrow P(n)$  because  $T_q \subset T_p$ .
5. False. If  $n = 1$  then  $2m + 1$  is odd for all  $n \in \mathbb{Z}$ , yet 4 is even.

□

#### 18.2.4 Practice Exam III

**Problem 18.2.25** Let  $A = \{1, 2, 3\}$ ,  $B = \{3, 4, 5, 6\}$ , and  $C = \{1, 2, 4, 6\}$ . Let the universe set be  $U = \{1, 2, 3, 4, 5, 6, 7\}$ .

1. Find all minsets generated by  $A, B, C$ .
2. Show that the nonempty minsets form a partition of  $U$ .
3. How many different sets in  $\mathcal{P}(U)$  can be generated by  $A, B, C$  via any combination of union, intersection, and complement?
4. Express  $\{1, 2, 3, 5, 7\}$  and  $\{5, 6, 7\}$  in minset normal form, if possible. If not, explain why.

*Solution.* The non-empty sets form a partition of  $U$  for their union is  $U$ , and there are no overlaps. There are 5 non-empty minsets, so  $2^5 = 32$  subsets of  $\mathcal{P}(U)$  can be generated from the minsets.  $\{1, 2, 3, 5, 7\} = (A^c \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A \cap B \cap C^c)$ .  $\{5, 6, 7\}$  cannot be expressed as a union of minsets, for 6 lies in the minset  $\{4, 6\}$ , and thus 4 would need to be included, but it is not.

$A$	$B$	$C$		
0	0	0	$A^c \cap B^c \cap C^c$	$\{7\}$
0	0	1	$A^c \cap B^c \cap C$	$\emptyset$
0	1	0	$A^c \cap B \cap C^c$	$\{5\}$
0	1	1	$A^c \cap B \cap C$	$\{4, 6\}$
1	0	0	$A \cap B^c \cap C^c$	$\emptyset$
1	0	1	$A \cap B^c \cap C$	$\{1, 2\}$
1	1	0	$A \cap B \cap C^c$	$\{3\}$
1	1	1	$A \cap B \cap C$	$\emptyset$

Table 18.4: The Minsets of  $A$ ,  $B$ , and  $C$ .

□

**Problem 18.2.26** Let  $A$  and  $B$  be subsets of  $U$ .

1. List all minsets generated by  $A, B$ .
2. Express the sets  $A$  and  $A^c \cup B$  in minset normal form.

*Solution.*

$$A = A \cap U = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c)$$

$$A^c \cup B = (A^c \cap U) \cup (B \cap U) = (A^c \cap (B \cup B^c)) \cup (B \cap (A \cup A^c)) = (A^c \cap B) \cup (A^c \cap B^c) \cup (B \cap A) \cup (B \cap A^c)$$

$A$	$B$	
0	0	$A^c \cap B^c$
0	1	$A^c \cap B$
1	0	$A \cap B^c$
1	1	$A \cap B$

Table 18.5: Minsets of  $A$  and  $B$ .

□

**Problem 18.2.27** Let:

$$A = \begin{bmatrix} 1 & 4 \\ 0 & 5 \\ 2 & -3 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 10 \\ -1 & 0 \\ 5 & -1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 3 \\ -2 & 4 \end{bmatrix} \quad D = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \quad F = \begin{bmatrix} 1 & 3 \\ -2 & -6 \end{bmatrix}$$

Compute the following, if possible.

1.  $AB$
2.  $3A - 2B$
3.  $7A + 2C$
4.  $A^2$
5.  $C^2$
6.  $D^4$
7.  $AC + BC$
8.  $CA$
9.  $C^{-1}$
10.  $C^{-2}$
11.  $F^{-1}$ .
12.  $XC = A$

*Proof.*

1.  $AB$  is not possible because  $A$  and  $B$  are both  $3 \times 2$  matrices. The number of columns of  $A$  must be the same as the number of rows of  $B$ .

2.  $3A - 2B = \begin{pmatrix} -1 & -8 \\ 2 & 15 \\ -4 & -7 \end{pmatrix}$

3.  $7A + 2C$  is not possible because  $A$  and  $C$  are of different dimensions.

4.  $A^2$  is not possible because  $A$  is not a square matrix.

5.  $C^2 = \begin{pmatrix} -5 & 15 \\ -10 & 10 \end{pmatrix}$

6.  $D^4 = \begin{pmatrix} 81 & 0 \\ 0 & 16 \end{pmatrix}$

7.  $AC + BC = (A + B)C = \begin{pmatrix} 3 & 14 \\ -1 & 5 \\ 7 & -4 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} -25 & 65 \\ -11 & 17 \\ 15 & 5 \end{pmatrix}$

8.  $C$  is a  $2 \times 2$  and  $A$  is a  $3 \times 2$ , so multiplication is undefined.

9.  $\det(C) = 10$ , so  $C^{-1} = \frac{1}{10} \begin{pmatrix} 4 & -3 \\ 2 & 1 \end{pmatrix}$

10.  $C^{-2} = (C^2)^{-1} = \frac{1}{100} \begin{pmatrix} 10 & -15 \\ 10 & -5 \end{pmatrix}$

11.  $\det(F) = 0$ , so  $F^{-1}$  does not exist.

12.  $X = AC^{-1} = \frac{1}{10} \begin{pmatrix} 12 & 1 \\ 10 & 5 \\ 2 & -9 \end{pmatrix}$

□

**Problem 18.2.28** Convert the following into  $AX = B$ , and then solve.

$$4x - y = 10$$

$$3x - 2y = 3$$

*Solution.*

$$\begin{bmatrix} 4 & -1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 10 \\ 3 \end{bmatrix} \Rightarrow X = A^{-1}B \Rightarrow X = -\frac{1}{5} \begin{bmatrix} -2 & -3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 29 \\ -22 \end{bmatrix} \Rightarrow X = \frac{1}{5} \begin{bmatrix} 17 \\ 18 \end{bmatrix}$$

□

**Problem 18.2.29** Prove or disprove the following for  $n \times n$  matrices  $A, B$ .

1.  $AB = BA$

2.  $(A + B)(A - B) = A^2 - B^2$

3. If  $A^2 = AB$ , and  $\det(A) = 4$ , then  $A = B$ .

4. If  $AB = 0$ , then  $A = 0$  or  $B = 0$

*Solution.*

1. False, for  $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ , but  $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ .

2. False, for  $(A+B)(A-B) = A(A-B)+B(A-B) = A^2 - AB + BA - B^2 = d(A^2 - B^2) + (BA - AB)$ . Since  $AB$  may not necessarily be equal to  $BA$ ,  $BA - AB$  may not zero.

3. True. If  $\det(A) = 4$ , then  $A^{-1}$  exists. Then  $A^2 = AB$ , and thus  $A^{-1}A^2 = A^{-1}AB \Leftrightarrow A = B$ .

4. False. For  $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = 0$ .

□

### 18.2.5 Exam III

**Problem 18.2.30** Prove that for sets  $A, B, C$ ,  $(A \cap B) \times C \subset B \times C$ .

*Solution.* For let  $(x, y) \in (A \cap B) \times C$ . Then  $x \in A \cap B$  and  $y \in C$ . But if  $x \in A \cap B$ , then  $x \in B$ . But if  $x \in B$  and  $y \in C$ , then  $(x, y) \in B \times C$ . Therefore,  $(A \cap B) \times C \subset B \times C$ .  $\square$

**Problem 18.2.31** For set  $A, B \subset U$ , prove that  $A \cup (A \cap B)^c = U$ , and find the dual of this.

*Solution.* For Let  $x \in U$ . If  $x \in A$ , then  $x \in A \cup (A \cap B)^c$ . Suppose not. Then  $x \in A^c$ . But if  $x \in A^c$ , then  $x \notin A \cap B$ . But if  $x \notin A \cap B$ , then  $x \in (A \cap B)^c$ . But then  $x \in A \cup (A \cap B)^c$ . Therefore  $U \subset A \cup (A \cap B)^c$ . But  $A \cup (A \cap B)^c \subset U$ , as  $U$  is the universe set. Therefore  $A \cup (A \cap B)^c = U$ . The dual is  $A \cap (A \cup B)^c = \emptyset$   $\square$

**Problem 18.2.32** Convert the following into the form  $AX = B$ . Find  $A^{-1}$ , and then solve for  $X$ :

$$\begin{aligned} 4x - 6y &= 5 \\ 3x - 7y &= -7 \end{aligned}$$

*Solution.*

$$\begin{bmatrix} 4 & -6 \\ 3 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ -7 \end{bmatrix} \Rightarrow X = A^{-1}B = -\frac{1}{10} \begin{bmatrix} -7 & 6 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ -7 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 83 \\ 47 \end{bmatrix}$$

$\square$

**Problem 18.2.33** Let  $A = \begin{pmatrix} 1 & 0 & -2 \\ 5 & 3 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 0 & 2 \\ 4 & -10 \\ 8 & -6 \end{pmatrix}$ ,  $C = \begin{pmatrix} 7 & -1 \\ -2 & 5 \\ -4 & 3 \end{pmatrix}$ ,  $E = \begin{pmatrix} 1 & -3 \\ 2 & 5 \end{pmatrix}$ . Compute the following, if possible:

1.  $3A + 5C$
2.  $2B - 3C$
3.  $CE$
4.  $EC$
5.  $E^2$
6.  $BA + 2CA$

*Solution.*

1.  $A$  and  $C$  do not have the same dimensions, so this can't be done.

$$2. \quad 2B - 3C = \begin{pmatrix} -21 & 7 \\ 14 & -35 \\ 28 & -21 \end{pmatrix}$$

$$3. \quad CE = \begin{pmatrix} 5 & -26 \\ 8 & 31 \\ 2 & 27 \end{pmatrix}$$

4.  $EC$  can't be done as  $E$  is a  $2 \times 2$  and  $C$  is a  $3 \times 2$ .

$$5. \quad E^2 = \begin{pmatrix} -5 & -18 \\ 12 & 19 \end{pmatrix}$$

$$6. \quad BA + 2CA = \begin{pmatrix} 13 & 0 & -28 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

□

**Problem 18.2.34** Let  $A = \{1, 3, 5\}$ ,  $B = \{2, 3, 4, 5\}$ , and  $U = \{1, 2, 3, 4, 5, 6\}$ .

1. Find all minsets generated by  $A$  and  $B$ .
2. How many different sets in the power set of  $U$  can be generated by  $A, B$  by any combination of union, intersection, and complement?
3. Express  $\{2, 4, 6\}$  in minset normal form.
4. Find the maxsets generated by  $A$  and  $B$ .
5. Express  $\{1, 3, 5\}$  in maxset normal form.

*Solution.*

1.  $A^c \cap B^c = \{6\}$ ,  $A \cap B^c = \{1\}$ ,  $A^c \cap B = \{2, 4\}$ ,  $A \cap B = \{3, 5\}$ .
2. There are 4 non-empty minsets, so  $2^4 = 16$ .
3.  $\{2, 4, 6\} = (A^c \cap B) \cup (A^c \cap B^c)$ .
4.  $A \cup B = \{1, 2, 3, 4, 5\}$ ,  $A^c \cup B = \{2, 3, 4, 5, 6\}$ ,  $A \cup B^c = \{1, 3, 5, 6\}$ ,  $A^c \cup B^c = \{1, 2, 4, 6\}$ .
5.  $\{1, 3, 5\} = (A \cup B) \cap (A \cup B^c)$ .

□

### 18.2.6 Final Exam

**Problem 18.2.35** Let  $A = \{1, 2\}$ ,  $B = \{2, 3, 4, 6\}$ ,  $C = \{4, 6, 7\}$ ,  $U = \{1, 2, 3, 4, 5, 6, 7, 8\}$ . Compute the following:

1.  $A^c \cap B$
2.  $A \cup C$
3.  $B \oplus C$
4.  $B \setminus C$
5.  $A \times C$
6.  $A^3$

7.  $\mathcal{P}(C)$

*Solution.*

1.  $A^c \cap B = \{3, 4, 6\}$ .
2.  $A \cup C = \{1, 2, 4, 6, 7\}$
3.  $B \oplus C = \{2, 3, 7\}$
4.  $B \setminus C = \{2, 3\}$
5.  $A \times C = \{(1, 4), (1, 6), (1, 7), (2, 4), (2, 6), (2, 7)\}$
6.  $A^3 = \{(1, 1, 1), (1, 1, 2), (1, 2, 1), (1, 2, 2), (2, 1, 1), (2, 1, 2), (2, 2, 1), (2, 2, 2)\}.$
7.  $\mathcal{P}(C) = \{\emptyset, \{4\}, \{6\}, \{7\}, \{4, 6\}, \{4, 7\}, \{6, 7\}, \{4, 6, 7\}\}.$

□

**Problem 18.2.36** Prove the following is false: If  $A \cap B = A \cap C$ , then  $B = C$ .

*Solution.* For let  $A = \{1\}$ ,  $B = \{1, 2\}$ , and  $C = \mathbb{R}$ . Then  $A \cap B = \{1\}$ ,  $A \cap C = \{1\}$ , but  $B \neq C$ . □

**Problem 18.2.37** Ten students are competing for a scholarship.

1. If there are three scholarships worth \$2000, how many ways can they be distributed?
2. If there are two scholarships worth \$5000 and three worth \$2000, how many ways can they be distributed?

3. Suppose that the group of ten students consists of six freshmen and four sophomores. In how many different ways can four equal scholarships be distributed if at least two of the scholarships should be awarded to freshmen?
4. Suppose the group of ten students consists of six freshmen and four sophomores. In how many different ways can two scholarships of \$5000 and two scholarships of \$2000 be distributed if at least three of the scholarships will be awarded to freshmen?

*Solution.*

1.  $\binom{10}{3} = \frac{10!}{3!(10-3)!} = 120$
2.  $\binom{10}{2}\binom{8}{3} = 2520$
3. If 2 scholarships are awarded to freshmen, we have  $\binom{6}{2}\binom{4}{2} = 90$ . If 3 scholarships are awards to freshmen, we have  $\binom{6}{3}\binom{4}{1} = 80$ . If 4 scholarships are awarded to freshmen, we have  $\binom{6}{4}\binom{4}{0} = 15$ . Adding them together, we get 185.
4. If 2 \$5000 scholarships are awarded to freshmen, and 1 \$2000 scholarship is awarded to a freshman, then there are  $\binom{6}{3}\binom{4}{1} = 80$ . Simiarly if 2 \$2000 scholarships are awarded to freshmen and 1 \$5000 scholarship is awarded to a freshamn. Finally, there are  $\binom{6}{4} = 15$  ways to give all scholarships to freshmen. In total, there are 175 total possible outcomes.

□

## 18.3 Quizzes

### 18.3.1 Quiz I

**Problem 18.3.1** Express 49 in binary.

*Solution.*

$$49 = 2 \cdot 24 + 1$$

$$6 = 2 \cdot 3 + 0$$

$$24 = 2 \cdot 12 + 0$$

$$3 = 2 \cdot 1 + 1$$

$$12 = 2 \cdot 6 + 0$$

$$1 = 2 \cdot 0 + 1$$

So,  $49 = 110001_2$

□

### 18.3.2 Quiz II

**Problem 18.3.2** Calculate  $\sum_{k=-1}^3 (2^k + 1)$ .

*Solution.*  $\sum_{k=-1}^3 (2^k + 1) = (2^{-1} + 1) + (2^0 + 1) + (2^1 + 1) + (2^2 + 1) + (2^3 + 1) = 20 + \frac{1}{2} = \frac{41}{2}$ .  $\square$

**Problem 18.3.3** Three men and three women are to be seated in a row.

1. How many different ways can the six people be seated?
2. How many different ways can the six people be seated if it is required that the genders alternate.

*Solution.*

1.  $6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$
2. It is  $MWMWMW$ , so  $3 \cdot 3 \cdot 2 \cdot 2 \cdot 1 \cdot 1 = 72$ . Or, there are 6 ways to seat the first person, 3 ways to seat the second person, 2 ways to seat the third person, 2 ways to seat the fourth person, and 1 way to seat the last two. So,  $6 \cdot 3 \cdot 2 \cdot 2 \cdot 1 \cdot 1 = 72$ .

$\square$

**Problem 18.3.4** Calculate  $P(7; 3)$

*Solution.*  $P(7; 3) = \frac{7!}{(7-3)!} = 7 \cdot 6 \cdot 5 = 210$ .  $\square$

**Problem 18.3.5** Let  $A$  be a set such that  $|A| = n$ .

1. Calculate  $|A^4|$
2. Calculate  $|\{\{a, b, c, d\} \subset A : \text{Each Term is Different}\}|$

*Solution.* 1.  $|A^4| = |A \times A \times A \times A| = n^4$

$$2. n \cdot (n - 1) \cdot (n - 2) \cdot (n - 3) = P(n; 4)$$

$\square$

### 18.3.3 Quiz III

**Problem 18.3.6** Let  $p, q, r$  be the following propositions:

$$p(x) : x = 1 \quad p(x) : x = -1 \quad r(x) : x^2 = 1$$

1. Express "If  $x^2 = 1$ , then  $x = 1$  and  $x = -1$ ," in symbolic form.
2. Write the converse of this in English, and symbolically.
3. Express  $\neg p \wedge \neg r$  in English.
4. Express  $r \leftrightarrow (q \vee p)$  in English.

*Solution.*

1.  $r \rightarrow (p \wedge q)$
2.  $(p \wedge q) \rightarrow r$ . If  $x = 1$  and  $x = -1$ , then  $x^2 = 1$ .
3.  $x \neq 1$  and  $x^2 \neq 1$
4.  $x^2 = 1$  if and only if  $x = 1$  or  $x = -1$ .

□

# CHAPTER 19

---

## Linear Algebra

---

### 19.1 Linear Algebra

#### Example 19.1.1: I

---

$V$  and  $W$  are 2-dimensional subspaces in  $\mathbb{R}^4$ , what are the possible dimensions of  $V \cap W$ . If  $V$  and  $W$  are subspaces, then  $V \cap W$  is subspace, and  $\dim(V \cap W) \leq \min(\{\dim(V), \dim(W)\})$  we have in our problem that  $\dim(V \cap W) \leq 2$ . We now must show that  $V \cap W$  can have dimensions 0, 1, or 2. If  $V = \{(x, y, 0, 0) : x, y \in \mathbb{R}\}$  and  $W = \{(0, 0, z, w) : z, w \in \mathbb{R}\}$ , then  $V \cap W = \{(0, 0, 0, 0)\}$  which has dimension 0. If  $V = \{(x, y, 0, 0) : x, y \in \mathbb{R}\}$  and  $W = \{(0, y, z, 0) : y, z \in \mathbb{R}\}$ , then  $V \cap W = \{(0, y, 0, 0) : y \in \mathbb{R}\}$  which has dimension 1. Finally, if  $V = W$  then  $V \cap W = V$ , which has dimension 2. So, the only possibilities are 0, 1, or 2. ■

A system of linear equations can be written using matrices. Suppose we have the following equations:

$$a_{0,0}x_0 + a_{0,1}x_1 + a_{0,2}x_2 = b_0 \quad (19.1.1)$$

$$a_{1,0}x_0 + a_{1,1}x_1 + a_{1,2}x_2 = b_1 \quad (19.1.2)$$

$$a_{2,0}x_0 + a_{2,1}x_1 + a_{2,2}x_2 = b_2 \quad (19.1.3)$$

$$(19.1.4)$$

where all the variables belong to some field  $(\mathbb{F}, +, \cdot)$ . We can express this system in terms of matrices as follows:

$$\begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} \\ a_{1,0} & a_{1,1} & a_{1,2} \\ a_{2,0} & a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \quad (19.1.5)$$

Better yet, if we let  $\mathbf{x}, \mathbf{b} \in \mathbb{F}^3$  be the points such that  $\mathbf{x}(k) = x_k$  and  $\mathbf{b}(k) = b_k$ , for  $k \in \mathbb{Z}_3$ , and if we let  $\mathbf{A} : \mathbb{F}^3 \rightarrow \mathbb{F}^3$  be the linear operator defined by this matrix, we can write:

$$\mathbf{A}(\mathbf{x}) = \mathbf{b} \quad (19.1.6)$$

Matrices can also be written as  $\mathbf{A} = (a_{ij})$ . The following rules are used to define matrix arithmetic.

1. Addition: To add two matrices, add their corresponding elements. That is, if  $\mathbf{A} = (a_{ij})$  and  $\mathbf{B} = (b_{ij})$ , then  $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$ . Matrix addition is only defined on matrices of the same size.
2. Scale multiplication: To multiply a matrix by a real or complex number  $c$ , multiply this number to every element. That is, if  $\mathbf{A} = (a_{ij})$ , then  $c\mathbf{A} = (c \cdot a_{ij})$
3. Matrix Multiplication: The product of an  $M \times N$  matrix with an  $N \times P$  matrix is defined by  $\mathbf{C} = \mathbf{AB}$ , where  $(c_{ij}) = (\sum_{k=1}^N a_{ik}b_{kj})$ . Note that it is possible for  $\mathbf{AB} \neq \mathbf{BA}$ . Indeed, it is possible for  $\mathbf{AB}$  to be defined, whereas  $\mathbf{BA}$  is undefined.

**Example 19.1.2** Let the following be true:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

Then, using the defined rules, we have:

$$\begin{aligned} A + B &= \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix} & 5A &= \begin{bmatrix} 5 & 10 \\ 15 & 20 \end{bmatrix} \\ AB &= \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} & BA &= \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \end{aligned}$$

Note that even in this trivial example,  $AB \neq BA$ .

**Definition 19.1.1** The  $n \times n$  identity matrix is the matrix  $I_n = (I_{ij})$ , where

$$I_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

**Definition 19.1.2** An inverse matrix of an  $n \times n$  matrix  $A$  is a matrix  $A^{-1}$  such that  $AA^{-1} = A^{-1}A = I_n$

Not every matrix has an inverse matrix. If one does exist, there are many properties it contains.

**Theorem 19.1.1.** *The following are true:*

1. If  $\mathbf{A}$  and  $\mathbf{B}$  are invertible  $n \times n$  matrices, then  $\mathbf{AB}$  is invertible and  $\mathbf{AB}^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
2. If  $\mathbf{A}$  is an invertible matrix, then  $\mathbf{A}^{-1}$  is an invertible matrix and  $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$

**Definition 19.1.3** The trace of an  $n \times n$  matrix  $A$  is the sum of its diagonal:  
 $\text{Tr}(A) = \sum_{i=1}^n a_{ii}$

**Example 19.1.3**

$$\text{Tr} \left( \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 8 & 8 & 3 \end{bmatrix} \right) = 4 + 2 + 3 = 9$$

**Definition 19.1.4** The determinant of a  $2 \times 2$  matrix  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  is  $\det(A) = ad - bc$

**Definition 19.1.5** The minor of the  $i^{th}$  row and  $j^{th}$  column of an  $n \times n$  matrix  $\mathbf{A}$ , denoted  $M_{ij}$ , is the determinant of the  $(n-1) \times (n-1)$  matrix formed by removing the  $i^{th}$  row and  $j^{th}$  column from  $\mathbf{A}$ .

**Definition 19.1.6** The cofactor of the minor  $M_{ij}$  of an  $n \times n$  matrix  $\mathbf{A}$ , denoted  $C_{ij}$ , is  $(-1)^{i+j} M_{ij}$ .

**Example 19.1.4**

$$A = \begin{bmatrix} 7 & 1 & 3 \\ 1 & 3 & 5 \\ 17 & 4 & 20 \end{bmatrix} \quad M_{11} = \det \left( \begin{bmatrix} 3 & 5 \\ 4 & 20 \end{bmatrix} \right) = 40 \quad C_{11} = (-1)^{1+1} M_{11} = 40$$

**Definition 19.1.7** The determinant of an  $n \times n$  matrix  $\mathbf{A}$  is  $\det(A) = \sum_{j=1}^n a_{1j} C_{1j}$

**Theorem 19.1.2.** If  $\mathbf{A}$  is an  $n \times n$  matrix and  $1 \leq i \leq n$ , then  $\det(A) = \sum_{j=1}^n a_{ij} C_{ij}$

**Definition 19.1.8** The transpose of an  $n \times m$  matrix  $\mathbf{A}$ , denoted  $\mathbf{A}^T$ , is the  $m \times n$  matrix formed by swapping the rows and columns of  $\mathbf{A}$  with each other. That is  $(a_{ij})^T = (a_{ji})$ .

**Definition 19.1.9** A symmetric matrix is a matrix  $\mathbf{A}$  such that  $\mathbf{A}^T = \mathbf{A}$

**Theorem 19.1.3.** If  $\mathbf{A}$  is an  $n \times m$  matrix and  $\mathbf{B}$  is an  $m \times p$  matrix, then the following are true:

1.  $(\mathbf{A}^T)^T = \mathbf{A}$
2.  $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
3.  $(k\mathbf{A})^T = k\mathbf{A}^T$
4.  $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$

**Definition 19.1.10** The adjoint of an  $n \times n$  matrix  $\mathbf{A}$ , denoted  $\text{adj}\mathbf{A}$ , is the matrix  $(C_{ij})^T$ .

**Theorem 19.1.4.** If  $\det(\mathbf{A}) \neq 0$ , then  $\mathbf{A}$  is invertible and  $\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj}\mathbf{A}$

**Theorem 19.1.5.** If  $\mathbf{A}$  and  $\mathbf{B}$  are  $n \times n$  matrices, then the following are true:

$$\begin{array}{lll} 1. \det(\mathbf{A}) = \det(\mathbf{A}^T) & 2. \det(k\mathbf{A}) = k^n \det(\mathbf{A}) & 3. \det(\mathbf{AB}) \\ & & = \det(\mathbf{A}) \det(\mathbf{B}) \end{array}$$

**Theorem 19.1.6.** A matrix  $\mathbf{A}$  is invertible if and only if  $\det(\mathbf{A}) \neq 0$

**Theorem 19.1.7.** If  $\mathbf{A}$  is invertible, then  $\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$

The differential equation  $\sum_{k=0}^n a_k y^{(k)}(x)$  can be expressed in terms of the characteristic polynomial  $\sum_{k=0}^n a_k D^k$ . Factoring this linear operator into  $\prod_{k=0}^n (D - r_k)$ , the general solution is  $y(x) = \sum_{k=1}^n c_k e^{r_k x}$ . If some of the  $r_k$  repeat, we have  $c_k x^{m_k - 1} e^{r_k x}$ , where  $m_k$  is the number of repetitions. In general, if we have  $\prod_{k=0}^n (D - r_k)^{m_k}$ , the general solution is  $y(x) = \sum_{k=1}^n c_k e^{r_k x} (\sum_{j=0}^{m_k - 1} x^j)$

**Example 19.1.5**

1.  $y''' - 4y'' + 4y' = 0$  has the characteristic polynomial  $D(D - 2)^2$ , so  $y(x) = c_1 + c_2 e^{2x} + c_3 x e^{2x}$

In linear algebra, the determinant  $\det(\mathbf{A} - \lambda I)$  is the characteristic polynomial of the square matrix  $\mathbf{A}$ .

**Definition 19.1.11** A vector space  $V$  over a Field (Set of scalars)  $F$  is a set  $V$  with two operations  $+$  and  $\cdot$  such that the following are true:

1.  $\forall \mathbf{a}, \mathbf{b} \in V \quad \mathbf{a} + \mathbf{b} \in V$
2.  $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
3.  $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$
4.  $\forall \mathbf{a} \in V \exists \mathbf{b} \in V : \mathbf{a} + \mathbf{b} = \mathbf{0}$
5.  $\forall k \in F, \mathbf{a} \in V \quad k\mathbf{a} \in V$
6.  $k(\mathbf{a} + \mathbf{b}) = k\mathbf{a} + k\mathbf{b}$
7.  $(k_1 + k_2)\mathbf{a} = k_1\mathbf{a} + k_2\mathbf{a}$
8.  $1\mathbf{a} = \mathbf{a}$
9.  $k_1(k_2\mathbf{a}) = (k_1 k_2)\mathbf{a}$

**Theorem 19.1.8.** If  $V$  is a vector space, then there is a  $\mathbf{0} \in V$  such that for all  $\mathbf{a} \in V$ ,  $\mathbf{a} + \mathbf{0} = \mathbf{a}$

**Definition 19.1.12** A linearly dependent subset of a vector space  $V$  (Over  $\mathbb{R}$ ) is a subset  $S \subset V$  such that there exists an  $N \in \mathbb{N}$ , a non-zero  $a_n : \mathbb{Z}_N \rightarrow \mathbb{R}$  and an injective function  $v_n : \mathbb{Z}_N \rightarrow V$  such that  $\sum_{k=1}^N a_n v_n = \mathbf{0}$

**Definition 19.1.13** A linearly independent subset of a vector space  $V$  is a subset  $S \subset V$  that is not linearly dependent.

**Theorem 19.1.9.** *If  $V \subset \mathbb{R}^n$  has more than  $n$  vectors, then  $V$  is linearly dependent.*

**Definition 19.1.14** The rank of a matrix is the number of linearly independent columns of the matrix.

**Example 19.1.6** Let  $\mathbf{A} = [A_1 \ A_2]$  where  $A_1 = (1, 2)^T$  and  $A_2 = (2, 4)^T$ . So  $2A_1 + (-1)A_2 = (0, 0)^T = \mathbf{0}$ . Therefore  $\{A_1, A_2\}$  is a linearly independent subset. Thus,  $\text{rk}(\mathbf{A}) = 1$ .

**Definition 19.1.15** A matrix with full rank is a square  $n \times n$  matrix  $\mathbf{A}$  such that  $\text{rk}(\mathbf{A}) = n$ .

**Theorem 19.1.10.** *If  $\mathbf{A}$  is a square matrix and  $\det(\mathbf{A}) \neq 0$ , then  $\mathbf{A}$  has full rank.*

**Theorem 19.1.11.** *If  $\mathbf{A}$  is a square matrix with full rank, then it is invertible.*

**Definition 19.1.16** A finite basis of a vector space  $V$  is a linearly independent subset  $S \subset V$  where  $S = \{\mathbf{v}_k\}_{k=0}^n$  and for all  $\mathbf{x} \in V$  there is an  $a_n : \mathbb{Z}_n \rightarrow \mathbb{R}$  such that  $\mathbf{x} = \sum_{k=1}^n a_k \mathbf{v}_k$

**Theorem 19.1.12.** *All bases of a vector space  $V$  have the same number of elements.*

**Definition 19.1.17** The dimension of a vector space  $V$  is the number of elements in any basis of  $V$ .

**Definition 19.1.18** An eigenvector of an  $n \times n$  matrix  $\mathbf{A}$  is a vector  $\mathbf{x} \in \mathbb{R}^n$  such that there exists a  $\lambda \in \mathbb{R}$  such that  $\mathbf{Ax} = \lambda \mathbf{x}$

**Definition 19.1.19** An eigenvalue of an  $n \times n$  matrix  $\mathbf{A}$  is a real number  $\lambda \in \mathbb{R}$  such that there is a vector  $\mathbf{x} \in \mathbb{R}^n$  such that  $\mathbf{Ax} = \lambda \mathbf{x}$

**Definition 19.1.20** The characteristic equation, or the characteristic polynomial, of an  $n \times n$  matrix  $\mathbf{A}$  is  $\det(\lambda I - \mathbf{A}) = 0$

**Definition 19.1.21** A diagonalizable matrix is an  $n \times n$  matrix  $\mathbf{A}$  such that there exists an invertible matrix  $\mathbf{B}$  such that  $\mathbf{A} = \mathbf{B}^{-1}\mathbf{AB}$

**Theorem 19.1.13.** *The following are true:*

1. *If  $\mathbf{A}$  is an  $n \times n$  diagonalable matrix, then  $\mathbf{A}$  has  $n$  linearly independent eigenvectors.*
2. *If  $\mathbf{A}$  is an  $n \times n$  matrix with  $n$  linearly independent eigenvectors, then  $\mathbf{A}$  is diagonalizable.*
3. *A symmetric matrix has all real eigenvalues.*

## 19.2 Miscellaneous Lecture Notes

### 19.2.1 Orthogonal Projections

**Definition 19.2.1** The span of  $\mathcal{W} = \{X_i\}_1^k \subset \mathbb{R}^n$  is the set  $\text{Span}(\mathcal{W}) = \{\sum_{i=1}^k a_i X_i : a_i \in \mathbb{R}\}$ .

**Definition 19.2.2** A linearly dependent subset of  $\mathbb{R}^n$  is a subset  $S \subset \mathbb{R}^n$  such that there exists a finite subset  $\{X_i\}_{i=1}^k \subset S$  and a subset  $\{a_i\}_{i=1}^k \subset \mathbb{R} \setminus \{0\}$  such that  $\sum_{i=1}^k a_i X_i = \mathbf{0}$

**Definition 19.2.3** A linearly independent subset of  $\mathbb{R}^n$  is a subset  $S \subset \mathbb{R}^n$  that is not linearly dependent.

**Theorem 19.2.1.** If  $S \subset \mathbb{R}^n$  is linearly independent, then  $|S| \leq n$ .

**Theorem 19.2.2.** If  $\mathcal{W} \subset \mathbb{R}^n$  is linearly independent and  $|\mathcal{W}| = k$ , then  $\text{Span}(\mathcal{W})$  is a  $k$  dimensional subspace of  $\mathbb{R}^n$ .

If we have a linearly independent subset  $\mathcal{W} = \{X_1, \dots, X_k\} \subset \mathbb{R}^n$ , and some other vector  $Y$ , we may wish to consider the orthogonal projection of  $Y$  onto the  $k$  dimensional subspace spanned by  $\mathcal{W}$ . That is, we may wish to find a vector  $Y' \in \text{Span}(X_1, \dots, X_k)$  such that  $Y - Y'$  is orthogonal to  $\text{Span}(X_1, \dots, X_k)$ .

**Theorem 19.2.3.** If  $\{X_1, \dots, X_k\} \subset \mathbb{R}^n$  is linearly independent,  $\mathcal{W} = \text{Span}(X_1, \dots, X_k)$  and if  $Y \in \mathbb{R}^n$  is such that  $Y \perp X_i$  for  $i = 1, 2, \dots, k$ , then  $\forall Z \in \mathcal{W}$ ,  $Y \perp Z$ .

*Proof.* For let  $Y \in \mathbb{R}^n$  be such that for  $i = 1, 2, \dots, k$ ,  $Y \perp X_i$ . Let  $Z \in \mathcal{W}$ . Then  $Z = \sum_{i=1}^k a_i X_i$ , where  $a_i \in \mathbb{R}$ . But then  $\langle Y, Z \rangle = \sum_{i=1}^k a_i \langle Y, X_i \rangle = \sum_{i=1}^k a_i \cdot 0 = 0$ .  $\square$

**Theorem 19.2.4.** If  $P$  is an  $n \times k$  matrix whose columns are linearly independent, then  $P^T P$  is invertible.

*Proof.* If  $P^T P X = 0$ , then  $P X$  is orthogonal to the columns of  $P$ . But  $P X$  is a linear combination of the columns of  $P$ , and thus  $P X$  is orthogonal to itself. Thus,  $P X = 0$ . But the columns of  $P$  are linearly independent, if  $P X = 0$ , then  $X = 0$ . Thus  $P^T P X = 0$  if and only if  $X = 0$ .  $P^T P$  is invertible.  $\square$

We need  $X_i^T(Y - Y') = 0$ . Let  $X_i = (x_{1i}, x_{2i}, \dots, x_{ni})^T$ ,  $P = (x_{ij})$ . Then  $P^T(Y - Y') = 0$ , so  $P^T Y = P^T Y'$ . But  $Y' \in \mathcal{W}$ , so  $Y' = \sum_{i=1}^k c_i X_i = P(c_1, \dots, c_k)^T = PC$ . Thus,  $C = (P^T P)^{-1} P^T Y$ . Therefore  $Y' = P(P^T P)^{-1} P^T Y$ .

**Definition 19.2.4** The projection matrix of  $\text{Span}(X_1, \dots, X_k)$  is  $P(P^T P)^{-1} P^T$ , where  $P = (x_{ij})$ .

**Theorem 19.2.5.** If  $Q = P(P^T P)^{-1} P^T$  is a projection matrix for a subspace  $\mathcal{W}$ , then  $Q^T = Q$ .

*Proof.*  $Q^T = (P(P^T P)^{-1} P^T)^T = (P^T)^T (P(P^T P)^{-1})^T = P((P^T P)^{-1} g)^T P^T = P(P^T P)^{-1} P^T = Q$   $\square$

**Theorem 19.2.6.** *If  $Q = P(P^T P)^{-1} P^T$  is a projection matrix for a subspace  $\mathcal{W}$ , then  $Q^2 = Q$ .*

*Proof.*  $Q^2 = P(P^T P)^{-1} P^T P(P^T P)^{-1} P^T = P((P^T P)^{-1}(P^T P))(P^T P)^{-1} P^T = P(P^T P)^{-1} P^T = Q$   $\square$

**Theorem 19.2.7.** *If  $Q$  is an  $n \times n$  matrix,  $Q = Q^2$ , and  $Q = Q^T$ , then there is a subspace  $\mathcal{W} \subset \mathbb{R}^n$  such that  $Q$  is the projection matrix of  $\mathcal{W}$ .*

## 19.2.2 Reflections

Let  $\mathcal{W}$  be a plane passing through the origin, and suppose we want to reflect a vector  $v$  across this plane. Let  $u$  be a unit vector along  $\mathcal{W}^\perp$ . That is,  $u$  is normal to the plane. The projection of  $v$  along the line through  $u$  is then given by  $\hat{v} = \text{Proj}_u(v) = u(u^T u)^{-1} u^T v$ . But  $u$  is a unit vector, and therefore  $u^T u = 1$ , so  $\hat{v} = uu^T v$ . Let  $Q_u = uu^T$ . The definition of the reflection of  $v$  across  $\mathcal{W}$  is the vector  $\text{Refl}_{\mathcal{W}}(v)$  such that has the same magnitude as  $v$  lying on the opposite side of  $\mathcal{W}$ . Thus  $v - \text{Refl}_{\mathcal{W}}(v) = 2Q_u v$ , and so we have:

$$\text{Refl}_{\mathcal{W}}(v) = v - 2Q_u v = (I - 2Q_u)v = (I - 2uu^T)v$$

**Definition 19.2.5**  $H_{\mathcal{W}} = I - 2uu^T$  is called the Reflection (Householder) Matrix for  $\mathcal{W}$ .

**Definition 19.2.6** An orthogonal matrix is a matrix  $P$  such that  $P^T P = I$ .

## 19.2.3 Lecture Notes on Orthogonal Matrices

**Definition 19.2.7** An orthogonal matrix is a  $n \times n$  matrix  $A$  such that  $A^T A = I$ .

**Theorem 19.2.8.** *If  $A$  is an orthogonal matrix, then  $A^T = A^{-1}$ .*

*Proof.* For  $A^T A = I$ , and inverses are unique. Thus  $A^T = A^{-1}$ .  $\square$

If we let  $A_i = Ae_i$ , then  $A^T A = (A_i^T A_j) = I$ . Therefore  $A_i^T A_j = \delta_{ij}$ .

**Theorem 19.2.9.** *If  $A$  is an  $n \times n$  real-valued matrix and  $A_i = Ae_i$ ,  $i = 1, 2, \dots, n$ , then  $A$  is orthogonal if and only if  $\{A_1, \dots, A_n\}$  is an orthonormal set of vectors.*

*Proof.* If  $A$  is orthogonal, then  $A_i^T A_j = \delta_{ij}$ , and from this we have orthonormality. If  $\{A_1, \dots, A_n\}$  is orthonormal, then  $A^T A = I$  and is therefore orthogonal.  $\square$

**Theorem 19.2.10.** *The following statements are equivalent:*

1.  $A$  is orthogonal      2.  $\forall_{X \in \mathbb{R}^n}, \|AX\| = \|X\|$  3.  $\forall_{X,Y \in \mathbb{R}^n}, \langle AX, AY \rangle = \langle X, Y \rangle$

*Proof.* We show  $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1$ .

1. If  $A^T A = I$ , then  $\|AX\|^2 = (AX)^T AX = X^T A^T AX = X^T X = \|X\|^2$ . Therefore,  $\|AX\| = \|X\|$ .
2. If  $A$  is a square matrix such that  $\forall_{X \in \mathbb{R}^n}, \|AX\| = \|X\|$ , then:

$$\|X+Y\|^2 = (X+Y)^T(X+Y) = X^T X + X^T Y + Y^T X + Y^T Y = \|X\|^2 + 2X^T Y + \|Y\|^2$$

But:

$$\|A(X+Y)\|^2 = \|AX+AY\|^2 = \|AX\|^2 + 2(AX)^T AY + \|AY\|^2 = \|X\|^2 + 2(AX)^T AY$$

Therefore  $(AX)^T AY = X^T Y$ . That is,  $\langle AX, AY \rangle = \langle X, Y \rangle$ .

3. If  $A$  is a square matrix such that  $\forall_{X,Y \in \mathbb{R}^n}, \langle AX, AY \rangle = \langle X, Y \rangle$ , then  $\langle Ae_i, Ae_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}$ . Therefore,  $A$  is orthogonal.

□

**Theorem 19.2.11.** *If  $A$  and  $B$  are  $n \times n$  orthogonal matrices, then  $AB$  is orthogonal.*

*Proof.* For if  $A^T A = I$  and  $B^T B = I$ , then  $(AB)^T AB = B^T A^T AB = B^T IB = B^T B = I$ .  $AB$  is orthogonal. □

**Theorem 19.2.12.** *If  $A$  is an  $n \times n$  orthogonal matrix, then  $\det(A) \pm 1$ .*

*Proof.* For  $\det(I) = \det(A^T A) = \det(A^T) \det(A) = \det(A)^2$ . Thus,  $\det(A) = \pm 1$ . □

The converse of Thm. 19.2.12 is false. Recall that if  $u \in \mathbb{R}^n$  is a unit vector and  $W = u^\perp$ , then  $H = 2uu^T$  is the reflection matrix for  $W$ . Reflections preserve distance, and therefore  $H$  must be orthogonal.

**Theorem 19.2.13.** *If  $A$  is an  $n \times n$  orthogonal matrix, then there exist  $k$   $n \times n$  reflection matrices  $H_1, \dots, H_k$ ,  $0 \leq k \leq n$ , such that  $A = \prod_{i=1}^k H_i$ .*

*Proof.* By induction. The base case is trivial. Suppose it holds for  $n - 1$ . Let  $z = Ae_n$ , and let  $H$  be the reflection matrix that exchanges  $z$  and  $e_n$ . Then  $HAe_n = Hz = e_n$ , so  $HA$  fixes  $e_n$ . But  $HA$  is an orthogonal matrix, and thus preserves distances and angles. Thus  $HA$  maps  $\mathbb{R}^{n-1}$  onto itself, and thus by induction there are  $H_2, \dots, H_k$  such that  $HA = \prod_{i=2}^k H_i$ . Letting  $H_1 = H$ , we have  $A = HHA = \prod_{i=1}^k H_i$ .  $\square$

**Theorem 19.2.14.** *If  $H$  is a reflection matrix, then  $\det(H) = -1$ .*

**Theorem 19.2.15.** *If  $A$  is an orthogonal matrix and  $A = \prod_{i=1}^k H_i$ , then  $\det(A) = (-1)^k$ .*

If  $A$  is an orthogonal  $2 \times 2$  matrix, then we know that columns must be unit vectors that are also orthogonal (Orthonormal). That is, the two columns must lie on the unit circle about the origin. So we may express the first column as  $(\cos(\theta), \sin(\theta))$  for some angle  $\theta$ . There are then two options for the second column:  $(-\sin(\theta), \cos(\theta))$  or  $(\sin(\theta), -\cos(\theta))$ . The first is the rotation matrix which rotates  $\mathbb{R}^2$  counterclockwise around the origin, and the second is the reflection matrix that makes a reflection across the line that makes an angle  $\frac{\theta}{2}$  with the  $x$ -axis.

**Theorem 19.2.16.** *If  $A$  is a  $3 \times 3$  orthogonal matrix and  $\det(A) = 1$ , then  $A$  is a rotation matrix.*

*Proof.*  $A$  must be the product of 0, 1, 2, or 3 reflection matrices. If  $\det(A) = 1$ , then  $A$  is the product of an even number of reflections, and thus either  $A = I$  or  $A$  is the product of two reflections, and is thus a rotation matrix.  $\square$

**Theorem 19.2.17.** *If  $A$  is a  $3 \times 3$  orthogonal matrix,  $\det(A) = -1$ , and  $A = A^T$ , then either  $A = -I$  or  $A$  is a reflection matrix.*

*Proof.* If  $\det(A) = -1$ , then  $A$  is the product of an odd number of reflections, either 1 or 3. If  $A$  is a single reflection, then  $A = H$  for some Householder matrix  $H$ . Thus  $A^T = A$ . Conversely, if  $A = A^T$  and  $\det(A) = -1$ , then  $\det(-A) = 1$ , and  $-A^T = -A = -A^{-1}$ . Therefore  $-A$  is a rotation whose square is the identity. If  $A \neq I$ , then  $A$  must be a rotation of  $\pi/2$  around some axis, and thus  $A$  is a reflection.  $\square$

**Theorem 19.2.18.** *If  $A$  is a  $3 \times 3$  orthogonal matrix,  $\det(A) = -1$ , and  $A \neq A^T$ , then  $A$  is the product of three reflections.*

*Proof.* If  $\det(A) = -1$ , and  $A \neq A^T$ , then  $A$  is not a rotation or a pure reflection, and is thus a product of 3 reflection matrices.  $\square$

**Theorem 19.2.19.** *If  $A$  and  $B$  are  $3 \times 3$  rotation matrices, then  $AB$  is a rotation matrix.*

*Proof.* For  $A$  and  $B$  must be orthogonal, and thus  $AB$  is orthogonal. But  $\det(AB) = \det(A)\det(B) = 1 \cdot 1 = 1$ , and thus  $AB$  is an orthogonal matrix with determinant equal to 1, and is therefore a rotation matrix.  $\square$

### 19.2.4 Rotations

The  $2 \times 2$  matrix  $A_\theta$  rotates the plane  $\mathbb{R}^2$  counterclockwise by  $\theta$  around the origin. The question that then arises is, “Is there a similar way to do this for  $\mathbb{R}^3$ ?“ The simple case would be rotating by an angle  $\theta$  about the  $z$ -axis, analogous the rotating the Earth by  $\theta$  about the North Pole. This fixes the  $z$ -axis and acts on the  $xy$  plane only. This can be represented by the matrix  $S_\theta$ .

$$A_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad S_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$S_\theta$  is an orthogonal matrix. That is,  $S_\theta S_\theta^T = I$ , and therefore  $S_\theta^T = S_\theta^{-1}$ . Suppose we want to rotate by an angle  $\theta$  about a different axis. Let  $\mathbf{u}$  be a unit vector pointing in the direction of the axis of rotation and let  $R_{\theta,\mathbf{u}}$  be the new rotation matrix. To compute  $R_{\theta,\mathbf{u}}$  choose any unit vector  $\mathbf{v}$  that is orthogonal to  $\mathbf{u}$ . Let  $\mathbf{w} = \mathbf{u} \times \mathbf{v}$ . Then  $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$  is an orthonormal basis of  $\mathbb{R}^3$  such that  $\mathbf{v} \times \mathbf{w} = \mathbf{u}$ . Let:

$$P = \begin{bmatrix} v_1 & w_1 & u_1 \\ v_2 & w_2 & u_2 \\ v_3 & w_3 & u_3 \end{bmatrix}$$

The columns of  $P$  form an orthonormal set, and therefore  $P$  is orthogonal. In particular:

$$P^T \mathbf{v} = e_1 \quad P^T \mathbf{w} = e_2 \quad P^T \mathbf{u} = e_3$$

**Theorem 19.2.20.** *If  $\theta \in [0, 2\pi]$  and  $\mathbf{u} \in \mathbb{R}^3$  is a unit vector, then  $R_{\theta,\mathbf{u}} = PS_\theta P^T$ .*

*Proof.* For  $PS_\theta P^T \mathbf{u} = \mathbf{u}$ ,  $PS_\theta P^T \mathbf{v} = \cos(\theta)\mathbf{v} + \sin(\theta)\mathbf{w}$ , and  $PS_\theta P^T \mathbf{w} = -\sin(\theta)\mathbf{v} + \cos(\theta)\mathbf{w}$ . Thus, if  $X = a\mathbf{v} + b\mathbf{w} + c\mathbf{u}$ , then  $PS_\theta P^T X = a(\cos(\theta)\mathbf{v} + \sin(\theta)\mathbf{w}) + b(-\sin(\theta)\mathbf{v} + \cos(\theta)\mathbf{w}) + c\mathbf{u} = R_{\theta,\mathbf{u}}X$   $\square$

From the orthogonality of  $P$  and  $S_\theta$  we have that  $R_{\theta,\mathbf{u}}$  is also orthogonal.

**Theorem 19.2.21.** *A rotation matrix  $R$  is an orthogonal matrix with determinant 1.*

*Proof.* For  $R^T R = (PS_\theta P^T)^T PS_\theta P^T = PS_\theta^T P^T PS_\theta P^T = PS_\theta^T S_\theta P^T = PP^T = I$ . But also we have  $\det(R) = \det(PS_\theta P^T) = \det(P) \det(S_\theta) \det(P^T) = \det(P) \det(P^{-1}) = 1$   $\square$

The converse of Thm. 19.2.21 is also true. We now turn to the question of how to compute the rotation of  $\mathbb{R}^3$  represented by a given orthogonal matrix. If  $R$  is an orthogonal matrix such that  $\det(R) = 1$ , how do we compute the angle of rotation? First recall that the trace of a matrix is the sum of the diagonal components,  $\text{Tr}(A) = \sum_{i=1}^n a_{ii}$ .

**Theorem 19.2.22.** *If  $A$  and  $B$  are  $n \times n$  matrices, then  $\text{Tr}(AB) = \text{Tr}(BA)$*

**Theorem 19.2.23.** *If  $R$  is a rotation matrix of angle  $\theta$ , then  $\cos(\theta) = \frac{\text{Tr}(R)-1}{2}$ .*

*Proof.* For  $\text{Tr}(R) = \text{Tr}(PS_\theta P^{-1}) = \text{Tr}(PP^{-1}S_\theta) = \text{Tr}(S_\theta) = 1 + 2\cos(\theta) \Rightarrow \cos(\theta) = \frac{\text{Tr}(R)-1}{2}$   $\square$

This doesn't tell us everything, as we still don't know  $\mathbf{u}$ , and  $\cos(\theta) = \cos(-\theta)$ , so we still don't know the sign of  $\theta$ . Since  $R$  is an orthogonal matrix,  $R^T = R^{-1}$ . So if  $\mathbf{u}$  lies on the axis of rotation, then  $(R - R^T)\mathbf{u} = (R - R^{-1})\mathbf{u} = 0$ . We can find the axis of rotation by determining the null space of  $R - R^T$ .

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \Rightarrow R - R^T = \begin{bmatrix} 0 & r_{12} - r_{21} & r_{13} - r_{31} \\ r_{21} - r_{12} & 0 & r_{23} - r_{32} \\ r_{31} - r_{13} & r_{32} - r_{23} & 0 \end{bmatrix} \equiv \begin{bmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{bmatrix}$$

This suggests that  $\mathbf{u}$  is parallel to  $(-\gamma, \beta, -\alpha)^T = (r_{32} - r_{23}, r_{13} - r_{31}, r_{21} - r_{12})^T$ .

**Theorem 19.2.24.** *If  $R$  is a rotation matrix such that  $R \neq R^T$ , then the axis of rotation of  $R$  is parallel to  $\mathbf{q} = (-\gamma, \beta, -\alpha)^T = 2\sin(\theta)\mathbf{u}$ , where  $\mathbf{u}$  is a unit vector about the axis of rotation.*

*Proof.* Let  $R = PS_\theta P^T$ . Then:

$$R - R^T = PS_\theta P^T - PS_\theta^T P^T = P(S_\theta - S_\theta^T)P^T = 2P \begin{bmatrix} 0 & -\sin(\theta) & 0 \\ \sin(\theta) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P^T = 2\sin(\theta)I$$

Where  $\mathbf{v}$  is orthogonal to  $\mathbf{u}$  and  $\mathbf{w} = \mathbf{u} \times \mathbf{v}$ . Thus,  $\mathbf{q} = (-\gamma, \beta, -\alpha)^T = 2\sin(\theta)\mathbf{v} \times \mathbf{w} = 2\sin(\theta)\mathbf{u}$   $\square$

What about the case when  $R - R^T = 0$ ? When this happens either  $\theta = 0$  or  $\theta = \pi$ . If  $\theta = 0$ , then this is the identity rotation and thus  $R = I$ , and we are done. If  $R \neq I$ , then  $\theta = \pi$ . To find out the axis of rotation, we have that:

$$R = PS_\pi P^T = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = -\mathbf{v}\mathbf{v}^T - \mathbf{w}\mathbf{w}^T + \mathbf{u}\mathbf{u}^T$$

But  $\mathbf{v}$ ,  $\mathbf{w}$ , and  $\mathbf{u}$  form an orthonormal basis, and therefore  $\mathbf{v}\mathbf{v}^T + \mathbf{w}\mathbf{w}^T + \mathbf{u}\mathbf{u}^T = I$ . Thus,  $R = -I + 2\mathbf{u}\mathbf{u}^T$ , so  $\mathbf{u}\mathbf{u}^T = \frac{1}{2}(R + I)$ . But the columns of  $\mathbf{u}\mathbf{u}^T$  are parallel to  $\mathbf{u}$ , and therefore we can determine  $\mathbf{u}$  by normalizing one of the columns of  $\frac{1}{2}(R + I)$ .

### 19.2.5 The Matrix Exponential

**Definition 19.2.8** If  $A$  is an  $n \times n$  matrix, then the exponential of  $A$  is  $e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$ .

Notationally, we write  $A^0 = I$ . For any complex-valued matrix  $A$  of finite dimension, it can be shown that this sum converges.

**Theorem 19.2.25.** If  $A$  and  $P$  are complex  $n \times n$  matrices and  $P$  invertible, then  $e^{P^{-1}AP} = P^{-1}e^A P$ .

*Proof.* For all  $m \in \mathbb{N}$ ,  $(P^{-1}AP)^m = P^{-1}A^mP$ . Thus:

$$e^{P^{-1}AP} = \sum_{k=0}^{\infty} P^{-1} \frac{A^k}{k!} P = P^{-1} \left( \sum_{k=0}^{\infty} \frac{A^k}{k!} \right) P = P^{-1} e^A P$$

□

**Theorem 19.2.26.** If  $0$  is the zero matrix, then  $e^0 = I$ .

**Theorem 19.2.27.** If  $A$  is an  $n \times n$  matrix and  $m \in \mathbb{N}$ , then  $A^m e^A = e^A A^m$ .

*Proof.* For  $A^m e^A = A^m \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \frac{A^{k+m}}{k!} = \left( \sum_{k=0}^{\infty} \frac{A^k}{k!} \right) A^m$ . □

**Theorem 19.2.28.** If  $A$  is an  $n \times n$  matrix, then  $e^{A^T} = (e^A)^T$ .

*Proof.* For  $e^{A^T} = \sum_{k=0}^{\infty} \frac{(A^T)^k}{k!} = \sum_{k=0}^{\infty} \frac{(A^k)^T}{k!} = \left( \sum_{k=0}^{\infty} \frac{A^k}{k!} \right)^T = (e^A)^T$ . □

**Theorem 19.2.29.** If  $A$  and  $B$  are  $n \times n$  matrices and if  $AB = BA$ , then  $Ae^B = e^B A$ .

*Proof.* For  $Ae^B = A \sum_{k=0}^{\infty} \frac{B^k}{k!} = \sum_{k=0}^{\infty} A \frac{B^k}{k!} = \sum_{k=0}^{\infty} \frac{B^k}{k!} A = \left( \sum_{k=0}^{\infty} \frac{B^k}{k!} \right) A = e^B A$ . □

**Theorem 19.2.30.** If  $A$  and  $B$  are  $n \times n$  matrices and  $AB = BA$ , then  $e^A e^B = e^B e^A$ .

*Proof.* For:

$$\begin{aligned} e^A e^B &= e^A \sum_{k=0}^{\infty} \frac{B^k}{k!} = \sum_{k=0}^{\infty} e^A \frac{B^k}{k!} = \sum_{k=0}^{\infty} \left( \sum_{j=0}^{\infty} \frac{A^j}{j!} \right) \frac{B^k}{k!} = \sum_{k=0}^{\infty} \left( \sum_{j=0}^{\infty} \frac{A^j}{j!} \frac{B^k}{k!} \right) \\ &= \sum_{k=0}^{\infty} \left( \sum_{j=0}^{\infty} \frac{B^k}{k!} \frac{A^j}{j!} \right) = \sum_{k=0}^{\infty} \left( \sum_{j=0}^{\infty} \frac{B^k}{k!} \right) \frac{A^j}{j!} = \sum_{k=0}^{\infty} \frac{B^k}{k!} \sum_{j=0}^{\infty} \frac{A^j}{j!} = e^B e^A \end{aligned}$$

□

It is NOT true that  $e^{A+B} = e^A e^B$ , in general. Matrix exponentiation lacks this feature.

**Theorem 19.2.31.** *If  $A$  is an  $n \times n$  matrix and  $s, t \in \mathbb{C}$ , then  $e^{A(s+t)} = e^{As} e^{At}$ .*

*Proof.* For  $e^{As} e^{At} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{A^{j+k} s^j t^k}{j! k!}$ . Letting  $n = j + k$ , so  $j = n - k$ , we have:

$$\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{A^n}{n!} \frac{n!}{k!(n-k)!} s^{n-k} t^k = \sum_{n=0}^{\infty} \frac{A^n}{n!} \left( \sum_{k=0}^{\infty} \frac{n!}{k!(n-k)!} s^{n-k} t^k \right)$$

From the binomial theorem, the expression inside the parenthesis is equal to  $(s+t)^n$ . So we have  $e^{As} e^{At} = \sum_{n=0}^{\infty} \frac{A^n (t+s)^n}{n!} = e^{A(s+t)}$ .  $\square$

**Theorem 19.2.32.** *If  $A$  is an  $n \times n$  matrix, then  $e^A$  is invertible and  $(e^A)^{-1} = e^{-A}$ .*

*Proof.* For  $I = e^0 = e^{A(1-1)} = e^A e^{-A}$ . Thus  $(e^A)^{-1} = e^{-A}$ .  $\square$

**Theorem 19.2.33.** *If  $A$  is an  $n \times n$  matrix and  $t \in \mathbb{R}$ , then  $\frac{d}{dt}(e^{At}) = Ae^{At}$ .*

*Proof.* For  $\lim_{h \rightarrow 0} \frac{e^{A(t+h)} - e^{At}}{h} = e^{At} \lim_{h \rightarrow 0} \frac{e^{Ah} - I}{h} = e^{At} \lim_{h \rightarrow 0} [A + \frac{A^2}{2!} h + \dots] = e^{At} A = Ae^{At}$ .  $\square$

**Theorem 19.2.34.** *If  $A$  and  $B$  are  $n \times n$  matrices and  $AB = BA$ , then  $e^{A+B} = e^A e^B$ .*

*Proof.* For let  $g(t) = e^{(A+B)t} e^{-Bt} e^{-At}$ . Then from commutativity of  $A$  and  $B$ ,  $g'(t) = 0$ . But then  $g(t)$  is a constant. From the definition,  $g(0) = I$ , and thus  $g(t) = I$ . So  $e^{(A+B)t} e^{-Bt} e^{-At} = I$ , and therefore  $e^{(A+B)t} = e^{At} e^{Bt}$ .  $\square$

**Theorem 19.2.35.** *If  $A^2 = 0$ , then  $e^A = I + A$ .*

*Proof.* For  $e^A = I + A + A^2(\frac{I}{2!} + \frac{A}{3!} + \dots) = I + A + 0 = I + A$ .  $\square$

## 19.2.6 Linear Systems of Ordinary Differential Equations

Consider the equation  $y' = ky$ , where  $k$  is some constant. We can solve this via calculus using separation of variables:

$$\frac{y'}{y} = k \Rightarrow \int \frac{y'}{y} dx = \int k dx \Rightarrow \ln(y) = kx + c \Rightarrow y = e^c e^{kx}$$

Setting  $x = 0$ , we have  $e^c = y_0$ . So  $y = y_0 e^{kx}$ . Let us solve this a different way: Let  $F(x) = e^{-kx} y$ , and let  $y' = kx$ . Differentiating we have:

$$F'(x) = -ke^{kx} y + e^{-kx} y' = -kye^{-kx} + e^{-kx} ky = 0$$

So  $F'(x) = 0$ , and therefore  $F(x)$  is a constant. Setting  $x = 0$ , we have  $F(x) = y_0$ . So  $y = y_0 e^{kx}$ . This shows us that  $y_0 e^{kx}$  is the *only* solution to this problem. Let:

$$Y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

Consider  $Y'(t) = AY(t)$ , where  $A$  is an  $n \times n$  matrix. Let  $F(t) = e^{-At}Y(t)$ . Then  $F'(t) = 0$ , and  $Y(t) = Y_0 e^{At}$ .

**Theorem 19.2.36.** *If  $Y : \mathbb{R} \rightarrow \mathbb{R}^n$  is a differentiable function such that  $Y'(t) = AY(t)$ , where  $A$  is a diagonalizable matrix with eigenvalues  $\lambda_1, \dots, \lambda_n$  and eigenvectors  $v_1, \dots, v_n$ , then  $Y(t) = \sum_{k=1}^n \lambda_k e^{\lambda_k t} v_k$*

## 19.3 Problem Sets

### 19.3.1 Problem Set I

**Problem 19.3.1** Find the point on the line  $y = 4x$  which is closest to the point  $(2, 5)$ .

*Solution 1.* Given a vector  $\mathbf{v}$  that is parallel to the line  $y$ , we know that the vector  $\mathbf{w}$  from  $(2, 5)$  to the point  $(x, y)$  that minimizes the distance from  $y = 4x$  to the point  $(2, 5)$  will satisfy  $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ . That is:

$$\langle (1, 4), (2 - x, 5 - y) \rangle = 0 \Rightarrow 2 - x + 4(5 - y) = 0 \Rightarrow 22 - x - 4y = 0$$

But  $y = 4x$ , and thus  $22 - 17x = 0 \Rightarrow x = \frac{22}{17}$ . The point of least distance is  $\frac{22}{17}(1, 4)$ .  $\square$

*Solution 2.* This point is the projection of the vector  $(2, 5)^T$  onto  $(1, 4)^T$ . That is:

$$\mathbf{P} = \frac{\begin{bmatrix} 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix}}{\begin{bmatrix} 1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix}} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \frac{22}{17} \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$\square$

**Problem 19.3.2** Show that  $\mathbf{x}\mathbf{y}^T + \mathbf{y}\mathbf{x}^T$  is symmetric.

*Solution.* Recall that a matrix is symmetric if it is equal to its transpose. Thus, we must show  $A = A^T$ . But for any  $n \times n$  matrices  $A$  and  $B$ ,  $(A + B)^T = A^T + B^T$ , and  $(AB)^T = B^T A^T$ , and  $(A^T)^T = A$ . Thus, given our matrix  $A = \mathbf{x}\mathbf{y}^T + \mathbf{y}\mathbf{x}^T$ , we have that  $A^T = (\mathbf{x}\mathbf{y}^T + \mathbf{y}\mathbf{x}^T)^T = (\mathbf{x}\mathbf{y}^T)^T + (\mathbf{y}\mathbf{x}^T)^T = (\mathbf{y}^T)^T \mathbf{x}^T + (\mathbf{x}^T)^T \mathbf{y}^T = \mathbf{y}\mathbf{x}^T + \mathbf{x}\mathbf{y}^T = \mathbf{x}\mathbf{y}^T + \mathbf{y}\mathbf{x}^T = A$   $\square$

**Problem 19.3.3** Compute the product  $\begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 & 3 & 1 \\ 2 & -2 & 1 & -1 \end{bmatrix}$

*Solution.*

$$\begin{aligned} \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 & 3 & 1 \\ 2 & -2 & 1 & -1 \end{bmatrix} &= \begin{bmatrix} 2(-1) + (-1)2 & 2 \cdot 2 + (-1)(-2) & 2 \cdot 3 + (-1)1 & 2 \cdot 1 \\ 3(-1) + 1 \cdot 2 & 3 \cdot 2 + 1(-2) & 3 \cdot 3 + 1 \cdot 1 & 3 \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} -4 & 6 & 5 & 3 \\ -1 & 4 & 10 & 2 \end{bmatrix} \end{aligned}$$

□

**Problem 19.3.4** Find the equation of the plane that contains  $P_1(2, 2, 1)$ ,  $P_2(2, 3, 2)$ , and  $P_3(-1, 3, 1)$ .

*Solution.* It suffices to find a vector normal to this plane. We have that:

$$\overrightarrow{P_1P_2} = (0, 1, 1)^T \quad \overrightarrow{P_1P_3} = (-3, 1, 0)^T$$

Then both vectors are parallel to the plane, and thus  $\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3} = (-1, 3, 3)^T$  is perpendicular to the plane. Suppose  $Q = (x, y, z)$  is a point in the plane. Then the relative position vector  $P_1Q = (x - 2, y - 2, z - 1)^T$  is orthogonal to  $(-1, 3, 3)^T$ . Thus:

$$\begin{aligned} (x - 2, y - 2, z - 1)(-1, 3, 3)^T &= 0 \\ \Rightarrow 2 - x + 3y - 6 + 3z - 3 &= 0 \\ \Rightarrow x - 3y - 3z + 7 &= 0 \end{aligned}$$

This is the equation of the plane. □

**Problem 19.3.5** Let  $S = \text{Span}(\mathbf{x}_1, \mathbf{x}_2)$ , where  $\mathbf{x}_1 = (1, -1, 2)^T$ ,  $\mathbf{x}_2 = (-1, 2, 2)^T$ . Find a basis for  $S^\perp$

*Solution.* We seek a vector in  $\mathbf{x}_3 \in \mathbb{R}^3$  such that  $\langle \mathbf{x}_3, \mathbf{x}_i \rangle = 0$ ,  $i = 1, 2$ . That is:

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

Solving gives us  $x_2 = -4x_3$ ,  $x_1 = -6x_3$ .  $\{(-6, -4, 1)\}$  is a basis. □

**Problem 19.3.6** For the matrix  $A = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -1 & 0 \end{bmatrix}$ , find a basis for the following:

$$1. \ R(A^T) \quad 2. \ N(A) \quad 3. \ R(A) \quad 4. \ N(A^T)$$

*Solution.* The row-echelon form of  $A$  and  $A^T$  are given below:

$$A' = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \quad (A^T)' = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}$$

1. The rows of  $A'$  give us a basis for  $R(A^T)$  of  $\{(1, 1, 0), (0, 1, 2)\}$
2.  $N(A) = \{x \in \mathbb{R}^3 : Ax = 0\}$ . Solving  $A'x = 0$  gives us a basis of  $\{(2, -2, 1)\}$
3. The non-zero rows of  $(A^T)'$  give us a basis of  $\{(1, 0), (0, -1)\}$ .
4.  $N(A^T) = \{x \in \mathbb{R}^2 : A^Tx = 0\}$ .  $A'x = 0$  gives us  $x_1 = 0$  and  $-x_2 = 0$ .  $N(A^T) = \{(0, 0)\}$ .

□

### 19.3.2 Problem Set II

**Problem 19.3.7** Find a point on the line  $y = 5x$  that is closest to the point  $(1, 3)$ .

*Solution.* Pick a point on the line, say  $\mathbf{w} = (1, 5)^T$ . The point  $P$  is the projection of  $\mathbf{v} = (1, 3)^T$  onto the line  $y = 5x$ , and thus:

$$P = \frac{\mathbf{v}^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}} = \frac{\begin{bmatrix} 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix}}{\begin{bmatrix} 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix}} (1, 5)^T = \frac{8}{13} (1, 5)^T$$

□

**Problem 19.3.8** Is  $A = xy^T - yx^T$  symmetric? ( $x$  and  $y$  are  $n \times 1$  vectors)

*Solution.* In general, no. For if it were, then  $A - A^T = 0$ . But:

$$\begin{aligned} 0 &= A - A^T = xy^T - yx^T - (xy^T - yx^T)^T = xy^T - yx^T - [(xy^T)^T - (yx^T)^T] \\ &= xy^T - yx^T - [yx^T - xy^T] = 2xy^T - 2yx^T = 2A \Rightarrow xy^T - yx^T = 0 \Rightarrow xy^T = yx^T \end{aligned}$$

As this is not, in general, true,  $A$  is not necessarily symmetric. □

**Problem 19.3.9** Compute the product  $\begin{bmatrix} -1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 & 2 & -2 \\ 2 & 3 & 1 & 1 \end{bmatrix}$

*Solution.*

$$\begin{bmatrix} -1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 & 2 & -2 \\ 2 & 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1+6 & -1+9 & -2+3 & 2+3 \\ -4+4 & 4+6 & 8+2 & -8+2 \end{bmatrix} = \begin{bmatrix} 7 & 8 & 1 \\ 0 & 10 & 10 \end{bmatrix}$$

□

**Problem 19.3.10** Find the equation of the plane that passes through  $P_1(2, 2, 2)$ ,  $P_2(2, 3, 1)$ .

*Solution.*  $\overrightarrow{P_1P_2} = (0, 1, 2)^T$ ,  $\overrightarrow{P_1P_3} = (-3, 1, 1)^T$ . So:

$$\vec{N} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 0 & 1 & 2 \\ -3 & 1 & 1 \end{vmatrix} = \hat{\mathbf{i}}(1-2) + \hat{\mathbf{j}}(0+6) + \hat{\mathbf{k}}(0+3) = \begin{bmatrix} -1 \\ -6 \\ 3 \end{bmatrix}$$

For a point  $P = (x, y, z)$  in the plane,  $\langle \overrightarrow{P_1P}, \vec{N} \rangle = 0$ . Thus,  $x+6y-3z=0$  □

**Problem 19.3.11** Let  $S = \text{Span}(\{(2, 1, 2)^T, (-2, -1, 3)^T\})$ . Find a basis for  $S^\perp$ .

*Solution.* Let  $A$  and it's row-echelon form be the matrices shown below. Then  $S^\perp = N(A)$ .

$$A = \begin{bmatrix} 2 & 1 & 2 \\ -2 & -1 & 3 \end{bmatrix} \quad A' = \begin{bmatrix} 2 & 1 & 2 \\ 0 & 0 & 5 \end{bmatrix}$$

Solving for  $A'x = 0$  gives us a basis of  $\{(1, -2, 0)\}$

□

**Problem 19.3.12** For the matrix  $A = \begin{bmatrix} 2 & 3 & 4 \\ -2 & -2 & 0 \end{bmatrix}$ , find a basis for the following:

1.  $R(A^T)$
2.  $N(A)$
3.  $R(A)$
4.  $N(A^T)$

*Solution.*  $A$  and  $A^T$  have the following row-echelon forms:

$$A' = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 4 \end{bmatrix} \quad (A^T)' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

1. Putting  $A$  into row-echelon form and reading off the rows, we obtain the basis  $\{(1, 1, 0), (0, 1, 4)\}$
2.  $N(A) = \{x \in \mathbb{R}^3 : Ax = 0\}$ . This gives us a basis of  $\{(4, -4, 1)\}$
3. The non-zero rows of  $(A^T)'$  give us a basis of  $\{(1, 0), (0, 1)\}$
4.  $N(A^T) = \{x \in \mathbb{R}^2 : A^Tx = 0\}$ . Solving  $A'^Tx = 0$  gives us  $x_1 = 0$  and  $x_2 = 0$ .  $N(A^T) = \{(0, 0)\}$

□

### 19.3.3 Problem Set III

**Problem 19.3.13** Let  $A, B, C$  be  $n \times n$  matrices. Is  $A = BC^T + CB^T$  symmetric?

*Solution.* A matrix is symmetric if  $A = A^T$ . If  $A = BC^T + CB^T$ , then:

$$A^T = (BC^T + CB^T)^T = (BC^T)^T + (CB^T)^T = (C^T)^T B^T + (B^T)^T C^T = CB^T + BC^T = A$$

$A$  is symmetric. □

**Problem 19.3.14** Compute  $\|x\|_1, \|x\|_2, \|x\|_3$  for  $x = (2, -3, 1)^T$

*Solution.* By definition, for  $x \in \mathbb{R}^n$ ,  $\|x\|_p = (\sum_{k=1}^n |x_k|^p)^{1/p}$ . So we have the following:

1.  $\|x\|_1 = |2| + |-3| + |1| = 2 + 3 + 1 = 6$
2.  $\|x\|_2 = (\lvert 2 \rvert^2 + \lvert -3 \rvert^2 + \lvert 1 \rvert^2)^{1/2} = (4 + 9 + 1)^{1/2} = \sqrt{14}$
3.  $\|x\|_3 = (\lvert 2 \rvert^3 + \lvert -3 \rvert^3 + \lvert 1 \rvert^3)^{1/3} = (8 + 27 + 1)^{1/3} = \sqrt[3]{36}$

□

**Problem 19.3.15** For the matrix  $A = \begin{bmatrix} 2 & -2 & 4 \\ -1 & 1 & -2 \end{bmatrix}$ , find a basis for the following:

1.  $R(A^T)$
2.  $N(A)$
3.  $R(A)$
4.  $N(A^T)$

*Solution.*  $A$  and  $A^T$  have the following row-echelon forms:

$$A' = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad (A^T)' = \begin{bmatrix} -2 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

1. The non-zero rows of  $A'$  give a basis of  $\{(1, -1, 2)\}$
2.  $N(A) = \{x \in \mathbb{R}^3 : Ax = 0\}$ . Solving  $A'x = 0$  gives a basis of  $\{(1, 1, 0), (-2, 0, 1)\}$
3. The non-zero rows of  $(A^T)'$  give a basis of  $\{(-2, 1)\}$
4.  $N(A^T) = \{x \in \mathbb{R}^2 : A^Tx = 0\}$ . Solving  $(A^T)'x = 0$  gives a basis of  $\{(1, 2)\}$

□

**Problem 19.3.16** Find the least-squares solution to the following system:

$$\begin{aligned} x_1 - x_2 &= 2 \\ x_1 + x_2 &= 0 \\ x_1 + 2x_2 &= -1 \end{aligned}$$

*Solution.* We want the solution to  $A^T A x = A^T b$ . We have:

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad b = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \quad A^T A x = A^T b \Rightarrow \begin{bmatrix} 3 & 1 \\ 1 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -6 \end{bmatrix}$$

The solution is  $x = \frac{1}{26}(15, -19)^T$

□

**Problem 19.3.17** Let  $\theta \in \mathbb{R}$  and let  $\mathbf{x}_1 = (\cos(\theta), \sin(\theta))^T$ ,  $\mathbf{x}_2 = (-\sin(\theta), \cos(\theta))^T$ . Show that  $\{\mathbf{x}_1, \mathbf{x}_2\}$  is an orthonormal basis for  $\mathbb{R}^2$ . Write  $\mathbf{y} = (-2, 3)^T$  as a linear combination  $\mathbf{y} = c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2$

*Solution.* They are orthonormal for  $\mathbf{x}_1^T \mathbf{x}_2 = -\cos(\theta)\sin(\theta) + \cos(\theta)\sin(\theta) = 0$ , and since  $\|\mathbf{x}_1\| = \|\mathbf{x}_2\| = (\sin^2(\theta) + \cos^2(\theta))^{1/2} = 1$ . Let  $c_1 = \langle \mathbf{y}, \mathbf{x}_1 \rangle$  and  $c_2 = \langle \mathbf{y}, \mathbf{x}_2 \rangle$ . Then  $c_1 = -2\cos(\theta) + 3\sin(\theta)$  and  $c_2 = 2\sin(\theta) + 3\cos(\theta)$ . Therefore,  $\mathbf{y} = (-2\cos(\theta) + 3\sin(\theta))\mathbf{x}_1 + (2\sin(\theta) + 3\cos(\theta))\mathbf{x}_2$

□

### 19.3.4 Problem Set IV

**Problem 19.3.18** Find the eigenvalues and associated eigenspaces of  $A = \begin{bmatrix} 4 & 5 \\ 2 & 1 \end{bmatrix}$

*Solution.* We need to compute  $\det(A - \lambda I) = 0$ . This gives us:

$$\begin{vmatrix} 4 - \lambda & 5 \\ 2 & 1 - \lambda \end{vmatrix} = (4 - \lambda)(1 - \lambda) - 10 = 0$$

The solutions to this are  $\lambda_1 = 6, \lambda_2 = -1$ . Solving  $Ax = \lambda x$  yields the eigenspaces. We have:

$$\begin{bmatrix} 4 & 5 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = - \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \begin{bmatrix} 4 & 5 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 6 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

These give solutions  $x_2(-1, 1)^T$  and  $x_2(\frac{5}{2}, 1)^T$ , where  $x_2$  is a free variable.  $\square$

**Problem 19.3.19** Show that for a  $2 \times 2$  matrix  $A$ ,  $\lambda^2 - \text{Tr}(A)\lambda + \det(A) = 0$ , where  $\lambda$  is an eigenvalue of  $A$ .

*Solution.* For we have that  $\det(A - \lambda I) = 0$ . But:

$$\det(A - \lambda I) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a+d)\lambda + ad - bc = \lambda^2 - \text{Tr}(A)\lambda + \det(A)$$

Therefore, if  $\lambda$  is an eigenvalue of  $A$ , then  $\lambda^2 - \text{Tr}(A)\lambda + \det(A) = 0$ .  $\square$

**Problem 19.3.20** Find the eigenvalues and associated eigenspaces for  $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$

*Solution.* Recall that the determinant expansion can be done along any row. Thus:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1 - \lambda & 1 & 1 \\ 0 & 2 - \lambda & 1 \\ 0 & 0 & 3 - \lambda \end{vmatrix} = 0 \begin{vmatrix} 1 & 1 \\ 2 - \lambda & 1 \end{vmatrix} - 0 \begin{vmatrix} 1 - \lambda & 1 \\ 0 & 1 \end{vmatrix} + (3 - \lambda) \begin{vmatrix} 1 - \lambda & 1 \\ 0 & 2 - \lambda \end{vmatrix} \\ &= (3 - \lambda)(1 - \lambda)(2 - \lambda) \end{aligned}$$

The solutions are  $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$ . The eigenspaces correspond to the solutions of the equation  $Ax = \lambda x$ . Thus we get:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

This gives 3 different equations for each value of  $\lambda$ .

$$Ax = x \Rightarrow x = (1, 0, 0)^T$$

$$Ax = 2x \Rightarrow x = (1, 1, 0)^T$$

$$Ax = 3x \Rightarrow x = (1, 2, 1)^T$$

□

### 19.3.5 Problem Set V

**Problem 19.3.21** Factor  $\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$  into the form  $PDP^T$ , where  $D$  is a diagonal and  $P$  is orthogonal.

*Solution.* The eigenvalues of  $A$  are the solutions to  $(4 - \lambda)(1 - \lambda) - 4 = 0$ :  $\lambda_1 = 0$ ,  $\lambda_2 = 5$ . The eigenvectors are solutions to:

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$$

Which gives us  $\frac{1}{\sqrt{5}}(2, 1)^T$  and  $\frac{1}{\sqrt{5}}(-1, 2)^T$ . Thus:

$$P = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} \quad P^T = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$$

□

**Problem 19.3.22** Solve the differential equation  $Y'(t) = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} Y(t)$  with  $Y(0) = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$

*Solution.* We know from the previous problem that the eigenvalues and eigenvectors are distinct, and thus  $Y(t) = \alpha V_1 e^{\lambda_1 t} + \beta V_2 e^{\lambda_2 t}$  where  $\lambda_i$  are the distinct eigenvalues, and  $V_i$  are the distinct eigenvectors. Solving for the initial condition:

$$\frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \Rightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 9 \\ 2 \end{bmatrix}$$

$$\text{Thus, } Y(t) = \frac{9}{5}(-1, 2)^T + \frac{2}{5}(2, 1)^T e^{5t}$$

□

**Problem 19.3.23** Solve the following:

1. Let  $A$  be an  $n \times n$  complex Hermitian matrix such that  $A^4 = I$ . What are the possible eigenvalues of  $A$ ?
2. If  $A$  is an  $n \times n$  complex matrix and  $A^4 = I$ , what are the possible eigenvalues?

**Problem 19.3.24** Using least squares, find the line in  $\mathbb{R}^2$  that best fits  $\{(2, 1), (3, 2), (4, 3)\}$ .

*Solution.* We want a line  $y = mx + b$  that best fits the points. Setting up the problem, we get:

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix}$$

This has no solution. Let  $A$  be the left-most matrix. Then:

$$A^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 5 \end{bmatrix} \Rightarrow A^T A = \begin{bmatrix} 4 & 14 \\ 14 & 54 \end{bmatrix}$$

We now solve  $A^T AX$ :

$$\begin{bmatrix} 4 & 14 \\ 14 & 54 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = A^T \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 8 \\ 31 \end{bmatrix}$$

The solution gives us  $y = \frac{3}{5}x - \frac{1}{10}$

□

**Problem 19.3.25** Find the projection matrix  $P$  that projects  $\mathbb{R}^4$  onto the line through the origin spanned by the vector  $(2, 1, -1, -1)$ .

**Problem 19.3.26** Consider the rotation matrix  $R$  shown below. Compute the axis vector  $\mathbf{u}$  and both the sine and cosine of the counterclockwise angle  $\theta$  such that  $R = R_{\theta, \mathbf{u}}$

$$R = \begin{bmatrix} -\frac{4}{9} & -\frac{7}{9} & \frac{4}{9} \\ \frac{1}{9} & \frac{4}{9} & \frac{8}{9} \\ -\frac{8}{9} & \frac{4}{9} & -\frac{1}{9} \end{bmatrix}$$

**Problem 19.3.27** Find an orthonormal basis for the column space of the matrix:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 1 \\ 2 & 2 & 2 \\ 2 & 4 & 3 \\ -1 & 2 & 0 \end{bmatrix}$$

*Solution.* We use Gram-Schmidt to do this. Let  $v_1 = (1, 0, 2, 2, -1)$ . Normalizing gives us:

$$e_1 = \frac{1}{\sqrt{10}}(1, 0, 2, 2, -1)^T$$

We then compute:

$$(1, 3, 2, 4, 2)^T - \frac{(1, 3, 2, 4, 2)^T (1, 0, 2, 2, -1)}{(1, 0, 2, 2, -1)^T (1, 0, 2, 2, -1)} (1, 0, 2, 2, -1)^T = (1, 3, 2, 4, 2)^T - \frac{11}{10} (1, 0, 2, 2, -1)^T \\ = \left(-\frac{1}{10}, 3, -\frac{2}{10}, \frac{18}{10}, \frac{33}{10}\right)^T = \frac{1}{10} (-1, 30, -2, 18, 33)$$

Thus:

$$e_2 = \frac{\frac{1}{10} (-1, 30, -2, 18, 33)}{\left\| \frac{1}{10} (-1, 30, -2, 18, 33) \right\|} = \frac{1}{\sqrt{2318}} (-1, 30, -2, 18, 33)$$

Finishing off, we compute:

$$\mathbf{v}_3 = (1, 1, 2, 3, 0)^T - \frac{(1, 1, 2, 3, 0)^T (1, 0, 2, 2, -1)}{10} (1, 0, 2, 2, -1)^T - \frac{(1, 1, 2, 3, 0)^T (1, 3, 2, 4, 2)}{34} (1, 3, 2, 4, 2, 0)^T$$

Finally,  $e_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|}$

□

**Problem 19.3.28** Eliminate crossterms and classify the conic section  $6x^2 - 4xy + 3y^2 = 1$

### 19.3.6 Problem Set VI

**Problem 19.3.29** Let  $\left[ \begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & -2 & 3 \\ 1 & 2 & 0 & 0 \end{array} \right]$  be an augmented matrix.

1. Solve the system using Gaussian elimination.
2. Express  $(1, 3, 0)^T$  as a linear combination of the column vectors of the coefficient matrix.
3. Use elementary matrices to find the LU decomposition of the coefficient matrix.

*Solution.* In order,

1.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & -2 & 3 \\ 1 & 2 & 0 & 0 \end{array} \right] \xrightarrow{r_2 \leftrightarrow r_3} \left[ \begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & -2 & 3 \end{array} \right] \xrightarrow{r_2 - r_1} \left[ \begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 1 & -2 & 3 \end{array} \right] \\ \xrightarrow{r_2 \div 2} \left[ \begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 1 & -2 & 3 \end{array} \right] \xrightarrow{r_3 - r_2} \left[ \begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{7}{2} \end{array} \right] \\ \xrightarrow{r_3 \cdot (-2)} \left[ \begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -7 \end{array} \right] \xrightarrow{r_1 - 3r_3} \left[ \begin{array}{ccc|c} 1 & 0 & 0 & 22 \\ 0 & 1 & 0 & -11 \\ 0 & 0 & 1 & -7 \end{array} \right]$$

2.  $(1, 3, 0)^T = 22(1, 0, 1)^T - 11(0, 1, 2)^T - 7(3, -2, 0)^T$

3.

$$A = \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{array} \right] \left[ \begin{array}{ccc} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array} \right]$$

□

**Problem 19.3.30** Let  $A = \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{array} \right]$ ,  $B = \left[ \begin{array}{ccc} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -4 & 1 \end{array} \right]$ , and  $C = \left[ \begin{array}{cc} 2 & 3 \\ -1 & 0 \\ 1 & 1 \end{array} \right]$ .

1. Solve  $AC + BC$       2. Solve  $AB$       3. Does  $A = B^{-1}$ ?

*Solution.* In order,

$$1. AC + BC = (A + B)C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 8 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ -1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ -2 & 0 \\ 18 & 26 \end{bmatrix}$$

$$2. AB = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -15 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3. No, for if  $A = B^{-1}$  then  $AB = I$ , but this is not true.

□

**Problem 19.3.31** If  $A$  and  $B$  are  $n \times n$  invertible matrices, what is  $(AB)^{-1}$ ?

*Solution.* As  $A^{-1}$  and  $B^{-1}$  exist, and as  $A$  and  $B$  are of the same dimension,  $B^{-1}A^{-1}$  exists. But  $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$ . As inverses are unique,  $(AB)^{-1} = B^{-1}A^{-1}$ . □

**Problem 19.3.32** If  $A$  and  $B$  are  $n \times n$  matrices, what is  $(A + B)^2$ ?

*Solution.*  $(A + B)^2 = (A + B)(A + B) = A(A + B) + B(A + B) = A^2 + AB + BA + B^2$ . Note: It is not true in general that  $AB = BA$ , and thus we cannot simplify further. □

**Problem 19.3.33** If  $A$  and  $A^T$  are  $n \times n$  invertible matrices, show that  $(A^T)^{-1} = (A^{-1})^T$

*Solution.* For  $A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$ . As inverses are unique,  $(A^T)^{-1} = (A^{-1})^T$ . □

**Problem 19.3.34** What are the solutions of:

$$1. \left[ \begin{array}{cccc|c} 1 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 \end{array} \right] \quad 2. \left[ \begin{array}{cccc|c} 1 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array} \right]$$

*Solution.* In order,

- No solution as the bottom two rows say  $x_3 + x_4 = 2$  and  $x_3 + x_4 = 1$ . An impossibility.
- The entire space  $S = \{(-4, 3, x, 1 - x) : x \in \mathbb{R}\}$ .

□

**Problem 19.3.35** If  $A, B$ , and  $C$  are  $n \times n$  invertible matrices, then solve the following equations for  $X$ :

$$1. \quad XA + B = C \quad 2. \quad AX + B = X \quad 3. \quad XA + C = X$$

*Proof.* In order,

1.  $XA + B = C \Rightarrow XA = C - B \Rightarrow X = (C - B)A^{-1}$
2.  $AX + B = X \Rightarrow AX - X = -B \Rightarrow (A - I)X = -B \Rightarrow X = -(A - I)^{-1}B$
3.  $XA + C = X \Rightarrow XA - X = -C \Rightarrow X(A - I) = -C \Rightarrow X = -C(A - I)^{-1}$

□

### 19.3.7 Problem Set VII

**Problem 19.3.36** Determine the basis of the given vector space over the given field.

1.  $V = \mathbb{R}$  over  $K = \mathbb{R}$
2.  $V = \mathbb{C}$  over  $K = \mathbb{C}$
3.  $V = \mathbb{C}$  over  $K = \mathbb{R}$

*Solution.* In order,

1. The set  $\{1\}$  is a basis. Let  $r \in \mathbb{R}$ . Then  $r = 1 \cdot r$ .
2. The set  $\{(1, 0)\}$  is a basis. Let  $z \in \mathbb{Z}$ . Then  $z \cdot (1, 0) = z$
3. The set  $\{(1, 0), (0, 1)\}$  is a basis. Let  $z = a + bi \in \mathbb{Z}$ . Then  $z = a(1, 0) + b(0, 1)$ .

□

**Problem 19.3.37** What is the nullspace of an  $n \times n$  matrix  $A$  with real entries?

*Solution.* The nullspace is the set  $N(A) = \{X \in \mathbb{R}^n : AX = 0\}$

□

**Problem 19.3.38** A matrix  $A$  and its row reduced form  $A'$  are shown below. What is the rank of  $A$ ?

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & -1 & -4 & -2 \\ 3 & 4 & 11 & 8 \end{bmatrix} \qquad A' = \begin{bmatrix} 1 & 0 & 5 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

*Solution.* The rank is the dimension of the space spanned by the column vectors of the matrix. Using the row-reduced form, we see that these columns span  $\mathbb{R}^2$  and thus the matrix has rank 2.

□

**Problem 19.3.39** What is the rank-nullity theorem?

*Solution.* For an  $n \times n$  matrix  $A$ ,  $\text{rk}(A) + \text{nul}(A) = n$ . □

### 19.3.8 Problem Set VIII

**Problem 19.3.40** Let  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  be defined by  $T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_3 \\ x_1 + x_2 \end{bmatrix}$ .

1. Determine  $\ker(T)$ .
2. Determine the dimensions of  $\ker(T)$ .
3. Using the Nullity Theorem, determine the dimension of  $\text{im}(T)$ .

*Solution.* In order,

1. If  $T(x_1, x_2, x_3)^T = 0$ , then  $x_3 = 0$  and  $x_1 + x_2 = 0$ .  $\ker(T) = \{(x, -x, 0) : x \in \mathbb{R}\}$
2. This is a line through the origin, so the dimension is 1
3. The Nullity Theorem states that  $\dim(\ker(T)) + \dim(\text{im}(T)) = \dim(\mathbb{R}^3) = 3$ . Thus  $\dim(\text{im}(T)) = 2$ .

□

**Problem 19.3.41** Find the matrix representation of  $T$  (Previous problem) in the standard basis of  $\mathbb{R}^3$ .

*Solution.*  $Te_1 = (0, 1)^T$ ,  $Te_2 = (0, 1)^T$ , and  $Te_3 = (1, 0)^T$ . The matrix representation is  $T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$  □

**Problem 19.3.42** Let  $P_n$  be the set of all polynomials with real coefficients of degree less than  $n$ . The standard basis is  $\{1, x, \dots, x^{n-1}\}$ . Let  $D : P_3 \rightarrow P_2$  be defined by  $D(p) = 5 \frac{dp}{dx}$ . Determine the matrix representation of  $D$  with respect to the standard basis.

*Solution.* We need only check how  $D$  acts on the basis vectors.  $D(1) = 0 + 0x$ ,  $D(x) = 1 + 0x$ ,  $D(x^2) = 0 + 2x$ . So, we have  $D = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}$  □

**Problem 19.3.43** Let  $V$  be a vector space over  $\mathbb{R}$  and let  $S$  be a subspace of  $V$ .

1. Define  $S^\perp$ .
2. If  $S = \text{Span}\{(1, 2, 1)^T, (1 - 1, 2)^T\}$ , what is  $S^\perp$ ?

*Solution.* In order,

1.  $S^\perp = \{x \in V : \forall y \in S, x^T y = 0\}.$
2. Using the definition, the equations below give us  $S^\perp = \{x_3(-\frac{5}{3}, \frac{1}{3}, 1) : x_3 \in \mathbb{R}\}$

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 0 & \frac{5}{3} \\ 0 & 1 & \frac{-1}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

□

**Problem 19.3.44**

1. Let  $V$  be a vector space over  $\mathbb{R}$ . Define an inner product.
2. What is the difference between the standard dot product in  $\mathbb{R}^n$  and an inner product? Can a vector space have more than one inner product?
3. If  $\langle x, y \rangle = xy$ , what is  $\|x\|$ ?

*Solution.* In order,

1. An inner product is a generalization of the standard dot product. The dot product is itself an inner product, but not all inner products are dot products. There are infinitely many inner products for  $\mathbb{R}$ . Let  $n \in \mathbb{N}$  be arbitrary, then  $\langle x|y \rangle = nxy$  is an inner product.
2.  $\|x\| = \sqrt{\langle x|x \rangle} = \sqrt{x^2} = |x|.$

□

**Problem 19.3.45** Let  $V = C[-1, 1]$  and let  $\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx$ .

1. Show that  $f(x) = 1$  and  $g(x) = x$  are orthogonal with respect to this inner product.
2. Determine  $\|f\|$  and  $\|g\|$ .
3. Show that  $f$  and  $g$  satisfy the Pythagorean Law.

*Solution.* In order,

1.  $\langle 1, x \rangle = \int_{-1}^1 x dx = 0$
2.  $\|1\| = \sqrt{\int_{-1}^1 dx} = \sqrt{2}, \|x\| = \sqrt{\int_{-1}^1 x^2 dx} = \sqrt{\frac{2}{3}}$
3.  $\|1+x\|^2 = \langle 1+x, 1+x \rangle = \langle 1, 1 \rangle + 2\langle 1, x \rangle + \langle x, x \rangle = \|1\|^2 + \|x\|^2$

□

**Problem 19.3.46** Let  $V$  be any inner product space. State and prove the Pythagorean Theorem for inner product spaces.

*Solution.* The Pythagorean Theorem for Inner Product Spaces state that if  $V$  is an inner product space with inner product  $\langle \cdot, \cdot \rangle$ , and if  $\langle x, y \rangle = 0$ , then  $\|x\|^2 + \|y\|^2 = \|x + y\|^2$ . For  $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle$ . But as  $x$  and  $y$  are orthogonal,  $\langle x, y \rangle = 0$ . Thus  $\|x + y\|^2 = \langle x, x \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2$ .  $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ . □

**Problem 19.3.47** Prove that if  $V$  is an inner product space and  $S$  is a subspace of  $V$ , then  $S^\perp$  is a subspace of  $V$ .

*Solution.* We must check that  $0 \in S^\perp$  and that  $S^\perp$  is closed under addition and scalar multiplication.

1. For all  $x \in S$ ,  $\langle 0, x \rangle = 0$ , and thus  $0 \in S^\perp$ .
2. If  $x, y \in S^\perp$  and  $z \in S$ , then  $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle = 0 + 0 = 0$ . Thus  $x + y \in S^\perp$ .
3. If  $x \in S^\perp$ ,  $y \in S$ , and  $\alpha$  is a scalar, then  $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle = \alpha \cdot 0 = 0$ . Thus  $\alpha x \in S^\perp$ .  $S^\perp$  is a subspace.

□

# CHAPTER 20

---

## Algebraic Geometry

---

### 20.1 Notes on Cox, Little, and O’Shea

**Theorem 20.1.1.** *If  $p$  is prime, then  $\mathbb{Z}_p \setminus \{0\}$  is a group under multiplication modulo  $p$ .*

#### Fields and Rings

We usually omit the multiplication symbol  $\cdot$  and just write  $ab$  instead of  $a \cdot b$

**Theorem 20.1.2.** *If  $-1$  is the additive inverse of  $1$ , then  $(-1)^2 = 1$*

In rings and fields,  $+$  is usually called addition and  $\cdot$  is usually called multiplication.

**Theorem 20.1.3.** *If  $R$  is a ring and  $a \in R$ , then  $a \cdot 0 = 0 \cdot a = 0$*

**Definition 20.1.1** An integral domain is a commutative ring such that  $ab = 0 \Rightarrow a = 0$  or  $b = 0$

**Definition 20.1.2** A divisor of zero in a ring  $R$  is an element  $a \in R$  such that  $\exists_{b \in R \setminus \{0\}} : ab = 0$

**Theorem 20.1.4.** *a divisor of zero in a ring  $R$  if and only if  $f : R \rightarrow R$ ,  $f(x) = ax$  is not injective.*

**Theorem 20.1.5.** *Any field  $k$  is an integral domain.*

**Definition 20.1.3** An ideal of a commutative ring is a set  $I \subset R$  such that:

1.  $0 \in I$  [Existence of Additive Inverse]
2.  $\forall_{a,b \in I}, a + b \in I$  [Closure Under Addition]
3.  $\forall_{a \in I, b \in R}, ab \in I$  [Absorption Property]

## Determinants

The elementary definitions from linear algebra are presumed. The set of all permutations of  $\mathbb{Z}_n$  is denoted  $S_n$ .  $S_n$  is a group under composition,  $\langle S_n, \circ \rangle$

**Definition 20.1.4** The permutation matrix of  $\sigma \in S_n$ , denoted  $P_\sigma$ , is the matrix formed by the image of the identity matrix  $I_n$  under the mapping  $(a_{ij}) \mapsto (a_{i\sigma(j)})$

**Example 20.1.1** Consider the permutation on  $\mathbb{Z}_3$  defined by the cycle  $1 \rightarrow 3 \rightarrow 2 \rightarrow 1$ . We can make this a matrix equation as follows:

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

The leftmost matrix is obtained by permuting the columns of the identity matrix  $I_3$  by  $\sigma$ .

**Definition 20.1.5** The sign of a permutation  $\sigma \in S_n$  is  $\text{sgn}(\sigma) = \det(P_\sigma)$ .

From the way  $P_\sigma$  is defined,  $\text{sgn}(\sigma) = \det(P_\sigma) = \pm 1$ , depending on  $\sigma$ .

**Theorem 20.1.6.** If  $A = (a_{ij})$  is an  $n \times n$  matrix, then  $\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{k=1}^n a_{k\sigma(k)}$

### 20.1.1 Geometry, Algebra, and Algorithms

This section introduces the basic ideas. Affine varieties and ideals in the polynomial ring  $k[x_1, \dots, x_n]$  are studied. Finally, polynomials in one variable are studied to introduce the role of algorithms.

#### Polynomials and Affine Space

To link algebra and geometry, we will study polynomials over a field. Fields are important because linear algebra works over any field  $k$ . There are three particular fields that will be used the most:

1.  $\mathbb{Q}$ : This field is used for computer examples.
2.  $\mathbb{R}$ : This field is used for drawing pictures of curves and surfaces.
3.  $\mathbb{C}$ : This field is used for proving many theorems.

**Definition 20.1.6** A monomial in  $x_1, \dots, x_n$  is a product  $\prod_{i=1}^n x_i^{\alpha_i}$ , where  $\alpha_1, \dots, \alpha_n \in \mathbb{N}_0$

**Definition 20.1.7** The total degree of a monomial  $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$  is the sum  $\sum_{i=1}^n \alpha_i$

For  $\alpha_1, \dots, \alpha_n \in \mathbb{N}_0$ , let  $\alpha = (\alpha_1, \dots, \alpha_n)$ . We write  $\prod_{i=1}^n x_i^{\alpha_i} = x^\alpha$

**Definition 20.1.8** A polynomial  $f$  in  $x_1, \dots, x_n$  is a finite linear combination of monomials over  $k$ .

The set of all polynomials in  $n$  variables with coefficients in  $k$  is denoted  $k[x_1, \dots, x_n]$

**Definition 20.1.9** For a polynomial  $f = \sum_\alpha a_\alpha x^\alpha \in k[x_1, \dots, x_n]$   $a_\alpha$  is called the coefficient of  $x^\alpha$

**Definition 20.1.10** A term of  $f = \sum_\alpha a_\alpha x^\alpha \in k[x_1, \dots, x_n]$  is a product  $a_\alpha x^\alpha$  where  $a_\alpha \neq 0$

**Definition 20.1.11** The total degree of  $f = \sum_\alpha a_\alpha x^\alpha$ , denoted  $\deg(f)$ , is  $\deg(f) = \max\{|\alpha| : a_\alpha \neq 0\}$

**Definition 20.1.12** The zero polynomial is the polynomial with all zero coefficients.

**Theorem 20.1.7.** *The sum and product of polynomials in  $k[x_1, \dots, x_n]$  is a polynomial in  $k[x_1, \dots, x_n]$*

**Definition 20.1.13** A divisor of  $f \in k[x_1, \dots, x_n]$ , is a  $g \in k[x_1, \dots, x_n]$  such that  $\exists h \in k[x_1, \dots, x_n] : f = gh$

**Theorem 20.1.8.** *For all  $n \in \mathbb{N}$ ,  $k[x_1, \dots, x_n]$  is a commutative ring.*

Because of this we call  $k[x_1, \dots, x_n]$  a polynomial ring.

**Definition 20.1.14** The  $n$ -dimensional affine space over  $k$  is the set  $k^n = \{(a_1, \dots, a_n) : a_1, \dots, a_n \in k\}$

A polynomial defines a function  $f : k^n \rightarrow k$ .

**Definition 20.1.15** A zero function  $f : k^n \rightarrow k$  is a function such that  $f(x) = 0$  for all  $x \in k^n$

A zero function and the zero polynomial are not necessarily the same thing. That is, there are fields  $k$  with non-zero polynomials that evaluate to zero at every point.

**Theorem 20.1.9.** *There exists fields  $k$ ,  $f \in k[x]$  such that  $f$  is a non-zero polynomial and  $\forall a \in k, f(a) = 0$*

**Theorem 20.1.10.** *If  $k$  is an infinite field,  $f \in k[x_1, \dots, x_n]$ , then  $f$  is a zero function if and only if it is the zero polynomial.*

**Theorem 20.1.11.** *If  $k$  is an infinite field and  $f, g \in k[x_1, \dots, x_n]$ , then  $f = g$  if and only if  $f : k^n \rightarrow k$  and  $g : k^n \rightarrow k$  give the same function.*

There is a special property for polynomials over the complex numbers  $\mathbb{C}$ .

**Theorem 20.1.12.** *Every non-constant polynomial  $f \in \mathbb{C}[x]$  has a root in  $\mathbb{C}$ .*

**Definition 20.1.16** An algebraically closed field is a field such that for non-constant  $f$ ,  $\exists_{x \in k} : f(x) = 0$ .

## Affine Varieties

**Definition 20.1.17** The affine variety of  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$  is  $\{x \in k^n : \forall_{1 \leq i \leq s}, f_i(x) = 0\}$

The affine variety of  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$  is denoted  $\mathbf{V}(f_1, \dots, f_s)$ . The affine variety of a finite set of polynomials is the solution set of the system of equations  $f_i(x) = 0$ .

**Example 20.1.2**  $\mathbf{V}(x^2 + y^2 - 1) \subset \mathbb{R}^2$  is the set of solutions to  $x^2 + y^2 - 1 = 0$ : The unit circle.

**Example 20.1.3** The conic sections (Circles, ellipses, parabolas, and hyperbolas) are affine varieties. The graphs of rational functions are also affine varieties. For if  $y = \frac{P(x)}{Q(x)}$ , where  $P, Q \in \mathbb{R}[x]$ , then  $\mathbf{V}(yQ(x) - P(x))$  is an affine variety equivalent to that graph.

**Example 20.1.4** The surfaces the represent affine varieties need not be smooth everywhere. Indeed,  $\mathbf{V}(z^2 - x^2 - y^2)$  is the graph of a cone with its apex at the origin. As such, the surface obtained is not smooth at the origin. Such points are called singular points.

**Example 20.1.5** The twisted cubic is  $\mathbf{V}(y - x^2, z - x^3)$ , with the parametrization  $\{(t, t^2, t^3) : t \in \mathbb{R}\}$

The notion of dimension is very subtle. In previous examples, if we have  $m$  polynomials in  $\mathbb{R}^n$ , we expect a surface of  $n - m$  dimension. This is not always the case, however.

**Example 20.1.6**  $\mathbf{V}(xz, yz)$  is the set of solutions to  $xy = yz = 0$ . If  $z = 0$ , then any point  $(x, y, 0) \in \mathbb{R}^3$  satisfies this. If  $z \neq 0$ , then  $x = y = 0$  and thus any point  $(0, 0, z) \in \mathbb{R}^3$  is a solution. Thus,  $\mathbf{V}(xz, yz)$  is the union of the  $xy$  plane and the  $z$  axis. So  $\mathbf{V}(xz, yz)$  is two dimensional, not one.

**Definition 20.1.18** A linear variety is an affine variety in which the defining polynomials are linear.

**Example 20.1.7** Let  $k$  be a field and consider the following polynomials:

$$a_{11}x_1 + \dots + a_{1n}x_n = b_1$$

⋮

$$a_{m1}x_1 + \dots + a_{mn}x_n = b_m$$

From linear algebra we know that the method of Gaussian Elimination and row-reduction gives us the solution set of the system of equations. We also know that the dimension of the solutions set is  $n - r$ , where  $r$  is the number of independent equations (Also known as the rank of the coefficient matrix). The dimension of an affine variety is also determined by the number of independent equations, however the term “Independent,” is much more subtle.

**Example 20.1.8** Find the maximum of  $f(x, y, z) = x^3 + 2xyz - z^2$  subject to  $g(x, y, z) = x^2 + y^2 + z^2 = 1$ . From multivariable calculus, specifically the method of Lagrange Multipliers, we know this occurs when  $\nabla(f) = \lambda \nabla(g)$ , for some  $\lambda \in \mathbb{R}$ . This gives us the following:

$$\begin{array}{ll} x^2 + 2yz = 2\lambda x & 2xy - 2z = 2\lambda z \\ 2xz = 2\lambda y & x^2 + y^2 + z^2 = 1 \end{array}$$

Solving this via algebraic means can be a nightmare. Various algorithms exist, however.

It is possible for an affine variety to be the empty set. Let  $k = \mathbb{R}$ , and  $f = x^2 + y^2 + 1$ . Then  $\mathbf{V}(f) = \emptyset$ . That is, there is no real solution to  $x^2 + y^2 = -1$ .

**Example 20.1.9** Consider a robot arm. The “Armpit,” is at the origin, and the “Elbow,” is at the point  $(x, y) \in \mathbb{R}^2$  where  $x^2 + y^2 = r^2$  ( $r$  is the length of “Bicep.”) The “Hand,” will then be at  $(z, w) \in \mathbb{R}^2$  where  $(x-z)^2 + (y-w)^2 = \ell^2$  ( $\ell$  is the length of the “Forearm.”) Not every point  $(x, y, z, w) \in \mathbb{R}^4$  represents a possible position of the robot arm, there are the following constraints:

$$\begin{aligned} x^2 + y^2 &= r^2 \\ (x-z)^2 + (y-w)^2 &= \ell^2 \end{aligned}$$

The solution set defines an affine variety in  $\mathbb{R}^4$ . For arms in  $\mathbb{R}^3$ , the solution set would be in  $\mathbb{R}^6$ .

**Theorem 20.1.13.** *If  $V, W \subset k^n$  are affine varieties, then so are  $V \cup W$  and  $V \cap W$ . Moreover:*

1.  $\mathbf{V}(f_1, \dots, f_s) \cap \mathbf{V}(g_1, \dots, g_t) = \mathbf{V}(f_1, \dots, f_s, g_1, \dots, g_t)$
2.  $\mathbf{V}(f_1, \dots, f_s) \cup \mathbf{V}(g_1, \dots, g_t) = \mathbf{V}(f_i g_j : 1 \leq i \leq s, 1 \leq j \leq t)$

**Example 20.1.10**  $\mathbf{V}(xz, yz) = \mathbf{V}(xy) \cup \mathbf{V}(z)$ .  $\mathbf{V}(xz, yz)$  is the union of the  $xy$  plane and the  $z$  axis.

**Example 20.1.11** For the twisted cubic:  $\mathbf{V}(y - x^2, z - x^3) = \mathbf{V}(y - x^2) \cap \mathbf{V}(z - x^3)$

Several problems arise concerning affine varieties:

1. Can we determine if  $\mathbf{V}(f_1, \dots, f_s) \neq \emptyset$ ? [Consistency]
2. Can we determine if  $\mathbf{V}(f_1, \dots, f_s)$  is finite? [Finiteness]
3. Can we determine the “Dimension,” of  $\mathbf{V}(f_1, \dots, f_s)$ ?

The answer to these questions is yes, although we must be careful in choosing the field we work with.

### Parametrizations of Affine Varieties

We now arrive at the problem of describing all of the points in an affine variety.

**Example 20.1.12** Consider the system in  $\mathbb{R}[x, y, z]$ :

$$\begin{aligned} x + y + z &= 1 \\ x + 2y - z &= 3 \end{aligned}$$

From linear algebra we get the row echelon matrix:

$$\left[ \begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 2 \end{array} \right]$$

Letting  $z = t$ , we get  $x = -3t - 1$  and  $y = 2 + 2t$ . The parametrization of the affine variety is thus  $\{(-3t - 1, 2t + 2, t) : t \in \mathbb{R}\}$ . We call  $t$  a parameter, and  $(-3t - 1, 2t + 2, t)$  a parametrization.

**Example 20.1.13** One way to parametrize the unit circle uses trigonometric functions:  $(\cos(t), \sin(t))$ . A rational way to do this is  $(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2})$ . This parametrizes the entire unit circle, with the exception of the point  $(-1, 0)$ . This point is  $\lim_{t \rightarrow \infty} (\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2})$ . So in a sense,  $(-1, 0)$  is a “Point at infinity.”

**Definition 20.1.19** A parametrization of an affine variety  $\mathbf{V}(f_1, \dots, f_s) \subset k^n$  is a set of  $j$  equations  $x_j = f_j(t_1, \dots, t_m)$ , whose solution set  $S$  is such that  $S \subset \mathbf{V}(f_1, \dots, f_s)$ , and for all  $g_1, \dots, g_t$  such that  $S \subset \mathbf{V}(g_1, \dots, g_t)$ ,  $\mathbf{V}(f_1, \dots, f_s) \subset \mathbf{V}(g_1, \dots, g_s)$

Solutions to  $x_k = f_k(t_1, \dots, t_m)$  lie in  $\mathbf{V}(f_1, \dots, f_s)$  and  $\mathbf{V}(f_1, \dots, f_s)$  is the smallest affine variety containing these points.

**Definition 20.1.20** A rational function in  $x_1, \dots, x_n$  is a quotient  $\frac{P(x)}{Q(x)}$  :  $P, Q \in k[x_1, \dots, x_n], Q \neq 0$ .

**Definition 20.1.21**  $k(x_1, \dots, x_n)$  is the set of all rational functions over a field  $k$  in  $x_1, \dots, x_n$ .

**Definition 20.1.22** Equal rational functions are functions  $\frac{P_1}{Q_1}, 5 \frac{P_2}{Q_2} \in k(x_1, \dots, x_n)$  where  $P_1 Q_2 = P_2 Q_1$ .

**Theorem 20.1.14.** *If  $k$  is a field, then  $k(x_1, \dots, x_n)$  is a field.*

**Definition 20.1.23** A rational representation of an affine variety is a rational parametrization.

**Definition 20.1.24** A polynomial representation of an affine variety is a polynomial parametrization.

Writing out an affine variety as  $V = \mathbf{V}(f_1, \dots, f_s)$  is called an implicit representation. There are two questions that arise from parametrization:

1. Does every affine variety have a rational parametric representation?
2. Given a parametric representation of an affine variety, can we find the implicit representation?

The answers are: No to the first question, yes to the second. Indeed, most affine varieties cannot be parametrized by rational functions.

**Example 20.1.14** Find the affine variety parametrized by:

$$\begin{aligned}x &= 1 + t \\y &= 1 + t^2\end{aligned}$$

We have that  $t = x - 1$ , and thus  $y = 1 + (x - 1)^2 = x^2 - 2x + 2$ .

The process described above involved eliminating the variable  $t$  and creating a polynomial in  $x$  and  $y$ . This illustrates the role played by elimination theory.

**Example 20.1.15** Let's parametrize the unit circle in a rational manner. Let  $(x, y)$  be a point on the unit circle and draw a line from the point  $(-1, 0)$  to  $(x, y)$ . This line intersects the  $y$ -axis at some point  $(0, t)$ . We have that the slope of this line is  $m = \frac{t-0}{0-(-1)} = \frac{y-0}{x-(-1)} = \frac{y}{x+1}$ . So  $y = t(x+1)$ . But  $x^2 + y^2 = 1$ , so  $x^2 + t^2(x+1)^2 = 1 \Leftrightarrow x^2 + x \frac{2t^2}{1+t^2} = \frac{1-t^2}{1+t^2} \Leftrightarrow (x + \frac{t^2}{1+t^2})^2 = \frac{1}{(1+t^2)^2} \Leftrightarrow x = \frac{-t^2 \pm 1}{1+t^2}$ . But  $x \in [-1, 1]$ , and thus we get  $x = \frac{1-t^2}{1+t^2}$ . But  $y = t(x+1)$ , and thus  $y = \frac{2t}{1+t^2}$ .  $(x, y) = (\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2})$ .

**Definition 20.1.25** The tangent surface of a smooth curve  $\Gamma : \mathbb{R} \rightarrow \mathbb{R}^n$  is  $\{\Gamma(t) + u\Gamma'(t) : t, u \in \mathbb{R}\}$

The tangent surface is obtained by taking the union of all of the tangent lines to every point on the curve.  $t$  tells us which point on the curve we are one, and  $u$  tells us how far along the tangent line we are.

**Example 20.1.16** The twisted cubic is the curve defined by  $\mathbf{r}(t) = (t, t^2, t^3)$ . Its tangent surface is  $\mathbf{r} + u\mathbf{r}'(t) = (t, t^2, t^3) + u(1, 2t, 3t^2) = (t+u, t^2+2ut, t^3+3ut^2)$ . One question that arises is “Is this an affine variety? If so, what are the defining polynomials.” The answer for this particular surface is yes. The graph of this surface is equal to  $\mathbf{V}(-4x^3z + 3x^2y^2 - 4y^3 + 6xyz - z^2)$ .

An application of this is in the design of complex objects such as automobile hoods and airplane wings. Engineers need curves and surfaces that are easy to describe, quick to draw, and varied in shape. Polynomials and rational functions satisfy this criteria. Complicated curves are usually formed by joining together simpler curves. Suppose a design engineer needs to draw a curve in the plane. The curves in question need to join smoothly, and thus the tangent directions need to match at the endpoints. The engineer must control the following:

1. The starting and ending points of the curve.
2. The tangent directions at the starting and ending points.

The Bézier Cubic does this.

**Definition 20.1.26** The Bézier Cubic in  $\mathbb{R}^2$  is defined by:

$$\begin{aligned}x &= (1-t)^3 x_0 + 3t(1-t)^2 x_1 + 3t^2(1-t)x_2 + t^3 x_3 \\y &= (1-t)^2 y_0 + 3t(1-t)^2 y_1 + 3t^2(1-t)y_2 + t^3 y_3\end{aligned}$$

Where  $x_0, x_1, x_2, x_3, y_0, y_1, y_2, y_3$  are input parameters.

When  $t = 0$ , we have  $x = x_0, y = y_0$ . Thus  $(x_0, y_0)$  is the starting point. Similarly  $(x_3, y_3)$  is the end point. The derivatives are:

$$\begin{aligned}x' &= -3(1-t)^2 x_0 + 3(1-t)(1-3t)x_1 + 3t(2-3t)x_2 + 3t^2 x_3 \\y' &= -3(1-t)^2 y_0 + 3(1-t)(1-3t)y_1 + 3t(2-3t)y_2 + 3t^2 y_3\end{aligned}$$

So  $(x'(0), y'(0)) = (3(x_1-x_0), 3(y_1-y_0))$  and  $(x'(1), y'(1)) = (3(x_3-x_2), 3(y_3-y_2))$ . Hence, choosing  $x_1, x_2$  and  $y_1, y_2$  carefully allows that designer to control the tangent of the curve at the endpoints. Moreover, choosing the point  $(x_1, y_1)$  makes the tangent at  $(x(0), y(0))$  point in the same direction as the line from  $(x_0, y_0)$  to  $(x_1, y_1)$ . Similarly, choosing  $(x_2, y_2)$  makes the tangent at  $(x(1), y(1))$  point in the same direction as the line from  $(x_2, y_2)$  to  $(x_3, y_3)$ .

**Definition 20.1.27** The control polygon of a Bézier Cubic in  $\mathbb{R}^2$  is the polygon formed by the lines  $(x_0, y_0) \rightarrow (x_1, y_1) \rightarrow (x_2, y_2) \rightarrow (x_3, y_3) \rightarrow (x_0, y_0)$ .

Interestingly enough, the Bézier Cubic always lies inside the control polygon. The final thing to control is the length of the tangents at the endpoint. But from equations 1.3 and 1.4, the lengths are three times the distance from  $(x_0, y_0)$  to  $(x_1, y_1)$  and  $(x_2, y_2)$  to  $(x_3, y_3)$ , respectively.

## Ideals

**Definition 20.1.28** An ideal of a polynomial ring  $k[x_1, \dots, x_n]$  is a set  $I \subset k[x_1, \dots, x_n]$  such that:

1.  $0 \in I$

2.  $\forall_{f,g \in I}, f + g \in I$

3.  $\forall_{f \in I, h \in k[x_1, \dots, x_n]}, hf \in I$

**Definition 20.1.29** The ideal generated by a set  $\{f_1, \dots, f_s\} \subset k[x_1, \dots, x_n]$  is the set  $\langle f_1, \dots, f_s \rangle = \{\sum_{i=1}^s h_i f_i : h_1, \dots, h_s \in k[x_1, \dots, x_n]\}$ .

**Theorem 20.1.15.** If  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ , then  $\langle f_1, \dots, f_s \rangle$  is an ideal.

The ideal  $\langle f_1, \dots, f_s \rangle$  has a nice interpretation. If  $x \in k$  such that  $f_1(x) = \dots = f_s(x) = 0$ , then for any set of polynomials  $h_1, \dots, h_s$ , we have  $h_1(x)f_1(x) = \dots = h_s(x)f_s(x) = 0$ , and adding the equations we get  $h_1(x)f_1(x) + \dots + h_s(x)f_s(x) = 0$ . Thus we can think of  $\langle f_1, \dots, f_s \rangle$  as the set of all “Polynomial consequences,” of the equations  $f_1 = \dots = f_s = 0$ .

**Example 20.1.17** Consider the following system:

$$x = 1 + t$$

$$y = 1 + t^2$$

We can eliminate  $t$  to obtain  $y = x^2 - 2x + 2$ . To see this, write

$$x - 1 - t = 0$$

$$-y + 1 + t^2 = 0$$

Multiplying this first by  $x - 1 + t$  and adding, we get  $(x - 1)^2 - y + 1 = 0$ . Thus  $y = x^2 - 2x + 2$ .

**Definition 20.1.30** A finitely generated ideal is an ideal such that  $\exists_{f_1, \dots, f_s} : I = \langle f_1, \dots, f_s \rangle$ .

**Definition 20.1.31** A basis of an ideal is a set  $\{f_1, \dots, f_s\} \subset k[x_1, \dots, x_n]$  such that  $I = \langle f_1, \dots, f_s \rangle$

Hilbert's Basis Theorem, to be proved later, states that every ideal in  $k[x_1, \dots, x_n]$  is finitely generated. An ideal in  $k[x_1, \dots, x_n]$  is similar to a subspace in linear algebra. Both must be closed under multiplication and addition, except that in a subspace we multiply by scalars and in an ideal we multiply by polynomials.

**Theorem 20.1.16.** If  $\langle f_1, \dots, f_s \rangle = \langle g_1, \dots, g_t \rangle$ , then  $\mathbf{V}(f_1, \dots, f_s) = \mathbf{V}(g_1, \dots, g_s)$ .

**Example 20.1.18**  $\langle 2x^2 + 3y^2 - 11, x^2 - y^2 - 3 \rangle = \langle x^2 - 4, y^2 - 1 \rangle$ . So,  $\mathbf{V}(2x^2 + 3y^2 - 11, x^2 - y^2 - 3) = \{(2, 1), (-2, 1), (2, -1), (-2, -1)\}$ . Changing basis simplifies the problem.

**Definition 20.1.32** The ideal of an affine variety  $V \subset k^n$  is  $\mathbf{I}(V) = \{f \in k[x_1, \dots, x_n] : \forall_{x \in V} f(x) = 0\}$

**Theorem 20.1.17.** If  $V \subset k^n$  is an affine variety, then  $\mathbf{I}(V)$  is an ideal of  $k[x_1, \dots, x_n]$ .

**Theorem 20.1.18.** For any field  $k$ ,  $\mathbf{I}(\{(0, 0)\}) = \langle x, y \rangle$ .

**Theorem 20.1.19.** For any infinite field  $k$ ,  $\mathbf{I}(k^n) = \{0\}$ .

**Theorem 20.1.20.** If  $V = \mathbf{V}(y-x^2, z-x^3) \subset \mathbb{R}^3$ ,  $f \in \mathbf{I}(V)$ , then  $\exists_{h_1, h_2, r(x) \in \mathbb{R}[x, y, z]}$  such that  $f = h_1(y - x^2) + h_2(z - x^3) + r$ .

**Theorem 20.1.21.** If  $V = \mathbf{V}(y-x^2, z-x^3) \subset \mathbb{R}^3$ , then  $\mathbf{I}(V) = \langle y-x^2, z-x^3 \rangle$

It is not always true that  $\mathbf{I}(\mathbf{V}(f_1, \dots, f_s)) = \langle f_1, \dots, f_s \rangle$ .

**Theorem 20.1.22.** If  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ , then  $\langle f_1, \dots, f_s \rangle \subset \mathbf{I}(\mathbf{V}(f_1, \dots, f_s))$ .

**Theorem 20.1.23.** There exists fields  $k$  and polynomials such that  $\langle f_1, \dots, f_s \rangle \neq \mathbf{I}(\mathbf{V}(f_1, \dots, f_s))$

**Theorem 20.1.24.** If  $k$  is a field,  $V, W \subset k^n$  are affine varieties, then  $V \subset W$  if and only if  $\mathbf{I}(W) \subset \mathbf{I}(V)$ .

**Theorem 20.1.25.** If  $k$  is a field,  $W, W \subset k^n$  are affine varieties, then  $V = W$  if and only if  $\mathbf{I}(W) = \mathbf{I}(V)$ .

Three questions arise concerning ideals in  $k[x_1, \dots, x_n]$ .

1. Can every ideal  $I \subset k[x_1, \dots, x_n]$  be written as  $\langle f_1, \dots, f_s \rangle$  for some  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ ?
2. If  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ ,  $f \in k[x_1, \dots, x_n]$ , is there an algorithm to see if  $f \in \langle f_1, \dots, f_s \rangle$ ?
3. Is there a relation between  $\langle f_1, \dots, f_s \rangle$  and  $\mathbf{I}(\mathbf{V}(f_1, \dots, f_s))$ ?

## Polynomials in One Variable

This section studies the division algorithm of polynomials in one variable.

**Definition 20.1.33** The leading term of  $f = \sum_{k=1}^n a_k x^k \in k[x]$ , where  $a_n \neq 0$ , is  $\text{LT}(f) = a_n x^n$ .

**Example 20.1.19** If  $f = 2x^3 - 4x + 3$ , then  $\text{LT}(f) = 2x^3$ .

**Theorem 20.1.26.** If  $k$  is a field and  $g \in k[x] \setminus \{0\}$ , then  $\forall_{f \in k[x]}: \exists_{q, r \in k[x]}: f = qg + r$ , where either  $r = 0$  or  $\deg(r) < \deg(g)$ . Furthermore,  $q$  and  $r$  are unique.

From the uniqueness of  $r$ , we call  $r$  the remainder of  $f$  with respect to  $g$ .

**Theorem 20.1.27.** If  $k$  is a field and  $f \in k[x]$  is a non-zero polynomial, then  $f$  has at most  $\deg(f)$  roots.

**Definition 20.1.34** A principal ideal is an ideal generated by a single element.

**Theorem 20.1.28.** *If  $k$  is a field, then every ideal of  $k[x]$  is principal.*

**Theorem 20.1.29.** *If  $\langle f \rangle = \langle g \rangle$  are ideals in  $k[x]$ , then there is a constant  $h$  such that  $f = hg$ .*

**Definition 20.1.35** A greatest common divisor of  $f, g \in k[x]$  is a polynomial  $h \in k[x]$  such that  $h$  divides  $f$  and  $g$  and  $\forall p \in k[x]$  such that  $p$  divides  $f$  and  $g$ ,  $p$  divides  $h$ .

**Theorem 20.1.30.** *If  $f, g \in k[x]$ , then there is a greatest common divisor of  $f$  and  $g$ .*

**Theorem 20.1.31.** *If  $f, g \in k[x]$ , and  $h_1, h_2$  are greatest common divisors of  $f$  and  $g$ , then there is a constant  $c \in k$  such that  $h_1 = ch_2$ .*

The Euclidean Algorithm is used for computational purposes to compute the greatest common divisor of two polynomials. Let  $f, g \in k[x]$ .

1. Let  $h_1 = f$
2. Let  $s_1 = g$
3. While  $s_n \neq 0$ , do the following:
  - (a)  $r_n = \text{remainder}(h_n, s_n)$
  - (b)  $h_{n+1} = s_n$
  - (c)  $s_{n+1} = r_n$

There is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $h_n = h_N$ . Letting  $h = h_N$ , this is the greatest common divisor of  $f$  and  $g$ . This comes from  $\text{GCD}(f, g) = \text{GCD}(f - qg, g) = \text{GCD}(r, g)$  and the fact that  $\deg(r) < \deg(g)$ . So  $\deg(r_{n+1}) < \deg(r_n)$ , and eventually  $\deg(r_N) = 0$ .

**Definition 20.1.36** A greatest common divisor of polynomials  $f_1, \dots, f_s \in k[x]$  is a polynomial  $h \in k[x]$  such that  $h$  divides  $f_1, \dots, f_s$  and if  $p \in k[x]$  such that  $p$  divides  $f_1, \dots, f_s$ , then  $p$  divides  $h$ .

**Theorem 20.1.32.** *If  $f_1, \dots, f_s \in k[x]$ , then there is a polynomial  $h \in k[x]$  that is a greatest common divisor of  $f_1, \dots, f_s$ .*

**Theorem 20.1.33.** *If  $f_1, \dots, f_s \in k[x]$ , and if  $h$  is a GCD of  $f_1, \dots, f_s$ , then  $\langle h \rangle = \langle f_1, \dots, f_s \rangle$*

## 20.1.2 Elimination Theory

### The Elimination and Extension Theorems

**Definition 20.1.37** If  $I = \langle f_1, \dots, f_s \rangle \subset k[x_1, \dots, x_n]$ , the  $\ell$ -th elimination ideal, denoted  $I_\ell$ , is the ideal defined as  $I_\ell = I \cap k[x_{\ell+1}, \dots, x_n]$ .

**Theorem 20.1.34.** For  $\ell \in \mathbb{Z}_{n-1}$ , if  $I = \langle f_1, \dots, f_s \rangle \subset k[x_1, \dots, x_n]$  is an ideal, then  $I_\ell$  is an ideal of  $k[x_1, \dots, x_n]$ .

**Theorem 20.1.35** (The Elimination Theorem). If  $I \subset k[x_1, \dots, x_n]$  is an ideal and  $G$  is a Groebner Basis of  $I$  with respect to the lexicographic ordering  $x_1 > x_2 > \dots > x_n$ , then for all  $\ell \in \mathbb{Z}_n$ ,  $G_\ell = G \cap k[x_{\ell+1}, \dots, x_n]$  is a Groebner Basis of  $I_\ell$ .

**Theorem 20.1.36** (The Extension Theorem). If  $I = \langle f_1, \dots, f_s \rangle \subset \mathbb{C}[x_1, \dots, x_n]$ , and if  $I_1$  is the first elimination ideal of  $I$ , and if for all  $i \in \mathbb{Z}_s$   $f_i = g(x_2, \dots, x_n)x_1^{N_i} + h$ , where the degree of the  $x_1$  component of  $h$  is less than  $N_i$ , and if  $(a_2, \dots, a_n) \notin V(g_1, \dots, g_s)$ , then there is an  $a_1 \in \mathbb{C}$  such that  $(a_1, \dots, a_n) \in V(I)$ .

The requirement that we work in  $\mathbb{C}$  is crucial. This theorem does not hold in  $\mathbb{R}$ .

**Theorem 20.1.37.** If  $I = \langle f_1, \dots, f_s \rangle \subset \mathbb{C}[x_1, \dots, x_n]$  if for some  $i$ ,  $f_i$  is of the form  $f_i = cx_1^N + g(x_2, \dots, x_n)$ , where the degree of the  $x_1$  term in  $g$  is less than  $N$ , and  $c \neq 0$ , and if  $(a_2, \dots, a_n) \in V(I_1)$ , then there is an  $a_1 \in \mathbb{C}$  such that  $(a_1, \dots, a_n) \in V(I)$ .

### The Geometry of Elimination

**Definition 20.1.38** The projection map  $\pi_\ell : \mathbb{C}^n \rightarrow \mathbb{C}^{n-\ell}$  is defined as  $\pi_\ell(a_1, \dots, a_n) = (a_{\ell+1}, \dots, a_n)$ .

**Theorem 20.1.38.** If  $V = V(f_1, \dots, f_s) \subset \mathbb{C}^n$ , and  $I_\ell$  is the  $\ell$ th elimination ideal of  $\langle f_1, \dots, f_s \rangle$ , then  $\pi_\ell(V) \subset V(I_\ell)$

**Theorem 20.1.39.** If  $V = V(f_1, \dots, f_s) \subset \mathbb{C}^n$ , and  $G_\ell$  is as defined in the extension theorem, then  $V(I_\ell) = \pi_\ell(V) \cup G_\ell$

**Theorem 20.1.40** (The First Closure Theorem). If  $V = V(f_1, \dots, f_s) \subset \mathbb{C}^n$  and  $I_\ell$  is the  $\ell$ th elimination ideal of  $\langle f_1, \dots, f_s \rangle$ , then  $V(I_\ell)$  is the smallest affine variety containing  $\pi_\ell(V) \subset \mathbb{C}^{n-\ell}$ .

**Theorem 20.1.41** (The Second Closure Theorem). If  $V = V(f_1, \dots, f_s) \subset \mathbb{C}^n$ ,  $V \neq \emptyset$ , and if  $I_\ell$  is the  $\ell$ -th elimination ideal of  $\langle f_1, \dots, f_s \rangle$ , then there is an affine variety  $W \underset{\text{Proper}}{\subset} V(I_\ell)$  such that  $V(I_\ell) \setminus W \subset \pi_\ell(V)$ .

**Theorem 20.1.42.** If  $V = V(f_1, \dots, f_s) \subset \mathbb{C}^n$  and if for some  $i$ ,  $f_i$  is of the form  $f_i = cx_1^N + g$ , where the  $x_1$  terms in  $g$  are of degreeless than  $N$ , and  $c \neq 0$ , then  $\pi_1(V) = V(I_1)$ .

## Implicitization

**Definition 20.1.39** A polynomial parametrization is a finite set of equations  $x_k = f_k(t_1, \dots, t_m) \in k[t_1, \dots, t_m]$ . The function  $F : k^m \rightarrow k^n$  is the image defined by  $(t_1, \dots, t_m) \mapsto (x_1, \dots, x_n)$

**Theorem 20.1.43** (The Polynomial Implicitization Theorem). *If  $k$  is an infinite field and  $F : k^m \rightarrow k^n$  is a function determined by some polynomial parametrization, and if  $I$  is an ideal  $I = \langle x_1 - f_1, \dots, x_n - f_n \rangle \subset k[t_1, \dots, t_m, x_1, \dots, x_n]$  then  $\mathbf{V}(I_m)$  is the smallest variety in  $k^n$  containing  $F(k^n)$ , where  $I_m$  is the  $m^{\text{th}}$  elimination ideal.*

**Definition 20.1.40** A rational parametrization is a finite set of equations  $x_k = f_k(t_1, \dots, t_m) \in k(t_1, \dots, t_m)$

**Theorem 20.1.44** (Rational Implicitization). *If  $k$  is an infinite field,  $f_k, g_k, k = 1, 2, \dots, n$  are a rational parametrization,  $W = \mathbf{V}(g_1, \dots, g_s)$ , and if  $F : k^m \setminus W \rightarrow k^n$  is the function determined by the rational parametrization, if  $J = \langle g_1 x_1 - g_1, \dots, g_n x_n - g_n, 1 - gy \rangle \subset k[y, t_1, \dots, t_m, x_1, \dots, x_n]$ , where  $g = g_1 \cdots g_n$ , and if  $J_{m+1}$  is the  $(m+1)^{\text{th}}$  elimination ideal, then  $\mathbf{V}(J_{m+1})$  is the smallest variety in  $k^n$  containing  $F(k^m \setminus W)$ .*

## Singular Points and Envelopes

**Definition 20.1.41** A singular point on an affine variety  $\mathbf{V}(f)$  is a point  $x \in k$  such that there exists no tangent line at  $x$ .

For curves in the plane, this usually happens when either the curve intersects itself or has a kink in it.

**Definition 20.1.42** If  $k \in \mathbb{N}$ , if  $(a, b) \in \mathbf{V}(f)$ , and if  $L$  is a line through  $(a, b)$ , then  $L$  meets  $\mathbf{V}(f)$  with multiplicity  $k$  at  $(a, b)$  if  $L$  can be linearly parametrized in  $x$  and  $y$  so that  $t = 0$  is a root of multiplicity  $k$  of the polynomial  $g(t) = f(a + ct, b + dt)$ .

**Theorem 20.1.45.** *If  $f \in k[x, y]$ ,  $(a, b) \in \mathbf{V}(f)$ , and if  $\nabla f(a, b) \neq (0, 0)$ , then there is a unique line through  $(a, b)$  which meets  $\mathbf{V}(f)$  with multiplicity  $k \geq 2$ .*

**Theorem 20.1.46.** *If  $f \in k[x, y]$ ,  $(a, b) \in \mathbf{V}(f)$ , and if  $\nabla f(a, b) = 0$ , then every line through  $(a, b)$  meets  $\mathbf{V}(f)$  with multiplicity  $k \geq 2$ .*

**Definition 20.1.43** If  $f \in k[x, y]$ ,  $(a, b) \in \mathbf{V}(f)$ , and if  $\nabla f(a, b) \neq (0, 0)$ , then the tangent line of  $\mathbf{V}(f)$  at  $(a, b)$  is the unique line through  $(a, b)$  with multiplicity  $k \geq 2$ . We say that  $(a, b)$  is a non-singular point of  $\mathbf{V}(f)$ .

**Definition 20.1.44** If  $f \in k[x, y]$ ,  $(a, b) \in \mathbf{V}(f)$ , and if  $\nabla f(a, b) = (0, 0)$ , then we say that  $(a, b)$  is a singular point of  $\mathbf{V}(f)$ .

**Definition 20.1.45** If  $\mathbf{V}(F_t)$  is a family of curves in  $\mathbb{R}^2$ , its envelope consists of all points  $(x, y) \in \mathbb{R}^2$  such that  $F(x, y, t) = 0$  and  $\frac{\partial}{\partial t} F(x, y, t) = 0$  for some  $t \in \mathbb{R}$ .

### Unique Factorization and Resultants

**Definition 20.1.46** If  $k$  is a field, then a polynomial  $f \in k[x_1, \dots, x_n]$  is said to be irreducible if  $f$  is non-constant and is not the product of two non-constant polynomials in  $k[x_1, \dots, x_n]$ .

**Theorem 20.1.47.** Every non-constant polynomial  $f \in k[x_1, \dots, x_n]$  can be written as a product of polynomials which are irreducible over  $k$

**Theorem 20.1.48.** If  $f, g \in k[x_1, \dots, x_n]$  have positive degree in  $x_1$ , then  $f$  and  $g$  have a common factor in  $k[x_1, \dots, x_n]$  of positive degree in  $x_1$  if and only if they have a common factor in  $k(x_2, \dots, x_n)[x_1]$

**Theorem 20.1.49.** Every non-constant  $f \in k[x_1, \dots, x_n]$  can be written as a product  $f = f_1 \cdots f_r$  of irreducibles of  $k$ . Furthermore, if  $f = g_1 \cdots g_s$ , where the  $g_k$  are irreducible, then  $r = s$  and there are constants  $\alpha_1, \dots, \alpha_n$  such that  $\{f_1, \dots, f_r\} = \{\alpha_1 g_1, \dots, \alpha_r g_r\}$ .

**Theorem 20.1.50.** If  $f, g \in k[x]$  are polynomials of degree  $\ell > 0$  and  $m > 0$ , respectively, then  $f$  and  $g$  have a common factor if and only if there are polynomials  $A, B \in k[x]$  such that  $A$  and  $B$  are not both zero,  $A$  has degree at most  $m - 1$  and  $B$  has degree at most  $\ell - 1$ , and  $Af + Bg = 0$ .

**Definition 20.1.47** If  $f = a_0 x^\ell + \dots + a_\ell$  and  $g = b_0 x^m + \dots + b_m$ , then the Sylvester Matrix is:

$$\begin{pmatrix} a_0 & 0 & 0 & 0 & b_0 & 0 & 0 & 0 \\ a_1 & a_0 & 0 & 0 & b_1 & b_0 & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & \vdots & \vdots & \ddots & 0 \\ \vdots & \vdots & \ddots & a_0 & \vdots & \vdots & \ddots & b_0 \\ a_\ell & \dots & \dots & a_1 & b_m & \dots & \dots & 0 \\ 0 & a_\ell & \dots & \vdots & 0 & b_m & \dots & \vdots \\ 0 & 0 & \ddots & 0 & 0 & \dots & \ddots & 0 \\ 0 & \dots & \dots & a_\ell & 0 & \dots & \dots & b_m \end{pmatrix}$$

**Theorem 20.1.51.** If  $f, g \in k[x]$ , then the resultant of  $f$  and  $g$  is the determinant of the Sylvester matrix of  $f$  and  $g$ .

**Theorem 20.1.52.** If  $f, g \in k[x]$  are polynomials of positive degree, then the resultant of  $f$  and  $g$  is an integer polynomial in the coefficients of  $f$  and  $g$ .

**Theorem 20.1.53.** *If  $f, g \in k[x]$  are polynomials of positive degree, then  $f$  and  $g$  have a common factor if and only if their resultant is zero.*

**Theorem 20.1.54.** *If  $f, g \in k[x]$  are of positive degree, then there are polynomials  $A, B \in k[x]$  such that  $Af + Bg = \text{Resultant}(f, g)$*

### 20.1.3 Groebner Bases

#### Introduction

There are three problems we wish to address:

1. Does every Ideal  $I \subset k[x_1, \dots, x_n]$  have a finite generating set?
2. Given  $f \in k[x_1, \dots, x_n]$ , and  $I = \langle f_1, \dots, f_s \rangle$ , can we determine if  $f \in I$ ?
3. For  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ , can we determine what  $\mathbf{V}(f_1, \dots, f_s)$  is?

We've already solved this in the case of one variable,  $n = 1$ . The case of  $n \in \mathbb{N}$  where  $f_1, \dots, f_s$  are linear functions is the subject of linear algebra. Both the Euclidean algorithm and the methods of linear algebra require a notion of ordering of terms. In the case of one variable, if  $n > m$  we write  $x^n > x^m$ . In the case of linear algebra we usually write  $x_n > x_{n-1} > \dots > x_2 > x_1$ .

#### Orderings on the Monomials in $k[x_1, \dots, x_n]$

**Definition 20.1.48** A monomial ordering on  $k[x_1, \dots, x_n]$  is any relation  $\succ$  on  $\mathbb{N}^n$  such that:

1.  $\succ$  is a total ordering.
2. If  $\alpha \succ \beta$  and  $\gamma \in \mathbb{N}^n$ , then  $\alpha + \gamma \succ \beta + \gamma$ .
3.  $\succ$  is a well-ordering on  $\mathbb{N}^n$ .

**Theorem 20.1.55.** *An ordering  $\prec$  on  $\mathbb{N}^n$  is a well-ordering if and only if for any monotonically decreasing sequence  $\{a_n\}_{n=1}^\infty$ , there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $a_n = a_N$ .*

*Proof.* For if  $\prec$  is a well ordering, then  $\{a_n\}_{n=1}^\infty$  contains a least element  $x$ . Suppose  $a_n$  contains a strictly decreasing subsequence. But  $\prec$  is a well ordering, and therefore  $\{a_n\}_{n=1}^\infty$  contains a least element  $x$ . But again  $\prec$  is a well ordering, and thus  $\{a_n\}_{n=1}^\infty \setminus \{x\}$  contains a least element  $y$ . But then  $x \prec y$ , and  $x$  is the least element of  $\{a_n\}_{n=1}^\infty$ . Therefore there is an  $a_n$  such that  $x \preceq a_n \preceq y$ . But  $a_n$  is strictly decreasing, and therefore  $a_{n+1} \preceq x$ , and thus  $a_{n+2} \prec x$ . But  $x$  is the least element of  $\{a_n\}_{n=1}^\infty$ , a contradiction. Therefore  $a_n$  contains no strictly increasing subsequence. Suppose every decreasing sequence

eventually terminates. Let  $E \subset \mathbb{N}^n$ . Suppose there is no least element. Then we can construct a strictly decreasing sequence. But every decreasing sequence eventually terminates, a contradiction. Therefore, etc.  $\square$

**Definition 20.1.49** If  $\alpha, \beta \in \mathbb{N}^n$ , then  $\alpha$  is said to be lexicographically greater than  $\beta$ , denoted  $\underset{Lex}{>}$ , if the left-most entry of  $\alpha - \beta$  is positive.

**Theorem 20.1.56.** *The Lexicographic Ordering is a monomial ordering.*

**Definition 20.1.50** The graded lexicographic ordering  $\underset{GrLex}{>}$  on  $\mathbb{N}^n$  is an ordering on  $\mathbb{N}^n$  such that  $\alpha \underset{GrLex}{>} \beta$  if and only if either  $|\alpha| > |\beta|$ , or  $|\alpha| = |\beta|$  and  $\alpha \underset{Lex}{>} \beta$ .

**Theorem 20.1.57.** *The graded lexicographic ordering is a monomial ordering.*

**Definition 20.1.51** For  $f = \sum_{\alpha} a_{\alpha} x^{\alpha} \in k[x_1, \dots, x_n]$ , and  $\prec$  a monomial ordering, the multidegree of  $f$  is  $\text{mutldeg}(f) = \max\{\alpha \in \mathbb{N}^n : a_{\alpha} \neq 0\}$ .

**Definition 20.1.52** For  $f = \sum_{\alpha} a_{\alpha} x^{\alpha} \in k[x_1, \dots, x_n]$  and monomial order  $>$ , the leading coefficient of  $f$  is  $LC(f) = a_{\text{mutldeg}(f)} \in k$

**Definition 20.1.53** For  $f = \sum_{\alpha} a_{\alpha} x^{\alpha} \in k[x_1, \dots, x_n]$ , and  $\prec$  a monomial ordering, the leading monomial of  $f$  is  $LM(f) = x^{\text{mutldeg}(f)}$

**Definition 20.1.54** For  $f = \sum_{\alpha} a_{\alpha} x^{\alpha} \in k[x_1, \dots, x_n]$ , and  $\prec$  a monomial ordering, the leading term of  $f$  is  $LT(f) = LC(f) \cdot LM(f)$ .

**Theorem 20.1.58.** *If  $f, g \in k[x_1, \dots, x_n]$  are non-zero, then  $\text{mutldeg}(fg) = \text{mutldeg}(f) + \text{mutldeg}(g)$*

**Theorem 20.1.59.** *If  $f, g \in k[x_1, \dots, x_n]$  are non-zero, and if  $f + g \neq 0$ , then  $\text{mutldeg}(f + g) \leq \max\{\text{mutldeg}(f), \text{mutldeg}(g)\}$ .*

**Theorem 20.1.60.** *If  $f, g \in k[x_1, \dots, x_n]$  are non-zero,  $f + g \neq 0$ , and if  $\text{mutldeg}(f) \neq \text{mutldeg}(g)$ , then  $\text{mutldeg}(f+g) = \max\{\text{mutldeg}(f), \text{mutldeg}(g)\}$ .*

**Theorem 20.1.61.** *If  $>$  is a monomial ordering on  $\mathbb{N}^n$ , and  $F = (f_1, \dots, f_s)$  is an ordered  $s$ -tuple of polynomials in  $k[x_1, \dots, x_n]$ , then every  $f \in k[x_1, \dots, x_n]$  can be written as  $f = r + \sum_{k=1}^s a_k f_k$ , where  $a_k, r \in k[x_1, \dots, x_n]$ , and either  $r = 0$  or  $r$  is a linear combination, with coefficients in  $k$ , of monomials, none of which is divisible by any of  $LT(f_1), \dots, LT(f_s)$ . We call  $r$  the remainder of  $f$  with respect to  $F$ .*

**Definition 20.1.55** An ideal  $I \subset k[x_1, \dots, x_n]$  is a monomial ideal if there is a subset  $A \subset \mathbb{N}^n$  such that  $I$  consists of all polynomials which are finite sums of the form  $\sum_{\alpha} h_{\alpha} x^{\alpha}$ , where  $h_{\alpha} \in k[x_1, \dots, x_n]$ .

**Theorem 20.1.62.** If  $I = \langle x^\alpha : \alpha \in A \rangle$  is a monomial ideal, then a monomial  $x^\beta$  lies in  $I$  if and only if  $x^\beta$  is divisible by  $x^\alpha$  for some  $\alpha \in A$ .

**Theorem 20.1.63.** If  $I$  is a monomial ideal, and  $f \in k[x_1, \dots, x_n]$ , then the following are equivalent:

1.  $f \in I$
2. Every term of  $f$  lies in  $I$ .
3.  $f$  is a  $k$ -linear combination of the monomials in  $I$ .

**Theorem 20.1.64** (Dickson's Lemma). If  $I = \langle x^\alpha : \alpha \in A \rangle$  is a monomial ideal, then  $I$  can be written as  $\langle x^{\alpha(1)}, \dots, x^{\alpha(s)} \rangle$ , where  $\alpha(1), \dots, \alpha(s) \in A$ .

**Theorem 20.1.65.** If  $>$  is a relation on  $\mathbb{N}^n$  such that  $>$  is a total ordering and for  $\alpha > \beta$  and  $\gamma \in \mathbb{N}^n$ ,  $\alpha + \gamma > \beta + \gamma$ , then  $>$  is a well-ordering if and only if for all  $\alpha \in \mathbb{N}^n$ ,  $\alpha \geq 0$ .

### The Hilbert Basis Theorem and Groebner Bases

**Definition 20.1.56** For a non-zero ideal  $I \subset k[x_1, \dots, x_n]$ ,  $\text{LT}(I)$  is the set of leading terms of elements of  $I$ .  $\langle \text{LT}(I) \rangle$  is the ideal generated by this set.

**Theorem 20.1.66.** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then  $\langle \text{LT}(I) \rangle$  is a monomial ideal.

**Theorem 20.1.67.** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then there are  $g_1, \dots, g_t \in I$  such that  $\langle \text{LT}(I) \rangle = \langle \text{LT}(g_1), \dots, \text{LT}(g_t) \rangle$

**Theorem 20.1.68** (Hilbert Basis Theorem). Every ideal  $I \subset k[x_1, \dots, x_n]$  has a finite generating set.

**Definition 20.1.57** For a monomial order  $>$ , a finite subset  $G = \{g_1, \dots, g_t\}$  of an ideal  $I$  is said to be a Groebner Basis if  $\langle \text{LT}(g_1), \dots, \text{LT}(g_t) \rangle = \langle \text{LT}(I) \rangle$

**Theorem 20.1.69.** If  $>$  is a monomial order, then every non-zero ideal  $I \subset k[x_1, \dots, x_n]$  has a Groebner basis.

**Theorem 20.1.70.** If  $I \subset k[x_1, \dots, x_n]$  is a non-zero ideal and  $G$  is a Groebner Basis, then  $G$  is also a generated set of  $I$ .

**Theorem 20.1.71** (The Ascending Chain Condition). If  $I_n$  is a sequence of ideals such that  $I_n \subset I_{n+1}$ , then there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $I_n = I_N$ .

**Definition 20.1.58** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then  $\mathbf{V}(I)$  is the set  $\{\alpha \in k^n : \forall f \in I, f(\alpha) = 0\}$

**Theorem 20.1.72.** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then  $\mathbf{V}(I)$  is an affine variety.

**Theorem 20.1.73.** If  $I = \langle f_1, \dots, f_s \rangle$ , then  $\mathbf{V}(I) = \mathbf{V}(f_1, \dots, f_s)$ .

## Properties of Groebner Bases

**Theorem 20.1.74.** *If  $G = \{g_1, \dots, g_t\}$  is a Groebner basis of  $I \subset k[x_1, \dots, x_n]$  and  $f \in k[x_1, \dots, x_n]$ , then there is a unique  $r \in k[x_1, \dots, x_n]$  such that  $r$  is not divisible by any of  $\text{LT}(g_1), \dots, \text{LT}(g_t)$ , and there is a  $g \in I$  such that  $f = g + r$ .*

We write  $\bar{f}^F$  for the remainder on division of  $f$  by  $F = (f_1, \dots, f_s)$

**Definition 20.1.59** If  $f, g \in k[x_1, \dots, x_n]$  are non-zero polynomials,  $\text{mutldeg}(f) = \alpha$ ,  $\text{mutldeg}(g) = \beta$ , and if  $\gamma = (\gamma_1, \dots, \gamma_n)$ , where  $\gamma_k = \max\{\alpha_k, \beta_k\}$ , then  $x^\gamma$  is the least common multiple of  $\text{LM}(f)$  and  $\text{LM}(g)$ , denoted  $x^\gamma = \text{LCM}(\text{LM}(f), \text{LM}(g))$ .

**Definition 20.1.60** If  $f, g \in k[x_1, \dots, x_n]$  are non-zero, then the  $S$ -polynomial of  $f$  and  $g$  is  $S(f, g) = \frac{x^\gamma}{\text{LT}(f)}f - \frac{x^\gamma}{\text{LT}(g)}g$

**Theorem 20.1.75** (Buchberger's Criterion). *If  $I$  is a polynomial ideal, then a basis  $G = \{g_1, \dots, g_t\}$  for  $I$  is a Groebner basis for  $I$  if and only if for all pairs  $i \neq j$ , the remainder on division of  $S(g_i, g_j)$  by  $G$  is zero.*

### 20.1.4 The Algebra-Geometry Dictionary

#### Hilbert's Nullstellensatz

**Theorem 20.1.76** (The Weak Nullstellensatz Theorem). *If  $k$  is an algebraically closed field,  $I \subset k[x_1, \dots, x_n]$  is an ideal, and  $\mathbf{V}(I) = \emptyset$ , then  $I = k[x_1, \dots, x_n]$ .*

**Theorem 20.1.77** (Hilbert's Nullstellensatz). *If  $k$  is an algebraically closed,  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ , and if  $f \in \mathbf{I}(\mathbf{V}(f_1, \dots, f_s))$ , then  $\exists_{m \in \mathbb{N}}$  such that  $f^m \in \langle f_1, \dots, f_s \rangle$ .*

#### Radical Ideals and the Ideal-Variety Correspondence

**Theorem 20.1.78.** *If  $V$  is an affine variety, and if  $f \in \mathbf{I}(V)$ , then  $f^m \in \mathbf{I}(V)$ .*

**Definition 20.1.61** An ideal  $I$  is said to be radical  $f^m \in I$  implies  $f \in I$  for some  $m \geq 1$ .

**Theorem 20.1.79.** *If  $V$  is an affine variety, then  $\mathbf{I}(V)$  is a radical ideal.*

**Definition 20.1.62** The radical of an ideal  $I \subset k[x_1, \dots, x_n]$  is the set  $\sqrt{I} = \{f : f^m \in I, m \in \mathbb{N}\}$ .

**Theorem 20.1.80.** *If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then  $\sqrt{I}$  is an ideal.*

**Theorem 20.1.81** (The Strong Nullstellensatz). *If  $k$  is an algebraically closed, and  $I \subset k[x_1, \dots, x_n]$  is an ideal, then  $\mathbf{I}(\mathbf{V}(I)) = \sqrt{I}$ .*

**Theorem 20.1.82** (The Ideal-Variety Correspondence). *If  $k$  is a field, then the maps affine varieties  $\xrightarrow{I}$  ideals and ideals  $\xrightarrow{V}$  affine varieties are inclusion reversing and for any affine variety  $V$ ,  $\mathbf{V}(I(V)) = V$ .*

**Theorem 20.1.83** (Radical Membership Theorem). *If  $k$  is a field and  $I = \langle f_1, \dots, f_s \rangle \subset k[x_1, \dots, x_n]$  is an ideal, then  $f \in \sqrt{I}$  if and only if the constant polynomial 1 belongs to  $\langle f_1, \dots, f_s, 1 - yf \rangle$ .*

**Theorem 20.1.84.** *If  $f \in k[x_1, \dots, x_n]$ , and  $I = \langle f \rangle$ , and if  $f = f_1^{\alpha_1} \cdots f_s^{\alpha_s}$ , then  $\sqrt{I} = \langle f_1 \cdots f_s \rangle$ .*

**Definition 20.1.63** The reduction of a polynomial  $f \in k[x_1, \dots, x_n]$  is the polynomial  $f_{red}$  such that  $\langle f_{red} \rangle = \sqrt{\langle f \rangle}$ .

**Definition 20.1.64** A square free polynomial is a polynomial  $f \in k[x_1, \dots, x_n]$  such that  $f = f_{red}$ .

**Definition 20.1.65** If  $f, g \in k[x_1, \dots, x_n]$ , then  $h \in k[x_1, \dots, x_n]$  is said to be the greatest common divisor of  $f$  and  $g$  if  $f$  divides  $h$  and  $g$ , and if  $p$  is any polynomial that divides  $f$  and  $g$ , then  $p$  divides  $h$ .

**Theorem 20.1.85.** *If  $k$  is a field such that  $\mathbb{Q} \subset k$ , and  $I = \langle f \rangle$  for some  $f \in k[x_1, \dots, x_n]$ , then  $\sqrt{I} = \langle f_{red} \rangle$ , where  $f_{red} = \frac{f}{GCD(f, \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n})}$*

### Sums, Products, and Intersections of Ideals

**Definition 20.1.66** If  $I$  and  $J$  are ideals of a the ring  $k[x_1, \dots, x_n]$ , then the sum of  $I$  and  $J$ , denoted  $I + J$ , is the set  $I + J = \{f + g : f \in I, g \in J\}$ .

**Theorem 20.1.86.** *If  $I$  and  $J$  are ideals in  $k[x_1, \dots, x_n]$ , then  $I + J$  is also an ideal in  $k[x_1, \dots, x_n]$ .*

**Theorem 20.1.87.** *If  $I$  and  $J$  are ideals in  $k[x_1, \dots, x_n]$ , then  $I + J$  is the smallest ideal containing  $I$  and  $J$ .*

**Theorem 20.1.88.** *If  $f_1, \dots, f_r \in k[x_1, \dots, x_n]$ , then  $\langle f_1, \dots, f_r \rangle = \sum_{k=1}^r \langle f_k \rangle$*

**Theorem 20.1.89.** *If  $I$  and  $J$  are ideals in  $k[x_1, \dots, x_n]$ , then  $\mathbf{V}(I + J) = \mathbf{V}(I) \cap \mathbf{V}(J)$ .*

**Definition 20.1.67** If  $I$  and  $J$  are two ideals in  $k[x_1, \dots, x_n]$ , then their product, denoted  $I \cdot J$ , is defined to be the ideal generated by all polynomials  $f \cdot g$ , where  $f \in I$ , and  $g \in J$ .

**Theorem 20.1.90.** *If  $I = \langle f_1, \dots, f_r \rangle$  and  $J = \langle g_1, \dots, g_s \rangle$ , then  $I \cdot J$  is generated by the set of all products  $\{f_i g_j : 1 \leq i \leq r, 1 \leq j \leq s\}$*

**Theorem 20.1.91.** If  $I, J \subset k[x_1, \dots, x_n]$  are ideals, then  $\mathbf{V}(I \cdot J) = \mathbf{V}(I) \cup \mathbf{V}(J)$ .

**Definition 20.1.68** If  $I, J \subset k[x_1, \dots, x_n]$  are ideals, then the intersection of  $I$  and  $J$ , denoted  $I \cap J$ , is the set of polynomials in both  $I$  and  $J$ .

**Theorem 20.1.92.** If  $I, J \subset k[x_1, \dots, x_n]$  are ideals, then  $I \cap J$  is an ideal.

### Zariski Closure and Quotients of Ideals

**Theorem 20.1.93.** If  $S \subset k^n$ , then the affine variety  $\mathbf{V}(I(S))$  is the smallest affine variety that contains  $S$ .

**Definition 20.1.69** The Zariski Closure of a subset  $S$ , denoted  $\overline{S}$ , of an affine space is the smallest affine algebraic variety containing the set.

**Theorem 20.1.94.** If  $k$  is an algebraically closed field and  $V = \mathbf{V}(f_1, \dots, f_s) \subset k^n$ , then  $\mathbf{V}(I_\ell)$  is the Zariski Closure of  $\pi_\ell(V)$ .

**Theorem 20.1.95.** If  $V$  and  $W$  are varieties such that  $V \subset W$ , then  $W = V \cup \overline{(W \setminus V)}$ .

**Definition 20.1.70** If  $I, J \subset k[x_1, \dots, x_n]$  are ideals, then  $I : J$  is the set,  $\{f \in k[x_1, \dots, x_n] : fg \in I \forall g \in J\}$  and is called the ideal quotient of  $I$  by  $J$ .

**Theorem 20.1.96.** If  $I, J \subset k[x_1, \dots, x_n]$  are ideals, then  $I : J$  is an ideal.

**Theorem 20.1.97.** If  $I, J \subset k[x_1, \dots, x_n]$  are ideals, then  $\overline{\mathbf{V}(I) \setminus \mathbf{V}(J)} \subset \mathbf{V}(I : J)$ .

**Theorem 20.1.98.** If  $I, J \subset k^n$  are affine varieties, then  $\mathbf{I}(V) : \mathbf{I}(W) = \mathbf{I}(V \setminus W)$

**Theorem 20.1.99.** If  $I, J, K \subset k[x_1, \dots, x_n]$ , then  $I : k[x_1, \dots, x_n] = I$ .

**Theorem 20.1.100.** If  $I, J, K \subset k[x_1, \dots, x_n]$  are ideals, then  $I \cdot J \subset K$  if and only if  $I \subset K : J$

**Theorem 20.1.101.** If  $I, J, K \subset k[x_1, \dots, x_n]$  are ideals, then  $J \subset I$  if and only if  $I : J = k[x_1, \dots, x_n]$

**Theorem 20.1.102.** If  $I$  is an ideal,  $g \in k[x_1, \dots, x_n]$ , and if  $\{h_1, \dots, h_p\}$  is a basis of the ideal  $I \cap \langle g \rangle$ , then  $\{h_1/g, \dots, h_p/g\}$  is a basis of  $I : \langle g \rangle$ .

### Irreducible Varieties and Prime Ideals

**Definition 20.1.71** An affine variety  $V \subset k^n$  is irreducible if there are no affine varieties  $V_1, V_2$ , such that  $V = V_1 \cup V_2$ ,  $V_1, V_2 \neq \emptyset$ , and  $V_1 \neq V, V_2 \neq V$ .

**Definition 20.1.72** An ideal  $I \subset k[x_1, \dots, x_n]$  is said to be prime if whenever  $f, g \in k[x_1, \dots, x_n]$  and  $fg \in I$ , either  $f \in I$  or  $g \in I$ .

**Theorem 20.1.103.** *If  $V \subset k^n$  is an affine variety, then  $V$  is irreducible if and only if  $\mathbf{I}(V)$  is a prime ideal.*

**Definition 20.1.73** An ideal  $I \subset k[x_1, \dots, x_n]$  is said to be maximal if  $I \neq k[x_1, \dots, x_n]$  and any ideal  $J$  containing  $I$  is such that either  $J = I$  or  $J = k[x_1, \dots, x_n]$ .

**Definition 20.1.74** An ideal  $I \subset k[x_1, \dots, x_n]$  is called proper if  $I$  is not equal to  $k[x_1, \dots, x_n]$ .

**Theorem 20.1.104.** *If  $k$  is a field and  $I = \langle x_1 - a_1, \dots, x_n - a_n \rangle$  is an ideal where  $a_1, \dots, a_n \in k$ , then  $I$  is maximal.*

**Theorem 20.1.105.** *If  $k$  is a field, then any maximal ideal is also a prime ideal.*

**Theorem 20.1.106.** *If  $k$  is an algebraically closed field, then every maximal ideal of  $k[x_1, \dots, x_n]$  is of the form  $\langle x_1 - a_1, \dots, x_n - a_n \rangle$  for some  $a_1, \dots, a_n \in k$ .*

**Definition 20.1.75** A primary decomposition of an ideal  $I$  is an expression of  $I$  as an intersection of primary ideals  $I = \cap_{i=1}^r Q_i$ .

**Definition 20.1.76** A primary decomposition of an ideal  $I$  is said to be minimal if  $\sqrt{Q_i}$  are all distinct and  $\cap_{j \neq i} Q_j \not\subset Q_i$ .

**Theorem 20.1.107.** *If  $I, J$  are primary and  $\sqrt{I} = \sqrt{J}$ , then  $I \cap J$  is primary.*

**Theorem 20.1.108** (Lasker-Noether Theorem). *Every ideal  $I \subset k[x_1, \dots, x_n]$  has a minimal primary decomposition.*

## 20.1.5 Polynomials and Rational Functions on a Variety

### Polynomial Mappings

**Definition 20.1.77** If  $V \subset k^m$ ,  $W \subset k^n$  are affine varieties, a function  $\phi : V \rightarrow W$  is said to be a polynomial mapping if there exist polynomials  $f_1, \dots, f_n \in k[x_1, \dots, x_m]$  such that  $\phi(a_1, \dots, a_m) = (f_1(a_1, \dots, a_m), \dots, f_n(a_1, \dots, a_m))$  for all  $(a_1, \dots, a_m) \in V$ . We say that  $(f_1, \dots, f_n)$  represents  $\phi$ .

**Theorem 20.1.109.** *If  $V \subset k^m$  is an affine variety, then  $f, g \in k[x_1, \dots, x_m]$  represent the same polynomial on  $V$  if and only if  $f - g \in \mathbf{I}(V)$ .*

**Theorem 20.1.110.** *If  $V \subset k^m$  is an affine variety, then  $(f_1, \dots, f_n)$  and  $(g_1, \dots, g_n)$  represent the same polynomial mapping if and only if  $f_i - g_i \in \mathbf{I}(V)$  for  $1 \leq i \leq n$ .*

The set of polynomial mappings from  $V$  to  $k$  is denoted  $k[V]$ .

**Theorem 20.1.111.** *If  $V \subset k^n$  is an affine variety, the the following are equivalent:*

1.  $V$  is irreducible.      2.  $\mathbf{I}(V)$  is a prime ideal. 3.  $k[V]$  is an integral domain.

### Quotients of Polynomial Rings

**Definition 20.1.78** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, if  $f, g \in k[x_1, \dots, x_n]$ , then  $f$  and  $g$  are congruent modulo  $I$ , denoted  $f \equiv g \pmod{I}$ , if  $f - g \in I$ .

**Theorem 20.1.112.** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then the congruence modulo  $I$  is an equivalence relation on  $k[x_1, \dots, x_n]$ .

**Theorem 20.1.113.** There exists a bijection from the set of distinct polynomial functions  $\phi : V \rightarrow k$  and the set of equivalence classes of polynomials under congruence modulo  $\mathbf{I}(V)$ .

**Definition 20.1.79** The quotient of  $k[x_1, \dots, x_n]$  modulo  $I$ , denoted  $k[x_1, \dots, x_n]/I$ , is the set of equivalence classes for congruence modulo  $I$ .

**Theorem 20.1.114.** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then  $k[x_1, \dots, x_n]/I$  is a commutative ring under the sum and product operations.

**Definition 20.1.80** A ring isomorphism of rings  $R$  and  $S$  is a bijective function  $\phi : R \rightarrow S$  such that:

1. For all  $a, b \in R$ ,  $\phi(a + b) = \phi(a) + \phi(b)$

**Theorem 20.1.115.** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then there is a bijection between the ideals in the quotient ring  $k[x_1, \dots, x_n]/I$  and the ideals of  $k[x_1, \dots, x_n]$  that contain  $I$ .

**Theorem 20.1.116.** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then every ideal of  $k[x_1, \dots, x_n]/I$  is finitely generated.

### The Coordinate Ring of an Affine Variety

**Definition 20.1.81** The coordinate ring of an affine variety  $V \subset k^n$  is the ring  $k[V]$ .

**Definition 20.1.82** If  $V \subset k^n$  is an affine variety, and if  $J = \langle \phi_1, \dots, \phi_s \rangle \subset k[V]$ , then  $\mathbf{V}_V(J) = \{x \in V : \forall_{\phi \in J}, \phi(x) = 0\}$  is called the subvariety of  $V$ .

**Theorem 20.1.117.** If  $V \subset k^n$  is an affine variety and if  $J \subset k[V]$  is an ideal, then  $W = \mathbf{V}_V(J)$  is an affine variety in  $k^n$  contained in  $V$ .

**Theorem 20.1.118.** If  $V \subset k^n$  is an affine variety, and if  $W \subset V$ , then  $\mathbf{V}_V(W)$  is an ideal of  $k[V]$ .

**Definition 20.1.83** If  $V$  is an irreducible variety in  $k^n$ , then the function field, denoted  $QF(k[V])$ , on  $V$  is the quotient field of  $k[V]$ .

**Definition 20.1.84** If  $V \subset k^m$  and  $W \subset k^n$  are irreducible affine varieties, then a rational mapping is a function  $\phi$  such that  $\phi(x_1, \dots, x_m) = \left( \frac{f_1(x_1, \dots, x_m)}{g_1(x_1, \dots, x_m)}, \dots, \frac{f_n(x_1, \dots, x_m)}{g_n(x_1, \dots, x_m)} \right)$ .

**Theorem 20.1.119.** Two rational mappings  $\phi, \psi : V \rightarrow W$  are equal if and only if there is a proper subvariety  $V' \subset V$  such that  $\phi$  and  $\psi$  are defined on  $V \setminus V'$  and  $\phi(p) = \psi(p)$  for all  $p \in V \setminus V'$ .

**Theorem 20.1.120** (The Closure Theorem). If  $k$  is an algebraically closed field,  $V = \mathbf{V}(I)$ ,  $V \neq \emptyset$ , then there is an affine variety  $W \underset{\text{Proper}}{\subset} \mathbf{V}(I_\ell)$  such that  $\mathbf{V}(I_\ell) \setminus W \subset \pi_\ell(V)$ .

## 20.2 Miscellaneous Notes

### 20.2.1 Groebner Bases

**Definition 20.2.1** A ring is a set  $R$  with two binary operations  $+$  and  $\cdot$ , called addition and multiplication, such that the following are true:

1.  $(R, +)$  is an Abelian Group.  $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
2.  $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$

**Definition 20.2.2** A commutative ring is a ring  $R$  such that  $\forall_{a,b \in R}, a \cdot b = b \cdot a$

**Definition 20.2.3** A ring with identity is a ring  $R$  such that  $\exists 1_R \in R : \forall_{a \in R}, 1_R \cdot a = a \cdot 1_R = a$

**Definition 20.2.4** A subring of a ring with identity  $R$  is a set  $S \subset R$  such that  $1_R \in S$ , and  $S$  is closed under the ring operations.

**Definition 20.2.5** A monomial in variables  $x_1, \dots, x_n$  over a ring  $R$  is a product  $x^\alpha = \prod_{k=1}^n x_1^{\alpha_1}$ , where  $(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ .

The set of monomials in  $n$  variables over  $R$  is denoted  $\text{Mon}_R(x_1, \dots, x_n)$ .

**Definition 20.2.6** If  $\alpha, \beta \in \mathbb{N}^n$  such that  $\alpha_i \leq \beta_i$ , then  $x^\alpha$  is said to divide  $x^\beta$ , denoted  $x^\alpha | x^\beta$ , if  $x^\beta = x^\alpha \cdot x^\gamma$  for some  $\gamma \in \mathbb{N}^n$ .

**Definition 20.2.7** A term is a monomial multiplied by a coefficient in  $R$ .

**Definition 20.2.8** A polynomial over  $R$  is a finite  $R$ -linear combination of monomials,  $f = \sum_{\alpha} a_{\alpha} \cdot x^{\alpha}$ .

The set of all polynomials in  $n$  variables over a ring  $R$  is denoted  $R[x_1, \dots, x_n]$ .

**Theorem 20.2.1.** *If  $R$  is a commutative ring with identity, then  $R[x_1, \dots, x_n]$  is a commutative ring with identity.*

**Definition 20.2.9** A polynomial  $f \in R[x_1, \dots, x_n]$  is called a constant polynomial if  $f \in R$ .

**Definition 20.2.10** A field  $k$  is a commutative ring with identity such that for all  $a \in k$ ,  $a \neq 0$ , there is a  $b \in k$  such that  $a \cdot b = 1$

We usually work with fields and consider polynomial rings of the form  $k[x_1, \dots, x_n]$ .

**Definition 20.2.11** A total ordering on a set  $A$  is a relation  $>$  such that  $\forall_{a,b \in A}$ , precisely one of the following true:

1.  $a < b$
2.  $a = b$
3.  $b < a$

**Definition 20.2.12** A relation  $\sim$  on a set  $A$  is said to be transitive if for all  $a, b, c \in A$ , if  $a \sim b$  and  $b \sim c$ , then  $a \sim c$ .

**Definition 20.2.13** A well ordering on a set  $A$  is a relation  $<$  such that for every subset  $E \subset A$ , there is an element  $x \in E$  such that for all  $y \in E$ ,  $y \neq x$ , we have  $x < y$ .

Equivalently, a well ordering on a set  $A$  is a relation  $<$  such that for every monotonically decreasing sequence  $\alpha_n$ , there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $\alpha_n = \alpha_N$ . That is, decreasing sequences terminate.

**Definition 20.2.14** A monomial ordering on  $\mathbb{N}^n$  is a relation  $>$  such that  $>$  is total, transitive, well ordering. A well ordering on  $k[x_1, \dots, x_n]$  is a well ordering on  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ .

**Definition 20.2.15** The lexicographic ordering on  $\mathbb{N}^n$  is defined as  $(\alpha_1, \dots, \alpha_n) >_{Lex} (\beta_1, \dots, \beta_n)$  if the left-most non-zero entry of  $(\alpha_1 - \beta_1, \dots, \alpha_n - \beta_n)$  is positive.

**Theorem 20.2.2.** *The lexicographic ordering is a monomial ordering.*

**Definition 20.2.16** The graded lexicographic ordering is defined as  $(\alpha_1, \dots, \alpha_n) >_{GrLex} (\beta_1, \dots, \beta_n)$  if  $|\alpha| > |\beta|$  or  $|\alpha| = |\beta|$  and  $\alpha >_{Lex} \beta$ .

**Theorem 20.2.3.** *The graded lexicographic ordering is a monomial ordering.*

**Theorem 20.2.4** (The Division Algorithm). *If  $f_1, \dots, f_s \in k[x_1, \dots, x_n]$  are non-zero polynomials and if  $>$  is a monomial ordering, then there are  $r, q_1, \dots, q_n \in k[x_1, \dots, x_n]$  such that the following are true:*

1.  $f = q_1 f_1 + \dots + q_s f_s + r$
2. No term of  $r$  is divisible by any of  $\text{LT}(f_1), \dots, \text{LT}(f_s)$ .
3.  $\text{LT}(f) = \max_{>} \{\text{LT}(q_i) \cdot \text{LT}(f_i) : q_i \neq 0\}$

**Definition 20.2.17** An ideal  $I = \langle x^\alpha : \alpha \in A \rangle = \{\sum_\alpha h_\alpha x^\alpha, h_\alpha \in k[x_1, \dots, x_n]\}$  is called a monomial ideal.

**Theorem 20.2.5.** If  $I = \langle x^\alpha : \alpha \in A \rangle$  is a monomial ideal,  $\beta \in \mathbb{N}^n$ , then  $x^\beta \in I$  if and only if there is an  $\alpha \in A$  such that  $x^\alpha$  divides  $x^\beta$ .

**Theorem 20.2.6.** If  $I$  is a monomial ideal,  $f \in k[x_1, \dots, x_n]$ , then the following are equivalent:

1.  $f \in I$
2. Every term of  $f$  lies in  $I$ .
3.  $f$  is a  $k$ -linear combination of monomials in  $I$ .

**Theorem 20.2.7** (Dickson's Lemma). Every monomial ideal of  $k[x_1, \dots, x_n]$  is finitely generated.

**Theorem 20.2.8** (Hilbert's Basis Theorem). Every ideal  $I \subset k[x_1, \dots, x_n]$  is finitely generated.

**Definition 20.2.18** If  $>$  is a monomial ordering on  $k[x_1, \dots, x_n]$ , then a Groebner Basis of  $I \subset k[x_1, \dots, x_n]$  is a set  $G = \{g_1, \dots, g_s\}$  such that  $\langle \text{LT}(I) \rangle = \langle \text{LT}(g_1), \dots, \text{LT}(g_s) \rangle$

**Theorem 20.2.9.** Every non-zero ideal  $I \subset k[x_1, \dots, x_n]$  has a Groebner Basis.

## 20.2.2 Elimination Theory

**Definition 20.2.19** If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then the  $i^{\text{th}}$  elimination ideal of  $I$ , denoted  $I_i$ , is the set  $I_i = I \cap k[x_{i+1}, \dots, x_n]$ , where  $1 \leq i \leq n$ , and  $I_0 = I$ .

**Theorem 20.2.10** (The Elimination Theorem). If  $I \subset k[x_1, \dots, x_n]$  is an ideal and  $G$  is a Groebner Basis of  $I$  with respect to the lexicographic ordering, and  $x_1 > \dots > x_n$ , then for all  $i = 0, 1, \dots, n$ , the set  $G_i \cap k[x_1, \dots, x_n]$  is a Groebner Basis of the  $i^{\text{th}}$  elimination ideal  $I_i$ .

Using the lexicographic ordering, and for some ideal  $I = \langle f_1, \dots, f_s \rangle \subset k[x_1, \dots, x_n]$ , to compute all elimination ideals  $I_i$ :

1. Compute a Groebner Basis  $G$  for  $I$  with respect to the lex order on  $k[x_1, \dots, x_n]$ .
2. For all  $i$ , the elements  $g \in G$  with  $\text{LT}(g) \in k[x_{i+1}, \dots, x_n]$  form a Groebner basis  $I_i$  with respect to the lexicographic ordering on  $k[x_{i+1}, \dots, x_n]$ .

$\tilde{A}$

**Definition 20.2.20** A monomial order on  $k[x_1, \dots, x_n, y_1, \dots, y_m]$  is an elimination order with respect to  $x_1, \dots, x_n$  if the following holds for all  $f \in k[x_1, \dots, x_n, y_1, \dots, y_m]$ :  $L(f) \in k[y_1, \dots, y_m] \Rightarrow f \in k[y_1, \dots, y_m]$

**Theorem 20.2.11** (The Extension Theorem). *If  $k$  is an algebraically closed field,  $I = \langle f_1, \dots, f_s \rangle$ ,  $I_1$  is the first elimination ideal of  $I$ , and if  $f_i = g_i(x_2, \dots, x_n)x_1^{N_i} + r_i$ , where  $r_i$  contains only terms where the degree of  $x_1$  is less than  $N_i$ , and if  $(a_2, \dots, a_n) \in k^{n-1}$  such that  $(a_2, \dots, a_n) \notin \mathbf{V}(g_1, \dots, g_s)$ , then there is an  $a_1 \in k$  such that  $(a_1, \dots, a_n) \in \mathbf{V}(I_1)$ .*

**Definition 20.2.21** The  $k^{\text{th}}$  projection map on  $k^n$  is  $\pi_k : k^n \rightarrow k^{n-k}$  defined by  $(a_1, \dots, a_n) = (a_{k+1}, \dots, a_n)$

If  $I \subset k[x_1, \dots, x_n]$  is an ideal,  $X = \mathbf{V}(I)$ , and  $f \in I_k$ , then  $f(X) = 0$ . Thus  $f(\pi_k(X)) = 0$ , and therefore  $\pi_k(X) \subset \mathbf{V}(I_k)$ . Also  $\pi_k(X)$  may NOT be Zariski closed.

**Theorem 20.2.12.** *If  $k$  is algebraically closed, then  $\overline{\pi_k(X)} = \mathbf{V}(I_k)$ .*

**Theorem 20.2.13.** *If  $k$  is an infinite field,  $F : k^m \rightarrow k^n$  a function determined by some parametrization  $x_j = f_j(t_1, \dots, t_m)$ , and if  $I = \langle x_1 - f_1, \dots, x_n - f_n \rangle$ , then  $\mathbf{V}(I_m)$  is the smallest algebraic set in  $k^n$  containing  $F(k^m)$ .*

$V(I_m)$  is the Zariski closure of  $F(k^m)$ .

### 20.2.3 Étale Cohomology

#### Review of Schemes

Limitations of Affine Varieties:

- One would like to construct spaces by gluing together simpler pieces, like in geometry and topology.
- Difficult over non-algebraically closed fields.
- Keeping track of multiplicities.

Grothendieck's Theory of Schemes gives solutions to these problems. Should  $x^2 + y^2 = -1$  and  $x^2 + y^2 = 3$  be regarded as the same over  $\mathbb{A}_{\mathbb{R}}^2$ ? They both

have no solution. The answer is no. An isomorphism should be given by an invertible transformation. In general, the affine variety  $X \subset \mathbb{A}_R^n$  is completely determined by the coordinate ring  $S = \mathcal{O}(X) = R[x_1, \dots, x_n]/(f_1, \dots, f_N)$ . Given a compact Hausdorff space  $X$ , let  $C(X)$  denote the set of continuous complex valued functions. This is a commutative ring with identity. With the supremum norm, it becomes a unital  $C^*$ -algebra.

**Theorem 20.2.14.** *The map  $X \rightarrow \max\{C(X)\}$  is a homeomorphism.*

Given a continuous map of spaces  $f : X \rightarrow Y$ , we get a homomorphism  $C(Y) \rightarrow C(X)$  given by  $g \mapsto g \circ f$ . Thus  $C(X)$  can be regarded as a contravariant functor.

**Theorem 20.2.15** (Gelfand). *The functor  $X \mapsto C(X)$  induces an equivalence between the category of compact Hausdorff spaces and the opposite category of commutative unital  $C^*$ -algebras.*

**Definition 20.2.22** The spectrum of  $R$ , denoted  $\text{Spec}(R)$ , is the set of prime ideals of  $R$ .

**Theorem 20.2.16.** *The Zariski topology on  $\text{Spec}(R)$  contains open sets  $D(f) = \{p \in \text{Spec}(R) : f \notin p\}$*

A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is  $C^\infty$  if and only if its restriction to the neighborhood of every point is  $C^\infty$ . That is,  $f \in C^\infty(X)$  if and only if for any open cover  $\{U_i\}$ ,  $f|_{U_i} \in C^\infty(U_i)$ .

**Definition 20.2.23** If  $X$  is a topological space, a presheaf of sets  $\mathcal{F}$  is a collection of sets  $\mathcal{F}(U)$  for each open set  $U \subset X$  together with maps  $\rho_{UV} : \mathcal{F}(U) \rightarrow \mathcal{F}(V)$  for each pair  $U \subset V$  such that  $\rho_{UU} = id$  and  $\rho_{WV} \circ \rho_{VU} = \rho_{WU}$  whenever  $U \subset V \subset W$ .

**Definition 20.2.24** A sheaf is a presheaf such that for any open cover  $\{U_i\}$  of an open  $U \subset X$  and section  $f_i \in \mathcal{F}(U_i)$  such that  $F_i|_{U_i \cap U_j} = f_j|_{U_i \cap U_j}$ , there is a unique  $f \in \mathcal{F}(U)$  such that  $f_i = f|_{U_i}$ .

**Definition 20.2.25** A ringed space is a pair  $(X, \mathcal{O}_X)$ , where  $X$  is a topological space and  $\mathcal{O}_X$  is a sheaf of commutative rings.

The collection of presheaves of a topological space form a category, denoted  $Sh(X)$ .

**Definition 20.2.26** A scheme is a ringed space  $(X, \mathcal{O}_X)$  which is locally an affine space.

**Theorem 20.2.17.** *A property of commutative rings extends to schemes if it is local.*

## Differential Calculus of Schemes

**Definition 20.2.27** The tangent space of an affine variety  $X = V(f_1, \dots, f_N) \subset \mathbb{A}_k^n$ , denoted  $T_{X,p}$ , is the set of points  $v \in k^n$  such that  $\sum \frac{\partial f_i}{\partial x_i}(p)v_i = 0$

**Definition 20.2.28** A domain  $R = k[x_1, \dots, x_n]/(f_1, \dots, f_N)$  or  $\text{Spec}(R)$  is smooth if and only if the rank of  $(\frac{\partial f_j}{\partial x_i}(p))$  is  $n = \dim(R)$  for all  $p \in \max(R)$ .

**Definition 20.2.29** An étale  $R$  algebra is smooth of relative dimension 0, where  $\det(\frac{\partial f_i}{\partial x_j})$  is a unit in  $S$ .

**Theorem 20.2.18.** *If  $k$  is a field, then an algebra over  $k$  is étale if and only if it is a finite Cartesian product of separable field extensions.*

**Theorem 20.2.19.** *The tensor product of two étale algebras is étale.*

**Theorem 20.2.20.** *If  $S$  is étale over  $R$  and  $T$  is étale over  $S$ , then  $T$  is étale over  $R$ .*

**Definition 20.2.30** If  $R$  is a commutative ring and  $S$  is an  $R$  algebra and  $M$  is an  $S$  module, then an  $R$  linear derivation from  $S$  to  $M$  is a map  $\delta : S \rightarrow M$  such that  $\delta(s_1 + s_2) = \delta(s_1) + \delta(s_2)$ ,  $\delta(s_1 s_2) = s_1 \delta(s_2) + s_2 \delta(s_1)$ , and  $\delta(r) = 0$  for all  $r \in R$ .

**Theorem 20.2.21.** *There exists an  $S$  module  $\Omega_{S/R}$  with a universal  $R$  linear derivation  $d : S \rightarrow \Omega_{S/R}$ .*

**Theorem 20.2.22.** *If  $M$  is a finitely generated module over a Noetherian ring  $R$ , then these are equivalent:*

1.  $M$  is locally free.
2.  $\forall_{p \in \text{Spec}(R)}, R_p \otimes M$  is free.

**Definition 20.2.31** If  $M$  is an  $S$ -module, then  $M$  is called flat if  $M \otimes i$  is injective for any  $i$ .

**Theorem 20.2.23.** *If  $S$  is an  $R$  algebra and  $M$  is an  $S$  module,  $f \in S$  is an element such that multiplication by  $f$  is injective on  $M \otimes k(m)$  for all  $m \in \max(R)$ , and if  $M$  is flat over  $R$ , then  $M/fM$  is flat over  $R$ .*

**Theorem 20.2.24.** *A smooth algebra is flat.*

**Theorem 20.2.25.** *If  $R$  is a Noetherian ring, then a homomorphism  $R \rightarrow S$  is étale if and only if:*

1.  $S$  is finitely generated as an algebra.
2.  $S$  is flat as an  $R$ -module.

$$3. \Omega_{S/R} = 0.$$

**Definition 20.2.32** A sheaf on a scheme is quasi-coherent if it is with respect to some affine open cover.

**Theorem 20.2.26.** *If  $f : X \rightarrow Y$  is a morphism, there exists a quasi-coherent sheaf  $\Omega_{X/Y}$  such that  $\Omega_{X/Y}|_{\text{Spec}(S_{ij})} = \Omega_{S_{ij}/R_i}$  for open affine covers  $\text{Spec}(R_i) = U_i$ .*

### The Fundamental Étale Group

**Definition 20.2.33** A topological group is a topological space  $(X, \tau)$  with a group structure  $(X, *)$  such that  $* : X \times X \rightarrow X$  is a continuous function with respect to the product topology.

**Theorem 20.2.27.** *A topological space is profinite if and only if it is compact Hausdorff and totally disconnected.*

**Definition 20.2.34** The topological fundamental group of a topological space  $X$ , denoted  $\pi_1(X)$ , is the group of homotopy classes of loops in  $X$  with a given base point.

**Theorem 20.2.28.** *Any étale morphism  $Y \rightarrow X$  is a finite to one covering space of  $X$  with the usual topology.*

**Theorem 20.2.29** (Grothendieck's Theorem). *If  $X$  is a scheme of finite type of  $\mathbb{C}$ , then  $\pi_1^{et}(X)$  is the profinite completion of  $\pi_1 X$ .*

### Étale Topology

Given a topological space  $(X, \tau)$ , the topology  $\tau$  (That is, the collection of open sets) forms a partially ordered set with respect to set inclusion. There also exists a notion of open covering  $U = \cup U_i$ .

**Definition 20.2.35** A Groethendieck Topology on a category  $C$  with fibre products is a collection of families of morphisms  $U_i \rightarrow U$  such that:

1. The family consisting of a single isomorphism  $\{U \sim U\}$  is a covering.
2. If  $\{U_i \rightarrow U\}$  and  $\{V_{ij} \rightarrow U_i\}$  are coverings, then so is the composition  $\{V_{ij} \rightarrow U\}$ .

**Definition 20.2.36** A site is a category with a Grothendieck Topology.

### 20.2.4 The Zariski Topology

#### The Zariski Topology

**Definition 20.2.37** A subset of  $k^n$  is closed in the Zariski Topology if it is an algebraic set. The Zariski Topology is formed by considering all such sets.

**Definition 20.2.38** A topological space  $X$  is called irreducible if for any closed subsets  $X_1, X_2 \subset X$  such that  $X = X_1 \cup X_2$ , either  $X = X_1$  or  $X = X_2$ . A topological space that is not irreducible is called reducible.

**Definition 20.2.39** A subset  $Y \subset X$  of a topological space is said to be irreducible if  $Y$  is irreducible with respect to the inherited, or the induced topology.

**Definition 20.2.40** A topological space  $X$  is said to be disconnected if there are two non-empty closed subsets  $X_1, X_2$  such that  $X = X_1 \cup X_2$ , and  $X_1 \cap X_2 = \emptyset$ .

**Theorem 20.2.30.** *If  $X$  is disconnected, then it is reducible.*

*Proof.* For if  $X$  is disconnected, there are two non-empty closed sets  $X_1, X_2 \subset X$  such that  $X_1 \cap X_2 = \emptyset$  and  $X = X_1 \cup X_2$ . But if  $X_1$  and  $X_2$  are non-empty and disjoint, then  $X_1 \neq X$  and  $X_2 \neq X$ . Therefore  $X$  is reducible.  $\square$

**Definition 20.2.41** An algebraic affine variety is an irreducible closed subset of  $k^n$ .

**Definition 20.2.42** An open subset of an affine variety is called a quasi-affine variety.

**Definition 20.2.43** If  $X \subset k^n$  is an algebraic set, then  $k[x_1, \dots, x_n]/\mathbb{I}(X)$  is called the coordinate ring of  $X$ .

**Definition 20.2.44** A set  $Y$  in a topological space  $X$  is said to be dense in  $X$  if for every non-empty open set  $\mathcal{O}$ ,  $\mathcal{O} \cap Y \neq \emptyset$ .

**Theorem 20.2.31.** *A topological space  $X$  is irreducible if and only if every non-empty open set is dense.*

**Definition 20.2.45** An irreducible component of  $X$  is a maximal irreducible subset of  $X$ .

**Theorem 20.2.32.** *If  $X$  is a closed topological space, then any irreducible subset  $Y \subset X$  is contained in a maximal component.*

**Theorem 20.2.33.** *If  $X$  is a topological space, then it is the union of irreducible components.*

**Definition 20.2.46** A topological space  $X$  is called Noetherian if every descending chain  $X_n \subset X_{n+1}$  of closed subsets stabilizes.

**Theorem 20.2.34.** *If  $X$  is a Noetherian Space, then every subset  $Y \subset X$  can be written as a finite union of irreducible closed subsets.*

**Theorem 20.2.35.** *Every algebraic set in  $k^n$  can be expressed uniquely as a union of varieties.*

**Theorem 20.2.36.** *If  $R$  is an Noetherian ring, then  $k[x_1, \dots, x_n]$  is Noetherian.*

**Theorem 20.2.37.** *A ring  $R$  is Noetherian if and only if every non-empty set of ideals in  $R$  has a maximal element.*

**Theorem 20.2.38** (Hilbert's Nullstellensatz). *If  $k$  is an algebraically closed field,  $I \subset R = k[x_1, \dots, x_n]$  is an ideal, and  $f \in R$  is a polynomial which vanishes on  $\mathbf{V}(I)$ , then there is an  $n \in \mathbb{N}$  such that  $f^n \in I$ .*

**Definition 20.2.47** The dimension of a topological space  $X$  is the supremum of all  $n \in \mathbb{N}$  such that there is a chain  $Z_0 \subset Z_1 \subset \dots \subset Z_n$  of distinct irreducible closed subsets of  $X$ .

**Theorem 20.2.39.** *If  $k$  is a field, and  $B$  is an integral domain which is finitely generated by a  $k$ -algebra, then the dimension of  $B$  is equal to the transcendence degree of the quotient field  $k(B)$  of  $B$  over  $k$ .*

**Theorem 20.2.40.** *The dimension of  $k^n$  is  $n$ .*

**Theorem 20.2.41.** *If  $Y$  is a quasi-affine variety, then  $\dim(Y) = \dim(\overline{Y})$ .*

## Problems

**Problem 20.2.1** Let  $f \in k[x]$  be a non-constant polynomial in one variable over a field  $k$ .  $f$  is called irreducible if  $f \notin k$  and if it is not the product of two polynomials of strictly smaller degree. Prove the following are equivalent:

1.  $k[x]/\langle f \rangle$  is a field.
2.  $k[x]/\langle f \rangle$  is an integral domain.
3.  $f$  is irreducible.

*Solution.* If  $k[x]/\langle f \rangle$  is a field, then it is an integral domain. If  $f$  is irreducible, then  $\langle f \rangle$  is maximal and thus  $k[x]/\langle f \rangle$  is a field. Finally, if  $k[x]/\langle f \rangle$  is an integral domain, then  $\langle f \rangle$  is prime. But if  $\langle f \rangle$  is prime, then it is maximal. And if  $\langle f \rangle$  is maximal, then  $f$  is irreducible.  $\square$

**Problem 20.2.2** Show every prime ideal is radical.

*Solution.* Let  $I$  be a prime ideal. Then if  $fg \in I$ , either  $f \in I$  or  $g \in I$ . Suppose  $f^n \in I$  for some  $f \in R$ . Then  $f^{n-1}f \in I$ . But then either  $f^{n-1} \in I$  or  $f \in I$ . If  $f \in I$ , we are done. If not, by induction  $f^{n-k} \in I$  and we obtain  $f \in I$ .  $\square$

**Problem 20.2.3** Show that any Noetherian Topological Space  $X$  is compact.

*Solution.* If  $X$  is Noetherian, then every ascending chain terminates. Suppose  $X$  is not compact. Then there is an open cover  $\Delta$  with no finite subcover. Let  $\mathcal{O}_1$  be a finite subcover. Then  $\cup_{U \in \mathcal{O}_1} U$  is not all of  $X$ , otherwise  $X$  would be compact. Thus there is an open subcover  $\mathcal{O}_2$  such that  $\mathcal{O}_1 \subset \mathcal{O}_2$ . Inductively, we have a sequence  $\mathcal{O}_n \subset \mathcal{O}_{n+1}$ . Let  $A_n = \cup_{k=1}^n \cup_{U \in \mathcal{O}_k} U$ . Then  $A_n \subset A_{n+1}$ . But by the Noetherian property, this chain must stabilize. But then there is an  $N \in \mathbb{N}$  such that  $\mathcal{O}_{N+1} = \mathcal{O}_N$ , a contradiction as we said  $X$  is not compact. Therefore, etc.  $\square$

This proof subtly requires the axiom of choice in the construction of such  $\mathcal{O}'s$ .

## 20.2.5 Notes on Varieties

### Affine Varieties

Let  $k$  denote an algebraically closed field.  $\mathbf{A}_k^n$  is the affine  $k$ -space in  $n$  dimensions. An element  $a = (a_1, \dots, a_n)$  is called a point, and  $a_i$  is called a coordinate.

**Definition 20.2.48** The zero set of a set of polynomials  $T = \{f_1, \dots, f_s\}$  is the set  $Z(T) = \{p \in \mathbf{A}_k^n \mid f_i(p) = 0, i = 1, \dots, s\}$ .

The set of polynomials in  $n$  variables over  $\mathbf{A}_k^n$  is denoted  $A$ .

**Definition 20.2.49** A subset  $Y \subset \mathbf{A}_k^n$  is an algebraic set if there exists a subset  $T \subset A$  such that  $Z(T) = Y$ .

**Theorem 20.2.42.** *The union of two algebraic sets is algebraic.*

**Theorem 20.2.43.** *The intersection of two algebraic sets is algebraic.*

**Definition 20.2.50** The Zariski topology  $\mathcal{Z}$  on  $\mathbf{A}_k^n$  is the set of complements of algebraic sets. That is, algebraic sets are closed.

**Definition 20.2.51** A non-empty subset  $Y$  of a topological space  $X$  is irreducible if it cannot be expressed as the union  $Y = Y_1 \cup Y_2$  of two proper subsets, each one of which is closed in  $Y$ .

**Definition 20.2.52** An affine algebraic variety is an irreducible subset of  $\mathbf{A}_k^n$  with respect to the induced topology.

**Definition 20.2.53** An open subset of an affine variety is called a quasi-affine variety.

If  $Y \subset \mathbf{A}_k^n$ ,  $I(Y) = \{f \in A : \forall p \in Y, f(p) = 0\}$ .

**Theorem 20.2.44.**

1. If  $T_1 \subset T_2$ , then  $Z(T_2) \subset Z(T_1)$

2. If  $Y_1 \subset Y_2 \subset \mathbf{A}_k^n$ , then  $I(Y_2) \subset I(Y_1)$
3.  $I(Y_1 \cup Y_2) = I(Y_1) \cap I(Y_2)$
4. If  $a \subset A$ , then  $I(Z(a)) = \sqrt{a}$  (The radical of  $a$ )
5. If  $Y \subset \mathbf{A}_k^n$ , then  $Z(I(Y)) = \bar{Y}$  (The closure of  $Y$ )

**Theorem 20.2.45** (Hilbert's Nullstellensatz). *If  $k$  is an algebraically closed field,  $a \subset A = k[x_1, \dots, x_n]$  is an ideal, and if  $f \in A$  is a polynomial which vanishes on  $Z(a)$ , then there is an  $r \in \mathbb{N}$  such that  $f^r \in a$ .*

**Definition 20.2.54** The affine coordinate ring of an affine algebraic set  $Y \subset \mathbf{A}_k^n$  is  $A/I(Y)$ .

**Definition 20.2.55** A topological space  $X$  is called Noetherian if it satisfies the descending chain condition for closed subsets.

**Theorem 20.2.46.** *A Noetherian Topological Space is compact.*

**Definition 20.2.56** If  $A$  is a ring, then height of a prime ideal  $p$  is the supremum of all integers  $n$  such that there is a chain  $p_0 \subset \dots \subset p_n = p$  of distinct prime ideals.

**Definition 20.2.57** The Krull dimension of a ring  $A$  is the supremum of the height of all ideals.

**Theorem 20.2.47** (Krull's Hauptidealsatz). *If  $A$  is a Noetherian Ring, and  $f \in A$  has neither a zero divisor nor a unit, then every minimal prime ideal  $p$  containing  $f$  has height 1.*

**Theorem 20.2.48.** *The dimension of  $\mathbf{A}_k^n$  is  $n$ .*

## Projective Varieties

**Definition 20.2.58** A subset  $Y$  of  $P^n$  is an algebraic set if there is a set  $T$  of homogeneous elements of  $S$  such that  $Y = Z(T)$ .

**Definition 20.2.59** The Zariski Topology on  $P^n$  is defined as the complements of algebraic sets. That is, algebraic sets are closed.

**Definition 20.2.60** A projective algebraic variety is an irreducible algebraic set in  $P^n$ .

## More Notes on Projective Varieties

**Definition 20.2.61** The projective  $n$ -space over  $\mathbb{A}$ , denoted  $\mathbb{P}^n$ , is the set of all one-dimensional linear subspaces of the vector space  $\mathbb{A}^{n+1}$ .

Equivalently, it is the set of all lines in  $\mathbb{A}^{n+1}$  through the origin.

**Definition 20.2.62** The projective  $n$  space  $\mathbb{P}^n$  over  $k$  is the set of all equivalence classes  $\mathbb{A}^{n+1}/\{0\}$ , where  $(a_1, \dots, a_n) \sim (b_1, \dots, b_n)$  if and only if there is a  $\lambda \in \mathbb{A} \setminus \{0\}$  such that  $b_i = \lambda a_i$ .

Elements of  $\mathbb{P}^n$  are called points.

**Definition 20.2.63** A homogenous polynomial of degree  $d$  is a polynomial  $f$  such that  $f(\lambda a_1, \dots, \lambda a_n) = \lambda^d f(a_1, \dots, a_n)$ .

**Theorem 20.2.49.** *If  $I \subset k[x_1, \dots, x_n]$  is an ideal, then the following are equivalent:*

1. *I can be generated by homogeneous polynomials.*
2. *For every  $f \in I$ , the degree  $d$  part of  $f$  is contained in  $I$*

**Definition 20.2.64** If  $I \subset k[x_1, \dots, x_n]$  is a homogeneous ideal, then  $\mathbf{V}(I) = \{(a_1 : \dots : a_n) \in \mathbb{P}^n : f(a_1, \dots, a_n) = 0, f \in I\}$ .

**Definition 20.2.65** An algebraic subset of  $\mathbb{P}^n$  is a set of the form  $\mathbf{V}(I)$ . These are called projective algebraic sets.

**Theorem 20.2.50.** *Every projective algebraic set can be written as the zero set of finitely many homogeneous polynomials of the same degree.*

**Definition 20.2.66** The projective close of an algebraic set  $X \subset \mathbb{A}^n$  is the Zariski closure in  $\mathbb{P}^n$  under the mapping  $\mathbb{A}^n \rightarrow \mathbb{P}^n$  by  $(x_1, \dots, x_n) \mapsto (1 : x_1, \dots, x_n)$ .

**Theorem 20.2.51.** *If  $f$  is the sum of forms  $f = \sum_d f^{(d)}$ , if  $P \in \mathbb{P}^n$  and  $f(x_1, \dots, x_n) = 0$  for every choice of homogeneous coordinates, then for each  $d$ ,  $f^{(d)}(x_1, \dots, x_n) = 0$ .*

**Definition 20.2.67** If  $F \in \mathbb{A}[x_1, \dots, x_n]$  is homogeneous of degree  $d$ , then its de-homogenization is the polynomial  $f(x_1, \dots, x_n) = F(1, x_1, \dots, x_n)$ .

**Theorem 20.2.52.** *Let  $X \subset \mathbb{A}^n$  be an affine algebraic set,  $\overline{X}$  the projective closure. Then  $\mathbb{I}(\overline{X}) \subset \mathbb{A}[x_1, \dots, x_n]$  is generated by the homogenization of all elements of  $\mathbb{I}(X)$ .*

**Theorem 20.2.53.** *An algebraic set  $X$  is irreducible if and only if the ideal  $\mathbb{I}(X)$  is prime.*

**Definition 20.2.68** An affine algebraic set  $X \subset \mathbb{A}^{n+1}$  is called a cone if it is not empty, and if for all  $\lambda \in k$ ,  $(x_1, \dots, x_n) \in X \Rightarrow (\lambda x_1, \dots, \lambda x_n) \in X$ .

**Theorem 20.2.54** (The Projective Nullstellensatz).

1. If  $X_1 \subset X_2$  are algebraic set in  $\mathbb{P}^n$ , then  $I(X_2) \subset I(X_1)$ .
2. For any algebraic set  $X \subset \mathbb{P}^n$ , we have  $\mathbf{V}(I(X)) = X$ .
3. For any homogeneous ideal  $I \subset k[x_1, \dots, x_n]$  such that  $\mathbf{V}(I) \neq \emptyset$ , we have  $\mathbb{I}(\mathbf{V}(I)) = \sqrt{I}$ .

## 20.3 Lie Algebras

For most purposes,  $\mathbb{F}$  will be either the field  $\mathbb{R}$  or  $\mathbb{C}$ , where  $+$  and  $\cdot$  are the usual notions of addition and multiplication.

### Definition 20.3.1: Lie Algebras

A Lie algebra over a field  $(\mathbb{F}, +, \cdot)$  is vector space  $(V, +, \cdot)$  over  $\mathbb{F}$  and a bilinear map  $[\cdot] : V \times V \rightarrow V$ , denoted  $(X, Y) \mapsto [X, Y]$ , satisfying:

1.  $[x, x] = 0$  [Alternating Property]
2.  $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$  [Jacobi Identity]

**Theorem 20.3.1.** If  $(\mathbb{F}, +, \cdot)$  is a field, if  $(V, [\cdot])$  is a Lie algebra over  $\mathbb{F}$ , and if  $x, y \in V$ , then:

$$[x, y] = -[y, x] \quad (20.3.1)$$

*Proof.* Applying bilinearity and the alternating property, we have:

$$\mathbf{0} = [x + y, x + y] = [x, x] + [x, y] + [y, x] + [y, y] \quad (20.3.2)$$

But from the alternating property,  $[x, x] = \mathbf{0}$  and  $[y, y] = \mathbf{0}$ , and thus:

$$\mathbf{0} = [x, y] + [y, x] \quad (20.3.3)$$

This completes the proof.  $\square$

**Theorem 20.3.2.** If  $(\mathbb{F}, +, \cdot)$  is a field, if  $(V, [\cdot])$  is a Lie algebra over  $\mathbb{F}$ , and if  $X, Y \in V$ , then:

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]] \quad (20.3.4)$$

*Proof.* Applying the Jacobi identity and Thm. 20.3.1 completes the proof.  $\square$

---

**Theorem 20.3.3: Lie Brackets Form a Derivation**


---

If  $(V, [ ])$  is a Lie algebra, if  $\mathbf{x} \in V$ , and if  $D : V \rightarrow V$  is defined by:

$$D(\mathbf{y}) = [\mathbf{x}, \mathbf{y}] \quad (20.3.5)$$

Then  $D$  is a derivation on  $V$ . ■

*Proof.* For by Thm. 20.3.2, we have:

$$D([\mathbf{y}, \mathbf{z}]) = [\mathbf{x}, [\mathbf{y}, \mathbf{z}]] = [[\mathbf{x}, \mathbf{y}], \mathbf{z}] + [\mathbf{y}, [\mathbf{x}, \mathbf{z}]] = [D(\mathbf{y}), \mathbf{z}] + [\mathbf{y}, D(\mathbf{z})] \quad (20.3.6)$$

And thus  $D$  is a derivation. □

---

**Example 20.3.1: Examples of Lie Algebras**


---

gain, the most elementary example is the cross product on  $\mathbb{R}^3$ , with scalar multiplication and vector addition having their usual definitions. The cross product is anti-commutative:

$$\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x} \quad (20.3.7)$$

And from this we obtain the alternating law. Similarly, it obeys the following identity:

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = (\mathbf{x} \times \mathbf{y}) \times \mathbf{z} + \mathbf{y} \times (\mathbf{x} \times \mathbf{z}) \quad (20.3.8)$$

And from this the Jacobi identity is recovered. If  $(V, +, \cdot)$  is a vector space and if  $[ ] : V \times V \rightarrow V$  is defined by  $[\mathbf{x}, \mathbf{y}] = \mathbf{0}$  for all  $\mathbf{x}, \mathbf{y} \in V$ , then  $(V, [ ])$  is a Lie algebra.

Given a ring  $R$ , the Heisenberg group  $H_3(R)$  is defined by considering matrices of the following form:

$$A = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \quad (20.3.9)$$

Where  $a, b, c \in R$ . If we consider this over  $\mathbb{R}$ , then  $A$  will be invertible for any such  $a, b, c \in \mathbb{R}$  since  $\det(A) = 1$ , and thus the set of all such matrices forms a group under matrix multiplication. The Lie algebra associated with the Heisenberg group is the set of matrices of the form  $A - I$ , where  $I$  is the  $3 \times 3$  identity matrix. We form as a basis the following three matrices:

$$X = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (20.3.10a)$$

$$Y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (20.3.10b)$$

$$Z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad (20.3.10c)$$

We then define the Lie bracket from this basis:

$$[X, Y] = Z \quad (20.3.11a) \quad [X, Z] = 0 \quad (20.3.11b) \quad [Y, Z] = 0 \quad (20.3.11c)$$

$(H_3(\mathbb{R}), [] )$  then forms a Lie algebra. ■

### Definition 20.3.2: Lie Algebra Homomorphism

A Lie algebra homomorphism from a Lie algebra  $(X, []_X)$  to a Lie algebra  $(Y, []_Y)$  is a linear function  $f : X \rightarrow Y$  such that, for all  $\mathbf{x}_1, \mathbf{x}_2 \in X$ , the following is true:

$$f([\mathbf{x}_1, \mathbf{x}_2]_X) = [f(\mathbf{x}_1), f(\mathbf{x}_2)]_Y \quad (20.3.12)$$

### Definition 20.3.3: Lie Algebra Isomorphism

A Lie algebra isomorphism between two Lie algebras  $G$  and  $H$  is a homomorphism  $f : G \rightarrow H$  such that  $f$  is a bijection.

### Theorem 20.3.4: Equivalent Definition of Isomorphism (Part I)

If  $(X, []_X)$  and  $(Y, []_Y)$  are Lie algebras, and if  $f : X \rightarrow Y$  is a Lie algebra isomorphism, then  $f$  is a Lie algebra homomorphism such that  $f^{-1} : Y \rightarrow X$  is a Lie algebra homomorphism. ■

*Proof.* For if  $f : X \rightarrow Y$  is a Lie algebra isomorphism, then it is a homomorphism and a bijection (Def. 20.3.3). But then the inverse function  $f^{-1} : Y \rightarrow X$  is well defined. It thus suffices to show that this a homomorphism as well. Let  $\mathbf{y}_1, \mathbf{y}_2 \in Y$ . Then:

$$f\left([f^{-1}(\mathbf{y}_1), f^{-1}(\mathbf{y}_2)]_Y\right) = \left[f(f^{-1}(\mathbf{y}_1)), f(f^{-1}(\mathbf{y}_2))\right]_Y = [\mathbf{y}_1, \mathbf{y}_2] \quad (20.3.13)$$

Since  $f$  is a bijection, we may take the inverse of this to obtain:

$$[f^{-1}(\mathbf{y}_1), f^{-1}(\mathbf{y}_2)]_X = f^{-1}\left([\mathbf{y}_1, \mathbf{y}_2]_Y\right) \quad (20.3.14)$$

Thus completing the proof. □

### Theorem 20.3.5: Equivalent Definition of Isomorphism (Part II)

If  $(X, [\cdot]_X)$  and  $(Y, [\cdot]_Y)$  are Lie algebras, and if  $f : X \rightarrow Y$  is a Lie algebra homomorphism such that  $f^{-1} : Y \rightarrow X$  exists and is a homomorphism, then  $f$  is a Lie algebra homomorphism. ■

*Proof.* For if  $f : X \rightarrow Y$  is a Lie algebra homomorphism such that  $f^{-1} : X \rightarrow Y$  exists and is a Lie algebra homomorphism, then  $f$  is bijective. Therefore,  $f$  is a Lie algebra isomorphism (Def. 20.3.3). □

#### Definition 20.3.4: Commutative Elements of a Lie Algebra

Commutative elements of a Lie algebra  $(X, [\cdot])$  are elements  $\mathbf{x}, \mathbf{y}$  such that:

$$[\mathbf{x}, \mathbf{y}] = \mathbf{0} \quad (20.3.15)$$

#### Definition 20.3.5: Abelian Lie Algebra

An Abelian Lie Algebra is a Lie algebra  $(X, [\cdot])$  such that for all  $\mathbf{x}, \mathbf{y} \in X$ ,  $\mathbf{x}$  and  $\mathbf{y}$  are commutative elements.

It will be seen that these come from Abelian Lie groups.

#### Example 20.3.2

Let  $(A, \cdot)$  be an associative  $\mathbb{F}$  algebra. We can define a Lie algebra  $\text{Lie}(A)$  as a vector space  $A$  and by setting the Lie bracket to be the commutator:

$$[\mathbf{x}, \mathbf{y}] = \mathbf{x} \cdot \mathbf{y} - \mathbf{y} \cdot \mathbf{x} \quad (20.3.16)$$

Where  $\cdot$  is the multiplicative operation that comes from the associative  $\mathbb{F}$  algebra  $(A, \cdot)$ . The commutator is anticommutative:

$$[\mathbf{x}, \mathbf{y}] = \mathbf{x} \cdot \mathbf{y} - \mathbf{y} \cdot \mathbf{x} = -(\mathbf{y} \cdot \mathbf{x} - \mathbf{x} \cdot \mathbf{y}) = -[\mathbf{y}, \mathbf{x}] \quad (20.3.17)$$

And thus the alternating property is satisfied. Similarly, the Jacobi identity is valid and thus  $\text{Lie}(A)$  is a Lie algebra. A special case of this is when we have  $A = M_n(\mathbb{F})$ . Then  $\text{Lie}(M_n(\mathbb{F}))$  is denoted by  $\text{GL}_n(\mathbb{F})$ . ■

We will use the following notation.

**Notation 20.3.1: Lie Bracket of Vector Subspaces**

Given a Lie Algebra  $(V, [] )$  and two vector subspaces  $W_1, W_2 \subseteq V$ , we write  $[W_1, W_2]$  to denote the following:

$$[W_1, W_2] = \text{Span}\{ [\mathbf{w}_1, \mathbf{w}_2] : \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2 \} \quad (20.3.18)$$

With this we can define Lie Subalgebras.

**Definition 20.3.6: Lie Subalgebra**

A Lie subalgebra of a Lie algebra  $(V, [] )$  is a subset  $W \subseteq V$  such that:

$$[W, W] \subseteq W \quad (20.3.19)$$

That is to say, a Lie subalgebra of a Lie algebra is a subspace that is closed under the Lie bracket.

**Definition 20.3.7: Ideal of a Subspace**

An ideal of a Lie algebra  $(V, [] )$  is a subset  $W \subseteq V$  such that:

$$[V, W] \subseteq W \quad (20.3.20)$$

By the anticommutativity of the Lie bracket, there is no distinction between left ideals and right ideals.

**Theorem 20.3.6.** *If  $(X, []_X)$  and  $(Y, []_Y)$  are Lie algebras, and if  $f : X \rightarrow Y$  is a Lie algebra homomorphism, then  $\ker(f)$  is an ideal of  $X$ .*

*Proof.* For if  $\mathbf{x} \in \ker(f)$  and  $\mathbf{y} \in X$ , then:

$$f([\mathbf{x}, \mathbf{y}]) = [f(\mathbf{x}), f(\mathbf{y})] = [\mathbf{0}, \mathbf{y}] = \mathbf{0} \quad (20.3.21)$$

And therefore  $[\mathbf{x}, \mathbf{y}] \in \ker(f)$ . Thus,  $\ker(f)$  is an ideal of  $X$ .  $\square$

**Theorem 20.3.7.** *If  $(V, [] )$  is a Lie algebra, and if  $W \subseteq V$  is an ideal of  $V$ , then  $G/H$  has a Lie algebra structure such that the projection mapping  $\pi : V \rightarrow V/W$ , defined by  $\mathbf{x} \mapsto \mathbf{x} + W$ , is a Lie algebra homomorphism.*

*Proof.* For define  $[\mathbf{x} + H, \mathbf{y} + H]$  as follows:

$$[\mathbf{x} + H, \mathbf{y} + H] = [\mathbf{x}, \mathbf{y}] + H \quad (20.3.22)$$

Then:

$$[\mathbf{x} + H, \mathbf{y} + H] = [\pi(\mathbf{x}), \pi(\mathbf{y})] = [\mathbf{x}, \mathbf{y}] + H = \pi([\mathbf{x}, \mathbf{y}]) \quad (20.3.23)$$

And therefore:

$$[\pi(\mathbf{x}), \pi(\mathbf{y})] = \pi([\mathbf{x}, \mathbf{y}]) \quad (20.3.24)$$

Let  $\overline{X} = X + H$ . If  $\overline{X}' = \overline{X}$  and  $\overline{Y}' = \overline{Y}$ , then  $\overline{[X, Y]} = \overline{[X', Y']}$ . Thus:

$$X - X' = H_1 \in H \quad Y - Y' = H_2 \in H \quad (20.3.25)$$

And:

$$[X, Y] = [X' + H_1, Y' + H_2] = [X', Y'] + [H_1, Y'] + [X', H_2] + [H_1, H_2] \quad (20.3.26)$$

And this last sum of three is in  $H$ .  $\square$

**Theorem 20.3.8.** *If  $(X, []_X)$  is a Lie algebra, and if  $W \subseteq X$  is an ideal of  $X$ , then there is a Lie algebra  $(Y, []_Y)$  and a homomorphism  $f : X \rightarrow Y$  such that  $W = \ker(f)$ .*

*Proof.* For let  $Y = X/W$  and let  $[]_Y$  be the Lie bracket:

$$[\mathbf{x} + W, \mathbf{y} + W]_Y = [\mathbf{x}, \mathbf{y}] + W \quad (20.3.27)$$

Let  $\pi : X \rightarrow X/W$  be the projection mapping  $\pi(x) = x + W$ . From the previous theorem,  $(X/W, []_Y)$  is a Lie algebra and  $\pi$  is a Lie algebra homomorphism. Moreover,  $\ker(\pi) = W$ . Therefore, etc.  $\square$

**Theorem 20.3.9.** *Let  $f : G \rightarrow H$  be a map of Lie algebras, and let  $K = \ker(f)$ . Then there is a unique map  $\bar{f} : G/K \rightarrow H$  such that  $\bar{f}$  is injective and such that some commutative diagram thing.*

**Example 20.3.3**    1.  $0, G(\text{TRIANGLE}))G$

2.  $Z(G) = \{Z \in G : [X, Z] = 0\}$  is called the center of  $G$ .
3.  $[G, G] = \text{Span}\{[X, Y] : X, Y \in G\}(G(\text{TRIANGLE}))G$
4. Fuck it.
5. For ideals  $a, b(G(\text{TRIANGLE}))G, a + b(G(\text{TRIANGLE}))G$ .
6. Similarly for subtraction.

**Example 20.3.4**    1.  $V$  finite dimensional vector space over  $\mathbb{F}$ ,  $A = \text{End}_{\mathbb{F}}(V)$  an associative  $\mathbb{F}$  algebra, then  $GL(V) = \text{Lie}(\text{End}_{\mathbb{F}}(V))$  with  $[X, Y] = XY - YX$ .

2.  $Tr : gl(V) \rightarrow \mathbb{F}$  abelian is a Lie algebra homomorphism.  $Tr([X, Y]) = Tr(X, Y - YX) = 0 = [Tr(X), Tr(Y)]$ .  $\ker(Tr) = SL(V) = \{X \in GL(V) : Tr(X) = 0\}$

If  $V = \mathbb{F}^n$ , denote  $GL(V)$  by  $GL_n(\mathbb{F})$  and  $SL(V)$  by  $SL_n(\mathbb{F})$ . Special cases:  $SL_2(\mathbb{F})$ . This has basis  $(E^{ij})_{k\ell} = \delta_k^i \delta_\ell^j$ .

**Example 20.3.5** Compute  $[X, Y] = [E^{12}, E^{21}] = E^{11} - E^{22} = H$ . And  $[H, X] = 2X$ ,  $[H, Y] = -2Y$ .

### Definition 20.3.8: Simple Lie Algebra

A Lie algebra is simple if it is non-abelian and its only ideals are 0 and itself.



**Theorem 20.3.10.**  $SL_2(\mathbb{F})$  is simple for  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{C}$ .

*Proof.* Bracket notation says that  $X$  is a 2-eigenvector of  $[H, \cdot]$  and  $Y$  is a -2-eigenvector of  $[H, \cdot]$ . Let  $A$  be a non-trivial ideal. We must show that  $A = SP_2(\mathbb{F})$ . Then  $[X, U] \in G$ . But:

$$[X, [X, U]] = [X, [X, xX + yY + hH]] = [X, gH - 2hX] = -2yX \quad (20.3.28)$$

So thus, either  $X \in A$  or  $Y = 0$ . Doing the same for  $[Y, [Y, U]]$  shows that  $-2xY = [Y, [Y, U]]$  and thus either  $Y \in A$  or  $X = 0$ . There are two cases now. If If  $x = y = 0$  then  $h \neq 0$  since  $U$  is non-zero. This would imply  $H \in A$ . But  $2X = [X, H] \in A$  so  $X \in A$ . Also,  $Y \in A$ . Similarly if  $y \neq 0$ .  $\square$

This can be generalized for  $SL_n(\mathbb{F})$  as well.

### Definition 20.3.9: Normalizer of a Lie Algebra Subspace

The Normalizer of  $H \subset G$  of a Lie Algebra  $G$  is the set:

$$N_G(H) = \{X \in G : [X, h] \in H\} \quad (20.3.29)$$



By the Jacobian identity, the normalizer of a subspace is also a subalgebra. Moreover, by def,  $H$  is an ideal of its normalizer.

### Definition 20.3.10: Centralizer of a Lie Algebra Subspace

The centralizer of  $H \subset G$  is the set:

$$C_G(H) = \{X \in G : [X, H] = 0\} \quad (20.3.30)$$



---

**Definition 20.3.11: Product Lie Algebras**


---

For Lie algebras  $G_1, G_2$ , their product  $G_1 \times G_2$ , also written  $G_1 \oplus G_2$ , is the set  $G_1 \oplus G_2$  as a vector space with the bracket:

$$[(X_1, X_2), (Y_1, Y_2)] = ([X_1, Y_1], [X_2, Y_2]) \quad (20.3.31)$$



$G_1 \oplus G_2$  has ideals  $\bar{G}_1 = G_1 \oplus 0$  and  $\bar{G}_2 = 0 \oplus G_2$ , and moreover  $\bar{G}_1 + \bar{G}_2 = G_1 \oplus G_2$

**Theorem 20.3.11.** *If  $G$  is a lie algebra,  $a, b$  ideals of  $G$  satisfying:*

1.  $a + b = G$
2.  $a \cap b = \emptyset$
3.  $[a, b] = 0$

*Then  $G$  is isomorphic to  $a \oplus b$ .*

*Proof.* Define  $\varphi : a \oplus b \rightarrow G$  by  $\varphi(A, B) = A + B$ . By vector space theory, this is a vector space isomorphism. We need to check that it preserves brackets. But:

$$\varphi([(A, B), (A', B')]) = \varphi([A, A'], [B, B']) = [A, A'] + [B, B'] \quad (20.3.32)$$

But also:

$$[\varphi(A, B), \varphi(A', B')] = [A + B, A' + B'] = [A, A'] + [B, A'] + [A, B'] + [B, B'] \quad (20.3.33)$$

But  $[B, A'] = [A, B'] = 0$ , completing the proof.  $\square$

## 20.4 Review of Differentiation

If  $\mathcal{U} \subseteq \mathbb{R}^n$  is open, and if  $f : \mathcal{U} \rightarrow \mathbb{R}^n$  is a function, then  $Df(p)$  (if it exists) is a linear map  $Df(p) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  such that:

$$\lim_{h \rightarrow 0} \frac{f(p + h) - (f(p) + Df(p)(h))}{\|h\|} = 0 \quad (20.4.1)$$

That is,  $Df(p)$  is the best linear affine approximation to  $f$  at  $p$ .

**Theorem 20.4.1.** *If  $f \in C^1$ , then  $Df(p)$  exists and:*

$$Df(p)(v) = \lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t} \quad (20.4.2)$$

Chain rule, if  $p \in \mathcal{U}$ ,  $f : \mathcal{U} \rightarrow \mathbb{R}$ , then  $D(g \circ f)(o) = Dg(f(p))Df(p)$

**Theorem 20.4.2.** If  $V_1, V_2, W$  are  $\mathbb{R}$  vector spaces, and if  $B : V_1 \times V_2 \rightarrow W$  is bilinear, then for  $p_1, v_1 \in V_1$ ,  $p_2, v_2 \in V_2$ :

$$DB(p_1, p_2)(v_1, v_2) = B(p_1, v_2) + B(v_1, p_2) \quad (20.4.3)$$

This can be generalized to multilinear maps.

**Theorem 20.4.3.** if  $B : V_1 \times \cdots \times V_n \rightarrow W$  is multilinear, then:

$$DB(p_1, \dots, p_n)(v_1, \dots, v_n) = \sum_{j=1}^n B(p_1, \dots, p_{j-1}, v_j, p_{j+1}, \dots, p_n) \quad (20.4.4)$$

**Example 20.4.1** Let  $\det : M_n(\mathbb{R}) \rightarrow \mathbb{R}$  be the determinant function. Then:

$$D(\det)(x)(H) = \text{Tr}(X^{Thing} H) \quad (20.4.5)$$

where  $X^{Thing}$  is the classical adjugate matrix:

$$(X^{Thing})_{ij} = (-1)^{i+j} \det(M_{ij}(X)) \quad (20.4.6)$$

Where  $M_{ij}(X)$  is the minor of  $X$  formed by crossing out the  $i^{th}$  row and  $j^{th}$  column.

**Example 20.4.2** Let  $F_k : M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R})$  be defined by  $F_k(X) = X^k$ . Then:

$$DF_k(X)(H) = X^{k-1}H + x^{k-2}HX + \cdots + XHX^{k-1} + HX^{k-1} \quad (20.4.7)$$

If  $X$  and  $H$  commute, then:

$$DF_k(X)(H) = kX^{k-1}H \quad (20.4.8)$$

**Theorem 20.4.4.** If  $V : GL_n(\mathbb{R}) \rightarrow M_n(\mathbb{R})$  is defined by  $V(X) = X^{-1}$ , then:

$$DV(X)(H) = -X^{-1}HX^{-1} \quad (20.4.9)$$

## 20.5 Lie Groups

Let  $\mathbb{F}$  denote either  $\mathbb{R}$  or  $\mathbb{C}$ .

### Definition 20.5.1: Real Lie Groups

---

A real Lie group is a  $C^\infty$  manifold  $G$  with a group structure such that the operation is smooth. That is,  $m : G \times G \rightarrow G$  which maps  $(x, y) \mapsto x * y$  is

smooth, and  $v : G \rightarrow G$  which maps  $x \mapsto x^{-1}$  is also smooth. ■

### Definition 20.5.2: Lie Group Homomorphism

---

If  $G$  and  $H$  are Lie groups, a Lie group homomorphism is a smooth group homomorphism  $f : G \rightarrow H$ . ■

**Example 20.5.1** A Complex Lie group is a complex manifold with a group structure whose operations are holomorphic. Let  $\mathbb{R}$  or  $\mathbb{F}$  be a finite dimensional  $\mathbb{F}$  vector space. Then  $G : (V) \subseteq \text{End}(V)$  is an open subset of the vector space  $\text{End}_{\mathbb{F}}(V)$ . So this is a Lie group (real if  $\mathbb{F} = \mathbb{R}$ , complex if  $\mathbb{F} = \mathbb{C}$ ) and  $GL_n(\mathbb{F}) = GL(\mathbb{F}^n)$ . As another example,  $SL_n(\mathbb{F}) = \{g \in GL_n(\mathbb{F}) : \det(g) = 1\}$ . We now show how to show that this is a Lie group. We have that  $SL_n(\mathbb{F}) = \det^{-1}(1)$ . By the implicit function theorem,  $\det^{-1}(1)$  is a smooth manifold provided that  $D(\det)(X) : M_n(\mathbb{F}) \rightarrow \mathbb{F}$  is a surjective map for any  $X \in \det^{-1}(1)$ . But  $D(\det)(X)H = \text{Tr}(X^{Thing}H)$ , so:

$$D(\det)(X)(X) = \text{Tr}(X^{Thing}X) = \text{Tr}(\det(X)I) = n \neq 0 \quad (20.5.1)$$

**Example 20.5.2** Let  $T_n(\mathbb{F})$  be the subgroup of  $GL_n(\mathbb{F})$  stabilizing the standard flag:

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = \mathbb{F}^n \quad (20.5.2)$$

This is often called the Borel subgroup. Let  $V_j$  be the span of  $\{e_1, \dots, e_j\}$ . There is a special case when  $\mathbb{F} = \mathbb{R}$  and  $n = 3$ . This is called the Heisenberg group, and is the set of all matrices of the following:

$$A = \begin{bmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{bmatrix} \quad (20.5.3)$$

Topologically this is homeomorphic to  $\mathbb{R}^3$ .

Let  $\langle \cdot, \cdot \rangle$  be a bilinear form on  $\mathbb{F}$ . Pick a basis  $B$  of  $V$ . Define  $G$  by  $G_{ij} = \langle v_i, v_j \rangle$ .

## 20.6 Solvability and Semisimplicity

### Theorem 20.6.1: Nondegenerate Splitting Lemma

---

If  $V$  is a finite dimensional  $\mathbb{F}$  vector space with a symmetric bilinear form  $\langle \cdot, \cdot \rangle$ , if  $W \subseteq V$  is a non-degenerate subspace, then  $V$  is isomorphic to  $W \oplus W^\perp$ . ■  
If  $V$  itself is nondegenerate, then  $W^\perp$  is nondegenerate.

---

**Theorem 20.6.2: Structure of Semisimple Lie algebras**


---

If  $\mathfrak{g}$  is a semisimple Lie algebra over either  $\mathbb{R}$  or  $\mathbb{C}$ , then there exists finitely many ideals  $\mathfrak{s}_k \subseteq \mathfrak{g}$  such that:

$$\mathfrak{g} = \bigoplus_{k=1}^N \mathfrak{s}_k \quad (20.6.1)$$

And if  $\mathfrak{h}$  is simple then it is one of the  $\mathfrak{s}_k$  above. ■

*Proof.* If  $\mathfrak{g}$  is simple, then we are done. If not then there is some non-zero proper ideal  $\mathfrak{h} \subsetneq \mathfrak{g}$ . But then  $\mathfrak{h}^\perp$  is an ideal of  $\mathfrak{g}$ . Let  $\mathfrak{a} = \mathfrak{h} \cap \mathfrak{h}^\perp$ . Then the killing form on  $\mathfrak{a}$  is:

$$B_{\mathfrak{a}} = B_{\mathfrak{g}}|_{\mathfrak{a} \times \mathfrak{a}} \quad (20.6.2)$$

But for  $X, Y \in \mathfrak{a}$ , we have  $B_{\mathfrak{g}}(X, Y) = 0$  since  $X \in \mathfrak{h}$  and  $Y \in \mathfrak{h}^\perp$ , and thus  $B_{\mathfrak{a}} = 0$ . Then by Cartan's Solvability criterion,  $\mathfrak{h}$  is a nondegenerate subspace of  $\mathfrak{g}$ . By the nondegenerate splitting lemma this means that  $\mathfrak{g}$  is isomorphic to  $\mathfrak{h} \oplus \mathfrak{h}^\perp$ . By Cartan's criterion for semisimplicity,  $B_{\mathfrak{g}}$  is nondegenerate and thus  $\mathfrak{h}^\perp$  is nondegenerate. But since  $\mathfrak{h}$  and  $\mathfrak{h}^\perp$  are ideals of  $\mathfrak{g}$  we have:

$$B_{\mathfrak{h}} = B_{\mathfrak{g}}|_{\mathfrak{h} \times \mathfrak{h}} \quad B_{\mathfrak{h}^\perp} = B_{\mathfrak{g}}|_{\mathfrak{h}^\perp \times \mathfrak{h}^\perp} \quad (20.6.3)$$

By Cartan's semisimplicity criterion,  $\mathfrak{h}$  and  $\mathfrak{h}^\perp$  are semisimple. But  $\mathfrak{h}$  and  $\mathfrak{h}^\perp$  are proper ideals of  $\mathfrak{g}$  and thus  $\dim(\mathfrak{h}) < \dim(\mathfrak{g})$ , and similarly for  $\mathfrak{h}^\perp$ . So by induction we have that  $\mathfrak{h}$  and  $\mathfrak{h}^\perp$  are the products of simple ideals, so  $\mathfrak{g}$  is the product of simple ideals. Moreover, any simple ideal is one of the  $\mathfrak{s}_k$ . For let  $\mathfrak{h}$  be a non-zero simple ideal of  $\mathfrak{g}$ . Then  $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$  is an ideal of  $\mathfrak{h}$ . But this bracket is either zero or all of  $\mathfrak{h}$ . But  $[\mathfrak{g}, \mathfrak{h}] = 0$  means that  $\mathfrak{h}$  is in the center of  $\mathfrak{g}$ ,  $Z(\mathfrak{g})$ , but since  $\mathfrak{h}$  is semisimple we have that  $[\mathfrak{g}, \mathfrak{h}] = \mathfrak{h}$ , a contradiction as  $\mathfrak{h}$  is nonzero. As such  $[\mathfrak{g}, \mathfrak{h}] = \mathfrak{h}$ . But then:

$$[\mathfrak{s}_k, \mathfrak{h}] \subseteq \left[ \bigoplus_{k=1}^N \mathfrak{s}_k, \mathfrak{h} \right] = \mathfrak{h} = [\mathfrak{g}, \mathfrak{h}] \quad (20.6.4)$$

If  $[\mathfrak{s}_k, \mathfrak{h}] = 0$  for all  $k$  then  $[\mathfrak{g}, \mathfrak{h}] = 0$ , a contradiction, and thus there is a  $k$  such that  $[\mathfrak{s}_k, \mathfrak{h}] \neq 0$ . But then we obtain:

$$0 \neq [\mathfrak{s}_k, \mathfrak{h}] \subseteq \mathfrak{s}_k \quad (20.6.5)$$

$$0 \neq [\mathfrak{s}_k, \mathfrak{h}] \subseteq \mathfrak{h} \quad (20.6.6)$$

From simplicity we obtain equality, and thus  $\mathfrak{h} = \mathfrak{s}_k$ . □

**Theorem 20.6.3.** *If  $\mathfrak{h}$  is a semisimple ideal in a Lie algebra  $\mathfrak{h}$ , then there is a complementary ideal  $\mathfrak{a}$  of  $\mathfrak{g}$  such that  $\mathfrak{h} = \mathfrak{h} \oplus \mathfrak{a}$ .*

*Proof.* By Cartan's semisimplicity criterion, since  $\mathfrak{h}$  is semisimple the killing form is nondegenerate. But then  $\mathfrak{h}$  is a nondegenerate subspace of  $\mathfrak{g}$ . By the nondegenerate splitting lemma,  $\mathfrak{g}$  is isomorphic as a vector space to  $\mathfrak{h} \oplus \mathfrak{h}^\perp$ .  $\square$

## 20.7 Semidirect Products

Let  $\mathfrak{g}$  be a Lie algebra with an ideal  $\mathfrak{k}$  and a subalgebra  $\mathfrak{h}$  such that  $\mathfrak{k} + \mathfrak{h} = \mathfrak{g}$  and  $\mathfrak{k} \cap \mathfrak{h} = 0$ . That is,  $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{h}$  as vector spaces. Then:

$$[K + H, K' + H'] = [K, K'] + [H, K'] + [K, H'] + [H, H'] \quad (20.7.1a)$$

$$= ([K, K'] + (\text{ad}H)(K') - (\text{ad}H')(K)) + [H, H'] \quad (20.7.1b)$$

The left term is in  $\mathfrak{k}$  and the right term is in  $\mathfrak{g}$ . Abstractly, given two Lie algebras  $\mathfrak{k}$  and  $\mathfrak{h}$  and a Lie algebra homomorphism  $\delta : \mathfrak{h} \rightarrow \text{Der}(\mathfrak{k})$ , define the Lie algebra on  $\mathfrak{k} \oplus \mathfrak{h}$  by the bracket:

$$[(K, H), (K', H')] = ([K, K] + \delta(H)(K') - \delta(H')(K), [H, H']) \quad (20.7.2)$$

This is a Lie algebra as the bracket will satisfy the Jacobi identity. Let  $\bar{\mathfrak{k}} = \mathfrak{k} \oplus 0$ , which is isomorphic to  $\mathfrak{k}$ , and let  $\bar{\mathfrak{h}} = 0 \oplus \mathfrak{h}$  be subalgebras such that  $\bar{\mathfrak{h}} + \bar{\mathfrak{k}} = \mathfrak{g}$  and  $\bar{\mathfrak{h}} \cap \bar{\mathfrak{k}} = 0$ . Recalling from the definition of a derivation, for any  $X \in \mathfrak{g}$ ,  $\text{ad}X$  is a derivation. A derivation of  $\mathfrak{g}$  that is of the form  $\text{ad}X$  for some  $X \in \mathfrak{g}$  is called an inner derivation. We denote the set of derivations on  $\mathfrak{g}$  by:

$$\text{Der}(\mathfrak{k}) = \{D \in \text{End}_{\mathbb{F}}(\mathfrak{g}) : D \text{ is a derivation}\} \quad (20.7.3)$$

Note that for  $D, D' \in \text{Der}(\mathfrak{g})$ , we have that:

$$[D, D'] = D \circ D' - D' \circ D \quad (20.7.4)$$

is also a derivation. Let  $D \in \text{Der}(\mathfrak{g})$ . Then for all  $X \in \mathfrak{g}$  we have  $[D, \text{ad}X] = \text{ad}(D(X))$ . That is  $\text{ad}(\mathfrak{g})$  is the set of inner derivations on  $\mathfrak{g}$ . It is also an ideal of  $\text{Der}(\mathfrak{g})$ .

**Theorem 20.7.1.** *If  $\mathfrak{g}$  is a semisimple Lie algebra, then all derivations of  $\mathfrak{g}$  are inner. That is,  $\text{Der}(\mathfrak{g}) = \text{ad}(\mathfrak{g})$ .*

*Proof.* For let  $D \in \text{Der}(\mathfrak{g})$ . Form  $\tilde{\mathfrak{g}} = \mathfrak{g} \times_{\delta} \mathbb{F}$ , where  $\delta : \mathbb{F} \rightarrow \text{Der}(\mathfrak{g})$ . Since  $\mathfrak{g}$  is a semisimple ideal in  $\tilde{\mathfrak{g}}$  there is a complementary ideal  $\mathfrak{a}$  such that  $\tilde{\mathfrak{g}} = \mathfrak{g} \oplus \mathfrak{a}$ . Thus  $\mathfrak{a} \cap \mathfrak{g} = 0$  and

$$\text{dim}(\mathfrak{a}) = \dim(\tilde{\mathfrak{g}}) - \dim(\mathfrak{g}) = 1 \quad (20.7.5)$$

Some more stuff.  $\square$

**Definition 20.7.1: Perfect Lie Algebra**

A perfect Lie algebra is a Lie algebra  $\mathfrak{g}$  such that  $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$ .

**Theorem 20.7.2.** *A semisimple Lie algebra is perfect.*

*Proof.* For if  $\mathfrak{g}$  is simple then it is perfect since otherwise  $[\mathfrak{g}, \mathfrak{g}] = 0$ . If not then  $\mathfrak{g}$  is the direct sum of simple ideals by the structure theorem. But then:

$$[\mathfrak{g}, \mathfrak{g}] = \left[ \bigoplus_{k=1}^n \mathfrak{s}_k, \bigoplus_{j=1}^n \mathfrak{s}_j \right] = \bigoplus_{k=1}^n \bigoplus_{j=1}^n [\mathfrak{s}_k, \mathfrak{s}_j] = \bigoplus_{i=1}^n [\mathfrak{s}_i, \mathfrak{s}_i] = \mathfrak{g} \quad (20.7.6)$$

□

## 20.8 Abstract Jordan-Chevelay Decomposition

Let  $\mathfrak{a}$  be a finite dimensional algebra. It need not be associative. Let  $D$  be a derivation of  $\mathfrak{a}$  and let  $D_s$  and  $D_n$  be the Jordan-Chevelay decomposition of  $D$ . That is,  $D = D_s + D_n$ . Then  $D_s$  and  $D_n$  are also derivations. Let  $\mathfrak{g}$  be a semisimple Lie algebra.

Let  $\varphi : \mathfrak{g} \rightarrow gl(V)$  be a representation of a Lie algebra  $\mathfrak{G}$  and let  $A_\varphi$  be the associative algebra defined by:

$$A_\varphi = \mathbb{F}[\varphi(\mathfrak{g})] \quad (20.8.1)$$

Then  $A_\varphi \subseteq \text{End}_{\mathbb{F}}(V)$ , and this is moreover generated by  $\varphi(\mathfrak{g})$ . Then the subrepresentations of  $V$ , that is the subspaces  $W \subseteq V$  that are invariant under all  $\varphi(x)$  in  $X(\mathfrak{g})$ , are just the  $A_\varphi$  submodules of  $V$ .

**Theorem 20.8.1: Schur's Lemma**

If  $A$  is a finite dimensional  $\mathbb{F}$  algebra and if  $V$  is an irreducible  $A$  module, then any non-zero  $A$  module endomorphism  $f : V \rightarrow V$  is an automorphism. ■

*Proof.* Since  $V$  is irreducible and  $f$  is non-zero, the kernel of  $f$  is zero and the image is  $V$ , and therefore  $f$  is bijective, and therefore  $\text{End}_A(V)$  is a division algebra. □

**Theorem 20.8.2.** *If  $\mathbb{F}$  is algebraically closed then there is a  $\lambda \in \mathbb{F}$  such that  $f = \lambda \cdot \text{Id}$ .*

*Proof.* If  $\mathbb{F}$  is algebraically closed then  $f$  has an eigenvalue  $\lambda \in \mathbb{F}$ . Then  $\ker(\lambda \cdot \text{Id} - f) \subseteq V$  is a non-zero submodule so is all of  $V$ . That is,  $f = \lambda \cdot \text{Id}$ . □

### Theorem 20.8.3: Weyl's Theorem

---

Let  $\mathfrak{g}$  be a semi-simple Lie algebra over  $\mathbb{C}$  and let  $\varphi : \mathfrak{g} \rightarrow gl(V)$  be a finite dimensional representation of  $\mathfrak{g}$ . Then  $V$  is completely reducible. █

**Theorem 20.8.4.** *If  $V$  is a vector space over  $\mathbb{C}$  and if  $\mathfrak{g}$  is a semisimple subalgebra of  $gl_{\mathbb{C}}(V)$ , then for all  $x \in \mathfrak{g}$ ,  $X_s, X_n \in \mathfrak{g}$ .*

**Theorem 20.8.5.** *If  $\mathfrak{g}$  is a semisimple Lie algebra over  $\mathbb{C}$ , if  $\varphi : \mathfrak{g} \rightarrow gl_{\mathbb{C}}(V)$  is a finite dimensional representation of  $\mathfrak{g}$ , if  $X = X_s + X_n$  is the abstract Jordan-Chevelay decomposition of  $X \in \mathfrak{g}$ , then:*

$$\varphi(X) = \varphi(X_n)\varphi(X_s) \quad (20.8.2)$$

*Is the Jordan-Chevelay decomposition of  $\varphi(X) \in End_{\mathbb{C}}(V)$ . That is,  $\varphi(X_s) = \varphi(X)_s$  and  $\varphi(X_n) = \varphi(X)_n$ .*

*Proof.* Apply a previous theorem to the semisimple subalgebra  $\overline{\mathfrak{G}} = \varphi(\mathfrak{g}) \subseteq g(V)$ . □

## 20.9 Representations of $sl_2(\mathbb{C})$

$sl_2(\mathbb{C})$  is the space of all  $X \in gl_2(\mathbb{C})$  such that  $\text{Tr}(A) = 0$ . This has the following basis:

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad (20.9.1)$$

The brackets are:

$$[H, X] = 2X \quad (20.9.2)$$

Let  $\varphi : sl_2(\mathbb{C}) \rightarrow gl(V)$  be a finite dimensional representation of  $sl_2(\mathbb{C})$  on a  $\mathbb{C}$  vector space  $V$ . Consider the eigenspaces of  $\varphi(H)$ . Recall that  $\varphi$  preserves the Jordan-Chevelay decomposition. Let  $V_{\lambda}$  be defined by:

$$V_{\lambda} = \{v \in V \mid Hv = \lambda v\} \quad (20.9.3)$$

If  $\lambda \in \text{spec}(\varphi(H))$ , and 0 otherwise. If  $V_{\lambda} \neq 0$  then  $\lambda$  is said to be a weight of  $H$  in  $V$  and  $V_{\lambda}$  is the weight space. Consider  $0 \neq V_0 = V$  and let  $V_1 = YV$ ,  $V_2 = Y^2V$ , and so on. Let  $V_{-1} = 0$  and let  $\mathcal{V} = \{V_k\}$ . Note that all weights are equivalent mod 2.

**Theorem 20.9.1.** *If  $k \geq 0$ ,  $HV_k = (\lambda - 2k)V_k$ .  $YV_k = V_{k+1}$ .  $XV_k = k(\lambda - k + 1)V_{k-1}$ .*

**Theorem 20.9.2.** *If  $V$  is a finite dimensional irreducible representation of  $sl_2(\mathbb{C})$  then:*

$$V = \bigoplus_{k=-m}^m V_k \quad (20.9.4)$$

*That is,  $V$  is the direct sum of one dimensional weight spaces.*

**Theorem 20.9.3.** *If  $V$  is a finite dimensional irreducible representation of  $sl_2(\mathbb{C})$ , then there is a vector  $v \in V$  with weight  $m$  such that  $m$  is the highest weight.*

**Theorem 20.9.4.** *If  $V$  is a finite dimensional irreducible representation of  $sl_2(\mathbb{C})$  then  $V$  has a basis  $\mathcal{V} = \{v_k\}$  such that  $v_k \in V_{m-2k}$  and such that the  $sl_2(\mathbb{C})$  action is something.*

**Theorem 20.9.5.** *If  $\varphi : sl_2(\mathbb{C}) \rightarrow gl(V)$  is a finite dimensional complex representation of  $sl_2(\mathbb{C})$ , then the eigenvalues of  $\varphi(H)$  are all integers. If  $\lambda$  is an eigenvalue of  $\varphi(H)$ , then  $-\lambda$  is an eigenvalue and has the same multiplicity. Lastly, if:*

$$V = \bigoplus_{k=1}^r W_k \quad (20.9.5)$$

*where  $W_k$  is an irreducible representation, then  $r = \dim(E_0(V)) + \dim(E_1(V))$ , where  $E_\lambda(V)$  is the  $\lambda$  eigenspace of  $\varphi(H)$ .*

*Proof.* By Weyl's theorem,  $V$  is completely reducible and each summand is described by some previous theorem, so the first two parts are done. If  $V$  is the direct sum over  $W_k$ , then each  $W_k$  is isomorphic to  $V(m)$  for some  $m$ .  $\square$

Let  $m = 1$ , then  $V(1) = V_{-1} \oplus V_1$ . This is the defining representation  $\mathbb{C}^2$  of  $sl_2(\mathbb{C})$ . If we let  $W = \mathbb{C}^2$  be the defining representation of  $sl_2(\mathbb{C})$  and consider  $\text{Sym}^2(W)$  where  $W$  is a  $k$  vector space. Something about tensor product. If  $\{x, y\}$  is the standard basis of  $W$  then a basis of  $\text{Sym}^2(W)$  is  $\{x^2, xy, y^2\}$ . We have:

$$\text{Sym}^2(W) = \mathbb{C}x^2 \oplus \mathbb{C}xy \oplus \mathbb{C}y^2 = W_{-2} \oplus W_0 \oplus W_2 = V(2) \quad (20.9.6)$$

The adjoint representation of  $sl_2(\mathbb{C})$  is irreducible, and so the adjoint representation is  $V(2)$ .

## 20.10 Root Space Decomposition

Let  $\mathfrak{g}$  be a nonzero semisimple Lie algebra. The idea is to immitate the case of  $sl_2(\mathbb{C})$  for the adjoint representation of  $\mathfrak{g}$  with a family  $\mathfrak{h}$  of semisimple elements instead of just  $H \in sl_2(\mathbb{C})$ .

### Definition 20.10.1: Toral Subalgebra

A Toral subalgebra is a subalgebra  $\mathfrak{h}$  of a Lie algebra  $\mathfrak{g}$  such that for every element  $H \in \mathfrak{h}$  we have that  $H$  is semisimple,  $H = H_s$ .

**Theorem 20.10.1.** *If  $H \in \mathfrak{g}$  is semisimple then  $\mathbb{C}H$  is a toral subalgebra of  $\mathfrak{g}$ . If  $\mathfrak{a}, \mathfrak{b}$  are subalgebras of  $\mathfrak{g}$  and if one of them is toral, then  $\mathfrak{a} + \mathfrak{b}$  is toral.*

Toral subalgebras exists since  $\mathfrak{g}$  has semisimple elements. For if not then every  $X \in \mathfrak{g}$  would be ad-nilpotent and thus  $\mathfrak{g}$  would be a nilpotent Lie algebra, by Engel's theorem.

**Theorem 20.10.2.** *If  $V$  is a finite dimensional vector space over  $\mathbb{C}$ , if  $T \in \text{End}_{\mathbb{C}}(V)$  is a semisimple and singular endomorphism, then  $\ker(T) \cap T(V) = 0$ .*

**Theorem 20.10.3.** *If  $\mathfrak{h} \subseteq \mathfrak{g}$  is a toral Abelian subalgebra, then  $\mathfrak{h}$  is Abelian.*

Since the elements of  $\mathfrak{h}$  are commuting semisimple elements, the operators  $\{\text{ad}H : H \in \mathfrak{h}\}$  are simultaneously diagonalizable. Recal that if  $X$  is a common eigenvector for all  $\{\text{ad}H\}_{H \in \mathfrak{h}}$ , then  $[H, X] = \lambda_H X$  for some  $\lambda_H \in \mathbb{C}$ . This  $\mathfrak{g}$  can be decomposed as the direct sum of subspaces called weight spaces or root spaces as:

$$\mathfrak{g} = \mathfrak{g}_0 \oplus (\bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha) \quad (20.10.1)$$

Where  $\alpha \in \mathfrak{h}^* \setminus \{0\}$  and:

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} : \forall H \in \mathfrak{h}, [H, X] = \alpha(H)X\} \quad (20.10.2)$$

and:

$$\Delta = \{\alpha \in \mathfrak{h}^* \setminus \{0\} : \mathfrak{g}_\alpha \neq 0\} \quad (20.10.3)$$

### Definition 20.10.2: Relative Roots

The roots of  $\mathfrak{g}$  relative to  $\mathfrak{h}$  are the nonzero  $\alpha \in \mathfrak{h}^*$  such that  $\mathfrak{g}_\alpha \neq 0$ .

**Theorem 20.10.4.** *For all  $\alpha, \beta \in \mathfrak{h}^*$ ,  $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$ .*

*Proof.* For let  $x \in \mathfrak{g}_\alpha$  and  $y \in \mathfrak{g}_\beta$ . Then for all  $H \in \mathfrak{h}$ :

$$[H, X] = \alpha(H)X \quad [H, Y] = \beta(H)Y \quad (20.10.4)$$

And therefore:

$$[H, [X, Y]] = [[H, X], Y] + [X, [H, Y]][\alpha(H)X, Y] + [X, \beta(H)Y] \quad (20.10.5)$$

□

**Theorem 20.10.5.** *If  $\mathfrak{g}$  is a semisimple Lie algebra over  $\mathbb{C}$ , if  $\mathfrak{h} \subseteq \mathfrak{g}$  is a maximal toral subalgebra, then:*

$$\mathfrak{g} \simeq C_{\mathfrak{g}}(\mathfrak{h}) \oplus \left( \bigoplus_{\alpha \in \Delta} \mathfrak{g}_{\alpha} \right) \quad (20.10.6)$$

Moreover  $C_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$ .

This is the root space decomposition of  $\mathfrak{g}$ . Let  $\mathfrak{g}$  be a semisimple complex Lie algebra, and let  $\mathfrak{h} \subseteq \mathfrak{g}$  be a maximal toral subalgebra. By the previous theorem, the killing form restricted to  $\mathfrak{h}$  is nondegenerate and so it defined a linear isomorphism  $b : \mathfrak{h} \rightarrow \mathfrak{h}^*$  by  $H \mapsto B_{\mathfrak{g}}(H, \cdot)|_{\mathfrak{h}}$ . Thus for all  $\psi \in \mathfrak{h}^*$  there is a unique  $T_{\psi} \in \mathfrak{h}$  such that  $\psi = b(T_{\psi}) = B_{\mathfrak{g}}(T_{\psi}, \cdot)$ . In particular roots  $\alpha \in \Delta$  correspond to vectors  $T_{\alpha} \in \mathfrak{h}$  called the coroots.

## 20.11 Geometric Properties of Root Space Decomposition

Let  $\mathfrak{g}$  be a semisimple Lie algebra over  $\mathbb{C}$  and let  $\mathfrak{h} \subseteq \mathfrak{g}$  be a maximal toral subalgebra.

**Theorem 20.11.1.** *The set of roots  $\Delta$  spans  $\mathfrak{h}^*$ .*

*Proof.* If not, let  $\{\alpha_1, \dots, \alpha_r\}$  be a basis for the span of the roots and enlarge it to a basis of  $\mathfrak{h}^*$ . Let  $\{H_1, \dots, H_n\}$  be the dual basis of  $\mathfrak{h}$ . That is,  $\alpha_i(H_j) = \delta_{ij}$ .  $\square$

Let  $\mathfrak{g}$  be a semisimple Lie algebra over  $\mathbb{C}$  and let  $\mathfrak{h} \subseteq \mathfrak{g}$  be a maximal toral subalgebra. Then  $\mathfrak{h}$  is Abelian and for any  $\alpha \in \mathfrak{g}^*$ , we have:

$$\mathfrak{g}_{\alpha} = \{x \in \mathfrak{g} : \forall_H, [H, X] = \alpha(H)X\} \quad (20.11.1)$$

And  $\mathfrak{g}_0 = \mathfrak{h}$ . The root space decomposition says:

$$\mathfrak{g} = \mathfrak{h} \oplus \left( \bigoplus_{\alpha} \mathfrak{g}_{\alpha} \right) \quad (20.11.2)$$

Also  $[\mathfrak{g}_{\alpha}, \mathfrak{g}_{\beta}] \subseteq \mathfrak{g}_{\alpha+\beta}$  and  $B_{\mathfrak{g}}(\mathfrak{g}_{\alpha}, \mathfrak{g}_{\beta}) = 0$  unless  $\alpha + \beta = 0$ . Let  $\alpha \in \Delta$ , and let  $Y \in \mathfrak{g}_{-\alpha}$ . Then:

$$[X, Y] = B_{\mathfrak{g}}(X, Y)T_{\alpha} \subseteq \mathfrak{g}_0 \quad (20.11.3)$$

**Theorem 20.11.2.** *If  $\alpha \in \Delta$ , then*

$$[\mathfrak{g}_{\alpha}, \mathfrak{g}_{-\alpha}] = \mathbb{C} \cdot T_{\alpha} \quad (20.11.4)$$

**Theorem 20.11.3.** *For all  $\alpha \in \Delta$ :*

$$\alpha(T_\alpha) = B_{\mathfrak{g}}(T_\alpha, T_\alpha) \neq 0 \quad (20.11.5)$$

*Proof.* Suppose  $\alpha(T_\alpha) = 0$ . Then for  $X \in \mathfrak{g}_\alpha$  and for any  $Y \in \mathfrak{g}_{-\alpha}$  we have:

$$[T_\alpha, X] = \alpha(T_\alpha)X = 0 \quad (20.11.6)$$

and similarly  $[T_\alpha, Y] = 0$ . But for any non-zero  $X \in \mathfrak{g}_\alpha$  there is a  $Y \in \mathfrak{g}_{-\alpha}$  such that  $B_{\mathfrak{g}}(X, Y) \neq 0$ . Replacing  $Y$  with a scalar multiple we can assume that  $B_{\mathfrak{g}}(X, Y) = 1$ . But then:

$$B_{\mathfrak{g}}(X, Y)T_\alpha = T_\alpha \quad (20.11.7)$$

Let  $\mathfrak{H}$  be the span of  $X, Y, T_\alpha$ . Then  $[X, Y] = T_\alpha$  and therefore  $\mathfrak{H}$  is isomorphic to the Heisenberg algebra.  $\square$

**Theorem 20.11.4.** *If  $\alpha \in \Delta$  and  $0 \neq X_\alpha \in \mathfrak{g}_\alpha$ , then there is a  $Y_\alpha$  such that the span over  $\mathbb{C}$  of  $X_\alpha, Y_\alpha, H_\alpha$  is a three dimensional subalgebra of  $\mathfrak{g}$  isomorphic to  $sl_2(\mathbb{C})$ .*

**Example 20.11.1** Let  $\mathfrak{g} = sl_2(\mathbb{C})$ . Let  $\mathfrak{h}$  be the subspace of all diagonal matrices with zero trace.

**Example 20.11.2** Hexagons are cool.

**Theorem 20.11.5.** *If  $\alpha$  and  $\beta$  are roots with  $\beta \neq \pm\alpha$ , then  $\beta(H_\alpha) \in \mathbb{Z}$  and the roots of the form  $\beta + k\alpha$ ,  $k \in \mathbb{Z}$ , form an unbroken sequence  $\beta - r\alpha, \dots, \beta + q\alpha$ , with  $r, q \in \mathbb{N}$ . For  $\alpha + \beta \in \Delta$ ,  $\beta(H_\alpha) = r - q$ .*

Let  $B : V \times V \rightarrow k$  be a nondegenerate symmetric bilinear form on a finite dimensional  $V$  over  $k$ . By the non-degeneracy, the flat map  $\flat : V \rightarrow V^*$  which maps  $\flat(v) = \langle v | \cdot \rangle$ , and the inverse map  $\sharp : V^* \rightarrow V$ , allow us to define a bilinear form  $B^* : V^* \times V^* \rightarrow k$  by:

$$B^*(\varphi, \psi) = B(\sharp_B(\varphi), \varphi_B(\psi)) \quad (20.11.8)$$

**Theorem 20.11.6.** *If  $\mathcal{A}$  is a basis of  $\mathfrak{h}^*$  consisted entirely of roots, then all of the roots are contained in the span of  $\mathcal{A}$ .*

For any  $H \in \mathfrak{h}$ , the root space decomposition is an eigenspace decomposition of  $\text{ad}(H)$ . Let  $\mathcal{B}$  be a basis obtained as the union of bases of root spaces. Then  $\text{ad}(H) = 0$ , and  $\text{ad}(H)|_{\mathfrak{g}} = \gamma(H)\text{Id}$ . Then  $B_{\mathfrak{g}}(H_\alpha, H_\beta)$  is the trace of the product of  $\text{ad}(H_\alpha)$  and  $\text{ad}(H_\beta)$ .

## 20.12 Root Systems

We have a vector space  $E_{\mathbb{Q}}$  over  $\mathbb{Q}$  such that  $E_{\mathbb{Q}} = \text{span}_{\mathbb{Q}}(\Delta)$  with an inner product  $\langle \cdot | \cdot \rangle$ . Let  $\langle \alpha | \beta \rangle$  be viewed as a form on  $E = \mathbb{R} \otimes_{\mathbb{Q}} E_{\mathbb{Q}}$ . Then  $\langle \cdot | \cdot \rangle : E \times E \rightarrow \mathbb{R}$ . Moreover,  $\Delta$  spans  $E$ ,  $\Delta$  is finite, and  $0 \notin \Delta$ .

Book Three

Topology



Part IX

# Point-Set Topology



# CHAPTER 21

---

## Topological Spaces

---

We've developed the notions of sets and order, but still cannot talk about some of the most fundamental notions that one comes across in mathematics and physics. Most notably is that of *continuity*. The structure on two lone sets  $A$  and  $B$  is not enough to describe whether or not a function  $f : A \rightarrow B$  is continuous. To do this requires the concept of a *topology*, and from this we get the notion of a *topological space*. This chapter aims to define topologies and show examples of such spaces.

### 21.1 Topologies

When one uses the notation developed for intervals on the real line  $[a, b]$  and  $(a, b)$ , the former is called the *closed* interval and the latter the *open* interval. The clearest difference is that  $[a, b]$  has its endpoints, whereas  $(a, b)$  does not. We can generalize this further. For every point  $x \in (a, b)$  there is some  $\varepsilon > 0$  such that we can fit  $(x - \varepsilon, x + \varepsilon)$  inside of the interval  $(a, b)$ . Namely, define  $\varepsilon$  to be:

$$\varepsilon = \frac{\min\{b - x, x - a\}}{2} \quad (21.1.1)$$

From this we have  $(x - \varepsilon, x + \varepsilon) \subseteq (a, b)$  (Fig. 21.1).

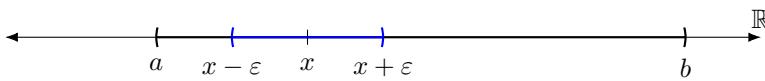


Fig. 21.1: The Open Interval  $(a, b)$  is an Open Subset of  $\mathbb{R}$

This cannot be done for the closed unit interval. If we let  $x = a$ , then for any  $\varepsilon > 0$  we have that  $(a - \varepsilon, a + \varepsilon)$  has points that fall outside of  $[a, b]$ . That is,

all of the points between  $a - \varepsilon$  and  $a$  (Fig. 21.3). This is the distinction we want to note between a closed interval and an open interval, and we wish to axiomatize these properties.

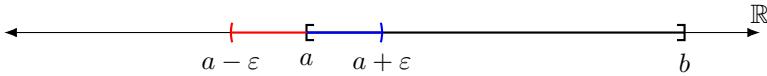


Fig. 21.2: The Closed Interval  $[a, b]$  is Not Open.

If we take the intersection of two intervals  $(a, b)$  and  $(c, d)$ , then the result is either the empty set or it is once again an open interval. More importantly, if the intersection is non-empty then we can find an  $\varepsilon > 0$  such that  $(x - \varepsilon, x + \varepsilon)$  fits inside the intersection. That is, letting  $\varepsilon_1$  and  $\varepsilon_2$  be such that  $(x - \varepsilon_1, x + \varepsilon_1) \subseteq (a, b)$  and  $(x - \varepsilon_2, x + \varepsilon_2) \subseteq (c, d)$ , taking  $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$  gives us the desired result. This is our first *axiom*: Topologies should be closed to the intersection of open sets (see Fig.). Since the intersection of  $(a, b)$  and  $(c, d)$  may be empty, we must consider the empty set  $\emptyset$  to be a member of the topology as well.

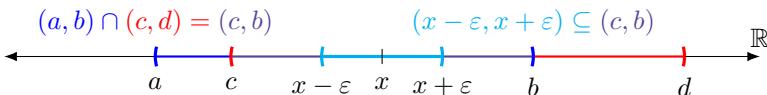


Fig. 21.3: The Intersection of Open Intervals is Open.

Lastly, if we take the union over an arbitrary collection of open intervals, then end result may not be an open interval, but it will still have the property that for any  $x$  in the union there is a  $\varepsilon > 0$  such that  $(x - \varepsilon, x + \varepsilon)$  fits inside the union. This is the property we will ascribe to *open* subsets of  $\mathbb{R}$ : an open subset of  $\mathbb{R}$  is a set  $\mathcal{U} \subseteq \mathbb{R}$  such that for all  $x \in \mathcal{U}$  there is an  $\varepsilon > 0$  such that  $(x - \varepsilon, x + \varepsilon) \subseteq \mathcal{U}$ . It is these two properties of open sets that we wish to capture: Closure to finite intersections and arbitrary unions. We've already noted that the intersection of two open intervals may be empty, and so we want to claim that the empty set is open as well. Lastly, we may regard the entire real line  $\mathbb{R}$  as the *interval*  $(-\infty, \infty)$ , and because of this we wish to regard all of  $\mathbb{R}$  to be open too. In generalizing these notions we can define a *topology* on a set  $X$ .

**Definition 21.1.1: Topology**

A **topology** on a **set**  $X$  is a **subset**  $\tau \subseteq \mathcal{P}(X)$  such that  $\emptyset \in \tau$ ,  $X \in \tau$ , and for any subset  $\mathcal{O} \subseteq \tau$ , it is true that:

$$\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \in \tau$$

And such that for all  $A, B \in \tau$ , it is true that  $A \cap B \in \tau$ .

There are many trivial examples of topologies on any set  $X$ , but the trivial examples often provide excellent counterexamples for various propositions. The two extreme topologies are the *discrete* topology and the *chaotic* topology. The chaotic topology is also called the trivial topology or the indiscrete topology.

**Theorem 21.1.1.** *If  $X$  is a set, the  $\mathcal{P}(X)$  is a topology on  $X$ .*

*Proof.* For  $\emptyset \subseteq X$  (Thm. 3.2.1) and  $X \subseteq X$  (Thm. 3.2.6), and thus it is true that  $\emptyset \in \mathcal{P}(X)$  and  $X \in \mathcal{P}(X)$  (Def. 3.1.12). If  $\mathcal{O} \subseteq \mathcal{P}(X)$ , then  $\bigcup \mathcal{O} \subseteq X$  (Def. 3.1.6). Lastly, if  $A, B \in \mathcal{P}(X)$ , then  $A \subseteq X$  and  $B \subseteq X$ , and thus  $A \cap B \subseteq X$  (Thm. 3.2.45). Thus,  $\mathcal{P}(X)$  is a topology on  $X$  (Def. 21.1.1).  $\square$

**Theorem 21.1.2.** *If  $X$  is a set, then  $\{\emptyset, X\}$  is a topology on  $X$ .*

*Proof.* For by definition  $\emptyset \in \{\emptyset\}$  and  $X \in \{\emptyset, X\}$  (Thm. 3.1.3). If  $\mathcal{O} \subseteq \{\emptyset, X\}$ , then by the law of the excluded middle either  $X \in \mathcal{O}$  or  $X \notin \mathcal{O}$ . If  $X \in \mathcal{O}$ , then  $\bigcup \mathcal{O} = X$ , and if not, then  $\bigcup \mathcal{O} = \emptyset$ . In either case,  $\bigcup \mathcal{O} \in \{\emptyset, X\}$ . If  $A, B \in \{\emptyset, X\}$ , then either one of these is the empty set or not. If so, then  $A \cap B = \emptyset$  (Thm. 3.2.42), and if not then  $A = X$  and  $B = X$ , and thus  $A \cap B = X$  (Thm. 3.2.44). Thus,  $A \cap B \in \{\emptyset, X\}$ , and  $\{\emptyset, X\}$  is a topology on  $X$  (Def. 21.1.1).  $\square$

**Example 21.1.1** The Discrete Topology on a set  $X$  is simply  $\tau = \mathcal{P}(X)$ . That is, the **power set** of  $X$ . This is indeed a topology as proved in Thm. 21.1.1. Since a topology is a subset of  $\mathcal{P}(X)$  (by definition), the discrete topology is thus the *largest* topology that one can consider on a given set. It is called the discrete topology since it arises from the *discrete metric*. This metric, which is of great importance to the theory of *metric spaces*, defines a fairly trivial distance between all of the points in  $X$ . If  $x, y \in X$  and  $x = y$ , the distance is  $d(x, y) = 0$ . If  $x \neq y$ , then  $d(x, y) = 1$ . The topology *generated* from this metric is precisely the discrete topology  $\mathcal{P}(X)$ . One of the strange properties of the discrete topology is that any function  $f : X \rightarrow A$  is *continuous* (to be defined) regardless of the topology on  $A$ .

**Example 21.1.2** The chaotic topology is  $\tau = \{\emptyset, X\}$ . This is also a topology, as demonstrated in Thm. 21.1.2, but we can also check this by simple computation. The empty set and the whole space are contained within  $\tau$ , and there are only four possible ways to perform unions and intersections. The commutativity of union and intersection reduces this to three. We have:

$$\emptyset \cap \emptyset = \emptyset \quad (21.1.2a) \qquad \emptyset \cup \emptyset = \emptyset \quad (21.1.2d)$$

$$\emptyset \cap X = \emptyset \quad (21.1.2b) \qquad \emptyset \cup X = X \quad (21.1.2e)$$

$$X \cap X = X \quad (21.1.2c) \qquad X \cup X = X \quad (21.1.2f)$$

The chaotic topology is so named for one of its strange properties. As we will see once we've defined continuity, the chaotic topology has the property that every single function  $f : A \rightarrow X$  will be continuous, regardless of what the topology chosen on  $A$  was. Unlike the discrete topology, the chaotic topology is almost never associated with a metric space. Indeed, the chaotic space is induced from a metric if and only if  $X$  has one point (in which case it is equivalent to the discrete topology).

**Example 21.1.3** The Sierpinski topology is a topology placed on  $\mathbb{Z}_2 = \{0, 1\}$ . It is defined as the set:

$$\tau_S = \{\emptyset, \{0\}, \mathbb{Z}_2\} \quad (21.1.3)$$

By definition  $\tau_S$  contains both the empty set and the entirety of  $\mathbb{Z}_2$ . Checking unions is trivial since  $\emptyset \subseteq \{0\} \subseteq \mathbb{Z}_2$ , and thus any union of any collection  $\mathcal{O} \subseteq \tau_S$  is simply the *largest* set in the collection, meaning  $\tau_S$  is closed to arbitrary unions. Moreover, all of the intersections are trivial: If  $A = \emptyset$ , then  $A \cap B = \emptyset$ , if  $A = \mathbb{Z}_2$ , then  $A \cap B = B$ , and lastly if  $A = \{0\}$  and  $B = \{0\}$ , then  $A \cap B = \{0\}$  meaning  $\tau_S$  is closed to finite intersections. Hence, it is a topology (Def. 21.1.1). This too serves as a test case for many plausible propositions.

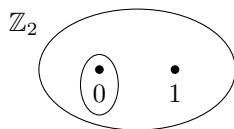


Fig. 21.4: The Sierpinski Topology

We can be a little more general, defining a Sierpinski-like topology on any non-empty set.

**Theorem 21.1.3.** *If  $X$  is a non-empty set and if  $x \in X$ , then the set  $\tau$  defined by  $\tau = \{\emptyset, \{x\}, X\}$  is a topology on  $X$ .*

*Proof.* For  $\emptyset, X \in \tau$ . If  $\mathcal{O} \subseteq \tau$ , either  $X \in \mathcal{O}$  or  $X \notin \mathcal{O}$ . If  $X \in \mathcal{O}$ , then  $\bigcup \mathcal{O} = X$ , which is contained in  $\tau$ . If  $X \notin \mathcal{O}$ , either  $\{x\} \in \mathcal{O}$  or not. If

$\{x\} \in \mathcal{O}$ , then  $\bigcup \mathcal{O} = \{x\}$ , which is contained in  $\tau$ . Lastly, if neither  $X$  nor  $\{x\}$  are elements of  $\mathcal{O}$ , then either  $\mathcal{O} = \emptyset$  or  $\mathcal{O} = \{\emptyset\}$ . In either case,  $\bigcup \mathcal{O} = \emptyset$ . Thus,  $\tau$  is closed to arbitrary unions. Similarly, if  $A, B \in \tau$ , and if either  $A$  or  $B$  are empty, then  $A \cap B = \emptyset$ . If neither  $A$  nor  $B$  are empty, then either  $A$  or  $B$  is the set  $\{x\}$ , or both are the whole space. In the former case  $A \cap B = \{x\}$  and in the latter  $A \cap B = X$ . In both scenarios the intersection is contained in  $\tau$ . Hence,  $\tau$  is a topology on  $X$  (Def. 21.1.1).  $\square$

If  $X$  is a single point,  $X = \{x\}$ , then the Sierpinski-like topology defined in Thm. 21.1.3 is simply the chaotic topology, which is also the discrete topology on one point. This is because, since  $\{x\} = X$ , the set  $\{\emptyset, \{x\}, X\}$  is equal to the set  $\{\emptyset, X\}$  (recall that sets have no notion of repetition). A further generalization goes as follows:

**Theorem 21.1.4.** *If  $X$  is a set, if  $\mathcal{U} \subseteq X$ , and if  $\tau = \{\emptyset, \mathcal{U}, X\}$ , then  $\tau$  is a topology on  $X$ .*

*Proof.* For  $\emptyset, X \in \tau$ . If  $\mathcal{O} \subset \tau$ , either  $X \in \mathcal{O}$  or not. If it is, then  $\bigcup \mathcal{O} = X$ . If not, then either  $\mathcal{U} \in \mathcal{O}$  or not. If it is, then  $\bigcup \mathcal{O} = \mathcal{U}$ . If not, then  $\bigcup \mathcal{O} = \emptyset$ . In any of the three cases, the union is contained in  $\tau$  and so we have closure with respect to arbitrary unions. Similarly, it is closed to intersections, and is therefore a topology (Def. 21.1.1).  $\square$

A set combined with a topology on it is called a topological space. These are the main objects, together with continuous functions, of study in topology.

### Definition 21.1.2: Topological Space

A **topological space**, denote  $(X, \tau)$  is a **set**  $X$  and a **topology**  $\tau$  on  $X$ .

We pause to briefly digress about metric spaces, which are the most important example of topological spaces. Metric spaces are spaces which generalize the notion of *distance* in Euclidean space. These were first introduced by Maurice Fréchet in 1906, and were the primary motivator for developing point-set topology in detail. A *metric* on a set  $X$  is a function  $d : X \times X \rightarrow \mathbb{R}$  with the following four properties:

$$d(x, y) \geq 0 \quad (\text{Positivity})$$

$$d(x, y) = 0 \implies x = y \quad (\text{Definiteness})$$

$$d(x, y) = d(y, x) \quad (\text{Symmetry})$$

$$d(x, z) \leq d(x, y) + d(y, z) \quad (\text{Triangle Inequality})$$

The problem of proving whether or not a given function is a metric is eased slightly since symmetry can be proved from definiteness and the triangle inequality. As stated before, this is a generalization of the notion of distance, and

as such the primary examples come from Euclidean space. In  $\mathbb{R}$  the distance function is the *absolute value* function:  $d(x, y) = |x - y|$ . In higher dimension,  $\mathbb{R}^n$ , we use the Pythagorean theorem to define distance:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{k \in \mathbb{Z}_n} (x_k - y_k)^2} \quad (21.1.4)$$

A subset  $\mathcal{U}$  of a metric space  $(X, d)$  is called *metrically open* if for all  $x \in \mathcal{U}$  there is an  $\varepsilon > 0$  such that, for all  $y \in X$  such that  $d(x, y) < \varepsilon$ , it is true that  $y \in \mathcal{U}$ . The collection of all metrically open subsets of a metric space form a topology on  $X$ , and this is called the topology induced by the metric. Metric spaces are extremely well-behaved spaces and many theorems in topology are dedicated to proving that certain topological spaces can be given compatible metrics.

There are two common ways in which to teach point-set topology: start with metric spaces or end with them. Starting with them is the pedagogical choice, since they are intuitive and any student who has taken a course in real analysis should have no problem solving problems about such spaces. The logical route is to end with them, since many of the fundamental theorems about metric spaces become short corollaries when one has developed a sufficient amount of topology. We will adopt the latter option, saving metric spaces for after we've developed connectedness, compactness, and the plethora of other *topological* concepts (concepts that do not need a notion of distance, just a topology). An attempt will be made to give intuition in the form of examples, usually with Euclidean spaces or by presenting pictures. It is hoped that the abstraction will not be too confusing, but the end result is well worth it. Indeed, theorems such as the *intermediate value theorem* and the *extreme value theorem* from calculus can be proven in just a few lines once one has the right tools.

One of the common problems in analysis, geometry, and topology is placing a *natural* topology on a given set. That is, some how concocting an *obvious* or most useful topology. The construction often goes by considering the smallest topology that has a certain property, smallest meaning the intersection of all other topologies with said property. It would be essential, then, to know that the intersection of topologies is again a topology. We prove this now.

### Theorem 21.1.5: The Intersection of Topologies is a Topology

---

If  $X$  is a set, if  $\Omega$  is a non-empty set such that for all  $\tau \in \Omega$  it is true that  $\tau$  is a topology on  $X$ , then  $\bigcap \Omega$  is a topology on  $X$ . ■

*Proof.* For since  $\Omega$  is non-empty there is a set  $\tau \in \Omega$  (Def. 3.1.1). But by hypothesis, if  $\tau \in \Omega$  then  $\tau$  is a topology on  $X$ , and thus  $\emptyset, X \in \tau$ . Suppose

$\emptyset \notin \bigcap \Omega$ . Then there exists  $\xi \in \Omega$ ,  $\xi \neq \tau$ , such that  $X \notin \xi$  (Def. 3.1.10). But if  $\xi \in \Omega$  then  $\xi$  is a topology on  $X$ , and thus  $X \in \xi$  (Def. 21.1.1), thus  $X \in \xi$ , and similarly  $\emptyset \in \xi$ . Suppose there is an  $\mathcal{O} \subseteq \bigcap \Omega$  such that  $\bigcup \mathcal{O} \notin \bigcap \Omega$ . But for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $\mathcal{U} \in \bigcap \Omega$  since  $\mathcal{O}$  is a subset of  $\bigcap \Omega$  (Def. 3.1.2). But then for all  $\tau \in \bigcap \Omega$  is it true that  $\mathcal{U} \in \tau$  (Def. 3.1.10). But  $\tau$  is a topology on  $X$ , and thus  $\bigcup \mathcal{O} \in \tau$  (Def. 21.1.1). And this is true of all such  $\tau$ , and thus  $\bigcup \mathcal{O} \in \bigcap \Omega$ , a contradiction. Thus  $\bigcap \Omega$  is closed to arbitrary unions. Lastly, if  $A, B \in \bigcap \Omega$ , then for all  $\tau \in \bigcap \Omega$  it is true that  $A, B \in \tau$ . But  $\tau$  is a topology, and thus  $A \cap B \in \tau$  (Def. 21.1.1). And since this is true of all  $\tau \in \bigcap \Omega$ , it is true that  $A \cap B \in \bigcap \Omega$  (Def. 3.1.10). Therefore  $\bigcap \Omega$  is a topology on  $X$  (Def. 21.1.1).  $\square$

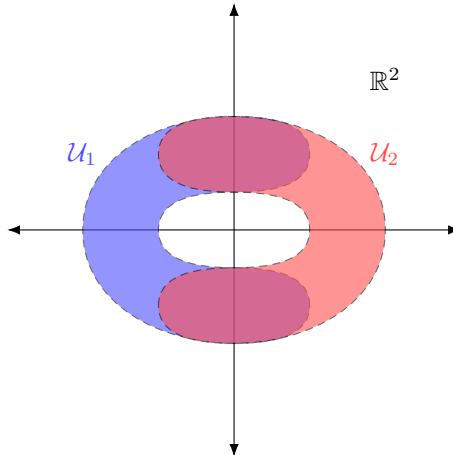


Fig. 21.5: The Union of Topologies Need Not be Closed to Finite Intersections

However, the union of topologies may not be a topology. For consider the two Sierpinski-like topologies on  $\mathbb{Z}_3$  defined by  $\tau_1 = \{\emptyset, \{0\}, \mathbb{Z}_3\}$  and  $\tau_2 = \{\emptyset, \{1\}, \mathbb{Z}_3\}$ . Both of these are topologies, however  $\tau_1 \cup \tau_2 = \{\emptyset, \{0\}, \{1\}, \mathbb{Z}_3\}$ , which is not a topology since it is not closed to unions. Note that  $\{0\} \in \tau_1 \cup \tau_2$ , as is the set  $\{1\}$ , however  $\{0, 1\}$  is not. This is why we chose  $\mathbb{Z}_3$  for our counterexample. Picking  $\mathbb{Z}_2$  would actually result in a topology since  $\{0, 1\} = \mathbb{Z}_2$ , and the whole space would be contained in the union. Thus we see that the union of topologies may not be closed to arbitrary unions, but it need not be closed to intersections either. Consider the topologies on  $\mathbb{R}^2$  defined by  $\tau_1 = \{\emptyset, \mathcal{U}_1, \mathbb{R}^2\}$  and  $\tau_2 = \{\emptyset, \mathcal{U}_2, \mathbb{R}^2\}$ , where  $\mathcal{U}_1$  and  $\mathcal{U}_2$  are the two blobs shown in Fig. 21.5. Since  $\mathcal{U}_1 \subseteq \mathbb{R}^2$ , as well as  $\mathcal{U}_2 \subseteq \mathbb{R}^2$ , by Thm. 21.1.4 we have that both  $\tau_1$  and  $\tau_2$  are topologies. However, as indicated by the figure, the union  $\tau_1 \cup \tau_2$  is not closed to intersections since  $\mathcal{U}_1 \cap \mathcal{U}_2$  is not contained in this union.

Moreover,  $\mathcal{U}_1 \cup \mathcal{U}_2$  is not either, and hence  $\tau_1 \cup \tau_2$  is closed to neither arbitrary unions nor finite intersections.

In some contexts it is necessary to compare topologies on a given set. For example, if one were to study measure theory, there are two topologies one could study on  $\mathbb{R}$ : The standard topology  $\tau_{\mathbb{R}}$  which is the one induced by the standard metric  $d(x, y) = |x - y|$ , and the topology associated with the *Lebesgue*  $\sigma$ -Algebra,  $\tau_L$ . As it turns out,  $\tau_{\mathbb{R}} \subseteq \tau_L$ . A topology  $\tau_1$  is often said to be *finer* than a topology  $\tau_0$  if  $\tau_0 \subseteq \tau_1$ . Thus, the Lebesgue topology is finer than the standard one. There's a *fine* line between a topology being too large and too small. The chaotic topology is essentially useless, and is the smallest possible topology, but the discrete topology is also rather poor, whereas it is the largest. The standard topology on  $\mathbb{R}$  is good middle ground and many of the wonderous results of calculus come from this structure. The Lebesgue topology  $\tau_{\mathbb{R}}$  being only slightly larger (having the same cardinality) is horrendous and leads to counterintuitive and undesirable results (to be elaborated on in Book [Four](#)).

We develop some more terminology to become acquainted with the language of point-set topology.

### Definition 21.1.3: Open Subset

An open subset of a [topological space](#)  $(X, \tau)$  is a [set](#)  $\mathcal{U} \in \tau$ .

That is, an open subset of a topological space  $(X, \tau)$  is simply an element of the topology.  $\mathcal{U} \subseteq X$  is open if and only if  $\mathcal{U} \in \tau$ . Thus we could, rather circularly, define the topology to be the collection of all open subsets of  $X$ .

**Theorem 21.1.6.** *If  $(X, \tau)$  is a topological space, then  $\emptyset$  is open.*

*Proof.* For since  $(X, \tau)$  is a topological space,  $\tau$  is a topology on  $X$  (Def. 21.1.2). But then  $\emptyset \in \tau$  (Def. 21.1.1) and thus  $\emptyset$  is open (Def. 21.1.3).  $\square$

**Theorem 21.1.7.** *If  $(X, \tau)$  is a topological space, then  $X$  is an open subset.*

*Proof.* For since  $(X, \tau)$  is a topological space,  $\tau$  is a topology on  $X$  (Def. 21.1.2). But then  $X \in \tau$  (Def. 21.1.1) and thus  $X$  is open. (Def. 21.1.3).  $\square$

Equally important is the notion of *closed sets*. Indeed, topologies may equivalently defined solely in terms of closed sets.

**Definition 21.1.4: Closed Subset**

A closed subset of a topological space  $(X, \tau)$  is a subset  $\mathcal{C} \subseteq X$  such that  $X \setminus \mathcal{C} \in \tau$ . That is, the complement of  $\mathcal{C}$  is an open subset of  $(X, \tau)$ .

**Theorem 21.1.8.** *If  $(X, \tau)$  is a topological space, then  $\emptyset$  is closed.*

*Proof.* For  $\emptyset = X \setminus X$  (Thm. 3.2.56) and  $X$  is open (Thm. 21.1.7). Thus,  $\emptyset$  is closed (Def. 21.1.4)  $\square$

**Theorem 21.1.9.** *If  $(X, \tau)$  is a topological space, then  $X$  is closed.*

*Proof.* For  $X = X \setminus \emptyset$  (Thm. 3.2.58), and  $\emptyset$  is open (Thm. 21.1.6). Thus,  $X$  is closed (Def. 21.1.4).  $\square$

**Theorem 21.1.10.** *If  $(X, \tau)$  is a topological space, and if  $\mathcal{U} \subseteq X$  is open, then  $X \setminus \mathcal{U}$  is closed.*

*Proof.* For if  $\mathcal{U} \subseteq X$  is a set, then  $X \setminus (X \setminus \mathcal{U}) = \mathcal{U}$ . But  $\mathcal{U}$  is open, and thus the complement of  $X \setminus \mathcal{U}$  is open. But then  $X \setminus \mathcal{U}$  is closed (Def. 21.1.4).  $\square$

Thus, by definition, a closed set is a set whose complement is open, and equivalently an open set is a set whose complement is closed. A common misunderstanding is that sets are either closed or open, but not both. This is false as we've just shown, the entire space  $X$  is both closed and open, and the empty set  $\emptyset$  is also both closed and open. Sets of this nature are called *clopen*. As it turns out, clopen sets are indeed rare and only occur in *disconnected* spaces and as such a discussion on clopen sets will be saved for when we discuss the notion of connectedness. It is further worth pointing out that it is possible for a set to be neither closed nor open. Indeed, consider any *large* set  $X$  (having at least a few points), and endow it with the chaotic (or trivial) topology  $\tau$ . Then every proper non-empty subset will be neither open nor closed, since the only open sets are  $X$  and  $\emptyset$ , and similarly these are the only closed sets. For a concrete example one can consider the real numbers. The subset  $\mathbb{Q}$  of rational numbers is neither open nor closed. It is not open since for any point  $x \in \mathbb{Q}$  and for any  $\varepsilon > 0$  the open interval  $(x - \varepsilon, x + \varepsilon)$  contains irrational numbers, and thus  $(x - \varepsilon, x + \varepsilon) \not\subseteq \mathbb{Q}$ . It is not closed since its complement is the set of irrational numbers, and by a similar argument this set is not open either. Thus  $\mathbb{Q}$  is neither open nor closed, and similarly the set of irrational numbers is neither open nor closed.

Much the way a topology can be defined by declaring certain sets to be open, we can do the same thing by declaring certain sets to be closed. Such schemes

are occasionally useful, such as in algebraic topology where one defines the *Zariski Topology* in terms of closed sets. A few theorems are then useful, all of which are simple applications of the DeMorgan laws.

**Theorem 21.1.11.** *If  $X$  is a set, if  $\tau$  is a topology on  $X$ , and if  $\mathcal{C} \subseteq \tau$  is such that for all  $\mathcal{U} \in \mathcal{C}$  it is true that  $\mathcal{U}$  is closed, then  $\bigcap \mathcal{C}$  is closed.*

### Definition 21.1.5: Relative Topology

The relative topology of a subset  $A \subseteq X$  in a topological space  $(X, \tau)$  is the set  $\tau_A$  defined by:

$$\tau_A = \{ A \cap \mathcal{U} : \mathcal{U} \in \tau \} \quad (21.1.5)$$

■

**Theorem 21.1.12.** *If  $(X, \tau)$  is a topological space, if  $A \subseteq X$ , and if  $\tau_A$  is the relative topology on  $A$ , then  $\tau_A$  is a topology on  $A$ .*

*Proof.* For since  $\emptyset \in \tau$ , and  $\emptyset \cap A = \emptyset$ , we have that  $\emptyset \in \tau_A$ . Similarly, since  $A \subseteq X$ , and since  $X \in \tau$ , we see that  $A = A \cap X \in \tau_A$ . If  $\mathcal{O}$  is a subset of  $\tau_A$ , then there is a subset  $\Delta \subseteq \tau$  such that:

$$\mathcal{O} = \{ A \cap \mathcal{U} : \mathcal{U} \in \Delta \} \quad (21.1.6)$$

Define  $\mathcal{D}$  by:

$$\mathcal{D} = \bigcup_{\mathcal{U} \in \Delta} \mathcal{U} \quad (21.1.7)$$

But then:

$$\bigcup_{\mathcal{V} \in \mathcal{O}} \mathcal{V} = \bigcup_{\mathcal{U} \in \Delta} (A \cap \mathcal{U}) = A \cap \left( \bigcup_{\mathcal{U} \in \Delta} \mathcal{U} \right) = A \cap \mathcal{D} \quad (21.1.8)$$

But  $\tau$  is a topology on  $X$ , and thus  $\mathcal{D} \in \tau$  (Def. 21.1.1). But then  $A \cap \mathcal{D} \in \tau_A$  (Def. 21.1.5). Thus,  $\tau_A$  is closed to arbitrary unions. □

**Definition 21.1.6** If  $(X, \tau)$  is a topological space and  $S \subset X$ , then a set  $\mathcal{U} \subset S$  is said to be open in  $S$  if and only if  $\mathcal{U} \in \mathcal{T}$ , where  $\mathcal{T}$  is the relative topology on  $S$ .

### Definition 21.1.7: Continuous Functions

A continuous function from a topological space  $(X, \tau_X)$  to a topological space  $(Y, \tau_Y)$  is a function  $f : X \rightarrow Y$  such that for all  $\mathcal{U} \in \tau_Y$  it is true that  $f^{-1}(\mathcal{U}) \in \tau_X$ . That is, the pre-image of open sets is open.

### Notation 21.1.1: Set of Continuous Functions

The set of continuous functions  $f : X \rightarrow Y$  is denoted  $C(X, Y)$ .

**Theorem 21.1.13.** *If  $(X, \tau)$  is a topological space, and if  $\text{id}_X$  is the identity function of  $X$ , then  $\text{id}_X$  is continuous.*

*Proof.* For suppose not. Then there is a  $\mathcal{U} \in \tau$  such that  $\text{id}_X(\mathcal{U}) \notin \tau$ . But if  $\mathcal{U} \in \tau$  then  $\mathcal{U} \subseteq X$ , and therefore  $\text{id}_X(\mathcal{U}) = \mathcal{U}$ . But then  $\text{id}_X(\mathcal{U}) \in \tau$ , a contradiction. Therefore,  $\text{id}_X$  is continuous.  $\square$

**Theorem 21.1.14.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, and if  $f : X \rightarrow Y$  is a constant mapping, then  $f$  is continuous.*

*Proof.* For suppose not. Then there is a  $\mathcal{U} \in \tau_Y$  such that  $f^{-1}(\mathcal{U}) \notin \tau_X$ . But if  $f$  is a constant mapping, then there is a  $y \in Y$  such that, for all  $x \in X$  it is true that  $f(x) = y$ . By the law of the excluded middle either  $y \in \mathcal{U}$  or  $y \notin \mathcal{U}$ . If  $y \notin \mathcal{U}$  then  $f^{-1}(\mathcal{U}) = \emptyset$ . But  $\tau_X$  is a topology and therefore  $\emptyset \in \tau_X$ . Thus  $y \in \mathcal{U}$ . But if  $y \in \mathcal{U}$  then  $f^{-1}(\mathcal{U}) = X$ . But  $\tau_X$  is a topology and thus  $X \in \tau_X$ . But then  $f^{-1}(\mathcal{U}) \in \tau_X$ , a contradiction. Therefore,  $f$  is continuous.  $\square$

**Theorem 21.1.15.** *If  $(X, \tau_X)$ ,  $(Y, \tau_Y)$ , and  $(Z, \tau_Z)$  are topological spaces, and if the functions  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  are continuous, then  $g \circ f : X \rightarrow Z$  is continuous.*

*Proof.* For if  $\mathcal{V} \in \tau_Z$  is an open set, then  $g^{-1}(\mathcal{V}) \in \tau_Y$ , since  $g$  is continuous. But then since  $f$  is continuous,  $f^{-1}(g^{-1}(\mathcal{V})) \in \tau_X$  (Def. 21.1.7). Thus  $g \circ f$  is continuous.  $\square$

### Definition 21.1.8: Convergent Sequences In Topological Spaces

A convergent sequence in a topological space  $(X, \tau)$  is a sequence  $a : \mathbb{N} \rightarrow X$  such that there is an  $x \in X$  such that, for all  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$ , there is an  $N \in \mathbb{N}$  such that, for all  $n > N$ , it is true that  $a_n \in \mathcal{U}$ . We denote this by  $a_n \rightarrow x$ .  $\blacksquare$

### Definition 21.1.9: Limits of Sequences in Topological Spaces

A limit of a sequence  $a : \mathbb{N} \rightarrow X$  in a topological space  $(X, \tau)$  is a point  $x \in X$  such that  $a_n \rightarrow x$ .  $\blacksquare$

**Theorem 21.1.16.** *There exist topological spaces with convergent sequences that do not have unique limits.*

*Proof.* For let  $X = \{1, 2, 3\}$ , and let  $\tau = \{\emptyset, \{1, 2\}, \{1, 2, 3\}\}$ . We see that  $\emptyset, X \in \tau$ , unions and intersections are in  $\tau$ , and thus  $\tau$  is a topology. Let:

$$x_n = \begin{cases} 1, & n \text{ odd} \\ 2, & n \text{ even} \end{cases} \quad (21.1.9)$$

Then  $x_n \rightarrow 1$  and  $x_n \rightarrow 2$ . To see this, let  $\mathcal{U}$  be an open set such that  $1 \in \mathcal{U}$ . Our choices are  $\{1, 2\}$  and  $\{1, 2, 3\}$ . Then for all  $n \in \mathbb{N}$ ,  $x_n \in \mathcal{U}$ , and thus  $x_n \rightarrow 1$ . Similarly,  $x_n \rightarrow 2$ . Convergence is not necessarily unique in topological spaces.  $\square$

**Theorem 21.1.17.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $a : \mathbb{N} \rightarrow X$  is a convergent sequence in  $X$ , and if  $f : X \rightarrow Y$  is a continuous function, then the sequence  $b : \mathbb{N} \rightarrow Y$  defined by  $b_n = f(a_n)$  is a convergent sequence in  $Y$ .*

*Proof.* For if  $a : \mathbb{N} \rightarrow X$  is a convergent sequence, then there is a point  $x \in X$  such that, for all  $\mathcal{U} \in \tau_X$  such that  $x \in \mathcal{U}$ , there is an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  with  $n > N$ , it is true that  $a_n \in \mathcal{U}$ . Suppose  $b : \mathbb{N} \rightarrow Y$  is not a convergent sequence. Then for all  $y \in Y$  there exists  $\mathcal{V} \in \tau_Y$  such that  $y \in \mathcal{V}$  and for all  $N \in \mathbb{N}$  there exists an  $n \in \mathbb{N}$  such that  $n > N$  and  $b_n \notin \mathcal{V}$ . Let  $y = f(x)$  and let  $\mathcal{V}_y$  be such that  $y \in \mathcal{V}_y$  and for all  $N \in \mathbb{N}$  there is an  $n \in \mathbb{N}$  such that  $n > N$  and  $b_n \notin \mathcal{V}_y$ . But  $f$  is continuous, and thus  $f^{-1}(\mathcal{V}_y) \in \tau_X$ . Moreover, since  $y = f(x)$ , it is true that  $x \in f^{-1}(\mathcal{V}_y)$ . But then there is an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  with  $n > N$ , it is true that  $a_n \in f^{-1}(\mathcal{V}_y)$ . But then for all such  $n$  it is true that  $f(a_n) \in \mathcal{V}_y$ . But  $b_n = f(a_n)$  and there exists an  $n > N$  such that  $b_n \notin \mathcal{V}_y$ , a contradiction. Therefore,  $b$  is a convergent sequence in  $Y$ .  $\square$

## 21.1.1 Separation Axioms

### Definition 21.1.10: Fréchet Topological Space

A Fréchet Topological Space is a topological space  $(X, d)$  such that for all distinct  $x, y \in X$  there is an open set  $\mathcal{U}$  such that  $x \in \mathcal{U}$  and  $y \notin \mathcal{U}$ .  $\blacksquare$

**Theorem 21.1.18.** *If  $(X, \tau)$  is a Fréchet Topological space, and if  $x \in X$ ,  $\{x\}$  is closed subset.*

*Proof.* For if  $x \in X$ , then for all  $y \in X$  such that  $y \neq x$ , there is a  $\mathcal{U}_y \in \tau$  such that  $x \notin \mathcal{U}_y$  and  $y \in \mathcal{U}_y$ . Define  $\mathcal{V}$  by:

$$\mathcal{V} = \bigcup_{y \in X \setminus \{x\}} \mathcal{U}_y \quad (21.1.10)$$

But then  $\mathcal{V} \in \tau$ , since  $\tau$  is a topology (Def. 21.1.1). And moreover, for all  $y \in X$  such that  $y \neq x$ , it is true that  $y \in \mathcal{V}$ . Lastly,  $x \notin \mathcal{V}$ . Therefore  $X \setminus \mathcal{V} = \{x\}$ . But if  $\mathcal{V}$  is open, then  $X \setminus \mathcal{V}$  is closed (Thm. 21.1.10). Therefore, etc.  $\square$

### Definition 21.1.11: Hausdorff Topological Space

A Hausdorff Topological space is a topological space  $(X, \tau)$  such that, for all distinct points  $x, y \in X$ , there are disjoint open subsets  $\mathcal{U}_x, \mathcal{U}_y \in \tau$  such that  $x \in \mathcal{U}_x$  and  $y \in \mathcal{U}_y$ .  $\blacksquare$

**Theorem 21.1.19.** *If  $(X, \tau)$  is a Hausdorff topological space, then it is a Fréchet topological space.*

*Proof.* For if  $x$  and  $y$  are distinct points in  $X$  and if  $(X, \tau)$  is a Hausdorff topological space, then there are disjoint open subsets  $\mathcal{U}_x$  and  $\mathcal{U}_y$  such that  $x \in \mathcal{U}_x$  and  $y \in \mathcal{U}_y$ . But then there is a open subset such that  $x \notin \mathcal{U}_y$  and  $y \in \mathcal{U}_y$ . Therefore,  $(X, \tau)$  is a Fr'echet topological space.  $\square$

**Theorem 21.1.20.** *Convergence in a Hausdorff Space  $(X, \tau)$  is unique.*

*Proof.*  $[x_n \rightarrow x \in X] \wedge [x_n \rightarrow y \in X] \wedge [x \neq y] \Rightarrow [\exists \mathcal{U}, \mathcal{V} : \mathcal{U} \cap \mathcal{V} = \emptyset \wedge x \in \mathcal{U} \wedge y \in \mathcal{V}]$ .  $[x_n \rightarrow x] \Rightarrow [\exists N_1 \in \mathbb{N} : n > N_1 \Rightarrow x_n \in \mathcal{U}]$ .  $[x_n \rightarrow y] \Rightarrow [\exists N_2 \in \mathbb{N} : n > N_2 \Rightarrow x_n \in \mathcal{V}]$ .  $[n > \max\{N_1, N_2\}] \Rightarrow [x_n \in \mathcal{U} \cap \mathcal{V}]$ , a contradiction. Therefore, etc.  $\square$

**Definition 21.1.12** A topological space  $(X, \tau)$  is said to be regular if for each closed subset  $E \subset X$  and for each point  $x \in E^c$ , there exist disjoint open sets  $\mathcal{U}$  and  $\mathcal{V}$  such that  $x \in \mathcal{U}$  and  $E \subset \mathcal{V}$ .

**Definition 21.1.13** In a topological space  $(X, \tau)$ , a point  $p$  is said to have a neighborhood  $S \subset X$  if and only if there is a set  $\mathcal{U} \subset S$  such that  $\mathcal{U} \in \tau$  and  $p \in \mathcal{U}$ .

**Definition 21.1.14** A  $T_3$  space is a regular  $T_1$  space.

**Theorem 21.1.21.** *A  $T_3$  space  $(X, \tau)$  is a  $T_2$  space.*

*Proof.* Let  $x, y \in X$  be distinct. As a  $T_3$  space is  $T_1$ ,  $\{x\}$  is closed. Thus  $\exists \mathcal{U}, \mathcal{V} \in \tau : \mathcal{U} \cap \mathcal{V} = \emptyset, \{x\} \subset \mathcal{U}$ , and  $y \in \mathcal{V}$ .  $\square$

**Definition 21.1.15** A topological space  $(X, \tau)$  is said to be normal if and only if for all disjoint closed subsets  $E, F \subset X$ , there are disjoint open sets  $\mathcal{U}$  and  $\mathcal{V}$  such that  $E \subset \mathcal{U}$  and  $F \subset \mathcal{V}$ .

**Definition 21.1.16** A  $T_4$  space is a normal  $T_1$  space.

**Theorem 21.1.22.** *A  $T_4$  space  $(X, \tau)$  is a  $T_3$  space.*

*Proof.* A  $T_4$  space is  $T_1$ . If  $E \subset X$  and  $x \in E^c$ , then  $\{x\}$  is closed. Thus  $\exists \mathcal{U}, \mathcal{V} \in \tau : \mathcal{U} \cap \mathcal{V} = \emptyset, \{x\} \subset \mathcal{U}$ , and  $E \subset \mathcal{V}$ .  $\square$

**Definition 21.1.17** A homeomorphism between two topological spaces  $(X, \tau)$  and  $(Y, \tau)$  is a continuous bijection  $f : X \rightarrow Y$  such that  $f^{-1} : Y \rightarrow X$  is continuous.

**Definition 21.1.18** If  $(X, \tau)$  is a topological space, and  $S \subset X$ , then an open cover  $\mathcal{O}$  of  $S$  is a set of open sets  $\mathcal{U}_\alpha$  such that  $S \subset \cup_{\alpha \in A} \mathcal{U}_\alpha$ , where  $A$  is some index set.

**Definition 21.1.19** A subcover of an open cover  $\mathcal{O}$  is a subset of  $\mathcal{O}$  that is also a cover.

**Definition 21.1.20** If  $(X, \tau)$  is a topological space and  $S \subset X$ , then  $S$  is said to be compact if and only if every open cover of  $S$  has a finite subcover.

**Theorem 21.1.23.** *If  $S$  is a compact subset of a Hausdorff space, then for all  $x \in S^c$  there are disjoint open sets  $\mathcal{U}$  and  $\mathcal{V}$  such that  $x \in \mathcal{U}$  and  $S \subset \mathcal{V}$ .*

*Proof.* For let  $x \in S^c$ . For all  $y \in S$  there are disjoint open sets  $\mathcal{U}_y$  and  $\mathcal{V}_y$  such that  $x \in \mathcal{U}_y$  and  $y \in \mathcal{V}_y$ . But then  $\cup_{y \in S} \mathcal{U}_y$  is an open cover of  $S$ . As  $S$  is compact, there is a finite subcover, that is sets  $\mathcal{V}_{y_1}, \dots, \mathcal{V}_{y_n}$  that cover  $S$ . But then  $\cap_{k=1}^n \mathcal{U}_{y_k}$  is open, contains  $x$  and is disjoint from  $\cup_{k=1}^n \mathcal{V}_{y_k}$ . Therefore, etc.  $\square$

**Theorem 21.1.24.** *Every compact subset of a Hausdorff space  $(X, \tau)$  is closed.*

*Proof.* Let  $S$  be a compact subset of a  $X$ .  $\forall x \in S^c, \exists \mathcal{U}_x \in \tau : \mathcal{U}_x \cap S = \emptyset : x \in \mathcal{U}_x$ . But then  $S^c \subset \cup_{x \in S^c} \mathcal{U}_x$ . But also  $S \cap (\cup_{x \in S^c} \mathcal{U}_x) = \emptyset$ . Thus  $S^c = \cup_{x \in S^c} \mathcal{U}_x$ , and therefore  $S^c$  is open. Thus  $S$  is closed.  $\square$

**Theorem 21.1.25.** *If  $S$  is a closed subset of a compact space  $(X, \tau)$ ,  $S$  is compact.*

*Proof.* For let  $\mathcal{O}$  be an open cover of  $S$ . As  $S$  is closed,  $S^c$  is open, and thus  $\{S^c\} \cup \mathcal{O}$  is an open cover  $X$ . As  $X$  is compact, there is a finite subcover, call it  $\mathcal{O}'$ . But then  $\mathcal{O}' \setminus \{S^c\}$  is a finite subcover  $\mathcal{O}$  that covers  $S$ . Thus, etc.  $\square$

**Theorem 21.1.26.** *If  $f : X \rightarrow Y$  is continuous and  $X$  is compact, then  $f(X) \subset Y$  is compact.*

*Proof.* Let  $\mathcal{O}$  be an open cover of  $f(X)$ . As  $f$  is continuous,  $\mathcal{U} \in \mathcal{O} \Rightarrow f^{-1}(\mathcal{U})$  is open in  $X$ . Thus  $\cup_{\mathcal{U} \in \mathcal{O}} f^{-1}(\mathcal{U})$  is an open cover of  $X$ . As  $X$  is compact, there is a finite subcover, say  $\mathcal{O}'$ . But then  $\cup_{\mathcal{V} \in \mathcal{O}'} \mathcal{V}$  is a finite subcover of  $\mathcal{O}$ . Therefore, etc.  $\square$

**Theorem 21.1.27.** *If  $f : X \rightarrow Y$  is a continuous bijection,  $X$  is compact and  $Y$  is Hausdorff, then  $f$  is a homeomorphism.*

*Proof.* It suffices to show that if  $\mathcal{U}$  is open in  $X$ , then  $f(\mathcal{U})$  is open in  $f(X)$ . Let  $\mathcal{U}$  be open in  $X$ . As  $X$  is compact and  $\mathcal{U}$  is open,  $\mathcal{U}^c$  is compact. But then  $f(\mathcal{U}^c) = f(X) \setminus f(\mathcal{U})$  is compact. Thus  $f(X) \setminus f(\mathcal{U}) \underset{\text{Closed}}{\subset} f(X) \Rightarrow f(\mathcal{U}) \underset{\text{Open}}{\subset} f(X)$ .  $\square$

**Definition 21.1.21** A topological space  $(X, \tau)$  is said to be disconnected if and only if there are two disjoint nonempty open sets  $\mathcal{U}$  and  $\mathcal{V}$  such that  $X = \mathcal{U} \cup \mathcal{V}$ .

**Theorem 21.1.28.** *A topological space  $(X, \tau)$  is disconnected if and only if there are two non-empty disjoint closed set  $\mathcal{C}$  and  $\mathcal{D}$  such that  $X = \mathcal{C} \cup \mathcal{D}$ .*

*Proof.*  $[\exists \mathcal{U}, \mathcal{V} \in \tau : [\mathcal{U} \cap \mathcal{V} = \emptyset] \wedge [X = \mathcal{U} \cup \mathcal{V}] \wedge [\mathcal{U}, \mathcal{V} \neq \emptyset]] \Rightarrow [X = \mathcal{U}^c \cup \mathcal{V}^c]$  thus,  $X$  is the union of disjoint, non-empty closed set.  $[\mathcal{C}^c, \mathcal{D}^c \in \tau] \wedge [\mathcal{C}, \mathcal{V} \neq \emptyset] \wedge [\mathcal{C} \cap \mathcal{D} = \emptyset] \wedge [\mathcal{C} \cup \mathcal{D} = X] \Rightarrow [X = \mathcal{C}^c \cup \mathcal{D}^c]$ . Thus  $X$  is disconnected.  $\square$

**Theorem 21.1.29.**  *$(X, \tau)$  is disconnected if and only if there is a proper, nonempty set  $A \subset X$  that is both open and closed.*

*Proof.*  $[\exists \mathcal{U}, \mathcal{V} \in \tau : [\mathcal{U} \cap \mathcal{V} = \emptyset] \wedge [X = \mathcal{U} \cup \mathcal{V}] \wedge [\mathcal{U}, \mathcal{V} \neq \emptyset]] \Rightarrow [\mathcal{U}^c = \mathcal{V}] \Rightarrow [\mathcal{U}^c \in \tau]$ . Thus,  $\mathcal{U}$  is open and closed.  $\square$

**Definition 21.1.22** A topological space is called connected if and only if it is not disconnected.

**Theorem 21.1.30.** *If  $f : X \rightarrow Y$  is a continuous function and  $X$  is connected, then  $f(X)$  is connected.*

*Proof.* For let  $f$  be continuous and  $X$  be connected. Suppose  $f(X)$  is disconnected. Then there are two nonempty open disjoint sets  $\mathcal{U}$  and  $\mathcal{V}$  such that  $f(X) = \mathcal{U} \cap \mathcal{V}$ . But then their preimage is open, and thus  $X = f^{-1}(\mathcal{U}) \cup f^{-1}(\mathcal{V})$ , and thus  $X$  is disconnected, a contradiction. Thus  $f(X)$  is connected.  $\square$

**Definition 21.1.23** If  $(X, \tau)$  and  $(Y, \tau')$  are topological spaces, then the product topology on the set  $X \times Y$  is the set  $\mathcal{T} = \{\mathcal{U} \times \mathcal{V} : \mathcal{U} \in \tau, \mathcal{V} \in \tau'\}$ .

**Theorem 21.1.31.** *The product topology is a topology.*

*Proof.*

1. As  $\emptyset \in \tau$  and  $\emptyset \in \tau'$ ,  $\emptyset = \emptyset \times \emptyset \in \mathcal{T}$ .
2. If  $\mathcal{U}_\alpha \in \mathcal{T}$ , then  $\cup_\alpha \mathcal{U}_\alpha = \cup_\alpha (\mathcal{U}_\alpha, \mathcal{V}_\alpha)$ . As  $\tau$  and  $\tau'$  are topologies,  $\cup_\alpha \mathcal{U}_\alpha \in \tau$  and  $\cup_\alpha \mathcal{V}_\alpha \in \tau'$ . Thus,  $\cup_\alpha \mathcal{U}_\alpha \in \mathcal{T}$ .
3.  $\cap_{k=1}^n \mathcal{U}_k = \cap_{k=1}^n (\mathcal{U}_k, \mathcal{V}_k)$ . As  $\tau$  and  $\tau'$  are topologies,  $\cap_{k=1}^n \mathcal{U}_k \in \tau$  and  $\cap_{k=1}^n \mathcal{V}_k \in \tau'$ . Thus  $\cap_{k=1}^n \mathcal{U}_k \in \mathcal{T}$

□

**Definition 21.1.24** The projection map  $\pi_1$  is defined as  $\pi_1 : X_1 \times X_2 \rightarrow X_1$  by  $(x_1, x_2) \mapsto x_1$ . Similarly for  $\pi_2$ .

**Theorem 21.1.32.** *The projection map is continuous.*

*Proof.* Let  $\pi_1 : X_1 \times X_2 \rightarrow X_1$  be the projection map,  $X \times Y$  having the project topology. Let  $\mathcal{U} \subset_{Open} X_1$ . Then  $f^{-1}(\mathcal{U}) = \{(x_1, x_2) : x_1 \in \mathcal{U}, x_2 \in X_2\}$ . But  $\mathcal{U}$  and  $X_2$  are open, and thus  $f^{-1}(\mathcal{U})$  is open (In the product topology). □

**Definition 21.1.25** An open mapping is a function  $f : X \rightarrow Y$  such that  $\mathcal{U} \subset_{Open} X \Rightarrow f(\mathcal{U}) \subset_{Open} Y$ .

**Theorem 21.1.33.** *The projection map is an open mapping.*

*Proof.* For let  $\mathcal{U}$  be an open set in  $X \times Y$  (With the product topology). That is, there are open sets  $\mathcal{U} \subset X$  and  $\mathcal{V} \subset Y$  such that  $\mathcal{U} = \{(x, y) : x \in \mathcal{U}, y \in \mathcal{V}\}$ . Then  $\pi_1(\mathcal{U}) = \mathcal{U}$ , which is open. Therefore, etc. □

**Theorem 21.1.34.** *If  $X$  and  $Y$  are compact, then  $X \times Y$  is compact with the product topology.*

*Proof.* For let  $\mathcal{O}$  be an open cover of  $X \times Y$ . Then  $\{\pi_X(\mathcal{U}) : \mathcal{U} \in \mathcal{O}\}$  is an open cover of  $X$  and  $\{\pi_Y(\mathcal{V}) : \mathcal{V} \in \mathcal{O}\}$  is an open cover of  $Y$ . As  $X$  and  $Y$  are compact, there exist finite subcovers of each, say  $\mathcal{O}_X$  and  $\mathcal{O}_Y$ . But then  $\{\pi_X^{-1}(\mathcal{U}) : \mathcal{U} \in \mathcal{O}_X\} \cup \{\pi_Y^{-1}(\mathcal{V}) : \mathcal{V} \in \mathcal{O}_Y\}$  is a finite subcover of  $\mathcal{O}$ . Thus,  $X \times Y$  is compact. □

**Theorem 21.1.35.** *If  $X, Y \subset Z$  are compact,  $X \cup Y$  is compact.*

*Proof.* Let  $\mathcal{O}$  be an open cover of  $X \cup Y$ . Then there is a finite subcover of  $X$  and a finite subcover of  $Y$ , and thus the combination of these subcovers is a cover of  $X \cup Y$ . □

## 21.2 Old Notes

An accumulation point of a set  $A$  is a point  $x$  such that, for all open neighborhoods  $U$  of  $A$ ,  $U \cap A \neq \emptyset$ .

**Theorem 21.2.1** (Bolzano-Weierstrass Theorem). *If  $X$  is a bounded, infinite subset of  $\mathbb{R}$ , then  $X$  has at least one accumulation point.*

**Theorem 21.2.2.** *The intersection of an arbitrary collection of closed sets is closed. The union of finitely many closed sets is closed.*

*Proof.* Apply De Morgan's theorem to the properties of a topological space.  $\square$

There's also something called the derived set of a set  $A$ .

**Theorem 21.2.3.** *If  $A$  is a set, then the interior of  $A$  is equal to  $(\overline{A^C})^C$*

**Definition 21.2.1** The neighborhood system of a point  $x$  in a topological space  $(X, \tau)$  is the set of all neighborhoods of  $x$ .

Limits of sequences in topological spaces are NOT necessarily unique. This is different from convergence in  $\mathbb{R}$ , where convergence is always unique.

**Definition 21.2.2** The relative topology of a topological space  $(X, \tau)$  with respect to a subset  $A \subset X$  is  $\tau_A = \{A \cap U : U \in \tau\}$

**Theorem 21.2.4.** *If  $(X, \tau)$  is a topological space and  $A \subset X$ , then  $(A, \tau_A)$  is a topological space.*

$(A, \tau_A)$  is called a subspace of  $(X, \tau)$ .

**Definition 21.2.3** A basis of a topological space  $(X, \tau)$  is a subset  $B$  of  $\tau$  such that every element of  $\tau$  is the union of some of the elements of  $B$ .

**Theorem 21.2.5.** *A subset  $B \subset \tau$  is a basis for  $\tau$  if and only if for all  $U \in \tau$  and all  $x \in U$ , there is a  $V \in B$  such that  $x \in V \subset U$ .*

**Theorem 21.2.6.** *If  $B$  is a basis of  $\tau$ , then  $U$  is open if and only if for all  $x \in U$  there is a  $V \in B$  such that  $x \in V \subset U$ .*

**Theorem 21.2.7.**  $\mathbb{R}$  has a countable basis.

*Proof.* For the set of open intervals  $(p, q)$ , where  $p$  and  $q$  are rational numbers, forms a basis for the standard topology on  $\mathbb{R}$ . Moreover, this is countable.  $\square$

**Definition 21.2.4** If  $(X, \tau)$  is a topological space and  $S \subset \tau$ , then  $S$  is a subbase if a finite intersection of elements of  $S$  forms a base of  $\tau$ .

**Definition 21.2.5** A local base for a point  $x$  in a topological space  $(X, \tau)$  is a set of open neighborhoods  $B_x$  of  $x$  such that for all open neighborhoods  $G$  of  $x$ , there is a  $G_x \in B_x$  such that  $x \in G_x \subset G$ .

**Theorem 21.2.8.** *If  $(X, \tau)$  is a topological space,  $x \in X$ , and if  $B$  is a base for  $\tau$ , then the set of elements  $G_x$  in  $B$  such that  $x \in G_x$  is a local base for  $x$ .*



# CHAPTER 22

---

## The Product Topology

---

### 22.1 Product Topology

It is common in the literature of mathematics to drop this ordered pair notation, and simply call  $X$  a topological space. To prevent confusion, we will distinguish between the two:  $X$  is a set,  $(X, \tau)$  is a topological space.

The notion of a topological space is a generalization of that of a *metric space*. We discard all properties of metric spaces, with the exception of the fact that open sets are closed under arbitrary unions and finite intersections. As such, we call the elements of a topology  $\tau$  on a set  $X$  the *open subsets* of  $X$ . We wish to talk about the *product space* formed by the Cartesian product of two sets and their respective topologies. We'll need to define the *generated topology*, so we prove the following:

**Theorem 22.1.1.** *If  $X$  is a set, and if  $T$  is a set of topologies on  $X$ , then:*

$$\tau = \bigcap_{t \in T} t \tag{22.1.1}$$

*Is a topology on  $X$ .*

*Proof.* First note that, since for all  $t \in T$ ,  $t$  is a topology, it is true that  $\emptyset \in t$ , and thus  $\emptyset$  is contained in the intersection. Therefore  $\emptyset \in \tau$ . Similarly,  $X \in \tau$ . Given a subset  $\mathcal{O} \subseteq \tau$ , it is true that  $\mathcal{O} \subseteq t$  for all  $t \in T$ . But for all  $t \in T$ ,  $t$  is a topology on  $X$ , and therefore the union of the elements of  $\mathcal{O}$  are contained in  $t$ , and thus this union is contained in  $\tau$ . Thus,  $\tau$  is closed to arbitrary unions. Similarly for finite intersections. Thus,  $\tau$  is a topology on  $X$ .  $\square$

---

### Definition 22.1.1: Generated Topology

---

The topology generated by a subset  $S \subseteq \mathcal{P}(X)$  is the set:

$$\tau = \bigcap \{\tau_S : \tau_S \text{ is a topology on } X \text{ and } S \subseteq \tau_S\} \quad (22.1.2)$$

That is, the smallest topology such that the elements of  $S$  are open. ■

By Thm. 22.1.1 we see that the topology generated by some collection of subsets of  $X$  is indeed a topology on  $X$ . This notion is similar to the one found when one studies measure theory. For arbitrary topologies it is often difficult, perhaps even impossible, to describe explicitly the elements of the topology. Analogously, consider the Borel  $\sigma$ -Algebra on  $\mathbb{R}$ . We describe this as the  $\sigma$ -Algebra generated by semi-intervals  $[a, b]$ . An explicit description of the elements of the Borel  $\sigma$ -Algebra is almost certainly impossible. With this, we can move to product spaces.

---

### Definition 22.1.2: Product of Two Topologies

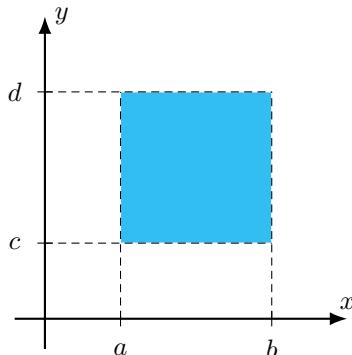
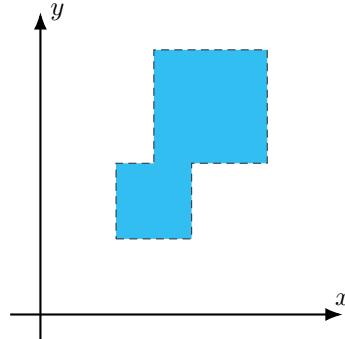
---

The product of two topological spaces  $(X, \tau_X)$  and  $(Y, \tau_Y)$  is the topological space  $(X \times Y, \tau)$ , where  $X \times Y$  is the Cartesian product of  $X$  and  $Y$ , and where  $\tau$  is the topology generated by the sets:

$$\mathcal{O} = \{\mathcal{U} \times \mathcal{V} : \mathcal{U} \in \tau_X, \mathcal{V} \in \tau_Y\} \quad (22.1.3)$$

■

It is important to note that we cannot simply set the topology  $\tau$  to be the set of all sets of the form  $\mathcal{O} = \{\mathcal{U} \times \mathcal{V}\}$ , where  $\mathcal{U} \in \tau_X$  and  $\mathcal{V} \in \tau_Y$ , for this will most likely **not** be a topology. The reason being that it may fail to be closed to unions.

22.1.1: The Open Rectangle  $(a, b) \times (c, d)$ .22.1.2: A Region That Cannot be Written as  $\mathcal{U} \times \mathcal{V}$ .Fig. 22.1: Examples of Open Subsets of  $\mathbb{R}^2$ .

For consider  $\mathbb{R}$ . The standard topology on  $\mathbb{R}^2$  is constructed by considering the collection of all open *rectangles*,  $(a, b) \times (c, d)$ . However, the set of open rectangles will not, by itself, be a topology on  $\mathbb{R}^2$ . For one, the union of two rectangles may not even be connected: Consider two disjoint non-empty open rectangles. This union will **not** be a rectangle. But even if two open rectangles are not disjoint, their union may not be a rectangle. See Fig. 22.1 for examples. As a final example, consider the open unit disc in  $\mathbb{R}^2$ . This is the set:

$$D^2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\} \quad (22.1.4)$$

This is not of the form  $\mathcal{U} \times \mathcal{V}$  for some pair of sets  $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}$ . However, seeing as we've called it the open unit disc, we would certainly like it to be open. And indeed it is, for it lies in the topology that is *generated* by open rectangles.

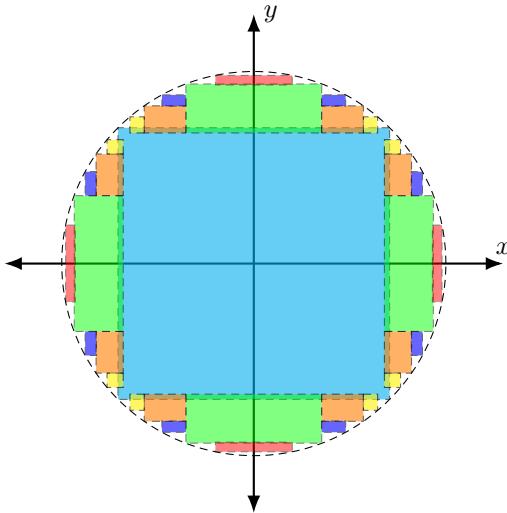


Fig. 22.2: Tiling of the Open Unit Disc by Rectangles.

Note that, in the tiling of the unit disc shown in Fig. 22.2, many of the rectangles overlap. This is to avoid excluding any points within the circle, and to give a clear picture. Such a tiling is allowed in the topology generated by open rectangles, since *arbitrary* unions are allowed. With this figure we have some evidence that the topology generated by open rectangles is most likely the same as the standard topology on  $\mathbb{R}^2$ . That is: The set of all sets  $\mathcal{U} \subseteq \mathbb{R}^2$  such that, for all  $\mathbf{x} \in \mathcal{U}$ , there is an  $r > 0$  such that, for all  $\mathbf{y} \in \mathbb{R}^2$  such that  $\|\mathbf{x} - \mathbf{y}\|_2 < r$ , it is true that  $\mathbf{y} \in \mathcal{U}$ . Here  $\|\cdot\|_2$  denotes the standard Euclidean norm, where we compute length by invoking the Pythagorean formula. Rather than carrying out complicated computations, we can simply note that open balls in the  $\|\cdot\|_2$  norm are of the form:

$$B_r^{(\mathbb{R}^2, \|\cdot\|_2)}(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^2 : \|\mathbf{x} - \mathbf{y}\|_2 < r\} \quad (22.1.5)$$

These are circles centered at  $\mathbf{x}$ . Contrast that with open balls in the  $\|\cdot\|_\infty$  metric:

$$B_r^{(\mathbb{R}^2, \|\cdot\|_\infty)}(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^2 : \max\{|x_1 - y_1|, |x_2 - y_2|\} < r\} \quad (22.1.6)$$

These are just squares centered at  $\mathbf{x}$ . And we know that the  $\|\cdot\|_2$  and  $\|\cdot\|_\infty$  metrics are equivalent, so these topologies must be the same. Thus we've found a slightly more inconvenient way of describing the topology on  $\mathbb{R}^2$ . The plus side is that this alternative notion generalizes to  $X \times Y$  when  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are more general topological spaces.

In defining the product topology of two topological spaces, we used the familiar notion of a Cartesian product. Elements of the Cartesian product  $X \times Y$  are ordered pairs  $(x, y)$ , where  $x \in X$  and  $y \in Y$ . We can continue to ordered triples  $(x, y, z)$  and the general  $n$  tuple  $(x_1, \dots, x_n)$  and similarly define the product topology of  $n$  topological spaces  $(X_1, \tau_1), \dots, (X_n, \tau_n)$ . But what if we wanted to define an *infinite* product of infinitely many spaces? If the product is countable, we have some intuition for we can think of *infinite* tuples  $(x_1, \dots, x_n, \dots)$ , but this lacks clarity. Rather, let's go back to the product of two topological spaces and redefine it. Let  $\mathbb{Z}_2 = \{1, 2\}$  and define:

$$\prod_{i=1}^2 X_i = \{f : \mathbb{Z}_2 \rightarrow \bigcup_{k=1}^2 X_k : f(1) \in X_1, f(2) \in X_2\} \quad (22.1.7)$$

That is, the set of all functions from  $\mathbb{Z}_2$  into  $X_1 \cup X_2$  with the property that 1 maps into  $X_1$  and 2 maps into  $X_2$ . There is a clear bijection between this new thing and  $X_1 \times X_2$ , simply map  $(x, y)$  to the function  $f$ , where  $f(1) = x$  and  $f(2) = y$ . But now we have a definition that really didn't depend on how many products we were making. Let  $\mathbb{Z}_n = \{1, \dots, n\}$ , and let  $X_1, \dots, X_n$  be sets. We can then define:

$$\prod_{i \in \mathbb{Z}_n} X_i = \{f : \mathbb{Z}_n \rightarrow \bigcup_{k \in \mathbb{Z}_n} X_i : \forall i \in \mathbb{Z}_n, f(i) \in X_i\} \quad (22.1.8)$$

And we can go further, defining the product for any collection of sets. Let's first introduce some notation. An indexing set for a collection of sets is some set  $I$  such that we can write all of the sets in our collection as  $X_i$ , for  $i \in I$ . To improve rigor, let's say that an indexing set for a collection of sets  $\mathcal{O}$  is some set  $I$  such that there is a surjective function  $X : I \rightarrow \mathcal{O}$ , and let's write  $X(i) = X_i$ , for all  $i \in I$ . That is, for all  $i \in I$ ,  $X_i$  is a set in  $\mathcal{O}$ . We can now define the general product of sets.

### Definition 22.1.3: Product of Sets

---

The product of a collection of sets indexed by a set  $I$  is the set:

$$\prod_{i \in I} X_i = \{f : I \rightarrow \bigcup_{i \in I} X_i : \forall i \in I, f(i) \in X_i\} \quad (22.1.9)$$



This notion is well-defined for arbitrary products, countable or not. It is important to note that the elements of the product space are *functions*.

**Example 22.1.1: N**

thing in the definition of an indexing set requires  $\mathcal{O}$  to contain many sets, so let  $\mathcal{O} = \{\mathbb{R}\}$  and let  $I = \mathbb{N}$ . Then the product is simply:

$$\prod_{n \in \mathbb{N}} \mathbb{R} = \{a : \mathbb{N} \rightarrow \mathbb{R}\} \quad (22.1.10)$$

That is, the set of all sequences of real numbers. Thus, the countable product of  $\mathbb{R}$  can be thought of in two ways: The set of all *infinite* tuples  $(x_1, \dots, x_n, \dots)$ , or the set of all *sequences* of real numbers. ■

All of this has been purely set theoretic: There is no topology yet. Given a collection of topological spaces  $(X_i, \tau_i)$ , is there a good topology to place on the product? That is, can we form a nice product topological space? There are two well established ways to do this: The *obvious* way, and the *correct* way. We first let intuition lead us astray, and define the obvious answer: The Box Topology.

**Definition 22.1.4: Box Topology**

The box topology on a collection of topological spaces  $(X_i, \tau_i)$  indexed by a set  $I$  is the topology  $\tau$  on the set  $X$  where:

$$X = \prod_{i \in I} X_i \quad (22.1.11)$$

And where  $\tau$  is the topology generated by the sets:

$$\mathcal{U} = \left\{ \prod_{i \in I} \mathcal{U}_i : \mathcal{U}_i \in \tau_i \right\} \quad (22.1.12)$$

That is,  $\tau$  is generated by all of the open sets in all of the  $X_i$ . ■

This is precisely what we did for  $\mathbb{R}^2$ . We took the topology to be the one generated by all of the open rectangles in the plane. Unfortunately, when the product is infinite, the box topology is horrible. Some problems with the box topology:

1. The product of compact spaces need not be compact.
2. The product of connected spaces need not be connected.
3. The product of metric spaces need not be metrizable.

Moreover, some functions that *look* continuous, and that we would obviously want to be continuous, are not. For example, let  $X$  be the set of sequences in

$\mathbb{R}$ , and define  $f : \mathbb{R} \rightarrow X$  by mapping  $x$  to the sequence  $a_n = x$ ,  $n \in \mathbb{N}$ . That is:

$$f(x) = x, x, x, x, \dots, x, x, \dots \quad (22.1.13)$$

This function is *nowhere* continuous in the box topology. So now we devise a plan to make a *better* topology with the following property: Suppose  $g : \mathbb{R} \rightarrow X$ , where  $X$  is again the space of real-valued sequences, and suppose  $g$  is of the form:

$$g(x) = g_1(x), g_2(x), \dots, g_n(x), \dots \quad (22.1.14)$$

Where  $g_k$  is continuous for all  $k \in \mathbb{N}$ . We **require** that  $g$  be continuous in the product space. We could simply make  $\tau$  be the chaotic topology,  $\tau = \{\emptyset, X\}$ , but then *every* function  $f : A \rightarrow X$  is continuous, for *any* topological space  $(A, \tau_A)$ , and this is rather boring. So we try another approach. Given a collection of topological spaces  $(X_i, \tau_i)$ , we require that the product space  $(X, \tau)$  is such that all of the projection mappings  $p_i : X \rightarrow X_i$  are continuous. The projection mappings can be defined set theoretically using the notation we've developed. Given a product  $X$  of sets  $X_i$ , the projection mapping  $p_i : X \rightarrow X_i$  is simply:

$$p_i(x) = x(i) \quad (22.1.15)$$

This looks strange, but remember that we've defined the product space to be a set of *functions*, and therefore  $x \in X$  is a function. Thus, the  $i^{\text{th}}$  projection mapping simply evaluates these functions in the  $i^{\text{th}}$  coordinate.

In the search for a topology on the product set  $X$  that makes all of the projection mappings  $p_i$  continuous, we could simply take  $\tau = \mathcal{P}(X)$ . Then, for *any* topological space  $(A, \tau_A)$ , and for *every* function  $f : X \rightarrow A$ ,  $f$  is continuous. This is overkill and we see that this is larger than the box topology. So all of the problems with the box topology still exist! So, we require that  $\tau$  is the *smallest* such topology. We now define the initial topology.

### Definition 22.1.5: Initial Topology

The initial topology on a set  $X$  generated by a set of functions  $f_i$  from  $X$  to topological spaces  $(X_i, \tau_i)$  is the set:

$$\tau = \bigcap \left\{ \tau_X : \tau_X \text{ is a topology on } X \text{ and } \forall_{i \in I}, f_i \text{ is continuous.} \right\} \quad (22.1.16)$$

■

This collection is non-empty, since  $\mathcal{P}(X)$  is contained in it, and by Thm. 22.1.1,  $\tau$  is a topology on  $X$ . We now define the product topology.

---

**Definition 22.1.6: Product Topology**


---

the product topology on a set  $X$  defined as the product of topological spaces  $(X_i, \tau_i)$  indexed over  $I$ :

$$X = \prod_{i \in I} X_i \quad (22.1.17)$$

Is the initial topology defined by the set:

$$\mathcal{F} = \{p_i : X \rightarrow X_i, p_i(x) = x(i)\} \quad (22.1.18)$$

That is, the set of projection mappings. ■

In this construction one might have noted that the projection mappings are continuous in the box topology. Thus one might very reasonably ask if the product topology and the box topology are the same thing. And indeed, for a *finite* product, they are! This makes sense, for in  $\mathbb{R}^2$  we can think of the topology generated by rectangles, or the topology generated by the projection mappings, and they are the same. What's crucial is that they differ for infinite products.

---

**Theorem 22.1.2: Product Topology Basis Theorem**


---

$f(X, \tau)$  is the product topological space formed by the topological spaces  $(X_i, \tau_i)$ , indexed by a set  $I$ , then:

$$\tau = \bigcup \prod_{i \in I} \{\mathcal{U}_i \in \tau_i : \mathcal{U}_i = X_i \text{ for all but finitely many sets.}\} \quad (22.1.19)$$

That is,  $\tau$  is the set of all products of open sets  $\mathcal{U}_i$ , such that all but finitely many of the  $\mathcal{U}_i$  are the entire space  $X_i$ . ■

This seems confusing, so we illustrate with some pictures. What's important to note is that, if the product is infinite, then the box topology and the product topology differ. To see this, in the box topology we allowed *all* products of open sets, whereas now we only allow the product of open sets  $\mathcal{U}_i$  where, for all but finitely many  $i$ , we have  $\mathcal{U}_i = X_i$ . It should then be clear that, if  $\tau_B$  is the box topology, and  $\tau_P$  is the product topology, then  $\tau_P \subseteq \tau_B$ , and for infinite products  $\tau_P$  is a proper subset.

Let's dumb down the theorem a bit, and imagine again a world where  $X = Y = \mathbb{R}$ . Let's consider the topology generated by sets  $\mathcal{U} \times \mathcal{V}$ , where  $\mathcal{U}$  is an open subset of  $\mathbb{R}$ , and  $\mathcal{V} = \mathbb{R}$ . That is, rather than allowing the product to be over finitely many arbitrary open sets, we allow it to be over one, and it must be in the  $x$  axis. In doing this we can get a sense of what the product topology

might look like.

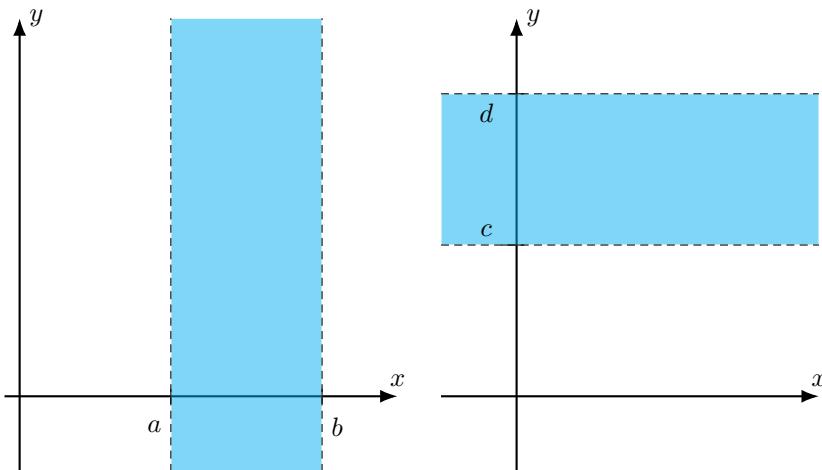


Fig. 22.3: Strips in the Plane.

We can do the same thing and consider sets of the form  $\mathbb{R} \times \mathcal{V}$ , where  $\mathcal{V}$  is open in  $\mathbb{R}$ . Recall that open sets in  $\mathbb{R}$  are intervals and arbitrary collections of intervals. Using this, we see that the topology generated by  $\mathcal{U} \times \mathbb{R}$  is the collection of all open vertical *strips*, and  $\mathbb{R} \times \mathcal{V}$  form the horizontal strips. See Fig. 22.3. We expand this game to  $\mathbb{R}^3$ , and think of sets of the form  $\mathcal{U} \times \mathcal{V} \times \mathbb{R}$ , or any permutation of the three coordinates. See Fig. 22.4

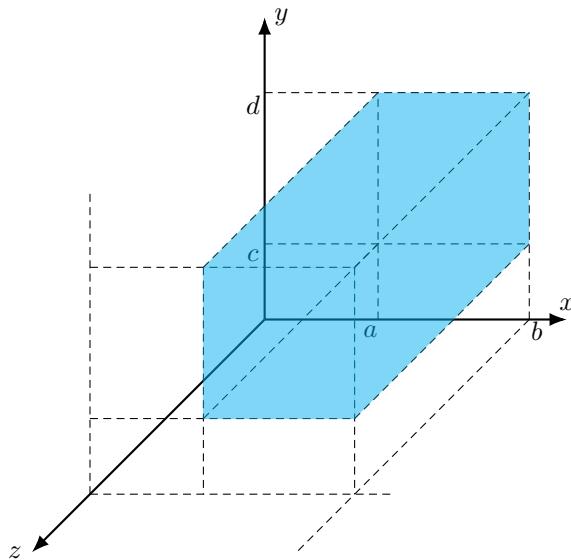


Fig. 22.4: Blocks in Space.

The product topology has the following wonderful features:

1. The product of compact topological spaces is compact.
2. The product of connected spaces is connected.
3. The product of metric spaces is metrizable.

## 22.2 Hocking and Young (Chapter 1)

**Example 22.2.1** Let  $\mathbb{R}$  be the real numbers with the usual topology and consider the subset  $A \subseteq \mathbb{R}$  defined by:

$$A = \left\{ \frac{1}{n} \mid n \in \mathbb{N} \right\} \quad (22.2.1)$$

$A$  has only one limit point and this is zero. Note that  $0 \notin A$ , and hence the derived set  $\text{Der}_\tau(A)$  is disjoint from  $A$  itself. That is:

$$\text{Der}(A) = \{0\} \quad (22.2.2)$$

And thence  $\text{Der}(A) \cap A = \emptyset$ .

**Example 22.2.2** Again consider  $\mathbb{R}$  and define  $B$  by:

$$B = \left\{ \frac{n+1}{n} + (-1)^n \frac{n-1}{n} \mid n \in \mathbb{N} \right\} \quad (22.2.3)$$

This oscillates between values that are asymptotically approaching zero and one. The derived set is then:

$$\text{Der}(B) = \{0, 1\} \quad (22.2.4)$$

Once again, the derived set is disjoint from the original set.

**Example 22.2.3** In the usual topology, the derived set of  $\mathbb{Q}$  is the entirety of  $\mathbb{R}$ . That is:

$$\text{Der}(\mathbb{Q}) = \mathbb{R} \quad (22.2.5)$$

So it is possible for the derived set to be strictly larger than the original set.

**Example 22.2.4** If  $(X, \tau)$  is the chaotic topological space on some set  $X$  with the trivial topology  $\tau = \{\emptyset, X\}$ , then for any subset  $A \subseteq X$ , it is true that  $\text{Der}_\tau(A) = X$ . That is, in the chaotic topology every point is a limit point of every other point.

**Example 22.2.5** If  $(X, \tau)$  is the discrete topology,  $\tau = \mathcal{P}(X)$ , then for any subset  $A \subseteq X$ , the derived set is empty:  $\text{Der}_\tau(A) = \emptyset$ . That is, no point is a limit point of any subset. This is because every element in the discrete topology is isolated since  $\{x\}$  is an open subset for every  $x \in X$ .

**Theorem 22.2.1.** *If  $(X, \tau)$  is a topological space, if  $A, B \subseteq X$ , and if  $A \subseteq B$ , then  $\text{Der}_\tau(A) \subseteq \text{Der}_\tau(B)$ , where  $\text{Der}_\tau(Y)$  denotes the derived set of  $Y$  in  $\tau$ .*

**Theorem 22.2.2.** *If  $(X, \tau)$  is a topological space, if  $C \subseteq X$ , then  $C$  is closed if and only if  $\text{Cl}_\tau(C) = C$ .*

**Theorem 22.2.3.** *If  $(X, \tau)$  is a topological space, if  $\mathcal{C} \subseteq \mathcal{P}(X)$  is such that for all  $C \in \mathcal{C}$  it is true that  $C$  is closed, then  $\bigcap \mathcal{C}$  is closed.*

**Theorem 22.2.4.** *If  $(X, \tau)$  is a topological space, if  $C, D \subseteq X$  are closed in  $X$ , then  $C \cup D$  is closed in  $X$ .*

**Definition 22.2.1** A basis for a topology  $\tau$  on a set  $X$  is a subset  $\mathcal{B} \subseteq \tau$  such that  $\bigcup \mathcal{B} = X$  and such that for all  $B_1, B_2 \in \mathcal{B}$  and for all  $p \in B_1 \cap B_2$ , there exists  $B_3 \in \mathcal{B}$  such that  $p \in B_3$  and  $B_3 \subseteq B_1 \cap B_2$ .

**Theorem 22.2.5.** *If  $(X, \tau)$  is a topological space, then  $\tau$  is a basis for  $\tau$ .*

*Proof.* For  $X \in \tau$ , and hence  $X \subseteq \bigcup \tau$ , but since  $\tau \subseteq \mathcal{P}(X)$ , for all  $\mathcal{U} \in \tau$  it is true that  $\mathcal{U} \subseteq X$ , and hence  $X = \bigcup \tau$ . If  $\mathcal{U}_1, \mathcal{U}_2$ , and if  $x \in \mathcal{U}_1 \cap \mathcal{U}_2$ , then  $\mathcal{U}_1 \cap \mathcal{U}_2 \in \tau$  is such that  $x \in \mathcal{U}_1 \cap \mathcal{U}_2$  and  $\mathcal{U}_1 \cap \mathcal{U}_2 \subseteq \mathcal{U}_1 \cap \mathcal{U}_2$ , and hence  $\tau$  is a basis for  $\tau$ .  $\square$

**Theorem 22.2.6.** *If  $(X, \tau)$  is a topological space, if  $\mathcal{B}$  is a basis for  $\tau$ , then:*

$$\tau = \left\{ \mathcal{U} \in \mathcal{P}(X) \mid \text{There exists } \mathcal{O} \subseteq \mathcal{B} \text{ such that } \mathcal{U} = \bigcup \mathcal{O} \right\} \quad (22.2.6)$$

**Theorem 22.2.7.** If  $X$  is a set, if  $\tau_1$  and  $\tau_2$  are topologies on  $X$ , if  $\tau_1 \subseteq \tau_2$ , if  $\mathcal{B}_1$  is a basis for  $\tau_1$ , if  $\mathcal{B}_2$  is a basis for  $\tau_2$ , if  $x \in X$ , and if  $B_1 \in \mathcal{B}_1$  is such that  $x \in B_1$ , there exists a  $B_2 \in \mathcal{B}_2$  such that  $x \in B_2$  and  $B_2 \subseteq B_1$ .

**Example 22.2.6** The previous theorem shows when two bases for a topology  $\tau$  are the same, but different bases may give rise to different topologies. Consider on  $\mathbb{R}$  the standard topology  $\tau_{\mathbb{R}}$  and the topology  $\tau$  generated by the set:

$$\mathcal{B} = \{ (x, \infty) \mid x \in \mathbb{R} \} \quad (22.2.7)$$

We can see that  $\tau \subseteq \tau_{\mathbb{R}}$  since every set in  $\mathcal{B}$  can be obtained as the union of intervals as follows:

$$(x, \infty) = \bigcup_{n \in \mathbb{N}} (x, x + n) \quad (22.2.8)$$

Since  $(x, x + n) \in \tau_{\mathbb{R}}$  for all  $n \in \mathbb{N}$  and for all  $x \in \mathbb{R}$ , and since topologies are closed under arbitrary unions, we thus have that every element of  $\mathcal{B}$  is contained in  $\tau_{\mathbb{R}}$  and thus every element of  $\tau$  is also contained in  $\tau_{\mathbb{R}}$ . However, these are not the same topology.

**Example 22.2.7** In  $\mathbb{R}^2$ , the set of all vertical open line segments forms the basis of a topology that is strictly finer than the standard Euclidean topology  $\tau_{\mathbb{R}^2}$ , and indeed this can be seen as the product topology on  $\mathbb{R}$  with the standard topology multiplied by  $\mathbb{R}$  with the discrete topology. The resulting space, being the product of metric spaces, is again a metric space. Since the discrete topology  $\tau_D$  is strictly finer than the standard topology on  $\mathbb{R}$ , we see that the product topology  $\tau_{\mathbb{R}} \times \tau_D$  is strictly finer than the standard one  $\tau_{\mathbb{R}^2}$ .

Def second countable. Intersection of topologies is topology. Topology generated by a collection.

**Theorem 22.2.8.** If  $X$  is a set, if  $\mathcal{O} \subseteq \mathcal{P}(X)$ , if  $\tau(\mathcal{O})$  is the topology generated by  $\mathcal{O}$ , and if  $\mathcal{B}$  is the set:

$$\mathcal{B} = \left\{ \mathcal{U} \in \mathcal{P}(X) \mid \exists_{n \in \mathbb{N}} \exists_{B: \mathbb{Z}_n \rightarrow \mathcal{O}} (\mathcal{U} = \bigcap B_k) \right\} \quad (22.2.9)$$

Then  $\mathcal{B}$  is a basis for  $\tau(\mathcal{O})$ .

That is, the collection of all finite intersections of elements of  $\mathcal{O}$  forms a basis for the topology that  $\mathcal{O}$  generates.

**Definition 22.2.2** A subbasis for a topological space  $(X, \tau)$  is a subset  $\mathcal{B} \subseteq \mathcal{P}(X)$  such that the topology generated by  $\mathcal{B}$  is equal to  $\tau$ .

**Example 22.2.8** Consider the set of all open rays on  $\mathbb{R}$ . That is, the set of all subsets of  $\mathbb{R}$  of the form:

$$\mathcal{U}_+ = (x, \infty) \quad (22.2.10a)$$

$$\mathcal{U}_- = (-\infty, x) \quad (22.2.10b)$$

Then this forms a subbasis for the standard topology on  $\mathbb{R}$ . To see this we simply need to show that the standard basis for  $\mathbb{R}$  can be obtained from finite intersections of elements of  $\mathcal{B}$ , and also note that  $\mathcal{B} \subseteq \tau_{\mathbb{R}}$ . Given an open interval  $(a, b) \subseteq \mathbb{R}$ , let  $B_1 = (a, \infty)$  and  $B_2 = (-\infty, b)$ . Then both  $B_1, B_2 \in \mathcal{B}$ , but  $(a, b) = B_1 \cap B_2$ . Thus the topology generated by the subbasis  $\mathcal{B}$  is the same as the topology generated by the standard basis of open intervals, and this is the standard topology on  $\mathbb{R}$ .

### Theorem 22.2.9: Birkhoff's Topology Lattice Theorem

---

If  $X$  is a set, if  $T$  is the set of all topologies of  $X$ , and if  $\subseteq$  is the inclusion relation, then  $(T, \subseteq)$  is a complete lattice. ■

**Theorem 22.2.10.** *If  $X$  is a set, if  $\tau_1, \tau_2$  are topologies on  $X$ , if  $\tau_1 \subseteq \tau_2$ , and if  $\text{id}_X : (X, \tau_1) \rightarrow (X, \tau_2)$  is the identity mapping, then it open.*

**Theorem 22.2.11.** *If  $X$  is a set, if  $\tau_1, \tau_2$  are topologies on  $X$ , if  $\tau_2 \subseteq \tau_1$ , and if  $\text{id}_X : (X, \tau_1) \rightarrow (X, \tau_2)$  is the identity mapping, then it continuous.*

**Theorem 22.2.12.** *If  $(Y, \tau_Y)$  is a topological space, if  $X$  is a set, if  $\tau_1, \tau_2$  are topologies on  $X$ , if  $\tau_1 \subseteq \tau_2$ , and if  $f : X \rightarrow Y$  is continuous with respect to  $\tau_1$ , then it is continuous with respect to  $\tau_2$ .*

**Theorem 22.2.13.** *If  $(X, \tau_X)$  is a topological space, if  $Y$  is a set, if  $\tau_1, \tau_2$  are topologies on  $Y$ , if  $\tau_1 \subseteq \tau_2$ , and if  $f : X \rightarrow Y$  is continuous with respect to  $\tau_2$ , then it is continuous with respect to  $\tau_1$ .*

Similar result for open maps. Def metric space, metrics, open balls.

**Example 22.2.9** Metrics are not topological entities, and two vastly different metrics on the same set may give the same topology. For example, given any metric  $d$  on a set  $X$ , the metric  $\tilde{d}$  formed by:

$$\tilde{d}(x, y) = \frac{d(x, y)}{1 + d(x, y)} \quad (22.2.11)$$

is bounded, yet forms the same topology. Hence boundedness is not a topological property but a metric one. Thus, there are questions about metric spaces that cannot be answered from the study of metrizable spaces. For example, if  $(X, d)$  is a metric space one can ask if it has the midpoint property: For all  $x, y \in X$  is there a  $z \in X$  such that  $d(x, z) = d(y, z) = d(x, y)/2$ ? This cannot be answer topologically, for consider the closed unit interval  $[0, 1] \subseteq \mathbb{R}$  and the closed upper half unit circle in  $\mathbb{R}^2$ . That is:

$$S = \{(x, y) \in \mathbb{S}_1 \mid y \geq 0\} \quad (22.2.12)$$

Then  $[0, 1]$  and  $S$  are homeomorphic under the function  $f : [0, 1] \rightarrow S$  defined by:

$$f(x) = (\cos(\pi x), \sin(\pi x)) \quad (22.2.13)$$

But  $[0, 1]$  does have the midpoint property, whereas the upper semi-circle does not. That is, closed intervals are convex where circles (the boundaries of discs) are not. To topologies this question we might ask if there exists a metric with the midpoint property or if every metric has it. That is, can a given topological space  $(X, \tau)$  be given a metric  $d$  with the midpoint property? Are there topological spaces  $(X, \tau)$  where every metric on  $X$  that induces  $\tau$  must have the midpoint property?

Def separable.

**Theorem 22.2.14.** *If  $(X, \tau)$  is a separable and metrizable topological space, then it is second countable.*

*Proof.* For if  $(X, \tau)$  is separable, there exists a countable dense subset  $A$ , and if  $(X, \tau)$  is metrizable, there exists a metric  $d$  on  $X$  that induces the topology  $\tau$ . Define  $\mathcal{B}$  by:

$$\mathcal{B} = \{B_q^{(X,d)}(x) \mid x \in A \text{ and } q \in \mathbb{Q}^+\} \quad (22.2.14)$$

□

**Example 22.2.10** A common *non-theorem* that confuses students is the belief that separable and first countable imply second countable, and this is not true. What we've shown is that first countable and separable imply second countable if we're working with metric spaces. For example, consider the particular point topology on  $\mathbb{R}$ . That is, a set  $\mathcal{U}$  is open if and only if it is either empty or contains the origin. Then  $(\mathbb{R}, \tau)$  is first countable. To see this, let  $x \in \mathbb{R}$  and consider:

$$\mathcal{B}_x = \{\{x, 0\}\} \quad (22.2.15)$$

this is a neighborhood basis for  $x$ , and hence the space is first countable. Moreover it is separable since  $\text{Cl}_\tau(\{0\}) = \mathbb{R}$ . To see that it is not second countable we'll show that it is not  $\sigma$  locally finite. For suppose  $\mathcal{O}$  is a locally finite collection. That 0 is contained in finitely many elements of  $\mathcal{O}$ , and thus  $\mathcal{O}$  must itself be finite. But then for any countable collection of locally finite sets we have that this can't be a basis since  $\mathbb{R}$  is uncountable. Hence  $(\mathbb{R}, \tau)$  is not  $\sigma$  locally finite, and therefore it is not second countable.

**Example 22.2.11** As another example, let  $H$  be the closed upper half plane in  $\mathbb{R}^2$ . That is, all point  $(x, y) \in \mathbb{R}^2$  such that  $y \geq 0$ . Consider the following basis: open balls in the interior of the upper half plane combined with open balls that lie tangent to the  $x$  axis together with the tangential point  $(x, 0)$ . Then the intersection of  $\mathbb{Q}^2$  with the upper half plane still forms a countable dense subset, but any basis needs to contain at least one set for every  $(x, 0)$ . Since  $\mathbb{R}$  is uncountable, this space cannot possibly be second countable.

Def cont func.

**Theorem 22.2.15.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, and if  $f : X \rightarrow Y$  is a continuous function, then for all  $x \in X$  and for all  $\mathcal{V} \in \tau_Y$  such that  $f(x) \in \mathcal{V}$ , there exists a  $\mathcal{U} \in \tau_X$  such that  $f[\mathcal{U}] \subseteq \mathcal{V}$ .*

$\varepsilon - \delta$  def of cont for metric space.

**Example 22.2.12** There are bijective continuous function  $f : X \rightarrow Y$  that are not homeomorphisms. Let  $\mathbb{S}^1$  be the unit circle and  $[0, 1)$  be the semi-open interval with the usual inherited metric topology from  $\mathbb{R}$ . Define  $f : [0, 1) \rightarrow \mathbb{S}^1$  by:

$$f(x) = (\cos(2\pi x), \sin(2\pi x)) \quad (22.2.16)$$

This is simply wrapping the semi-open interval up into a circle. Since the left endpoint 0 is included,  $f$  is bijective. It's also continuous, however  $f^{-1}$  is not continuous at the point  $(1, 0) \in \mathbb{S}^1$ . The function  $f^{-1}$  effectively tears  $\mathbb{S}^1$  at the point, and thus it is not continuous. Rigorously, we see that  $\mathbb{S}^1$  is compact and  $[0, 1)$  is not (Both are true by the Heine-Borel theorem), and so  $f^{-1}$  can't be continuous since continuous functions preserve compactness.

Def connected (open def, closed def, clopen def).

**Theorem 22.2.16.** *If  $(\mathbb{R}, \tau)$  is the standard topological space on  $\mathbb{R}$ , then it is connected.*

**Theorem 22.2.17.** *If  $(X, \tau)$  is a topological space, then it is disconnected if and only if there exists non-empty subsets  $A, B \subseteq X$  such that:*

$$(\text{Cl}_\tau(A) \cap B) \cup (A \cap \text{Cl}_\tau(B)) = \emptyset \quad (22.2.17)$$

**Theorem 22.2.18.** *If  $(X, \tau)$  is a topological space, if  $\mathcal{C} \subseteq \mathcal{P}(X)$  is a collection of connected subsets of  $X$ , and if  $\bigcap \mathcal{C} \neq \emptyset$ , then  $\bigcup \mathcal{C}$  is connected.*

This theorem provides a rather crude way of showing that  $\mathbb{R}^n$  is connected. Consider  $\mathbb{R}^n$  as the union of all lines through the origin. Since we know that  $\mathbb{R}$  is connected, and since all of these lines have non-empty intersection (they intersect at the origin), their union must again be connected. But this union is simply the entirety of  $\mathbb{R}^n$ , and so Euclidean space is connected.

**Theorem 22.2.19.** *If  $n \in \mathbb{N}$ ,  $n > 1$ , if  $(\mathbb{R}^n, \tau)$  is the standard Euclidean topological space, and if  $\mathbf{x} \in \mathbb{R}^n$ , then  $\mathbb{R}^n \setminus \{\mathbf{x}\}$  is connected.*

*Proof.* For it is path connected, and hence connected (connect two dots).  $\square$

**Theorem 22.2.20.** *If  $n \in \mathbb{N}$ , if  $n > 0$ , and if  $(\mathbb{S}^n, \tau)$  is the usual subspace topology on the sphere, then it is connected.*

*Proof.* For  $\mathbb{R}^n n + 1 \setminus \{\mathbf{0}\}$  is connected by the previous theorem, and the function  $f : \mathbb{R}^n n + 1 \setminus \{\mathbf{0}\} \rightarrow \mathbb{S}^n$  defined by:

$$f(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad (22.2.18)$$

is surjective and continuous, and hence it's image is connected. But since it is surjective, it's image is the entire sphere, and thus  $\mathbb{S}^n$  is connected.  $\square$

IVP. Comp of cont is cont. Intervals are connected.

**Theorem 22.2.21.** *If  $n \in \mathbb{N}$ ,  $n > 1$ , and if  $(\mathbb{R}^n n + 1, \tau)$  is the standard Euclidean topological space, then  $\mathbb{R}^n n + 1 \setminus \mathbb{S}^n$  has two open connected components.*

*Proof.* For the open ball  $\mathbb{B}^n n + 1$  is path connected, and  $\mathbb{R}^n n + 1 \setminus \text{Cl}_\tau(\mathbb{B}^n n + 1)$  is path connected.  $\square$

Removing a hyperplane leaves two open connected components.

**Theorem 22.2.22.** *If  $(\mathbb{T}^n, \tau)$  is the  $n$  torus with the product topology, then it is connected.*

*Proof.* For the product of connected is connected, and  $\mathbb{S}^1$  is connected.  $\square$

Def cover, open cover, compactness, compact subset (compact in subspace topology).

**Theorem 22.2.23.** *if  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then  $A$  is compact if and only if for every subset  $\mathcal{O} \subseteq \tau$  such that  $\mathcal{O}$  is a cover of  $A$ , there is a finite subcover  $\Delta \subseteq \mathcal{O}$ .*

*Proof.* For if  $A$  is a compact subset, then  $(A, \tau_A)$  is a compact space where  $\tau_A$  is the subspace topology. But then:

$$\mathcal{O}_A = \{ A \cap \mathcal{U} \mid \mathcal{U} \in \mathcal{O} \} \quad (22.2.19)$$

is an open cover of  $A$  in the subspace topology. But  $A$  is compact, and thus there is a finite subcover. Yadda yadda.  $\square$

### Definition 22.2.3: Finite Intersection Property

A topological space with the finite intersection property is a topological space  $(X, \tau)$  such that for any set  $\mathcal{C} \subseteq \mathcal{P}(X)$  of closed sets such that for any finite subsets  $\Delta \subseteq \mathcal{C}$  it is true that  $\bigcap \Delta \neq \emptyset$ , it is also true that  $\bigcap \mathcal{C} \neq \emptyset$ .

**Theorem 22.2.24.** *If  $(X, \tau)$  is a topological space, then it is compact if and only if it has the finite intersection property.*

*Proof.* For suppose  $(X, \tau)$  is compact and does not have the finite intersection property. Then there exists a sets  $\mathcal{C} \subseteq \mathcal{P}(X)$  such that  $\bigcap \mathcal{C} = \emptyset$ , yet for every finite subset  $\Delta \subseteq \mathcal{C}$  it is true that  $\bigcap \Delta \neq \emptyset$ . But if  $\bigcap \mathcal{C} = \emptyset$ , then for all  $x \in X$  there exists  $C \in \mathcal{C}$  such that  $x \notin C$ . But then the set  $\mathcal{O}$  defined by:

$$\mathcal{O} = \{ \mathcal{U} \in \tau \mid \exists_{C \in \mathcal{C}} (\mathcal{U} = X \setminus C) \} \quad (22.2.20)$$

is an open cover of  $X$ . But  $X$  is compact, and thus there is a finite subcover  $\Lambda \subseteq \mathcal{O}$ . But then the set  $\Delta \subseteq \mathcal{C}$  defined by the complements of  $\Lambda$  is finite and has empty intersection, a contradiction. Thus,  $X$  has the finite intersection property. Next, suppose  $X$  has the finite intersection property but is not compact. Then there is an open cover  $\mathcal{O}$  of  $X$  with no finite subcover. But then the complements  $\mathcal{C}$  are a collection of closed sets such that  $\bigcap \mathcal{C} = \emptyset$ . But  $X$  has the finite intersection property, and thus there exists a finite subset  $\Delta \subseteq \mathcal{C}$  such that  $\bigcap \Delta = \emptyset$ . But then the set of complements of  $\Delta$  is a finite subcover of  $\mathcal{O}$ , a contradiction. Thus,  $X$  is compact.  $\square$

#### Definition 22.2.4: Limit Point Compact

A limit point compact topological space is a topological space  $(X, \tau)$  such that for every infinite subset  $A \subseteq X$ , the derived set  $\text{Der}_\tau(A)$  is non-empty.

Limit point compact is also often called *weakly countably compact*. We've adopted the name limit point compact since this seems common in analysis, and the great theorems about compactness in metric spaces (for example, the Bolzano-Weierstrass theorem or the generalized Heine-Borel theorem) are often stated in terms of sequences or limit points. Moreover, some authors choose to make no distinction between limit point compact and countably compact, which we shall describe in a moment. The reason being that in an accessible space (a  $T_1$  topological space), limit point compact and countably compact are equivalent. Since most spaces that are studied in the wild are Hausdorff, they are automatically accessible, and hence limit point compact and countably compact are usually the same thing. As the name suggests, countably compact is a weakening of compact.

**Definition 22.2.5: Countably Compact**

A countably compact topological space is a topological space  $(X, \tau)$  such that for every countable open cover  $\mathcal{O}$  of  $X$  there exists a finite subcover.

**Theorem 22.2.25.** *If  $(X, \tau)$  is a compact topological space, then it is countably compact.*

*Proof.* For if  $\mathcal{O}$  is a countable open cover of  $X$ , then it is an open cover of  $X$ , and since  $X$  is compact there exists a finite subcover. Hence,  $(X, \tau)$  is countably compact.  $\square$

In metric spaces, this result reverses, which is quite astounding. This is tied into the fact that in metric spaces, sequentially compact and compact are one in the same. In general, a countably compact space is compact if and only if it is Lindelöf. This can be seen quite easily since given a cover, the Lindelöf property can reduce this down to a countable subcover, and countable compactness then extracts a finite subcover. The reverse direction is true since compact implies both Lindelöf and countably compact.

**Theorem 22.2.26.** *If  $(X, \tau)$  is a topological space, then it is compact if and only if it is countably compact and Lindelöf.*

*Proof.* For if  $\mathcal{O}$  is an open cover, then since  $X$  is Lindelöf there exists a countable open subcover  $\Delta$ . But  $X$  is countably compact, and thus if  $\Delta$  is a countable open cover, then there exists a finite subcover  $\Lambda$ . Hence,  $X$  is compact.  $\square$

Countably compact always implies limit point compact as well.

**Theorem 22.2.27.** *If  $(X, \tau)$  is a countably compact topological space, then it is limit point compact.*

*Proof.* For suppose not and let  $A \subseteq X$  be an infinite set with no limit point. Then since  $A$  is infinite, there exists a countable subset  $N \subseteq A$ . But then there is a bijection  $a : \mathbb{N} \rightarrow N$ . But if  $A$  has no limit point and  $A \subseteq N$ , then  $N$  has no limit point. But then for all  $n \in \mathbb{N}$  there is a  $\mathcal{U}_n$  such that  $N \cap \mathcal{U}_n = \{a_n\}$ . But then for all  $x \in X$  there exists an open subset  $\mathcal{U}_x$  such that  $\mathcal{U}_x \cap N$  is finite. Let  $\mathcal{O} = \{\mathcal{U}_x\}$ . Let  $\mathcal{V}_n$  be the union of the  $\mathcal{U}_x$  that contain  $a_n$ . This is a countable cover, and since  $X$  is countably compact there is a finite subcover. From this we conclude that  $N$  is finite, a contradiction. Thus,  $(X, \tau)$  is limit point compact.  $\square$

**Theorem 22.2.28.** *If  $(X, \tau)$  is a compact topological space, then it is limit point compact.*

*Proof.* For compact spaces are countably compact, and countably compact spaces are limit point compact.  $\square$

**Theorem 22.2.29.** *If  $(X, \tau)$  is an accessible  $(T_1)$  limit point compact topological space, then it is countably compact.*

The fact that limit point compact is very weak in a general topological space can be seen by studying the extreme value. If  $(X, \tau_X)$  is compact and  $(Y, \tau_Y)$  is an order topology, then for any continuous function  $f : X \rightarrow Y$  the range must be bounded. This still holds for countably compact sets, but fails in the case of limit point compact.

**Example 22.2.13** Let  $(\mathbb{Z}_2, \tau)$  be the trivial topological space on 2 points, and let  $[\mathcal{P}(\cdot)]\mathbb{Z}$  be the discrete topology. The product topology on  $\mathbb{Z}_2 \times \mathbb{Z}$  is then limit point compact and the projection mapping  $\pi_2 : \mathbb{Z}_2 \times \mathbb{Z} \rightarrow \mathbb{Z}$  is continuous. But the discrete topology on  $\mathbb{Z}$  is the same as its order topology, but  $\pi_Z$  is not bounded. Indeed, the range of  $\pi_Z$  is the entirety of  $\mathbb{Z}$ , and hence the range isn't even limit point compact.

If we consider accessible spaces  $(T_1)$ , all of our problems disappear. The product topology of a trivial space (with at least two points) with a discrete space will not be accessible, and therein lies the issue.

Compact subset of Hausdorff is closed, countable finite intersection property. Product space,  $\mathbb{R}^n$ , torus  $\mathbb{T}^2$  as the product of  $\mathbb{S}^1$  with itself. Draw pictures. Box vs. Product topology. Axiom of Choice, Volterra set  $[x - y \in \mathbb{Q}]$ .

**Theorem 22.2.30.** *Assuming axiom of choice, if  $S$  is a non-empty set of disjoint non-empty sets, then there is a set  $A$  such that for all  $B \in S$ ,  $A \cap B$  contains one point.*

Partial order, total order, well order, well ordering theorem. Hausdorff maximality theorem.

### Definition 22.2.6: Compact Subsets

---

A compact subset of a topological space  $X$  is a set  $A \subseteq X$  such that for every open cover  $\mathcal{O}$  of  $A$ , there is a finite subcover  $\Delta \subseteq \mathcal{O}$ .  $\blacksquare$

### Definition 22.2.7: Disconnected Sets

---

A disconnected subset of a topological space  $X$  is a set  $S \subseteq X$  such that there exist disjoint non-empty open sets  $X_1, X_2$  such that  $S = X_1 \cup X_2$ .  $\blacksquare$

**Definition 22.2.8: Connected Sets**

A connected subset of a topological space is a subset that is not disconnected.

■

**Theorem 22.2.31.** *If  $X$  is compact and  $S \subset X$  is closed, then  $S$  is compact.*

*Proof.* For let  $\mathcal{O}$  be an open cover of  $S$ . Then since  $S$  is closed,  $S^C$  is open. But then  $\mathcal{O} \cup \{S^C\}$  is an open cover of  $X$ . But  $X$  is compact and therefore there is an open subcover  $\Delta \subseteq \mathcal{O} \cup \{S^C\}$ . But then  $\Delta \setminus \{S^C\}$  is an finite subcover of  $S$ . □

**Theorem 22.2.32.** *If  $a, b \in \mathbb{R}$  and  $a < b$ , then  $[a, b]$  is compact.*

*Proof.* For suppose not. Then there is an open cover  $\mathcal{O}$  of  $[a, b]$  with no finite subcover. Let  $A$  be the set  $A = \{r \in \mathbb{R} : [a, r] \text{ has a finite subcover}\}$ . As  $\mathcal{O}$  is an open cover, there is an open subset  $\mathcal{U}_1 \in \mathcal{O}$  such that  $a \in \mathcal{U}_1$ . Therefore  $A$  is not empty. Moreover, since  $[a, b]$  is not compact, for all  $r \in A$ ,  $r < b$ . Therefore  $A$  is bounded above. By the least upper bound property there is a  $\gamma \in \mathbb{R}$  such that for all  $r \in A$ ,  $r \leq \gamma$ . But, as  $\mathcal{U}_1$  is open and  $a \in \mathcal{U}_1$ ,  $a < \gamma \leq b$ . But then  $\gamma \in [a, b]$ , and thus there is a  $\mathcal{U}_2 \in \mathcal{O}$  such that  $\gamma \in \mathcal{U}_2$ . But as  $\mathcal{U}_2$  is open, there is an  $\varepsilon > 0$  such that  $(\gamma - \varepsilon, \gamma + \varepsilon) \subset \mathcal{U}_2$ . But then  $[a, \gamma + \varepsilon/2]$  has a finite subcover, a contradiction since  $\gamma$  is the least upper bound of  $A$ . □

**Theorem 22.2.33.** *If  $A$  and  $B$  are compact, then  $A \times B$  is compact.*

*Proof.* For let  $\mathcal{O}$  be an open cover of  $A \times B$ . Then  $\{\pi_A(\mathcal{U}) : \mathcal{U} \in \mathcal{O}\}$ , that is, the set of projections of open sets in  $\mathcal{O}$  onto  $A$ , is an open cover of  $A$ . Similarly for  $B$ . But  $A$  and  $B$  are compact, and therefore there exists finite subcovers. Taking the union of these two gives a finite subcover of  $A \times B$ . □

**Theorem 22.2.34.** *If  $A_1, \dots, A_n$  are compact, then  $A_1 \times \cdots \times A_n$  is compact.*

*Proof.* Apply induction to Thm. 22.2.33. □

The finiteness of the product in Thm. 22.2.34 is unnecessary. Tychonoff's Theorem, which is equivalent to the axiom of choice, says given an arbitrary collection of compact sets, the space formed by the product of these sets is also compact with respect to the product topology. We can now prove our main result.

**Theorem 22.2.35: Heine-Borel Theorem**

A subset  $S \subset \mathbb{R}^n$  is compact if and only if it closed and bounded.

*Proof.* Suppose  $S$  is compact and suppose it is unbounded. Then the set of open balls about the origin  $B_r(0) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| < r\}$  is an open cover of  $S$ , since it is an open cover of  $\mathbb{R}^n$ , and yet no finite subcover exists. For if one did, then there is a least  $N \in \mathbb{N}$  such that  $S \subset B_N(0)$ , a contradiction as  $S$  is unbounded. Therefore  $S$  is bounded. Furthermore, suppose  $S$  is not closed. Then there exists a point  $\mathbf{x} \in S^C$  such that, for all  $r > 0$ ,  $B_r(\mathbf{x}) \cap S \neq \emptyset$ , where  $B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < r\}$ . Let  $\overline{B}_r(\mathbf{x})$  be the closure of these sets (That is, the closed ball about  $\mathbf{x}$ ). Then the set of complements  $\overline{B}_r(\mathbf{x})^C$  is an open cover of  $S$ , for it is an open cover of  $\mathbb{R}^n \setminus \{\mathbf{x}\}$ , but no finite subcover exists, a contradiction. Thus  $S$  is closed. Therefore, if  $S$  is compact then it is closed and bounded. If  $S$  is bounded, then there is an  $r \in \mathbb{R}$  such that  $S \subset [-r, r]^n$ . But  $[-r, r]^n$  is the product of compact sets, and is therefore compact. But  $S$  is closed, and closed subsets of compact spaces are compact. Therefore  $S$  is compact.  $\square$

### Theorem 22.2.36: Homeomorphisms Preserve Compactness

If  $X$  and  $Y$  are homeomorphic, and if  $X$  is compact, then  $Y$  is compact.  $\blacksquare$

*Proof.* For if  $X$  and  $Y$  are homeomorphic, then there is a continuous bijection  $f : X \rightarrow Y$ . Let  $\mathcal{O}$  be an open cover of  $Y$ . Then  $\{f^{-1}(\mathcal{U}) : \mathcal{U} \in \mathcal{O}\}$  is an open cover of  $X$ . But  $X$  is compact, and therefore there is a finite subcover  $\Delta$ . But, since  $f$  is surjective,  $\{\mathcal{U} \in \mathcal{O} : f^{-1}(\mathcal{U}) \in \Delta\}$  is a finite subcover of  $Y$ .  $\square$

### Theorem 22.2.37: Homeomorphisms Preserve Connectedness

If  $X$  and  $Y$  are homeomorphic and  $X$  is connected, then  $Y$  is connected.  $\blacksquare$

*Proof.* Suppose not. If  $Y$  is disconnected, then there are disjoint non-empty open sets  $Y_1, Y_2$  such that  $Y = Y_1 \cup Y_2$ . But as  $X$  and  $Y$  are homeomorphic, there is a continuous bijection  $f : X \rightarrow Y$ . But then  $f^{-1}(Y_1)$  and  $f^{-1}(Y_2)$  are non-empty, as  $f$  is a bijection, and moreover they are disjoint open subsets of  $X$ , as  $f$  is continuous. But then  $X$  is disconnected, a contradiction. Thus,  $Y$  is connected.  $\square$

## 22.3 A Review of Topology

Fig. ?? shows how both  $S^2$  with three points removed and  $T^2$  with one point removed are homotopy equivalent. Recall that  $\mathbb{R}^2 \setminus \{(0, 0)\}$  is homotopy equivalent to  $S^1$ . In a similar manner, the plane with two points removed is homotopy equivalent to two circles whose intersection contains a single points (That is, a figure-8). While the “Proof,” given was hand wavy, the fact that the sphere is

not homeomorphic to the torus comes from the fact that these two objects have different boundary components, something preserved by homeomorphism. As we saw before, we can remove a circle from the torus, leaving one connected surface, but removing a circle from the sphere results in two connected components. “What about compact manifolds without boundary?”

### Theorem 22.3.1: Generalized Poincare Conjecture

---

If  $X$  is an  $n$  dimensional manifold that is homotopy equivalent to  $S^n$ , then  $X$  is homeomorphic to  $S^n$ . █

Recall that the boundary of a topological space is what remains when you remove its interior. That is:

$$\partial X = \overline{X} \setminus \text{Int}(X) \quad (22.3.1)$$

Where  $\overline{X}$  is the closure of  $X$ . A topological space without boundary is one such that  $\partial X = \emptyset$ . With this we can now define closed and rigid manifolds.

### Definition 22.3.1: Closed Manifolds

---

A closed manifold of dimension  $n \in \mathbb{N}$  is a compact topological manifold  $M$  of dimension  $n$  without boundary. █

### Definition 22.3.2: Closed Rigid Manifolds

---

A closed rigid manifold of dimension  $n$  is a closed topological manifold  $M$  such that, for all closed homotopy equivalent manifolds  $N$  of dimension  $n$ ,  $M$  is homeomorphic to  $N$ . █

The question then becomes “Which manifolds are rigid, and which are not?” From the Poincare theorem,  $S^n$  is topologically rigid for all  $n \in \mathbb{N}$ .  $S^n$  is not differentially rigid. That is, we cannot necessarily relax the definition of rigidity to include diffeomorphisms. The first example of a non-rigid closed manifold came in the 1930’s from Franz, Reidemeister, and de Rham, and is called a Lens Space. Let  $p$  and  $q$  be coprime positive integers. Divide  $S^3$  into  $p$  equal parts, and then divide this into its northern and southern hemispheres. Take a piece of the northern hemisphere and move it over  $q$  slices, and then glue this to the southern hemisphere. Take the piece that is already there and move it over  $q$  pieces, and then glue that to the northern hemisphere. Repeat this process until all slices are done. This is called the Lens Space  $L(p, q)$ .  $L(1, 1)$  is simply the 3-sphere,  $L(2, 1)$  is the real projective space  $\mathbb{RP}^3$ . See Fig. 22.5 to see how this construction occurs.

**Theorem 22.3.2.** *If  $p, q_1, q_2 \in \mathbb{N}$ , then  $L(p, q_1)$  and  $L(p, q_2)$  are homotopy*

equivalent if and only if there is an  $n \in \mathbb{N}$  such that:

$$q_1 q_2 = n^2 \quad (22.3.2)$$

**Theorem 22.3.3.** If  $p, q_1, q_2 \in \mathbb{N}$ , then  $L(p, q_1)$  and  $L(p, q_2)$  are homotopy equivalent if and only if  $q_1 = q_2$ .

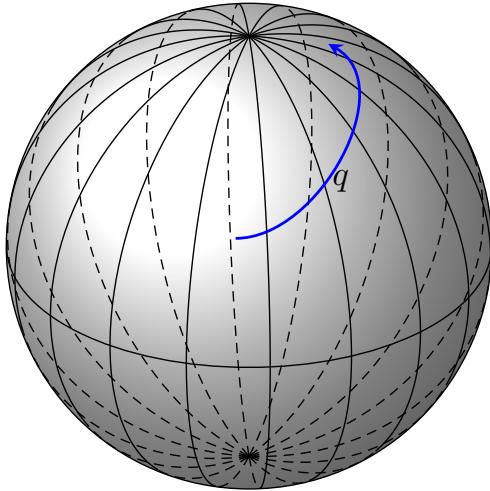


Fig. 22.5: How to construct  $L(p, q)$ .

We move on to the structure set of topological spaces, in particular closed topological manifolds  $\mathcal{M}$ .

**Definition 22.3.3** Equivalent homotopies are homotopy equivalences  $f_1 : X_1 \rightarrow Y$ ,  $f_2 : X_2 \rightarrow Y$ , denoted  $f_1 \sim f_2$ , such that there exists a continuous function  $g : X_1 \rightarrow X_2$  and  $f_2 \circ g \simeq f_1$ .

The equivalent classes of  $Y$  is called the structure set of  $Y$ , denoted  $\mathcal{S}(Y)$ . This set contains maps like  $f_1$ ,  $f_2$ . If  $g$  is a homeomorphism, then  $f_1 = f_2$ .

### Example 22.3.1

From the generalized Poincare Conjecture:

$$\mathcal{S}^{\text{Top}}(S^n) = \{S^n\} \quad (22.3.3)$$

From before, if  $p, q \in \mathbb{N}$  are co-prime, then:

$$\text{curl}(\mathcal{S}^{\text{Top}}(L(p, q))) > 1 \quad (22.3.4)$$

Moreover, from the definition of rigid manifolds, if  $\text{curl}(\mathcal{S}(X)) > 1$ , then  $X$  is

non-rigid. If  $T^n$  denotes the  $n$  torus, then:

$$\text{curl}(\mathcal{S}^{\text{Top}}(T^n)) = 2^n \quad (22.3.5)$$

However, not all surgery structure sets are finite. ■

A few questions naturally arise from the definition of the structure set. Is there a natural group structure that can be placed on  $\mathcal{S}^{\text{Top}}(X)$ ? Is it possible for  $\mathcal{S}^{\text{Top}}(X)$  to be infinite? By studying the real projective spaces  $\mathbb{RP}^n$ , we arrive at our first example of a space whose surgery structure set is infinite.

$$n \bmod 4 = 3 \implies \text{curl}(\mathcal{S}^{\text{Top}}(\mathbb{RP}^n)) = \aleph_0 \quad (22.3.6)$$

A review of some concepts from algebraic topology.

#### Definition 22.3.4: Paths

A path in a topological space  $X$  is a continuous function  $f : I \rightarrow X$  ■

By further requiring that  $f(0) = f(1)$ , we call  $f$  a loop. This is equivalent to the following:

#### Definition 22.3.5: Loops

A loop in a topological space  $X$  a continuous function  $f : S^1 \rightarrow X$ . ■

#### Definition 22.3.6: Fundamental Group

The fundamental group of a topological space  $X$  with base point  $p$  is the set:

$$\pi_1(X, p) = \{f \in C(I, X) : p = f(0) = f(1)\}/h \quad (22.3.7)$$

where  $h$  is the modulo of homotopy, equipped with the concatenation operation:

$$(f * g)(t) = \begin{cases} f(2t), & 0 \leq t < \frac{1}{2} \\ g(2t - 1), & \frac{1}{2} \leq t < 1 \end{cases} \quad (22.3.8)$$
■

#### Theorem 22.3.4: Homeomorphisms Preserve the Fundamental Group

If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $f : X \rightarrow Y$  is a homeomorphism, if  $p_X \in X$  and if  $p_Y = f(p_X)$ , then  $\pi_1(X, p_X)$  is isomorphic to  $\pi_1(Y, p_Y)$ . ■

*Proof.* If  $X$  and  $Y$  are homeomorphic, then there is a continuous bijective function  $f : X \rightarrow Y$  such that  $f^{-1}$  is continuous. Define  $\phi : \pi_1(X) \rightarrow \pi_1(Y)$

by:

$$\phi(x) = f \circ x \quad (22.3.9)$$

But  $f$  and  $x$  are continuous and the composition of continuous functions is continuous, and therefore  $\phi(x)$  is continuous. Moreover:

$$\phi(x(0)) = (f \circ x)(0) = f(p_X) = p_Y \quad (22.3.10)$$

$$\phi(x(1)) = (f \circ x)(1) = f(p_X) = p_Y \quad (22.3.11)$$

Therefore  $\phi(x) \in \pi_1(X, p_Y)$ . But if  $x_1, x_2 \in \pi_1(X)$ , then:

$$\phi(x_1(t) * x_2(t)) = \phi(x_1(t)) * \phi(x_2(t)) \quad (22.3.12)$$

Thus  $\phi$  is a homomorphism. But as  $f$  is a bijection, so is  $\phi$ , and therefore  $\phi$  is an isomorphism. Thus,  $\pi_1(X, p_X)$  and  $\pi_1(Y, p_Y)$  are isomorphic.  $\square$

**Theorem 22.3.5.** *If  $X$  and  $Y$  are topological spaces, and if  $\pi_1(X)$  and  $\pi_1(Y)$  are not isomorphic, then  $X$  and  $Y$  are not homeomorphic.*

Using this theorem we can tell whether or not certain spaces are homeomorphic. That is, the fundamental group is a *topological invariant*.

### Definition 22.3.7: Order of an Element of a Group

---

The order of an element  $g$  in a group  $G$  with neutral element  $e$  is:

$$\text{Ord}_G(g) = \inf\{n \in \mathbb{N} : a^n = e\} \quad (22.3.13)$$

If there is no such  $n \in \mathbb{N}$ , then we write  $\text{Ord}_G(g) = \infty$ . ■

### Definition 22.3.8: Group with Torsion

---

A group with torsion is a group  $(G, *)$  such that there exists  $g \in \{G\}$  such that  $g$  is not an identity of  $(G, *)$  and  $\text{Ord}_G(g) < \infty$ . ■

### Example 22.3.2

---

If  $n \in \mathbb{N}$  and  $n > 1$ , then  $S^n$  is simply connected. Any loop in  $S^n$  can be continuously transformed down to a point, and thus all loops in  $S^n$  are homotopic. Therefore, for all  $p \in S^n$ :

$$\pi_1(S^n, p) \simeq \{e\} \quad (22.3.14)$$

The fundamental group of spheres is isomorphic to the trivial group. For the case of  $n = 1$ , this is not true. Intuitively we can see that loops are uniquely determined by how many times they wrap around the circle. While the actual

computation is difficult, for all  $p \in S^1$  we have:

$$\pi_1(S^1, p) \simeq \mathbb{Z} \quad (22.3.15)$$

Thus, for all  $n \in \mathbb{N}$ , and for all  $p \in S^n$ ,  $\pi_1(X, p)$  is not a group with torsion. There are topological spaces, and base points in the spaces, such that their fundamental groups have torsion. The first two examples are the real projective plane  $\mathbb{RP}^n$  and the Lens' Spaces  $L(p, q)$ . We have:

$$\pi_1(\mathbb{RP}^n) = \mathbb{Z}_2 \quad (22.3.16a) \quad \pi_1(L(p, q)) = \mathbb{Z}_p \quad (22.3.16b)$$

If  $n > 1$ , then  $\mathbb{Z}_n$  is a group with torsion since 1 has order  $n - 1$ . Thus, these fundamental groups are groups with torsion. ■

**Theorem 22.3.6.** *If  $n \geq 5$ ,  $n \equiv 3 \pmod{4}$ , and  $\pi_1(X)$  is a torsion group, then:*

$$\text{curl}(S(X^n)) = \infty \quad (22.3.17)$$

Some other gems:  $S(\mathbb{CP}^n) = \mathbb{Z}_2$ . Chern Manifolds are a thing.

### The Unsolvable Word Problem

**Definition 22.3.9** A presentation of a group  $G$  is a set  $H \subset G$  of generators and a set  $R$  of relations on  $H$ . This is denoted  $G = \langle H | S \rangle$ .

#### Example 22.3.3

1.  $\langle a | a^n = e \rangle$  is a the cyclic group of order  $n$  generated by  $a$ .
2.  $\langle g, h | hg = gh \rangle = \mathbb{Z}^2$
3.  $\langle g, h | g^2 = e, h^2 = e \rangle = \mathbb{Z}_2 * \mathbb{Z}_n$
4.  $\langle g, h | f^2 = e, h^2 = e, gh = h^{-1}g \rangle = D_{2n}$

The word problem on unsolvability: Given two group presentations, there is no algorithm to show that they are isomorphic.

**Definition 22.3.10** A finitely presented group is a group with a presentation  $\langle H | R \rangle$  such that  $H$  and  $R$  are finite.

**Theorem 22.3.7.** *If  $n \geq 5$  and  $G$  is finitely presented, then there is a closed  $n$  dimensional manifold  $\mathcal{M}$  such that  $\pi_1(\mathcal{M}) = G$ .*

### Exact Sequences and Surgery Exact Sequences

**Definition 22.3.11** An exact sequence  $\cdots \rightarrow G_3 \xrightarrow{f_3} G_2 \xrightarrow{f_2} G_1 \xrightarrow{f_1} G_0$  is a sequence  $f_n$  of homomorphisms and a sequence  $G_n$  of groups such that  $\text{Im}(f_{n+1}) = \ker(f_n)$

Note, the definition requires that the  $f_n$  are *homomorphisms*, not homeomorphisms. Homeomorphism is a topological notion, not an algebraic one.

**Example 22.3.4**  $O \xrightarrow{f} G \xrightarrow{g} H$ .  $\text{Im}(f) = 0 \Rightarrow \ker(g) = 0$ . So  $g$  is injective.

**Example 22.3.5**  $G \xrightarrow{f} H \xrightarrow{g} O$ ,  $\ker(g) = H \Rightarrow \text{Im}(f) = H$ . So  $f$  is surjective.

**Definition 22.3.12** A short exact sequence is an exact sequence  $0 \xrightarrow{f} G \xrightarrow{g} H \xrightarrow{h} L \xrightarrow{\ell} 0$

We have, from the previous examples, that in a short exact sequence  $f$  must be injective and  $g$  must be surjective. We now move onto surgery exact sequences (See Wall et. al). Let  $n \geq 5$ , and  $\mathcal{M}$  be a closed manifold of dimension  $n$ . Let  $\pi = \pi_1(\mathcal{M})$ . Let Cat have the following meaning:

- Top: Category of continuous maps. That is, the topological category.
- PL: Piece-Wise linear category. Maps are piece-wise linear.
- Diff: Differentiable category. Maps are diffeomorphisms.

### Example 22.3.6

- |                                    |                                       |
|------------------------------------|---------------------------------------|
| 1. $S^{Top}(S^n) = \{S^n\}$        | 4. $S^{PL}(T^n) = \{S^n\}$ - Rigid    |
| 2. $S^{PL}(S^n) = \{S^n\}$         | 5. $ S^{PL}(T^n)  = 2^n$ - Non-Rigid. |
| 3. $ S^{Diff}(S^2)  = 28$ (Milnor) | 6. $S^{Diff}(T^n)$ - Difficult.       |

A surgery exact sequence is a sequence of the form:

$$\begin{aligned} S^{Cat}(M \times S') &\rightarrow [M \times S', G/Cat] \rightarrow L_{n+1}(\pi_1(\mathcal{M})) \rightarrow S^{Cat}(\mathcal{M}) \rightarrow \cdots \\ &\cdots \rightarrow [M, G/Cat] \rightarrow L_n(\pi_1(\mathcal{M})) \end{aligned}$$

Here,  $L_n(X)$  is a *Wall Group*, and  $[A, B]$  is a type of classifying space.



**Part X**

# **Metric Spaces**



**Part XI**

# **Homotopy**



# CHAPTER 23

---

## Stuff

---

### 23.1 Lecture 7-ish, Maybe

We've computed the following:

$$\pi_1(\mathbb{Z}, x_0) = \begin{cases} \mathbb{Z}, & n = 1 \\ \{e\}, & n \geq 2 \end{cases} \quad (23.1.1)$$

Where  $\{e\}$  denotes the trivial group. Let's now prove this.

**Theorem 23.1.1.** *If  $n \geq 2$ , and if  $x_0 \in S^n$ , then  $\pi_1(S^n, x_0) \simeq \{e\}$ .*

*Proof.* What we want to do is take any loop  $\gamma$ , remove a point that  $\gamma$  doesn't map to, and then use the fact that  $S^n \setminus \{x\}$  is homeomorphic to  $\mathbb{R}^n$ . Since  $\mathbb{R}^n$  is contractible, we are done. However, there exist space filling curves. So now we need to show that space filling curves can still be contracted. Pick  $x \in S^n$  such that  $x \neq x_0$ . Then there is an open ball  $B$  centered about  $x$  such that  $x_0 \notin B$ . But  $B$  is open and  $\gamma$  is continuous. Thus,  $\gamma^{-1}$  is open in  $[0, 1]$ . But since  $x_0 \notin B$ ,  $0, 1 \notin \gamma^{-1}(B)$ , and therefore the pre-image is an open subset of  $\mathbb{R}$  as well. But every open subset of  $\mathbb{R}$  is the disjoint union of open intervals. Let  $(a, b)$  be one of these intervals. But then  $f : (a, b) \rightarrow B$  is such that  $f(a), f(b) \notin B$ . But  $\partial B \simeq S^{n-1}$  for  $n \geq 2$ , and  $S^{n-1}$  is path connected. Lift the image continuously to the bounded. This works if there are finitely many such intervals  $(a, b)$ , but there could be countably infinitely many. But  $f^{-1}(x)$  is closed and is a subset of  $S^n$ , and is therefore closed and bounded and thus compact, by Heine-Borel. thus finitely many of the  $(a_i, b_i)$  cover  $f^{-1}(x)$ . Move these ones. Thus, we can remove  $x$  and complete the proof.  $\square$

We next move to Van Kampen's theorem. As an aside we must talk about free products of groups. The direct product of two groups  $G_1$  and  $G_2$  is defined by:

$$(x_1, x_2) * (y_1, y_2) = (x_1 * y_1, x_2 * y_2) \quad (23.1.2)$$

There is a universal property on such groups that goes as follows:

**Theorem 23.1.2.** *If  $G_1, G_2$ , and  $H$  are groups such that there are group homomorphisms  $\varphi_i : H \rightarrow G_i$ , then there is a unique group homomorphism  $\psi : H \rightarrow G_1 \times G_2$  such that:*

$$\psi(h) = (\varphi_1(h), \varphi_2(h)) \quad (23.1.3)$$

We want to flip this picture and define something called the *free product* of  $G_1$  and  $G_2$ . This is some group  $G_1 * G_2$  such that it contains, disjointly, all of the elements of  $G_1$  and  $G_2$ , with no relations between elements of  $G_1$  and  $G_2$ . This is the set of *reduced words*. Here, a word is something in an *alphabet*, or a set. We take as our alphabet the disjoint union of  $G_1$  and  $G_2$ . A word is a finite ordered sequence from this alphabet. This includes the empty word, which is simply the empty set. A reduced word is a word such that  $a_n$  and  $a_{n+1}$  are not from the same group, and also none of the letters  $a_n$  are the identity element in either  $G_1$  or  $G_2$ . The product is then to concatenate two reduced words, and then further reduce the concatenation. Similarly, for a set of groups  $\{G_\alpha : \alpha \in I\}$ , the free product:

$$*_\alpha G_\alpha = \{\text{Reduced Word in } \coprod_{\alpha \in I} G_\alpha\} \quad (23.1.4)$$

Again, there is a unique group that satisfies the universal property mentioned earlier. As a special example, consider  $\mathbb{Z} \times \mathbb{Z}$ . This has a Cayley group, and the Cayley actually reveals the group structure. That is, there are two generators,  $a = (0, 1)$  and  $b = (1, 0)$ , and has the relation that  $ab = ba$ . The Cayley graph of  $\mathbb{Z} * \mathbb{Z}$  shows that this is the free group on 2 generators. It is a non-trivial question to show whether or not  $F_n$  is isomorphic to  $\mathbb{Z}^n$ .

## 23.2 Van Kampen's Theorem

### Theorem 23.2.1: Van Kampen's Theorem

---

If  $X$  is path connected, if  $x_0 \in X$ , and if  $X$  is such that:

$$X = \bigcup_{\alpha \in J} A_\alpha \quad (23.2.1)$$

Where  $x_0 \in A_\alpha$  for all  $\alpha \in I$ ,  $A_\alpha$  is path-connected,  $A_a \cap A_b$  is path-connected, and  $A_a \cap A_b \cap A_c$  is path-connected, then for all  $[f] \in \pi_1(X, x_0)$ ,  $[f]$  can be factored in  $\pi_1(X, x_0)$  as the product:

$$[f] = [f_1] \cdot [f_2] \cdots [f_m] \quad (23.2.2)$$

Where  $f_j : I \rightarrow A_{\alpha_j}$  for each  $j \in \mathbb{Z}_m$ . ■

This is intuitively clear, given a loop in the space  $X$  we can write it as a concatenation of different loops contained in each of the  $A_{\alpha_j}$ . But this theorem relates topology to algebra by the use of these free groups and free products.

### 23.2.1 Examples

**Example 23.2.1** Let  $X$  be a torus with two internal circles identified. This is equivalent to two spheres attached by two lines, which is homotopic equivalent to two spheres and two loops, all joined at one point. That is,

$$X \approx S^1 \vee S^1 \vee S^2 \vee S^2 \quad (23.2.3)$$

The fundamental group of wedge sums has the following property:

$$\pi_1(\vee_\alpha X_\alpha) = \star_\alpha \pi_1(X_\alpha) \quad (23.2.4)$$

The condition is that the base point in each  $X_\alpha$  must have a neighborhood  $\mathcal{U}_\alpha \in X_\alpha$  that deformation retracts onto the base point.

Given a graph, infinite or not, what is the fundamental group of the graph? First fint a maximal spanning tree. Tree's are contractible, and thus we have:

$$\pi_1(Graph) \simeq \pi_1(graph/tree) \quad (23.2.5)$$

This last thing is the free group with generators of the edges the are not removed in the modulo. A know is a diffeomorphic copy of  $S^1$  in  $\mathbb{R}^3$ . Links are several knots put together. Moving on to finite CW complex, consider  $X$ , which is  $n \in \mathbb{N}$  dimensional, and has skeletons  $X^0, \dots, X^n$ . Then  $\pi_1(X^1)$  is a free group since  $X^1$  is a graph. What about  $\pi_1(X^2)$ ? We use Van Kampen's theorem to simplify this problem.  $X^2$  is obtained from  $X^1$  by attaching 2-cells  $e_\alpha^2$ ,  $\alpha \in J$ , where  $J$  is some index set, via attaching maps:  $\varphi_\alpha : S^1 \rightarrow X^1$ .

**Theorem 23.2.2.**  $\pi_1(X^2) = \pi_1(X^1)/N$  Where  $N$  is the normal subgroup generated by  $[\gamma_\alpha \varphi_\alpha \gamma_\alpha]$  in  $\pi_1(X^1)$ .

**Theorem 23.2.3.**  $\pi_1(X) \simeq \pi_1(X^2)$ .

So what is  $\pi_1$  of a two torus? We can write the two torus as an octagon with relations on the sides. There are four such relations, and we get that the fundamental group is the group generated by four elements with the condition that:

$$aba^{-1}b^{-1}cdc^{-1}d^{-1} = e \quad (23.2.6)$$

**Theorem 23.2.4.** If  $X$  is a CW complex, and if  $x_0$  is a vertex, or a point in the one skeleton of  $X$ , then:

$$\pi_1(X, x_0) = \pi_1(X^2, x_0) = \pi_1(X, x_0)/N \quad (23.2.7)$$

There are higher homotopy groups,  $\pi_k(X, x_0)$ . These are homotopy classes of maps  $f : S^k \rightarrow X$  such that  $f(0) = x_0$ .

**Theorem 23.2.5.** *If  $X$  and  $Y$  are CW complexes, if  $f : X \rightarrow Y$  is continuous, and if  $f$  is a homotopy equivalence, then there are group homomorphisms  $f_* : \pi_k(X, x_0) \rightarrow \pi_k(Y, y_0)$  for all  $k \in \mathbb{N}$ .*

Whitehead's theorem says the converse of this is true as well.

### Theorem 23.2.6: Whitehead's Theorem

---

he converse of the previous theorem is true. ■

**Theorem 23.2.7.** *If  $X$  is a topological space, then there exists a CW complex  $X'$  and a map  $f : X' \rightarrow X$  such that  $f_* : \pi_k(X') \simeq \pi_k(X)$ .*

This says that every topological space has a unique, up to homotopy equivalence, CW complex that approximates the space very well.

## 23.3 Covering Spaces

There's a deep link between covering spaces and Galois theory. When studying the fundamental group of a circle we came across loops  $f : I \rightarrow S^1$  such that  $f(0) = f(1)$ . We studied the map  $p(t) = \exp(2\pi it)$  which spiralled  $\mathbb{R}$  into the circle, and used this to define the winding number and calculate the fundamental group. This notion generalizes to other spaces.

### Definition 23.3.1: Covering Space

---

covering space of  $X$  is a space  $\tilde{X}$  such that, for all  $x \in X$ , there is an open neighborhood  $\mathcal{U}_x \subseteq X$  such that  $p^{-1}(\mathcal{U})$  is the disjoint union of open sets in  $\tilde{X}$  and such that the image under  $p$  of one of these sets is a homeomorphism with  $\mathcal{U}$ . ■

The first example is  $\mathbb{R}$  and  $S^1$ . We will prove that nice enough topological spaces  $X$  have a covering space  $\tilde{X}$  such that  $\tilde{X}$  is contractible. Such covers are unique up to homotopy. These are called the universal covers of the space. We will also show that there is a one-to-one correspondence between subgroups of  $\pi_1(X)$  and covering spaces. That is, there is a function  $\rho_* : \pi_1(\tilde{X}) \rightarrow \pi_1(X)$  that is injective.

## 23.4 Universal Cover

We now discuss the existence of universal covers. This notion relations to Lie groups, representation theory, the special orthogonal group  $SO(n)$ , and it

occurs in modern physics.  $SO(n)$  is not simply connected, but its universal cover, the spin group, is simply connected.

### Definition 23.4.1: Simply Connected Space

---

simply connected topological space is a path connected topological space  $(X, \tau)$  such that for any two points  $x, y \in X$  and any two paths  $\gamma_1, \gamma_2$  between  $x$  and  $y$ ,  $\gamma_1$  and  $\gamma_2$  are homotopic. ■

That is,  $\pi_0(X)$  is trivial, there is only one path component, and  $\pi_1(X, x_0)$  is also trivial for all  $x_0 \in X$ .

**Example 23.4.1**  $\mathbb{R}^n$  is simply connected for all  $n \in \mathbb{N}$ . Given two paths between the same two points, the straight line homotopy is a homotopy between such paths. For  $n \geq 2$ ,  $S^n$  is also simply connected. However,  $S^1$  is not simply connected. For  $\mathbb{RP}^2$ , we know that  $\pi_1(\mathbb{RP}^2) = \mathbb{Z}/2\mathbb{Z}$ , which is a two element group. Thus there is, up to homotopy, only one non-trivial loop in  $\mathbb{RP}^2$ . Therefore the real projective plane is not simply connected. There's a clear candidate for the universal cover of  $\mathbb{RP}^2$ , and that is the sphere  $S^2$ .

**Example 23.4.2** We know that the torus  $T^2 = S^1 \times S^1$  is not simply connected, since the fundamental group of  $T^2$  is the product of the fundamental group of  $S^1$  with itself, that is,  $\mathbb{Z}^2$ . The universal cover of  $T^2$  is  $\mathbb{R}^2$ . If we had a surface with  $n$  holes, it is not automatically clear what the universal cover is, but this turns out to be  $\mathbb{R}^2$  as well.

One question that arises is which topological spaces have a universal cover? It is true for all CW complexes, and there is the following necessary condition. If  $X$  is a topological space, and if there is a universal cover  $\tilde{X}$ , then for all  $x \in X$ , if  $\tilde{x}$  is such that  $p(\tilde{x}) = x$ , then  $\mathcal{U} \subseteq X$ , where  $x \in \mathcal{U}$ , evenly covered and  $\tilde{x} \in \tilde{\mathcal{U}}$ , which is homeomorphic to  $\mathcal{U}$ . Given any loop  $\gamma : I \rightarrow \mathcal{U}$ , where  $\gamma(0) = x$ , this lifts to a loop  $\tilde{\gamma}$  in  $\tilde{\mathcal{U}}$ . But  $\tilde{\mathcal{U}}$  is simply connected, and thus  $\tilde{\gamma}$  is null homotopic in  $\tilde{X}$ . Thus, we have the following definition:

### Definition 23.4.2: Semi-Locally Simply Connected Space

---

semi-locally connected space is a topological space  $(X, \tau)$  such that for all  $x \in X$  there is a neighborhood  $\mathcal{U}$  such that for loop in  $\mathcal{U}$  contracts in  $X$ . ■

When speaking of a local property, one usually requires that a certain property holds in every open set about a certain point. For example, locally compact spaces are topological spaces such that for every point  $x$  and every open neighborhood of  $x$  there is a compact subset  $V$  contained in this neighborhood. We have the following chain:

$$\text{CW} \implies \text{Locally Contractible} \implies \text{Locally Simply-Connected} \implies \dots \quad (23.4.1)$$

$$\dots \implies \text{Semi-Locally Simply-Connected}$$

There are two examples of spaces that look very similar, but not homeomorphic. One is the Hawaiian earrings, and the other is the countable  $\vee S^1$ . The Hawaiian earrings are not semi-locally simply connected, whereas the other set is locally contractible. Thus these two spaces are not homeomorphic. Now for a sufficient condition for  $X$  to have a universal cover.

**Theorem 23.4.1.** *If  $X$  is path connected, locally path connected, and semi-locally path connected, then there exists a simply connected cover of  $X$ .*

*Proof.* Let  $\tilde{X}$  be the set of homotopy equivalence classes of paths  $\gamma : I \rightarrow X$  such that  $\gamma(0) = x_0$ . We need a function  $p : \tilde{X} \rightarrow X$  such that  $p([\gamma]) = \gamma(1)$ . Next we need a topology on  $\tilde{X}$ . Recall that for a basis for a topology of a topological space  $Y$  is a collection  $\mathcal{B}$  of open subsets of  $Y$  with the property that for all  $y \in Y$ ,  $\mathcal{U} \subseteq Y$  is open and  $y \in \mathcal{U}$ , there is a  $V \in \mathcal{B}$  such that  $y \in V \subseteq \mathcal{U}$ . For a set  $Z$ , a collection  $\mathcal{B}$  of subsets of  $Z$  can be used as the basis of a topology on  $Z$  if for all  $U, V \in \mathcal{B}$ , and for all  $x \in U \cap V$ , there is a  $W \in \mathcal{B}$  such that  $x \in W \subseteq U \cap V$ . Define  $\mathcal{B}$  as follows:

$$\mathcal{B} = \{B \subseteq \tilde{X} : B \in \tau, B \text{ path connected and } \pi_1(B) \rightarrow \pi_1(X) \text{ is trivial}\} \quad (23.4.2)$$

$\mathcal{B}$  is a basis for the topology of  $\tilde{X}$ . For since  $X$  is semi-locally simply connected, for all  $x \in X$  there is a  $V \subseteq X$  such that  $x \in V$  and  $\pi_1(V) \rightarrow \pi_1(X)$  is trivial. But  $X$  is locally path-connected, and thus there is an open  $W \subseteq U \cap V$  such that  $x \in W$  and  $W$  is path connected. From this we need to make a basis for  $\tilde{X}$ . For  $\mathcal{U} \in \mathcal{B}$ , let  $\gamma : I \rightarrow X$  be such that  $\gamma(0) = x \in \mathcal{U}$  and  $\gamma(1) = x_0$ . Define  $\mathcal{U}_\gamma$  by:

$$\mathcal{U}_\gamma = \{[\gamma\eta] : \eta(0) = \gamma(1)\} \quad (23.4.3)$$

The set of all  $\mathcal{U}_\gamma$  define a basis for a topology on  $\tilde{X}$ .  $\square$

**Theorem 23.4.2.** *If  $X$  is a path connected CW complex, then  $X$  has a simply connected cover.*

**Theorem 23.4.3.** *If  $X$  is a manifold, then  $X$  has a simply connected cover.*

## 23.5 Group Actions

Let  $G$  be a discrete group, and  $X$ . There is a notion called a group action of  $G$  on  $X$ . There are two ways of viewing this, as a function  $p : G \times X \rightarrow X$ , or as a function  $p : G \rightarrow \text{Perm}(X)$ , where  $\text{Perm}$  denotes the set of permutations of  $X$ . A third definition is  $\rho(g) : X \rightarrow X$  for  $g \in G$ . A group action obeys the group law:

$$p(g_1, g_2) = p(g_1)p(g_2) \quad (23.5.1)$$

Given a topological space  $X$ ,  $\rho(g)$  continuous would imply that it is a homeomorphism, since it has a continuous inverse.

**Example 23.5.1**  $\mathbb{Z}^2$  acts on  $\mathbb{R}^2$ :

$$\rho : \mathbb{Z}^2 \rightarrow \text{Homeo}(\mathbb{R}^2) \quad (23.5.2)$$

This is defined by  $\rho(m, n) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  by  $(x, y) \mapsto (x+m, y+n)$ . This translates all points in  $\mathbb{R}^2$  by the vector  $(m, n)$ . It suffices to define  $\rho(g)$  on the generators  $g$  of  $G$ .  $\mathbb{Z}^2$  has the generators  $(1, 0)$  and  $(0, 1)$ . So:

$$\rho((1, 0))(x, y) = (x + 1, y) \quad (23.5.3a)$$

$$\rho((0, 1))(x, y) = (x, y + 1) \quad (23.5.3b)$$

We now need to check that  $\rho(a)\rho(b) = \rho(b)\rho(a)$ .

**Example 23.5.2** Let  $G$  be the group defined by the presentation:

$$G = \langle a, b | ab = ba^{-1} \rangle \quad (23.5.4)$$

This acts on  $\mathbb{R}^2$  by:

$$\rho((1, 0))(x, y) = (x + 1, y) \quad (23.5.5a)$$

$$\rho((0, 1))(x, y) = (-x, y + 1) \quad (23.5.5b)$$

We now need to check:

$$\rho(a)(\rho(b)(x, y)) = \rho(a)(-x, y + 1) = (-x + 1, y + 1) \quad (23.5.6)$$

And also:

$$\rho(b)(\rho(a^{-1})(x, y)) = \rho(b)(x - 1, y) = (-x + 1, y + 1) \quad (23.5.7)$$

And thus this is a valid Group action.

Given a group  $G$  and a set  $X$ , we consider the *orbit* space  $G/X$  as a set. The elements are called the orbits. Given a topology on  $X$ , we take the topology on the orbit space to be the quotient topology.

**Example 23.5.3** Taking the group action discussed earlier of  $\mathbb{Z}^2$  over  $\mathbb{R}^2$ , the orbit space  $\mathbb{R}^2/\mathbb{Z}^2$ . and this is homeomorphic to  $T^2$ . In the other example, with  $G = \langle a, b | ab = ba^{-1} \rangle$ , we have that the orbit space is  $\mathbb{R}^2/G$  is the Klein bottle.

In general, if  $\pi_1(X, x_0) = G$ , then  $G$  acts on the covering space  $p : \tilde{X} \rightarrow X$  with  $\pi_1(X) = 0$ . In particular, if  $X$  is a CW complex.

**Example 23.5.4** Let  $p : \mathbb{R} \rightarrow S^1$  be defined by  $p(t) = \exp(2\pi it)$ . We know that  $\pi_1(S^1) = \mathbb{Z}$ . The generator of  $\mathbb{Z}$  acts by mapping  $x \mapsto x + 1$ . Note that:

$$\mathbb{R}/\pi_1(S^1) \simeq \mathbb{R}/\mathbb{Z} \simeq S^1 \quad (23.5.8)$$

This is, in general, true. Consider the Klein bottle. The universal cover of the Klein bottle is  $\mathbb{R}^2$ .

**Theorem 23.5.1.** *If  $X$  is a topological space that is path connected and a CW complex, then for every subgroup  $H \subseteq \pi_1(X)$ , there exists a cover  $p : \tilde{X}_H \rightarrow X$  such that:*

$$p_*(\pi_1(\tilde{X}_H, \tilde{x}_0)) = H \quad (23.5.9)$$

For suitable choice of  $\tilde{x}_0$ .

*Proof.* Let  $\tilde{X} \rightarrow X$  be a simply connected cover. Since  $\pi_1(X, x_0) = G$  acts on  $\tilde{X}$ ,  $H$  also acts on  $\tilde{X}$ . Now the quotient space  $\tilde{X}/H$  has the desired properties.  $\square$

### 23.5.1 Lifting Criterion

Given a space  $X$  and a covering space  $\tilde{X}$  with covering  $p : \tilde{X} \rightarrow X$ , and given a space  $Y$  with a map  $f : Y \rightarrow X$ , when is there a *lift* of the map  $f$  to the covering space  $\tilde{X}$ ? it seems natural to impose certain criterion on  $Y$ , and we require that  $Y$  is path connected and locally path connected. 1.33 in Hatcher.

**Theorem 23.5.2.** *If  $p : \tilde{X} \rightarrow X$  is a covering space, if  $f : Y \rightarrow X$  is continuous, then a lift  $\tilde{f}$  exists if and only if  $f_* : \pi_1(Y, y_0) \rightarrow \pi_1(X, x_0)$  is such that  $\text{Ran}(f_*) \subseteq \text{Ran}(p_*)$ , where  $\text{Ran}$  denotes that range of the functions.*

*Proof.* If  $\tilde{f}$  exists, then  $f = \tilde{f} \circ p$ , and thus  $f_* = p_* \circ f_*$ , so the range of  $f_*$  is contained within the range of  $p_*$ . The hard part is going the other way. Now we need to construct such a lift.  $\square$

## 23.6 Homology

### 23.6.1 History

The basic root of the theory stems from the Euler number of a graph in the plane. That is, from his classic formula:

$$F - E + V = 1 \quad (\text{Planar Graphs})$$

$$F - E + V = 2 \quad (\text{Platonic Solids})$$

Similarly, if one triangulates a sphere, we again obtain  $F - E + V = 2$ . So there's nothing to do with the nature of the platonic solids here, but rather that all of these objects are homeomorphic to the unit sphere. What about the torus? There's a formula for any surface with genus  $g$ , and we obtain:

$$F - E + V = 2 - 2g \quad (23.6.1)$$

So, for a torus we get zero. Riemann furthered the theory in the 1850's when he introduced the notion of connectivity number. Betti and Riemann continue

this work around 1870 and introduced the idea of Betti numbers. This is the number of cuts of dimension  $k$  are needed to disconnect a space. For a genus two surface, we see that the first Betti number is four, since we need two cuts for each hole. Moving on, Emmy Noether added more to the theory in 1925 when she related this to groups. For example the Betti number 4 somehow relates to the group  $\mathbb{Z}^4$ , with Torsion. The first modern definition of homology groups comes from Mayer and Vietoris in 1928. Between 1930 and 1950, the theory began to develop into its current form. It now extends beyond topology and into algebra and group theory. There is also the notion of cohomology that came along with this.

## 23.6.2 Singular Homology

### Definition 23.6.1: Simplex

simplex of degree  $n$  is a subset of  $\mathbb{R}^n$  formed by  $n$  points that are affinely independent  $v_0, \dots, v_n$ , defined by the convex hull:

$$[v_0, \dots, v_n] = \left\{ \sum_{k=0}^n t_k v_k : \sum_{k=0}^n t_k = 1, t_k \geq 0 \right\} \quad (23.6.2)$$

■

The standard simplex, denote  $\Delta^n$ , is the convex hull of  $e_1, \dots, e_{n+1}$ . These are triangles, or tetrahedron's, etc. Every  $n$ -simplex with ordered vertices  $[v_0, v_1, \dots, v_n]$  is canonically homeomorphic to  $\Delta^n$ . Take  $t = (t_0, t_1, \dots, t_n)$  and map this to  $t_0 v_0 + \dots + t_n v_n$ . In particular, each of the faces of an  $n$  simplex comes with a map:

$$[v_0, \dots, v_j, \dots, v_n] \mapsto [v_0, \dots, v_{j-1}, v_{j+1}, \dots, v_n] \quad (23.6.3)$$

This is equivalent to the face mapping  $\Delta^{n-1} \mapsto \Delta^n$ .

### Simplicial Homology

Defined for  $\Delta$  complexes.

### Definition 23.6.2: $\Delta$ Complex

topological space  $X$  together with the following structure:

1. A collection of continuous maps  $\sigma : \Delta^n \rightarrow X$ .
2.  $\sigma(\Delta^n \setminus \partial\Delta^n) = e_\alpha^n$
3. The topology of  $X$  is a CW topology. That is,  $A \subseteq X$  is open if and only if  $\sigma_\alpha^{-1}(A)$  is open in  $\Delta^n$  for all  $\alpha \in J$ .

4. For  $\sigma_\alpha : \Delta^n \rightarrow X$  in the structure, so is  $\sigma_\alpha \circ F_j^n$ , where  $F_j^n$  is the face map.

■

### Definition 23.6.3: Chain Complex

sequence of a Abelian groups  $C_n$ , one for each  $n \in \mathbb{Z}$ , together with group homomorphisms  $\partial_n : C_n \rightarrow C_{n-1}$ . Such that  $\partial_{n-1} \circ \partial_n = 0$ . Equivalently,  $\text{Im}(\partial_n) \subseteq \ker(\partial_{n-1})$ .

We can also let:

$$C_{\cdot} = \bigoplus_{n \in \mathbb{Z}} C_n \quad (23.6.4)$$

With a homomorphism:

$$\partial : C_{\cdot} \rightarrow C_{\cdot} \quad (23.6.5)$$

That is a degree -1 homomorphism such that  $\partial^2 = 0$ . The homology of a chain complex is the collection of groups:

$$Z_n = \ker(\partial_n) \subseteq C_n \quad (23.6.6)$$

$$B_n = \text{Im}(\partial_{n+1}) \subseteq C_n \quad (23.6.7)$$

$$H_n = Z_n / B_n \quad (23.6.8)$$

### Simplicial Homology

Let  $X$  be a  $\Delta$  complex. That is, the disjoint union of open simplices. We want to create a chain complex:

$$\dots \longrightarrow \Delta_n(X) \xrightarrow{\partial_n} \Delta_{n-1}(X) \xrightarrow{\partial_{n-1}} \dots \xrightarrow{\partial_1} \Delta_0(X) \longrightarrow 0 \quad (23.6.9)$$

Form a free abelian group generated by  $\rho_\alpha^n$ .

$$\partial(\sigma) = \sum (-1)^j \sigma | [v_0, \dots, v_{j-1}, v_{j+1}, \dots, v_n] \quad (23.6.10)$$

The  $(-1)^j$  term gives an orientation to simplices.

**Theorem 23.6.1.**  $\partial_{n-1} \circ \partial_n = 0$ .

### 23.6.3 Singular Homology

Let  $X$  be a topological space. A singular  $n$  simplex in  $X$  is a continuous map  $\sigma : \Delta^n \rightarrow X$ .

**Theorem 23.6.2.** *If  $X$  has  $k$  path-connected components, then  $H_0(X) = \mathbb{Z}^k$ .*

**Theorem 23.6.3.** *If  $X$  is the union of  $X_\alpha$ , where  $X_\alpha$  is path connected, then:*

$$H_n(X) = \bigoplus H_n(X_\alpha) \quad (23.6.11)$$

**Theorem 23.6.4.** *If  $X$  is path connected, then  $H_0(X) = \mathbb{Z}$ .*

What is an  $n$  cycle? The abstract definition is some finite combination of singular simplices:  $\sigma : \Delta^n \rightarrow X$ , allowing repetitions.

## 23.7 More Homology

**Theorem 23.7.1.**

$$H_n(S^k) = \begin{cases} \mathbb{Z}, & n = k, n = 0 \\ 0, & \text{Otherwise} \end{cases} \quad (23.7.1)$$

*Proof.* We'll need to use the properties (Axioms) of singular homology. Let  $X = D^k$  be the closed  $k$  disk in  $\mathbb{R}^k$ . Let  $A = \partial D^k = S^{k-1}$ . Then  $X \cdot A \simeq S^k$ . For CW pairs we have a long exact sequence as well as strong excision:

$$\cdots \xrightarrow{A} H_n(S^{k-1}) \xrightarrow{X} H_n(D^k) \xrightarrow{X/A} \tilde{H}_n(S^k) \xrightarrow{} H_{n-1}(S^{k-1}) \xrightarrow{} \cdots \quad (23.7.2)$$

For  $n \geq 2$ , we have:

$$0 \longrightarrow H_n(S^k) \xrightarrow{\partial} H_{n-1}(S^{k-1}) \longrightarrow 0 \quad (23.7.3)$$

The end of the sequence is then:

$$H_1(D^k) \longrightarrow H_1(S^k) \xrightarrow{\partial} H_0(S^{k-1}) \longrightarrow H_0(D^k) \longrightarrow \tilde{H}_0(S^k) \longrightarrow 0 \quad (23.7.4)$$

But  $H_1(D^k) = 0$ , and so this reduces. When  $k \geq 2$  we also have that  $\tilde{H}_0(S^k) = 0$ . So, for the case of  $k \geq 2$  we have:

$$0 \longrightarrow H_1(S^k) \longrightarrow H_0^{k-1} \longrightarrow H_0(D^k) \longrightarrow 0 \quad (23.7.5)$$

And this reduces to:

$$0 \longrightarrow 0 \longrightarrow 0 \longrightarrow \mathbb{Z} \longrightarrow \mathbb{Z} \longrightarrow 0 \quad (23.7.6)$$

The next case is  $k = 1$ . We obtain the short exact sequence:

$$0 \longrightarrow H_1(S^1) \longrightarrow \mathbb{Z}^2 \longrightarrow \mathbb{Z} \longrightarrow 0 \quad (23.7.7)$$

From this we obtain  $H_1(S^1) = \mathbb{Z}$ . Using induction, if  $n > k \geq 1$ , then  $H_n(S^k) \simeq H_1(S^{k-n+1})$ , and this is the trivial group. If  $k = n \geq 1$ , then  $H_n(S^k) \simeq H_1(S^1) = \mathbb{Z}$ . Similarly for  $k < n$ .  $\square$

**Theorem 23.7.2: Milnor's Uniqueness Theorem**

If:

$$H_n\left(\coprod_{\alpha} X_{\alpha}\right) = \bigoplus_{\alpha} H_n(X_{\alpha}) \quad (23.7.8)$$

Then all homology theories on the category of CW complexes are equivalent.

**Theorem 23.7.3.**  $H_n^{\Delta}(X) \simeq H_n(X)$ .

$$H_n^{\Delta}(X^k, X^{k-1}) = \begin{cases} \text{Free Abelian Group Generated by n-cells,} & n = k \\ 0, & \text{Otherwise} \end{cases} \quad (23.7.9)$$

Induction on the dimension of the skeleton. If  $k = 0$ , and if  $n > 0$ , then  $H_n^{\Delta}(X^0) = 0$  and  $H_n(X^0) = 0$ . If  $n = 0$ , then we obtain  $\mathbb{Z}^{\oplus c}$ , where  $c$  is the number of connected components. The Five Lemma handles this.

## 23.8 Mayer-Vietoris Sequence

Mayer-Vietoris sequences are useful computational tools for computing homology. It is analogous to the van Kampen theorem for homotopy. Often used in proofs by induction on the number of dimensions.

**Theorem 23.8.1: Mayer-Vietoris Theorem**

iven a space  $X$  and two subsets  $A, b \subseteq X$  such that:

$$X = \text{Int}(A) \cup \text{Int}(B) \quad (23.8.1)$$

Then the Mayer-Vietoris sequence:

$$\cdots \rightarrow H_n(A \cap B) \xrightarrow{\varphi} H_n(A) \oplus H_n(B) \xrightarrow{\psi} H_n(X) \rightarrow H_{n-1}(A \cap B) \rightarrow \cdots \quad (23.8.2)$$

Is an exact sequence.

The maps are  $\phi(x) = (-x, x)$ , and  $\psi(x, y) = x + y$ .  $\varphi : C_n(A \cap B) \rightarrow C_n(A) \oplus C_n(B)$  and  $\psi : C_n(A) \oplus C_n(B) \rightarrow C_n(X)$ . To define  $\partial$  connecting maps, we have a short exact sequence of chain complexes:

$$0 \rightarrow C_n(A \cap B) \xrightarrow{\varphi} C_n(A) \oplus C_n(B) \xrightarrow{\psi} C_n(A + B) \rightarrow 0 \quad (23.8.3)$$

$n$  chains in  $X$  such that each simplex  $\sigma : \Delta^n \rightarrow X$  is entirely in  $A$  or in  $B$ , or both. The proof is similar in technical details to excision.

**Theorem 23.8.2.** *If  $(X, A)$  is a CW pair, then  $A$  has a neighborhood  $A \subseteq \mathcal{U} \subseteq X$  such that  $A$  is a deformation retract of  $\mathcal{U}$ .*

**Example 23.8.1** What is  $H_n(S^2 \times S^1)$ ? Write  $S^2$  as  $S_+^2 \cup S_-^2$ , where  $S_\pm^2 \simeq D^2$ . Let  $A = S_+^2 \times S^1$  and  $B = S_-^2 \times S^1$ . Then  $A$  and  $B$  are solid torii, and  $A \cap B = S^1 \times S^1$ , which is a torus. If you have a closed, smooth manifold, then the highest degree that has homology is the degree of the manifold. So, we can start the Mayer-Vietoris sequence at 3. But  $A \cap B$  is a torus, and thus  $H_3(A \cap B) = 0$ . Similarly,  $H_3(A) = H_3(B) = 0$ . So, we have:

$$0 \rightarrow H_3(X) \rightarrow H_2(S^1 \times S^1) \rightarrow H_2(S^1) \oplus H_2(S^1) \quad (23.8.4)$$

But  $H_2(S^1 \times S^1) = \mathbb{Z}$ , and  $H_2(S^1) = 0$ , so we obtain:

$$0 \rightarrow H_3(X) \rightarrow \mathbb{Z} \rightarrow 0 \rightarrow H_2(X) \rightarrow \cdots \quad (23.8.5)$$

Now  $H_1(S^1 \times S^1) = \mathbb{Z}^2$ , so we can simplify further to obtain:

$$0 \rightarrow H_2(X) \rightarrow \mathbb{Z}^2 \rightarrow H_1(S^1) \oplus H_1(S^1) \rightarrow H_1(X) \quad (23.8.6)$$

Thus,  $H_0(X) = \mathbb{Z}$ . Now for  $H_1$  and  $H_2$ :

$$0 \rightarrow H_2(X) \rightarrow \mathbb{Z}^2 \rightarrow \mathbb{Z} \oplus \mathbb{Z} \rightarrow H_1(X) \rightarrow 0 \quad (23.8.7)$$

We get  $H_1(X) = H_2(X) = \mathbb{Z}$ , and thus  $S^2 \times S^1 \neq S^3$ .

## 23.9 Cohomology

Let  $G$  be an Abelian group, and let  $\text{Hom}$  be a functor from Abelian groups to Abelian groups:  $\text{Hom}(A, G)$  is the set of group homomorphism. This is again a group,  $(f_1 + f_2)(a) = f_1(a) + f_2(a)$ .

**Theorem 23.9.1.** *If  $A$ ,  $B$ , and  $G$  are Abelian groups, then:*

$$\text{Hom}(A \oplus B, G) \simeq \text{Hom}(A, G) \oplus \text{Hom}(B, G) \quad (23.9.1)$$

**Theorem 23.9.2.** *If  $G$  is an Abelian group, then  $\text{Hom}(\mathbb{Z}, G) \simeq G$ .*

*Proof.* For let  $f \mapsto f(1) \in G$ . □

**Theorem 23.9.3.** *If  $G$  is an Abelian group, and if  $m \in \mathbb{N}$ , then:*

$$\text{Hom}(Z^m, G) \simeq G^m \quad (23.9.2)$$

$\text{Hom}$  is a functorial contravariant. The functorial properties are:

$$\text{id}^* = \text{id} \quad (23.9.3a)$$

$$(\varphi\psi)^* = \psi^*\varphi^* \quad (23.9.3b)$$

$$0^* = 0 \quad (23.9.3c)$$

Given any chain complex of Abelian groups with the boundary map  $\partial^2 = 0$  (Simplicial, singular, cellular, etc.), applying the  $\text{Hom}$  functor, we get:

$$\cdots \leftarrow \text{Hom}(C_n, G) \leftarrow \text{Hom}(C_{n-1}, G) \leftarrow \text{Hom}(C_{n-2}, G) \leftarrow \cdots \quad (23.9.4)$$

With maps  $\delta$  called the *coboundary*. If  $\varphi \in C_n^* = \text{Hom}(C_n, G)$ ,  $\varphi : C_n \rightarrow G$ , then  $\delta\varphi \in C_{n+1}^*$ .  $\delta\varphi = \varphi\delta$ . The cohomology groups are:

$$H^n(X, G) = \ker(\delta)/\text{Im}(\delta) \quad (23.9.5)$$

## 23.10 Len's Spaces

Take  $\mathbb{C}^n$ , and consider the unit sphere  $S^{2n-1} \subseteq \mathbb{C}^n$ . Define the map  $p$  on  $\mathbb{C}^n$  by  $(z_1, \dots, z_n) \mapsto (\zeta^{j_1}z_1, \dots, \zeta^{j_n}z_n)$  Where  $\zeta = \exp(2\pi i/m)$ , and  $0 < j_i < m$  for all  $i$ , and furthermore  $j$  and  $m$  are coprime. The lens space is  $S^{2n-1}/\mathbb{Z}_m$ , and is denoted  $L_m(j_1, \dots, j_n)$ . This is the first example of two closed manifolds without boundary, of the same dimension, that are homotopy equivalent, but not homeomorphic. For  $L_7(1, 1)$  is homotopy equivalent to  $L_7(1, 2)$ , but not homeomorphic. The homology groups are  $\mathbb{Z}, \mathbb{Z}_m, 0, \mathbb{Z}$ , and then 0 for all others. The cohomology is  $\text{Hom}(G, \mathbb{Z})$ , where  $G$  are the groups from the chain complex from homology. We obtain the same sequence, but in reverse. Next is the universal coefficient theorem. We can compute  $H^n(X; G)$  from  $H(X)$  (Singular homology).

**Theorem 23.10.1.** *If  $H_n(X) \simeq \mathbb{Z}^m \oplus T_n$  where  $T_n$  is a torsion group, then:*

$$H^n(X; \mathbb{Z}) \simeq \mathbb{Z}^m \oplus T_{n-1} \quad (23.10.1)$$

Taking  $\mathbb{R}P^2$  as an example, the homology groups are  $\mathbb{Z}, \mathbb{Z}_2, 0, 0, \dots$ . Thus, using this theorem, the cohomology is  $\mathbb{Z}, 0, \mathbb{Z}_2, 0, 0, \dots$ . As another case, if  $G = \mathbb{R}$  or  $\mathbb{G}$ , then:

$$H_n(X; \mathbb{R}) \simeq \mathbb{R}^m \quad (23.10.2)$$

And also:

$$H^n(X; \mathbb{R}) \simeq \mathbb{R}^m \quad (23.10.3)$$

These are dual vector spaces. From UTC, we have:

$$H^1(X; G) \simeq \text{Hom}(H_1(X); G) \quad (23.10.4)$$

If  $X$  is path-connected, then  $H_1(X)$  is the Abelianization of  $\pi_1(X)$ .

### Theorem 23.10.2: Thom's Theorem

Every  $n$  cycle  $\xi \in H_n(X)$  for a smooth manifold  $X$  is realizable as a triangulated oriented submanifold  $M$  of dimension  $n$ , kinda. ■

This theorem leads to the following idea: To define cocycles, we need a procedure that assigns elements in  $G$  to every  $n$  cycle in  $X$ . If  $G = \mathbb{R}$ , then the integer  $k$  from Thom's theorem becomes irrelevant. Thom's theorem says that it suffices to assign an element in  $\mathbb{R}$  to every  $n$  dimensional submanifold. A differential  $n$  form  $\alpha$  on  $X$ ,  $X \mapsto TX$ , where  $TX$  is the tangent bundle, and  $TX \mapsto T^*X$ , where  $T^*X$  is the tangent co-bundle.

## 23.11 Cohomology Rings

Recall that  $H^*(X; R)$ , where  $R$  is the coefficient ring ( $\mathbb{Z}, \mathbb{Z}_n, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ , etc.), is the direct product:

$$H^*(X; R) = \bigoplus_{n \in \mathbb{Z}} H^n(X; R) \quad (23.11.1)$$

This is the cup product. It is distributive and associative, and forms a graded ring. The unit,  $1$ , is the class in  $\text{Hom}(C_0(X); R)$  such that  $1(x) = 1_R$ , for all  $x \in X$ . There is also a contravariant functor from topological spaces to unital graded rings.

**Theorem 23.11.1.** *If  $f : X \rightarrow Y$  is continuous, then there is a pullback  $f^* : H^*(Y, R) \rightarrow H^*(X, R)$  which is a ring homomorphism for these unital graded rings.*

The proof comes directly from the various definitions involved in the theorem. That is, there are no surprises in the proof. The implications are quite strong, however, and limit the possibilities.

**Theorem 23.11.2.** *If  $\alpha \in H(X, R)$  and if  $\beta \in H^\ell(X, R)$ , if  $R$  is a commutative ring, then the cohomology ring is graded commutative, that is  $\alpha \cup \beta = (-1)^{k\ell} \beta \cup \alpha$ , where  $\cup$  is the cup product of  $\alpha$  and  $\beta$ .*

**Example 23.11.1** Let's compute  $H^*(\mathbb{T}^2, \mathbb{Z})$ , where  $\mathbb{T}^2 = S^1 \times S^1$ . We'll use simplicial homology to do this. Using the planar representation, we get the following:

$$\Delta_0 = \mathbb{Z} = \langle v \rangle \quad (23.11.2a)$$

$$\Delta_1 = \mathbb{Z}^3 = \langle a, b, c \rangle \quad (23.11.2b)$$

$$\Delta_2 = \mathbb{Z}^2 = \langle u, L \rangle \quad (23.11.2c)$$

After magic, we get:

$$H^0 = \mathbb{Z} = \langle v^* \rangle \quad (23.11.3a)$$

$$H^1 = \mathbb{Z}^2 = \langle a^* + c^*, b^* + c^* \rangle \quad (23.11.3b)$$

$$H^2 = \mathbb{Z} = \langle u^* \rangle \quad (23.11.3c)$$

What is the ring structure? So,  $H^*(\mathbb{T}^2, \mathbb{Z})$  is generated by four elements, and thus we have  $\mathbb{Z}^4$ .

**Example 23.11.2** Consider  $H^*(\mathbb{R}\mathbb{P}^n, \mathbb{Z}_2)$ . We can compute the groups via cellular chain complexes. The groups  $H^k(\mathbb{R}\mathbb{P}^n, \mathbb{Z}_2)$  are isomorphic to  $\mathbb{Z}_2$  for  $k = 0, \dots, n$ .

**Theorem 23.11.3** (Hatcher, Page 212).  $H^*(\mathbb{R}\mathbb{P}^n, \mathbb{Z}_2)$ , as a graded ring, has one generator  $x$  and is isomorphic to:

$$H^*(\mathbb{R}\mathbb{P}^n, \mathbb{Z}_2) \simeq \mathbb{Z}_2[x]/\langle x^{n+1} \rangle \quad (23.11.4)$$

Where  $x$  has degree 1,  $|x| = 1$ .

## 23.12 Cap Product

The cap product of a ring  $R$ , the coefficient ring, is a thing.  $C_k(X; R)$  is the singular  $k$  chains with coefficients in  $R$ .  $C^\ell(X; R)$  is the dual complex  $\text{Hom}(C_\ell(X; R))$ . The cross product is mapped to  $C_{k-\ell}(X; R)$  by the cap product. Take a generator  $\sigma : \Delta^K \rightarrow X$  and  $\varphi : C_\ell \rightarrow R$ , and cap them:  $(\sigma \cap \varphi)$ . We should then get an element of  $C_{k-\ell}(X; R)$ . We define this by:

$$\sigma \cap \varphi = \varphi(\sigma|[v_0, \dots, v_\ell])\sigma|[v_\ell, \dots, v_k] \quad (23.12.1)$$

The image of  $\varphi$  is an element of  $R$ , and is simply the coefficient. The restriction of  $\sigma$  gives another map  $\Delta^{k-\ell} \rightarrow X$ . The cap product is well defined for cohomology and homology, so the is a map:

$$H_k(X; R) \times H^\ell(X; R) \rightarrow H_{k-\ell}(X; R) \quad (23.12.2)$$

Moreover, if  $\xi$  is a homology element and  $\alpha$  and  $\beta$  are cohomology elements, then:

$$(\xi \cap \alpha) \cap \beta = \xi \cap (\alpha \cup \beta) \quad (23.12.3)$$

In other words:

$$H_*(X; R) = \bigoplus_{j \in \mathbb{Z}} H_j(X; R) \quad (23.12.4)$$

Which is a graded module over the graded cohomology ring. The same is true in two other theories. The theory of bordism and cobordism, as well as in K-Theory (K-Homology and H-Theory).

**Theorem 23.12.1: Poincaré Duality**

If  $X$  is a closed and oriented manifold that is compact and without boundary,  $[X] \in H_n(X; \mathbb{R})$  (Fundamental cycle of  $X$ ), then homology is a free module over cohomology, with generator  $[X]$ . ■

If  $R$  is a ring, a free module that is finitely generated is just copies of  $R$ :  $R^n = R \oplus \cdots \oplus R$ . As modules, we have an isomorphism:  $H_*(X; \mathbb{R}) \cong H^*(X; \mathbb{R})$ , where have the map  $\alpha \mapsto [X] \cap \alpha$ . Also, from the universal coefficient theorem,  $H^k(X; \mathbb{R}) \cong H_k(X; \mathbb{R})$ , and  $H_{n-k} \cong H_k$ , ignoring torsion. This can be translated to Betti numbers as follows:  $\beta_{n-k} = \beta_k$ .



**Part XII**

# **Homology**



**Part XIII**

# Cohomology



Book Four

Analysis



**Part XIV**

# **Measure Theory**



## CHAPTER 24

---

### Infinite Series

---



# CHAPTER 25

---

## Real Analysis

---

When we studied ordered sets we saw that the real numbers possess the *least upper bound property*, and could therefore said to be complete. In the context of metric spaces we used the notion of Cauchy sequences, and then showed that this was equivalent to the nested intervals property. There are other equivalent notions, such as the Bolzano-Weierstrass theorem and the monotone convergence theorem. This property is fundamental to many theorems involved in a standard calculus or real analysis course. For example, the concepts of differentiation and convergence rely on completeness, and the intermediate value theorem may fail without it. On the other hand,  $\mathbb{Q}$  is not complete. The rationals are, however, *dense* in the reals. That is, elements of  $\mathbb{R}$  can be approximated arbitrarily well by elements of  $\mathbb{Q}$ .  $\mathbb{R}$  is also something called a *field*. From algebra, a field is just a set with two operations (Usually called addition and multiplication) that are defined in such a way as to give rise to the usual notions of addition, subtraction, multiplication, and non-zero division, and to give them the basic properties of associativity, commutativity, and the distributive law that is found in arithmetic.  $\mathbb{Q}$  is also a field. Moreover,  $\mathbb{R}$  and  $\mathbb{Q}$  are *ordered fields* with respect to their standard ordering. What makes  $\mathbb{R}$  special is that it is a complete ordered field. In fact,  $\mathbb{R}$  is the *only* complete ordered field (Up to isomorphism). Completeness in  $\mathbb{R}$  can be stated by fact that the real numbers have the least upper bound property.

**Definition 25.0.1** A bounded above subset of  $\mathbb{R}$  is a nonempty subset  $S \subseteq \mathbb{R}$  such that there exists an  $M \in \mathbb{R}$  such that for all  $x \in S$ ,  $x \leq M$ .

**Definition 25.0.2** An upper bound of a bounded above subset  $S \subseteq \mathbb{R}$  is a real number  $M \in \mathbb{R}$  such that for all  $x \in S$ ,  $x \leq M$ .

If  $S \subseteq \mathbb{R}$  is bounded above, then there exists infinitely many bounds. Completeness says that every bounded above subset has a smallest upper bound.

**Theorem 25.0.1** (Least Upper Bound Theorem). *If  $S \subseteq \mathbb{R}$  is bounded above, then there exists an  $s \in \mathbb{R}$ , called the least upper bound, such that  $s$  and an upper bound and for all upper bounds  $M$  of  $S$ ,  $s \leq M$ .*

The proof of Thm. 25.0.1 depends on how one defines the real numbers. This is often done via Dedekind cuts or equivalence classes of Cauchy sequences in  $\mathbb{Q}$ .

**Theorem 25.0.2.** *There exist bounded above sets  $S \subset \mathbb{Q}$  such that for all upper bounds  $M$  there exists an  $s \in \mathbb{Q}$  such that  $s$  is an upper bound of  $S$  and  $s < M$ .*

*Sketch of Proof.* For let  $S = \{x \in \mathbb{Q} : x^2 \leq 2\}$ . This set has no least upper bound. Loosely speaking this is because the least upper bound wants to be  $\sqrt{2}$ , but  $\sqrt{2}$  is not a rational number. Thus there is no rational number to fill the gap.

The least upper bound property gives rise to many theorems, many of which are equivalent to this axiom. Recall that a sequence is a function  $x : \mathbb{N} \rightarrow X$ . That is, a sequence is a function whose domain is the natural numbers, and whose image lies in some set  $X$ . A sequence of real numbers is thus a function  $x : \mathbb{N} \rightarrow \mathbb{R}$ , and a sequence of rational numbers is a function  $x : \mathbb{N} \rightarrow \mathbb{Q}$ . Often times sequences are denoted  $x_n$ , but also the image of  $n$  is usually denoted  $x(n) = x_n$  which may be a cause for confusion. That is, when we write  $x_n$  we often mean  $x(n)$ , so  $x_0, x_1, x_2$  can be written as  $x(0), x(1), x(2)$  but nobody does this. Similarly, we may mean  $x_n = x$  since nobody writes a sequence as  $x$ . For consistency, we will.

**Definition 25.0.3** A sequence in a set  $X$  is a function  $x : \mathbb{N} \rightarrow X$ . We write the image of  $n \in \mathbb{N}$  as  $x(n) = x_n$ .

The notion of *convergence* of a sequence in  $\mathbb{R}$  is defined as follows.

**Definition 25.0.4** A convergent sequence in  $S \subseteq \mathbb{R}$  is a sequence  $x : \mathbb{N} \rightarrow S$  such that there exists an  $a \in \mathbb{R}$  such that  $|a - x_n| \rightarrow 0$  as  $n \rightarrow \infty$ . We write  $x_n \rightarrow a$ .

**Definition 25.0.5** A limit of a sequence  $x$  in a subset  $S \subseteq \mathbb{R}$  is an element  $a \in \mathbb{R}$  such that  $|a - x_n| \rightarrow 0$ .

**Theorem 25.0.3.** *If  $S \subseteq \mathbb{R}$  and  $a$  and  $b$  are limits of  $x : \mathbb{N} \rightarrow S$ , then  $a = b$ .*

*Proof.* Suppose not. Then  $|a - b| > 0$ . But as  $a$  is a limit of  $x$ , there is an  $N_1 \in \mathbb{N}$  such that, for all  $n > N_1$ ,  $|a - x_n| < |a - b|/4$ . But, as  $b$  is a limit of  $x$ , there is an  $N_2 \in \mathbb{N}$  such that for all  $n > N_2$ ,  $|b - x_n| < |a - b|/4$ . Let  $N = \max\{N_1, N_2\} + 1$ . But from the triangle inequality:  $|a - b| \leq |a - x_N| + |b - x_N| < |a - b|/2$ , a contradiction. Therefore,  $a = b$ .  $\square$

The next notion to discuss is that of *subsequences*. There are two ways to define a subsequence rigorously. A subsequence of a sequence  $x : \mathbb{N} \rightarrow X$  is a sequence  $y : \mathbb{N} \rightarrow X$  such that there exists a strictly increasing sequence  $k : \mathbb{N} \rightarrow \mathbb{N}$  such that, for all  $n \in \mathbb{N}$ ,  $y_n = (x \circ k)(n)$ . Here,  $(x \circ k)$  is the *composition* of the two functions  $x$  and  $k$ . This is often written  $x_{k_n}$ , but this can occasionally be confusing. We can also just define a subsequence to be any strictly increasing sequence  $k : \mathbb{N} \rightarrow \mathbb{N}$ . Given a sequence  $x : \mathbb{N} \rightarrow X$ , since  $k$  is strictly increasing the ordering of  $x \circ k$  remains the same, and we've simply skipped over some points. Recall that monotonic sequences are sequences such that, for all  $n \in \mathbb{N}$ , either  $x_{n+1} \leq x_n$  (Monotonically decreasing), or  $x_n \leq x_{n+1}$  (Monotonically increasing). Strictly monotonic means either  $x_{n+1} < x_n$  or  $x_n < x_{n+1}$  for all  $n \in \mathbb{N}$ .

**Definition 25.0.6** A subsequence is a strictly increasing sequence  $k : \mathbb{N} \rightarrow \mathbb{N}$

**Definition 25.0.7** A convergent subsequence of a sequence  $x : \mathbb{N} \rightarrow S$  in a subset  $S \subseteq \mathbb{R}$  is a subsequence  $k$  such that  $x \circ k$  is a convergent sequence in  $S$ .

**Definition 25.0.8** A monotonic subsequence of a sequence  $x : \mathbb{N} \rightarrow S$  in a subset  $S \subseteq \mathbb{R}$  is a subsequence  $k : \mathbb{N} \rightarrow \mathbb{N}$  such that  $x \circ k$  is a monotonic sequence.

**Example 25.0.1** If  $x : \mathbb{N} \rightarrow \mathbb{N}$  is the sequence defined by  $x_n = n$ , and if  $k : \mathbb{N} \rightarrow \mathbb{N}$  is the subsequence defined by  $k_n = 2n$ , then  $x_{k_n} = 2n$ . This is the subsequence of all even numbers. If  $k_n = 2n - 1$ , then  $x_{k_n} = 2n - 1$ . This is the subsequence of all odd numbers. As a boring example, let  $k_n = n$ . Then  $x_{k_n} = n$ . This is the identity subsequence.

**Theorem 25.0.4.** *If  $S \subseteq \mathbb{R}$ ,  $x : \mathbb{N} \rightarrow \mathbb{R}$  is a convergent sequence, and if  $k : \mathbb{N} \rightarrow \mathbb{N}$  is a subsequence, then  $x \circ k$  is a convergent sequence.*

*Proof.* For let  $\varepsilon > 0$ . As  $x$  is a convergent sequence there is an  $a \in \mathbb{R}$  such that  $x_n \rightarrow a$ . Thus, there is an  $N \in \mathbb{N}$  such that, for all  $n > N$ ,  $|a - x_n| < \varepsilon$ . But  $k$  is a subsequence and is therefore strictly increasing, so for all  $n \in \mathbb{N}$ ,  $k_n \geq n$ . But then for all  $n > N$ ,  $k_n > N$ , and thus  $|x_{k_n} - a| < \varepsilon$ . Therefore,  $x_{k_n} \rightarrow a$ .  $\square$

There is an important theorem about subsequences of bounded sequences called the Bolzano-Weierstrass theorem. It states that every bounded sequence has a convergent subsequence, and is an equivalent definition of the completeness of  $\mathbb{R}$ . There are a few theorems needed before we can prove it.

**Theorem 25.0.5.** *If  $x : \mathbb{N} \rightarrow \mathbb{R}$  is a bounded monotonic sequence, then  $x$  is a convergent sequence.*

*Proof.* Let  $x$  be a bounded monotonic sequence that is increasing in  $\mathbb{R}$ . If  $x$  is decreasing, we replace the least upper bound with the greatest lower bound

and the proof is symmetric. Then  $S = \{x_n : n \in \mathbb{N}\}$  is a non-empty subset of  $\mathbb{R}$ . But  $x$  is a bounded sequence, and therefore  $S$  is a bounded subset of  $\mathbb{R}$ . By the least upper bound property there exists a least upper bound  $s \in \mathbb{R}$  of  $S$ . We now show that  $x_n \rightarrow s$ . Let  $\varepsilon > 0$  be given. Since  $s$  is the least upper bound,  $s - \varepsilon$  is not an upper bound of  $S$ , since  $s - \varepsilon < s$ . Therefore there exists an  $N \in \mathbb{N}$  such that  $s - \varepsilon < x_N$ . But  $x$  is monotonically increasing, and therefore for all  $n > N$ ,  $x_N \leq x_n$ . But, as  $s$  is a least upper bound of  $S$ ,  $x_n \leq s$ . But then, for all  $n > N$ ,  $0 \leq s - x_n \leq s - x_N < \varepsilon$ . Therefore,  $x_n \rightarrow s$ .  $\square$

The least upper bound is, in a sense, the reason why decimal expansions of real numbers work. For example, let  $x$  be the sequence 3, 3.1, 3.14, 3.141, 3.1415, 3.14159, and so forth. This sequence, which is the decimal expansion of  $\pi$ , is bounded by 4. Therefore it has a least upper bound. We define the number  $\pi$  to be the least upper bound of this sequence. Completeness is a very important property but so far it relies on the ordering of the real numbers. We want to find an equivalent definition of completeness that does not rely on ordering so that we may speak of complete spaces, or sets, which have no notion of order. We start with a different definition for the completeness of  $\mathbb{R}$ .

**Definition 25.0.9** A Cauchy sequence in a subset  $S \subseteq \mathbb{R}$  is a sequence  $x : \mathbb{N} \rightarrow S$  such that for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that for all  $n, m > N$ ,  $|x_n - x_m| < \varepsilon$ . That is:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} : n, m > N \Rightarrow |x_n - x_m| < \varepsilon \quad (25.0.1)$$

**Theorem 25.0.6.** *If  $S \subseteq \mathbb{R}$  and if  $x : \mathbb{N} \rightarrow S$  is a convergent sequence, then it is a Cauchy sequence.*

*Proof.* For let  $x$  be a convergent sequence and let  $a$  be its limit. Let  $\varepsilon > 0$  be given. Then, as  $x_n \rightarrow a$ , there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $|x_n - a| < \varepsilon/2$ . But by the triangle inequality, for all  $n, m > N$ :

$$|x_n - x_m| \leq |x_n - a| + |x_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad (25.0.2)$$

Therefore,  $x$  is a Cauchy sequence.  $\square$

The converse of Thm. 25.0.6 turns out to be a more general notion of completeness. That is, we can apply this to spaces that do not have a notion of order, but do have a notion of completeness. There are Cauchy sequences  $x : \mathbb{N} \rightarrow \mathbb{Q}$  that do not converge. This is again related to the fact that  $\mathbb{Q}$  is not complete. For sequences  $x : \mathbb{N} \rightarrow \mathbb{R}$ , if  $x$  is Cauchy then it must converge.

**Theorem 25.0.7.** *If  $S \subseteq \mathbb{R}$ ,  $x : \mathbb{N} \rightarrow S$  is a Cauchy sequence, and if  $k : \mathbb{N} \rightarrow \mathbb{N}$  is a subsequence, then  $x \circ k$  is a Cauchy sequence.*

*Proof.* For let  $\varepsilon > 0$ . As  $x$  is a Cauchy sequence, there is an  $N \in \mathbb{N}$  such that, for all  $n, m > N$ ,  $|x_n - x_m| < \varepsilon$ . But, as  $k$  is a subsequence it is strictly increasing, and thus for all  $n \in \mathbb{N}$ ,  $k_n \geq n$ . But then for all  $n, m > N$ ,  $k_n, k_m > N$ , and thus  $|x_{k_n} - x_{k_m}| < \varepsilon$ . Thus,  $x \circ k$  is Cauchy.  $\square$

**Theorem 25.0.8.** *If  $S \subseteq \mathbb{R}$  and if  $x : \mathbb{N} \rightarrow S$  is a Cauchy sequence, then  $x$  is a bounded sequence.*

*Proof.* For as  $x$  is a Cauchy sequence there is an  $N \in \mathbb{N}$  such that, for all  $n, m > N$ ,  $|x_n - x_m| < 1$ . Then, for all  $n > N + 1$ ,  $x_{N+1} - 1 < x_n < x_{N+1} + 1$ . Let  $M = \max\{|x_0|, |x_1|, \dots, |x_{N+1}| + 1\}$ . Then for all  $n \in \mathbb{N}$ ,  $|x_n| \leq M$ .  $\square$

We cannot replace the requirement that, for all  $n, m > N$ ,  $|x_n - x_m| < \varepsilon$  with  $n, n + k$  for some fixed  $k \in \mathbb{N}$ . There are sequences such that  $x_{n+1} - x_n \rightarrow 0$ , yet  $x$  is not Cauchy. Indeed, there are such sequences that are bounded.

**Example 25.0.2** There are unbounded sequences  $x$  such that  $x_{n+1} - x_n \rightarrow 0$ . For let  $x : \mathbb{N} \rightarrow \mathbb{R}$  be the sequence defined by  $x_n = \sqrt{n}$ . Then:

$$|x_{n+1} - x_n| = |\sqrt{n+1} - \sqrt{n}| = \frac{1}{\sqrt{n+1} + \sqrt{n}} < \frac{1}{2\sqrt{n}} \rightarrow 0 \quad (25.0.3a)$$

But  $\sqrt{n} \rightarrow \infty$ . Moreover, there are bounded sequences  $x$  such that  $x_{n+1} - x_n \rightarrow 0$ , yet  $x$  is not Cauchy. For let  $x : \mathbb{N} \rightarrow \mathbb{R}$  be defined by  $x_n = \cos(\pi\sqrt{n})$ . Then  $x$  is bounded, and:

$$x_{n+1} - x_n = \cos(\pi\sqrt{n+1}) - \cos(\pi\sqrt{n}) \quad (25.0.3b)$$

$$= -2 \sin\left(\pi \frac{\sqrt{n+1} - \sqrt{n}}{2}\right) \sin\left(\pi \frac{\sqrt{n+1} + \sqrt{n}}{2}\right) \quad (25.0.3c)$$

But we saw from the previous example that  $\sqrt{n+1} - \sqrt{n} \rightarrow 0$ , and therefore  $x_{n+1} - x_n \rightarrow 0$ .  $x$  is not Cauchy, however, for let  $k : \mathbb{N} \rightarrow \mathbb{N}$  be the subsequence defined by  $k_n = n^2$ . Then:

$$x_{k_n} = \cos(\pi n) = (-1)^n \quad (25.0.3d)$$

And this is not a Cauchy sequence. By Thm. 25.0.7,  $x$  is not a Cauchy sequence.

**Theorem 25.0.9.** *Every sequence in  $\mathbb{R}$  has a monotonic subsequence.*

*Proof.* Let  $x$  be a sequence in  $\mathbb{R}$ . Call  $n$  a “peak point” if  $x_n \geq x_m$  for all  $m \geq n$ . If there are infinitely many of these peak points, then we have obtained a decreasing sequence, since the  $n^{th}$  peak point will be greater than or equal to the  $(n+1)^{th}$  peak point. We have thus obtained a monotonically decreasing subsequence. If there are finitely many, there are either zero or there is a last one,  $x_{n_0}$ . Then  $x_{n_0+1}$  is not a peak point. But then there is a  $k \in \mathbb{N}$

such that  $k > n_0 + 1$  and  $x_k \geq x_{n_0+1}$ , for otherwise  $x_{n_0+1}$  would be a peak point. But  $x_k$  is also not a peak point, and so there is a  $k_1$  such that  $k_1 > k$  and  $x_{k_1} \geq x_k$ . This pattern continues, and we thus have a monotonically increasing subsequence. If there are zero peak points, repeat the argument above argument with  $x_{n_0} = x_1$ .  $\square$

There's probably some axiom of choice stuff going on here.

**Theorem 25.0.10** (Bolzano-Weierstrass Theorem). *If  $x : \mathbb{N} \rightarrow \mathbb{R}$  is a bounded sequence, then there is a convergent subsequence  $k : \mathbb{N} \rightarrow \mathbb{N}$  of  $x$ .*

*Proof.* By Thm. 25.0.9, if  $x : \mathbb{N} \rightarrow \mathbb{R}$  is a sequence, then there is a monotonic subsequence  $k : \mathbb{N} \rightarrow \mathbb{N}$ . But by Thm. 25.0.5, bounded monotonic sequences converge. Thus  $x \circ k$  converges. Therefore  $k$  is a convergent subsequence of  $x$ .  $\square$

This notion is so important it has a name. A sequentially compact space is a space such that every sequence has a convergent subsequence. The Bolzano-Weierstrass Theorem is equivalent to saying that every closed and bounded subset of  $\mathbb{R}$  is sequentially compact. The general notion of *compactness* is a topological one, but as it turns out sequential compactness and compactness are identical concepts in a *metric space*. Metric spaces will be one of the primary subjects of study in functional analysis. In  $\mathbb{R}^n$  there is another equivalent, and perhaps more intuitive, definition of compactness. A subset of  $\mathbb{R}^n$  is compact if and only if it is closed and bounded. A set  $S \subseteq \mathbb{R}$  is closed if for all convergent sequences  $x : \mathbb{N} \rightarrow S$ , the limit also lies in  $S$ . Compactness will be discussed later in the context of continuous functions on a compact set.

**Example 25.0.3** Find a subsequence  $k$  of the identity  $x : \mathbb{N} \rightarrow \mathbb{R}$  defined by  $x_n = n$  for which both  $\sin(x \circ k)$  and  $\cos(x \circ k)$  converge. First note that for any subsequence  $k$ ,  $(x \circ k)(n) = k_n$ . In degrees this is simple:

$$k_n = 360n + 45 \Rightarrow \sin(k_n) = \cos(k_n) = \frac{1}{\sqrt{2}} \quad (25.0.4a)$$

In radians we need to be a little more careful. Let  $y : \mathbb{N} \rightarrow \mathbb{R}$  be defined by  $y_n = \sin(n)$ . Then  $y$  is bounded and by the Bolzano-Weierstrass theorem, there is a convergent subsequence  $k$ . Let  $z : \mathbb{N} \rightarrow \mathbb{R}$  be defined by  $z_n = \cos(k_n)$ . Then  $z$  is bounded and by the Bolzano-Weierstrass theorem there is a convergent subsequence  $j$ . Let  $k_j$  denote the subsequence  $k \circ j$ . But any subsequence of a convergent sequence converges to the same limit, and therefore  $\sin(k_j)$  is a convergent sequence. Thus,  $\sin(k_j)$  and  $\cos(k_j)$  are convergent sequences. It's also possible to make them converge to the same limit. We need to know that  $\{n \bmod \alpha : n \in \mathbb{N}\}$  is dense in  $(0, \alpha)$  when  $\alpha$  is irrational. Thus there is a subsequence such that  $k_n \bmod 2\pi \rightarrow \pi/4$ . Then  $\sin(k_n)$  and  $\cos(k_n)$  both

converge to  $1/\sqrt{2}$ . Let's first try to find a subsequence such that  $\cos(k_n) \rightarrow 1$ . If we can do that, we simply need to modify the argument so that  $\cos(k_n) \rightarrow 1/\sqrt{2}$ . Let  $k$  be a sequence of integers such that  $0 < n - 2\pi k_n < 2\pi$ . Let  $\varepsilon > 0$  and let  $N \in \mathbb{N}$  be such that  $N > \frac{2\pi}{\varepsilon}$ . Now consider the set:

$$A_N = \{n - 2\pi k_n : n = 1, 2, \dots, N + 1\} \quad (25.0.4b)$$

Then  $A_N$  has  $N + 1$  elements and by the pigeon-hole principle there are elements that are within  $2\pi/\frac{2\pi}{\varepsilon} = \varepsilon$  of each other. Let  $n_1$  and  $n_2$  be such numbers. Then:

$$\cos(n_2 - n_1) = \cos(n_2 - n_1 - 2\pi(k_2 - k_1)) \quad (25.0.4c)$$

$$= \cos((n_2 - 2\pi k_2) - (n_1 - 2\pi k_1)) \quad (25.0.4d)$$

$$= \cos(\xi) \quad (25.0.4e)$$

Where  $\xi$  is a number such that  $0 < |\xi| < \varepsilon$ . But then  $|1 - \cos(\xi)| < \frac{\varepsilon^2}{2}$ . And  $n_2 - n_1$  is a natural number, so we can find a subsequence  $k$  such that  $\cos(k_n) \rightarrow 1$ . Modifying this with  $\pi/4$  and  $1/\sqrt{2}$  gives the result.

**Theorem 25.0.11.** *If  $x : \mathbb{N} \rightarrow \mathbb{R}$  is a Cauchy sequence, then it converges.*

*Proof.* If  $x$  is Cauchy, then it is bounded. By the Bolzano-Weierstrass theorem there is a convergent subsequence  $k$ . But then there is an  $a \in \mathbb{R}$  such that  $x_{k_n} \rightarrow a$ . We now must show that  $x_n \rightarrow a$ . Let  $\varepsilon > 0$  be given. As  $x_{k_n} \rightarrow a$ , there is an  $N_1 \in \mathbb{N}$  such that for all  $n > N_1$ ,  $|x_{k_n} - a| < \frac{\varepsilon}{2}$ . But as  $x$  is a Cauchy sequence, there is an  $N_2$  such that for all  $n, m > N_2$ ,  $|x_n - x_m| < \frac{\varepsilon}{2}$ . Let  $N = \max\{N_1, N_2\}$ . But  $k$  is a subsequence, and thus for all  $n > N$ ,  $k_n > N$ . But then if  $n > N$ ,  $|x_{k_n} - x_n| < \frac{\varepsilon}{2}$ . By the triangle inequality,  $|a - x_n| \leq |a - x_{k_n}| + |x_{k_n} - x_n| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ .  $\square$

Real numbers can be constructed by considering *equivalence classes* of Cauchy sequences of rational numbers. Two Cauchy sequences  $x_n$  and  $y_n$  are equivalent if  $x_n - y_n \rightarrow 0$ . By considering the set of all such equivalent sequences, we can give a more rigorous construction of the real numbers.

**Example 25.0.4** Let  $x : \mathbb{N} \rightarrow \mathbb{Q}$  be the sequence:

$$x_n = \frac{2n+3}{n} \quad (25.0.5a)$$

Let  $\varepsilon > 0$  and let  $N = \lceil 6/\varepsilon \rceil + 1$ . Then, for  $n, m > N$ , we have:

$$|x_n - x_m| = \left| \frac{2n+3}{n} - \frac{2m+3}{m} \right| = 3 \left| \frac{n-m}{nm} \right| < \frac{6}{\min\{n, m\}} < \frac{6}{N} < \varepsilon \quad (25.0.5b)$$

Therefore  $x$  is a Cauchy sequence of rational numbers. It has a standard representation of 2 since  $x_n \rightarrow 2$ . To see this:

$$\left| 2 - \frac{2n+3}{n} \right| = \left| \frac{3}{n} \right| \rightarrow 0 \quad (25.0.5c)$$

There are other elements of the equivalence class for 2. Indeed the sequence  $x : \mathbb{N} \rightarrow \mathbb{Q}$  defined by  $x_n = 2$  for all  $n \in \mathbb{N}$  is such an example. The equivalence classes also define the irrational numbers as well. For let  $x : \mathbb{N} \rightarrow \mathbb{Q}$  be defined by:

$$x_n = \sum_{k=0}^n \frac{(-1)^k}{n!} \quad (25.0.5d)$$

The ratio test, or the alternating series test, can be applied to show that this converges. Convergent sequences are Cauchy sequence, and thus  $x$  can be used to represent some real number. The number it represents is  $e^{-1}$ , which is irrational. If one recalls the history of  $e$ , we know that the equivalence class for  $e^{-1}$  also contains the following sequence:

$$x_n = \left(1 - \frac{1}{n}\right)^n \quad (25.0.5e)$$

We have seen that the least upper bound axiom, together with the ordered field structure that  $\mathbb{R}$  possesses, implies that Cauchy sequence in  $\mathbb{R}$  converge. This can be reversed, showing that we now have two equivalent definitions of completeness.

**Theorem 25.0.12.** *If  $x : \mathbb{N} \rightarrow \mathbb{R}$  is a bounded monotonic sequence, then  $x$  is a Cauchy sequence.*

*Proof.* For suppose not. Suppose  $x$  is monotonically increasing. If  $x$  is not Cauchy then there exists an  $\varepsilon > 0$  such that, for all  $N \in \mathbb{N}$  there exists  $n, m > N$  such that  $|x_n - x_m| \geq \varepsilon$ . But if  $x$  is bounded, there is an  $s$  such that, for all  $n \in \mathbb{N}$ ,  $|x_n| \leq s$ . From the Archimedean principle, as  $\varepsilon > 0$  there is an  $N_1 \in \mathbb{N}$  such that  $x_1 + N_1\varepsilon > s$ . Let  $X = \{x_n : n \in \mathbb{N}\}$ . For all  $N \in \mathbb{N}$ ,  $N \geq 2$ , there exists a function  $z : \mathbb{Z}_N \rightarrow X$  such that, for all  $n \in \mathbb{Z}_{N-1}$ ,  $z_n < z_{n+1}$ , and  $\min(\{|z_n - z_m| : n, m \in \mathbb{Z}_N\}) \geq \varepsilon$ . By induction, let  $z_1 = x_1$ . As  $x$  is not Cauchy, there are  $n, m > 1$  such that  $|x_n - x_m| \geq \varepsilon$ . But from monotonicity,  $x_m \geq x_1$ , and thus  $|x_n - x_1| \geq \varepsilon$ . Let  $z_2 = x_n$ . Suppose it is true for  $N \in \mathbb{N}$ . Let  $z : \mathbb{Z}_N \rightarrow X$  be such a function. As  $x$  is not Cauchy and monotonic, there is an  $n > N$  such that  $|x_n - z_N| \geq \varepsilon$ . Let  $z' : \mathbb{Z}_{N+1} \rightarrow X$  be defined by:

$$z'_k = \begin{cases} z_k, & 1 \leq k \leq N \\ x_n, & k = N+1 \end{cases} \quad (25.0.6a)$$

From monotonicity, for all  $n \in \mathbb{Z}_N$ ,  $z'_{N+1} - z'_n \geq \varepsilon$ . Moreover,  $z'_{N+1} > z'_N$ . Thus  $z'$  satisfies the criterion. Thus, there is a function  $z : \mathbb{Z}_{N_1+1} \rightarrow X$  such that  $z$  is increasing and  $\min(\{|z_n - z_m| : n, m \in \mathbb{Z}_{N_1}\}) \geq \varepsilon$ . But then:

$$z_{N_1+1} - z_1 = \sum_{n=1}^{N_1} (z_{n+1} - z_n) \geq N_1\varepsilon \quad (25.0.6b)$$

But then:

$$z_{N_1+1} > z_1 + N_1 \varepsilon \quad (25.0.6c)$$

But  $z_1 \in X$ , and thus  $z_1 \geq x_1$ . But then  $z_{N_1+1} > x_1 + N\varepsilon$ . But  $s < x_1 + N\varepsilon$ , a contradiction as  $s \geq x_n$  for all  $n \in \mathbb{N}$ . Therefore,  $x$  is Cauchy.  $\square$

This shows that the monotone convergence theorem can be proved without the least upper bound principle. The proof becomes messier, however. We now prove the equivalence of completeness and the least upper bound axiom.

**Theorem 25.0.13.** *If every Cauchy sequence in  $\mathbb{R}$  is a convergent sequence, then every bounded above subset of  $\mathbb{R}$  has a least upper bound.*

*Proof.* For if  $L \subseteq \mathbb{R}$  is non-empty and bounded then there is an  $a \in L$  and an  $s \in \mathbb{R}$  such that, for all  $y \in L$ ,  $y \leq s$ . If  $s \in L$ , then  $s$  is a least upper bound of  $L$ . Suppose not. Let  $S = \{y \in \mathbb{R} : \forall_{x \in L} x \leq y\}$ . Then  $S$  is non-empty, as  $s \in S$ . Suppose  $s$  is not the least upper bound of  $L$  and define the following:

$$A_k = \left\{ s - \frac{n}{2^k} : n \in \mathbb{N} \right\} \cap S \quad (25.0.7a)$$

There exists an  $N \in \mathbb{N}$  such that, for all  $k > N$ ,  $A_k \neq \emptyset$ , for otherwise  $s$  would be a least upper bound. Moreover, for all  $k > N$ ,  $A_k$  is finite for by the Archimedean property there is an  $n \in \mathbb{N}$  such that  $s - n/2^k < x$ , and thus for all  $m > n$ ,  $s - m/2^k \notin A_k$ . Lastly,  $A_k \subseteq A_{k+1}$ . Let  $x : \mathbb{N} \rightarrow \mathbb{R}$  be defined by:

$$x_n = \min(A_{n+N}) \quad (25.0.7b)$$

This is well defined since, for all  $n > N$ ,  $A_n$  is finite. Then, since  $A_n \subseteq A_{n+1}$  for all  $n > N$ ,  $x$  is a monotonically decreasing sequence. But then  $x$  is monotonic and bounded, and is therefore Cauchy. But Cauchy sequences converge, and thus there is a  $c \in \mathbb{R}$  such that  $x_n \rightarrow c$ . For all  $y \in L$ ,  $y \leq c$ . For if there is a  $y \in L$  such that  $c < y$ , then there is an  $N$  such that  $x_N < y$ , a contradiction as  $x_N \in S$ . Thus,  $c$  is an upper bound of  $L$ . Suppose  $c$  is not the least upper bound, and thus there is a  $d \in S$  such that  $d < c$ . But then there is an  $k \in \mathbb{N}$  such that  $c - d < 2^{-k}$ . But then  $x_{k+1} < c$ , a contradiction as  $x$  is monotonically decreasing and  $x_n \rightarrow c$ . Thus,  $c$  is the least upper bound.  $\square$

We've now used the Archimedean property twice. This says that for any  $\varepsilon > 0$  and any  $x > 0$ , there is an  $N \in \mathbb{N}$  such that  $N\varepsilon > x$ . It is equivalent to saying the real numbers have no "infinitesimals." We now have two ways to talk about the completeness of  $\mathbb{R}$ . The monotone convergence theorem can also be taken as axiom, and shown that it implies completeness, as well as the Bolzano-Weierstrass theorem. Lastly, there is the Nested Interval Theorem which will be proved later in the context of more general metric spaces.

## 25.1 Old stuff

### 25.1.1 Continuity

We now discuss continuity, uniform continuity, and related theorems.

**Definition 25.1.1** A function  $f : S \rightarrow \mathbb{R}$  on a subset  $S \subseteq \mathbb{R}$  continuous at a point  $x \in S$  is a function such that for all  $\varepsilon > 0$  there is a  $\delta > 0$  such that for all  $x_0 \in S$ ,  $|x - x_0| < \delta$  implies  $|f(x) - f(x_0)| < \varepsilon$ . That is:

$$\forall \varepsilon > 0 \exists \delta > 0 : x \in S, |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon \quad (25.1.1)$$

**Theorem 25.1.1.** *If  $S \subseteq \mathbb{R}$ ,  $x \in S$ ,  $f : S \rightarrow \mathbb{R}$  is continuous at  $x$ , and if  $a : \mathbb{N} \rightarrow S$  is a convergent sequence such that  $a_n \rightarrow x$ , then  $f(a_n) \rightarrow f(x)$ .*

*Proof.* For let  $\varepsilon > 0$ . As  $f$  is continuous there is a  $\delta > 0$  such that, for all  $x_0 \in S$  such that  $|x - x_0| < \delta$ ,  $|f(x) - f(x_0)| < \varepsilon$ . But  $a_n \rightarrow x$ , and thus there is an  $N \in \mathbb{N}$  such that, for all  $n > N$ ,  $|x - a_n| < \delta$ . But then, for all  $n > N$ ,  $|f(x) - f(a_n)| < \varepsilon$ . Therefore,  $f(a_n) \rightarrow f(x)$ .  $\square$

The converse of this theorem is true, giving us an equivalent definition of continuity.

**Theorem 25.1.2.** *If  $S \subseteq \mathbb{R}$ ,  $x \in S$  and  $f : S \rightarrow \mathbb{R}$  is a function such that for all sequences  $a : \mathbb{N} \rightarrow \mathbb{R}$  such that  $a_n \rightarrow x$ ,  $f(a_n) \rightarrow f(x)$ , then  $f$  is continuous at  $x$ .*

*Proof.* For suppose not. Then there is an  $\varepsilon > 0$  such that, for all  $\delta > 0$ , there is an  $x_0 \in S$  such that  $|x - x_0| < \delta$  and  $|f(x) - f(x_0)| \geq \varepsilon$ . Let  $a : \mathbb{N} \rightarrow \mathbb{R}$  be a sequence such that, for all  $n \in \mathbb{N}$ ,  $|a_n - x| < 1/n$ , but  $|f(x) - f(a_n)| \geq \varepsilon$ . But then  $a_n \rightarrow x$ . But for all sequences  $a$  such that  $a_n \rightarrow x$ ,  $f(a_n) \rightarrow f(x)$ . But, for all  $n$ ,  $|f(x) - f(a_n)| \geq \varepsilon$ , a contradiction. Therefore,  $f$  is continuous at  $x$ .  $\square$

**Theorem 25.1.3.** *If  $x \in \mathbb{R}$  and  $a : \mathbb{N} \rightarrow \mathbb{R}$  is a convergent sequence such that  $a_n \rightarrow x$  and for all  $n \in \mathbb{N}$ ,  $a_n \geq 0$ , then  $x \geq 0$ .*

*Proof.* For suppose not. Suppose  $x < 0$ . Let  $\varepsilon = |x|/2$ . Then, as  $\varepsilon > 0$ , there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $|x - a_n| < \varepsilon$ . But then  $a_{N+1} < x + \varepsilon = x/2 < 0$ , a contradiction as  $a_{N+1} \geq 0$ .  $\square$

**Theorem 25.1.4.** *If  $S \subseteq \mathbb{R}$ ,  $x \in S$ ,  $f : S \rightarrow \mathbb{R}$  is continuous at  $x$ , and if  $a : \mathbb{N} \rightarrow \mathbb{R}$  is a sequence such that  $a_n \rightarrow x$  and  $f(a_n) > 0$  for all  $n \in \mathbb{N}$ , then  $f(x) \geq 0$ .*

*Proof.* For suppose not. Let  $r = f(x) < 0$ , and let  $\varepsilon = |r|/2$ . Then  $\varepsilon > 0$ . But from continuity, there is a  $\delta > 0$  such that for all  $x_0 \in S$  such that  $|x - x_0| < \delta$ ,  $|f(x) - f(x_0)| < \varepsilon$ . But  $a_n \rightarrow x$ , and thus there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $|x - a_n| < \delta$ . Thus  $|f(x) - f(a_{N+1})| < \varepsilon$ . But then  $f(a_n) < f(x) + \varepsilon = f(x)/2 < 0$ , a contradiction as  $f(a_{N+1}) > 0$ . Therefore, etc.  $\square$

**Theorem 25.1.5.** *If  $x \in \mathbb{R}$ ,  $f : \mathbb{R} \rightarrow \mathbb{R}$  is continuous at  $x$ , and if  $f(x) > 0$ , then there is an open interval  $\mathcal{U}$  such that  $x \in \mathcal{U}$ , and for all  $y \in \mathcal{U}$ ,  $f(y) > 0$ .*

*Proof.* For let  $\varepsilon = f(x)/2$ . Then  $\varepsilon > 0$ , and thus there is a  $\delta > 0$  such that for all  $x_0 \in \mathbb{R}$  such that  $|x - x_0| < \delta$ ,  $|f(x) - f(x_0)| < \varepsilon$ . Let  $\mathcal{U} = (x - \delta, x + \delta)$ . Then  $\mathcal{U}$  is an open intervals and if  $y \in \mathcal{U}$ , then  $|x - y| < \delta$ , and therefore:

$$|f(y) - f(x)| < \varepsilon \Rightarrow f(y) > f(x) - \varepsilon = \frac{f(x)}{2} > 0 \quad (25.1.2)$$

Thus, for all  $y \in \mathcal{U}$ ,  $f(y) > 0$ .  $\square$

**Definition 25.1.2** A continuous function on  $S \subseteq \mathbb{R}$  is a function  $f : S \rightarrow \mathbb{R}$  such that  $f$  is continuous at all  $x \in S$ . That is:

$$\forall_{x \in S} \forall_{\varepsilon > 0} \exists_{\delta > 0} : x_0 \in S, |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon \quad (25.1.3)$$

This definition comes from the fact that continuity is a point-wise property, and not a “curve” property. Continuous functions are functions that have point-wise continuity at every point. The statement “A continuous function is a curve that you can draw,” which many have heard in calculus, is slightly misleading. There are functions that are continuous at one point and nowhere else. There are functions that are continuous on the irrationals and discontinuous on the rationals. For example, if  $x$  is rational write it as  $x = p/q$  where  $p$  and  $q$  are integers and relatively prime. Define  $f$  as follows:

$$f(x) = \begin{cases} \frac{1}{q}, & x \in \mathbb{Q} \\ 0, & x \notin \mathbb{Q} \end{cases} \quad (25.1.4)$$

This function, which is known as Dirichlet’s Function, but also as the Popcorn Function, or Thomae’s Function, is continuous at every irrational number and discontinuous at every rational number.

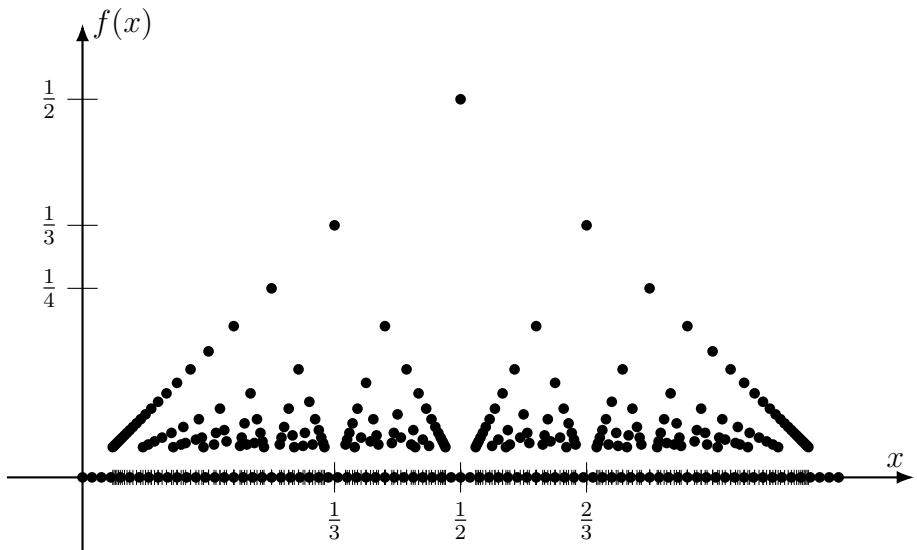


Fig. 25.1: Dirichlet's Popcorn Function

There is no “reverse,” of this function. That is, there is no function which is continuous on  $\mathbb{Q}$  and discontinuous at every irrational number. Uniform continuity is a property of all points in the domain of a function. Point-wise continuity says that given a point  $x$  and a positive number  $\varepsilon$ , one can find a  $\delta$  satisfying a certain property. The key part is that the point  $x$  must be specified first. That is, the  $\delta$  may be dependent on  $x$ . Uniform continuity occurs when a  $\delta > 0$  can be chosen regardless of  $x$ .  $\delta$  is only dependent on  $\varepsilon$ .

**Definition 25.1.3** A uniformly continuous function on a subset  $S \subseteq \mathbb{R}$  is a function  $f : S \rightarrow \mathbb{R}$  such that:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in S : \forall x_0 \in S, |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$$

Continuity is a point-wise property. There are functions that are continuous at one point and nowhere else. Uniform continuity, however, is a set property. You can't have uniform continuity at a single point, but rather on a set of points. Unless, of course, your domain  $S$  is a single point. But that's rather boring.

**Theorem 25.1.6.** *A function  $f : S \rightarrow \mathbb{R}$  is uniformly continuous if and only if for all sequences  $x, y : \mathbb{N} \rightarrow \mathbb{R}$  such that  $x_n - y_n \rightarrow 0$ ,  $f(x_n) - f(y_n) \rightarrow 0$ .*

*Proof.* Let  $\varepsilon > 0$ . If  $f$  is uniformly continuous, then there is a  $\delta > 0$  such that for all  $x, x_0 \in S$  such that  $|x - x_0| < \delta$ , we have that  $|f(x) - f(x_0)| < \varepsilon$ . But

if  $x_n - y_n \rightarrow 0$ , then there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $|x_n - y_n| < \delta$ . But then, for all  $n > N$ ,  $|f(x_n) - f(y_n)| < \varepsilon$ . Therefore,  $f(x_n) - f(y_n) \rightarrow 0$ . Proving the converse, suppose not. If  $f$  is not uniformly continuous, then there exists  $\varepsilon > 0$  such that for all  $\delta > 0$  there exists  $x, x_0 \in S$  such that  $|x - x_0| < \delta$  and yet  $|f(x) - f(x_0)| \geq \varepsilon$ . Let  $x_n$  and  $y_n$  be points such that  $|x_n - y_n| < \frac{1}{n}$  and yet  $|f(x_n) - f(y_n)| \geq \varepsilon$ . Then  $x_n - y_n \rightarrow 0$ . But if  $x_n - y_n \rightarrow 0$ , then  $f(x_n) - f(y_n) \rightarrow 0$ . But for all  $n$ ,  $|f(x_n) - f(y_n)| \geq \varepsilon$ , a contradiction.  $\square$

The requirement of uniform continuity is crucial. Let  $f : (0, 1) \rightarrow \mathbb{R}$  be defined by  $f(x) = x^{-1}$ . Then  $f$  is continuous, but not uniformly continuous. Let  $x_n = n^{-1}$  and  $y_n = 2n^{-1}$ . Then  $|y_n - x_n| = n^{-1} \rightarrow 0$ , but  $|f(y_n) - f(x_n)| = n/2$ , which diverges. Point-wise continuity says  $f(x_n) - f(x) \rightarrow 0$ , whereas uniform continuity allows the target to vary as well. Point-wise continuity can not guarantee this. The set under consideration is also crucial to uniform continuity. Indeed, the function  $f(x) = x^{-1}$  is uniformly continuous on  $(1, \infty)$ . For if  $x, y \in (1, \infty)$ :

$$|f(x) - f(y)| = \left| \frac{1}{x} - \frac{1}{y} \right| = \left| \frac{x - y}{xy} \right| \leq |x - y| \quad (25.1.5)$$

Choosing  $\delta = \varepsilon/2$  gives the result.

**Theorem 25.1.7.** *If  $f : [a, b] \rightarrow \mathbb{R}$  is continuous, then  $f$  is uniformly continuous.*

*Proof.* Suppose not. Then, by Thm. 25.1.6, there are sequences  $x, y : \mathbb{N} \rightarrow [a, b]$  such that  $x_n - y_n \rightarrow 0$ , yet there is an  $\varepsilon > 0$  such that, for all  $N \in \mathbb{N}$ , there is an  $n > N$  such that  $|f(x_n) - f(y_n)| \geq \varepsilon$ . Let  $k : \mathbb{N} \rightarrow \mathbb{N}$  be a subsequence such that, for all  $n \in \mathbb{N}$ ,  $|f(x_{k_n}) - f(y_{k_n})| \geq \varepsilon$ . By the Bolzano-Weierstrass theorem there is a convergent subsequence  $j : \mathbb{N} \rightarrow \mathbb{N}$  of  $x \circ k$ . Let  $\alpha$  be the limit. But for all  $n \in \mathbb{N}$ :

$$y_{j_{k_n}} = x_{j_{k_n}} - (y_{j_{k_n}} - x_{j_{k_n}}) \Rightarrow y_{j_{k_n}} \rightarrow \alpha \quad (25.1.6a)$$

Let  $X, Y : \mathbb{N} \rightarrow [a, b]$  be sequences defined by  $X_n = x_{j_{k_n}}$  and  $Y_n = y_{j_{k_n}}$ , respectively. Then we have:

$$f(X_n) - f(Y_n) = (f(X_n) - f(\alpha)) - (f(Y_n) - f(\alpha)) \quad (25.1.6b)$$

From continuity,  $f(X_n) \rightarrow f(\alpha)$  and  $f(Y_n) \rightarrow f(\alpha)$ , and thus  $f(X_n) - f(Y_n) \rightarrow 0$ . But for all  $n$ ,  $|f(x_{k_n}) - f(y_{k_n})| \geq \varepsilon$ , a contradiction. Therefore, etc.  $\square$

The above theorem relies on the fact that  $[a, b]$  is closed and bounded. Indeed, this is the only thing it relies on, the fact that it's an interval (Or connected) is unnecessary. We can write a more general result.

**Definition 25.1.4** A closed subset of  $\mathbb{R}$  is a subset  $S \subseteq \mathbb{R}$  such that for all convergent sequences  $x : \mathbb{N} \rightarrow S$ , the limit of  $x$  is an element of  $S$ .

**Definition 25.1.5** A compact subset of  $\mathbb{R}$  is a subset that is closed and bounded.

**Theorem 25.1.8.** *If  $S \subseteq \mathbb{R}$  is a compact subset of  $\mathbb{R}$  and if  $f : S \rightarrow \mathbb{R}$  is continuous, then  $f$  is uniformly continuous.*

Proving this more general result requires the equivalence of sequential compactness and regular compactness in  $\mathbb{R}$ . This will be shown to be true for any metric space. We can lessen the requirement that the subset be compact to being a half-open interval  $[a, \infty)$  provided that the limit of  $f(x)$  exists as  $x \rightarrow \infty$ .

**Theorem 25.1.9.** *If  $a \in \mathbb{R}$  and  $f : [a, \infty) \rightarrow \mathbb{R}$  is a continuous function such that  $\lim_{x \rightarrow \infty} f(x)$  exists, then  $f$  is uniformly continuous.*

*Proof.* For let  $\varepsilon > 0$ . As  $\lim_{x \rightarrow \infty} f(x)$  exists, there is a  $c \in \mathbb{R}$  such that, for all  $\varepsilon > 0$  there is an  $x_0 \in [a, \infty)$  such that, for all  $x > x_0$ ,  $|f(x) - c| < \varepsilon/2$ . Let  $b = x_0 + 1$ . By Thm. 25.1.7  $f$  is uniformly continuous on  $[a, b]$ , and thus there is a  $\delta > 0$  such that, for all  $x_1, x_2 \in [a, b]$  such that  $|x_1 - x_2| < \delta$ ,  $|f(x_1) - f(x_2)| < \varepsilon/2$ . But for all  $x_1, x_2 \in (b, \infty)$ :

$$|f(x_1) - f(x_2)| \leq |f(x_1) - c| + |f(x_2) - c| < \varepsilon \quad (25.1.7a)$$

And if  $x_1 \in [a, b]$  and  $x_2 \in (b, \infty)$  are such that  $|x_1 - x_2| < \delta$ , then:

$$|f(x_1) - f(x_2)| \leq |f(x_1) - f(b)| + |f(x_2) - f(b)| < \varepsilon \quad (25.1.7b)$$

Thus,  $f$  is uniformly continuous.  $\square$

**Theorem 25.1.10** (Intermediate Value Theorem). *If  $f : [a, b] \rightarrow \mathbb{R}$  is continuous and  $f(a) < f(b)$ , then for all  $z \in \mathbb{R}$  such that  $f(a) < z < f(b)$ , there is a  $c \in (a, b)$  such that  $f(c) = z$ .*

*Proof.* For if  $z \in \mathbb{R}$ , let  $g[a, b] \rightarrow \mathbb{R}$  be defined by  $g(x) = z - f(x)$  for all  $x \in [a, b]$ . Then, since  $f(a) < z$ ,  $g(z) < 0$ . But then there is an  $\varepsilon > 0$  such that, for all  $x \in [a, a + \varepsilon)$ ,  $g(x) < 0$ . Define the following:

$$\mathcal{U} = \{r > 0 : \forall_{s < r} g(a + s) < 0\} \quad (25.1.8)$$

Then  $\mathcal{U}$  is non-empty, for  $\varepsilon \in \mathcal{U}$ . But  $\mathcal{U}$  is bounded above for  $b - a \notin \mathcal{U}$ , for  $f(b) > 0$ , and thus for all  $r > b - a$ ,  $r \notin \mathcal{U}$ . But then  $\mathcal{U}$  is a non-empty bounded above subset, and by the least upper bound property, there exists a least upper bound  $c$  of  $\mathcal{U}$ . As  $c \leq b - a$ ,  $a + c \in [a, b]$ . By trichotomy, either  $g(a + c) < 0$ ,  $g(a + c) = 0$ , or  $g(a + c) > 0$ . Suppose  $g(a + c) > 0$ . Then there is a  $\varepsilon_1 > 0$  such that, for all  $x \in (a + c - \varepsilon_1, a + c]$ ,  $g(x) > 0$ . But this is a contradiction, as  $c$  is the least upper bound of  $\mathcal{U}$ . Suppose  $g(a + c) < 0$ . Then there is a  $\varepsilon_2 > 0$  such that, for all  $x \in [a + c, a + c + \varepsilon_2)$ ,  $g(x) < 0$ . Again, this is a contradiction as  $c$  is the least upper bound of  $\mathcal{U}$ . Therefore,  $g(a + c) = 0$ , ad thus  $f(a + c) = z$ .  $\square$

Another way that is commonly used to prove this theorem is the method of bisection. Start with  $x_1 = a$ ,  $x_2 = b$ , and then let  $x_3 = \frac{1}{2}(x_1 + x_2)$ . Check whether  $f(x_3) = z$  or not. If  $f(x_3) < z$ , let  $x_4 = \frac{1}{2}(x_2 + x_3)$ , otherwise let  $x_4 = \frac{1}{2}(x_1 + x_3)$ . Continuing dividing the region of interest in half, obtaining a Cauchy sequence  $x$ . The final part is to show that the limit  $c$  of  $x$  is such that  $f(c) = z$ . This theorem fails in  $\mathbb{Q}$ , for it relies on the completeness of  $\mathbb{R}$ . For example,  $f(x) = x^2$  defined on  $[0, 4]$ . Then  $2 \in [0, 4]$ , but there is no rational such that  $x^2 = 2$ . Another way to phrase this, in a more topological sense, is that the image of  $[a, b]$ , which is an interval, or a connected subset of  $\mathbb{R}$ , is again an interval, or a connected subset of  $\mathbb{R}$ . The proof that continuous functions take connected sets (Intervals) to connected sets (Again, intervals) is a lot easier than the one presented here, but relies on notions from topology. We'll revisit this when we discuss the topology of metric spaces. Another commonly used theorem in calculus is the extreme value theorem. The extreme value is used to prove Rolle's theorem, which says that if  $f$  is differentiable on  $(a, b)$  and if  $f(a) = f(b)$ , then there is a point  $c \in (a, b)$  such that  $f'(c) = 0$ . This is used to prove the mean value theorem, which says that if  $f$  is differentiable on  $(a, b)$ , then there is a point  $c \in (a, b)$  such that  $f'(c) = \frac{f(b)-f(a)}{b-a}$ . This is in turn used to prove the Fundamental Theorem of Calculus. First we prove that continuous functions on closed and bounded sets (That is, compact sets) are bounded. We stick to closed intervals for now.

**Theorem 25.1.11.** *If  $f : [a, b] \rightarrow \mathbb{R}$  is continuous, then it is bounded.*

*Proof.* Suppose not. Then for all  $n \in \mathbb{N}$ , there is an  $\alpha \in [a, b]$  such that  $f(\alpha) > n$ . Invoking choice and using the sequence  $x : \mathbb{N} \rightarrow [a, b]$  such that  $f(x_n) > n$ , we have that  $x$  is a bounded sequence, and thus by Bolzano-Weierstrass there is a convergent subsequence  $k$  of  $x$ . Let  $a$  be the limit of  $x \circ k$ . But then  $f(x_{k_n}) \rightarrow f(a)$ . But  $f(x_{k_n}) \rightarrow \infty$ , a contradiction. Therefore, etc.  $\square$

**Theorem 25.1.12** (Extreme Value Theorem). *If  $f : [a, b] \rightarrow \mathbb{R}$  is continuous, then there exists  $c \in [a, b]$  such that for all  $x \in [a, b]$ ,  $f(x) \leq f(c)$*

*Proof.* By the previous theorem,  $\{f(x) : x \in [a, b]\}$  is bounded. By completeness, there is a least upper bound. Let  $s$  be such a bound. If  $s$  is the least upper bound, then for all  $n \in \mathbb{N}$ ,  $s - \frac{1}{n}$  is not the least upper bound. Thus, for all  $n \in \mathbb{N}$  there is an  $\alpha \in [a, b]$  such that  $s - \frac{1}{n} < f(\alpha)$ . Invoking choice and choosing a sequence  $x : \mathbb{N} \rightarrow [a, b]$  such that, for all  $n \in \mathbb{N}$ ,  $s - \frac{1}{n} < f(x_n)$ . But then  $x$  is a bounded sequence, and bounded sequences have a convergent subsequence. Let  $a$  be the limit of this subsequence. From continuity,  $f(a) = \lim_{n \rightarrow \infty} f(x_{k_n})$ . But  $s - \frac{1}{n} \leq f(x_{k_n}) \leq s$ , and therefore  $f(x_{k_n}) \rightarrow s$ . Thus,  $f(a) = s$ .  $\square$

Much the way the intermediate value theorem can be generalized to say that the continuous image of connected sets is connected, the extreme value theorem can

be generalized to say that the continuous image of a compact set is compact. The proof is rather easy, but again requires topology, so we'll return to this later. The requirement of these previous theorems on continuity is crucial. Without continuity, functions on  $[a, b]$  need not be bounded and functions on  $(a, b)$  can just "jump," right over other points.

### 25.1.2 Sequences of Functions

**Definition 25.1.6** A sequence of functions from a set  $X$  to a set  $Y$  is a function  $F : \mathbb{N} \times X \rightarrow Y$ . We often write the image of  $(n, x) \in \mathbb{N} \times X$  as  $F(n, x) = F_n(x)$ .

**Definition 25.1.7** A point-wise convergent sequence of real-valued functions on a subset  $S \subseteq \mathbb{R}$  is a function  $F : \mathbb{N} \times S \rightarrow \mathbb{R}$  such that there exists a function  $f : S \rightarrow \mathbb{R}$  such that, for all  $\varepsilon > 0$  and for all  $x \in S$ , there exists an  $N \in \mathbb{N}$  such that, for all  $n > N$ ,  $|F_n(x) - f(x)| < \varepsilon$ . That is:

$$\forall_{\varepsilon > 0} \forall_{x \in S} \exists_{N \in \mathbb{N}} : n > N \Rightarrow |f(x) - F_n(x)| < \varepsilon \quad (25.1.9)$$

That is, a sequence  $F$  converges point-wise to  $f$  if, for all  $x \in \mathbb{R}$ ,  $F_n(x) \rightarrow f(x)$ . Uniform continuity requires that all of the points of the domain converge to  $f(x)$  at the same speed. That is, given any  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  that works for all points. Point-Wise convergence may not have this property.

**Definition 25.1.8** A uniformly convergent sequence of real-valued functions on a subset  $S \subseteq \mathbb{R}$  is a function  $F : \mathbb{N} \times S \rightarrow \mathbb{R}$  such that there exists a function  $f : S \rightarrow \mathbb{R}$  such that, for all  $\varepsilon > 0$  there exists an  $N \in \mathbb{N}$  such that, for all  $x \in S$  and for all  $n > N$ ,  $|F_n(x) - f(x)| < \varepsilon$ . That is:

$$\forall_{\varepsilon > 0} \exists_{N \in \mathbb{N}} \forall_{x \in S} : n > N \Rightarrow |f(x) - F_n(x)| < \varepsilon \quad (25.1.10)$$

That is,  $F_n \rightarrow f$  point-wise if for all  $x$ ,  $|F_n(x) - f(x)| \rightarrow 0$  and  $F_n \rightarrow f$  uniformly if  $\sup\{|F_n(x) - f(x)|\} \rightarrow 0$ . It is worthwhile spotting the very subtle difference between expressions 25.1.9 and 25.1.10.

**Definition 25.1.9** A limit function on a subset  $S \subseteq \mathbb{R}$  of a convergent sequence of real-valued functions  $F : \mathbb{N} \times S \rightarrow \mathbb{R}$  is a function  $f : S \rightarrow \mathbb{R}$  such that, for all  $x \in S$ ,  $F_n(x) \rightarrow f(x)$ .

**Theorem 25.1.13.** *If  $S \subseteq \mathbb{R}$ , if  $F : \mathbb{N} \times S \rightarrow \mathbb{R}$  is a convergent sequence of real-valued function, and if  $f, g : S \rightarrow \mathbb{R}$  are limit functions of  $F$ , then  $f = g$ .*

*Proof.* For suppose not. Suppose there is an  $x \in S$  such that  $f(x) \neq g(x)$ . But  $F_n(x) \rightarrow f(x)$  and  $F_n(x) \rightarrow g(x)$ . From the uniqueness of limits,  $f(x) = g(x)$ , a contradiction. Therefore, etc.  $\square$

**Example 25.1.1** Let  $F : \mathbb{N} \times [0, 1] \rightarrow \mathbb{R}$  be defined by  $F_n(x) = nx \exp(-nx)$ .  $F_n(x) \rightarrow 0$  for all  $x \in [0, 1]$ , and therefore  $F$  converges point-wise to zero. Note

that  $F'_n(x) = (n - n^2x) \exp(-nx)$ . This has a zero at  $x = n^{-1}$ , so  $F_n(x)$  has a maximum of  $e^{-1}$ . But then  $\sup |F_n(x) - f(x)| = \sup |F_n(x)| = e^{-1}$ . So  $F_n(x)$  does not converge *uniformly* to 0. The convergence is only *point-wise*.

**Example 25.1.2** Let  $F_n(x) = n^2x \exp(-nx)$ . Then  $F_n(x) \rightarrow 0$  for all  $x \geq 0$ . But  $F_n(x)$  has a maximum of  $ne^{-1}$  at  $x = n^{-1}$ . Thus  $F_n(n^{-1}) \rightarrow \infty$ . It is possible for a sequence of functions to converge point-wise to zero and for there to be a sequence such that  $F_n(x_n) \rightarrow \infty$ . Uniform convergence does not allow this.

**Definition 25.1.10** A sequence of continuous real-valued functions on a subset  $S \subseteq \mathbb{R}$  is a function  $F : \mathbb{N} \times S \rightarrow \mathbb{R}$  such that, for all  $n \in \mathbb{N}$ , the function  $g : S \rightarrow \mathbb{R}$  defined by  $g(x) = F_n(x)$  for all  $x \in S$ , is continuous.

**Theorem 25.1.14.** *If  $S \subseteq \mathbb{R}$  and if  $F : \mathbb{N} \times S \rightarrow \mathbb{R}$  is a uniformly convergent sequence of real-valued continuous functions and if  $f$  is the limit function of  $F$ , then  $f$  is continuous.*

*Proof.* For let  $x \in S$  and let  $\varepsilon > 0$ . As  $F$  converges uniformly to  $f$ , there is an  $N_0 \in \mathbb{N}$  such that, for all  $n > N_0$  and for all  $x_0 \in S$ ,  $|F_n(x_0) - f(x_0)| < \varepsilon/3$ . Let  $N = N_0 + 1$ . But, for all  $n \in \mathbb{N}$ ,  $F_n$  is a continuous function, and thus there is a  $\delta > 0$  such that, for all  $x_0 \in S$  such that  $|x - x_0| < \delta$ ,  $|F_N(x) - F_N(x_0)| < \varepsilon/3$ . But then, from the triangle inequality, for all  $x_0 \in S$  such that  $|x - x_0| < \delta$ :

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - F_N(x)| + |F_N(x) - F_N(x_0)| + |F_N(x_0) - f(x_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned} \tag{25.1.11}$$

□

The word “uniformly,” is crucial. This theorem is not necessarily true of point-wise converging functions. Let  $F$  be defined on  $[0, 1]$  as  $F_n(x) = x^n$ . Then  $F$  converges to 0 if  $x \neq 1$ , and 1 if  $x = 1$ . Thus, the limit function is discontinuous. This is possible because the convergence is point-wise and not uniform.

**Theorem 25.1.15.** *If  $f : [0, 1] \rightarrow \mathbb{R}$  is continuous, and if  $f(0) = f(1) = 0$ , then there is a sequence of polynomials  $F$  such that  $F_n \rightarrow f$  uniformly on  $[0, 1]$ .*

*Proof.* Extend  $f$  to be zero outside of  $[0, 1]$ . Let  $Q_n(x) = c_n(1-x^2)^n$  on  $[-1, 1]$ ,

and choose  $c_n$  such that  $\int_{-1}^1 Q_n(x) dx = 1$ . So we have:

$$c_n \int_{-1}^1 (1-x^2)^n dx = 2c_n \int_0^1 (1-x^2)^n dx \quad (25.1.12a)$$

$$= 2c_n \int_0^1 (1-x)^n (1+x)^n dx \quad (25.1.12b)$$

$$\geq 2c_n \int_0^1 (1-x)^n dx \quad (25.1.12c)$$

$$= \frac{2}{n+1} c_n \quad (25.1.12d)$$

From this we have that  $c_n \leq n+1$ . Let  $f_n(x) = \int_0^1 f(t)Q_n(x-t) dt$ . Then  $f_n(x)$  is a polynomial. Note that  $f(t)Q_n(x-t)$  is roughly zero when  $t$  differs from  $x$  and  $n$  is large enough. So we have:

$$f_n(x) = \int_0^1 f(t)Q_n(x-t) dt \approx f(x) \int_0^1 Q_n(x-t) dt = f(x) \quad (25.1.12e)$$

The remainder of the proof is to quantify this. Since  $f$  is zero outside of  $[0, 1]$ , if we let  $s = t - x$ , then:

$$f_n(x) = \int_{-x}^{1-x} f(s+x)Q_n(s) ds \quad (25.1.12f)$$

$$= \int_{-1}^1 f(s+x)Q_n(s) ds \quad (25.1.12g)$$

Using this, we obtain:

$$|f_n(x) - f(x)| = \left| \int_{-1}^1 f(x+t)Q_n(t) dt - \int_{-1}^1 f(x)Q_n(t) dt \right| \quad (25.1.12h)$$

$$\leq \int_{-1}^1 |f(x+t) - f(x)|Q_n(t) dt \quad (25.1.12i)$$

This comes for the fact that  $\int_{-1}^1 Q_n(t) dt = 1$  and from the integral version of the triangle inequality. Suppose  $\varepsilon > 0$ . Since  $f$  is continuous on  $[0, 1]$ , it is uniformly continuous. But if  $f$  is uniformly continuous then there exists a  $\delta > 0$  such that  $|f(x+t) - f(x)| < \frac{\varepsilon}{2}$  for all  $t < \delta$ . So we have:

$$\begin{aligned} |f_n(x) - f(x)| &\leq \int_{-1}^{-\delta} |f(x+t) - f(x)|Q_n(t) dt \\ &\quad + \int_{-\delta}^{\delta} |f(x+t) - f(x)|Q_n(t) dt \\ &\quad + \int_{\delta}^1 |f(x+t) - f(x)|Q_n(t) dt \end{aligned} \quad (25.1.12j)$$

But  $f$  is continuous on a closed and bounded set and therefore  $f$  is bounded. Let  $M$  be such a bound. Then  $|f(x+t) - f(x)| \leq 2M$ . We have:

$$|f_n(x) - f(x)| \leq 2M \int_{-1}^{-\delta} Q_n(t) dt + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} Q_n(t) dt + 2M \int_{\delta}^1 Q_n(t) dt \quad (25.1.12k)$$

But for all  $t \in [-1, -\delta]$ ,  $Q_n(t) \leq Q_n(-\delta)$ . Similarly for  $t$  in  $[\delta, 1]$ . Since  $Q_n(t)$  is an even function:

$$|f_n(x) - f(x)| \leq 4MQ_n(\delta) + \frac{\varepsilon}{2} \int_{-1}^1 Q_n(t) dt = 4MQ_n(\delta) + \frac{\varepsilon}{2} \quad (25.1.12l)$$

But since  $\delta > 0$ ,  $Q_n(\delta) \rightarrow 0$ . Therefore, there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $|Q_n(\delta)| < \frac{\varepsilon}{8M}$ . But then  $4MQ_n(\delta) < \frac{\varepsilon}{2}$ . Therefore, etc.  $\square$

Another way to put this is that if  $f$  is continuous on  $[a, b]$  and if  $\varepsilon > 0$ , then there is a polynomial  $P$  such that for all  $x \in [0, 1]$ ,  $|f(x) - P(x)| < \varepsilon$ . There is a generalization of this and the set of functions need not be polynomials. The set needs to be closed under addition, multiplication, and scalar multiplication, it must separate points, and must not take every point to zero. This is the Stone-Weierstrass theorem. This shows that continuous functions on compact sets can be approximated arbitrarily well by polynomials. Furthermore, any continuous function on a compact set can be approximated arbitrarily well by polynomials with rational coefficients. To see this, let  $f[0, 1] \rightarrow \mathbb{R}$  be continuous, and let  $\varepsilon > 0$ . Then there is a polynomial  $P$  such that  $\sup |P(x) - f(x)| < \varepsilon/2$ . Suppose  $P$  is of degree  $n$ . As  $\mathbb{Q}$  is dense in  $\mathbb{R}$ , for each coefficient  $c_k$  of  $P$  there is a  $d_k \in \mathbb{Q}$  such that  $|c_k - d_k| < \frac{\varepsilon}{2n}$ . Let  $Q(x) = \sum d_k x^k$ . Then:

$$|P(x) - Q(x)| \leq \sum_{k=0}^n |c_k - d_k| |x|^k < \frac{\varepsilon}{2} \quad (25.1.13)$$

Thus, by the triangle inequality:  $\sup |Q(x) - f(x)| < \varepsilon$ . A set is called *separable* if it contains a countable dense subset.  $\mathbb{R}$  is separable since  $\mathbb{Q}$  is dense in  $\mathbb{R}$ , and  $\mathbb{Q}$  is countable. The set of all continuous functions from  $[0, 1]$  to  $\mathbb{R}$ , which we label as  $C(I, \mathbb{R})$ , is also separable. Since any continuous function can be approximated arbitrarily well by a polynomial with rational coefficients, we can say the set of polynomials with rational coefficients is *dense* in  $C(I, \mathbb{R})$ . But the set of polynomials with rational coefficients is countable. For all  $N \in \mathbb{N}$ , define  $P_N$  as:

$$P_N = \left\{ \sum_{k=0}^N q_k x^k : q_k \in \mathbb{Q}, q_N \neq 0 \right\} \quad (25.1.14)$$

This is the set of all rational polynomials of degree  $N$ . It is countable since there is a one-to-one correspondence with the set  $\mathbb{Q}^N$ , and  $\mathbb{Q}^N$  is countable for

all  $N \in \mathbb{N}$ . But the set of all rational polynomials is simply the union over all  $P_N$ . And the countable union of countably many disjoint sets is countable. Therefore, the set of all polynomials with rational coefficients is countable. Thus  $C(I, \mathbb{R})$  is *separable*. We need to be careful when we say *dense* and *separable*, for we are implicitly speaking of some sort of notion of *closeness* on the sets. This all comes from the notion of *metrics* and *metric spaces*, and the more general *topological space*.

**Theorem 25.1.16.** *If  $f : [0, 1] \rightarrow \mathbb{R}$  is continuous, then there is a sequence of polynomials  $F$  such that  $F_n \rightarrow f$  uniformly on  $[0, 1]$ .*

*Proof.* If  $f : [0, 1] \rightarrow \mathbb{R}$  be continuous, let  $g(x) = xf(1) + (1 - x)f(0)$ . Then  $h(x) = f(x) - g(x)$  is a continuous function such that  $h(0) = h(1) = 0$  and thus by Thm. 25.1.15 there is a sequence of polynomials  $P_n(x)$  such that  $P_n(x) \rightarrow h(x)$  uniformly on  $[a, b]$ . But  $g(x)$  is a polynomial and  $f(x) = h(x) + g(x)$ . Therefore  $F_n(x) = P_n(x) + g(x)$  is a sequence of polynomials and  $F_n(x) \rightarrow f(x)$  uniformly on  $[0, 1]$ .  $\square$

**Theorem 25.1.17** (Weierstrass Approximation Theorem). *If  $f : [a, b] \rightarrow \mathbb{R}$  is a continuous function, then there is a sequence of polynomials  $P$  such that  $P_n \rightarrow f$  uniformly.*

*Proof.* If  $f : [a, b] \rightarrow \mathbb{R}$  is continuous, define  $g : [0, 1] \rightarrow \mathbb{R}$  by  $g(x) = f(\frac{x-a}{b-a})$ . Then, since the composition of continuous functions is continuous,  $g$  is a continuous function on  $[0, 1]$ . But by the Weierstrass approximation theorem there is a sequence of polynomials  $P_n(x)$  such that  $P_n(x) \rightarrow g(x)$ . Let  $F_n(x) = P_n(bx + (1 - x)a)$ . Then  $F_n(x)$  is a sequence of polynomials on  $[a, b]$ , and  $F_n(x) \rightarrow f(x)$ .  $\square$

An application of this is in the uniform approximation of continuous periodic functions by Cosines.

**Theorem 25.1.18.** *If  $f \in C[0, \pi]$  and  $\varepsilon > 0$ , then there exists  $a_0, \dots, a_n \in \mathbb{R}$  such that, for all  $x \in [0, \pi]$ :*

$$|f(x) - \sum_{k=0}^n a_k \cos(kx)| < \varepsilon \quad (25.1.15)$$

*Proof.* For  $\cos(x)$  is a bijective function on the interval  $[0, \pi]$ . Thus we can consider the function  $f(\cos^{-1}(x))$ . But since  $\cos(x)$  is continuous on  $[0, \pi]$ ,  $\cos^{-1}(x)$  is continuous on  $[-1, 1]$ . And the composition of continuous functions is continuous. So  $f(\cos^{-1}(x))$  is continuous. By the Weierstrass Approximation Theorem, there is a sequence of polynomials  $P$  such that  $P_n(x) \rightarrow f(\cos^{-1}(x))$ . But then  $P_n(\cos(x)) \rightarrow f(x)$ . But  $P_n(x)$  is a polynomial of the form  $\sum_{k=0}^n a_k x^k$ , and thus:

$$P_n(\cos(x)) = \sum_{k=0}^n a_k \cos^k(x) \quad (25.1.16a)$$

It now suffices to show that  $\cos^k(x) = \sum_{m=0}^N c_m \cos(mx)$  for suitable  $c_m$ . We prove by induction. The base case is trivial. Suppose it holds for some  $k \in \mathbb{N}$ . Then:

$$\cos^{k+1}(x) = \cos(x) \cos^k(x) = \cos(x) \sum_{k=0}^N c_k \cos(kx) \quad (25.1.16b)$$

Note that  $\cos(x) \cos(kx) = \frac{1}{2} \cos((k-1)x) + \frac{1}{2} \cos((k+1)x)$ . So we have:

$$\cos^{k+1}(x) = \frac{1}{2} \sum_{k=0}^N c_k \left( \cos((k+1)x) + \cos((k-1)x) \right) \quad (25.1.16c)$$

This completes the theorem.  $\square$

### 25.1.3 Inequalities

**Definition 25.1.11** Hölder Conjugates are non-zero real numbers  $p, q \in \mathbb{R}$  where:

$$p^{-1} + q^{-1} = 1 \quad (25.1.17)$$

**Theorem 25.1.19** (Young's Inequality). *If  $x, y \geq 0$ ,  $p > 1$ , and if  $p$  and  $q$  are Hölder Conjugates, then:*

$$xy \leq \frac{1}{p} x^p + \frac{1}{q} y^q \quad (25.1.18)$$

*Proof.* If  $x$  or  $y$  are zero, then we are done. Suppose  $x, y > 0$ . Let  $t = p^{-1}$ . As  $p$  and  $q$  are Hölder Conjugates,  $1 - t = q^{-1}$ . But then, as  $p > 1$ ,  $t$  and  $1 - t$  are positive, and thus:

$$\ln(tx^p + (1-t)y^q) \geq t \ln(x^p) + (1-t) \ln(y^q) = \ln(x) + \ln(y) = \ln(xy) \quad (25.1.19)$$

Where the inequality comes from the fact that  $\ln$  is a concave function. Exponentiating completes the proof.  $\square$

There is another way to prove this without using the concavity of the logarithmic function. Let  $y > 0$  and define  $f : (0, \infty)$  by:

$$f(x) = p^{-1}x^p - xy \quad (25.1.20a)$$

Then, differentiating, we have:

$$f'(x) = x^{p-1} - y \quad (25.1.20b)$$

This has an extremum at  $x = y^{\frac{1}{p-1}}$ . We also have:

$$f''(x) = (p-1)x^{p-2} \quad (25.1.20c)$$

And this is positive for all  $x \in (0, \infty)$ , and thus  $y^{\frac{1}{p-1}}$  is a global minimum. Using the fact that  $p$  and  $q$  are Hölder Conjugates, we have:

$$\frac{1}{p-1} = q - 1 \quad (25.1.20d)$$

Applying some algebra obtains the result. We also see that equality happens when  $x^p = y^q$ . For  $p = q = 2$  this is easy, for:

$$0 \leq \frac{(x-y)^2}{2} = \frac{x^2 + y^2}{2} - xy$$

Using this, we can prove the Peter-Paul inequality:

**Theorem 25.1.20** (Peter-Paul Inequality). *If  $x, y \in \mathbb{R}$  and  $\varepsilon > 0$ , then:*

$$ab \leq \frac{x^2}{2\varepsilon} + \frac{\varepsilon y^2}{2} \quad (25.1.21)$$

*Proof.* For we have:

$$0 \leq \left( \frac{x}{\sqrt{\varepsilon}} - \varepsilon y \right)^2 = \frac{x^2}{\varepsilon} - 2xy + \varepsilon y^2 \quad (25.1.22)$$

Bringing  $2xy$  to the left side and dividing by 2 completes the proof.  $\square$

**Theorem 25.1.21** (Hölder's Inequality). *If  $a : \mathbb{N} \rightarrow \mathbb{R}$  and  $b : \mathbb{N} \rightarrow \mathbb{R}$  are nonnegative sequences,  $p > 1$ , and if  $p$  and  $q$  are Hölder Conjugates, then:*

$$\sum_{n=1}^{\infty} a_n b_n \leq \left( \sum_{n=1}^{\infty} a_n^p \right)^{1/p} \left( \sum_{n=1}^{\infty} b_n^q \right)^{1/q} \quad (25.1.23)$$

*Proof.* If  $\sum_{n=1}^{\infty} a_n b_n = 0$ , then  $a$  and  $b$  are both the zero sequence and we are done. Suppose the sum is positive. If either  $\sum_{n=1}^{\infty} a_n^p$  or  $\sum_{n=1}^{\infty} b_n^q$  diverges, then the must diverge to  $+\infty$  since  $a$  and  $b$  are non-negative sequences, and we would again be done. Suppose they both converge. Define the following:

$$A = \left( \sum_{n=1}^{\infty} a_n^p \right)^{1/p} \quad (25.1.24a) \qquad B = \left( \sum_{n=1}^{\infty} b_n^q \right)^{1/q} \quad (25.1.24b)$$

Then, by Young's inequality, for all  $n \in \mathbb{N}$ :

$$\frac{a_n b_n}{AB} \leq \frac{1}{p} \left( \frac{a_n}{A} \right)^p + \frac{1}{q} \left( \frac{b_n}{B} \right)^q \quad (25.1.24c)$$

Summing both sides, we have:

$$\frac{1}{AB} \sum_{n=1}^{\infty} a_n b_n \leq \frac{1}{p} + \frac{1}{q} = 1 \quad (25.1.24d)$$

As  $p$  and  $q$  are Hölder Conjugates. Multiplying by  $AB$  proves the result.  $\square$

When  $p = q = 2$  this is often called the Cauchy-Schwartz inequality. That is,  $|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \|\mathbf{b}\|$ . It holds for the integrals of continuous functions, as well as for sequences.

**Theorem 25.1.22** (Minkowski's Inequality). *If  $a : \mathbb{N} \rightarrow \mathbb{R}$  and  $b : \mathbb{N} \rightarrow \mathbb{R}$  are non-negative sequences, and if  $p > 1$ , then:*

$$\left( \sum_{n=1}^{\infty} (a_n + b_n) \right)^{1/p} \leq \left( \sum_{n=1}^{\infty} a_n^p \right)^{1/p} + \left( \sum_{n=1}^{\infty} b_n^p \right)^{1/p}$$

## 25.2 Notes from Rosenlicht

### 25.2.1 Sets

Give a function  $f : X \rightarrow Y$ , the distinction between the image of a subset  $S \subseteq X$  and a point  $x \in X$  is:

$$f(\{x\}) = \{f(x)\} \quad (25.2.1)$$

Similarly for the pre-image:

$$\{f^{-1}(y)\} = f^{-1}(\{y\}) \quad (25.2.2)$$

One definition of an infinite set is that it contains a bijection between itself and a proper subset. Such sets are called Dedekind infinite, and countable choice is needed here. The following are true:

$$(A^C)^C = A \quad (25.2.3)$$

$$A \cup A = A \cap A = A \cup \emptyset = A \quad (25.2.4)$$

$$A \cap \emptyset = \emptyset \quad (25.2.5)$$

$$A \times \emptyset = \emptyset \quad (25.2.6)$$

In addition, there are De Morgan's laws and the distributive laws. Some more identities:

$$(A \setminus B) \cap C = (A \cap C) \setminus B \quad (25.2.7)$$

$$(A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A) \quad (25.2.8)$$

$$(A \setminus (B \setminus C)) = (A \setminus B) \cup (A \cap B \cap C) \quad (25.2.9)$$

$$(A \setminus B) \times C = (A \times C) \setminus (B \times C) \quad (25.2.10)$$

Given any collection of sets  $X_i$ ,  $i \in I$ , and a set  $B$ , we have:

$$B \cap \left( \bigcup_{i \in I} A_i \right) = \bigcup_{i \in I} (B \cap A_i) \quad (25.2.11)$$

Composition is a commutative operation. That is, given  $f : X \rightarrow Y$ ,  $g : Y \rightarrow Z$ , and  $h : Z \rightarrow W$ , we have:

$$h \circ (g \circ f) = (h \circ g) \circ f \quad (25.2.12)$$

The following is also true of functions:

$$f(A \cup B) = f(A) \cup f(B) \quad (25.2.13a)$$

$$f(A \cap B) \subseteq f(A) \cap f(B) \quad (25.2.13b)$$

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B) \quad (25.2.13c)$$

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B) \quad (25.2.13d)$$

$$A \subseteq f^{-1}(f(A)) \quad (25.2.13e)$$

$$f(f^{-1}(A)) \subseteq A \quad (25.2.13f)$$

**Theorem 25.2.1.** *If  $f : X \rightarrow Y$  is injective, then:*

$$f^{-1}(f(A)) = A \quad (25.2.14a)$$

$$f(A \cap B) = f(A) \cap f(B) \quad (25.2.14b)$$

**Theorem 25.2.2.** *If  $f : X \rightarrow Y$  is surjective, then:*

$$f(f^{-1}(A)) = A \quad (25.2.15)$$

## 25.2.2 The Real Number System

The real numbers are a set  $\mathbb{R}$  with several properties. These properties make  $\mathbb{R}$  a complete ordered field, and indeed the only complete ordered field. That is, the real numbers are unique up to *isomorphism*. There are two functions  $+, \cdot : \mathbb{R}^2 \rightarrow \mathbb{R}$ , called addition and multiplication, respectively, that satisfy the following *field axioms*:

$$a + b = b + a \qquad a \cdot b = b \cdot a \quad (\text{Commutativity})$$

$$a + (b + c) = (a + b) + c \qquad a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad (\text{Associativity})$$

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad (\text{Distributive Law})$$

$$\exists_{0 \in \mathbb{R}} : 0 + a = a \qquad \exists_{1 \in \mathbb{R}} : a \cdot 1 = a \quad (\text{Neutral Elements})$$

$$\forall_{a \in \mathbb{R}} \exists_{b \in \mathbb{R}} : a + b = 0 \qquad \forall_{a \in \mathbb{R}, a \neq 0} \exists_{a^{-1}} : a \cdot a^{-1} = 1 \quad (\text{Inverse Elements})$$

By inductively using the associative laws and the commutative laws, we see that adding  $n$  elements does not depend on the order in which they are added. Similarly for multiplication. For a general field, we write  $(F, +, \cdot)$ .

**Theorem 25.2.3.** *If  $(F, +, \cdot)$  is a field, and if  $a \in F$ , then the additive inverse of  $a$  is unique.*

*Proof.* For suppose  $b$  and  $b'$  are additive inverses. Then:

$$b = b + 0 = b + (a + b') = (b + a) + b' = 0 + b' = b' \quad (25.2.16)$$

And therefore  $b$  is unique.  $\square$

We denote the additive inverse of an element  $a$  by writing  $-a$ .

**Theorem 25.2.4.** *If  $(F, +, \cdot)$  is a field, if  $a, b \in F$ , then there is a unique  $x \in F$  such that  $x + a = b$ .*

*Proof.* For let  $x = a - b$ . Then:

$$x + a = (b - a) + a = b + (-a + a) = b + 0 = b \quad (25.2.17)$$

Moreover, of  $x'$  is a solution, then:

$$x' = x' + 0 = x' + (a + (-a)) = (x' + a) + (-a) = b + (-a) = x \quad (25.2.18)$$

Thus,  $x' = x$ .  $\square$

Instead of writing  $b + (-a)$ , we denote this by  $b - a$ . This new operation is called subtraction. Note that it is not commutative, nor is it associative. Indeed, for any  $a, b \in \mathbb{R}$ , suppose  $a - b = b - a$ , and let  $y = a - b$ . Then we have that  $y = -y$ , and thus  $y + y = 2y = 0$ . This is only possible in  $\mathbb{R}$  if  $y = 0$ , and thus we'd require that  $a = b$ . So subtraction is not commutative in  $\mathbb{R}$ . There are fields such that  $y + y = 0$  and such that  $y \neq 0$ , but such fields can't have a notion of *order* on them. We'll discuss these later. Note that the notion is not associative either. Again, let  $a = 2$  and  $b = c = 1$ . Then  $a - (b - c) = 2$ , but  $(a - b) - c = 0$ . Again we come to the conclusion that either  $2 = 0$ , or subtraction is not associative. In an ordered field, which is what  $\mathbb{R}$  is, we cannot have  $2 = 0$ . In finite fields, this is possible.

**Theorem 25.2.5.** *If  $(F, +, \cdot)$  is a field and if  $a \in F$  is non-zero, then the multiplicative inverse of  $a$  is unique.*

*Proof.* For suppose  $b$  and  $b'$  are multiplicative inverses of  $a$ . Then:

$$b = b \cdot 1 = b \cdot (a \cdot b') = (b \cdot a) \cdot b' = 1 \cdot b' = b' \quad (25.2.19)$$

And therefore  $b$  is unique.  $\square$

We write the multiplicative inverse of a non-zero element by  $a^{-1}$ .

**Theorem 25.2.6.** *If  $(F, +, \cdot)$  is a field, if  $a, b \in F$ , and if  $a \neq 0$ , then there is a unique  $x \in F$  such that  $x \cdot a = b$ .*

*Proof.* For let  $x = b \cdot a^{-1}$ . Then:

$$x \cdot a = (b \cdot a^{-1}) \cdot a = b \cdot (a^{-1} \cdot a) = b \cdot 1 = b \quad (25.2.20)$$

Moreoever, if  $x'$  is a solution, then:

$$x' = x' \cdot 1 = x' \cdot (a \cdot a^{-1}) = (x' \cdot a) \cdot a^{-1} = b \cdot a^{-1} = x \quad (25.2.21)$$

Thus,  $x' = x$ .  $\square$

We define division by non-zero numbers by writing  $\frac{a}{b} = a \cdot b^{-1}$ . Other symbols are used for this, like  $a \div b$ , or simply  $a/b$ . Similar to subtraction, division is neither commutative nor associative.

**Theorem 25.2.7.** *If  $(F, +, \cdot)$  is a field, if  $a, b, c \in F$ , and if  $a + c = b + c$ , then  $a = b$ .*

*Proof.* For:

$$a = a + 0 = a + (c - c) = (a + c) - c = (b + c) - c = b + (c - c) = b \quad (25.2.22)$$

Therefore, etc.  $\square$

**Theorem 25.2.8.** *If  $(F, +, \cdot)$  is a field,  $a, b, c \in F$ , if  $c \neq 0$ , and if  $a \cdot c = b \cdot c$ , then  $a = b$ .*

*Proof.* For:

$$a = a \cdot 1 = a \cdot (c \cdot c^{-1}) = (a \cdot c) \cdot c^{-1} = (b \cdot c) \cdot c^{-1} = b \cdot (c \cdot c^{-1}) = b \cdot 1 = b \quad (25.2.23)$$

Therefore, etc.  $\square$

**Theorem 25.2.9.** *If  $(F, \cdot, +)$  is a field, and if  $a \in F$ , then  $a \cdot 0 = 0$ .*

*Proof.* For:

$$a \cdot 0 + a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 = a \cdot 0 + 0 \quad (25.2.24)$$

And therefore from the cancellation laws,  $a \cdot 0 = 0$ .  $\square$

**Theorem 25.2.10.** *If  $(F, +, \cdot)$  is a field, and  $a \in F$ , then  $-a = (-1) \cdot a$*

*Proof.* For:

$$(-1) \cdot a + a = (-1 + 1) \cdot a = 0 \cdot a = 0 \quad (25.2.25)$$

From the uniqueness of inverses,  $-a = (-1) \cdot a$ .  $\square$

**Theorem 25.2.11.** *If  $(F, +, \cdot)$  is a field and  $a \in F$ , then  $-(-a) = a$ .*

*Proof.* For:

$$-(-a) + (-a) = (-1) \cdot (-a) + (-a) = (-1 + 1) \cdot (-a) = 0 \cdot (-a) = 0 \quad (25.2.26)$$

From the uniqueness of inverses, etc.  $\square$

**Theorem 25.2.12.** If  $(F, +, \cdot)$  is a field, and if  $a, b \in F$  are non-zero, then  $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$ .

*Proof.* For:

$$(a \cdot b) \cdot (b^{-1} \cdot a^{-1}) = a \cdot (b \cdot b^{-1}) \cdot a^{-1} = a \cdot 1 \cdot a^{-1} = a \cdot a^{-1} = 1 \quad (25.2.27)$$

From the uniqueness of inverses, etc.  $\square$

**Theorem 25.2.13.** If  $(F, +, \cdot)$  is a field,  $a \in F$  is non-zero, then  $(a^{-1})^{-1} = a$ .

*Proof.* For:

$$(a^{-1})^{-1} \cdot a^{-1} = (a \cdot a^{-1})^{-1} = 1^{-1} = 1 \quad (25.2.28)$$

From uniqueness, etc.  $\square$

**Theorem 25.2.14.** If  $(F, +, \cdot)$  is a field, and  $a, b, c, d \in F$ , and if  $b, d \neq 0$ , then:

$$(a \cdot b^{-1}) + (c \cdot d^{-1}) = (a \cdot d + b \cdot c) \cdot (b \cdot d)^{-1} = \frac{ad + bc}{bd} \quad (25.2.29)$$

As stated before, the axioms of a field are not enough to uniquely define the real numbers. Indeed, the rational numbers  $\mathbb{Q}$  define a field, as do the complex numbers  $\mathbb{C}$ . To see a finite field, consider the set  $\mathbb{F}_2 = \{0, 1\}$ , and consider the following arithmetic:

+	0	1
0	0	1
1	1	0

Table 25.1: Addition in  $\mathbb{F}_2$

*	0	1
0	0	0
1	0	1

Table 25.2: Multiplication in  $\mathbb{F}_2$

Then  $(F, +, \cdot)$  is a field. It's a very strange field, since we have  $1 + 1 = 0$ , but alas it satisfies all of the properties of a field, and all of the theorem's we have proved still apply. Interesting, it is the only field with two elements. We have no choice in deciding what  $a \cdot b$  means in the field, since multiplication by zero must give zero, and multiplication by one must give back the original number. Similarly for addition. Adding zero must not change anything, and so all we are left with is deciding what  $1 + 1$  equals. But to be a field, there must be an additive inverse element. Thus we are forced to set  $1 + 1 = 0$ . There is also a field with three elements. For let  $\mathbb{F}_3 = \{0, 1, 2\}$  and define:

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Table 25.3: Addition in  $\mathbb{F}_3$ 

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Table 25.4: Multiplication in  $\mathbb{F}_3$ 

Intuition tells us that  $1 + 1 > 1 > 0$ , and thus  $1 + 1$  cannot be equal to zero. Thus, to exclude finite fields we need to introduce the notion of order.

- There is a subset  $\mathbb{R}^+$  of  $\mathbb{R}$  such that, for all  $a, b \in \mathbb{R}^+$ , we have  $a \cdot b \in \mathbb{R}^+$  and  $a + b \in \mathbb{R}^+$ .
- For all  $a \in \mathbb{R}$ , one and only one of the following statements is true:
  - $a \in \mathbb{R}^+$
  - $a = 0$
  - $-a \in \mathbb{R}^+$

$\mathbb{R}^+$  is called the set of positive numbers, and the elements such that  $-a \in \mathbb{R}^+$  are called negative. We define less than by writing  $a < b$  if  $b - a \in \mathbb{R}^+$ . Similarly, we define greater than by writing  $a > b$  is  $a - b \in \mathbb{R}^+$ . The less than or equal to and greater than or equal to symbols, denoted  $\leq$  and  $\geq$ , respectively, are such that  $a \leq b$  if  $a < b$  or  $a = b$ , and similarly  $a \geq b$  if  $a > b$  or  $a = b$ . This defines  $\mathbb{R}$  to be an ordered field.

**Theorem 25.2.15.** *If  $a, b \in \mathbb{R}$ , then either  $a = b$ ,  $a < b$ , or  $a > b$ .*

*Proof.* For either  $a - b \in \mathbb{R}^+$ ,  $a - b = 0$ , or  $-(a - b) \in \mathbb{R}^+$ . If  $a - b \in \mathbb{R}^+$ , then  $a > b$ . If  $a - b = 0$ , then  $a = b$ . Finally, if  $-(a - b) \in \mathbb{R}^+$ , then  $b - a \in \mathbb{R}^+$ , and thus  $b > a$ .  $\square$

**Theorem 25.2.16.** *If  $a, b, c \in \mathbb{R}$ , if  $a < b$ , and if  $b < c$ , then  $a < c$ .*

*Proof.* For if  $a < b$ , then  $b - a \in \mathbb{R}^+$ . But if  $b < c$ , then  $c - b \in \mathbb{R}^+$ . But then:

$$c - a = (c - b) + (b - a) \in \mathbb{R}^+ \quad (25.2.30)$$

Therefore, etc.  $\square$

**Theorem 25.2.17.** *If  $a, b, c, d \in \mathbb{R}$ , if  $a < b$ , and if  $c \leq d$ , then  $a + c < b + d$ .*

*Proof.* If  $a < b$ , then  $b - a \in \mathbb{R}^+$ . If  $c \leq d$ , then either  $d - c \in \mathbb{R}^+$ , or  $d - c = 0$ . Thus:

$$(b + d) - (a + c) = (b - a) + (d - c) \in \mathbb{R}^+ \quad (25.2.31)$$

Therefore, etc.  $\square$

**Theorem 25.2.18.** *If  $a, b, c, d \in \mathbb{R}^+$ , if  $a < b$ , and if  $c \leq d$ , then  $a \cdot c < b \cdot d$ .*

*Proof.* For if  $a < b$ , then  $b - a \in \mathbb{R}^+$ . But if  $c \leq d$ , then  $d - c \in \mathbb{R}^+$ , or  $d - c = 0$ . But then  $b \cdot c - a \cdot c = c \cdot (b - a) \in \mathbb{R}^+$ . Similarly,  $a \cdot d - a \cdot c = a \cdot (d - c)$ , and thus this is either positive or zero. Therefore:

$$bd - ac = (bd - ad) + (ad - ac) = d(b - d) + a(d - c) \in \mathbb{R}^+ \quad (25.2.32)$$

□

**Theorem 25.2.19.** *If  $a, b \in \mathbb{R}$  are negative, then  $a + b$  is negative.*

*Proof.* For if  $a$  and  $b$  are negative, then  $-a$  and  $-b$  are positive. But then  $(-a) + (-b) \in \mathbb{R}^+$ . But:

$$(-a) + (-b) = (-1) \cdot a + (-1) \cdot b = (-1) \cdot (a + b) = -(a + b) \in \mathbb{R}^+ \quad (25.2.33)$$

Thus,  $a + b$  is negative. □

**Theorem 25.2.20.** *If  $a, b \in \mathbb{R}$ , if  $a$  is positive, and if  $b$  is negative, then  $a \cdot b$  is negative.*

*Proof.* For if  $b$  is negative, then  $-b$  is positive, and thus:

$$-(a \cdot b) = a \cdot (-b) \in \mathbb{R}^+ \quad (25.2.34)$$

Thus,  $a \cdot b$  is negative. □

**Theorem 25.2.21.** *If  $a, b \in \mathbb{R}$  are negative, then  $a \cdot b$  is positive.*

*Proof.* For if  $a$  and  $b$  are negative, then  $-a, -b \in \mathbb{R}^+$ . But then:

$$a \cdot b = 1 \cdot (a \cdot b) = ((-1) \cdot (-1)) \cdot (a \cdot b) = (-a) \cdot (-b) \in \mathbb{R}^+ \quad (25.2.35)$$

And thus  $a \cdot b$  is positive. □

**Theorem 25.2.22.** *If  $a \in \mathbb{R}$ , then  $a^2 \geq 0$ .*

*Proof.* For if  $a$  is positive, then  $a \cdot a$  is positive. If  $a$  is zero, then  $a \cdot a = 0$ . Finally, from the previous theorem, the product of two negative numbers is positive, and therefore if  $a$  is negative, then  $a \cdot a$  is positive. □

From this we have that  $1 = 1^1 > 0$ . This generalized to the sum of any number of squares.

**Theorem 25.2.23.** *If  $a > 0$ , then  $a^{-1} > 0$ .*

*Proof.* Suppose not. Then either  $a^{-1}$  is negative or it is zero. But it is not zero, for zero has no multiplicative inverse, and  $a$  is an inverse of  $a^{-1}$ . Thus  $a^{-1}$  is negative. But  $a \cdot a^{-1} = 1 > 0$ , a contradiction. Therefore,  $a^{-1}$  is positive. □

**Theorem 25.2.24.** *If  $0 < a < b$ , then  $0 < b^{-1} < a^{-1}$ .*

*Proof.* For:

$$0 < a < b \implies 0 < a \cdot (a^{-1}b^{-1}) < b \cdot (a^{-1}b^{-1}) \implies 0 < b^{-1} < a^{-1} \quad (25.2.36)$$

□

**Theorem 25.2.25.** *If  $a < b < 0$ , then  $b^{-1} < a^{-1}$ .*

*Proof.* For if  $a < b < 0$ , then  $0 < b - a$  and  $0 < a \cdot b$ . But then  $0 < a^{-1} \cdot b^{-1}$ . Thus:

$$0 < (b - a) \cdot a^{-1}b^{-1} = a^{-1} - b^{-1} \quad (25.2.37)$$

And thus  $b^{-1} < a^{-1}$ .

□

**Theorem 25.2.26.** *If  $a, b, c \in \mathbb{R}$ , then  $-(a - b) = b - a$ .*

*Proof.* For:

$$(b - a) + (a - b) = b + (-a + a) - b = b + 0 - b = b - b = 0 \quad (25.2.38)$$

From the uniqueness of inverses, etc.

□

**Theorem 25.2.27.** *If  $a, b, c, d \in \mathbb{R}$ , then:*

$$(a - b) \cdot (c - d) = (a \cdot c + b \cdot d) - (a \cdot d + b \cdot c) \quad (25.2.39)$$

*Proof.* For:

$$(a - b) \cdot (c - d) = a \cdot (c - d) - b \cdot (c - d) \quad (25.2.40a)$$

$$= (a \cdot c - a \cdot d) - (b \cdot c - b \cdot d) \quad (25.2.40b)$$

$$= (a \cdot c - a \cdot d) + (b \cdot d - b \cdot c) \quad (25.2.40c)$$

$$= (a \cdot c + b \cdot d) - (a \cdot d + b \cdot c) \quad (25.2.40d)$$

Therefore, etc.

□

We thus have a way to distinguish  $\mathbb{R}$  from finite fields. We define the natural numbers to be  $2 = 1+1$ ,  $3 = 2+1$ ,  $4 = 3+1$ , and so on. Order also excludes the complex numbers,  $\mathbb{C}$ , since the complex numbers are not ordered. However, the rational numbers,  $\mathbb{Q}$ , still satisfy all of these properties and are too an ordered field. We need another property to distinguish  $\mathbb{Q}$  from  $\mathbb{R}$ . First, a discussion of exponentiation and the absolute value function. Given a positive integer  $n$ , we define the exponentiation of a real number  $r$  by  $r^n = r \cdots r$ , where multiplication is carried out  $n$  times. From this, we get:

$$a^n \cdot a^m = a^{n+m} \quad (25.2.41a)$$

$$(a^m)^n = a^{mn} \quad (25.2.41b)$$

$$(ab)^n = a^n b^n \quad (25.2.41c)$$

The absolute value of a real number is defined as:

$$|a| = \begin{cases} a, & a \geq 0 \\ -a, & a < 0 \end{cases} \quad (25.2.42)$$

**Theorem 25.2.28.** *If  $a \in \mathbb{R}$ , then  $|a| \geq 0$ .*

**Theorem 25.2.29.** *If  $a, b \in \mathbb{R}$ , then  $|a \cdot b| = |a| \cdot |b|$ .*

**Theorem 25.2.30.** *If  $a \in \mathbb{R}$ , then  $a^2 = |a|^2$ .*

### Theorem 25.2.31: Triangle Inequality

---

f  $a, b \in \mathbb{R}$ , then  $|a + b| \leq |a| + |b|$ . ■

### Theorem 25.2.32: Reverse Triangle Inequality

---

f  $a, b \in \mathbb{R}$ , then  $|a - b| \geq ||a| - |b||$ . ■

**Theorem 25.2.33.** *If  $a, b \in \mathbb{R}$ , then:*

$$\max\{a, b\} = \frac{a + b + |a - b|}{2} \quad (25.2.43)$$

*Proof.* If  $a = b$ , then we are done. If  $a < b$ , then  $|a - b| = b - a$ , and thus:

$$\frac{a + b + |a - b|}{2} = \frac{a + b + b - a}{2} = b \quad (25.2.44)$$

And this is the max of  $a$  and  $b$ . similarly if  $b < a$ . □

**Theorem 25.2.34.** *If  $a, b \in \mathbb{R}$ , then:*

$$\min\{a, b\} = \frac{a + b - |a - b|}{2} \quad (25.2.45)$$

*Proof.* For if  $a = b$ , then we are done. If  $a < b$ , then  $|a - b| = b - a$ , and thus:

$$\frac{a + b - |a - b|}{2} = \frac{a + b - (b - a)}{2} = a \quad (25.2.46)$$

And this is the minimum of  $a$  and  $b$ . Similarly for  $b < a$ . □

**Theorem 25.2.35.** *If  $a, b, x, y \in \mathbb{R}$ , if  $a < x < b$  and if  $a < y < b$  then:*

$$|x - y| < b - a \quad (25.2.47)$$

*Proof.* For:

$$a - b = -(b - a) < x - b < x - y < b - y < ba \quad (25.2.48)$$

And therefore:

$$-(b - a) < x - y < b - a \quad (25.2.49)$$

Therefore, etc.  $\square$

Note that  $|x - a| < \varepsilon$  implies that  $\varepsilon - a < x < \varepsilon + a$ . Thus, the solution set to this inequality is all of the points that lie in the interval  $(a - \varepsilon, a + \varepsilon)$ . Now, to separate  $\mathbb{R}$  from  $\mathbb{Q}$  we need to introduce the idea of *completeness*. We will do this in the form of the Least Upper Bound axiom.

**Definition 25.2.1** An upper bound for a subset  $S \subseteq \mathbb{R}$  is a real number  $r$  such that, for all  $x \in S$ , we have  $x \leq r$ .

A bounded above subset is a subset with an upper bound.

**Definition 25.2.2** A least upper bound for a subset  $S \subseteq \mathbb{R}$  is a real number  $r$  such that  $r$  is an upper bound for  $S$ , and for all upper bounds  $s$ , we have  $r \leq s$ .

From this definition we have that least upper bounds are unique for a given bounded above set.

**Theorem 25.2.36.** *If  $S$  is a subset of  $\mathbb{R}$ , if  $s$  is a least upper bound of  $S$ , and if  $x \in \mathbb{R}$  is such that  $x < s$ , then there is a  $y \in S$  such that  $x < y$ .*

*Proof.* For suppose not. Then  $x$  is an upper bound of  $S$ , a contradiction as  $s$  is the least upper bound.  $\square$

Any non-empty finite subset will have a least upper bound. Infinite subsets need not have a least upper bound, and indeed  $\mathbb{R}$  does not have one. If the least upper bound of  $S$  exists, it may not belong to  $S$ . For example, the set of all negative numbers has zero as its least upper bound, but zero is not a negative number. The real numbers satisfy the following property:

1. For any non-empty set of real numbers that is bounded from above, there is a least upper bound.

This axiom distinguishes the rational numbers from the real numbers. That is, there are bounded above subsets of  $\mathbb{Q}$  with no least upper bound. We can justify the least upper bound axiom by considering the decimal expansion of real numbers. That is, we write out  $x = n + 0.x_1x_2x_3\dots$  where  $n$  is an integer, and  $x_i$  is an integer between zero and nine. If  $S$  is bounded above, then there is a least integer  $n$  such that, for all  $x \in S$ ,  $x \leq n$ . This simply comes from the Archimedean principle and the well-ordering principle of the real numbers. But

then there is a least  $x_1$  such that  $x_1$  is an integer between zero and nine and such that, for all  $x \in S$ ,  $x \leq n.x_1$  where this indicates the usual representation of  $n + x_1 \times 10^{-1}$ . We can continue on for  $x_2$  and so on, and this decimal expansion will represent the least upper bound of  $S$ . The least upper bound of a set  $S$  is often denoted  $\sup S$ , where  $\sup$  denotes the latin word *supremum*. Similarly, the greatest lower bound of a set is denoted  $\inf S$ , where  $\inf$  stands for *infinum*.

**Theorem 25.2.37.** *If  $S \subseteq \mathbb{R}$  is bounded from below, then there exists a greatest lower bound of  $S$ .*

*Proof.* For if  $S$  is bounded below, then  $-S = \{-x : x \in S\}$  is bounded from above. But sets that are bounded above have a least upper bound. Let  $s$  be the least upper bound of  $-S$ . Then  $-s$  is the greatest lower bound of  $S$ . Therefore, etc.  $\square$

The real numbers have the property that any real number can be approximated arbitrarily well by a rational number. The rational numbers, however, have certain gaps that are filled in by the real numbers. In a sense, the real numbers are *complete* whereas the rational numbers are not.

**Theorem 25.2.38: The Archimedean Property**

f  $x$  is a real number, then there is an integer  $n$  such that  $x < n$ .  $\blacksquare$

*Proof.* For suppose not. Then there is an  $x \in \mathbb{R}$  such that, for all  $n \in \mathbb{N}$ ,  $n \leq x$ . But then  $\mathbb{N}$  is bounded above, and then there exists a least upper bound. Let  $s$  be such a bound. But if  $s$  is a bound, then for all  $n \in \mathbb{N}$ ,  $n \leq s$ . But if  $n \in \mathbb{N}$ , then  $n + 1 \in \mathbb{N}$  and thus  $n + 1 \leq s$ . But then, for all  $n \in \mathbb{N}$ ,  $n \leq s - 1$ , a contradiction as  $s$  is a least upper bounded, and  $s - 1 < s$ . Therefore, etc.  $\square$

**Theorem 25.2.39.** *If  $\varepsilon > 0$ , then there is an  $n \in \mathbb{N}$  such that  $n^{-1} < \varepsilon$ .*

*Proof.* Since  $\varepsilon > 0$   $\varepsilon^{-1}$  is well defined and positive. But then there is an  $n \in \mathbb{N}$  such that  $n > \varepsilon^{-1}$ . But then  $n^{-1} < \varepsilon$ . Therefore, etc.  $\square$

**Theorem 25.2.40.** *If  $x \in \mathbb{R}$ , then there is an integer  $n \in \mathbb{Z}$  such that  $n \leq x < n + 1$ .*

*Proof.* For if  $x \in \mathbb{R}^+$ , there is an  $N \in \mathbb{N}$  such that  $x < N$ . But then, from the well-ordering of  $\mathbb{N}$ , there is a least  $k \in \mathbb{N}$  such that  $x < k$ . Let  $n = k - 1$ . But then  $n \in \mathbb{Z}$  and  $n \leq x < n + 1$ . If  $-x \in \mathbb{R}^+$ , negate this and repeat the process. If  $x = 0$ , let  $n = 0$ .  $\square$

**Theorem 25.2.41.** *If  $x \in \mathbb{R}$  and  $N \in \mathbb{N}$ , and there is an  $n \in \mathbb{Z}$  such that:*

$$\frac{n}{N} \leq x < \frac{n+1}{N} \quad (25.2.50)$$

*Proof.* For let  $y = N \cdot x$ . Then there is an  $n \in \mathbb{Z}$  such that  $n \leq N \cdot x < n + 1$ . Dividing by  $N$  proves the result.  $\square$

**Theorem 25.2.42.** *If  $\varepsilon > 0$  and  $r \in \mathbb{R}$ , then there is a  $q \in \mathbb{Q}$  such that  $|r - q| < \varepsilon$ .*

*Proof.* For let  $\varepsilon > 0$ . Then there is an  $N \in \mathbb{N}$  such that  $N^{-1} < \varepsilon$ . But then there is an  $n \in \mathbb{Z}$  such that  $n \leq N \cdot r < n + 1$ . Let  $q = n \cdot N^{-1}$ . Then  $|q - r| < \varepsilon$ .  $\square$

This final theorem shows that any real number can be approximated arbitrarily well by any rational number, as was claimed. Let's return to the discussion of the decimal expansion of real numbers. First, we consider finite decimals. Let  $n \in \mathbb{N}$  and let  $a_1, \dots, a_n$  be a sequence of integers between zero and nine. Let  $a_0$  be any integer. If  $m < n$ , then:

$$\begin{aligned} a_0.a_1a_2 \dots a_m &\leq a_0.a_1a_2 \dots a_ma_{m+1} \dots a_n \\ &\leq a_0.a_1a_2 \dots a_m + 9 \times 10^{-(m+1)} + \dots + 9 \times 10^{-n} \end{aligned} \quad (25.2.51)$$

If we add  $10^{-n}$ , this reduces to the following:

$$a_0.a_1a_2 \dots a_m \leq a_0.a_1a_2 \dots a_n \leq a_0.a_1a_2 \dots a_m + 10^{-m} \quad (25.2.52)$$

We can thus view an *infinite decimal* as a sequence  $a : \mathbb{N} \rightarrow \mathbb{Z}$  such that  $a_1 \in \mathbb{Z}$ , and for all  $k > 1$ ,  $a_k$  is an integer between zero and nine. Using the decimal expansion we can find real numbers that are not rational. For let  $x = 0.101001000100001000001\dots$ . This can't be rational since  $Nx$  is not an integer for any positive integer  $N$ . Another classic example of a real number that is not rational is  $\sqrt{2}$ .

**Theorem 25.2.43.** *If  $r > 0$ , then there is a unique number  $a > 0$ , called the square root of  $r$ , such that  $a^2 = r$ .*

*Proof.* For uniqueness, first note that if  $0 < a < b$ , then  $0 < a^2 < b^2$ , and thus any positive real number can have, at most, one positive square root. Define the following:

$$S = \{x \in \mathbb{R}^+ : x^2 \leq r\} \quad (25.2.53)$$

Then  $S$  is bounded above, since  $\max\{1, r\}$  is such a bound. Let  $a$  be the least upper bound of  $S$ . First, note that  $s > 0$  since:

$$(\min\{1, r\})^2 \leq \min\{1, r\} \cdot 1 = \min\{1, r\} \leq r \quad (25.2.54)$$

And therefore,  $\min\{1, r\} \leq s$ . Given  $\varepsilon > 0$ , we have:

$$(s - \varepsilon)^2 < r < (s + \varepsilon)^2 \quad (25.2.55)$$

And therefore:

$$|s^2 - r| < 4s\varepsilon \quad (25.2.56)$$

But  $\varepsilon$  is arbitrary, and thus this difference is zero. Therefore  $r = s^2$ .  $\square$

The value  $s$  is called the square root of  $r$ , and we denote it by  $s = \sqrt{r}$ . Note that, for any positive real number, there are two square roots:  $\pm\sqrt{r}$ . When we write  $\sqrt{r}$ , we mean the positive value. This theorem shows that positive real numbers are the squares of non-zero real numbers. Thus, the set  $\mathbb{R}^+$  described earlier is unique, further justifying the use of this set to order the real numbers. The real numbers are the only arithmetic system, up to isomorphism, that is a complete ordered field. Here, complete means that the least upper bound axiom holds. If  $(\mathbb{R}, +, \cdot)$  and  $(\mathbb{R}', +', \cdot')$  are complete ordered fields, we may as well consider them to be the exact same object. They are essentially a relabelling of each other.

### Example 25.2.1: F

---

nd the greatest lower bound and least upper bound of the set:

$$A = \left\{ \frac{1}{n} : n \in \mathbb{N} \right\} \quad (25.2.57)$$

The least upper bound is 1, since for all  $n \geq 1$ , we have  $1 \leq n^{-1}$ . There is no bound less, since  $1 \in A$ . The greatest lower bound is zero. It is indeed a bound, since for all  $n > 0$ ,  $n^{-1} > 0$ . Moreover, if  $s > 0$ , there is an  $N \in \mathbb{N}$  such that  $N^{-1} < s$ , and thus  $s$  cannot be a lower bound. Consider the set:

$$B = \left\{ \frac{1}{3}, \frac{4}{9}, \frac{13}{27}, \frac{40}{81}, \dots \right\} \quad (25.2.58)$$

The denominator's of this set are powers of three, and the numerators are sums of powers of three. That is, we can write:

$$B = \left\{ \frac{1}{3^n} \sum_{k=0}^{n-1} 3^k : n \in \mathbb{N} \right\} \quad (25.2.59)$$

We can use the geometric series to simplify the sum, noting that:

$$\frac{1}{3^n} \sum_{k=0}^{n-1} 3^k = \frac{1}{3^n} \frac{1 - 3^n}{1 - 3} = \frac{3^n - 1}{2 \cdot 3^n} \quad (25.2.60)$$

Splitting this into two parts, we get:

$$\frac{1}{3^n} \sum_{k=0}^{n-1} 3^k = \frac{1}{2} - \frac{1}{2 \cdot 3^n} \quad (25.2.61)$$

And this decays to zero. Thus, we see that the least upper bound is  $\frac{1}{2}$ , since for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that  $(2 \cdot 3^N)^{-1} < \varepsilon$ , and thus there are elements of the set that are arbitrarily close to  $\frac{1}{2}$ . Moreover,  $\frac{1}{2}$  is an upper

bound, since every element of the set is strictly less than it. Using the final equation, we see that the elements are strictly increasing as  $n$  increasing, and thus the greatest lower bound is simply the first element,  $\frac{1}{3}$ . Lastly, find the greatest lower bound and least upper bound for:

$$C = \{\sqrt{2}, \sqrt{2 + \sqrt{2}}, \sqrt{2 + \sqrt{2 + \sqrt{2}}}, \dots\} \quad (25.2.62)$$

We see that the pattern is:

$$x_{n+1} = \sqrt{2 + \sqrt{x_n}} \quad (25.2.63)$$

Thus, this sequence is strictly increasing as  $n$  increases. From this we know that  $\sqrt{2}$  is the greatest lower bound. Now we must show that the set has a least upper bound. Let  $x$  be the solution to the equation  $x = \sqrt{2 + \sqrt{x}}$ . We know such a solution exists since this equation simplifies to a quartic polynomial with roots. Then:

$$|x - x_{n+1}| = |\sqrt{2 + \sqrt{x}} - \sqrt{2 + \sqrt{x_n}}| \quad (25.2.64)$$

$$= \left| \frac{\sqrt{x} - \sqrt{x_n}}{\sqrt{2 + \sqrt{x}} + \sqrt{2 + \sqrt{x_n}}} \right| \quad (25.2.65)$$

$$< \left| \frac{\sqrt{x} - \sqrt{x_n}}{2\sqrt{2}} \right| \quad (25.2.66)$$

$$= \left| \frac{x - x_n}{2\sqrt{2}(\sqrt{x} + \sqrt{x_n})} \right| \quad (25.2.67)$$

$$(25.2.68)$$

But we note that  $\sqrt{x} + \sqrt{x_n} > 2\sqrt{x_n} > 2\sqrt{2}$ , and obtain:

$$|x - x_{n+1}| < \frac{1}{8}|x - x_n| \quad (25.2.69)$$

Similarly:

$$|x - x_{n+2}| < \frac{1}{8}|x - x_{n+1}| < \frac{1}{8^2}|x - x_n| \quad (25.2.70)$$

By induction:

$$|x - x_{n+k}| < \frac{1}{8^k}|x - x_n| \quad (25.2.71)$$

And this tends to zero, and therefore  $x_n \rightarrow x$ . Thus,  $x$  is the least upper bound. ■

## 25.3 Old Notes

The real line, or real number system, is a complete ordered field. That is, it is complete in the sense that all Cauchy sequences converge, has a total order structure on it, and has a field structure (That of addition, multiplication, subtraction, and division). An open subset of the real line is a set  $S$  such that for all  $x \in S$  there is an  $\varepsilon > 0$  such that  $(x - \varepsilon, x + \varepsilon) \subset S$ . The entire space  $\mathbb{R}$  is open, as is the empty set  $\emptyset$ . The union of an arbitrary collection of open sets is open, and the intersection of finitely many open sets is open. The intersection of infinitely many open sets may not be open, however. A set is closed if its complement is open. The Euclidean plane is the set of all ordered pairs  $(a, b)$ . That is,  $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ . Euclidean space, or 3-space, is  $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ . This is the set of all ordered triplets  $(x, y, z)$ . Similarly,  $n$  dimensional Euclidean space is the set of all  $n$  tuples. This is denoted  $\mathbb{R}^n$ . The distance between two points  $\mathbf{x}$  and  $\mathbf{y}$  is defined by the generalized Pythagorean Theorem:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$$

**Definition 25.3.1** A metric on a set  $X$  is a function  $d : X \times X \rightarrow \mathbb{R}$  such that:

1.  $d(x, y) \geq 0$  for all  $x, y \in X$ .
2.  $d(x, y) = 0$  if and only if  $x = y$ .
3.  $d(x, y) = d(y, x)$  for all  $x, y \in X$ .
4.  $d(x, z) \leq d(x, y) + d(y, z)$  for all  $x, y, z \in X$ .

There are two types of integrals defined for functions of a real variable: Riemann Integration and Lebesgue Integration. Lebesgue integration requires the notion of *measure*.

### 25.3.1 Definitions

**Definition 25.3.2** The tangent line of a differentiable function  $y : \mathbb{R} \rightarrow \mathbb{R}$  at a point  $x_0 \in \mathbb{R}$  is the function  $y_T : \mathbb{R} \rightarrow \mathbb{R}$  defined by  $y_T(x) = y'(x_0)(x - x_0) + y(x_0)$

**Definition 25.3.3** If  $\Gamma(t) = (x(t), y(t))$ , for  $a \leq t \leq b$ , and  $\Gamma'(t) = (x'(t), y'(t))$  exists for  $a < t < b$ , then the length of  $\Gamma$  from  $a$  to  $b$  is:

$$L = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt \quad (25.3.1)$$

**Definition 25.3.4** The dimension of a vector space is the cardinality of any basis of the space.

By the Dimension Theorem, all bases of a vector space have the same cardinality.

### 25.3.2 Theorems

**Theorem 25.3.1** (Mean Value Theorem). *If  $f : (a, b) \rightarrow \mathbb{R}$  is continuous and bounded, and if  $x \in (a, b)$ , then there is a  $c \in (a, x)$  such that  $\int_a^x f = (x-a)f(c)$ .*

**Theorem 25.3.2** (Generalized Fundamental Theorem of Calculus). *If  $\mathcal{U}$  is an open non-empty subset of  $\mathbb{R}$ ,  $a \in \mathcal{U}$ , and if  $f : \mathcal{U} \rightarrow \mathbb{R}$  is bounded and continuous, then  $F : \mathcal{U} \rightarrow \mathbb{R}$  defined by  $F(x) = \int_{\mathcal{U} \cap (a, x)} f$  is differentiable and  $F'(x) = f(x)$*

*Proof.* For let  $x \in \mathcal{U}$ . Let  $\{x_n\}_{n=1}^{\infty} \subset \mathcal{U}$  be a sequence such that  $x_n \rightarrow x$ ,  $x \notin \{x_n\}_{n=1}^{\infty}$ . As  $\mathcal{U}$  is open and  $x \in \mathcal{U}$ , there is an  $\varepsilon > 0$  such that  $B_{\varepsilon}(x) \subset \mathcal{U}$ . But, as  $x_n \rightarrow x$ , there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $x_n \in B_{\varepsilon}(x)$ . But then for all  $n > N$ :

$$\int_{\mathcal{U} \cap (a, x)} f - \int_{\mathcal{U} \cap (a, x_n)} f = \int_{x_n}^x f \quad (25.3.2)$$

But, as  $f$  is continuous, by the mean value theorem for all  $n > N$  there is a  $c_n \in (x_n, x)$  such that  $\int_{x_n}^x f = (x - x_n)f(c_n)$ . But then

$$\left| \frac{\int_{x_n}^x f}{x - x_n} - f(x) \right| = |f(c_n) - f(x)| \quad (25.3.3)$$

But  $c_n \in (x_n, x)$ , and  $x_n \rightarrow x$ , and therefore  $c_n \rightarrow x$ . But  $f$  is continuous, and therefore  $f(c_n) \rightarrow f(x)$ . Therefore, by the definition of the derivative of  $F$  at  $x$ ,  $F'(x) = f(x)$ .  $\square$

**Theorem 25.3.3.** *If  $V$  is a vector space and  $A, B \subset V$  are subspaces, then  $A \cap B$  is a subspace and  $\dim(A \cap B) \leq \min\{\dim(A), \dim(B)\}$*

**Theorem 25.3.4.** *If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is differentiable and  $f'(x) > 0$  for all  $x$ , then  $f$  is strictly increasing.*

**Theorem 25.3.5.** *If  $f : (a, b) \rightarrow \mathbb{R}$  is continuous and  $f(a) < 0 < f(b)$ , then there is a  $c \in (a, b)$  such that  $f(c) = 0$ .*

**Theorem 25.3.6.** *If  $f$  is integrable on  $(a, b)$ , and if  $c \in (a, b)$ , then  $\int_a^b f = \int_a^c f + \int_c^b f$*

### 25.3.3 Metric Spaces

**Definition 25.3.5** A metric space is a set  $X$  with a function  $d : X \times X \rightarrow \mathbb{R}$  with the following properties:

1. For all  $x, y \in X$ ,  $d(x, y) = 0 \Leftrightarrow x = y$ . [Identity of Indiscernables]
2. For all  $x, y, z \in X$ ,  $d(x, y) \leq d(x, z) + d(y, z)$  [Modified Triangle Inequality]

They are denoted  $(X, d)$ .  $d$  is called a *metric* or *distance* function.

**Theorem 25.3.7.** *A metric space  $(X, d)$  has the following properties:*

1.  $d(x, y) = 0 \Leftrightarrow x = y$  [Identity of Indiscernibles]
2.  $d(x, y) = d(y, x)$  [Symmetry]
3.  $d(x, y) \geq 0$  [Positivity]
4.  $d(x, y) \leq d(x, z) + d(z, y)$  [Triangle Inequality]

*Proof.* In order,

1. This is part of the definition.
2. For  $d(x, y) \leq d(x, x) + d(y, x) = d(y, x)$ . But  $d(y, x) \leq d(y, y) + d(x, y) = d(x, y)$ . Thus  $d(x, y) \leq d(y, x)$  and  $d(y, x) \leq d(x, y)$ , and therefore  $d(x, y) = d(y, x)$
3. For  $0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y)$ . Thus,  $0 \leq d(x, y)$
4.  $d(x, y) \leq d(x, z) + d(y, z) = d(x, z) + d(z, y)$

□

It is most common, almost universal, that textbooks state theorem 1.8.1 as the definition of a metric space. However, when proving something is a metric space, it is nicer to prove two things rather than four.

**Theorem 25.3.8.** *If  $V$  is a vector space with a norm and  $d$  is the induced metric, then  $(V, d)$  is a metric space.*

*Proof.* In order,

1.  $\|x - y\| = 0$  if and only if  $x - y = 0$ . Thus  $d(x, y) = 0 \Leftrightarrow x = y$ .
2.  $d(x, y) = \|x - y\| \leq \|x - z\| + \|y - z\| = d(x, z) + d(y, z)$

□

**Definition 25.3.6** If  $(X, d)$  is a metric space,  $x \in X$ , then the open ball of radius  $r > 0$  is  $B_r(x) = \{y \in X : d(x, y) < r\}$ .

**Definition 25.3.7** In a metric space,  $\mathcal{U}$  is metrically open if and only if for all  $x \in \mathcal{U}$  there is an  $r > 0$  such that  $B_r(x) \subset \mathcal{U}$ .

For metric spaces, metrically open and topologically open are the same thing, as we will see.

**Theorem 25.3.9.** *The empty set is open.*

*Proof.* For suppose not. Then there is some  $x \in \emptyset$  such that for all  $r > 0$ ,  $B_r(x) \not\subset \emptyset$ . A contradiction. Therefore, etc.  $\square$

**Theorem 25.3.10.** *The whole space  $X$  is open.*

*Proof.* For let  $x \in X$  and  $r > 0$ . Then  $B_r(x) = \{y \in X : d(x, y) < r\}$ , and thus  $B_r(x) \subset X$ . Therefore, etc.  $\square$

**Theorem 25.3.11.** *If  $\mathcal{U} \subset X$  is open, then it is the union of open balls.*

*Proof.* For let  $\mathcal{U} \subset X$  be open. Then, for all  $x \in \mathcal{U}$  there is a  $r(x) > 0$  such that  $B_{r(x)}(x) \subset \mathcal{U}$ . But then  $\cup_{x \in \mathcal{U}} B_{r(x)}(x) \subset \mathcal{U}$ . But, as for all  $y \in \mathcal{U}$ ,  $y \in \cup_{x \in \mathcal{U}} B_{r(x)}(x)$ ,  $\mathcal{U} \subset \cup_{x \in \mathcal{U}} B_{r(x)}(x)$ . Thus,  $\mathcal{U} = \cup_{x \in \mathcal{U}} B_{r(x)}(x)$ .  $\square$

**Definition 25.3.8** If  $(X, d)$  is a metric space, then the metric space topology is the set  $\tau = \{\mathcal{U} : \mathcal{U} \underset{\text{Open}}{\subset} X\}$

**Theorem 25.3.12.** *The metric space topology is a topology.*

*Proof.* In order,

1.  $\emptyset, X \in \tau$
2. Let  $\mathcal{U}_\alpha$  be a family of open sets and let  $x \in \mathcal{U}_\alpha$  be arbitrary. Then there is an open set  $\mathcal{U} \in \{\mathcal{U}_\alpha : \alpha \in A\}$  such that  $x \in \mathcal{U}$ . But then there is an  $r > 0$  such that  $B_r(x) \subset \mathcal{U}$ . But then  $B_r(x) \subset \cup_{\alpha \in A} \mathcal{U}_\alpha$ .
3. Let  $\mathcal{U}_k, 1 \leq k \leq n$  be open sets, and let  $x \in \cap_{k=1}^n \mathcal{U}_k$ . Then, for each  $\mathcal{U}_k$  there is an  $r_k$  such that  $B_{r_k}(x) \subset \mathcal{U}_k$ . Let  $r = \min\{r_k : 1 \leq k \leq n\}$ . Then  $B_r(x) \in \cap_{k=1}^n \mathcal{U}_k$ .

$\square$

**Theorem 25.3.13.** *If  $(X, d_X)$  and  $(Y, d_Y)$  are metric space, then  $f : X \rightarrow Y$  is a continuous function (With respect to their metric space topologies) if and only if  $\forall \varepsilon > 0, \forall x \in X, \exists \delta > 0 : y \in B_\delta(x) \Rightarrow f(y) \in B_\varepsilon(f(x))$ .*

*Proof.* For let  $x \in X$  and  $\varepsilon > 0$  be given. As  $B_\varepsilon(f(x))$  is open and  $f$  is continuous, the preimage is open. But as  $x \in f^{-1}(B_\varepsilon(f(x)))$ , there is a  $\delta > 0$  such that  $B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x)))$ . Thus, for all  $y \in B_\delta(x)$ ,  $f(y) \in B_\varepsilon(f(x))$ . Now suppose for all  $x \in X$  and for all  $\varepsilon > 0$ , there is a  $\delta > 0$  such that  $y \in B_\delta(x) \Rightarrow f(y) \in B_\varepsilon(f(x))$ . Let  $\mathcal{U}$  be open in  $f(X)$ . If  $f^{-1}(\mathcal{U})$  is empty, we are done. Suppose not. Let  $x \in f^{-1}(\mathcal{U})$ . As  $\mathcal{U}$  is open, there is a  $\varepsilon > 0$  such that  $B_\varepsilon(f(x))$  is open in  $\mathcal{U}$ . But then there is a  $\delta > 0$  such that if  $y \in B_\delta(x)$ , then  $f(y) \in B_\varepsilon(f(x))$ . But then  $f^{-1}(\mathcal{U})$  is open. Therefore, etc.  $\square$

**Definition 25.3.9** If  $S \subset (X, d)$ , then  $x$  is said to be a limit point of  $S$  if and only if for all  $\varepsilon > 0$ ,  $B_\varepsilon(x) \cap S \neq \emptyset$ .

**Definition 25.3.10** If  $S \subset (X, d)$ , then the closure of  $S$ , denoted  $\overline{S}$ , is the set of all limit points of  $S$ .

**Definition 25.3.11** If  $S \subset (X, d)$ , then  $x \in S$  is an interior point if and only if  $\exists r > 0 : B_r(x) \subset S$ .

**Definition 25.3.12** If  $S \subset (X, d)$ , the interior of  $S$ , denoted  $\text{Int}(S)$ , is the set of all interior points.

**Definition 25.3.13** If  $S \subset (X, d)$ , the relative interior of  $S$  is  $\text{ri}(S) = \{x \in S : \exists \varepsilon > 0 : B_\varepsilon(x) \cap \text{aff}(S) \subset S\}$ .

**Definition 25.3.14** The boundary of  $S \subset V$  is  $S \setminus \text{ri}(S)$ .

**Theorem 25.3.14.** A subset  $S$  of a metric space  $(X, d)$  is closed if and only if every limit point of  $S$  is in  $S$ .

*Proof.* For let  $S$  be closed and let  $x$  be a limit point of  $S$ . Suppose  $x \in S^c$ . But  $S^c$  is open, as  $S$  is closed, and thus there is a  $r > 0$  such that  $B_r(x) \subset S^c$ . But then  $B_r(x) \cap S = \emptyset$ , a contradiction. Thus,  $x \in S$ . Now suppose  $S$  contains all of its limit points. Suppose  $S^c$  is not open. Then there is a  $y \in S^c$  such that for all  $r > 0$ ,  $B_r(y) \not\subset S^c$ . Then for all  $r > 0$ ,  $B_r(y) \cap S \neq \emptyset$ . But then  $y \in S$ , as  $S$  contains all of its limit points. Thus  $S^c$  is open, and therefore  $S$  is closed.  $\square$

**Theorem 25.3.15.** Metric spaces, with the metric space topology, are  $T_4$  spaces.

*Proof.* Let  $(X, d)$  be a metric space and let  $\tau$  be the metric space topology.  $(X, \tau)$  is  $T_1$ , for let  $x, y \in X$ ,  $x \neq y$ , and let  $r = \frac{d(x,y)}{2}$ . Then  $x \in B_r(x)$  and  $y \notin B_r(x)$ .  $(X, \tau)$  is normal, for let  $E$  and  $V$  be closed, nonempty, disjoint subsets of  $X$ . As  $V$  is closed, and as  $E$  and  $V$  are disjoint, for all  $x \in E$  there is an  $r(x) > 0$  such that  $B_{r(x)}(x) \cap V = \emptyset$  (Otherwise  $x$  is a limit point of  $V$ , and thus in  $V$ ). Similarly, for all  $y \in V$  there is an  $r(y) > 0$  such that  $B_{r(y)}(y) \cap E = \emptyset$ . Let  $\mathcal{U} = \cup_{x \in E} B_{\frac{r(x)}{4}}(x)$  and  $\mathcal{V} = \cup_{y \in V} B_{\frac{r(y)}{4}}(y)$ . Then  $E \subset \mathcal{U}$

and  $E \subset \mathcal{V}$ , and  $\mathcal{U}$  and  $\mathcal{V}$  are disjoint. For suppose not. Let  $z \in \mathcal{U} \cap \mathcal{V}$ . Then, there is an  $x \in E$  and a  $y \in V$  such that  $d(x, z) \leq \frac{r(x)}{4}$  and  $d(y, z) \leq \frac{r(y)}{4}$ . But then  $d(x, y) \leq d(x, z) + d(y, z) = \frac{r(x)+r(y)}{4} \leq \frac{\max\{r(x), r(y)\}}{2}$ . Thus  $x \in B_{r(y)}(y)$ , or  $y \in B_{r(x)}(x)$ , a contradiction. Therefore, etc.  $\square$

**Definition 25.3.15** A subset  $S \subset (X, d)$  is said to be bounded if and only if  $\exists M \in \mathbb{R} : x, y \in S \Rightarrow d(x, y) \leq M$ .

**Definition 25.3.16** A Cauchy Sequence in a metric space is a sequence  $x_n : \forall \varepsilon > 0, \exists N \in \mathbb{N} : n, m > N \Rightarrow d(x_n, x_m) < \varepsilon$ .

**Theorem 25.3.16.** *Convergence in a metric space is unique.*

*Proof.* As metric spaces are  $T_4$ , they are Hausdorff, and thus limits are unique.  $\square$

**Theorem 25.3.17.** *In a metric space  $(X, d)$ , a sequence  $x_n \rightarrow x$  if and only if  $\forall \varepsilon > 0, \exists N \in \mathbb{N} : n > N \Rightarrow d(x, x_n) < \varepsilon$ .*

*Proof.* For any open set  $\mathcal{U}$ ,  $x \in \mathcal{U}$ , there is an  $N \in \mathbb{N} : n > N \Rightarrow x_n \in \mathcal{U}$ . Let  $\mathcal{U} = B_\varepsilon(x)$ . Then  $n > N \Rightarrow d(x, x_n) < \varepsilon$ . Now, let  $\mathcal{U}$  be open and  $x \in \mathcal{U}$ .  $\exists \varepsilon > 0 : B_\varepsilon(x) \subset \mathcal{U}$ . But  $\exists N \in \mathbb{N} : n > N \Rightarrow d(x, x_n) < \varepsilon$ . Thus,  $n > N \Rightarrow x_n \in \mathcal{U}$ .  $\square$

**Theorem 25.3.18.**  $f : (X, d_X) \rightarrow (Y, d_Y)$  is continuous if and only if for all  $x \in X$ ,  $x_n \rightarrow x \Rightarrow f(x_n) \rightarrow f(x)$ .

*Proof.*  $\forall \varepsilon > 0, \forall x \in X, \exists \delta > 0 : d_X(x, x_0) < \delta \Rightarrow d_Y(f(x), f(x_0)) < \varepsilon$ . Let  $x_n \rightarrow x$ . Then,  $\exists N \in \mathbb{N} : n > N \Rightarrow d_X(x_n, x) < \delta$ . But then  $d_Y(f(x), f(x_n)) < \varepsilon$ . Thus  $f(x_n) \rightarrow f(x)$ . Now suppose  $x_n \rightarrow x \Rightarrow f(x_n) \rightarrow f(x)$  for all such sequences, and suppose  $f$  is discontinuous. Then there is a  $\varepsilon > 0$  such that for all  $n \in \mathbb{N}$ , there is an  $x_{n_k} \in B_{\frac{1}{k}}(x)$  such that  $d_Y(f(x), f(x_{n_k})) > \varepsilon$ . But then  $d_X(x, x_{n_k}) \rightarrow 0$ , and thus  $d_Y(f(x), f(x_n)) \rightarrow 0$ , a contradiction. Therefore, etc.  $\square$

**Definition 25.3.17** A metric space  $(X, d)$  is said to be complete if and only if every Cauchy sequence in  $X$  converges.

**Definition 25.3.18** An inner product space is called a Hilbert Space if and only if it is complete (Induced Metric).

**Definition 25.3.19** A normed space is called a Banach Space if and only if it is complete (Induced Metric).

**Definition 25.3.20** A subset  $S$  of a metric space  $(X, d)$  is said to be sequentially compact if and only if every sequence  $x_n$  in  $S$  has a convergent subsequence whose limit is in  $S$ .

**Definition 25.3.21** If  $x_n$  is a sequence in a metric space  $(X, d)$ , then  $x$  is said to be an accumulation point of  $x_n$  if and only if for all  $\varepsilon > 0$ ,  $B_\varepsilon(x) \cap \{x_n\}_{n=1}^\infty$  is infinite.

**Definition 25.3.22** A subset  $S$  of a metric space  $(X, d)$  is said to be limit point compact if and only if every sequence in  $S$  has an accumulation point in  $S$ .

**Theorem 25.3.19.** *A subset  $S$  of  $(X, d)$  is sequentially compact if and only if it is limit point compact.*

*Proof.* For suppose  $S$  is sequentially compact, and let  $x_n$  be a sequence. As  $S$  is sequentially compact there is a convergent subsequence with a limit in  $S$ . But then this limit is an accumulation point in  $S$ . Now, suppose  $S$  is limit point compact. Let  $x_n$  be a sequence in  $S$ . As  $S$  is limit point compact, there is an accumulation point of  $x_n$ , call it  $x$ . But then for all  $n \in \mathbb{N}$ ,  $B_{\frac{1}{n}}(x) \cap \{x_k\}_{k=1}^\infty \neq \emptyset$ . By the axiom of choice, we may construct a subsequence  $x_{n_k}$  of points contained in each open ball. But then  $d(x_{n_k}, x) \rightarrow 0$ . Thus, there is a convergent subsequence.  $\square$

**Definition 25.3.23** A subset  $S$  of a metric space is said to be totally bounded if and only if for all  $r > 0$  there are finitely many points  $x_k$  such that  $S \subset \bigcup_{k=1}^n B_r(x_k)$ .

**Theorem 25.3.20.** *If  $(X, d)$  is a metric space and  $S \subset X$  is sequentially compact, then it is closed and bounded.*

*Proof.*

Suppose it is unbounded. That is, if  $x \in S$  and  $n \in \mathbb{N}$ , there is a  $y \in S$  such that  $d(x, y) > n$ . Let  $x_n$  be a such a sequence such that  $d(x, x_n) > n$  for all  $n \in \mathbb{N}$  (The existence of such a sequence requires the axiom of choice. My apologies). This sequence has no convergent subsequence, as suppose it does, say  $s \in S$ . But  $d(x, x_n) \leq d(s, x) + d(s, x_n)$ , and thus  $d(x, x_n) - d(s, x) \leq d(s, x_n)$ . Thus  $d(s, x_n) \not\rightarrow 0$ . Thus  $S$  is not unbounded, and is therefore bounded. Suppose it is not closed. Then there is a point  $x$  such that  $x$  is a limit point of  $S$  but  $x \notin S$ . Let  $x_n$  be a sequence that converges to  $S$ . (Such a sequence exists as  $x$  is a limit point, and the axiom of choice). But then  $x_n \rightarrow x$ , and thus  $x \in S$ . Therefore  $S$  is closed.  $\square$

**Theorem 25.3.21.** *Every subset of a totally bounded space is totally bounded.*

*Proof.* For let  $S \subset X$ , and suppose  $X$  is totally bounded. Let  $r > 0$ . As  $X$  is totally bounded, there are finitely many points such that  $\bigcup_{k=1}^n B_{\frac{r}{2}}(x_k)$  contains all of  $X$ . Let  $s_k$  be the points such that  $S \subset \bigcup_{k=1}^m B_{\frac{r}{2}}(s_k)$  and  $S \cap B_{\frac{r}{2}}(s_k) \neq \emptyset$ . If the  $s_k$  are in  $S$ , we are done. Suppose not. As  $s_k \notin S$  and  $B_{\frac{r}{2}}(s_k) \cap S \neq \emptyset$ , there is an  $\ell_k \in S$  such that  $d(\ell_k, s_k) < \frac{r}{2}$ . But then  $S \subset \bigcup_{k=1}^m B_r(\ell_k)$ . Therefore, etc.  $\square$

**Theorem 25.3.22.** *If  $(X, d)$  is a metric space and  $S \subset X$  is sequentially compact, then  $S$  is totally bounded.*

*Proof.* For let  $r > 0$  and suppose that  $S$  is not totally bounded. Let  $s_1 \in S$ . There must be a point  $s_2$  such that  $s_2 \notin B_r(s_1)$ , as  $S$  is not contained inside the entire open disc. Similarly, there is a point  $s_3 \in S$  such that  $s_3 \notin B_r(s_1) \cup B_r(s_2)$ . In this manner we obtain  $s_1, s_2, \dots, s_n, \dots$ , such that  $s_n \in S$  and  $s_n \notin \bigcup_{k=1}^{n-1} B_r(s_k)$ . But as  $s_n \notin B_r(s_{n-1})$ ,  $d(s_n, s_{n-1}) \geq r$ . But as  $s_n$  is a sequence in  $S$ , and as  $S$  is compact, the sequence must have a convergent subsequence whose limit is some  $s \in S$ . Thus the ball  $B_{\frac{r}{2}}(s) \cap \{x_n\}_{n=1}^{\infty}$  is infinite. But then there are points  $s_l$  and  $s_m$  such that  $d(s_l, s_m) < r$ , a contradiction. Thus  $S$  is totally bounded.  $\square$

**Theorem 25.3.23** (Heine-Borel-Lebesgue Theorem). *In a metric space  $(X, d)$ , with the metric space topology, if  $S \subset X$  then the following are equivalent:*

1.  $S$  is compact.
2.  $S$  is complete and totally bounded.
3.  $S$  is sequentially compact.
4.  $S$  is limit point compact.

*Proof.* We have seen that (3) and (4) imply each other. We now show (1)  $\Rightarrow$  (2)  $\Rightarrow$  (3)  $\Rightarrow$  (1).

1. Suppose  $S$  is compact. Let  $r > 0$  be arbitrary. Then  $\bigcup_{x \in S} B_r(x)$  is an open cover of  $S$ , and thus there is a finite subcover. Thus  $S$  is contained in a finite collection of open balls of radius  $r$ . Let  $x_n$  be a Cauchy sequence in  $S$ . Suppose it does not converge. Then, for all  $x \in S$  there is a  $\varepsilon(x) > 0$  such that  $B_{\varepsilon(x)}(x) \cap \{x_n\}_{n=1}^{\infty}$  is finite. But  $\bigcup_{x \in S} B_{\varepsilon(x)}(x)$  is a cover of  $S$ , and thus there is a finite subcover, suppose with  $N$  open sets. But then  $\{x_n\}_{n=1}^{\infty} \subset \bigcup_{k=1}^N B_{\varepsilon(x_k)}(x_k)$ , a contradiction as each  $B_{\varepsilon(x_k)}(x_k)$  contains but finitely many elements of  $\{x_k\}_{n=1}^{\infty}$ , and thus a finite union cannot contain all of  $\{x_n\}_{n=1}^{\infty}$ . Thus,  $x_n$  converges. Therefore  $S$  is complete.
2. Let  $x_n$  be a sequence in  $S$ . As  $S$  is totally bounded, for all  $n \in \mathbb{N}$  there is a finite set of points  $a(n)$  such that  $S \subset \bigcup_{a(n)} B_{\frac{1}{n+1}}(a(n))$ . Thus there is a point  $a(1)$  such that  $B_{1/2}(a(1)) \cap \{x_n\}_{n=1}^{\infty}$  is infinite. As  $B_{1/2}(a(1)) \subset S$ , it is totally bounded. Thus there is a covering of finitely many points of  $B_{1/2}(a(1))$  of radius  $1/3$ . By induction, for all  $n \in \mathbb{N}$  there is a point  $a(n+1)$  such that  $B_{\frac{1}{n+1}}(a(n+1)) \subset B_{\frac{1}{2^n}}(a(n))$ , and  $B_{\frac{1}{2^{n+1}}}(a(n+1)) \cap \{x_n\}_{n=1}^{\infty}$  is infinite. By the axiom of choice, we may choose a subsequence of points  $x_{n_k}$  that lie in each of these sets. But then this set is Cauchy, as for  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that  $n > N$  implies  $\frac{1}{n} < \frac{\varepsilon}{2}$ , and thus

for  $j, k > N$ ,  $d(x_{n_j}, x_{n_k}) \leq d(x_{n_j}, a(n+1)) + d(x_{n_k}, a(n+1)) < \varepsilon$ . But Cauchy sequences converge as  $S$  is complete. Therefore, etc.

3. Let  $\mathcal{O}$  be an open cover and suppose no finite subcover exist. But as  $S$  is sequentially compact, it is totally bounded and thus there are finitely many points such that  $S \subset \bigcup_k B_1(a_k(1))$ . But then one of these open balls must have no finite subcover, as the entirety of  $S$  has no finite subcover. Let  $a(1)$  be the center of such a set. But as  $B_1(a(1)) \cap S \subset S$ , it is totally bounded as well. Thus there are finitely many points such that  $B_1(a(1)) \subset \bigcup_k B_{1/2}(a_k(2))$ , and again there is at least one open ball that has no finite subcover, as  $B_1(a(1))$  has no finite subcover. Inductive, we obtain a sequence of points  $a(n)$  such that  $B_{\frac{1}{n+1}}(a(n+1)) \subset B_{\frac{1}{n}}(a(n))$  and  $B_{\frac{1}{n}}(a(n))$  has no finite subcover of  $\mathcal{O}$ . By the axiom of choice, we may choose a sequence  $a(n)$  of points of in the ball. But as  $S$  is sequentially compact, there is a convergent subsequence  $a(n_k)$  with some limit in  $S$ , call it  $x$ . But as  $x \in S$ ,  $x$  is covered by  $\mathcal{O}$ , and thus there is some open set such that  $x \in U$ . But as  $U$  is open, there is an  $\varepsilon > 0$  such that  $B_r(x) \subset U$ . But as the subsequence converges, there is an  $N \in \mathbb{N}$  such that for all  $k > N$ ,  $d(a(n_k), x) < \varepsilon < \varepsilon$ . But then for any point  $y \in B_{\frac{1}{n_{N+1}}}^{-1}(a(n_{N+1}))$ ,  $d(x, y) \leq d(x, a(n_{N+1})) + d(y, a(n_{N+1}))$ . But then  $B_{\frac{1}{n_{N+1}}} \in \mathcal{U}$ , a contradiction as  $B_{\frac{1}{n_{N+1}}}$  has no finite subcover. Thus,  $S$  is compact.

□

**Theorem 25.3.24** (Heine-Borel Theorem). *A set  $S \subset \mathbb{R}$  is compact if and only if it is closed and bounded.*

*Proof.* As  $S$  is compact, it is sequentially compact and thus closed and bounded. Suppose  $S \subset \mathbb{R}$  is closed and bounded and let  $\mathcal{O}$  be an open cover. Suppose no finite subcover exists. Denote  $\Delta$  as the set of elements  $x \in S$  such that for all elements  $s < x$  and  $s \in S$ , there are indeed finitely many open sets in  $\mathcal{O}$  that cover them. This set is not empty, as the greatest lower bound of  $S$  is contained in it. It is also bounded by the least upper bound of  $S$ . Let  $r$  be the least upper bound of  $\Delta$ . Suppose,  $r \neq l.u.b.(S)$ . As  $r \in S$ , there must be some open set  $U_1 \in \mathcal{O}$  such that  $r \in U_1$ . But then there is an  $\varepsilon > 0$  such that  $B_\varepsilon(r) \subset U_1$ . As  $r$  is the least upper bound of  $\Delta$ ,  $[r, r + \varepsilon) \cap S = \emptyset$ . Let  $r' = g.l.b.\{x \in S : x > r\}$ . Then  $r' \in S$  and there is a set  $U_2 \in \mathcal{O}$  such that  $r \in U_2$ . But then  $r' \in \Delta$  and  $r' > r$ , a contradiction. Thus,  $r = b$ . But then every element of  $S$  is covered by finitely many elements of  $\mathcal{O}$ . Therefore every open cover of  $S$  has a finite subcover. □

**Theorem 25.3.25.** *A subset of  $\mathbb{R}^n$  is compact if and only if it is closed and bounded.*

*Proof.* For if  $\mathcal{U} \subset \mathcal{R}^n$  is continuous, then  $\pi_j(\mathcal{U})$  is compact for all  $1 \leq j \leq n$ . But then  $\mathcal{U}$  is the product closed and bounded spaces and is thus closed and bounded. If  $\mathcal{R}^n$  is closed and bounded, then  $\pi_j(\mathcal{U})$  is as well and is thus compact. But the product of compact spaces is compact. Therefore, etc.  $\square$

**Definition 25.3.24** The unit sphere  $\mathbb{S}^{n-1}$  is defined as  $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$

The set of all compact subset of  $\mathbb{R}^n$  is denoted  $\mathcal{C}_n$

**Theorem 25.3.26.** *If  $S$  is a metric space  $T \subset S$  is compact, and  $f : T \rightarrow \mathbb{R}$  is continuous, then  $f$  attains its maximum in  $T$ .*

*Proof.* As  $f$  is continuous and  $T$  is compact,  $f(T)$  is compact and therefore  $f(T)$  is bounded. Let  $r$  be its least upper bound. For  $n \in \mathbb{N}$ , let  $x_n$  be a point such that  $|r - f(x_n)| < \frac{1}{n}$ . Such a point exist as otherwise  $r$  is not a least upper bound. As  $T$  is compact it is limit point compact and thus there is an accumulation point in  $x \in T$ . From continuity,  $f(x) = r$ .  $\square$

**Definition 25.3.25** A subset  $S \subset X$  of a topological space  $(X, \tau)$  is said to be path-connected if and only if for every pair of points  $x, y \in S$ , there is a continuous function  $f : [0, 1] \rightarrow S$  such that  $f(0) = x$  and  $f(1) = y$ .

**Theorem 25.3.27.** *A path-connected set is connected.*

*Proof.* For suppose not. Let  $S$  be path-connected and suppose  $T \subset S$  is both open and closed and non-empty. Let  $x \in T$  and  $y \in S/T$ . Let  $f : [0, 1] \rightarrow S$  be a continuous path. Let  $A = \{0 \leq x \leq 1 : f(x) \in T\}$ . This set is non-empty as  $0 \in A$ . As it is bounded, it has a least upper bound, call it  $r$ . Either  $f(r) \in T$  or  $f(r) \in T^c$ . If  $f(r) \in T$ , then there is a ball  $B_\varepsilon(f(r))$  that is contained in  $T$ . But then from continuity of  $f$ ,  $r$  is not the least upper bound of  $A$ . Thus  $r \notin T$ . In a similar manner,  $f(r) \notin T^c$ , a contradiction. Thus,  $S$  is connected.  $\square$

### 25.3.4 The Real Numbers

We construct the "God-Given" positive integers  $\mathbb{N}$ , then the whole numbers  $\mathbb{Z}$ , rational numbers  $\mathbb{Q}$ , and real numbers  $\mathbb{R}$ .

**Definition 25.3.26** (Peano's Axioms)  $\mathbb{N}$  is a set with equality, a total order  $\leq$ , and a successor function  $s$  such that:

1.  $1 \in \mathbb{N}$
2. For all  $n \in \mathbb{N}$ ,  $1 \leq n < s(n)$ .
3. If  $n, m \in \mathbb{N}$  and  $n \leq m \leq s(n)$ , then either  $m = n$  or  $m = s(n)$ .
4. Given any set  $K$ , if  $1 \in K$  and  $s(n) \in K$  for all  $n \in \mathbb{N}$ , then  $\mathbb{N} \subset K$ .

**Theorem 25.3.28.** *There is no element  $n \in \mathbb{N}$  such that  $s(n) = 1$ .*

*Proof.*  $[s(n) = 1] \Rightarrow [1 \leq n < s(n) = 1] \Rightarrow [1 < 1]$ , a contradiction.  $\square$

**Theorem 25.3.29.** *If  $n < m$ , then  $s(n) < s(m)$ .*

*Proof.*  $[n < m] \Rightarrow [s(n) \leq m] \Rightarrow [s(n) < s(m)]$ .  $\square$

**Theorem 25.3.30.** *For  $n, m \in \mathbb{N}$ ,  $s(n) = s(m)$  if and only if  $n = m$ .*

*Proof.*  $[n = m] \Rightarrow [s(n) = s(m)]$ .  $[[s(n) = s(m)] \wedge [n < m]] \Rightarrow [s(n) < s(m)]$ , a contradiction.  $\square$

The successor function  $s$  is the  $+1$  function,  $s(n) = n + 1$ . We freely write  $2 = 1 + 1$ ,  $3 = 1 + 2$ , ...

**Theorem 25.3.31.** *Every nonempty subset of  $\mathbb{N}$  has a least element.*

*Proof.* Suppose not. Let  $E \subset \mathbb{N}$ ,  $E \neq \emptyset$ .  $[n \in E] \Rightarrow [1 \leq n] \Rightarrow [1 \in E^c]$ .  $[k \in E^c] \Rightarrow [s(k) \in E^c] \Rightarrow [\mathbb{N} \subset E^c] \Rightarrow [E = \emptyset]$ .  $\square$

**Theorem 25.3.32** (Principle of Mathematical Induction). *If  $P$  is a proposition on the positive integers, if  $P(1)$  is true and the truthfulness of  $P(n)$  implies the truthfulness of  $P(n + 1)$ , then  $P(n)$  is true for all  $n \in \mathbb{N}$ .*

*Proof.* For suppose not. Then there is a least element  $n$  such that  $P(n)$  is false. As  $P(1)$  is true,  $n \neq 1$ . But then  $P(n - 1)$  is true. But the truthfulness of  $P(n - 1)$  implies the truthfulness of  $P(n)$ . Thus  $P(n)$  is true. A contradiction.  $\square$

**Definition 25.3.27** An  $n$ -tuple is inductively defined by  $(a_1, \dots, a_{n+1}) = (a_1, \dots, a_n) \cup \{a_1, \dots, a_{n+1}\}$ .

**Definition 25.3.28** We now inductively define for any set  $A$ ,  $A^n = A \times \underset{n-times}{\cdots} \times A$  by  $A^{n+1} = A^n \times A$ .

**Definition 25.3.29** The whole numbers  $\mathbb{Z}$  are a group with operation  $+$  with the following properties.

1.  $0$  is the identity element.
2.  $\mathbb{N} \subset \mathbb{Z}$ .
3. If  $0 < n$ ,  $n \in \mathbb{N}$ .

We have thus added all of the negative integers and  $0$ . The whole numbers are also called integers.

**Definition 25.3.30** The rational numbers  $\mathbb{Q}$  are an ordered field with operations  $+$  and  $\cdot$  such that  $\mathbb{Z} \subset \mathbb{Q}$ .

This gives us all of the fractions. If  $x \in \mathbb{Q}$  we may write  $x = \frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0$ .

**Definition 25.3.31** The greatest common divisor of two positive integers  $p, q \in \mathbb{N}$  is the smallest positive integer,  $r$ , such that there are integers  $n$  and  $m$  such that  $n \cdot r = p$  and  $m \cdot r = q$ . This number is denoted  $\text{g.c.d.}(p, q)$ .

**Theorem 25.3.33.** *If  $x \in \mathbb{Q}$  is positive, then there are unique positive integers  $p, q$  such that  $\text{g.c.d.}(p, q) = 1$  and  $x = \frac{p}{q}$ .*

*Proof.* This is proved via application of the fundamental theorem of arithmetic and will be one of the few omissions.  $\square$

**Definition 25.3.32** A subset  $A$  of  $\mathbb{Q}$  is said to be bounded above if and only if  $\exists M \in \mathbb{Q} : \forall x \in A, x \leq M$ .

**Definition 25.3.33** A subset  $A$  of  $\mathbb{Q}$  is said to be bounded below if and only if  $\exists M \in \mathbb{Q} : \forall x \in A, M \leq x$ .

**Definition 25.3.34** A subset  $A$  of  $\mathbb{Q}$  is said to be bounded if and only if it is both bounded below and bounded above.

**Definition 25.3.35** If  $A \subset \mathbb{Q}$  is bounded above, then  $r$  is said to be a least upper bound of  $A$  if and only if  $r$  is an upper bound for  $A$  and for all  $s < r$  there is an element  $x \in A$  such that  $s < x$ . We write  $\text{l.u.b.}(A)$ .

**Definition 25.3.36** If  $A \subset \mathbb{Q}$  is bounded below, then  $r$  is said to be a greatest lower bound of  $A$  if and only if  $r$  is a lower bound for  $A$  and for all  $r < s$  there is an element  $x \in A$  such that  $x < s$ . We write  $\text{g.l.b.}(A)$ .

**Definition 25.3.37** A number  $n \in \mathbb{N}$  is said to be even if and only if there is a number  $k \in \mathbb{N}$  such that  $n = 2k$ . A number  $m \in \mathbb{N}$  is said to be odd if and only if there is a number  $k \in \mathbb{N}$  such that  $m = 2k - 1$ .

**Theorem 25.3.34.** *If  $n \in \mathbb{N}$  and  $n^2$  is even, then  $n$  is even.*

*Proof.*  $[[n^2 \text{ even}] \wedge [n \text{ odd}]] \Rightarrow [\exists k \in \mathbb{N} : n = 2k - 1] \Rightarrow [n^2 = 4k(k - 1) + 1] \Rightarrow [n^2 \text{ odd}]$ , a contradiction. Thus,  $n$  is even.  $\square$

**Theorem 25.3.35.** *There is no rational number  $q$  such that  $q^2 = 2$ .*

*Proof.*  $[[x \in \mathbb{Q}] \wedge [x^2 = 2]] \Rightarrow [x = \frac{p}{q} : \text{g.c.d.}(p, q) = 1] \Rightarrow [\frac{p^2}{q^2} = 2] \Rightarrow [p^2 = 2q^2] \Rightarrow [p \text{ even}] \Rightarrow [\exists k \in \mathbb{N} : p = 2k] \Rightarrow [\frac{4k^2}{q^2} = 2] \Rightarrow [q^2 = 2k^2] \Rightarrow [q \text{ even}] \Rightarrow [\text{g.c.d.}(p, q) \geq 2]$ , a contradiction.  $\square$

**Theorem 25.3.36.** *There exist bounded subsets of  $\mathbb{Q}$  that contain no least upper bound.*

*Proof.* Let  $E = \{x \in \mathbb{Q} : x^2 < 2\}$ . It is bounded above by 2. Suppose  $s \in \mathbb{Q}$  is the least upper bound of  $E$ . Let  $x = s - \frac{s^2-2}{s+2}$ .  $[x \in \mathbb{Q}] \wedge [x^2 = 2\frac{s^2-2}{(s+2)^2} + 2]$ .  $[s^2 < 2] \Rightarrow [[x^2 < 2] \Rightarrow [x \in E]] \wedge [s < x]$ , a contradiction as  $s$  is an upper bound of  $E$ .  $[s^2 > 2] \Rightarrow [[2 < x^2] \wedge [x < s]] \Rightarrow$ , a contradiction as  $s$  is the least upper bound. Therefore, etc.  $\square$

**Definition 25.3.38**  $\mathbb{R}$  is an ordered field,  $\mathbb{Q} \subset \mathbb{R}$  : every nonempty, bounded above subset has a least upper bound.

**Theorem 25.3.37.** Least upper bounds are unique.

*Proof.* If  $A$  is a bounded set,  $r \neq r'$  are least upper bounds, then either  $r < r'$  or  $r' < r$ , a contradiction.  $\square$

**Theorem 25.3.38.** If  $A$  is a bounded below set, then there is a greatest lower bound.

*Proof.* Let  $-M < 0$  be a bound and define  $-A = \{-x : x \in A\}$ .  $[-x \in -A] \Rightarrow [x \in A] \Rightarrow [-x \leq M] \Rightarrow [-A \text{ is bounded above}] \Rightarrow [\exists l.u.b.(-A)]$ .  $[-x \in -A] \Rightarrow [x \in A] \Rightarrow [x \leq -l.u.b.(A)] \Rightarrow [-l.u.b.(A) \leq -x] \Rightarrow [g.l.b.(-A) = -l.u.b.(A)]$   $\square$

**Theorem 25.3.39** (The Archimedean Principle). For every  $x \in \mathbb{R}$  there is a least  $n \in \mathbb{N}$  such that  $x < n$ .

*Proof.* If  $x \leq 1$ , let  $n = 1$ . Let  $x > 1$  and  $E = \{i \in \mathbb{Z} : 0 \leq i \leq x\}$ .  $[0 \in E] \Rightarrow [E \neq \emptyset]$ .  $[i \in E] \Rightarrow [i \leq x] \Rightarrow [\exists l.u.b.(E)]$ . Let  $l.u.b.(E) = s$ .  $[s-1 < s] \Rightarrow [\exists i \in E : s-1 \leq i \leq s]$ .  $[i < s] \Rightarrow [\exists m \in E : i < m \leq s] \Rightarrow [0 < m-i \leq s-1 < 1]$ . But  $[m-i \in \mathbb{Z}] \Rightarrow [0 < m-i < 1 \text{ is false}] \Rightarrow [i = s] \Rightarrow s \in \mathbb{N}$ . If  $x = s$ ,  $n = s+1$ . Otherwise,  $n = s$ .  $\square$

**Theorem 25.3.40.** For every  $x \in \mathbb{R}$  there is a least  $n \in \mathbb{N}$  such that  $-n < x$ .

*Proof.* There is a least  $n \in \mathbb{N}$  such that  $(-x) < n$ . But then  $-n < -(-x) = x$ .  $\square$

**Theorem 25.3.41.** If  $x, y \in \mathbb{R}$  and  $x > 0$ , then there is an  $n \in \mathbb{N}$  such that  $nx > y$ .

*Proof.* If  $y \leq 0$ ,  $n = 1$ . If  $y > 1$ , let  $r = \frac{y}{x}$ .  $[x, y > 0] \Rightarrow [\frac{y}{x} > 0] \Rightarrow [\exists n \in \mathbb{N} : n > r] \Rightarrow [nx > rx = \frac{y}{x}x = y] \Rightarrow [nx > y]$ .  $\square$

**Theorem 25.3.42.** If  $0 \leq y$ , there is a unique number  $x > 0$  such that  $x^2 = y$ .

*Proof.*  $[(x^2 = y) \wedge (x'^2 = y) \wedge (x \neq x')] \Rightarrow [(x < x') \vee (x' < x)] \Rightarrow [(2 = x^2 < xx' < x'^2 = 2) \vee (2 = x'^2 < x'x < x^2 = 2)]$ , a contradiction. Thus, uniqueness is proved. For existence,  $[y = 0] \Rightarrow [x = 0]. [y = 1] \Rightarrow [x = 1]$ . Let  $0 < y < 1$  and define  $A = \{x \geq 0 : x^2 \leq y\}$ .  $[0 \in A] \Rightarrow [A \neq \emptyset]$ .  $[y < 1] \Rightarrow [A \text{ is bounded above}]$ . Let  $r$  be the least upper bound. Suppose  $r^2 \neq y$ .

1.  $[y < r^2] \Rightarrow [\frac{r^2-y}{2} > 0] \Rightarrow [r - \frac{r^2-y}{2} < r] \wedge [(r - \frac{r^2-y}{2})^2 = r^2 - (r^2 - y) + (\frac{r^2-y}{2})^2 = y + (\frac{r^2-y}{2})^2 < y]$ . A contradiction.
2.  $[r^2 < y] \Rightarrow [0 < \frac{y-r^2}{2r+1} < 1] \Rightarrow [r^2 + 2r\frac{y-r^2}{2r+1} + (\frac{y-r^2}{2r+1})^2 \leq r^2 + 2r\frac{y-r^2}{2r+1} + \frac{y-r^2}{2r+1} = r^2 + \frac{y-r^2}{2r+1}(2r+1) = y]$ . A contradiction.

Thus,  $r^2 = y$ .  $\square$

**Definition 25.3.39** If  $x > 0$ , then  $\sqrt{x}$  is the unique positive number such that  $(\sqrt{x})^2 = x$ . This is the *square-root* of  $x$ .

**Theorem 25.3.43.**  $1 < \sqrt{2}$

*Proof.* For  $\sqrt{2} \neq 1$ , as  $1^2 = 1 \neq 2$ . If  $\sqrt{2} < 1$ , then  $2 = (\sqrt{2})^2 < 1$ , a contradiction. Thus  $1 < \sqrt{2}$ .  $\square$

**Definition 25.3.40** An irrational number is a real number that is not rational.

**Theorem 25.3.44.**  $\frac{1}{\sqrt{2}}$  is irrational.

*Proof.* For if  $\frac{1}{\sqrt{2}} = \frac{p}{q}$ ,  $p, q \in \mathbb{N}$ , then  $\sqrt{2} = \frac{q}{p}$ , a contradiction.  $\square$

**Theorem 25.3.45.** If  $q$  is a rational number not equal to zero, and  $r$  is irrational, then  $rq$  is irrational.

*Proof.* As  $q \neq 0$ , let  $q = \frac{n}{m}$  be in reduced form. Suppose  $rq = \frac{x}{y} \in \mathbb{Q}$ . Then  $r = \frac{xm}{yn}$ , a contradiction.  $\square$

**Theorem 25.3.46.** Given a rational number  $q$ , and for any  $\varepsilon > 0$ , there is an irrational number  $r$  such that  $|r - q| < \varepsilon$ .

*Proof.*  $[\varepsilon > 0] \Rightarrow [\frac{1}{\varepsilon} > 0] \Rightarrow [\exists N \in \mathbb{N} : \frac{1}{\varepsilon} < N] \Rightarrow [\frac{1}{\varepsilon} < \sqrt{2}N] \Rightarrow [\frac{1}{\sqrt{2}N} < \varepsilon]$ .  $[r \equiv q + \frac{1}{\sqrt{2}N}] \Rightarrow [r \notin \mathbb{Q}] \wedge [|r - q| = |\frac{1}{\sqrt{2}N}| < \varepsilon]$ .  $\square$

**Theorem 25.3.47.** If  $r$  is an irrational number, and  $\varepsilon > 0$ , then there is a rational number  $q$  such that  $|r - q| < \varepsilon$ .

*Proof.*  $[0 < r] \Rightarrow [[\exists n \in \mathbb{N} : \frac{1}{n} < \varepsilon] \wedge [\exists m \in \mathbb{N} : m - 1 \leq nr \leq m]] \Rightarrow [|r - \frac{m}{n}| \leq \frac{1}{n} < \varepsilon]$ . Similarly if  $r < 0$ .  $\square$

**Definition 25.3.41** The absolute value function is defined on  $\mathbb{R}$  as  $|x| = \begin{cases} x, & 0 \leq x \\ -x, & x < 0 \end{cases}$

**Theorem 25.3.48.**  $|x| = \sqrt{x^2}$ .

*Proof.* If  $0 \leq x$ , we are done. If  $x < 0$ , then  $|x| = (-x) = \sqrt{(-x)^2} = \sqrt{x^2}$ .  $\square$

**Theorem 25.3.49.** For  $x, y \in \mathbb{R}$ ,  $|x + y| \leq |x| + |y|$

*Proof.*  $[0 \leq x, y] \Rightarrow [|x + y| = x + y = |x| + |y|]$ .  $[x, y \leq 0] \Rightarrow [|x + y| = -(x + y) = (-x) + (-y) = |x| + |y|]$ .  $[x \leq 0 \leq y] \Rightarrow [0 \leq x + y] \Rightarrow [|x + y| = x + y \leq (-x) + y = |x| + |y|]$ . Similarly for  $y \leq 0 \leq x$ .  $\square$

**Theorem 25.3.50** (Triangle Inequality). If  $x, y, z \in \mathbb{R}$ , then  $|x - y| \leq |x - z| + |y - z|$ .

*Proof.* For  $|x - y| = |x + 0 - y| = |x - z + z - y| = |(x - z) + (z - y)| \leq |x - z| + |z - y| = |x - y| + |y - z|$ .  $\square$

**Theorem 25.3.51.** If  $\varepsilon > 0$  and  $|x| < \varepsilon$ , then  $-\varepsilon < x < \varepsilon$ .

*Proof.* For if  $0 \leq x$ , then  $-\varepsilon < x = |x| < \varepsilon$ . If  $x \leq 0$ , then  $x \leq 0 < \varepsilon$  and  $(-x) = |x| < \varepsilon$ , thus  $-\varepsilon < -(-x) = x$ .  $\square$

**Definition 25.3.42** A sequence in  $A$  is a function  $x_n$  is a function  $x_n : \mathbb{N} \rightarrow A$ . We write the image of  $n \in \mathbb{N}$  as  $n \mapsto x_n$ .

**Definition 25.3.43** If  $x_n$  is a sequence, a subsequence is a subset of  $x_n$ , denoted  $x_{n_k}$ , where  $n_k \in \mathbb{N}$  is strictly increasing.

**Definition 25.3.44** A sequence  $x_n$  converges to  $x$  if and only if  $\forall \varepsilon > 0$ ,  $\exists N \in \mathbb{N} : n > N \Rightarrow |x - x_n| < \varepsilon$ . We write  $x_n \rightarrow x$ .

**Theorem 25.3.52.** Convergence in  $\mathbb{R}$  is unique.

*Proof.* For suppose not. Let  $x_n \rightarrow x$  and  $x_n \rightarrow x'$  and suppose  $x \neq x'$ . Let  $\varepsilon = \frac{|x - x'|}{2}$ .  $[x \neq x'] \Rightarrow [\varepsilon > 0] \Rightarrow [\exists N \in \mathbb{N} : |x_n - x| < \varepsilon \wedge |x_n - x'| < \varepsilon] \Rightarrow [|x - x'| = |x - x_n + x_n - x'| \leq |x_n - x| + |x_n - x'| < 2\varepsilon = |x - x'|]$ , a contradiction.  $\square$

**Definition 25.3.45** A sequence is Cauchy if and only if  $\forall \varepsilon > 0$ ,  $\exists N \in \mathbb{N} : n, m > N \Rightarrow |x_n - x_m| < \varepsilon$ .

**Definition 25.3.46** A closed interval  $[a, b]$  is a subset of  $\mathbb{R}$  defined as  $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ .

**Definition 25.3.47** A sequence is said to be monotonically increasing if and only if for all  $n \in \mathbb{N}$ ,  $x_n \leq x_{n+1}$ , monotonically decreasing if and only if for all  $n \in \mathbb{N}$ ,  $x_{n+1} \leq x_n$ , and monotonic if and only if it is either monotonically decreasing or monotonically increasing. Strictly increasing or strictly decreasing means  $x_n < x_{n+1}$  and  $x_{n+1} < x_n$ , respectively.

**Theorem 25.3.53.** *Every sequence in  $\mathbb{R}$  has a monotonic subsequence.*

*Proof.* If  $n \in \mathbb{N} : n < m \Rightarrow x_m \leq x_n$ , call  $n$  a peak point. If there is a sequence of peak points  $n_k$ , then the subsequence  $x_{n_k}$  is a sequence of peak points and is thus monotonic. If none such subsequence exist, there is a greatest peak point, call it  $n_1$ . Then there is an  $n_2 \in \mathbb{N}$  such that  $n_1 < n_2$  and  $x_{n_1} < x_{n_2}$ , otherwise there is a peak point greater than  $n_1$ . Inductively, there is a strictly increasing sequence  $n_k$  such that  $x_{n_k} < x_{n_{k+1}}$ . Therefore, etc.  $\square$

**Definition 25.3.48** A Dedekind Cut is a combination of two sets  $A$  and  $B$  such that for all  $x \in A$  and all  $y \in B$ ,  $x < y$ ,  $A \cap B = \emptyset$ , and  $\mathbb{Q} \subset A \cup B$ . A real number  $r$  is said to produce a Dedekind cut if and only if  $\forall a \in A \wedge \forall b \in B, a \leq r \leq b$ .

**Theorem 25.3.54.** *The following are equivalent characterizations of the completeness of  $\mathbb{R}$ .*

1. *Dedekind Cuts are produced by a unique real number.* [Dedekind Completeness]
2. *Bounded monotonic sequences converge.* [Monotone Convergence Theorem]
3. *If  $x_n$  is a bounded sequence, then there exist a convergent subsequence.* [Bolzano-Weierstrass Theorem]
4. *Cauchy Sequences Converge.* [Cauchy Completeness]
5. *If  $I_n = [a_n, b_n] \subset [a_{n+1}, b_{n+1}] = I_{n+1}$  are a sequence of non-empty closed intervals and  $b_n - a_n \rightarrow 0$ , then there is a unique point  $x$  that is contained in all intervals  $I_n$ .* [Cantor's Nested Intervals Theorem]

*Proof.* We show that (1)  $\Rightarrow$  (2)  $\Rightarrow$  (3)  $\Rightarrow$  (4)  $\Rightarrow$  (5)  $\Rightarrow$  (1), where  $\Rightarrow$  means implies.

1. For let  $A$  and  $B$  be a Dedekind Cut of  $\mathbb{Q}$ . It suffices to show that the least upper bound of  $A$  is equal to the greatest lower bound of  $B$ . Let  $r$  be the least upper bound of  $A$  and  $s$  the greatest lower bound of  $B$ . If  $x \in \mathbb{Q}$  and  $x < s$ , then  $x \notin B$ . But as  $A \cup B = \mathbb{Q}$ ,  $x \in A$ . Similarly, if  $r < x$ , then  $x \in B$ . Thus  $s = r$ .

2. For let  $x_n$  be a bounded monotonic sequence, suppose increasing, in  $\mathbb{R}$  and let  $A = \{M \in \mathbb{R} : x_n \leq M \text{ for all } n \in \mathbb{N}\}$ . Then  $A$  and  $A^c$  form a Dedekind cut of  $\mathbb{Q}$  and is thus produced by a real number, call it  $r$ . Then  $r$  is a greatest lower bound of  $A$ . Let  $\varepsilon > 0$  be arbitrary. Then, as  $r$  is a greatest lower bound, there is an  $N \in \mathbb{N}$  such that  $r - \varepsilon < x_N$ . But as  $x_n$  is monotonic, for all  $k \in \mathbb{N}$ ,  $r - \varepsilon < x_{N+k} \leq r$ . Thus,  $x_n \rightarrow r$
3. For let  $x_n$  be a bounded sequence. As all sequence have a monotonic subsequence, let  $x_{n_k}$  be such a subsequence. But then  $x_{n_k}$  is a bounded monotonic subsequence and thus converges.
4. Let  $x_n$  be a Cauchy sequence and let  $\varepsilon > 0$  be arbitrary. Then there is an  $N \in \mathbb{N}$  such that for all  $n, m > N$ ,  $|x_n - x_m| < \frac{\varepsilon}{2}$ . Then  $-\varepsilon < x_n - x_{N+1} < \varepsilon$ , and thus  $x_{N+1} - \varepsilon < x_n < \varepsilon + x_{N+1}$ . Thus,  $x_n$  is a bounded sequence. But bounded sequences have a convergent subsequence  $x_{n_k}$ . Let  $x$  be the limit. Then, for  $n > N$ ,  $|x - x_n| \leq |x - x_{n_k}| + |x_{n_k} - x_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ .
5. Let  $x_n = \begin{cases} a_n, & n \text{ is even.} \\ b_n, & n \text{ is odd.} \end{cases}$ . As  $b_n - a_n \rightarrow 0$  and  $a_n$  and  $b_n$  are monotonic (As  $I_n \subset I_{n+1}$ ),  $x_n$  is a Cauchy sequence and thus converges. Let  $x$  be the limit. But then  $a_n \leq x \leq b_n$  for all  $n \in \mathbb{N}$ , and for any  $x' \neq x$ , let  $\varepsilon = \frac{|x - x'|}{2}$ . As  $b_n - a_n \rightarrow 0$ , there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $|b_n - a_n| < \varepsilon$ , thus  $x' \notin [a_{N+1}, b_{N+1}]$ .  $x$  is unique.
6. Finally, (5)  $\Rightarrow$  (1). Let  $A$  and  $B$  be a Dedekind Cut of  $\mathbb{Q}$ . Let  $x_1 \in A$  be arbitrary and  $x_2 \in B$  be arbitrary and defined  $x_3 = \frac{x_1+x_2}{2}$ . Define  $x_n = \begin{cases} \frac{x_{n-1}+x_{n-2}}{2}, & \text{The Previous Two Terms are in Different Cuts} \\ \frac{x_{n-1}+x_{n-3}}{2}, & \text{The Previous Two Terms are in the Same Cut} \end{cases}$  For all  $n \in \mathbb{N}$ , define the following:

$$(a) a_n = \begin{cases} x_n, & x_n \in A \\ a_{n-1}, & x_n \notin A \end{cases}$$

$$(b) b_n = \begin{cases} x_n, & x_n \in B \\ b_{n-1}, & x_n \notin B \end{cases}$$

Then  $I_{n+1} = [a_{n+1}, b_{n+1}] \subset [a_n, b_n] = I_n$ , and  $b_n - a_n \rightarrow 0$ . Thus there is a unique point  $x \in I_n$  for all  $n \in \mathbb{N}$ . This produces the Dedekind cut, as for all  $a \in A$ ,  $a \leq x$  and for all  $b \in B$ ,  $x \leq b$ . Therefore, etc.

□

### 25.3.5 Vector Spaces and Euclidean Spaces

**Definition 25.3.49** A vector space is a set  $V$  of vectors and a field  $K$  of scalars with the following properties: For all  $a, b \in K$ ,  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ :

1.  $a\mathbf{v} \in V$ . [Closure of Scalar Multiplication]
2.  $\mathbf{v} + \mathbf{u} \in V$ . [Closure of Vector Addition]
3.  $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$  [Vector Addition is Associative]
4.  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$  [Vector Addition is Commutative]
5. There exists a  $\mathbf{0} \in V$  such that  $\mathbf{0} + \mathbf{v} = \mathbf{v}$  for all  $\mathbf{v} \in V$ . [Existence of Zero Vector]
6.  $a(b\mathbf{v}) = (ab)\mathbf{v}$ . [Associativity of Scalar Multiplication]
7.  $1\mathbf{v} = \mathbf{v}$ . [Multiplication by Scalar Identity]
8.  $a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u}$ . [Scalar Multiplication Distributes of Vector Addition]
9.  $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$ . [Scalar Multiplication Distributes over Field Addition]

**Theorem 25.3.55.**  $\mathbf{0}$  is unique.

*Proof.* For  $\mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0}$ . □

**Theorem 25.3.56.**  $0\mathbf{v} = \mathbf{0}$ .

*Proof.* For  $\mathbf{v} + 0\mathbf{v} = (1 + 0)\mathbf{v} = 1\mathbf{v} = \mathbf{v}$ . As  $\mathbf{0}$  is unique,  $0\mathbf{v} = \mathbf{0}$ . □

**Theorem 25.3.57.** For all  $\mathbf{v} \in V$ , there exists a  $\mathbf{u}$  such that  $\mathbf{v} + \mathbf{u} = \mathbf{0}$ . That is, additive inverses exist.

*Proof.* For let  $\mathbf{u} = (-1)\mathbf{v}$ . Then  $\mathbf{v} + \mathbf{u} = \mathbf{v} + (-1)\mathbf{v} = (1 + (-1))\mathbf{v} = 0\mathbf{v} = \mathbf{0}$ . □

**Theorem 25.3.58.** Inverses are unique.

*Proof.* For  $-\mathbf{v}' = -\mathbf{v}' + \mathbf{0} = -\mathbf{v}' + \mathbf{v} - \mathbf{v} = -\mathbf{v}$  □

**Definition 25.3.50** A subspace of a vector space  $V$  over a field  $K$  is a vector space  $W$  with the following properties:

1.  $\mathbf{0} \in W$
2. If  $\mathbf{u}, \mathbf{v} \in W$ , then  $\mathbf{u} + \mathbf{v} \in W$
3. For all  $a \in K$  and  $\mathbf{v} \in W$ ,  $a\mathbf{v} \in W$

**Definition 25.3.51** An affine subspace of a vector space  $V$  over a field  $K$  is a subset  $\xi \subset V$  such that  $\xi = \{v + w : w \in W\}$ , where  $v$  is a fixed vector in  $V$ , and  $W$  is a fixed subspace of  $V$ . That is, they are translations of subspaces.

**Definition 25.3.52** An inner product on a vector space  $V$  over a subfield  $K$  of  $\mathbb{R}$  is a function  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$  with the following properties: For all  $x, y, z \in V$ , and  $\alpha \in K$ ,

1.  $\langle x, y \rangle = \langle y, x \rangle$  [Symmetry]
2.  $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$  [Linearity]
3.  $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$  [Linearity]
4. If  $x \neq \mathbf{0}$ , then  $\langle x, x \rangle > 0$  [Positiveness]

**Definition 25.3.53** An inner product space is a vector space with an inner product.

**Theorem 25.3.59.**  $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$

*Proof.* For  $\langle x, y + z \rangle = \langle y + z, x \rangle = \langle y, x \rangle + \langle z, x \rangle = \langle x, y \rangle + \langle x, z \rangle$   $\square$

**Theorem 25.3.60.**  $\langle x, x \rangle = 0$  if and only if  $x = \mathbf{0}$ .

*Proof.* For  $\langle \mathbf{0}, \mathbf{0} \rangle = \langle \mathbf{00}, \mathbf{0} \rangle = 0 \langle \mathbf{0}, \mathbf{0} \rangle = 0$ . Suppose  $\langle x, x \rangle = 0$  but  $x \neq \mathbf{0}$ . But then  $\langle x, x \rangle > 0$ , a contradiction.  $\square$

**Theorem 25.3.61.** For all  $x \in V$ ,  $\langle \mathbf{0}, x \rangle = 0$

*Proof.* For  $\langle \mathbf{0}, x \rangle = \langle \mathbf{00}, x \rangle = 0 \langle \mathbf{0}, x \rangle = 0$ .  $\square$

**Theorem 25.3.62** (Cauchy-Bunyakovsky-Schwarz Inequality). In an inner product space  $V$ ,  $x, y \in V \Rightarrow \langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle$

*Proof.*  $[y = \mathbf{0}] \Rightarrow [\langle x, y \rangle = 0]$ . Suppose  $y \neq \mathbf{0}$ , and let  $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$ . Then  $[0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle] \Rightarrow [0 \leq \langle x, x \rangle - 2\frac{\langle x, y \rangle^2}{\langle y, y \rangle} + \frac{\langle x, y \rangle^2}{\langle y, y \rangle} = \langle x, x \rangle - \frac{\langle x, y \rangle^2}{\langle y, y \rangle}] \Rightarrow [\frac{\langle x, y \rangle^2}{\langle y, y \rangle} \leq \langle x, x \rangle] \Rightarrow [\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle]$ .  $\square$

**Definition 25.3.54** A norm on a vector space  $V$  over a subfield  $K$  of  $\mathbb{R}$  is a function  $\|\cdot\| : V \rightarrow \mathbb{R}$  with the following properties: For all  $x \in V$  and  $\alpha \in K$ :

1.  $\|\alpha x\| = |\alpha| \|x\|$  [Absolute Homogeneity]
2.  $\|x + y\| \leq \|x\| + \|y\|$  [Triangle Inequality]
3.  $\|x\| = 0$  if and only if  $x = \mathbf{0}$ . [Definiteness]

**Definition 25.3.55** A normed space is a vector space with a norm.

**Theorem 25.3.63.** *If  $V$  is a normed space, then for all  $x \in V$ ,  $0 \leq \|x\|$*

*Proof.* For  $0 = \|0\| = \|\frac{x-x}{2}\| \leq \|\frac{x}{2}\| + \|\frac{-x}{2}\| = \frac{1}{2}\|x\| + \frac{1}{2}\|x\| = \|x\|$ .  $\square$

**Definition 25.3.56** If  $V$  is an inner product space, then the induced norm is  $\|x\| = \sqrt{\langle x, x \rangle}$ .

**Theorem 25.3.64.** *The induced norm is a norm.*

*Proof.* In order,

1.  $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha^2 \langle x, x \rangle} = |\alpha| \sqrt{\langle x, x \rangle} = |\alpha| \|x\|$
2.  $\|x + y\|^2 = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \Rightarrow \|x + y\| \leq \|x\| + \|y\|$
3. If  $x = \mathbf{0}$ , then  $\sqrt{\langle x, x \rangle} = \sqrt{0} = 0$ . If  $\sqrt{\langle x, x \rangle} = 0$  then  $\langle x, x \rangle = 0$ , and thus  $x = \mathbf{0}$ .

$\square$

**Theorem 25.3.65** (Cauchy-Schwarz Inequality). *If  $V$  is an inner product space, then for all  $x, y \in V$ ,  $|\langle x, y \rangle| \leq \|x\|\|y\|$ .*

*Proof.* For  $\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle = \|x\|^2 \|y\|^2 = (\|x\|\|y\|)^2$ . Thus,  $|\langle x, y \rangle| \leq \|x\|\|y\|$ .  $\square$

**Definition 25.3.57** Euclidean  $n$ -space is defined as  $\mathbb{R}^n = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}\}$ , and has the following arithmetic: For all  $x, y \in \mathbb{R}^n$ ,  $\alpha \in \mathbb{R}$ ,

1.  $x + y = (x_1 + y_1, \dots, x_n + y_n)$
2.  $\alpha x = (\alpha x_1, \dots, \alpha x_n)$ .

**Theorem 25.3.66.**  $\mathbb{R}^n$ , with its usual arithmetic, is a vector space over  $\mathbb{R}$ .

*Proof.* In order (Laboriously): Let  $\alpha, \beta \in \mathbb{R}$ ,  $x, y, z \in \mathbb{R}^n$ ,

1.  $\alpha x = \alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n)$ . As  $\alpha x_i \in \mathbb{R}$ ,  $\alpha x \in \mathbb{R}^n$ .
2.  $x + y = (x_1 + y_1, \dots, x_n + y_n)$ . As  $x_i + y_i \in \mathbb{R}$ ,  $x + y \in \mathbb{R}^n$ .
3.  $x + (y + z) = (x_1 + (y_1 + z_1), \dots, x_n + (y_n + z_n)) = ((x_1 + y_1) + z_1, \dots, (x_n + y_n) + z_n) = (x + y) + z$
4.  $x + y = (x_1 + y_1, \dots, x_n + y_n) = (y_1 + x_1, \dots, y_n + x_n) = y + x$
5.  $\mathbf{0} + x = (0 + x_1, \dots, 0 + x_n) = (x_1, \dots, x_n) = x$ .
6.  $\alpha(\beta x) = \alpha(\beta x_1, \dots, \beta x_n) = \alpha\beta(x_1, \dots, x_n) = (\alpha\beta)x$

7.  $1x = (x_1, \dots, x_n) = x.$

8.

$$\begin{aligned} \alpha(x + y) &= \alpha(x_1 + y_1, \dots, x_n + y_n) &= (\alpha x_1, \dots, \alpha x_n) + (\alpha y_1, \dots, \alpha y_n) \\ &= (\alpha(x_1 + y_n), \dots, \alpha(x_n + y_n)) &= \alpha(x_1, \dots, x_n) + \alpha(y_1, \dots, y_n) \\ &= (\alpha x_1 + \alpha y_1, \dots, \alpha x_n + \alpha y_n) &= \alpha x + \alpha y \end{aligned}$$

9.

$$\begin{aligned} (\alpha + \beta)x &= ((\alpha + \beta)x_1, \dots, (\alpha + \beta)x_n) &= \alpha(x_1, \dots, x_n) + \beta(x_1, \dots, x_n) \\ &= (\alpha x_1 + \beta x_1, \dots, \alpha x_n + \beta x_n) &= \alpha x + \beta x \\ &= (\alpha x_1, \dots, \alpha x_n) + (\beta x_1, \dots, \beta x_n) \end{aligned}$$

□

**Definition 25.3.58** The dot product is a function  $\cdot : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  defined as  $x \cdot y = \sum_{i=1}^n x_i y_i$

**Theorem 25.3.67.** *The dot product is an inner product on  $\mathbb{R}^n$ .*

*Proof.* In order,

1.  $x \cdot y = \sum_{i=1}^n n x_i y_i = \sum_{i=1}^n y_i x_i = y \cdot x$
2.  $\alpha x \cdot y = \sum_{i=1}^n \alpha x_i y_i = \alpha \sum_{i=1}^n x_i y_i = \alpha x_i \cdot y_i$
3.  $(x+y) \cdot z = \sum_{i=1}^n (x_i + y_i) z_i = \sum_{i=1}^n (x_i z_i + y_i z_i) = \sum_{i=1}^n x_i z_i + \sum_{i=1}^n y_i z_i = x \cdot z + y \cdot z$
4.  $x \cdot x = \sum_{i=1}^n x_i^2 \geq 0$ . Indeed, if  $x \cdot x = 0$ , then  $x_i = 0$  for all  $i = 1, \dots, n$ , and thus  $x = \mathbf{0}$ .

□

The induced norm is thus  $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ .

**Definition 25.3.59** A linear combination of vectors  $\mathbf{v}_i$  in a vector space  $V$  over a field  $K$  is a sum  $\sum_{i=1}^n a_i \mathbf{v}_i$ ,  $a_i \in K$ .

**Definition 25.3.60** If  $V$  is a vector space over  $K$ , and if  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ , then they are said to be linearly dependent if and only if there are scalars  $a_1, \dots, a_n$ , not all equal to zero, such that  $\sum_{i=1}^n a_i \mathbf{v}_i = \mathbf{0}$ .

**Definition 25.3.61** If  $V$  is a vector space over  $K$ , and if  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ , then they are said to be linearly independent if and only if they are not linearly dependent.

**Definition 25.3.62** A set of vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  is said to span a vector space  $V$  if and only if every element  $x \in V$  can be written as a linear combination  $x = \sum_{i=1}^n a_i \mathbf{v}_i$  for some scalars  $a_i \in K$ .

**Definition 25.3.63** A basis of a vector space  $V$  of a field  $K$  is a set of linearly independent vectors that span  $V$ .

**Definition 25.3.64** A vector space  $V$  is said to be finite if it has a basis of finitely many elements.

**Definition 25.3.65** The dimension of a vector space  $V$ , denoted  $\dim(V)$ , is the cardinality of the smallest basis of  $V$ .

**Theorem 25.3.68** (The Dimension Theorem). *If  $V$  is a vector space and  $\dim(V) = n$ , then every basis of  $V$  has  $n$  vectors.*

**Definition 25.3.66** If  $V$  is a normed space, then an affine combination of vectors is  $\sum_{k=1}^n \lambda_k v_k$ , where  $\lambda_k \in K$ ,  $v_k \in V$ , and  $\sum_{k=1}^n \lambda_k = 1$ .

**Definition 25.3.67** A set of  $n$  vectors  $\{v_k : k \in \mathbb{Z}_n\}$  is said to be affinely dependent if and only if there exists scalars  $\lambda_k \in K$ , not all equal to zero, such that  $\sum_{k=1}^n \lambda_k v_k = \mathbf{0}$  and  $\sum_{k=1}^n \lambda_k = 0$ .

**Definition 25.3.68** A set of vectors is said to be affinely independent if and only if there are not affinely dependent.

**Theorem 25.3.69.** *A set  $\{v_k : k \in \mathbb{Z}_n\}$  is affinely independent if and only if  $\{v_k - v_1 : k \in \mathbb{Z}_n, k > 1\}$  is linearly independent.*

*Proof.* Suppose the latter set is linearly independent. Then  $\sum_{k=1}^n \lambda_k (v_k - v_1) \neq 0$  if at least one  $\lambda_k \neq 0$ . Let  $\lambda_k$  be any sequence such that  $\sum_{k=1}^n \lambda_k = 0$ , but not all  $\lambda_k$  are zero. Then  $\sum_{k=1}^n \lambda_k (v_k - v_1) \neq 0$ . Let this sum be  $c_\lambda$ . Then  $\sum_{k=1}^n \lambda_k v_k = \sum_{k=1}^n \lambda_k v_1 + c_\lambda = c_\lambda$ . Thus, the set  $v_k$  is affinely independent. Suppose the former set is affinely independent. Then  $\sum_{k=1}^n \lambda_k v_k = \mathbf{0} \Rightarrow \sum_{k=1}^n \lambda_k \neq 0$ . But then  $\sum_{k=1}^n \lambda_k (v_k - v_1) = -\sum_{k=1}^n \lambda_k v_1 \neq 0$ . Thus, the latter set is linearly independent.  $\square$

**Theorem 25.3.70.** *For any  $n$ -dimensional vector space  $V$ , any set of affinely independent vectors has at most  $n + 1$  vectors.*

*Proof.* If  $v_k$  are affinely independent, then  $v_k - v_1$  is linearly independent. The latter has at most  $n$  vectors. Thus, etc.  $\square$

**Theorem 25.3.71.** *If  $v_k$  are affinely independent and  $\sum_{k=1}^n \lambda_k v_k = \sum_{k=1}^n \sigma_k v_k$ , then  $\lambda_k = \sigma_k$  for all  $k$ .*

*Proof.*  $[\sum_{k=1}^n (\lambda_k - \sigma_k) v_k = 0] \wedge [\sum_{k=1}^n (\lambda_k - \sigma_k) = 0] \Rightarrow [\lambda_k - \sigma_k = 0]$ . Therefore, etc.  $\square$

**Definition 25.3.69** The Affine Hull of a set  $S \subset V$  of some normed space  $V$  is  $\text{aff}(S) = \{\sum_{i=1}^m \lambda_i x_i : x_i \in S \wedge \sum_{i=1}^m \lambda_i = 1\}$ .

**Definition 25.3.70** If  $V$  is a normed space, a convex combination is  $\sum_{i=1}^n |\lambda_i| v_i$ , where  $v_i \in V$ ,  $\sum_{i=1}^n |\lambda_i| = 1$ .

**Definition 25.3.71** For a normed space  $V$ , the Convex Hull of  $S \subset V$  is  $\text{conv}(S) = \{\sum_{i=1}^n |\lambda_i| x_i : x_i \in S \wedge \sum_{i=1}^n |\lambda_i| = 1\}$ .

**Definition 25.3.72** In an inner product space  $V$ ,  $v, w \in V$  are said to be orthogonal,  $v \perp w$ , if and only if  $\langle v, w \rangle = 0$ .

In an inner product space  $V$ ,  $W \subset V$ ,  $x \in V$ , and  $y \in W \Rightarrow \langle x, y \rangle = 0$ , we write  $x \perp W$ . Similarly for orthogonal subsets  $W$  and  $U$  of  $V$ , we write  $W \perp U$ .

**Definition 25.3.73** A line in  $\mathbb{R}^n$  containing the points  $x, y \in \mathbb{R}^n$  is the set  $\{\lambda y + (1 - \lambda)x : \lambda \in \mathbb{R}\}$ .

**Definition 25.3.74** A line segment in  $\mathbb{R}^n$  that begins at  $x$  and terminates at  $y$  is the set  $\{\lambda y + (1 - \lambda)x : 0 \leq \lambda \leq 1\}$ .

**Definition 25.3.75** If  $W \underset{\text{Subspace}}{\subset} \mathbb{R}^n$ ,  $K \subset \mathbb{R}^n$ , then the orthogonal projection is  $K_W \equiv \{x \in W : \exists y \in K : y - x \perp W\}$ .

**Theorem 25.3.72.** If  $W$  is a subspace of  $\mathbb{R}^n$ ,  $K \subset \mathbb{R}^n$ , and  $x \in K_W$ , then there is a line through  $x$  and a point  $y \in K$  such that, for any  $\alpha, \beta$  contained on said line,  $\beta - \alpha \perp W$ .

*Proof.* For let  $x \in W$ . Then there is a  $y \in K$  such that, for all  $z \in W$ ,  $\langle y - x, z \rangle = 0$ . Let  $\Gamma$  be the line  $\lambda y + (1 - \lambda)x$ . If  $\alpha, \beta \in \Gamma$ , there are values  $\lambda_1$  and  $\lambda_2$  such that  $\alpha = \lambda_1 y + (1 - \lambda_1)x$  and  $\beta = \lambda_2 y + (1 - \lambda_2)x$ . But then  $\beta - \alpha = y(\lambda_2 - \lambda_1) - x(\lambda_2 - \lambda_1) = (\lambda_2 - \lambda_1)(x - y)$ . But then  $\langle \beta - \alpha, z \rangle = \langle (\lambda_2 - \lambda_1)(x - y), z \rangle = (\lambda_2 - \lambda_1)\langle x - y, z \rangle = 0$ .  $\square$

From the Cauchy-Schwarz inequality, we have that  $|\langle x, y \rangle| \leq \|x\| \|y\|$ , and thus  $|\frac{\langle x, y \rangle}{\|x\| \|y\|}| \leq 1$ . We define the *angle*  $\theta$  between two non-zero vectors  $x, y \in \mathbb{R}^n$  as  $\theta = \cos^{-1} \left( \frac{\langle x, y \rangle}{\|x\| \|y\|} \right)$ . We omit rigorous definition of the cosine function.

**Definition 25.3.76** If  $\mathcal{U}, \mathcal{V} \subset V$ , then  $\mathcal{U} + \mathcal{V} = \{x + y : x \in \mathcal{U}, y \in \mathcal{V}\}$ .

**Definition 25.3.77** If  $V$  is a vector space over  $K$ ,  $\mathcal{U} \subset V$ , and  $\alpha \in K$ , then  $\alpha \mathcal{U} = \{\alpha x : x \in \mathcal{U}\}$ .

## 25.4 Definitions and Theorems

### 25.4.1 Definitions

**Definition 25.4.1** A binary relation on a set  $X$  is a subset  $R$  of  $X \times X$ .

**Definition 25.4.2** Comparable elements in a set  $X$  with respect to a binary relation  $R$  are elements  $x, y \in X$  such that either  $(x, y) \in R$ , denoted  $xRy$ , or  $(y, x) \in R$ , denoted  $yRx$ .

**Definition 25.4.3** A connex relation on a set  $X$  is a binary relation  $R$  on  $X$  such that  $\forall_{x,y \in X}$ , either  $xRy$  or  $yRx$ .

**Definition 25.4.4** A transitive relation on a set  $X$  is a binary relation  $R$  on  $X$  such that  $\forall_{x,y,z \in X}$ ,  $xRy \wedge yRz \Rightarrow xRz$ .

**Definition 25.4.5** An antisymmetric relation is a binary relation  $R$  on a set  $X$  such that  $\forall_{x,y \in X}$ ,  $xRy \wedge yRx \Rightarrow x = y$ .

**Definition 25.4.6** A total order on a set  $X$  is a binary relation  $R$  on  $X$  that is a transitive relation, an antisymmetric relation, and a connex relation.

**Definition 25.4.7** A trichotomous relation on a set  $X$  is a binary relation  $R$  on  $X$  such that  $\forall_{x,y \in X}$  precisely one of the following are true:  $xRy$ ,  $yRx$ , or  $x = y$ .

**Definition 25.4.8** A function  $f$  from a set  $A$  to a set  $B$ , denoted  $f : A \rightarrow B$ , is a subset  $f \subset A \times B$  such that  $\forall_{a \in A}$  there is a unique  $b \in B$  such that  $(a, b) \in f$ .

**Definition 25.4.9** The image of an element  $x \in A$  with respect to a function  $f : A \rightarrow B$ , denoted  $f(x)$ , is the unique element  $b \in B$  such that  $(a, b) \in f$ .

**Definition 25.4.10** A binary operation on a set  $A$  is a function  $* : A \times A \rightarrow A$

**Definition 25.4.11** The product of elements  $a, b \in A$  with respect to a binary operation  $*$  on  $A$ , denoted  $a * b$ , is the image of  $(a, b)$  with respect to  $*$ .

**Definition 25.4.12** A commutative operation on a set  $A$  is a binary operation  $*$  on  $A$  such that  $\forall_{x,y \in A}$ ,  $a * b = b * a$ .

**Definition 25.4.13** An associative operation on a set  $A$  is a binary operation  $*$  on  $A$  such that  $\forall_{a,b,c \in A}$ ,  $a * (b * c) = (a * b) * c$ .

**Definition 25.4.14** A binary operation that right distributes over a binary operation  $+$  on a set  $A$  is a binary operation  $\cdot$  on  $A$  such that  $\forall_{a,b,c \in A}$ ,  $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$ .

**Definition 25.4.15** A binary operation that left distributes over a binary operation  $+$  on a set  $A$  is a binary operation  $\cdot$  on  $A$  such that  $\forall_{a,b,c \in A}$ ,  $(b+c) \cdot a = (b \cdot a) + (c \cdot a)$ .

**Definition 25.4.16** A binary operation that distributes over a binary operation  $+$  on a set  $A$  is a binary operation  $\cdot$  on  $A$  such that  $\cdot$  both left and right distributes over  $+$ .

**Definition 25.4.17** A left identity in a set  $A$  with respect to a binary operation  $*$  is an element  $e \in A$  such that  $\forall_{a \in A}, e * a = a$ .

**Definition 25.4.18** A right identity in a set  $A$  with respect to a binary operation  $*$  is an element  $e \in A$  such that  $\forall_{a \in A}, a * e = a$ .

**Definition 25.4.19** An identity element in a set  $A$  with respect to a binary operation  $*$  on  $A$  is an element  $e \in A$  that is both a right and a left identity.

**Definition 25.4.20** A left inverse of an element  $a \in A$  with respect to a binary operation  $*$  on  $A$  and an identity  $e \in A$  is an element  $b \in A$  such that  $b * a = e$ .

**Definition 25.4.21** A right inverse of an element  $a \in A$  with respect to a binary operation  $*$  on  $A$  and an identity  $e \in A$  is an element  $b \in A$  such that  $a * b = e$ .

**Definition 25.4.22** A field is a commutative ring with unity  $(A, \cdot, +)$  such that  $\forall_{a \in A}$  such that  $a$  is not an identity with respect to  $\cdot$ ,  $\exists_{b \in A}$  such that  $b$  is an inverse of  $a$  with respect to  $\cdot$ .

## 25.4.2 Theorems

**Theorem 25.4.1.** If  $a, b \in \mathbb{R}^+$  and  $a < b$ , then  $a^2 < b^2$

*Proof.* If  $a < b$ , and  $0 < a$ , then  $a \cdot a < a \cdot b$  (Multiplicative Property of Ordered Fields)

But  $a \cdot a = a^2$ , and thus  $a^2 < a \cdot b$  (Definition of Exponents)

But if  $a < b$  and  $0 < b$ , then  $a \cdot b < b \cdot b$  (Multiplicative Property of Ordered Fields)

Therefore  $a \cdot b < b^2$  (Definition of Exponents)

But if  $a^2 < a \cdot b$  and  $a \cdot b < b^2$ , then  $a^2 < b^2$  (Transitive Property of Inequalities)

Therefore,  $a^2 < b^2$  □

**Theorem 25.4.2.** If  $a, b \in \mathbb{R}^+$  and  $a^2 = b^2$ , then  $a = b$ .

*Proof.* If  $a^2 = b^2$ , then  $b^2 - a^2 = 0$  (Additive Property of Ordered Fields)

If  $a < b$ , then  $0 < b - a$  (Additive Property of Ordered Fields)

If  $a, b \in \mathbb{R}^+$ , then  $0 < b + a$  (Closure of Addition in  $\mathbb{R}^+$ )

If  $0 < b - a$  and  $0 < b + a$ , then  $0 < (b - a) \cdot (b + a)$  (Closure of Multiplication in  $\mathbb{R}^+$ )

But  $(b - a) \cdot (b + a) = b^2 - a^2$  (Elementary Algebra)

But  $b^2 - a^2 = 0$ , and thus  $0 < b^2 - a^2$  (Trichotomous Property of Inequalities)

Therefore,  $a \not< b$ . Similarly,  $b \not< a$  (Proof by Contradiction)

But if  $a \not< b$  and  $b \not< a$ , then  $a = b$  (Trichotomous Property of Inequalities)

Therefore,  $a = b$  □

**Theorem 25.4.3.** If  $y \in (0, 1)$ , then there is an  $x \in (0, 1)$  such that  $y = x^2$ .

*Proof.* Let  $A = \{\xi \in \mathbb{R}^+ : \xi^2 \leq y\}$ .

But  $y \in (0, 1)$  and therefore  $y < 1$

(Definition of  $(0, 1)$ )

Therefore  $A$  is bounded above by 1.

Thus there exists a least upper bound  $x$  of  $A$

(Completeness of  $\mathbb{R}$ )

If  $x^2 > y$ , then  $\frac{x^2-y}{2} > 0$ .

(Additive Property of Ordered Fields)

Then  $(x - \frac{x^2-y}{2})^2 = a + (\frac{x^2-y}{2})^2 > y$  (Additive Property of Ordered Fields)

Thus  $x - \frac{x^2-y}{2}$  is an upper bound of  $A$ . (Definition of Upper Bounds)

But  $x$  is the least upper bound, a contradiction.

Therefore  $x^2 \not> y$ .

(Proof by Contradiction)

If  $x^2 < y$ , then  $0 < \frac{y-x^2}{2x+1}$

(Additive Property of Ordered Fields)

Let  $\epsilon = \min\{\frac{y-x^2}{2x+1}, 1\}$

Then  $(x + \epsilon)^2 = x^2 + 2x\epsilon + \epsilon^2 < x^2 + 2x\epsilon + \epsilon$

(Thm. 25.4.1)

But  $x^2 + 2x\epsilon + \epsilon = x^2 + (2x + 1)\epsilon = y$  (Elementary Algebra)

Therefore  $(x + \epsilon)^2 < y$  (Transitive Property of Total Orders)

But  $\epsilon > 0$ , and thus  $x + \epsilon > x$ . (Additive Property of Ordered Fields)

Thus  $x + \epsilon \in A$ , a contradiction as  $x$  is a least upper bound of  $A$ .

Therefore  $x^2 \not< y$ .

(Proof by Contradiction)

Therefore  $x^2 = y$

(Trichotomous Property of Inequalities)

□

## 25.5 Cheat Sheet

### 25.5.1 Series

**Theorem 25.5.1** (The Comparison Test). If  $a_n$  and  $b_n$  are sequences of non-negative real numbers, if there is an  $N_0 \in \mathbb{N}$  such that for all  $n > N_0$ ,  $a_n \leq b_n$ , and if  $\sum_{n=0}^N b_n$  converges, then  $\sum_{n=0}^N a_n$  converges.

*Proof.* Let  $B_N = \sum_{n=0}^N b_n$ ,  $S = \sum_{n=0}^{N_0-1} a_n$  and  $B = \lim_{n \rightarrow \infty} B_n$ . As  $b_n$  is non-negative, for all  $N \in \mathbb{N}$ ,  $B_N \leq B$ . Then for  $N > N_0$ ,  $\sum_{n=0}^N a_n = S + \sum_{n=N_0}^N a_n \leq S + \sum_{n=N_0}^N b_n \leq S + B_N \leq S + B$ . That is,  $\sum_{n=0}^N a_n$  bounded by  $S + B$ . And as  $a_n$  is non-negative,  $\sum_{n=0}^N a_n$  increasing monotonically, and is thus a monotonically increases sequence that is bounded above, and therefore converges. □

**Theorem 25.5.2** (Generalized Geometric Series Theorem). If  $r$  is a real number, then  $\sum_{n=0}^N r^n = \frac{1-r^{N+1}}{1-r}$

*Proof.* By induction. The base case is  $1 + r = (1 + r) \frac{1-r}{1-r} = \frac{1-r^2}{1-r}$ . Suppose

it is true for  $N \in \mathbb{N}$ . Then  $\sum_{n=0}^{N+1} r^n = \sum_{n=0}^N r^n + r^{N+1} = \frac{1-r^{N+1}}{1-r} + r^{N+1} = \frac{1-r^{N+1}+r^{N+1}(1-r)}{1-r} = \frac{1-r^{N+2}}{1-r}$ .  $\square$

**Theorem 25.5.3** (Geometric Series Theorem). *If  $|r| < 1$ , then  $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$*

*Proof.* For  $\sum_{n=0}^N r^n = \frac{1-r^{N+1}}{1-r}$ . As  $|r| < 1$ ,  $\lim_{N \rightarrow \infty} r^N = 0$ . Therefore  $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$ .  $\square$

**Theorem 25.5.4** (First Root Test Theorem). *If  $a_n$  is positive and  $\lim_{n \rightarrow \infty} \sqrt[n]{a_n} < 1$ , then  $\sum_{n=0}^{\infty} a_n$  converges.*

*Proof.* If  $\lim_{n \rightarrow \infty} \sqrt[n]{a_n} < 1$ , then there is a  $c \in (0, 1)$  and an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $\sqrt[n]{a_n} < c$ . But then  $a_n < c^n < 1$ , so  $\sum_{n=0}^{\infty} a_n = \frac{1}{1-c}$ . Thus, by the comparison test,  $\sum_{n=0}^N a_n$  converges.  $\square$

**Theorem 25.5.5** (Second Root Test Theorem). *If  $a_n$  is positive and  $\lim_{n \rightarrow \infty} \sqrt[n]{a_n} > 1$ , then  $\sum_{n=0}^{\infty} a_n$  diverges.*

*Proof.* If  $\lim_{n \rightarrow \infty} \sqrt[n]{a_n} > 1$ , then there is a  $c \in (1, \infty)$  and an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $\sqrt[n]{a_n} > c$ . But then  $a_n > c^n > 1$ , so  $\sum_{n=0}^{\infty} a_n$  diverges. By the comparison test,  $\sum_{n=0}^N a_n$  diverges.  $\square$

**Theorem 25.5.6.** *There exists sequences of positive real numbers  $a_n$  such that  $\sqrt[n]{a_n} \rightarrow 1$  and  $\sum_{n=1}^N a_n$  diverges.*

*Proof.* Let  $a_n = \frac{1}{n}$ . Then  $\ln(\sqrt[n]{a_n}) = \frac{\ln(\frac{1}{n})}{n} = -\frac{\ln(n)}{n} \rightarrow 0$ . So  $\sqrt[n]{a_n} \rightarrow 1$ . But  $\sum_{n=1}^N \frac{1}{n}$  diverges.  $\square$

**Theorem 25.5.7.** *There exists positive sequences  $a_n$  such that  $\sqrt[n]{a_n} \rightarrow 1$  and  $\sum_{n=1}^N a_n$  converges.*

*Proof.* Let  $a_n = \frac{1}{n^2}$ . Then  $\ln(\sqrt[n]{a_n}) = -2\frac{\ln(n)}{n} \rightarrow 0$ . Therefore  $\sqrt[n]{a_n} \rightarrow 1$ . But  $\sum_{n=1}^N \frac{1}{n^2}$  converges.  $\square$

**Theorem 25.5.8.** *If  $a_n$  is a sequences of positive real numbers, and if  $a(x)$  is a monotonically decreasing function such that for all  $n \in \mathbb{N}$ ,  $a(n) = a_n$ , then  $\sum_{n=1}^{\infty} a_n$  converges if and only if  $\int_1^{\infty} a(x)dx$  converges.*

*Proof.* As  $a(x)$  is decreasing, for all  $n \in \mathbb{N}$  and for all  $x \in [n, n+1]$ ,  $a_{n+1} \leq a(x) \leq a_n$ . Therefore  $\int_n^{n+1} a_{n+1} dx \leq \int_n^{n+1} a(x) dx \leq \int_n^{n+1} a_n dx \Rightarrow a_{n+1} \leq \int_n^{n+1} a(x) dx \leq a_n$ . But for all  $N \in \mathbb{N}$ ,  $\sum_{n=1}^N \int_n^{n+1} a(x) dx = \int_{n=1}^N a(x) dx$ , and thus  $\sum_{n=1}^N a_{n+1} \leq \int_{n=1}^N a(x) dx \leq \sum_{n=1}^N a_n$ . Let  $S_N = \sum_{n=1}^N a_n$ . Then  $S_{N+1} - a_1 \leq \int_{n=1}^N a(x) dx \leq S_N$ . If  $\int_{n=1}^N a(x) dx$  converges, say to  $I$ , then  $S_{N+1} \leq I + a_1$ . But  $S_{N+1}$  is a monotonically increasing sequence that is

bounded, and therefore converges. That is, if  $\int_{n=1}^N a(x)dx$  converges, then  $\sum_{n=1}^N a_n$  converges. If  $\sum_{n=1}^N a_n$  converges, say to  $S$ , then  $\int_{n=1}^N a(x)dx \leq S$ . But  $a(x)$  is monotonically decreasing and positive, and thus  $I_N = \int_{n=1}^N a(x)dx$  is a bounded monotonically increasing sequence, and therefore converges. That is, if  $\sum_{n=1}^N a_n$  converges, then  $\int_1^N a(x)dx$  converges.  $\square$

**Theorem 25.5.9.** *If  $a_n$  is a positive, monotonically decreasing, and if  $a_n \rightarrow 0$ , then  $\sum_{n=1}^{\infty} (-1)^n a_n$  converges.*

*Proof.* Let  $S_N = \sum_{n=1}^N (-1)^n a_n$ . If  $N$  is even, then  $S_N = \sum_{n=1}^{N/2} (a_{2n} - a_{2n-1})$ . But  $a_n$  is decreasing monotonically, so  $a_{N+2} - a_{N+1} \leq 0$ . That is, if  $N$  is even then  $S_N \geq S_{N+2}$ . So  $S_{2k}$  is a monotonically decreasing subsequence of  $S_N$ . Moreover,  $S_{2k} \geq a_2 - a_1$ . Thus,  $S_{2k}$  converges, say to  $S_1$ . If  $N$  is odd, then  $S_N = -a_N + \sum_{n=1}^{(N-1)/2} (a_{2n} - a_{2n-1})$ . But then  $S_{N+2} = -(a_{N+2} - a_{N+1}) + S_N \geq S_N$ . That is,  $S_{2k-1}$  is a monotonically increasing subsequence. Moreover,  $S_{2k-1} \leq a_2$ . So,  $S_{2k-1}$  converges, say to  $S_2$ . Let  $\epsilon > 0$  be given. As  $a_n \rightarrow 0$ , there is an  $N_0 \in \mathbb{N}$  such that for all  $n \geq N_0$ ,  $a_n < \frac{\epsilon}{2}$ . There is also an  $N_1$  such that for  $n > N_1$ ,  $|S_1 - S_{2n}| < \frac{\epsilon}{4}$ . Finally there is an  $N_2$  such that for  $n > N_2$ ,  $|S_2 - S_{2n-1}| < \frac{\epsilon}{4}$ . Let  $N = \max\{N_0, N_1, N_2\}$ . Then for  $n > N$ , we have  $|S_1 - S_2| \leq |S_1 - S_{2m}| + |S_2 - S_{2m-1}| + |S_{2m} - S_{2n-1}| < \frac{\epsilon}{4} + \frac{\epsilon}{4} + |a_{2n}| < \epsilon$ . But  $S_1$  and  $S_2$  are real numbers, and  $\epsilon$  is arbitrary, so  $S_1 = S_2$ . Therefore  $S_N \rightarrow S_1$ .  $\square$

**Theorem 25.5.10.** *There exists positive sequences  $a_n$  such that  $a_n \rightarrow 0$  and  $\sum_{n=1}^N (-1)^n a_n$  diverges.*

*Proof.* For let  $a_n = \frac{1}{n}$  for odd  $n$  and  $a_n = \frac{1}{n^2}$  for even  $n$  and let  $S_N = \sum_{n=1}^N (-1)^n a_n$ . Then  $S_{2N} = \sum_{n=1}^N (\frac{1}{n^2} - \frac{1}{n}) = \sum_{n=1}^N \frac{1}{n^2} - \sum_{n=1}^N \frac{1}{n}$ . But  $\sum_{n=1}^N \frac{1}{n^2}$  converges and  $\sum_{n=1}^N \frac{1}{n}$  diverges, and therefore  $S_{2N}$  diverges. But if  $S_{2N}$  diverges, then  $S_N$  diverges.  $\square$

**Theorem 25.5.11.** *If  $a_n$  is a sequence of real numbers and  $\sum_{n=1}^N |a_n|$  converges, then  $\sum_{n=1}^N a_n$  converges.*

*Proof.* So let  $T_N = \sum_{n=1}^N |a_n|$  and let  $S_N = \sum_{n=1}^N a_n$ . As  $T_N$  converges, it is a Cauchy Sequence. That is, for all  $\epsilon > 0$  there is an  $N_0 \in \mathbb{N}$  such that for all  $n, m > N_0$ ,  $|T_n - T_m| < \epsilon$ . But then for  $n, m > N_0$ ,  $|S_n - S_m| = |\sum_{n=m+1}^n a_n| \leq \sum_{n=m+1}^n |a_n| = |T_n - T_m| < \epsilon$ . That is  $S_N$  forms a Cauchy sequence. But Cauchy sequences converge. Therefore  $S_N$  converges.  $\square$

## 25.5.2 Complex Variables

**Theorem 25.5.12.** *If  $x \in \mathbb{C}$ , then  $e^{ix} = \cos(x) + i \sin(x)$*

*Proof.* For  $e^{ix}$  is the solution to  $y' = iy$ ,  $y(0) = 1$ . But  $\frac{d}{dx}(\cos(x) + i \sin(x))i(\cos(x) + i \sin(x))$ . Moreover,  $\cos(0) + i \sin(0) = 1$ . From the uniqueness of solutions,  $e^{ix} = \cos(x) + i \sin(x)$ .  $\square$

**Theorem 25.5.13.** If  $x \in \mathbb{C}$ , then  $\cos(x) = \frac{1}{2}(e^{ix} + e^{-ix})$

*Proof.* For  $\cos(x) = \cos(-x)$  and  $\sin(x) = -\sin(-x)$ . From the previous theorem  $e^{ix} + e^{-ix} = 2 \cos(x)$ .  $\square$

**Definition 25.5.1** The hyperbolic cosine of  $x \in \mathbb{R}$  is  $\cosh(x) = \cos(ix)$ .

**Theorem 25.5.14.** If  $x \in \mathbb{R}$ , then  $\cosh(x) = \frac{e^{ix} + e^{-ix}}{2}$ .

*Proof.* Apply the previous theorem to  $ix$ .  $\square$

**Theorem 25.5.15.** If  $x \in \mathbb{C}$ , then  $\sin(x) = \frac{1}{2i}(e^{i\theta} - e^{-i\theta})$

*Proof.* For  $e^{ix} = \cos(x) + i \sin(x)$ , and thus  $e^{ix} - e^{-ix} = 2i \sin(x)$   $\square$

**Definition 25.5.2** The hyperbolic sine of  $x \in \mathbb{R}$  is  $\sinh(x) = -i \sin(ix)$

**Theorem 25.5.16.** If  $x \in \mathbb{R}$ , then  $\sinh(x) = \frac{e^x - e^{-x}}{2}$

*Proof.* Apply the previous theorem to  $ix$  and multiply by  $-i$ .  $\square$

**Definition 25.5.3** A complex function that is differentiable at a point  $z_0 \in \mathbb{C}$  is a function  $f : \mathbb{C} \rightarrow \mathbb{C}$  such that  $\lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$

**Definition 25.5.4** A differentiable complex function is a function  $f : \mathbb{C} \rightarrow \mathbb{C}$  that is differentiable for all  $z \in \mathbb{C}$ .

**Theorem 25.5.17** (The Cauchy-Riemann Theorem). A function  $f : \mathbb{C} \rightarrow \mathbb{C}$ ,  $f(z) = u(x, y) + iv(x, y)$ , where  $u, v : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is differentiable if and only if  $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$  and  $\frac{\partial u}{\partial y} = 0 \frac{\partial v}{\partial x}$ .

**Theorem 25.5.18.** If  $z \in \mathbb{R}$ , then there is a unique  $r > 0$  and  $\theta \in [0, 2\pi)$  such that  $z = re^{i\theta}$ .

**Theorem 25.5.19** (Cauchy's Integral Theorem). If  $\mathcal{U} \subset \mathbb{C}$  is simply connected, if  $f : \mathcal{U} \rightarrow \mathbb{C}$  is differentiable, and if  $\gamma : I \rightarrow \mathcal{U}$  is a closed path of finite measure (Length), then  $\oint_{\gamma} f(z) dz = 0$ .

**Theorem 25.5.20** (Cauchy's Integral Formula). If  $\mathcal{U} \subset \mathbb{C}$  is open,  $z_0 \in \mathcal{U}$ , and if  $B_r(z_0) \subset \mathcal{U}$ , then for all  $a \in B_r(z_0)$ ,  $\oint_{\partial B_r(z_0)} \frac{f(z)}{z-a} dz = f(a)$

### 25.5.3 Matrices

**Definition 25.5.5** The transpose of a matrix  $A = (a_{ij})$  is  $A^T = (a_{ji})$ .

**Definition 25.5.6** The complex transpose of a matrix  $A = (a_{ij})$  is  $A^\dagger = (\overline{a_{ji}})$ .

**Definition 25.5.7** An orthogonal matrix is a matrix  $A$  such that  $A^T = A^{-1}$ .

**Definition 25.5.8** A Hermitian Matrix is a matrix  $A$  such that  $A^\dagger = A$ .

**Definition 25.5.9** A unitary matrix is a matrix  $A$  such that  $A^\dagger = A^{-1}$ .

**Definition 25.5.10** A singular matrix is a matrix with no inverse.

**Definition 25.5.11** The Commutator of two matrices  $A$  and  $B$  is  $[A, B] = AB - BA$ .

**Definition 25.5.12** A normal matrix is a matrix  $A$  such that  $[A, A^\dagger] = 0$ .

**Definition 25.5.13** The Levi-Civita symbol is defined as:

$$\varepsilon_{ijk} = \begin{cases} 1, & ijk \text{ is an even permutation} \\ 0, & i = j \text{ OR } i = k \text{ OR } j = k \\ -1, & ijk \text{ is an odd permutation} \end{cases}$$

**Definition 25.5.14** The Matrix Representation of  $a+ib$  is the matrix  $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

**Theorem 25.5.21.**  $\mathbb{C}$ , with it's usual arithmetic, is homomorphic to  $\mathbb{R}^{2 \times 2}$ , with it's usual arithmetic.

*Proof.* For let  $f : \mathbb{C} \rightarrow \mathbb{R}^{2 \times 2}$  be defined by  $f(z) = f(a+ib) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

$$f(z+w) = \begin{bmatrix} a+b & -(b+d) \\ b+d & a+c \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = f(z) + f(w)$$

$$f(z \cdot w) = \begin{bmatrix} ac-bd & -(ad+bc) \\ ad+bc & ac-bd \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = f(z) \cdot f(w)$$

□

**Definition 25.5.15** The  $ij$  minor of an  $n \times m$  matrix  $A$  is the matrix  $M_{ij} = \{(k, \ell, a_{k\ell}), k \neq i, j \neq \ell\}$ . That is, it is the matrix formed by removing the  $i^{th}$  row and  $j^{th}$  column from  $A$ .

**Definition 25.5.16** The cofactor matrix of an  $n \times n$  matrix  $A$  is the matrix  $C = (C_{ij})$ , where  $C_{ij} = (-1)^{i+j} M_{ij}$ .

**Theorem 25.5.22.** If  $A$  is an  $n \times n$  matrix, and  $\det(A) \neq 0$ , then  $A^{-1} = \frac{1}{\det(A)} C^T$ .

### 25.5.4 Vectors

**Theorem 25.5.23.** If  $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^3$ , then  $\langle \mathbf{A}, \mathbf{B} \times \mathbf{C} \rangle = \langle \mathbf{B}, \mathbf{C} \times \mathbf{A} \rangle$ .

*Proof.* For  $\mathbf{B} \times \mathbf{C} = (B_y C_z - B_z C_y, B_z C_x - B_x C_z, B_x C_y - B_y C_x)$ , and thus:

$$\begin{aligned}\langle \mathbf{A}, \mathbf{B} \times \mathbf{C} \rangle &= A_x(B_y C_z - B_z C_y) + A_y(B_z C_x - B_x C_z) + A_z(B_x C_y - B_y C_x) \\&= A_x B_y C_z - A_x B_z C_y + A_y B_z C_x - A_y B_x C_z + A_z B_x C_y - A_z B_y C_x \\&= B_x(A_z C_y - A_y C_z) + B_y(A_x C_z - A_z C_x) + B_z(A_y C_x - A_x C_y) \\&= B_x(C_y A_z - C_z A_y) + B_y(C_z A_x - C_x A_z) + B_z(C_x A_y - C_y A_x) \\&= \langle \mathbf{B}, (C_y A_z - C_z A_y, C_z A_x - C_x A_z, C_x A_y - C_y A_x) \rangle \\&= \langle \mathbf{B}, \mathbf{C} \times \mathbf{A} \rangle\end{aligned}$$

□

**Theorem 25.5.24.** If  $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^3$ , then  $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \langle \mathbf{A}, \mathbf{C} \rangle \mathbf{B} - \langle \mathbf{A}, \mathbf{B} \rangle \mathbf{C}$

*Proof.* For  $\mathbf{B} \times \mathbf{C} = (B_y C_z - B_z C_y, B_z C_x - B_x C_z, B_x C_y - B_y C_x)$ , and thus:

$$\begin{aligned}\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= \mathbf{A} \times (B_y C_z - B_z C_y, B_z C_x - B_x C_z, B_x C_y - B_y C_x) \\&= (A_y(B_x C_y - B_y C_x) - A_z(B_z C_x - B_x C_z))\hat{\mathbf{x}} \\&\quad + (A_z(B_y C_z - B_z C_y) - A_x(B_x C_y - B_y C_x))\hat{\mathbf{y}} \\&\quad + (A_x(B_z C_x - B_x C_z) - A_y(B_y C_z - B_z C_y))\hat{\mathbf{z}} \\&= (B_x(A_x C_x + A_y C_y + A_z C_z) - C_x(A_x B_x + A_y B_y + A_z B_z))\hat{\mathbf{x}} \\&\quad + (B_y(A_x C_x + A_y C_y + A_z C_z) - C_y(A_x B_x + A_y B_y + A_z B_z))\hat{\mathbf{y}} \\&\quad + (B_z(A_x C_x + A_y C_y + A_z C_z) - C_z(A_x B_x + A_y B_y + A_z B_z))\hat{\mathbf{z}} \\&= (A_x C_x + A_y C_y + A_z C_z)(B_x, B_y, B_z) \\&\quad - (A_x B_x + A_y B_y + A_z B_z)(C_x, C_y, C_z) \\&= \langle \mathbf{A}, \mathbf{C} \rangle \mathbf{B} - \langle \mathbf{A}, \mathbf{B} \rangle \mathbf{C}\end{aligned}$$

□

**Theorem 25.5.25** (Divergence Theorem). If  $\Omega$  is a closed bounded subset of  $\mathbb{R}^n$  with a smooth boundary  $\partial\Omega$ , and if  $\mathbf{A}$  is a smooth vector field, then  $\iiint_{\Omega} (\nabla \cdot \mathbf{A}) d\tau = \oint_{\partial\Omega} \mathbf{A} \cdot d\sigma$

**Theorem 25.5.26** (Stokes' Theorem). If  $\Sigma$  is a compact, simply connected subset of  $\mathbb{R}^3$  bounded by a smooth Jordan Curve  $\partial\Sigma$ , and if  $\mathbf{A}$  is a smooth vector field, then  $\iint_{\Sigma} (\nabla \times \mathbf{A}) \cdot d\sigma = \oint_{\partial\Sigma} \mathbf{A} \cdot d\ell$

**Theorem 25.5.27.** If  $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$  is a smooth function, then  $\nabla \times \nabla(\phi) = \mathbf{0}$ .

*Proof.* For  $\nabla(\phi) = \frac{\partial\phi}{\partial x}\hat{\mathbf{x}} + \frac{\partial\phi}{\partial y}\hat{\mathbf{y}} + \frac{\partial\phi}{\partial z}\hat{\mathbf{z}}$ , and therefore:

$$\nabla \times \nabla(\phi) = \left( \frac{\partial^2\phi}{\partial y\partial z} - \frac{\partial^2\phi}{\partial z\partial y} \right) \hat{\mathbf{x}} + \left( \frac{\partial^2\phi}{\partial z\partial x} - \frac{\partial^2\phi}{\partial x\partial z} \right) \hat{\mathbf{y}} + \left( \frac{\partial^2\phi}{\partial x\partial y} - \frac{\partial^2\phi}{\partial y\partial x} \right) \hat{\mathbf{z}}$$

But, as  $\phi$  is smooth,  $\frac{\partial^2\phi}{\partial x_i\partial x_j} = \frac{\partial^2\phi}{\partial x_j\partial x_i}$ . Therefore,  $\nabla \times \nabla(\phi) = \mathbf{0}$ .  $\square$

**Theorem 25.5.28.** *If  $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  is smooth, then  $\nabla \cdot (\nabla \times \mathbf{A}) = 0$ .*

*Proof.* For  $\nabla \times \mathbf{A} = (\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z})\hat{\mathbf{x}} + (\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x})\hat{\mathbf{y}} + (\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y})\hat{\mathbf{z}}$ , and thus:

$$\begin{aligned} \nabla \cdot (\nabla \times \mathbf{A}) &= \nabla \cdot \left( \left( \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{\mathbf{x}} + \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{\mathbf{y}} + \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{\mathbf{z}} \right) \\ &= \left( \frac{\partial^2 A_z}{\partial x \partial y} - \frac{\partial^2 A_y}{\partial x \partial z} \right) + \left( \frac{\partial^2 A_x}{\partial y \partial z} - \frac{\partial^2 A_z}{\partial y \partial x} \right) + \left( \frac{\partial^2 A_y}{\partial z \partial x} - \frac{\partial^2 A_x}{\partial z \partial y} \right) \\ &= \left( \frac{\partial^2 A_z}{\partial x \partial y} - \frac{\partial^2 A_z}{\partial y \partial x} \right) + \left( \frac{\partial^2 A_y}{\partial z \partial x} - \frac{\partial^2 A_y}{\partial x \partial z} \right) + \left( \frac{\partial^2 A_x}{\partial y \partial z} - \frac{\partial^2 A_x}{\partial z \partial y} \right) \end{aligned}$$

But  $A_x, A_y$ , and  $A_z$  are smooth, and thus  $\frac{\partial^2 A_k}{\partial x_i \partial x_j} = \frac{\partial^2 A_k}{\partial x_j \partial x_i}$ . Therefore,  $\nabla \cdot (\nabla \times \mathbf{A}) = \mathbf{0}$ .  $\square$

# CHAPTER 26

---

## Measurable Spaces

---

### 26.1 Set Rings

Given a set  $\Omega$ ,  $\mathcal{P}(\Omega)$  is the set of all subsets of  $\Omega$ . Often this is too much, and too difficult to handle. Indeed, even  $\mathcal{P}(\mathbb{R})$  is quite large and hard to get a grasp on. We wish to speak of collections of sets that have some structure on them. The first thing we will talk about is a set ring.

#### Definition 26.1.1: Set Ring

set ring of a set  $\Omega$  is a nonempty subset  $\mathcal{R} \subseteq \mathcal{P}(\Omega)$  such that for all  $A, B \in \mathcal{R}$ ,  $A \cup B \in \mathcal{R}$ , and  $A \setminus B \in \mathcal{R}$ .

#### Example 26.1.1: Example of Set Rings

If  $\Omega$  is a set, then  $\mathcal{P}(\Omega)$  is a set ring of  $\Omega$ . So is the set  $R = \{\emptyset\}$ . For any  $A \subset \Omega$ , the set  $R = \{A\}$  is also a set ring. If  $\Omega = \{1, 2, 3\}$ , then  $R = \{\emptyset, \{1\}, \{2, 3\}, \{1, 2, 3\}\}$  is a set ring on  $\Omega$ .

#### Example 26.1.2: I

---

$\Omega$  is an infinite set, and if  $\mathcal{E} = \{\{x\} : x \in \Omega\}$ , then the smallest set ring that contains  $\mathcal{E}$  is the set of all finite subsets of  $\Omega$ . For the union of two finite sets is finite, as is the set difference of two finite sets, and thus this satisfies a set ring. Moreover, if  $\mathcal{R}$  is a set ring that contains  $\mathcal{E}$  then it contains the union of

any finite collection of elements in  $\mathcal{E}$ . But  $\mathcal{E}$  is the set of all of the singletons of  $\Omega$ , and any finite subset of  $\Omega$  can be written as the union of finitely many singletons. Thus,  $\mathcal{R}$  is the smallest set ring that contains  $\mathcal{E}$ . ■

**Theorem 26.1.1.** *If  $\Omega$  is a set, if  $R$  is a set ring on  $\Omega$ , and if  $A$  is a finite subset of  $R$ , then  $\cup_{\alpha \in A} \alpha$  is an element of  $R$ .*

*Proof.* Apply induction to the closure of unions. □

**Theorem 26.1.2.** *If  $X$  is a set, if  $R$  is a set ring on  $X$ , and if  $A, B \in R$ , then  $A \cup B \in R$ .*

*Proof.* For  $A \cap B = A \setminus (A \setminus B)$ , and from the closure of set difference,  $A \cap B \in R$ . □

**Theorem 26.1.3.** *If  $X$  is a set, if  $R$  is a set ring on  $X$ , and if  $A$  is a finite subset of  $R$ , then  $\cap_{\alpha \in A} \alpha$  is an element of  $R$ .*

*Proof.* Apply induction to the closure of intersections. □

**Theorem 26.1.4.** *If  $\Omega$  is a set, if  $R$  is a set ring on  $\Omega$ , if  $A, B \subset \Omega$ , and if  $A \setminus B, B \setminus A$ , and  $A \cap B$  are elements of  $R$ , then  $A, B \in R$ .*

Thus, the set ring generated by the set  $\{A, B\}$  and the set ring generated by  $\{A \setminus B, B \setminus A, A \cap B\}$  are the same.

**Theorem 26.1.5.** *If  $\Omega$  is a set and  $R$  is a set ring of  $\Omega$ , then  $\emptyset \in R$ .*

*Proof.* For as  $R$  is non-empty, there is an element  $A \in R$ . If  $A = \emptyset$ , then we are done. If not, as  $R$  is closed under set difference,  $A \setminus A \in R$ . But  $A \setminus A = \emptyset$ . □

From this, if we have a collection  $R$  of subsets of  $\Omega$  and we wish to check if  $R$  is a set ring of  $\Omega$ , there are several redundant operations we don't need to check. Since, for any set  $A$ , we have:

$$A \setminus \emptyset = A \tag{26.1.1}$$

$$A \setminus A = \emptyset \tag{26.1.2}$$

$$\emptyset \setminus A = \emptyset \tag{26.1.3}$$

$$A \cup A = A \tag{26.1.4}$$

$$A \cup \emptyset = A \tag{26.1.5}$$

$$\emptyset \cup \emptyset = \emptyset \tag{26.1.6}$$

Using our previous example  $\Omega = \{1, 2, 3\}$ , we can check laboriously that  $R = \{\emptyset, \{1\}, \{2, 3\}, \{1, 2, 3\}\}$  is a set ring on  $\Omega$ . The set  $\{\emptyset, \{1\}, \{2\}, \{1, 2, 3\}\}$  is not a set ring, for  $\{1, 2\} = \{1\} \cup \{2\}$  is not an element.

**Theorem 26.1.6.** *If  $\Omega$  is a set, and if  $A$  and  $B$  are disjoint subsets of  $\Omega$ , then  $R = \{\emptyset, A, B, A \cup B\}$  is a set ring on  $\Omega$ .*

**Theorem 26.1.7.** *If  $\Omega$  is a set, if  $A$  and  $B$  are disjoint subsets of  $\Omega$ , and if  $R$  is a set ring such that  $A, B \in R$ , then  $\{\text{emptyset}, A, B, A \cup B\} \subset R$ .*

As such, the set ring  $\{\emptyset, A, B, A \cup B\}$  is called the set ring generated by  $A$  and  $B$ . We can continue and consider the case of three mutually disjoint subsets.

**Theorem 26.1.8.** *If  $\Omega$  is a set, and  $A_1, A_2, A_3$  are mutually disjoint subsets of  $\Omega$ , then:*

$$R = \{\emptyset, A_1, A_2, A_3, A_1 \cup A_2, A_1 \cup A_3, A_2 \cup A_3, A_1 \cup A_2 \cup A_3\} \quad (26.1.7)$$

*is a set ring on  $\Omega$ .*

Indeed, we may generalize further.

**Theorem 26.1.9.** *If  $\Omega$  is a set and if  $A$  is a subset of  $\mathcal{P}(\Omega)$  of  $n$  elements such that, for all  $a, b \in A$ ,  $a \cap b = \emptyset$ , then:*

$$R = \{\cup_{i \in I} A_i : I \in \mathcal{P}(\mathbb{Z}_n)\} \quad (26.1.8)$$

*Is a set ring on  $\Omega$ .*

**Theorem 26.1.10.** *If  $\Omega$  is a set, then the set of all finite subsets of  $\Omega$  is a set ring on  $\Omega$ .*

A left semi-interval of  $\mathbb{R}$  is an interval of the form  $[a, b)$  where  $a \leq b$ . If  $a = b$ , this is the empty set. The set of all left semi-intervals is not a set ring on  $\mathbb{R}$  since the union of two semi-intervals need not be a semi-interval. We need to add more sets to allow this to be a set ring. The collection of all finite unions of semi-intervals of  $\mathbb{R}$  is a set ring. First, note the following:

$$\left( \bigcup_{n=1}^N [a_n, b_n) \right) \setminus [c, d) = \bigcup_{n=1}^N ([a_n, b_n) \setminus [c, d]) \quad (26.1.9)$$

This is again the finite union of intervals. By induction we see that this collection is a ring on  $\mathbb{R}$ . We have seen that a set ring is closed to unions and set differences, and this implies that rings are closed under intersections and closed under symmetric differences. As it turns out, this is an equivalent definition of a set ring.

**Theorem 26.1.11.** *If  $\Omega$  is a set and  $R \subset \mathcal{P}(\Omega)$ , then  $R$  is a set ring of  $\Omega$  if and only if  $R$  is closed under symmetric differences and intersections.*

If  $R$  is a set ring on  $\Omega$ , and if  $A \in R$ , let  $\chi_A : \Omega \rightarrow [0, 1]$  be the indicator function defined as follows:

$$\chi_A(\omega) = \begin{cases} 0, & \omega \notin A \\ 1, & \omega \in A \end{cases} \quad (26.1.10)$$

Then we have:

$$\chi_{A \cap B}(\omega) = \chi_A(\omega)\chi_B(\omega) \quad (26.1.11)$$

$$\chi_{A \oplus B} = (\chi_A(\omega) + \chi_B(\omega)) \mod 2 \quad (26.1.12)$$

These two operations satisfy the axioms of a ring. That is, a ring in the algebraic sense of the word: A set with two operations that behave certain properties. It is because of this that set rings have earned their name.

## 26.2 Set Algebras

**Definition 26.2.1** A set algebra on a set  $\Omega$  is a set ring on  $\Omega$  such that  $\Omega \in \mathcal{A}$ .

**Example 26.2.1** Let  $\Omega = \{1, 2, 3, 4\}$  and  $R = \{\emptyset, \{1\}, \{2, 3\}\}$ . Then  $R$  is a set ring, but it is not a set algebra since  $\Omega \notin R$ .

### Example 26.2.2: I

$\Omega$  is an infinite set, and if  $\mathcal{E} = \{\{x\} : x \in \Omega\}$ , then the smallest set algebra that contains  $\mathcal{E}$  is the set of all finite and co-finite subsets of  $\Omega$ . There are a few cases to check. The finite union of finite subsets is finite, the finite union of co-finite subsets is co-finite, and the finite union of finite and co-finite is again co-finite. For set difference, the difference of finite with finite is again finite, and the difference of co-finite with co-finite is either co-finite or finite. The difference of co-finite with finite is co-finite, and the difference of finite with co-finite is finite. Thus, this set is a set algebra on  $\Omega$ . Moreoever it is the smallest set algebra that will contain  $\mathcal{E}$ . ■

From the definition, we see that a set algebra is closed under complements. indeed, this creates and equivalent definition for set algebras.

**Theorem 26.2.1.** *If  $\Omega$  is a set and  $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ , then  $\mathcal{A}$  is a set algebra on  $\Omega$  if and only if  $\Omega \in \mathcal{A}$ , and  $\mathcal{A}$  is closed under union and complement.*

**Theorem 26.2.2.** *If  $\Omega$  is a set and  $R$  is a set ring on  $\Omega$ , and if  $\mathcal{A}$  is a set algebra on  $\Omega$  such that  $R \subset \mathcal{A}$ , then for all  $A \in R$ ,  $A \in \mathcal{A}$  and  $A^C \in \mathcal{A}$ .*

This then defines the *smallest* set algebra that contains a set ring.

**Theorem 26.2.3.** *If  $\Omega$  is a set and  $R$  is a set ring on  $\Omega$ , then:*

$$\mathcal{A} = \{A, A^C : A \in R\} \quad (26.2.1)$$

*Is a set algebra on  $\Omega$ .*

**Theorem 26.2.4.** *If  $\Omega$  is a set and  $A$  and  $B$  are disjoint subset of  $A$ , then:*

$$\mathcal{A} = \{\emptyset, A, B, A \cup B, \Omega, A^C, B^C, A^C \cap B^C\} \quad (26.2.2)$$

*is a set algebra on  $\Omega$ .*

For non-disjoint  $A$  and  $B$ , the smallest set algebra becomes more complicated. We saw that the collection of all finite subsets of a set is a set ring on the set. The collection of all finite subsets, and their complements, is a set algebra.

## 26.3 $\sigma$ -Rings

If  $\Omega$  is a set, then  $R \subset \mathcal{P}(\Omega)$  is called a set ring on  $\Omega$  if, for all  $A, B \in R$ ,  $A \cup B \in R$  and  $A \setminus B \in R$ . From this, given a ring  $R$  on  $\Omega$ , the empty set is included, that is  $\emptyset \in R$ , and if  $A, B \in R$ , then  $A \cap B \in R$ . By induction, for any finite collection of elements in  $R$ , the union of these subsets is also contained in  $R$ , as well as the intersection. A set algebra on  $\Omega$  is a ring  $\mathcal{A}$  on  $\Omega$  such that  $\Omega \in \mathcal{A}$ . That is,  $\mathcal{A} \subset \mathcal{P}(\Omega)$ , and  $\mathcal{A}$  is closed under union, set difference, and  $\Omega \in \mathcal{A}$ . There is an equivalent definition:  $\Omega \in \mathcal{A}$ , for all  $A \in \mathcal{A}$ ,  $A^C \in \mathcal{A}$ , and for all  $A, B \in \mathcal{A}$ ,  $A \cup B \in \mathcal{A}$ . The complement of  $A$ ,  $A^C$ , is defined as  $\Omega \setminus A$ . The equivalence of the two definitions comes from DeMorgan's laws, since  $A \setminus B = A \cap B^C = (A^C \cup B)^C$ . We now talk about  $\sigma$ -Ring.

**Definition 26.3.1** A  $\sigma$ -Ring on a set  $\Omega$  is a set  $\sigma \subset \mathcal{P}(\Omega)$  such that, for all countable subsets of  $\sigma$ , the union  $\bigcup_{i=1}^{\infty} A_i \in \sigma$ , and for all  $A, B \in \sigma$ ,  $A \setminus B \in \sigma$ .

The requirement that the collections be countable is important to note. A *topology* is a subset of  $\mathcal{P}(\Omega)$  with the property that it is closed under arbitrary unions.  $\sigma$ -Rings need only be closed under countable unions.

**Example 26.3.1** Every  $\sigma$ -Ring is a set ring, but not every set ring is a  $\sigma$ -ring. Let  $\Omega$  be uncountable, and let  $R$  be the set of all finite subsets of  $\Omega$ . Then  $R$  is a ring, but it is not a  $\sigma$ -ring. For, as  $\Omega$  is uncountably infinite, it has a countable subset  $A$ , and we may subscript the elements as  $a_n$ . But  $\bigcup_{n=1}^{\infty} \{a_n\}$  is not a finite subset of  $\Omega$ , and is therefore not contained in  $R$ . Thus,  $R$  is not closed under countable unions and  $R$  is not a  $\sigma$ -ring. However, if we let  $\sigma$  be the set of all *countable* subsets of  $\Omega$ , the  $\sigma$  is indeed a  $\sigma$ -ring.

**Example 26.3.2: T**

e collection of all semi-intervals and finite unions of semi-intervals defines a ring on  $\mathbb{R}$ . It is tempting to think tha the collection of all countable unions of semi-intervals is a  $\sigma$ -ring on  $\mathbb{R}$ , but this is not the case. The Cantor set is an example of a subset that can be constructed by a countable number of steps of removing intervals from a given interval, but the resulting set is not the countable union of semi-intervals. To construct the Cantor set, consider the interval  $[0, 1]$ . From this, remove  $(\frac{1}{3}, \frac{2}{3})$ . Continuing removing the middle third from each sub-interval obtained. The resulting set contains no interval as a subset, and thus cannot be the union of countably many intervals, or semi-intervals. ■

## 26.4 $\sigma$ -Algebras

In an analogous manner to how set rings and set algebras were defined, there is something called a  $\sigma$ -algebra. This notion will be one of the central themes of measure theory.

**Definition 26.4.1** A  $\sigma$ -algebra on a set  $\Omega$  is a  $\sigma$ -ring on  $\Omega$  such that  $\Omega \in \mathcal{A}$ .

That is, given any countable collection of elements in  $\mathcal{A}$ , the union is also contained in  $\mathcal{A}$ . In addition,  $\mathcal{A}$  is closed under set differences and  $\Omega \in \mathcal{A}$ .

**Example 26.4.1** The first trivial example is the power set  $\mathcal{P}(\Omega)$ . Also the set  $\{\emptyset, \Omega\}$  defines a  $\sigma$ -algebra on  $\Omega$ . The set of all countable subsets defines a  $\sigma$ -ring, and the set of all countable and co-countable (Sets whose complement is countable) will define a  $\sigma$ -algebra.

We can equivalently define a  $\sigma$ -algebra to be a collection of sets  $\mathcal{A}$  such that  $\Omega \in \mathcal{A}$ , for all  $A \in \mathcal{A}$ ,  $A^C \in \mathcal{A}$ , and  $\mathcal{A}$  is closed under countable unions. Being closed under countable unions implies that it is closed under finite unions as well. For let  $A_1 = A$ , and for all  $n > 1$ , let  $A_n = B$ . Then  $\bigcup_{n=1}^{\infty} A_n = A \cup B$ . By induction, a  $\sigma$ -algebra is closed under any finite union.

### 26.4.1 Dynkin System

A Dynkin system on a set  $\Omega$  is a subset  $\mathcal{D} \subset \mathcal{P}(\Omega)$  such that  $\Omega \in \mathcal{D}$ , if  $A, B \in \Omega$  and if  $A \subseteq B$ , then  $A \setminus B \in \mathcal{D}$ , and for all countable collections of elements of  $\mathcal{D}$  such that  $A_1 \subset A_2 \subset \dots, \bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$ . There is an equivalent defintion for Dynkin Systems.  $\Omega \in \mathcal{D}$ ,  $A \in \mathcal{D}$  implies  $A^C \in \mathcal{D}$ , and for all countable disjoint collections of elements in  $\mathcal{D}$ , the union is also contained in  $\mathcal{D}$ . These requirements are weaker than those of a  $\sigma$ -algebra. Any  $\sigma$ -algebra is a Dynkin system, but not every Dynkin system is a  $\sigma$ -algebra.

### Definition 26.4.2: Dynkin System

---

Dynkin System on a set  $\Omega$  is a subset  $\mathcal{D} \subseteq \mathcal{P}(\Omega)$  such that:

1.  $\Omega \in \mathcal{D}$ .
2. For all  $A, B \in \mathcal{D}$  such that  $A \subseteq B$ ,  $B \setminus A \in \mathcal{D}$ .
3. For any sequence  $A_n \in \mathcal{D}$  such that  $A_n \subseteq A_{n+1}$ ,  $\cup_{n=1}^{\infty} A_n \in \mathcal{D}$



**Theorem 26.4.1.** *If  $\Omega$  is a set and  $\mathcal{D} \subseteq \mathcal{P}(\Omega)$  is such that  $\Omega \in \mathcal{D}$ , for all  $A \in \mathcal{D}$ ,  $A^C \in \mathcal{D}$ , and if for all sequences  $A_n \in \mathcal{D}$  such that  $A_n \cap A_m = \emptyset$  for all  $n \neq m$ ,  $\cup_{n=1}^{\infty} A_n \in \mathcal{D}$ , then  $\mathcal{D}$  is a Dynkin System on  $\Omega$ .*

**Theorem 26.4.2.** *If  $\mathcal{D}$  is a Dynkin system on a set  $\Omega$ , and if  $\mathcal{D}$  is closed with respect to intersections, then  $\mathcal{D}$  is a  $\sigma$ -algebra on  $\Omega$ .*

**Theorem 26.4.3.** *If  $\Omega$  is a set, if  $\mathcal{E} \subset \mathcal{P}(\Omega)$  is closed to intersections, and if  $\mathcal{D}$  is the Dynkin System generated by  $\mathcal{E}$ , then  $\mathcal{D}$  is a  $\sigma$ -algebra.*

**Theorem 26.4.4** (Dynkin's Theorem). *If  $\Omega$  is a set,  $\mathcal{C} \subseteq \mathcal{P}(\Omega)$  is intersection-stable, and if  $\mathcal{D}$  is the smallest Dynkin system that contains  $\mathcal{C}$ , then  $\mathcal{D}$  is also intersection-stable.*

*Proof.* For let:

$$\mathcal{D}_1 = \{D \in \mathcal{D} : \forall C \in \mathcal{C}, D \cap C \in \mathcal{D}\} \quad (26.4.1)$$

Then  $\mathcal{D}_1$  is a Dynkin system, and thus  $\mathcal{D}_1 = \mathcal{D}$ . Now define:

$$\mathcal{D}_2 = \{D \in \mathcal{D} : \forall A \in \mathcal{D}, D \cap A \in \mathcal{D}\} \quad (26.4.2)$$

Then  $\mathcal{C} \subseteq \mathcal{D}_2$  and  $\mathcal{D}_2$  is a Dynkin System, and thus  $\mathcal{D}_2 = \mathcal{D}$ . □

**Theorem 26.4.5.** *If  $\Omega$  is a set,  $\mathcal{C} \subseteq \mathcal{P}(\Omega)$  is intersection-stable, and if  $\mathcal{D}$  is the smallest Dynkin system that contains  $\mathcal{C}$ , then  $\mathcal{D}$  is a  $\sigma$ -Algebra on  $\Omega$ .*

Since semi-intervals are closed to intersections, the Borel  $\sigma$ -algebra is equivalently the Dynkin system generated by semi-intervals.

### 26.4.2 Borel $\sigma$ -Algebra

One of the most important types of  $\sigma$ -algebras is the Borel  $\sigma$ -algebra. We first define the Borel  $\sigma$ -algebra on  $\mathbb{R}$ .

**Definition 26.4.3** The Borel  $\sigma$ -algebra on  $\mathbb{R}$ , denoted  $\mathcal{B}$ , is the  $\sigma$ -algebra generated by the set  $\{[a, b] : a, b \in \mathbb{R}\}$ .

That is, the Borel  $\sigma$ -algebra on  $\mathbb{R}$  is the *smallest*  $\sigma$ -algebra that contains all of the semi-intervals. We can equivalently say that  $\mathcal{B}$  is the  $\sigma$ -algebra generated by all *open* intervals. If we know that every open subset of  $\mathbb{R}$  is the countable union of open subsets, than we can equivalently say that  $\mathcal{B}$  is the  $\sigma$ -algebra generated by all *open subsets* of  $\mathbb{R}$ . Writing  $[a, b)$  as the countable intersection of open intervals, or  $(a, b)$  as the countable union of semi-intervals comes from the Archimedean property of the real numbers. Thus, the smallest  $\sigma$ -algebra that contains all semi-intervals is the smallest  $\sigma$ -algebra that contains all open intervals, which is the smallest  $\sigma$ -algebra that contains all open subsets of  $\mathbb{R}$ . Similarly, this will contain all of the *closed* intervals, intervals of the form  $[a, b]$ . We say that a set  $\mathcal{U} \subset \mathbb{R}$  is open if, for all  $x \in \mathcal{U}$ , there is an  $r > 0$  such that  $(x - r, x + r) \subset \mathcal{U}$ . That is, every point in  $\mathcal{U}$  can be surrounded by an interval that is entirely contained in  $\mathcal{U}$ . Thus, any open set can be written as:

$$\mathcal{U} = \bigcup_{x \in \mathcal{U}} (\alpha_x, \beta_x) \quad (26.4.3)$$

This union is not countable, for any open set must contain an interval, an intervals are uncountable large. This is simply because  $(a, b)$  is equivalent to  $(0, 1)$ . By adjusting the size of  $\alpha_x$  and  $\beta_x$  to be rational numbers, we can written  $\mathcal{U}$  as the union of intervals with rational endpoints. But there are only countably many such intervals, and thus  $\mathcal{U}$  is the union of countably many open intervals. Thus, any open set is the union of countably many open intervals. From this, the smallest  $\sigma$ -algebra that contains open intervals will contain all open sets, since  $\sigma$ -algebras are closed under countable unions. Borel sets are elements of the Borel  $\sigma$ -algebra  $\mathcal{B}$ . Since all open sets are Borel sets, and as  $\sigma$ -algebras are closed under complements, all closed sets are also Borel sets. This is because the complement of an open set is a closed set, and vice versa. Thus, equivalently,  $\mathcal{B}$  is the smallest  $\sigma$ -algebra containing all closed sets. A  $G_\delta$  sets is a subset that is the countable intersection of open sets. As open sets are not necessarily closed under countable intersections, not all  $G_\delta$  sets are open. There is another notion,

# CHAPTER 27

---

## Measurable Functions

---

### 27.1 Definitions and Properties

We wish to eventually talk about what it means for a function to be *measurable*. First we do a quick review of functions.

#### Example 27.1.1: I

---

we let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $f(x) = x^2$ , then  $f([1, 2]) = [1, 4]$ , and  $f^{-1}([1, 4]) = [1, 2] \cup [-2, -1]$ . As another example we can consider  $f(x) = \sin(x)$ . Then  $f^{-1}(\{0\}) = \{n\pi : n \in \mathbb{N}\}$  and  $f^{-1}([-1, 1]) = \mathbb{R}$ . ■

**Theorem 27.1.1.** *If  $X$  and  $Y$  are sets,  $f : X \rightarrow Y$ , and if  $A, B \subset X$ , then:*

$$f(A \cup B) = f(A) \cup f(B) \tag{27.1.1}$$

**Theorem 27.1.2.** *If  $X$  and  $Y$  are sets,  $f : X \rightarrow Y$ , and if  $A, B \subset X$ , then:*

$$f(A \cap B) \subseteq f(A) \cap f(B) \tag{27.1.2}$$

For pre-images, we get equality:

**Theorem 27.1.3.** *If  $X$  and  $Y$  are sets,  $f : X \rightarrow Y$ , and if  $A, B \subset X$ , then:*

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B) \tag{27.1.3}$$

**Theorem 27.1.4.** *If  $X$  and  $Y$  are sets,  $f : X \rightarrow Y$ , and if  $A, B \subset X$ , then:*

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B) \tag{27.1.4}$$

**Theorem 27.1.5.** If  $X$  and  $Y$  are sets,  $f : X \rightarrow Y$ , and if  $A \subset X$ , then:

$$f^{-1}(A^C) = f^{-1}(A)^C \quad (27.1.5)$$

**Theorem 27.1.6.** If  $X$  and  $Y$  are sets, if  $f : X \rightarrow Y$  is a function, and if  $A \subseteq X$ , then:

$$A \subseteq f^{-1}(f(A)) \quad (27.1.6)$$

**Theorem 27.1.7.** If  $X$  and  $Y$  are sets, if  $f : X \rightarrow Y$  is an injective function, and if  $A \subseteq X$ , then:

$$A = f^{-1}(f(A)) \quad (27.1.7)$$

Recalling some definitions, a measure on a  $\sigma$ -Algebra  $\mathcal{A}$  is a function  $\mu : \mathcal{A} \rightarrow \mathbb{R}$  such that  $\mu(A) \geq 0$ ,  $\mu(\emptyset) = 0$ , and given a countable collection of disjoint sets  $A_i \in \mathcal{A}$ ,  $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ .

**Theorem 27.1.8.** If  $\Omega$  is a set and  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ , and if  $\mu : \mathcal{A} \rightarrow \mathbb{R}$  is a function such that:

1.  $\mu(A) \geq 0$
2.  $\mu(\emptyset) = 0$
3.  $\mu$  is finitely additive
4.  $\mu$  is continuous from below

Then  $\mu$  is a measure.

*Proof.* All that is necessary to show is countable additivity. Let  $A_n$  be a countable collection of disjoint elements of  $\mathcal{A}$ , and let  $B_n = \cup_{k=1}^n A_k$ . Then:

$$\mu(B_n) = \mu(\cup_{k=1}^n A_k) \quad (27.1.8)$$

$$= \sum_{k=1}^n \mu(A_k) \quad (27.1.9)$$

But by definition, for all  $n \in \mathbb{N}$ ,  $B_n \subseteq B_{n+1}$ , and therefore by continuity from below, we have:

$$\mu(\cup_{n=1}^{\infty} B_k) = \lim_{n \rightarrow \infty} \mu(\cup_{k=1}^n B_k) \quad (27.1.10)$$

And therefore

$$\mu(\cup_{n=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k) \quad (27.1.11)$$

□

The Borel  $\sigma$ -Algebra on  $\mathbb{R}$  is the smallest set that makes open sets, elements of the standard topology on  $\mathbb{R}$ , measurable. In an analogous manner to how continuous functions are defined for topological spaces, measurable functions can also be defined.

### 27.1.1 Measurable Functions

#### Definition 27.1.1: Measurable Functions

---

measure function from a measurable space  $(A, \mathcal{A})$  to a measurable space  $(B, \mathcal{B})$  is a function  $f : A \rightarrow B$  such that, for all  $\mathcal{U} \in \mathcal{A}$ ,  $f^{-1}(\mathcal{U}) \in \mathcal{B}$ . ■

That is, the pre-image of measurable sets is measurable. This is similar to continuous functions where the pre-image of open sets is open. Such functions are also called  $\mathcal{A} - \mathcal{B}$  measurable.

#### Example 27.1.2: I

---

$\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ ,  $\Omega$  and if  $\mathcal{B} = \{\emptyset, \Omega\}$ , then any function  $f : \omega \rightarrow \Omega$  will be  $\mathcal{A} - \mathcal{B}$  measurable. If  $\mathcal{A} = \mathcal{P}(\Omega)$  and if  $\mathcal{B}$  is a  $\sigma$ -Algebra on  $\Omega$ , then again any function  $f : \Omega \rightarrow \Omega$  will be  $\mathcal{A} - \mathcal{B}$  measurable. There are similar notions in topology called the discrete and chaotic topologies which make all functions continuous. ■

**Theorem 27.1.9.** *If  $A$  and  $B$  are sets, if  $\mathcal{B}$  is a  $\sigma$ -Algebra on  $B$ , and if  $f : A \rightarrow B$  is a function, then the set  $\mathcal{A}$  defined by:*

$$\mathcal{A} = \{f^{-1}(\mathcal{U}) : \mathcal{U} \in \mathcal{B}\} \quad (27.1.12)$$

*Is a  $\sigma$ -Algebra on  $A$ .*

*Proof.* It is true that  $\emptyset \in \mathcal{A}$ , since  $\emptyset \in \mathcal{B}$  and  $f^{-1}(\emptyset) = \emptyset$ . Also,  $B \in \mathcal{B}$ , and  $A = f^{-1}(B)$ , and therefore  $A \in \mathcal{A}$ . If  $A \in \mathcal{A}$ , then there is a  $B \in \mathcal{B}$  such that  $A = f^{-1}(B)$ , and thus:

$$A^C = f^{-1}(B)^C = f^{-1}(B^C) \quad (27.1.13)$$

But if  $B \in \mathcal{B}$ , then  $B^C \in \mathcal{B}$ , and thus  $A^C \in \mathcal{A}$ . Finally, for any countable collection of sets  $A_n \in \mathcal{A}$ , there is a countable collection of sets  $B_n$  such that  $A_n = f^{-1}(B_n)$  But then:

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} f^{-1}(B_n) = f^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right) \quad (27.1.14)$$

But  $\mathcal{B}$  is a  $\sigma$ -Algebra, and thus  $\bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$ . Therefore  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ . □

It's worth noting that  $\mathcal{A}$  is the smallest  $\sigma$ -Algebra on  $A$  that will make  $f$   $\mathcal{A} - \mathcal{B}$  measurable. Removing any set from  $\mathcal{A}$  will result in  $f : A \rightarrow B$  being non-measurable with respect to  $\mathcal{A}$  and  $\mathcal{B}$ .

**Theorem 27.1.10.** *If  $A$  and  $B$  are sets,  $f : A \rightarrow B$  a function, and if  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $A$ , then the set  $\mathcal{B}$  defined by:*

$$\mathcal{B} = \{B \subset B : f^{-1}(B) \in \mathcal{A}\} \quad (27.1.15)$$

Is a  $\sigma$ -Algebra on  $B$ .

*Proof.*  $\emptyset$  and  $B$  are elements since  $f^{-1}(\emptyset) = \emptyset \in \mathcal{A}$ , and  $f^{-1}(B) = A \in \mathcal{A}$ .  $\square$

**Theorem 27.1.11.** *If  $f : \Omega \rightarrow \mathbb{R}$ , if  $a \in \mathbb{R}$ , if  $\mathcal{B}$  is the Borel  $\sigma$ -Algebra on  $\mathbb{R}$ , and if  $\mathcal{A}$  is defined by:*

$$\mathcal{A} = \{\omega \in \Omega : f(\omega) < a\} \quad (27.1.16)$$

*then  $f$  is  $\mathcal{A} - \mathcal{B}$  measurable.*

**Theorem 27.1.12.** *If  $\Omega$  is a set, and if  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ , and if  $f : \Omega \rightarrow \mathbb{R}$  is a function such that, for all  $a \in \mathbb{R}$ ,  $\{\omega \in \Omega : f(\omega) < a\} \in \mathcal{A}$ , then  $f$  is  $\mathcal{A} - \mathcal{B}$  measurable, where  $\mathcal{B}$  is the Borel  $\sigma$ -Algebra.*

**Theorem 27.1.13.** *If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is continuous, then it is Borel measurable.*

**Theorem 27.1.14.** *If  $\Omega$  is a set,  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ , and if  $f : \Omega \rightarrow \mathbb{R}$  is  $\mathcal{A} - \mathcal{B}$  measurable, where  $\mathcal{B}$  is the Borel  $\sigma$ -Algebra, and if  $g : \mathbb{R} \rightarrow \mathbb{R}$  is  $\mathcal{B} - \mathcal{B}$  measurable, then  $g \circ f : \Omega \rightarrow \mathbb{R}$  is  $\mathcal{A} - \mathcal{B}$  measurable.*

*Proof.* For if  $B \in \mathcal{B}$ , then  $g^{-1}(B) \in \mathcal{B}$ , for  $g$  is  $\mathcal{B} - \mathcal{B}$  measurable. but then, as  $f$  is  $\mathcal{A} - \mathcal{B}$  measurable,  $f^{-1}(g^{-1}(B)) \in \mathcal{A}$ . Therefore,  $g \circ f$  is  $\mathcal{A} - \mathcal{B}$  measurable.  $\square$

In particular, if we have two measurable functions on  $\mathbb{R}$ , then the composition of these two function is also measurable. This is analogous to the fact that the composition of continuous functions is continuous. The sum, difference, and product of measurable functions is also measurable. We now define the Borel  $\sigma$ -Algebra for  $\mathbb{R}^2$ . This is denoted  $\mathcal{B}_2$ . It is defined similarly to  $\mathcal{B}$ : It is the smallest  $\sigma$ -Algebra that contains all open subsets of  $\mathbb{R}^2$ . We can also limit this to all open rectangles in the plane, or all open discs.

**Theorem 27.1.15.** *If  $\Omega$  is a set,  $\mathcal{A}$  a  $\sigma$ -Algebra on  $\Omega$ , if  $f, g : \Omega \rightarrow \mathbb{R}$  are  $\mathcal{A} - \mathcal{B}$  measurable functions, and if  $\vec{h} : \Omega \rightarrow \mathbb{R}^2$  is defined by  $\vec{h}(\omega) = (f(\omega), g(\omega))$ , then  $\vec{h}$  is  $\mathcal{A} - \mathcal{B}_2$  measurable.*

This theorem goes the other way as well.  $\vec{h}$  is measurable if and only if  $f$  and  $g$  are measurable.

**Theorem 27.1.16.** *If  $\Omega$  is a set,  $\mathcal{A}$  a  $\sigma$ -Algebra on  $\Omega$ , if  $\mathcal{B}$  is the Borel  $\sigma$ -Algebra on  $\mathbb{R}$ , and if  $f, g : \Omega \rightarrow \mathbb{R}$  are  $\mathcal{A} - \mathcal{B}$  measurable functions, then  $f + g$  is  $\mathcal{A} - \mathcal{B}$  measurable.*

*Proof.* For let  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$  be defined by  $\varphi(x, y) = x + y$ . Then  $\varphi$  is continuous, and is therefore  $\mathcal{B}_2 - \mathcal{B}$  measurable. Let  $h : \Omega \rightarrow \mathbb{R}^2$  be defined by  $h(\omega) = (f(\omega), g(\omega))$ . Then  $h$  is  $\mathcal{A} - \mathcal{B}_2$  measurable. But by taking the composition, we have that  $f + g = \varphi \circ h$  is  $\mathcal{A} - \mathcal{B}$  measurable.  $\square$

We can do the same thing with multiplication by defining  $\varphi(x, y) = x \cdot y$ .

**Theorem 27.1.17.** If  $f, g : \Omega \rightarrow \mathbb{R}$  are  $\mathcal{A} - \mathcal{B}$  measurable, and if:

$$A = \{\omega \in \Omega : f(\omega) < g(\omega)\} \quad (27.1.17)$$

$\Pi$  open set (Half plane along diagonal. Draw this). Do the same thing with the line  $L$ . They're measurable, yadda yadda.

### 27.1.2 Sequences of Measurable Functions

If  $f_n : \omega \rightarrow \mathbb{R}$  is a sequence of  $\mathcal{A} - \mathcal{B}$  measurable functions, and if  $f_n \rightarrow f$ , then  $f$  is  $\mathcal{A} - \mathcal{B}$  measurable.

**Theorem 27.1.18.** Let  $F(\omega) = \sup\{f_n(\omega) : n \in \mathbb{N}\}$ . Then  $F$  is measurable.

*Proof.* For let  $a \in \mathbb{R}$ . Then:

$$\{\omega : F(\omega) \leq a\} = \bigcap_{n=1}^{\infty} \{\omega : f_n(\omega) \leq a\} \quad (27.1.18)$$

Then  $F(\omega) \leq a$  if and only if  $f_n(\omega) \leq a$  for all  $n \in \mathbb{N}$ .  $\square$

Similarly,  $F(\omega) = \inf\{f_n(\omega)\}$  is measurable.

**Theorem 27.1.19.** If  $f_n : \Omega \rightarrow \mathbb{R}$  are  $\mathcal{A} - \mathcal{B}$  functions, then  $\overline{\lim}_{n \rightarrow \infty} f_n$  and  $\underline{\lim}_{n \rightarrow \infty} f_n$  are  $\mathcal{A} - \mathcal{B}$  measurable.

## 27.2 Convergence of Measurable Functions

### Definition 27.2.1: Measure Space

Measure Space on a set  $\Omega$ , denoted  $(\Omega, \mathcal{A}, \mu)$ , is a  $\sigma$ -Algebra on  $\Omega$  and a measure  $\mu : \Omega \rightarrow \mathbb{R}$ .  $\blacksquare$

A function  $f : \Omega \rightarrow \mathbb{R}$  is  $\mathcal{A} - \mathcal{B}$  measurable, or simply measurable, if for all  $B \in \mathcal{B}$ , the pre-image is in  $\mathcal{A}$ . That is,  $f^{-1}(B) \in \mathcal{A}$ .

**Theorem 27.2.1.** If  $f, g : \Omega \rightarrow \mathbb{R}$  are measurable functions, and if  $A$  is defined by:

$$A = \{\omega \in \Omega : f(\omega) \neq g(\omega)\} \quad (27.2.1)$$

then  $A$  is  $\mathcal{A}$  measurable.

**Definition 27.2.2** Two functions are equal  $\mu$  almost everywhere if:

$$\mu(\omega \in \Omega : f(\omega) \neq g(\omega)) = 0 \quad (27.2.2)$$

**Definition 27.2.3** A sequence of measurable functions  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  converges to a function  $f : \Omega \rightarrow \mathbb{R}$  if there is a set  $E$  such that  $\mu(E^C) = 0$ , and  $f_n(\omega) \rightarrow f(\omega)$  for all  $\omega \in E$ .

**Theorem 27.2.2.** If  $f_n \rightarrow f$  almost every and  $f_n \rightarrow g$  almost everywhere, then  $f = g$  almost everywhere.

**Theorem 27.2.3.** If  $f_n \rightarrow f$  almost everywhere, and if  $f = g$  almost everywhere, then  $f_n \rightarrow g$  almost everywhere.

**Definition 27.2.4** A function  $f : \Omega \rightarrow \mathbb{R}$  converges to  $f : \Omega \rightarrow \mathbb{R}$  uniformly if for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that, for all  $\omega \in \Omega$ ,  $|f(\omega) - f_n(\omega)| < \varepsilon$ .

**Example 27.2.1** Consider  $f_n(\omega) = \omega^n$  for  $\omega \in [0, a]$ , where  $a < 1$ . Then this converges uniformly to zero since, all  $\omega \in [0, a]$ :

$$|\omega^n - 0| = \omega^n \leq a^n \quad (27.2.3)$$

But since  $0 \leq a < 1$ ,  $a^n$  converges to zero. Thus  $f_n \rightarrow 0$  uniformly.

We can define non-uniform convergence by considering the logical negation of the definition for uniform convergence, but we can simplify this as well.

**Theorem 27.2.4.** A sequence of functions  $f_n$  converges non-uniformly to a function  $f$  if  $f_n \rightarrow f$  point-wise, and there exists a  $\delta > 0$ , a strictly increasing sequence  $n_k$ , and a sequence  $\omega_k$  such that  $|f_{n_k}(\omega_k) - f(\omega_k)| > \delta$

**Example 27.2.2** If we define  $f_n(\omega) = \omega^n$  on the interval  $[0, 1]$ , then the convergence is no longer uniform. Indeed, the limit is discontinuous.

**Definition 27.2.5** A sequence  $f_n$  converges to  $f$  almost uniformly if, for all  $\varepsilon > 0$  there is a set  $E$  such that  $\mu(E^C) < \varepsilon$ , and  $f_n$  converges to  $f$  uniformly on  $E$ .

**Example 27.2.3** If we again let  $f_n(\omega) = \omega^n$  on  $[0, 1]$ , then  $f_n \rightarrow 0$  almost uniformly. For let  $\varepsilon > 0$ . Then  $f_n \rightarrow 0$  uniformly on the set  $[0, 1 - \varepsilon]$ , and the measure of the complement of this is less than  $\varepsilon$ .

There is a difference between convergence almost everywhere and convergence almost uniformly. For convergence almost everywhere, we may remove a set of measure zero and expect that there is point-wise convergence on the remaining set. For almost uniform convergence we may remove a set of arbitrarily small measure, but not necessarily measure zero, and expect uniform convergence on the remaining set.

**Theorem 27.2.5.** If  $f_n \rightarrow f$  almost uniformly, then  $f_n \rightarrow f$  almost everywhere.

*Proof.* For all  $n \in \mathbb{N}$  there is a set  $E_n$  such that  $\mu(E_n^C) < 1/n$ , and  $f_n \rightarrow f$  uniformly on  $E_n$ . But then  $f_n \rightarrow f$  on  $\cup_{n=1}^{\infty} E_n$ . But the complement of this set has measure zero. Therefore, etc.  $\square$

The converse is not true, in general. For let  $f_n(\omega)$  be defined as follows:

$$f_n(\omega) = \begin{cases} 0, & \omega \leq n \\ 1, & \omega > n \end{cases} \quad (27.2.4)$$

The  $f_n \rightarrow 0$  almost everywhere, and indeed  $f_n \rightarrow 0$  point-wise. But the convergence is not uniform, nor is it almost uniform. There is no way to remove a set of finite measure and have uniform convergence on the resulting set. Similar to where continuity from above failed, the fact that  $\mu(\mathbb{R})$  is infinite is why this failed. If we can limit the measure on the set, then convergence almost everywhere implies almost uniform convergence.

### Theorem 27.2.6: Egorov's Theorem

If  $(\Omega, \mathcal{A}, \mu)$  is a measure space and if  $\mu(\Omega) < \infty$ , then convergence  $\mu$ -almost everywhere implies  $\mu$ -almost uniform convergence.

*Proof.* For if  $f_n(\omega) \rightarrow f(\omega)$   $\mu$ -almost everywhere, then there is a set  $E$  such that  $\mu(E^C) = 0$  and  $f_n(\omega) \rightarrow f(\omega)$  for all  $\omega \in E$ . This means that for all  $\delta > 0$  and for all  $\omega \in E$  there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $|f_n(\omega) - f(\omega)| < \delta$ . Let  $A_{nm}$  be defined as:

$$A_{nm} = \bigcup_{n=N}^{\infty} \left\{ \omega \in \Omega : |f_n(\omega) - f(\omega)| \geq \frac{1}{m} \right\} \quad (27.2.5)$$

Define  $B_m$  as:

$$B_m = \bigcap_{N=1}^{\infty} A_{Nm} \quad (27.2.6)$$

But since  $\mu(\Omega) < \infty$ , the measure  $\mu$  is continuous from above. Therefore:

$$\mu(B_m) = \lim_{N \rightarrow \infty} \mu(A_{Nm}) \quad (27.2.7)$$

But  $B_m \subseteq E^c$ , and thus  $\mu(B_m) = 0$ . But then  $\mu(A_{Nm}) \rightarrow 0$ .  $\square$

So we have shown that, even though convergence almost everywhere and almost uniform convergence are different concepts, on sets of finite measure they are equivalent. In probably the total measure of the entire set is 1, and so finite. Thus, in probability spaces, almost everywhere convergence and almost uniform convergence will always be equivalent. Thus it is common to use

the term convergence almost surely, and forgot the differences between the two properties. There is a third type of convergence called convergence in measure.

### Definition 27.2.6: Convergence in Measure

sequence of functions  $f_n$  convergence in measure to  $f$  is, for all  $\delta > 0$ , the following is true:

$$\lim_{n \rightarrow \infty} \mu\left(\{\omega : |f_n(\omega) - f(\omega)| < \delta\}\right) = 0 \quad (27.2.8)$$

■

**Example 27.2.4** Let  $\Omega = [0, 1]$ , and let  $\mathcal{B}$  be the Borel  $\sigma$ -Algebra on  $[0, 1]$ . Finally, let  $\mu$  be the standard Lebesgue-Measure. Define the following:

$$f_1 = \begin{cases} 1, & x < \frac{1}{2} \\ 0, & x \geq \frac{1}{2} \end{cases} \quad (27.2.9)$$

Define  $f_2 = 1 - f_1$ . The split the interval into fourths and define  $f_3$  as 1 in  $[0, 1/4)$  and zero otherwise, and continue the pattern for  $f_4, f_5, f_6$ , and  $f_7$ . This sequence of functions converges nowhere since there will be 1's and 0's oscillating back and forth, and thus there is no limit. However,  $f_n$  converges in measure to 0.

**Theorem 27.2.7.** *If  $f_n \rightarrow f$  in measure  $\mu$ , and if  $g = f$  almost everywhere, then  $f_n \rightarrow g$  in measure  $\mu$ .*

*Proof.* For all  $\delta > 0$ ,  $\mu(\{\omega : |f_n(\omega) - f(\omega)| > \delta\})$  tends to zero as  $n \rightarrow \infty$ . But:

$$\begin{aligned} \{\omega : |f_n(\omega) - f(\omega)| > \delta\} &= \left( \{\omega : |f_n(\omega) - f(\omega)| > \delta\} \cap \{\omega : f(\omega) = g(\omega)\} \right) \\ &\quad \cup \left( \{\omega : |f_n(\omega) - f(\omega)| > \delta\} \cap \{\omega : f(\omega) \neq g(\omega)\} \right) \end{aligned} \quad (27.2.10)$$

□

**Theorem 27.2.8.** *If  $f_n \rightarrow f$  in measure  $\mu$  and if  $f_n \rightarrow g$  in measure  $\mu$ , and  $f = g$   $\mu$  almost everywhere.*

*Proof.* For let  $A = \{\omega : f(\omega) \neq g(\omega)\}$ . Then  $A = \{\omega : |f(\omega) - g(\omega)| > 0\}$ . Thus we may write:

$$A = \bigcup_{n=1}^{\infty} \left\{ \omega : |f(\omega) - g(\omega)| > \frac{1}{n} \right\} \quad (27.2.11)$$

We now show that  $\{\omega : |f(\omega) - g(\omega)| > \frac{1}{n}\}$  has measure zero for all  $n \in \mathbb{N}$ . By subadditivity, this will imply  $A$  has measure zero. From the triangle inequality:

$$|f(\omega) - g(\omega)| \leq |f(\omega) - f_n(\omega)| + |g(\omega) - f_n(\omega)| \quad (27.2.12)$$

If  $|f(\omega) - g(\omega)| \geq 1/m$ , then at least one of the two numbers here must be greater than  $1/2m$ . Thus, either  $|f(\omega) - f_n(\omega)| \geq \frac{1}{2m}$  or  $|g(\omega) - f_n(\omega)| \geq \frac{1}{2m}$ , or both. Therefore:

$$\{\omega : |f(\omega) - g(\omega)| > \frac{1}{n}\} \subseteq \{\omega : |f(\omega) - f_n(\omega)| > \frac{1}{2n}\} \cup \{\omega : |g(\omega) - f_n(\omega)| > \frac{1}{2n}\} \quad (27.2.13)$$

But the two sets on the left have measures that tend to zero as  $n \rightarrow \infty$ , and thus the set on the left has measure zero. Thus, by subadditivity their union has measure zero, and therefore  $\mu(A) = 0$ . Thus,  $f = g$   $\mu$  almost everywhere.  $\square$

**Theorem 27.2.9.** *If  $f_n \rightarrow f$  in measure  $\mu$  and  $g_n \rightarrow g$  in measure  $\mu$ , then  $f_n + g_n \rightarrow f + g$  in measure  $\mu$ .*

**Theorem 27.2.10.** *If  $f_n \rightarrow f$  in measure  $\mu$ , then there exists a subsequence of  $f_n$  that converges to  $f$  almost uniformly.*

*Proof.* For all  $\delta > 0$  the limit of  $\mu(\{|f_n - f| \geq \delta\})$  tends to zero as  $n \rightarrow \infty$ . Thus, there is an index  $n_1$  such that  $\mu(\{|f_{n_1} - f| \geq 1\}) < 1$ . Choosing  $\delta = 1/2$ , we find an index  $n_2$  such that  $\mu(\{|f_{n_2} - f| \geq 1/2\}) < 1/2$ . Carrying on, we obtain a sequence  $n_k$  such that, for all  $k \in \mathbb{N}$ ,  $\mu(\{|f_{n_k} - f| \geq 1/k\}) < 1/2^k$ . Let  $E_k = \{|f_{n_k} - f| \geq 1/k\}^C$ .  $\mu(E_k^C) < 1/2^k$ , and thus for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that, for all  $n > N$ ,  $\mu(E_n^C) < \varepsilon$ .  $\square$

The  $\sigma$ -Algebra generated by a set  $\mathcal{E}$  is the intersection of all possible  $\sigma$ -Algebra's that contain all elements of  $\mathcal{E}$ . Given a function  $f : \Omega \rightarrow \mathbb{R}$  and a  $\sigma$ -Algebra  $\mathcal{A}$  on  $\Omega$ , it is often a good strategy to look at the set:

$$\mathcal{B}_{f,\mathcal{A}} = \{B \subseteq \mathbb{R} : f^{-1}(B) \in \mathcal{A}\} \quad (27.2.14)$$

And then show that all intervals of the form  $(a, b)$  are contained within  $\mathcal{B}_{f,\mathcal{A}}$ , thus implying that  $f$  is  $\mathcal{A}-\mathcal{B}$  measurable, where  $\mathcal{B}$  is the Borel  $\sigma$ -Algebra on  $\mathbb{R}$ . Recapping, we have now discussed three different types of convergence: Almost uniform convergence, convergence almost everywhere, and convergence in measure.

**Theorem 27.2.11.** *If  $f_n \rightarrow f$  almost uniformly, then  $f_n \rightarrow f$  in measure.*



# CHAPTER 28

---

## Measures

---

### 28.1 Measures

#### 28.1.1 A Review Infinite Series

Given a sequence of real numbers,  $a : \mathbb{N} \rightarrow \mathbb{R}$ , the sum of this sequence is defined as the limit of finite partial sums. That is:

$$\sum_{n=1}^{\infty} a_n = \lim_{N \rightarrow \infty} \sum_{n=1}^N a_n \quad (28.1.1)$$

In general, this limit may not in general exists. If it does, we say the series converges. If the limit does not exists, we do not define the sum and instead just have a meaningless combination of symbols. If the sequence is positive, then the sequence of partial sums will be increasing. If this sequence is bounded, then the limit exists. This comes from the fact that bounded monotonic sequences converge, a result that stems from the least upper bound property of  $\mathbb{R}$ . Moreover, if  $a : \mathbb{N} \rightarrow \mathbb{R}$  is a sequence of positive real numbers, and if  $f : \mathbb{N} \rightarrow \mathbb{N}$  is any bijective function, then the following is true:

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{f(n)} \quad (28.1.2)$$

We can also split the sequence into a grid, and take the double sum, obtaining the same result. If  $A_1, A_2, \dots$  are disjoint sets whose union is  $\mathbb{N}$ , and if  $b_{nm}$  is the  $n^{th}$  element of  $A_m$ , then:

$$\sum_{i=1}^{\infty} a_i = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} b_{nm} \quad (28.1.3)$$

We should be precise in what we mean. The double sum is the *limit of a limit*.

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm} = \lim_{N \rightarrow \infty} \sum_{n=1}^N \left( \lim_{M \rightarrow \infty} \sum_{m=1}^M a_{nm} \right) \quad (28.1.4)$$

We use infinite series to define *measures* on  $\sigma$ -algebra.

### 28.1.2 Measure Functions

A set function on a collection of sets  $\mathcal{E}$  is a function  $\mu : \mathcal{E} \rightarrow \mathbb{R}$ . For example, if we consider the set of all semi-intervals of the form  $[a, b)$ , where  $a, b \in \mathbb{R}$  and  $a \leq b$ , then we can define  $\mu([a, b)) = b - a$ . This gives rise to the notion of a measure function. A measure function on a collection of set  $\mathcal{E}$  is a function  $\mu : \mathcal{E} \rightarrow \mathbb{R}$  such that:

1. If  $\emptyset \in \mathcal{E}$ , then  $\mu(\emptyset) = 0$
2. For all  $A \in \mathcal{E}$ ,  $\mu(A) \geq 0$
3. For any countable collection of pair-wise disjoint sets whose union also lies in  $\mathcal{E}$ ,  $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$

It helps if we don't have to consider the case where  $\mu(\emptyset)$  is undefined, or where we don't have closure under countable unions, so we discuss measure functions on  $\sigma$ -algebras.

**Definition 28.1.1** A measure on a  $\sigma$ -algebra  $\mathcal{A}$  is a function  $\mu : \mathcal{A} \rightarrow \mathbb{R}$  such that:

1.  $\mu(\emptyset) = 0$
2. For all  $A \in \mathcal{A}$ ,  $\mu(A) \geq 0$
3. For any countable collection of pairwise disjoint elements of  $\mathcal{A}$ ,  $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$

**Example 28.1.1** Let  $\Omega$  be a set, and let  $\mathcal{A} = \mathcal{P}(\Omega)$ . Then  $\mathcal{A}$  is a  $\sigma$ -algebra on  $\Omega$ . If  $\omega_1, \dots, \omega_n \in \Omega$  and if  $p_1, \dots, p_n \in \mathbb{R}^+$ , then:

$$\mu(A) = \sum_{k=1}^n p_k \chi_A(\omega_k) \quad (28.1.5)$$

Where  $\xi_A$  is the indicator function:

$$\chi_A(\omega) = \begin{cases} 0, & \omega \notin A \\ 1, & \omega \in A \end{cases} \quad (28.1.6)$$

This is an example of a *point measure* on  $\mathcal{A}$ . It defines a measure function.

A  $\sigma$ -Algebra on a set  $\Omega$  is a subset  $\mathcal{A}$  of  $\mathcal{P}(\Omega)$  such that  $\Omega \in \mathcal{A}$  and for any countable collection of elements  $A_i \in \mathcal{A}$ , the union  $\bigcup_{i=1}^{\infty} A_i$  is also contained in  $\mathcal{A}$ .  $\mathcal{A}$  does not have to consist of countably many elements. The sequence of subset  $A_i$  does not have to exhaust the entirety of  $\mathcal{A}$ , much the way that any sequence of real numbers will not exhaust the entire of  $\mathbb{R}$ . Going in the other direction,  $\sigma$ -Algebras can be finite. If  $\Omega$  is a set, and if  $A \subset \Omega$  is non-empty, then  $\mathcal{A} = \{\emptyset, A, A^C, \Omega\}$  defines a  $\sigma$ -algebra on  $\Omega$ . A measure on a  $\sigma$ -Algebra  $\mathcal{A}$  is a function  $\mu : \mathcal{A} \rightarrow \mathbb{R}$  such that, for all  $A \in \mathcal{A}$ ,  $\mu(A) \geq 0$ ,  $\mu(\emptyset) = 0$ , and given a mutually disjoint countable collection of elements of  $\mathcal{A}$ , the following holds:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{n=1}^{\infty} \mu(A_i) \quad (28.1.7)$$

**Example 28.1.2** A pure point measure is a measure that assigns to a collection of elements  $\omega_j \in \Omega$  a positive real number  $p_j$ , and then the measure of any set  $A$  is:

$$\mu(A) = \sum_{j: \omega_j \in A} p_j \quad (28.1.8)$$

### 28.1.3 Properties of Measure

#### Monotonicity

If  $A$  and  $B$  are elements of a  $\sigma$ -Algebra  $\mathcal{A}$ , if  $\mu$  is a measure on  $\mathcal{A}$ , and if  $A \subseteq B$ , then  $\mu(A) \leq \mu(B)$ . This is the monotonic property of measures.

**Theorem 28.1.1.** *If  $\Omega$  is a set,  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ , if  $\mu$  is a measure on  $\mathcal{A}$ , and if  $A, B$  are elements of  $\mathcal{A}$  such that  $A \subseteq B$ , then  $\mu(A) \leq \mu(B)$ .*

*Proof.* For as  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ , and as  $A, B \in \mathcal{A}$ ,  $B \setminus A \in \mathcal{A}$ . But, as  $A \subseteq B$ ,  $B = (B \setminus A) \cup A$ . But then, as measures are additive and positive:

$$\mu(B) = \mu((B \setminus A) \cup A) \quad (28.1.9)$$

$$= \mu(B \setminus A) + \mu(A) \quad (28.1.10)$$

$$\geq \mu(A) \quad (28.1.11)$$

□

**Theorem 28.1.2.** *If  $\Omega$  is a set,  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ , if  $\mu$  is a measure on  $\mathcal{A}$ , and if  $A, B$  are elements of  $\mathcal{A}$  such that  $A \subseteq B$  and  $\mu(A), \mu(B) < \infty$ , then  $\mu(B \setminus A) = \mu(B) - \mu(A)$ .*

#### Continuity Theorems

**Theorem 28.1.3** (Continuity from Below). *If  $\Omega$  is a set,  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ , if  $\mu$  is a measure on  $\mathcal{A}$ , and if  $A_i$  is a sequence of elements in  $\mathcal{A}$  such that,*

for all  $i \in \mathbb{N}$ ,  $A_i \subseteq A_{i+1}$ , then:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{N \rightarrow \infty} \mu(A_N) \quad (28.1.12)$$

*Proof.* For let  $A = \cup_{n=1}^{\infty} A_n$  and let  $B_n = A_{n+1} \setminus A_n$ . Then, as  $A_n \subseteq A_{n+1}$ , for all  $i, j \in \mathbb{N}$ ,  $B_i \cap B_j = \emptyset$ . But  $A = A_1 \cup (\cup_{n=1}^{\infty} B_n)$  and this is the countable union of mutually disjoint sets, and therefore, using the telescoping series:

$$\mu(A) = \mu(A_1) + \sum_{n=1}^{\infty} \mu(B_n) \quad (28.1.13)$$

$$= \mu(A_1) + \sum_{n=1}^{\infty} (\mu(A_{n+1}) - \mu(A_n)) \quad (28.1.14)$$

$$= \mu(A_1) + \lim_{N \rightarrow \infty} (\mu(A_N) - \mu(A_1)) \quad (28.1.15)$$

$$= \lim_{N \rightarrow \infty} \mu(A_N) \quad (28.1.16)$$

□

**Theorem 28.1.4** (Continuity from Above). *If  $\Omega$  is a set,  $\mathcal{A}$  is a  $\sigma$ -Algebra on  $\Omega$ , if  $\mu$  is a measure on  $\mathcal{A}$ , and if  $A_i$  is a sequence of elements in  $\mathcal{A}$  such that, for all  $i \in \mathbb{N}$ ,  $A_{i+1} \subseteq A_i$  and there exists an  $n \in \mathbb{N}$  such that  $\mu(A_n)$  is finite, then:*

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{N \rightarrow \infty} \mu(A_N) \quad (28.1.17)$$

*Proof.* For let  $A = \cap_{n=1}^{\infty} A_n$  and let  $B_n = A_n \setminus A_{n+1}$ . Then:

$$A_1 = A \cup \left( \bigcup_{n=1}^{\infty} B_n \right) \quad (28.1.18)$$

And this is the union of countably many disjoint sets. Therefore:

$$\mu(A_1) = \mu(A) + \sum_{n=1}^{\infty} \mu(B_n) \quad (28.1.19)$$

$$= \mu(A) + \sum_{n=1}^{\infty} (\mu(A_n) - \mu(A_{n+1})) \quad (28.1.20)$$

$$= \mu(A) + \mu(A_1) - \lim_{N \rightarrow \infty} \mu(A_N) \quad (28.1.21)$$

Subtracting by  $\mu(A_1)$  obtains the result. □

If  $\mu(A_i) = \infty$  for all  $i \in \mathbb{N}$ , then the above theorem may not be true. For consider the collection of sets  $A_n = [n, \infty)$ . The measure of each  $A_n$  is infinite, but the intersection of the entire collection is empty. Thus the measure of the intersection is zero.

**Theorem 28.1.5** (Countable Sub-Additivity). *If  $\Omega$  is a set,  $\mathcal{A}$  a  $\sigma$ -Algebra on  $\mathcal{A}$ , and if  $A_i$  is a countable collection of elements of  $\mathcal{A}$ , then:*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n) \quad (28.1.22)$$

*Proof.* For if  $A_1, A_2 \in \mathcal{A}$ , then:

$$\mu(A_1 \cup A_2) = \mu(A_1 \setminus A_2) + \mu(A_2 \setminus A_1) + \mu(A_1 \cap A_2) \quad (28.1.23)$$

But also:

$$\mu(A_1) = \mu(A_1 \setminus A_2) + \mu(A_1 \cap A_2) \quad (28.1.24)$$

$$\mu(A_2) = \mu(A_2 \setminus A_1) + \mu(A_1 \cap A_2) \quad (28.1.25)$$

And therefore:

$$\mu(A_1) + \mu(A_2) = \mu(A_1 \cup A_2) + \mu(A_1 \cap A_2) \quad (28.1.26)$$

We now prove by induction. Suppose this is true of a collection of  $N$  elements. Given a collection  $A_i$  of  $N + 1$  elements, let  $B = \bigcup_{n=1}^{N+1} A_n$ . But then:

$$\mu(A_{N+1} \cup B) \leq \mu(A_{N+1}) + \mu(B) \quad (28.1.27)$$

$$\leq \mu(A_{N+1}) + \sum_{n=1}^N A_n \quad (28.1.28)$$

$$= \sum_{n=1}^{N+1} \mu(A_n) \quad (28.1.29)$$

□

## 28.2 Lebesgue-Stieltjes Measures

A Lebesgue-Stieltjes measure is any measure on the Borel  $\sigma$ -Algebra  $\mathcal{B}$  such that, for any finite semi-interval  $[a, b]$ ,  $\mu([a, b]) < \infty$ .  $\mu(\mathbb{R})$  may be infinite. Recall that the Borel  $\sigma$ -Algebra is the smallest  $\sigma$ -Algebra on  $\mathbb{R}$  that contains all semi-intervals  $[a, b]$ . A pure point measure on  $\mathbb{R}$ , indexing over the rational numbers, would be such a measure. If we have a Lebesgue-Stieltjes measure, we wish to find a function  $F_\mu : \mathbb{R} \rightarrow \mathbb{R}$  such that, for all semi-intervals  $[a, b]$ ,

$\mu([a, b)) = F_\mu(b) - F_\mu(a)$ . In probability, this is called the cumulative probability function. For now we wish to show that there is indeed such a function that does this. Consider the case when  $\mu(\mathbb{R}) < \infty$ . Let  $F_\mu(x) = \mu(-\infty, x)$ . Then:

$$\mu([a, b)) = \mu((-\infty, b) \setminus (-\infty, a)) \quad (28.2.1)$$

$$= \mu((-\infty, b)) - \mu((-\infty, a)) \quad (28.2.2)$$

$$= F_\mu(b) - F_\mu(a) \quad (28.2.3)$$

In the more general case when that measure of the entire real line is infinite we still want to find a function such that:

$$\mu([0, x)) = F_\mu(x) - F_\mu(0) \quad x > 0 \quad (28.2.4)$$

$$\mu([x, 0)) = F_\mu(0) - F_\mu(x) \quad x < 0 \quad (28.2.5)$$

We can define the following:

$$F_\mu(x) = \begin{cases} \mu([0, x)) + C, & x > 0 \\ -\mu([x, 0)) + C, & x < 0 \\ C, & x = 0 \end{cases} \quad (28.2.6)$$

Then  $F_\mu$  is a function that satisfies our criterion. Indeed,  $F_\mu$  is defined uniquely up to an additive constant. Any such function is non-decreasing since, for any  $x < y$ ,  $F_\mu(y) - F_\mu(x) = \mu([x, y)) \geq 0$ . In addition,  $F_\mu$  is left-continuous. That is, for all  $a \in \mathbb{R}$ :

$$\lim_{x \rightarrow a^-} F_\mu(x) = F_\mu(a) \quad (28.2.7)$$

**Theorem 28.2.1.** *If  $\mu$  is a Lebesgue-Stieltjes measure on the Borel  $\sigma$ -Algebra of  $\mathbb{R}$ , and if  $F_\mu$  is the function thing, then  $F_\mu$  is left-continuous.*

*Proof.* For let  $a \in \mathbb{R}$  and let  $x : \mathbb{N} \rightarrow \mathbb{R}$  be a monotonic increasing sequence such that  $x_n \rightarrow a$ . But then, for all  $n \in \mathbb{N}$ ,  $[x_{n+1}, a) \subset [x_n, a)$ . But then:

$$\mu\left(\bigcap_{n=1}^{\infty} [x_n, a)\right) = \lim_{N \rightarrow \infty} \mu([x_N, a)) \quad (28.2.8)$$

But  $\bigcap_{n=1}^{\infty} [x_n, a) = \emptyset$  as  $x_n \rightarrow a$ . Therefore:

$$\lim_{N \rightarrow \infty} \mu([x_N, a)) = 0 \quad (28.2.9)$$

But from the definition of  $F_\mu$ ,

$$\mu([x_n, a)) = F_\mu(a) - F_\mu(x_n) \quad (28.2.10)$$

Thus,  $F_\mu(x_n) \rightarrow F_\mu(a)$ . □

Such a function may not be right-continuous. The requirement that the sequence  $x$  be increasing was necessary for the proof. However, the right-hand limit does exist.

**Theorem 28.2.2.** *If blah blah, right hand limit exists.*

*Proof.* For:

$$\{a\} = \bigcap_{n=1}^{\infty} [a, a + \frac{1}{n}) \quad (28.2.11)$$

And thus, as  $\mu$  is a Lebesgue-Stieltjes measure, and thus  $\mu([a, b)) < \infty$  for all finite semi-intervals, we may apply continuity from above and obtain:

$$\mu(\{a\}) = \lim_{N \rightarrow \infty} \mu([a, a + \frac{1}{n}) \quad (28.2.12)$$

$$= \lim_{x \rightarrow a^+} F_\mu(x) - F_\mu(a) \quad (28.2.13)$$

□

If  $\mu$  has no points of positive measure, then  $F_\mu$  will be continuous.

### Theorem 28.2.3: Carathéodory Extension Theorem

If  $F : \mathbb{R} \rightarrow \mathbb{R}$  is a non-decreasing left-continuous function then there exists a unique Lebesgue-Stieltjes measure  $\mu$  such that, for all  $a, b \in \mathbb{R}$ ,  $a < b$ :

$$\mu([a, b)) = F(b) - F(a) \quad (28.2.14)$$

In particular, using  $F(x) = x$ , we see that there is a unique measure on the Borel  $\sigma$ -Algebra such that  $\mu([a, b)) = b - a$ . This measure is called the Lebesgue measure on  $\mathbb{R}$ . We define  $\mu^*$  on a set  $A \subseteq \mathbb{R}$  to be:

$$\mu^*(A) = \inf \left\{ \sum_{i=1}^{\infty} (b_i - a_i) : A \subseteq \bigcup_{i=1}^{\infty} [a_i, b_i] \right\} \quad (28.2.15)$$

If  $A$  is countable, then  $\mu^*(A)$  is zero. For let  $a : \mathbb{N} \rightarrow A$  be bijection, and let  $\varepsilon > 0$ . Then:

$$\mu^*(A) \leq \sum_{n=1}^{\infty} \left( (a_n + \frac{\varepsilon}{2^{n+1}}) - (a_n - \frac{\varepsilon}{2^{n+1}}) \right) \quad (28.2.16)$$

$$= \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} \quad (28.2.17)$$

$$= \varepsilon \quad (28.2.18)$$

Taking the infinimum, we see that  $\mu^*(A) = 0$ . This function is defined on all of  $\mathcal{P}(\mathbb{R})$ , however it is not a measure. The restriction of  $\mu^*$  to the Borel  $\sigma$ -Algebra is a measure.

# CHAPTER 29

---

## Product Measures

---

### 29.1 Product Measures

Let  $(\Omega_1, \mathcal{A}_1, \mu_1)$  and  $(\Omega_2, \mathcal{A}_2, \mu_2)$  be measure spaces. We wish to define a *natural* measure space on the Cartesian product  $\Omega_1 \times \Omega_2$ . Let  $\mathcal{P}$  be defined by:

$$\mathcal{P} = \{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\} \quad (29.1.1)$$

Then  $\mathcal{P}$  is a semi-ring, but is not a  $\sigma$ -Algebra on  $\Omega_1 \times \Omega_2$ . This is because the union of two rectangles may not be a rectangle. Similarly, the difference of two rectangles may not be a rectangle. However, the intersection of two rectangles is a rectangle, and hence this is a semi-ring. We defined the product  $\sigma$ -Algebra to be the  $\sigma$ -Algebra that is generated by  $\mathcal{P}$ .

#### Theorem 29.1.1: Carathéodory Extension Theorem

---

If  $(\Omega_1, \mathcal{A}, \mu_1)$  and  $(\Omega_2, \mathcal{A}_2, \mu_2)$  are measure spaces, if  $\mathcal{A}$  is the product  $\sigma$ -Algebra on  $\Omega_1 \times \Omega_2$ , then there is a unique measure  $\mu$  on  $\mathcal{A}$  such that, for all  $A_1 \in \mathcal{A}_1$  and all  $A_2 \in \mathcal{A}_2$ :

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \quad (29.1.2)$$

■

#### Theorem 29.1.2: Funini's Theorem

---

If  $f : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$  is a non-negative function that is  $\mathcal{A} - \mathcal{B}$  measurable, where  $\mathcal{A}$  is the product  $\sigma$ -Algebra, then:

$$\int_{\Omega_1 \times \Omega_2} f \, d\mu = \int_{\Omega_1} \left( \int_{\Omega_2} f \, d\mu_2 \right) \, d\mu_1 = \int_{\Omega_2} \left( \int_{\Omega_1} f \, d\mu_1 \right) \, d\mu_2 \quad (29.1.3)$$

■

As a summary, when is the following true?

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n \, d\mu \stackrel{?}{=} \int_{\Omega} \lim_{n \rightarrow \infty} f_n \, d\mu \quad (29.1.4)$$

There are two special cases when equality can be guaranteed. The first is monotone convergence. If  $f_n \rightarrow f$ , where  $f_{n+1}(x) \leq f_n(x)$  for all  $n$ , and if  $f_n(x) \geq F$ , where  $F$  is a summable minorant, or if  $f_n \rightarrow f$ ,  $f_{n+1}(x) \leq f_n(x)$ , and if  $f_n(x) \leq F$ , where  $F$  is a summable majorant, then equality holds. The next case is by dominated convergence. If the limit of  $f_n$  exists almost everywhere, and if  $|f_n| \leq F$ , where  $F$  is summable, then by Fatou's Lemma:

$$\varliminf_{n \rightarrow \infty} \int_{\Omega} f_n \, d\mu \geq \int_{\Omega} \varliminf_{n \rightarrow \infty} f_n \, d\mu \quad (29.1.5)$$

And also:

$$\overline{\lim}_{n \rightarrow \infty} \int_{\Omega} f_n \, d\mu \leq \int_{\Omega} \overline{\lim}_{n \rightarrow \infty} f_n \, d\mu \quad (29.1.6)$$

Part XV

# Probability Theory



## 29.2 Product Measures

Let  $(\Omega_1, \mathcal{A}_1, \mu_1)$  and  $(\Omega_2, \mathcal{A}_2, \mu_2)$  be measure spaces. We wish to define a *natural* measure space on the Cartesian product  $\Omega_1 \times \Omega_2$ . Let  $\mathcal{P}$  be defined by:

$$\mathcal{P} = \{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\} \quad (29.2.1)$$

Then  $\mathcal{P}$  is a semi-ring, but is not a  $\sigma$ -Algebra on  $\Omega_1 \times \Omega_2$ . This is because the union of two rectangles may not be a rectangle. Similarly, the difference of two rectangles may not be a rectangle. However, the intersection of two rectangles is a rectangle, and hence this is a semi-ring. We defined the product  $\sigma$ -Algebra to be the  $\sigma$ -Algebra that is generated by  $\mathcal{P}$ .

### Theorem 29.2.1: Carathéodory Extension Theorem

If  $(\Omega_1, \mathcal{A}, \mu_1)$  and  $(\Omega_2, \mathcal{A}_2, \mu_2)$  are measure spaces, if  $\mathcal{A}$  is the product  $\sigma$ -Algebra on  $\Omega_1 \times \Omega_2$ , then there is a unique measure  $\mu$  on  $\mathcal{A}$  such that, for all  $A_1 \in \mathcal{A}_1$  and all  $A_2 \in \mathcal{A}_2$ :

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \quad (29.2.2)$$



### Theorem 29.2.2: Funini's Theorem

If  $f : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$  is a non-negative function that is  $\mathcal{A} - \mathcal{B}$  measurable, where  $\mathcal{A}$  is the product  $\sigma$ -Algebra, then:

$$\int_{\Omega_1 \times \Omega_2} f \, d\mu = \int_{\Omega_1} \left( \int_{\Omega_2} f \, d\mu_2 \right) \, d\mu_1 = \int_{\Omega_2} \left( \int_{\Omega_1} f \, d\mu_1 \right) \, d\mu_2 \quad (29.2.3)$$



As a summary, when is the following true?

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n \, d\mu \stackrel{?}{=} \int_{\Omega} \lim_{n \rightarrow \infty} f_n \, d\mu \quad (29.2.4)$$

There are two special cases when equality can be guaranteed. The first is monotone convergence. If  $f_n \rightarrow f$ , where  $f_{n+1}(x) \leq f_n(x)$  for all  $n$ , and if  $f_n(x) \geq F$ , where  $F$  is a summable minorant, or if  $f_n \rightarrow f$ ,  $f_{n+1}(x) \leq f_n(x)$ , and if  $f_n(x) \leq F$ , where  $F$  is a summable majorant, then equality holds. The next case is by dominated convergence. If the limit of  $f_n$  exists almost everywhere, and if  $|f_n| \leq F$ , where  $F$  is summable, then by Fatou's Lemma:

$$\varliminf_{n \rightarrow \infty} \int_{\Omega} f_n \, d\mu \geq \int_{\Omega} \varliminf_{n \rightarrow \infty} f_n \, d\mu \quad (29.2.5)$$

And also:

$$\overline{\lim}_{n \rightarrow \infty} \int_{\Omega} f_n \, d\mu \leq \int_{\Omega} \overline{\lim}_{n \rightarrow \infty} f_n \, d\mu \quad (29.2.6)$$

## 29.3 Probability Spaces

To add later:

1. Probability space
2. Independent  $\sigma$ -Algebras.
3. Independent sets.
4. Infinite sequence of  $\sigma$ -Algebras.
5. Tail  $\sigma$ -Algebra.
6. Terminal  $\sigma$ -Algebra.
7. Kologorov zero-one law.
8. If  $\mathcal{A}_j$  independent,  $F$  is self-independent.
9.  $E_1, E_2 \in F$ ,  $\mu(E_1 \cap E_2) = \mu(E_1)\mu(E_2)$ , then  $\mu(E_1) = 0$  or  $\mu(E_1) = 1$ .
10. Uniting  $\sigma$ -Algebras lemma.
11.  $A_1, \dots, A_n$  independent, then  $A_k, F$  independent, where  $F$  is the tail  $\sigma$ -Algebra.

## 29.4 Random Variables

Let  $(\Omega, \mathcal{A}, \mu)$  be a probability space. A probability space is a measure space such that  $\mu(\Omega) = 1$ . Let  $f : \Omega \rightarrow \mathbb{R}$  be  $\mathcal{A} - \mathcal{B}$  measurable, where  $\mathcal{B}$  is the Borel  $\sigma$ -Algebra. Such functions are called random-variables on  $\Omega$ . While there's nothing random about this, we use such functions to model problems in probability theory. The probability of an event  $A \in \mathcal{A}$  is simply  $\mu(A)$ . The associated  $\sigma$ -Algebra is defined as:

$$\mathcal{A}_f = \{f^{-1}(B) : B \in \mathcal{B}\} \quad (29.4.1)$$

This is also called the  $\sigma$ -Algebra of events bearing on  $f$ . This is a  $\sigma$ -Algebra on  $\Omega$ .

### **Definition 29.4.1: Distribution of a Random Variable**

he distribution of a random variable  $f : \Omega \rightarrow \mathbb{R}$  on a probability space  $(\Omega, \mathcal{A}, \mu)$  is the image measure  $\mu_f$  of  $f$ . ■

The image measure is the measure:

$$\mu_f(B) = \mu(f^{-1}(B)) = \mu(\{\omega \in \Omega : f(\omega) \in B\}) \quad (29.4.2)$$

This is a Lebesgue-Stieljes Measure on the Borel  $\sigma$ -Algebra on  $\mathbb{R}$ .

$$\mu_f(\mathbb{R}) = \mu(f^{-1}(\mathbb{R})) = \mu(\Omega) = 1 \quad (29.4.3)$$

### Definition 29.4.2: Cumulative Distribution Function

he Cumulative Distribution Function of a random variable  $f : \Omega \rightarrow \mathbb{R}$  on a probability space  $(\Omega, \mathcal{A}, \mu)$  is the function  $F : \mathbb{R} \rightarrow \mathbb{R}$  defined by:

$$F(x) = \mu_f((-\infty, a]) \quad (29.4.4)$$

Where  $\mu_f$  is the distribution of  $f$ . ■

Some facts about the cumulative distribution function: It is non-decreasing on  $\mathbb{R}$ , left continuous, and  $F(-\infty) - F(\infty) = 1$ . By the Caratheodory extension theorem, and function  $F$  that satisfies these three conditions is the cumulative distribution function of some Lebesgue-Stieljes probability measure on  $\mathbb{R}$ . From this we also have that every Lebesgue-Stieljes probability measure on  $\mathbb{R}$  is a distribution for a random variable.

**Example 29.4.1** Let  $\Omega = \mathbb{R}$ , let  $\mathcal{A} = \mathcal{B}$ , where  $\mathcal{B}$  is the Borel  $\sigma$ -Algebra, and let  $\mu$  be a Lebesgue-Stieljes probability measure on  $\mathbb{R}$ . Define the random variable  $f : \Omega \rightarrow \mathbb{R}$  by  $f(\omega) = \omega$ . The inverse of any Borel set is itself, and thus we see that the distribution and the random variable coincide.

**Example 29.4.2** Let  $\Omega = [0, 1]$ ,  $\mathcal{B}$  be the Borel  $\sigma$ -Algebra, and define  $f_1, f_2 : \Omega \rightarrow \mathbb{R}$  by:

$$f_1(\omega) = \omega \quad f_2(\omega) = 1 - \omega \quad (29.4.5)$$

These two functions, while different, will have the same cumulative distribution function. For we have:

$$F_1(u) = \mu_{f_1}((-\infty, u]) = \mu(f_1^{-1}(-\infty, u)) \quad (29.4.6)$$

We can evaluate this case by case to get:

$$F_1(u) = \begin{cases} \mu(\emptyset) = 0, & u \leq 0 \\ \mu([0, u]) = u, & 0 < u < 1 \\ \mu([0, 1]) = 1, & 1 \leq u \end{cases} \quad (29.4.7)$$

Looking at  $F_2$ , we have:

$$F_2(u) = \mu_{f_2}((-\infty, u]) = \mu(f_2^{-1}(-\infty, u)) \quad (29.4.8)$$

Again, evaluating case by case, we get:

$$F_2(u) = \begin{cases} \mu(\emptyset) = 0, & u \leq 0 \\ \mu((1-u, 1])u, & 0 < u < 1 \\ \mu([0, 1]) = 1, & 1 \leq u \end{cases} \quad (29.4.9)$$

Thus,  $F_1 = F_2$ .

### Definition 29.4.3: Random Vector

---

random vector on a probability space  $(\Omega, \mathcal{A}, \mu)$  is an  $\mathcal{A} - \mathcal{B}_n$  measurable function  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ , where  $\mathcal{B}_n$  is the Borel  $\sigma$ -Algebra on  $\mathbb{R}^n$ . ■

As a comment, if  $f : \Omega \rightarrow \mathbb{R}$  is  $\mathcal{A} - \mathcal{B}$  measurable, then  $\mathcal{A}_f \subseteq \mathcal{A}$ . The associated  $\sigma$ -Algebra of a random vector  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  is:

$$\mathcal{A}_{\mathbf{f}} = \{\mathbf{f}^{-1}(B) : B \in \mathcal{B}_n\} \quad (29.4.10)$$

**Theorem 29.4.1.** If  $(\Omega, \mathcal{A}, \mu)$  is a probability space,  $\mathcal{B}_n$  is the Borel  $\sigma$ -Algebra on  $\mathbb{R}^n$ , and if  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  is a random vector such that:

$$\mathbf{f}(\omega) = (f_1(\omega), \dots, f_n(\omega)) \quad (29.4.11)$$

Then:

$$\mathcal{A}_{\mathbf{f}} = \sigma(\mathcal{A}_{f_1}, \dots, \mathcal{A}_{f_n}) \quad (29.4.12)$$

Where this is the  $\sigma$ -Algebra generated by these sets.

*Proof.* For any  $f_j$ ,  $\mathcal{A}_{f_j} \subseteq \mathcal{A}_{\mathbf{f}}$ , and thus the generated  $\sigma$ -Algebra is contained in  $\mathcal{A}_{\mathbf{f}}$ . Going the other way, let  $\tilde{\mathcal{B}}$  be the set of subsets  $B \subseteq \mathbb{R}^n$  such that:

$$\mathbf{f}^{-1}(B) \in \sigma(\mathcal{A}_{f_1}, \dots, \mathcal{A}_{f_n}) \quad (29.4.13)$$

But then for any sequence  $B_1, \dots, B_n \in \mathcal{B}$ ,  $B_1 \times \dots \times B_n$  is contained in  $\tilde{\mathcal{B}}$ . But  $\mathcal{B}_n$  is the smallest such  $\sigma$ -Algebra to contain such sets, and thus  $\mathcal{B}_n \subseteq \tilde{\mathcal{B}}$ . □

### Definition 29.4.4: Distribution of a Random Vector

---

he distribution of a random vector  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  on a measure space  $(\Omega, \mathcal{A}, \mu)$  is the measure:

$$\mu_{\mathbf{f}}(B) = \mu(\mathbf{f}^{-1}(B)) \quad (29.4.14)$$

Which is the joint distribution of  $f_1, \dots, f_n$ , where:

$$\mathbf{f}(\omega) = (f_1(\omega), \dots, f_n(\omega)) \quad (29.4.15)$$



The individual distributions can be computed in terms of the joint distribution. This is because:

$$\mu_{f_1}(B) = \mu(f_1^{-1}(B)) = \mu(\mathbf{f}^{-1}(B \times \mathbb{R}^{n-1})) = \mu_{\mathbf{f}}(B \times \mathbb{R}^{n-1}) \quad (29.4.16)$$

The joint distribution can not, in general, be computed in terms of the individual distributions. There is a special exception to this rule, and that is when the random variables are independent. That is, if the associated  $\sigma$ -Algebras are independent. So events that bear on  $f_1, \dots, f_n$  are independent. If  $E_j \in \mathcal{A}_{f_j}$ , then:

$$\mu\left(\bigcap_{k=1}^n E_k\right) = \prod_{k=1}^n \mu(E_k) \quad (29.4.17)$$

**Theorem 29.4.2.** *A sequence of random variables  $f_1, \dots, f_n$  are independent if and only if the joint distribution is the product measure of the individual distributions.*

*Proof.* For let  $B_k \in \mathcal{B}$  and let:

$$E_k = f_k^{-1}(B_k) \quad (29.4.18)$$

But then:

$$\mu\left(\bigcap_{k=1}^n E_k\right) = \mu\left(\bigcap_{k=1}^n f_k^{-1}(B_k)\right) \quad (29.4.19a)$$

$$= \mu(\mathbf{f}^{-1}(B_1 \times \cdots \times B_n)) \quad (29.4.19b)$$

$$= \mu_{\mathbf{f}}(B_1 \times \cdots \times B_n) \quad (29.4.19c)$$

$$= \prod_{k=1}^n \mu(E_k) \quad (29.4.19d)$$

$$= \prod_{k=1}^n \mu(f_k^{-1}(B_k)) \quad (29.4.19e)$$

$$= \prod_{k=1}^n \mu_{f_k}(B_k) \quad (29.4.19f)$$

□

Let  $\mu_1, \dots, \mu_n$  be probability Lebesgue-stieljes measures on  $\mathbb{R}$ , and let  $\mu$  be the product measure. Consider the probability space  $(\mathbb{R}^n, \mathcal{B}_n, \mu)$  and the projection mappings  $\pi_k : \mathbb{R}^n \rightarrow \mathbb{R}$ :

$$\pi_k(\omega_1, \dots, \omega_n) = \omega_k \quad (29.4.20)$$

**Theorem 29.4.3.** Let  $f_n$  be an infinite sequence of random variables on a probability space  $(\Omega, \mathcal{A}, \mu)$ . Let  $\mathcal{A}_{f_n}$  be the associated  $\sigma$ -Algebras. For every  $\omega \in \Omega$ , let:

$$F_{\inf}(\Omega) = \underline{\lim}_{n \rightarrow \infty} f_n(\omega) \quad F_{\sup}(\Omega) = \overline{\lim}_{n \rightarrow \infty} f_n(\omega) \quad (29.4.21)$$

Then  $F_{\inf}$  and  $F_{\sup}$  are measurable with respect to the terminal  $\sigma$ -Algebra.

*Proof.* For  $F_{\inf}$  is measurable if and only if for all  $u \in \mathbb{R}$ , we have  $F^{-1}((-\infty, u)) \in \mathcal{F}$ . But:

$$F^{-1}((-\infty, u)) = \{\omega : F(\omega) \leq u\} \quad (29.4.22a)$$

$$= \{\omega : \underline{\lim} f_n(\omega) \leq u\} \quad (29.4.22b)$$

$$= \{\omega : \sup_{n \in \mathbb{N}} \liminf_{k \geq n} f_k(\omega) \leq u\} \quad (29.4.22c)$$

$$= \bigcap_{n=1}^{\infty} \left\{ \omega : \inf_{n \geq k} f_k(\omega) \leq u \right\} \quad (29.4.22d)$$

$$= \bigcap_{n=N}^{\infty} \left\{ \omega : \inf_{n \geq k} f_k(\omega) \leq u \right\} \quad (29.4.22e)$$

□

**Theorem 29.4.4.** If  $\mathcal{F}$  is a self-independent  $\sigma$ -Algebra, if  $F$  is measurable with respect to  $\mathcal{F}$ , then  $F$  is constant almost everywhere.

*Proof.* For since  $\mathcal{F}$  is self independent:

$$\mu(\{\omega : F(\omega) < u\}) = 0 \quad \text{or} \quad \mu(\{\omega : F(\omega) < u\}) = 1 \quad (29.4.23)$$

Define  $A$  and  $B$  as follows:

$$A = \{u \in \mathbb{R} : \mu(\{\omega : F(\omega) < u\}) = 0\} \quad (29.4.24)$$

$$B = \{u \in \mathbb{R} : \mu(\{\omega : F(\omega) < u\}) = 1\} \quad (29.4.25)$$

$$(29.4.26)$$

This separates the real line into two parts. By Dedekind's Axiom there is a  $c \in \mathbb{R}$  such that, for all  $a \in A$ , and for all  $b \in B$ ,  $a \leq c \leq b$ . But then:

$$\mu(\{u : F(u) < c + \frac{1}{n}\}) = 1 \quad (29.4.27)$$

From continuity from above, we're done. □

**Theorem 29.4.5.** If  $(\Omega, \mathcal{A}, \mu)$  is a probability space,  $f_n$  is a sequence of independent random variables, then the limit inferior and the limit superior are constants  $\mu$  almost everywhere.

*Proof.* For the limit inferior and limit superior are measurable with respect to the terminal  $\sigma$ -Algebra. By the Kolmogorov zero-one law,  $\mathcal{F}$  is self-independent if  $\mathcal{A}_{f_n}$  are independent. Thus, by the previous theorem, these functions are constants almost everywhere.  $\square$

Thus the limit of random-variables is entirely not random, but constant functions.

**Theorem 29.4.6.** *If  $f_n$  is a sequence of random variables, then the limit of  $f_n$  almost surely exists, or almost never exists.*

*Proof.* For since the limit superior and limit inferior are constants almost everywhere, then either they agree, in which there's convergence almost surely, or they do not agree, in which there's convergence almost never.  $\square$

### Definition 29.4.5: Expectation Value

---

The expectation value of a summable random variable  $f : \Omega \rightarrow \mathbb{R}$  on a measure space  $(\Omega, \mathcal{A}, \mu)$  is the real number  $E(f)$  defined by:

$$E(f) = \int_{\Omega} f \, d\mu \quad (29.4.28)$$



The expectation can be expressed in terms of the distribution by using the measure transformation theorem. If  $g : \mathbb{R} \rightarrow \mathbb{R}$  is a real valued function, then:

$$\int_{\Omega} g \, d\mu = \int_{\mathbb{R}} g \circ f \, d\mu_f \quad (29.4.29)$$

Now we apply this in the simple case when  $g(u) = u$ . Then:

$$E(f) = \int_{\Omega} f \, d\mu = \int_{\mathbb{R}} u \, d\mu_f \quad (29.4.30)$$

Where we assume that  $f$  is summable against  $\mu$ . Thus,  $u$  is summable against  $\mu_f$ . So, we have that:

$$\int_{\mathbb{R}} |u| \, d\mu_f < \infty \quad (29.4.31)$$

### Definition 29.4.6: Variance

---

The variance of a random variable  $f : \Omega \rightarrow \mathbb{R}$  on a measure space  $(\Omega, \mathcal{A}, \mu)$ , is

the real number  $Var(f)$  defined by:

$$Var(f) = E(f - E(f))^2 = \int_{\Omega} (f - E(f))^2 d\mu \quad (29.4.32)$$

■

**Theorem 29.4.7.**

$$Var(f) = E(f^2) - E(f)^2 \quad (29.4.33)$$

## 29.5 Lecture 8-ish Maybe

If  $(\Omega, \mathcal{A}, \mu)$  is a measure space,  $f : \Omega \rightarrow \mathbb{R}$  is a Borel measurable function, then the expectation is:

$$E(f) = \int_{\Omega} f d\mu \quad (29.5.1)$$

The functions  $f_1, \dots, f_n$  are independent if the associated  $\sigma$ -Algebras are independent,  $\mathcal{A}_{f_1}, \dots, \mathcal{A}_{f_n}$ , where the associated  $\sigma$ -Algebra is defined as:

$$\mathcal{A}_f = \{f^{-1}(B) : B \in \mathcal{B}\} \quad (29.5.2)$$

Where  $\mathcal{B}$  is the Borel  $\sigma$ -Algebra. A random vector is a function  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ . The distribution of  $\mathbf{f}$  is defined as:

$$\mu_{\mathbf{f}}(B) = \mu(\mathbf{f}^{-1}(B)) \quad (29.5.3)$$

This is also called the joint distribution. We then proved that  $f_1, \dots, f_n$  are independent if and only if the joint distribution is the product of the individual distributions.

**Theorem 29.5.1.** *If  $(\Omega, \mathcal{A}, \mu)$  is a probability space, and if  $f_1, \dots, f_n$  are independent functions, then:*

$$E\left(\prod_k f_k\right) = \prod_k E(f_k) \quad (29.5.4)$$

*Proof.* For define  $g : \Omega \rightarrow \mathbb{R}$  by:

$$g(\omega) = \prod_{k=1}^n f_k(\omega) \quad (29.5.5)$$

Let  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$  be defined by:

$$\mathbf{f}(\omega) = (f_1(\omega), \dots, f_n(\omega)) \quad (29.5.6)$$

Then using the measure transformation, we have:

$$\int_{\Omega} \prod_{k=1}^n f_k d\mu = \int_{\Omega} g(\mathbf{f}(\omega)) d\mu \quad (29.5.7)$$

$$\int_{\mathbb{R}^n} g(u_1, \dots, u_n) \mu_{\mathbf{f}} \quad (29.5.8)$$

$$= \int_{\mathbb{R}^n} \prod_{k=1}^n u_k \mu_{\mathbf{f}} \quad (29.5.9)$$

□

Suppose  $n = 2$ . Then, since  $f_1$  and  $f_2$  are independent,  $\mu_{(f_1, f_2)}$  is the product of the measures  $\mu_{f_1}$  and  $\mu_{f_2}$ . Thus by Fubini's theorem:

$$\int_{\mathbb{R}^2} u_1 u_2 \mu_{(f_1, f_2)} = \int_{\mathbb{R}} \left( \int_{\mathbb{R}} u_1 u_2 \mu_{f_2} \right) \mu_{f_1} = \int_{\mathbb{R}} u_1 \left( \int_{\mathbb{R}} u_1 \mu_{f_1} \right) = \int_{\mathbb{R}} u \mu_{f_1} \int_{\mathbb{R}} u_2 \mu_{f_2} = \int_{\Omega} f_1 f_2 \quad (29.5.10)$$

From a course in integral calculus, one should be very surprised by this result, for it says that if  $f_1, \dots, f_n$  are independent, then:

$$\int_{\Omega} \prod_{k=1}^n f_k d\mu = \prod_{k=1}^n \int_{\Omega} f_k d\mu \quad (29.5.11)$$

This is almost never true for a given set of functions, but if they are independent then the result holds.

### 29.5.1 Covariance

The covariance of  $f_1$  and  $f_2$  is:

$$E((f_1 - E(f_1))(f_2 - E(f_2))) = \int_{\Omega} (f_1 - E(f_1))(f_2 - E(f_2)) d\mu \quad (29.5.12)$$

We can simplify this down to:

$$E(f_1 f_2) - E(f_1)E(f_2) \quad (29.5.13)$$

If  $f_1$  and  $f_2$  are independent, then:

$$Cov(f_1, f_2) = 0 \quad (29.5.14)$$

The converse is not true. It does not imply that  $f_1$  and  $f_2$  are independent. For let:

$$\Omega = \{1, 2, 3\} \quad (29.5.15)$$

Let  $\mathcal{A} = \mathcal{P}(\Omega)$  and let  $\mu$  be the counting measure on  $\Omega$ . That is:

$$\mu(A) = \frac{\text{curl}(A)}{3} \quad (29.5.16)$$

Then  $(\Omega, \mathcal{A}, \mu)$  is a probability measure. Define  $f_1$  and  $f_2$  as follows:

$$f_1(\omega) = \begin{cases} 1, & \omega = 1 \\ 0, & \omega = 0 \\ 1, & \omega = 2 \end{cases} \quad (29.5.17)$$

$$f_2(\omega) = \begin{cases} 1, & \omega = 1 \\ 0, & \omega = 0 \\ -1, & \omega = 2 \end{cases} \quad (29.5.18)$$

Then we compute and get:

$$E(f_1) = \int_{\Omega} f_1 \, d\mu = 0 \quad (29.5.19)$$

And also:

$$E(f_2) = \frac{2}{3} \quad (29.5.20)$$

But if we multiply, we see that  $f_1 f_2 = f_1$ , and therefore:

$$E(f_1 f_2) = E(f_1) = 0 \quad (29.5.21)$$

But then:

$$E(f_1 f_2) - E(f_1)E(f_2) = 0 \quad (29.5.22)$$

And thus  $f_1$  and  $f_2$  are uncorrelated. But they are dependent. We may expect this since  $f_2 = f_1^2$ . Let's compute the associated  $\sigma$ -Algebras. We have:

$$f_1^{-1}(\{1\}) = \{1\} \quad (29.5.23)$$

$$f_2^{-1} = \{1, 3\} \quad (29.5.24)$$

But then:

$$\mu(f_1^{-1}(\{1\})) = \frac{1}{3} \quad (29.5.25)$$

$$\mu(f_2^{-1}(\{1\})) = \frac{2}{3} \quad (29.5.26)$$

But the product measure is:

$$\mu_{(f_1, f_2)}(\{1\}) = \frac{1}{3} \quad (29.5.27)$$

And this is not the product of the two measure, and therefore it they are not independent. If  $Cov(f_1, f_2) = 0$ , we say that  $f_1$  and  $f_2$  are uncorrelated.

**Theorem 29.5.2.** If  $f_1, \dots, f_n$  are random variables that are pairwise uncorrelated, then:

$$\text{Var}\left(\sum_{k=1}^n f_k\right) = \sum_{k=1}^n \text{Var}(f_k) \quad (29.5.28)$$

*Proof.* For:

$$\int_{\Omega} \left( \sum_{k=1}^n f_k - E\left(\sum_{k=1}^n f_k\right) \right) d\mu = \sum_{i,j} \int_{\Omega} (f_i - E(f_i))(f_j - E(f_j)) d\mu \quad (29.5.29)$$

But the  $f_i$  are pairwise uncorrelated, and thus this product is zero if  $i \neq j$ . Thus, we get:

$$\int_{\Omega} \left( \sum_{k=1}^n f_k - E\left(\sum_{k=1}^n f_k\right) \right) d\mu = \sum_{k=1}^n \text{Var}(f_k) \quad (29.5.30)$$

□

## 29.6 Laws of Large Numbers

Consider a fair coin and toss it  $n$  times. We would expect that, as  $n$  gets large, the number of times heads occurs and the number of times tails occurs is roughly the same. That is:

$$\frac{|\text{Heads}| - |\text{Tails}|}{n^2} \rightarrow 0 \quad (29.6.1)$$

And also:

$$\frac{|\text{Heads}| \times |\text{Tails}|}{n} \rightarrow \frac{1}{2} \quad (29.6.2)$$

We want to build a more rigorous notion from this idea and create a mathematical model out of this. We use probability spaces as this model. Let  $(\Omega, \mathcal{A}, \mu)$  be a probability space and let  $f_j : \Omega \rightarrow \mathbb{R}$  be random variables take on the values -1 and 1, and such that they are independent. Then the associated  $\sigma$ -Algebra are:

$$\mathcal{A}_{f_j} = \{\emptyset, f^{-1}(\{-1\}), f^{-1}(\{1\}), \Omega\} \quad (29.6.3)$$

The measure on the space is such that:

$$\mu\left(f^{-1}(\{-1\})\right) = \mu\left(f^{-1}(\{1\})\right) = \frac{1}{2} \quad (29.6.4)$$

Define a new function by:

$$F_N(\omega) = \frac{1}{N} \sum_{k=1}^N f_k(\omega) \quad (29.6.5)$$

Then  $F_N(\omega)$  is the number of times 1 occurs minus the number of times -1 occurs, divided by  $N$ . It seems likely that this function should converge to zero for large  $N$ . Recall that there are three different types of convergence. We say  $g_n \rightarrow g$  almost everywhere if there is a set of measure 0 such that  $g_n \rightarrow g$  on the complement of this set. We say that  $g_n \rightarrow g$  almost uniformly if there is a set of arbitrarily small measure  $\varepsilon$  such that  $g_n \rightarrow g$  uniformly on the complement. Finally,  $g_n \rightarrow g$  in measure if for all  $\delta > 0$ :

$$\mu\left(\{\omega : |g_n(\omega) - g(\omega)| \geq \delta\}\right) \rightarrow 0 \quad (29.6.6)$$

We have seen the almost uniform convergence is the strongest and implies the other two. By Egorov, since  $\mu(\Omega) = 1$  in a probability space, convergence almost everywhere implies convergence almost uniformly. Lastly, convergence in measure implies there is a subsequence that converges almost uniformly.

### Definition 29.6.1: Strong Law of Large Numbers

sequence that obeys the Strong Law of Large Numbers in a probability space  $(\Omega, \mathcal{A}, \mu)$  is a sequence  $f_n$  such that:

$$\frac{1}{N} \sum_{n=1}^N [f_n(\omega) - E(f_n)] \rightarrow 0 \quad (29.6.7)$$

$\mu$  almost everywhere. ■

### Definition 29.6.2: Weak Law of Large Numbers

sequence that obeys the Weak Law of Large Numbers in a probability space  $(\Omega, \mathcal{A}, \mu)$  is a sequence  $f_n$  such that:

$$\frac{1}{N} \sum_{n=1}^N [f_n(\omega) - E(f_n)] \rightarrow 0 \quad (29.6.8)$$

Where the convergence is in measure. ■

### Theorem 29.6.1: Khinchin's Weak Law of Large Numbers

If  $(\Omega, \mathcal{A}, \mu)$  is a probability space, if  $f_j$  is a sequence of random variables that are pair-wise uncorrelated such that:

$$\frac{1}{n^2} \sum_{j=1}^n \text{Var}(f_k) \rightarrow 0 \quad (29.6.9)$$

Then  $f_j$  obeys the Weak Law of Large Numbers. ■

*Proof.* For:

$$\int_{\Omega} \left( \frac{1}{n} \sum_{k=1}^n (f_k(\omega) - E(f_k)) \right)^2 d\mu = \frac{1}{n^2} \sum_{k=1}^n Var(f_k) \quad (29.6.10)$$

Let:

$$\Omega_{\delta,n} = \{\omega : \left| \frac{1}{n} \sum_{k=1}^n (f_k(\omega) - E(f_k)) \right| \geq \delta\} \quad (29.6.11)$$

But by the Chebyshev inequality, we have:

$$\int_{\Omega} \left( \frac{1}{n} \sum_{k=1}^n (f_k(\omega) - E(f_k)) \right)^2 d\mu \geq \int_{\Omega_{\delta}} \left( \frac{1}{n} \sum_{k=1}^n (f_k(\omega) - E(f_k)) \right)^2 d\mu \geq \delta^2 \int_{\Omega_{\delta}} d\mu \quad (29.6.12)$$

But then:

$$\mu(\Omega_{\delta,n}) \leq \frac{1}{\delta^2} \frac{1}{n} \sum_{j=1}^n V(f_k) \quad (29.6.13)$$

But this last part tends to zero. Therefore, etc.  $\square$

### Example 29.6.1: I

all of the  $f_i$  have the same distribution, or if they are uniformly bounded, then the theorem applies. This can be used to show that our model for a fair coin toss obeys the weak law of large numbers.  $\blacksquare$

Suppose  $g_n(\omega) \rightarrow g(\omega)$  almost everywhere. Then, for all  $\delta > 0$  there is an  $N$  such that, for all  $n > N$ , we have:

$$|g_n(\omega) - g(\omega)| < k^{-1} \quad (29.6.14)$$

For some  $k$ . Consider the negation of this claim. Then there exists  $k \in \mathbb{N}$  such that, for all  $N \in \mathbb{N}$  there is an  $n > N$  such that:

$$|g_n(\omega) - g(\omega)| \geq k^{-1} \quad (29.6.15)$$

Consider the following set:

$$B = \bigcup_{n=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{k=N}^{\infty} \{\omega : |g_n(\omega) - g(\omega)| \geq k^{-1}\} \quad (29.6.16)$$

This is the set of  $\omega$  such that  $g_n(\omega) \not\rightarrow g(\omega)$ . We wish to show that  $\mu(B) = 0$ . This will happen if and only if for all  $k \in \mathbb{N}$ :

$$\mu \left( \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \{\omega : |g_n(\omega) - g(\omega)| \geq k^{-1}\} \right) = 0 \quad (29.6.17)$$

Consider a collection of set  $A_n$  and define:

$$\bar{A} = \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n \quad (29.6.18)$$

If the  $A_n$  are independent, then  $\bar{A}$  is a terminal event, and thus by the Kormogorov zero-one law, either  $\mu(\bar{A}) = 1$  or  $\mu(\bar{A}) = 0$ .

**Theorem 29.6.2.** *If:*

$$\sum_{n=1}^{\infty} \mu(A_n) < \infty \quad (29.6.19)$$

*Then:*

$$\mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n\right) = 0 \quad (29.6.20)$$

*Proof.* For:

$$\mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n\right) \leq \mu\left(\bigcup_{n=n}^{\infty} A_n\right) \leq \sum_{n=N}^{\infty} \mu(A_n) \quad (29.6.21)$$

But this sum converges, and thus the tail end can be made arbitrarily small.  $\square$

### **Theorem 29.6.3: Borel-Cantelli Lemma**

If  $A_n$  are pair-wise independent and are such that:

$$\sum_{k=1}^{\infty} \mu(A_k) = \infty \quad (29.6.22)$$

Then:

$$\mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n\right) = 1 \quad (29.6.23)$$

■

*Proof.* For if:

$$\mu\left(\bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} A_n^C\right) = 0 \quad (29.6.24)$$

Then, for all  $N$ :

$$\mu\left(\bigcap_{n=N}^{\infty} A_n^C\right) = 0 \quad (29.6.25)$$

So it suffices to show that this is true. For let  $N \in \mathbb{N}$ , and define:

$$B = \bigcap_{n=N}^{\infty} A_n^C \quad (29.6.26)$$

Also define:

$$B_M = \bigcap_{n=N}^M A_n^C \quad (29.6.27)$$

It then follows from continuity from below that:

$$\mu(B) = \lim_{M \rightarrow \infty} \mu(B_M) = \lim_{M \rightarrow \infty} \mu\left(\bigcap_{n=N}^M A_n^C\right) \quad (29.6.28)$$

But from independence, we obtain:

$$\mu(B) = \lim_{M \rightarrow \infty} \prod_{n=N}^M \mu(A_n^C) = \lim_{M \rightarrow \infty} \prod_{n=N}^M \mu(1 - A_n) \quad (29.6.29)$$

Using the exponential function, we note that  $1 - x \leq \exp(-x)$ , and so:

$$\mu(B) \leq \lim_{M \rightarrow \infty} \prod_{n=N}^M \exp(-\mu(A_n)) = \lim_{M \rightarrow \infty} \exp\left(\sum_{n=N}^M \mu(A_n)\right) = 0 \quad (29.6.30)$$

□

The independence of the  $A_n$  is indeed necessary for this theorem. For let  $A_n = A_0$ , and let  $\mu(A_0) = \frac{1}{2}$ . Then the sum will indeed diverge, but the measure of final set is still  $\frac{1}{2}$ . The Borel-Cantelli lemma thus complements the Kormogrov Zero-One law by giving the precise criterion for when the measure is either one or zero. Given a sequence of random events, the terminal event has measure one if and only if the sum of the individual measures converges, and is equal to one otherwise.

#### Theorem 29.6.4: Borel's Strong Law of Large Numbers

If  $f_n$  is a sequence of random variables such that:

$$\int_{\Omega} |f_n|^4 d\mu \leq M \quad (29.6.31)$$

For all  $n \in \mathbb{N}$ , then  $f_n$  obeys the strong law of large numbers. ■

*Proof.* It suffices to show that, for all  $\varepsilon > 0$ :

$$\mu\left(\bigcap_{N=1}^{\infty} \bigcup_{N=n}^{\infty} \left\{\omega : \left|\frac{1}{n} \sum_{k=1}^n f_k(\omega)\right| \geq \varepsilon\right\}\right) = 0 \quad (29.6.32)$$

Denote the sequence of centered random variables by:

$$\overset{\circ}{f}_n(\omega) = f_n(\omega) = E(f_n(\omega)) \quad (29.6.33)$$

To show this, we need to show that:

$$\sum_{n=1}^{\infty} \mu\left(\left\{\omega : \left|\frac{1}{n} \sum_{k=1}^n \overset{\circ}{f}_n(\omega)\right| \geq \varepsilon\right\}\right) < \infty \quad (29.6.34)$$

Define:

$$\Omega_{n,\varepsilon} = \left\{\omega : \left|\frac{1}{n} \sum_{k=1}^n \overset{\circ}{f}_n(\omega)\right| \geq \varepsilon\right\} \quad (29.6.35)$$

But then:

$$\int_{\Omega} \left|\frac{1}{n} \sum_{k=1}^n \overset{\circ}{f}_n\right|^4 d\mu \geq \int_{\Omega_{n,\varepsilon}} \left|\frac{1}{n} \sum_{k=1}^n \overset{\circ}{f}_n\right|^4 d\mu \geq \varepsilon^4 \mu(\Omega_{n,\varepsilon}) \quad (29.6.36)$$

Combining this together, we have:

$$\mu(\Omega_{n,\varepsilon}) \leq \frac{1}{\varepsilon^4} \frac{1}{n^4} \int_{\Omega} \left(\sum_{k=1}^n \overset{\circ}{f}_k(\omega)\right)^4 d\mu = \frac{1}{\varepsilon^4} \frac{1}{n^4} \sum_{i,j,k,\ell} \int_{\Omega} \overset{\circ}{f}_i \overset{\circ}{f}_j \overset{\circ}{f}_k \overset{\circ}{f}_{\ell} d\mu \quad (29.6.37)$$

But the  $f_n$  are independent, and thus the  $\overset{\circ}{f}_n$  are independent. But then  $\mathcal{A}_{\overset{\circ}{f}_n}$  are independent, and thus  $\overset{\circ}{f}_i$  is independent from the product  $\overset{\circ}{f}_j \overset{\circ}{f}_k \overset{\circ}{f}_{\ell}$ . But if they are independent, then:

$$\int_{\Omega} \overset{\circ}{f}_i \overset{\circ}{f}_j \overset{\circ}{f}_k \overset{\circ}{f}_{\ell} d\mu = \int_{\Omega} \overset{\circ}{f}_i d\mu \int_{\Omega} \overset{\circ}{f}_j \overset{\circ}{f}_k \overset{\circ}{f}_{\ell} d\mu = 0 \quad (29.6.38)$$

There are two cases left, when the indices are equal in pairs, and when all of the indices are equal. In the cases where all are equal, we have:

$$\sum_{i=1}^n \int_{\Omega} |\overset{\circ}{f}_i|^4 d\mu \leq Mn \quad (29.6.39)$$

For the case of pairs, we have  $n^2 - n$  possibilities, and thus:

$$\sum_{i,j} \int_{\Omega} |\overset{\circ}{f}_i^2 \overset{\circ}{f}_j^2| d\mu \leq M(n^2 - n) \quad (29.6.40)$$

Therefore, we have:

$$\frac{1}{\varepsilon^4} \frac{1}{n^4} \sum_{i,j,k,\ell} \int_{\Omega} \overset{\circ}{f}_i \overset{\circ}{f}_j \overset{\circ}{f}_k \overset{\circ}{f}_{\ell} d\mu \leq \frac{M}{\varepsilon^4} \frac{1}{n^2} \quad (29.6.41)$$

But:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty \quad (29.6.42)$$

Thus, the measure is zero.  $\square$

The  $|f_n|^4$  are called the fourth moments of the  $f_n$ . There are sequences that obey the weak law but not the strong law. Borel's theorem shows that uniformly bounded sequences of random variables automatically obey the strong law of strong numbers, since a uniformly bounded sequence will have uniformly bounded fourth moments. To find a sequence that obeys the weak law but not the strong law, we will need to consider sequences that take on arbitrarily large values.

### Example 29.6.2: L

Let  $f_n$  be a sequence of random variables such that the following are true:

$$\mu(\{\omega : f_n(\omega) = n\}) = \frac{P_n}{2} \quad (29.6.43a)$$

$$\mu(\{\omega : f_n(\omega) = -n\}) = \frac{P_n}{2} \quad (29.6.43b)$$

$$\mu(\{\omega : f_n(\omega) = 0\}) = 1 - P_n \quad (29.6.43c)$$

We need to find a sequence  $P_n$  such that the  $f_n$  will obey the weak law but not the strong law. Choosing the  $P_n$  to be small will most likely result in the sequence obeying the strong law. Indeed, if  $P_n = 0$ , then the  $f_n$  will obey the strong law. In fact, if:

$$P_n \leq \frac{1}{n^2} \quad (29.6.44)$$

Then  $f_n$  will obey the strong law. This is a consequence of the Borel-Cantelli lemma. If  $P_n$  is too large, it may not be true that the  $f_n$  obeys the weak law. For example, suppose:

$$P_n = \frac{1}{n} \quad (29.6.45)$$

We cannot apply Khinchin's theorem, since:

$$\frac{1}{n^2} \sum_{j=1}^n V(f_j) = \frac{n(n+1)}{2n^2} \quad (29.6.46)$$

And this does not converge to zero. Let:

$$P_n = \frac{1}{n \ln(n+2)} \quad (29.6.47)$$

Let's now show that  $f_n$  will obey the weak law. It does. But it does not obey the strong law. We will need to use the Borel-Cantelli lemma. But the sum:

$$\sum_{n=1}^{\infty} \frac{1}{n \ln(n+2)} = \infty \quad (29.6.48)$$

Therefore:

$$\mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \{\omega : |f_n(\omega)|\}\right) = 1 \quad (29.6.49)$$

This contradicts the strong law of large numbers. For suppose not. Then, for almost every  $\omega$ , and for all  $N$ , there is an  $N > N$  such that  $|f_n(\omega)| = n$ . Then there exists a sequence  $n_k$  such that  $|f_{n_k}(\omega)| = n_k$ . But:

$$\frac{1}{n_k} \sum_{j=0}^{n_k} f_j \rightarrow 0 \quad (29.6.50)$$

And therefore:

$$\frac{1}{n_k - 1} \sum_{j=0}^{n_k} f_j \rightarrow 0 \quad (29.6.51)$$

Taking the difference, we get:

$$\frac{1}{n_k} f_{n_k}(\omega) \rightarrow 0 \quad (29.6.52)$$

But  $|f_{n_k}(\omega)| = n_k$ , a contradiction. So the  $f_n$  do not obey the strong law. ■

## 29.6.1 Borel Numbers

Let  $0 \leq x \leq 1$  and suppose  $x$  has the representation  $x = 0.x_1x_2\dots$  and exclude numbers with two representations. For example,  $1 = 0.999\dots$ . The measure of the set of these numbers is zero. Let  $0 \leq a \leq 9$ . Let  $C_n(x)$  be the number of  $a$  among the first  $n$  digits. Then:

$$\frac{C_n(x)}{n} \rightarrow \frac{1}{10} \quad (29.6.53)$$

For almost every  $x$ . Let  $\Omega = [0, 1]$  and  $\mathcal{B}$  be the Borel  $\sigma$ -Algebra. Also, let  $\mu$  be the Lebesgue measure. Consider the functions:

$$f_j(x) = \begin{cases} 1, & x_j = a \\ 0, & x_j \neq a \end{cases} \quad (29.6.54)$$

Then:

$$\frac{C_n(x)}{n} = \frac{1}{n} \sum_{k=1}^n f_k(x) \quad (29.6.55)$$

If the  $f_k$  obey the strong law of large numbers, then:

$$\frac{1}{n} \sum_{k=1}^n (f_k - E(f_k)) \rightarrow 0 \quad (29.6.56)$$

$\mu$  almost everywhere. We have that  $f_j$  are bounded, and thus it suffices to show that they are also independent. Define:

$$\mathcal{A}_{f_i} = \{\emptyset, A_j, A_j^C, \Omega\} \quad (29.6.57)$$

Where:

$$A_j = \{x : f_j(x) = 1\} \quad (29.6.58)$$

The  $A_j$  are the set of elements  $x$  such that  $x = 0.x_1x_2\dots ax_{j+1}x_{j+2}\dots$  Using this there are  $10^{j-1}$  options for the first  $j-1$  digits. This set is covered by  $10^{j-1}$  intervals, each of length  $10^{-j}$ . Thus, the Lebesgue measure of  $A_j$  is  $\frac{1}{10}$ . We now need to show that, for distinct  $j, k$ , that the measure of the intersection is  $\frac{1}{100}$ . Suppose  $j < k$ . Then  $A_j \cap A_k$  is the set of numbers with  $a$  in the  $j^{th}$  decimal and  $a$  in the  $k^{th}$  decimal. There are  $10^{j-1}$  ways to choose the first  $j-1$  digits, and  $10^{k-j-1}$  ways to choose the next  $k-j-1$  digits. Total, there are  $10^{k-2}$  digits to choose. So we can cover this set with  $10^{k-2}$  intervals, each of length  $10^{-k}$ . Thus, the measure of the intersection is  $10^{-2}$ . Therefore, the  $f_j$  are independent. By Borel's Strong Law of Large Numbers, the  $f_j$  obey the strong law of large numbers.

Let  $\Omega = \mathbb{Z}_n$ , let  $\mathcal{A}$  be the power set, and let  $\mu$  be the counting measure on  $\Omega$ . Taking the product  $\Omega^n$ , and considering the product measure, we see that every point has measure  $10^{-n}$ . Thus, if we consider the infinite product, points will have measure zero. This isn't too strange since the Lebesgue measure is such that points have measure zero. Let  $\tilde{\Omega}$  be the infinite product and let  $B_n \in \Omega^n$ . Then  $B_n \times \Omega_{n+1} \times \dots$  is contained in  $\tilde{\Omega}$ . Let  $\tilde{\mathcal{B}}$  be the smallest  $\sigma$ -Algebra on the product space that contains all of these types of sets, and let  $\tilde{\mu}$  be the extension measure. Then  $f_j(\omega) = \omega_j$  are independent by construction of the product measure, and also:

$$\mu(f_j = a) = \frac{1}{10} \quad (29.6.59)$$

There is a map  $\tilde{\Omega} \mapsto [0, 1]$  by sending  $(\omega_1, \dots)$  to  $0.\omega_1\omega_2\dots$ . The image measure of the product measure  $\tilde{\mu}$  is the Lebesgue measure. So we have an equivalent model of  $[0, 1]$  with the Lebesgue measure.

## 29.7 Central Limit Theorem

We now wish to discuss convergence of measures and distributions. We restrict ourselves to Lebesgue-Stieljes measures on the Borel  $\sigma$ -Algebra of  $\mathbb{R}$ . What does it mean for a sequence of measures  $\mu_n$  to converge to a measure  $\mu$ ? For all  $B \in \mathcal{B}$ :

$$\mu_n(B) \rightarrow \mu(B) \quad (29.7.1)$$

This is reminiscent of point-wise convergence of functions of a real variable, but turns out to be too much. What if we restrict ourselves to sets of the form  $[a, b)$ ? Let:

$$f_n(x) = \begin{cases} 0, & |x| \geq \frac{1}{n} \\ n(1 - |x|), & |x| < \frac{1}{n} \end{cases} \quad (29.7.2)$$

And define:

$$\mu_n([a, b]) = \int_a^b \rho_n(x) dx \quad (29.7.3)$$

Then by the Caratheodory extension theorem, there is a measure  $\nu_n$  that agrees with  $\mu_n$  on all such intervals. Then  $\nu_n$  converges to the Dirac measure, which is an example of an atomic measure:

$$\delta(B) = \begin{cases} 1, & 0 \in B \\ 0, & 0 \notin B \end{cases} \quad (29.7.4)$$

However:

$$\mu_n([0, b]) \rightarrow \frac{1}{2} \quad (29.7.5)$$

And:

$$\mu_n([a, 0]) \rightarrow \frac{1}{2} \quad (29.7.6)$$

However:

$$\delta([0, b)) = 1 \quad (29.7.7)$$

$$\delta([a, 0)) = 0 \quad (29.7.8)$$

This leads us to the correct definition of measure:

---

**Definition 29.7.1: Convergence of Measure**


---

sequence of measure  $\nu_n$  converges to a measure  $\nu$  if, for all measure sets  $B$  such that  $\nu(\partial B) = 0$ , it is true that  $\nu_n(B) \rightarrow \nu(B)$ . ■

Given:

$$\int_{\mathbb{R}} \chi_{[a,b)} d\nu_n \rightarrow \int_{\mathbb{R}} \chi_{[a,b)} d\nu \quad (29.7.9)$$

We have that  $\nu(\{a\}) = \nu(\{b\}) = 0$ , and thus  $\chi_{[a,b)}$  is continuous  $\nu$  almost everywhere. Suppose  $\nu_n$  and  $\nu$  are probability Lebesgue-Stieltjes measure on  $\mathbb{R}$ . Then we get the equivalent form:

**Theorem 29.7.1.** *If for every continuous bounded function  $g(\omega)$ , we have that:*

$$\int_{\mathbb{R}} g d\mu_n \rightarrow \int_{\mathbb{R}} g d\mu \quad (29.7.10)$$

*Then  $\mu_n \rightarrow \mu$ .*

**Theorem 29.7.2.** *If for every continuous function with bounded support, if:*

$$\int_{\mathbb{R}} g d\mu_n \rightarrow \int_{\mathbb{R}} g d\mu \quad (29.7.11)$$

*Then  $\nu_n \rightarrow \nu$ .*

**Theorem 29.7.3.** *If:*

$$\int_{\mathbb{R}} \exp(itu) d\nu_n \rightarrow \int_{\mathbb{R}} \exp(itu) d\nu \quad (29.7.12)$$

*Then  $\nu_n \rightarrow \nu$ .*

Suppose  $(\Omega, \mathcal{A}, \mu)$  is a probability space, and suppose  $f_n : \Omega \rightarrow \mathbb{R}$  is a sequence of random variables that are  $\mathcal{A} - \mathcal{B}$  measure. Consider the distributions  $\mu_{g_n}$ . If  $g_n \rightarrow g$  in measure, then the distributions converge to  $\mu_g$ .

**Theorem 29.7.4.** *If  $(\Omega, \mathcal{A}, \mu)$  is a probability space, if  $h_n : \Omega \rightarrow \mathbb{R}$  is a sequence of random variables, if  $\mu_{h_n}$  are the distributions of  $g_n$ , and if  $h_n \rightarrow h$  in measure, then  $\mu_{h_n} \rightarrow \mu_h$ .*

*Proof.* For let  $g$  be a continuous function with compact support. Then, applying the measure transformation theorem, we have:

$$\left| \int_{\mathbb{R}} g d\mu_n - \int_{\mathbb{R}} g d\mu_h \right| = \left| \int_{\Omega} g(h_n) d\mu - \int_{\Omega} g(h) d\mu \right| \leq \int_{\Omega} |g_n(h) - g(h)| d\mu \quad (29.7.13)$$

But  $g$  is continuous on a compact set, and is therefore uniformly continuous. Thus, for all  $\varepsilon > 0$  there is a  $\delta > 0$  such that, for all  $|u' - u''| < \delta$ , we have that  $|g(u') - g(u'')| < \varepsilon$ . Define the following:

$$E_{1,n,\varepsilon} = \{\omega : |h_n(\omega) - h(\omega)| \geq \delta\} \quad (29.7.14)$$

$$E_{2,n,\varepsilon} = \{\omega : |h_n(\omega) - h(\omega)| < \delta\} \quad (29.7.15)$$

Then:

$$\int_{\Omega} |g_n(h) - g(h)| d\mu = \int_{E_{1,n,\varepsilon}} |g_n(h) - g(h)| d\mu + \int_{E_{2,n,\varepsilon}} |g_n(h) - g(h)| d\mu \quad (29.7.16)$$

$$\leq 2M\mu(E_{1,n,\varepsilon}) + \varepsilon\mu(E_{2,n,\varepsilon}) \quad (29.7.17)$$

And this converges to  $\varepsilon$ .  $\square$

The converse of this theorem is not true in general, since vastly different functions can have the same distributions. There is a special case, however, where the converse holds. Consider a function  $h$  such that it's distribution is the Dirac distribution. That is:

$$\mu(\{\omega : h(\omega) = a\}) = \mu_h(\{a\}) = \delta_a(\{a\}) = 1 \quad (29.7.18)$$

Then  $h(\omega) = a$   $\mu$  almost everywhere, or if we are in a probability space, almost surely.

**Theorem 29.7.5.** *If  $h_n$  is a sequence of random variables such that  $\mu_{h_n} \rightarrow \delta_a$ , where  $\delta_a$  is the Dirac measure centered at  $a$ , then  $h_n \rightarrow a$  almost surely.*

*Proof.* For:

$$\mu(\{\omega : |h_n(\omega) - a| \geq \delta\}) = \mu_{h_n}(\mathbb{R} \setminus (a - \delta, a + \delta)) = 1 - \mu_{h_n}((a - \delta, a + \delta)) \quad (29.7.19)$$

$$\rightarrow 1 - \delta_a((a - \delta, a + \delta)) \quad (29.7.20)$$

$$= 0 \quad (29.7.21)$$

$\square$

Thus, the weak law of large numbers can be restated by saying that, if:

$$\mu_{\frac{1}{n} \sum_{j=1}^n f_j} \rightarrow \delta_0 \quad (29.7.22)$$

Then  $f_j$  obeys the weak law of large numbers.

### 29.7.1 Convergence of Distributions

A distribution is an arbitrary probability Lebesgue-Stieljes measure. That is, a Lebesgue-Stieljes measure such that the measure of the entire space is one. We say that a sequence of distributions  $\nu_n$  converges to a measure  $\nu$  if any of the following equivalent statements holds:

1.  $\nu_n([a, b)) \rightarrow \nu([a, b))$  for all  $a < b$ .
2.  $\nu_n((-\infty, c)) \rightarrow \nu(-\infty, c))$  for all  $c$  such that  $\nu(\{c\}) = 0$ . This requirement implies that  $\nu$  is continuous at  $c$ . That is, if  $F_\nu$  is the cumulative distribution function, then  $F_\nu$  is continuous at  $c$ .
3. For every bounded continuous function  $h$ ,  $\int_{\mathbb{R}} h \, d\nu_n \rightarrow \int_{\mathbb{R}} h \, d\nu$ .
4. For every continuous function with compact support:  $\int_{\mathbb{R}} h \, d\nu_n \rightarrow \int_{\mathbb{R}} h \, d\nu$ .
5.  $\int_{\mathbb{R}} \exp(itu) \, d\nu_n = \int_{\mathbb{R}} \exp(itu) \, d\nu$

**Theorem 29.7.6.** *A sequence of random variables  $f_j$  obeys the weak law of large numbers if and only if:*

$$\mu_{\frac{1}{n} \sum_{j=1}^n f_j} \rightarrow \delta_0 \quad (29.7.23)$$

*Proof.* For:

$$\mu \left( \left\{ \omega : \left| \frac{1}{n} \sum_{j=1}^n \overset{\circ}{f}_k(\omega) \right| \geq \delta \right\} \right) = \mu_{\frac{1}{n} \sum_{k=1}^n \overset{\circ}{f}_j} \left( (-\delta, \delta)^C \right) \quad (29.7.24)$$

□

**Theorem 29.7.7.** *If  $f_j$  is a sequence of random variables such that the second moments are finite, then the first moments are finite.*

*Proof.* For:

$$\int_{\Omega} |f_j| \, d\mu \leq \int_{\Omega} (1 + |f_j|^2) \, d\mu = \int_{\Omega} d\mu + \int_{\Omega} |f_j|^2 \, d\mu = 1 + \int_{\Omega} |f_j|^2 \, d\mu \quad (29.7.25)$$

Therefore, etc. □

### Theorem 29.7.8: Central Limit Theorem

If  $f_j$  are independent and identically distributed, with standard deviation  $\sigma$ , then:

$$\mu_{\frac{1}{\sigma\sqrt{n}} \sum_{j=1}^n \overset{\circ}{f}_j} \rightarrow \nu_{0,1} \quad (29.7.26)$$

Where  $\nu_{0,1}$  is the Gaussian distribution:

$$\nu_{0,1}(B) = \frac{1}{\sqrt{2\pi}} \int_B \exp(-u^2/2) du \quad (29.7.27)$$

*Proof.* We will use the Fourier transform to prove this. We have:

$$\int_{\mathbb{R}} \exp(iut) d\nu_{0,1} = \int_{\mathbb{R}} \exp(itu) \exp(-\frac{u^2}{2}) du = \exp(-t^2/2) \quad (29.7.28)$$

That is, the Fourier transform of a Gaussian is itself. We will use this to make the computation easier. Using the measure transformation theorem, we have:

$$\int_{\mathbb{R}} \exp(iut) \mu_{\frac{1}{\sigma\sqrt{n}} \sum_{j=1}^n \overset{\circ}{f}_j} d\mu = \int_{\Omega} \exp\left(\frac{i}{\sigma\sqrt{n}} \sum_{j=1}^n \overset{\circ}{f}_j(\omega)t\right) d\mu \quad (29.7.29)$$

We invoke independence to get:

$$\int_{\Omega} \exp\left(\frac{i}{\sigma\sqrt{n}} \sum_{j=1}^n \overset{\circ}{f}_j(\omega)t\right) d\mu = \int_{\Omega} \prod_{j=1}^n \exp\left(\frac{i}{\sigma\sqrt{n}} \overset{\circ}{f}_j\right) d\mu \quad (29.7.30a)$$

$$= \prod_{j=1}^n \int_{\Omega} \exp\left(\frac{i}{\sigma\sqrt{n}} \overset{\circ}{f}_j\right) d\mu \quad (29.7.30b)$$

But the distributions are identically distributed, and thus we have:

$$\int_{\Omega} \exp\left(\frac{i}{\sigma\sqrt{n}} \sum_{j=1}^n \overset{\circ}{f}_j(\omega)t\right) d\mu = \left[ \int_{\Omega} \exp\left(iu\frac{t}{\sqrt{n}}\right) d\mu \right]^n \quad (29.7.31)$$

We now need to prove that for an arbitrary Lebesgue-Stieltjes Measure on the Borel  $\sigma$ -Algebra of  $\mathbb{R}$ , such that:

$$\int_{\mathbb{R}} d\mu = 0 \quad \int_{\mathbb{R}} u d\mu = 0 \quad \int_{\mathbb{R}} u^2 d\mu = 1 \quad (29.7.32)$$

Then:

$$\left[ \int_{\mathbb{R}} \exp\left(iu\frac{t}{\sqrt{n}}\right) d\mu \right]^n \rightarrow \exp(-t^2/2) \quad (29.7.33)$$

Consider the function:

$$\varphi_\nu(t) = \int_{\mathbb{R}} \exp(iut) d\nu \quad (29.7.34)$$

In analysis this is the Fourier transform, whereas in probability this is called the characteristic function of  $\nu$ . We are tasked with showing that:

$$\left[ \varphi_\mu \left( \frac{t}{\sqrt{n}} \right) \right]^n \rightarrow \exp(-t^2/2) \quad (29.7.35)$$

If  $\mu$  is a Lebesgue-Stieltjes measure, and if the second moment is finite, and if:

$$\varphi_\nu(t) = \int_{\mathbb{R}} \exp(itu) d\nu \quad (29.7.36)$$

then the first two derivatives of  $\varphi_\nu$  exist and are continuous. Moreover:

$$\varphi_\nu(t) = \varphi_\nu(0) + \varphi'_\nu(0)t + \varphi''_\nu(0)t^2 + h(t) \quad (29.7.37)$$

Where  $h$  is such that:

$$\lim_{t \rightarrow 0} \frac{h(t)}{t^2} = 0 \quad (29.7.38)$$

First, it is continuous. For let  $t_k$  be sequence such that  $t_k \rightarrow t$  and let  $g_k = \exp(it_k u)$ . Then  $|g_k| = 1$ , and is therefore summable. Moreover,  $g_k$  tends to  $\exp(itu)$ . Thus, by the dominated convergence theorem:

$$\lim_{n \rightarrow \infty} \varphi_\nu(t_k) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \exp(it_k u) d\mu = \int_{\mathbb{R}} \lim_{n \rightarrow \infty} \exp(it_k u) d\mu = \varphi_\nu(t) \quad (29.7.39)$$

And thus we have continuity. For differentiability, suppose  $\Delta t_k$  is a sequence that tends to zero, and consider:

$$\frac{\varphi_\nu(t + \Delta t_k) - \varphi_\nu(t)}{\Delta t_k} = \int_{\mathbb{R}} \frac{\exp(iu\Delta t_k) - 1}{\Delta t_k} \exp(iut) d\mu \quad (29.7.40)$$

Again, we want to apply the dominated convergence theorem. Thus we need to find a summable majorant. Consider  $f(s) = (\exp(s) - 1)/s$ . On the real axis, this function has finite limit at zero and has zero limit at infinity, and therefore  $f(s)$  is bounded on the real axis by some  $K$ . Thus,  $K \exp(iut)$  serves as a summable majorant. Applying the dominated convergence theorem shows that the limit exists, and thus  $\varphi_\nu$  is differentiable. We obtain:

$$\varphi'_\nu(t) = \int_{\mathbb{R}} iu \exp(iut) d\nu \quad (29.7.41)$$

Moreover, this is differentiable and:

$$\varphi''_\nu(t) = - \int_{\mathbb{R}} u^2 \exp(iut) d\mu \quad (29.7.42)$$

From Taylor, we have:

$$h(t) = \varphi_\nu(t) - \varphi_\nu(0) - \varphi'_\nu(0)t - \varphi''_\nu(0)\frac{t^2}{2} \quad (29.7.43)$$

Thus  $h''(t)$  exists and is continuous,  $h(0) = 0$ ,  $h'(0) = 0$ , and  $h''(0) = 0$ . By the mean value theorem, we have:

$$h(t) = h'(t_1)t \quad (29.7.44)$$

For some  $t_1 \in (0, t)$ . Moreover:

$$h(t) = h''(t_2)t^2 \quad (29.7.45)$$

Where  $0 < t_1 < t_2 < t$ . Thus:

$$\frac{h(t)}{t^2} = h''(t_2) \quad (29.7.46)$$

And from the continuity of  $h''(t)$ , this converges to zero as  $t$  tends to zero. Thus we have that  $\varphi_\nu(0) = 1$ ,  $\varphi'_\nu(0) = 0$ , and  $\varphi''_\nu(0) = -1$ . Now we need to finally justify the following limit:

$$\left[ \varphi_\mu \left( \frac{t}{\sqrt{n}} \right) \right]^n \rightarrow \exp(-t^2/2) \quad (29.7.47)$$

We have:

$$\varphi_\nu(t) = 1 - \frac{t^2}{2} + h(t) \quad (29.7.48)$$

Where  $h(t)/t^2 \rightarrow 0$  as  $t \rightarrow 0$ . Thus:

$$\left[ \varphi_\nu \left( \frac{t}{\sqrt{n}} \right) \right]^n = \left[ 1 - \frac{t^2}{2n} + h \left( \frac{t}{\sqrt{n}} \right) \right]^n \quad (29.7.49)$$

Define:

$$w_n(t) = h(t/\sqrt{n}) - \frac{t^2}{2n} \quad (29.7.50)$$

Then we have:

$$\left[ \varphi_\nu \left( \frac{t}{\sqrt{n}} \right) \right]^n = \left( \left[ 1 + w_n(t) \right]^{w_n(t)} \right)^{\frac{n}{w_n(t)}} \quad (29.7.51)$$

The inner part is the definition of  $e$ , so we now need to show that  $n/w_n(t)$  converges to  $-t^2/2$ .  $\square$

**Part XVI**

# **Complex Analysis**



# CHAPTER 30

---

## Complex Numbers

---

The theory of complex analysis extends the study of calculus of a single real variable to that of a *complex* variable. The complex numbers have many interesting and counter-intuitive properties, many of which are used regularly in physics.

### 30.1 Complex Numbers

A **complex number** is a point in the plane  $z = (x, y)$ , but we often write:

$$z = x + iy \tag{30.1.1}$$

and call  $i$  the *imaginary unit*. We call  $x$  the *real part* and  $y$  the *imaginary part*, denoted  $\Re(z)$  and  $\Im(z)$ , respectively. The planar representation is shown in Fig. 30.1. The arithmetic goes as follows:

$$(a + ib) + (c + id) = (a + c) + i(b + d) \tag{30.1.2a}$$

$$(a + ib) \cdot (c + id) = (ac - bd) + i(bc + ad) \tag{30.1.2b}$$

We'd hope this definition preserves the arithmetic of the *real* numbers, and indeed it does. Setting  $b$  and  $d$  to zero, we see that elementary arithmetic is recovered.

The arithmetic of the complex numbers arises when one studies equation like  $y(x) = x^2 + 1$ . For a real variable  $x$ , there is no root to this equation. That is, there is no real number  $x$  such that  $x^2 + 1 = 0$ . We can invent such a number and give that the property that its square is  $-1$ . This is what the imaginary unit does.

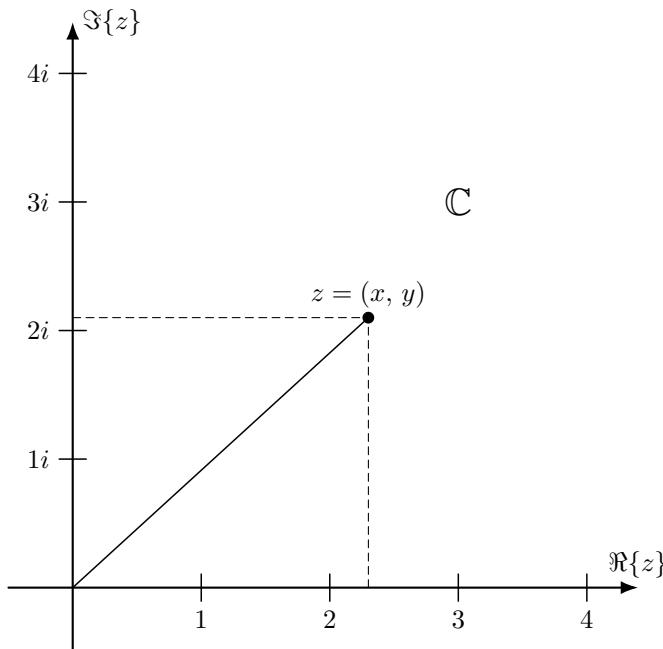


Fig. 30.1: Cartesian Representation of Complex Numbers

It should be clear that addition and multiplication are commutative operations ( $z + w = w + z$  and  $z \cdot w = w \cdot z$ ). That addition is associative is also straightforward. What is not obvious is the associativity of multiplication.

**Theorem 30.1.1.** *If  $z$ ,  $w$ , and  $v$  are complex numbers, then:*

$$z \cdot (w \cdot v) = (z \cdot w) \cdot v \quad (30.1.3)$$

*That is, complex multiplication is associative.*

*Proof.* For let  $z = a + ib$ ,  $w = c + id$ , and  $v = e + if$ . Then:

$$z \cdot (w \cdot v) = (a + ib) \cdot ((c + id) \cdot (e + if)) \quad (30.1.4a)$$

$$= (a + ib) \cdot ((ce - df) + i(cf + de)) \quad (30.1.4b)$$

$$= a(ce - df) - b(cf + de) + i(a(cf + de) + b(ce - df)) \quad (30.1.4c)$$

$$= (ace - adf - bcf - bde) + i(acf + ade + bce - bdf) \quad (30.1.4d)$$

$$= (ac - bd)e - (ad + bc)f + i((ad + bc)e + (ac - bd)f) \quad (30.1.4e)$$

$$= ((a + ib) \cdot (c + id)) \cdot (e + if) \quad (30.1.4f)$$

This completes the proof. □

**Theorem 30.1.2.** *If  $i$  is the imaginary unit, then  $i^2 = -1$ .*

*Proof.* For  $i = 0 + 1i$ , and thus by Eqn. 30.1.2b:

$$i^2 = (0 + 1i) \cdot (0 + 1i) = (0 \cdot 0 - 1 \cdot 1) + i(1 \cdot 0 + 0 \cdot 1) = -1 + i \cdot 0 = -1 \quad (30.1.5)$$

This completes the proof.  $\square$

The complex numbers are *algebraically closed*: Every non-constant polynomial has a *root*, or a zero. Moreover, given a polynomial of degree  $n$  there are at most  $n$  roots. This result is called the *Fundamental Theorem of Algebra*. The real numbers lack this feature, for consider the graph of  $y(x) = x^2 + 1$ . Many attempts at proving this theorem were made between 1608 and 1799, and the likes of Euler, Lagrange, Laplace, Gauss, and d'Alambert failed in their attempts. In 1806 Jean Robert-Argand published a rigorous proof, and due to this the complex plane is occasionally called the Argand plane.

There are two fundamental notions worth mentioning: The complex conjugate and the modulus of a complex number.

### Definition 30.1.1: Complex Conjugate

---

The **complex conjugate** a complex number  $z = x + iy$  is:

$$\bar{z} = x - iy \quad (30.1.6)$$

That is, the reflection of  $z$  across the  $x$  axis.  $\blacksquare$

A visual for the complex conjugate of a complex number is given in Fig. 30.2. There are various arithmetic properties of the complex conjugate that ease the process of computation.

**Theorem 30.1.3.** *If  $z$  is a complex number, then  $z \cdot \bar{z}$  is a non-negative real number.*

*Proof.* For let  $z = x + iy$ , where  $x$  and  $y$  are real numbers. Then, by the definition of the complex conjugate (Def. 30.1.1) and of complex multiplication (Eqn. 30.1.2b):

$$z \cdot \bar{z} = (x + iy) \cdot (x - iy) = x^2 + y^2 \quad (30.1.7)$$

This is the sum of the squares of two real numbers, and is therefore a real and non-negative number. Therefore, etc.  $\square$

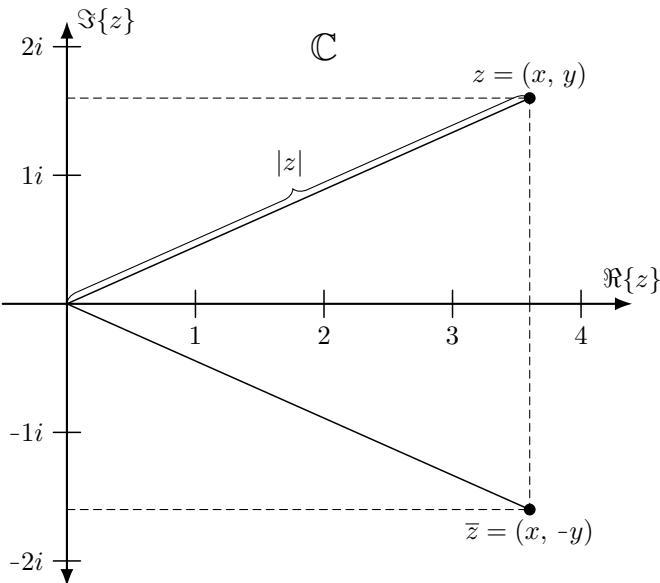


Fig. 30.2: Modulus and Conjugate of a Complex Number

**Theorem 30.1.4.** If  $z$  and  $w$  are complex numbers, then:

$$\overline{z + w} = \bar{z} + \bar{w} \quad (30.1.8)$$

*Proof.* For let  $z = a + ib$  and  $w = c + id$ . Then, by Eqn. 30.1.2a, we have:

$$\overline{z + w} = \overline{(a + ib) + (c + id)} = \overline{(a + c) + i(b + d)} \quad (30.1.9)$$

Invoking the definition of complex conjugate (Def. 30.1.1), we obtain:

$$\overline{z + w} = (a + c) - i(b + d) = (a - ib) + (c - id) = \bar{z} + \bar{w} \quad (30.1.10)$$

Therefore, etc.  $\square$

**Theorem 30.1.5.** If  $z$  and  $w$  are complex numbers, then:

$$\overline{z \cdot w} = \bar{z} \cdot \bar{w} \quad (30.1.11)$$

*Proof.* For let  $z = a + ib$  and  $w = c + id$ . Then, by Eqn. 30.1.2b, we obtain:

$$\overline{z \cdot w} = \overline{(a + ib) \cdot (c + id)} = \overline{(ac - bd) + i(ad + bc)} \quad (30.1.12)$$

Invoking Def. 30.1.1, we have:

$$\overline{z \cdot w} = (ac - bd) - i(ad + bc) = (a - ib) \cdot (c - id) = \bar{z} \cdot \bar{w} \quad (30.1.13)$$

Therefore, etc.  $\square$

From the geometry shown in Fig. 30.2, one would expect adding a complex number to its conjugate would eliminate the imaginary component, and subtracting would eliminate the real part. This is indeed true.

**Theorem 30.1.6.** *If  $z$  is a complex number, then:*

$$z + \bar{z} = 2\Re(z) \quad (30.1.14)$$

*Proof.* For let  $z = x + iy$ . Then, by Def. 30.1.1 and Eqn. 30.1.2a:

$$z + \bar{z} = (x + iy) + (x - iy) = (x + x) + i(y - y) = 2x = 2\Re(z) \quad (30.1.15)$$

Therefore, etc.  $\square$

**Theorem 30.1.7.** *If  $z$  is a complex number, then:*

$$z - \bar{z} = 2i\Im(z) \quad (30.1.16)$$

*Proof.* For let  $z = x + iy$ . Then, by Def. 30.1.1 and Eqn. 30.1.2a:

$$z - \bar{z} = (x + iy) - (x - iy) = (x - x) + i(y + y) = 2iy = 2i\Im(z) \quad (30.1.17)$$

Therefore, etc.  $\square$

Lastly, taking the complex conjugate twice is equivalent to performing two reflection across the  $x$  axis and thus should result in no change.

**Theorem 30.1.8.** *If  $z$  is a complex number, then  $\bar{\bar{z}} = z$ .*

*Proof.* For let  $z = x + iy$ . Then:

$$\bar{\bar{z}} = \overline{\overline{(x + iy)}} = \overline{(x - iy)} = x + iy = z \quad (30.1.18)$$

Therefore, etc.  $\square$

The complex conjugate can be used to define the modulus, or absolute value, of a complex number by simply taking the (positive) square root of  $z\bar{z}$ .

### Definition 30.1.2: Modulus of a Complex Number

The **modulus** of a complex number  $z = x + iy$  is:

$$|z| = \sqrt{x^2 + y^2} \quad (30.1.19)$$

We can also write  $|z| = \sqrt{z\bar{z}}$ , where  $\bar{z}$  is the complex conjugate of  $z$ .  $\blacksquare$

This is the size, or magnitude, of a complex number in the plane, using the Euclidean notion of distance: We compute the length via the Pythagorean formula.

**Theorem 30.1.9.** *If  $z$  a complex number, then  $|z| = |\bar{z}|$ .*

*Proof.* For let  $z = x + iy$ . Then:

$$|z| = \sqrt{x^2 + y^2} = \sqrt{x + (-y)^2} = |\bar{z}| \quad (30.1.20)$$

Therefore, etc.  $\square$

There is one particular theorem that is vital to all areas of mathematical analysis which dates back to Euclid: The Triangle Inequality. To prove this we will need a few results about the modulus of a complex number. Firstly, it is preserved by products, and secondly that the modulus of the real part of complex number is not greater than the entire modulus. That is, the projection of a complex number  $z$  onto the  $x$  axis is less than or equal to the magnitude of  $z$ .

**Theorem 30.1.10.** *If  $z$  and  $w$  are complex numbers, then:*

$$|z \cdot w| = |z| \cdot |w| \quad (30.1.21)$$

*Proof.* For let  $z = a + ib$  and  $w = c + id$ . By Eqn. 30.1.2b, we have:

$$|z \cdot w| = |(a + ib) \cdot (c + id)| = |(ac - bd) + i(ad + bc)| \quad (30.1.22)$$

Using Def. 30.1.2, we obtain:

$$|z \cdot w| = \sqrt{(ac - bd)^2 + (ad + bc)^2} = \sqrt{(ac)^2 + (bd)^2 + (ad)^2 + (bc)^2} \quad (30.1.23)$$

Factoring this gives us the result:

$$|z \cdot w| = \sqrt{(a^2 + b^2)(c^2 + d^2)} = \sqrt{a^2 + b^2} \sqrt{c^2 + d^2} = |z| \cdot |w| \quad (30.1.24)$$

Therefore, etc.  $\square$

**Theorem 30.1.11.** *If  $z$  is a complex number, then:*

$$|\Re(z)| \leq |z| \quad (30.1.25)$$

*Proof.* For let  $z = a + ib$ . Using Def. 30.1.2, we have:

$$|\Re(z)|^2 = |\Re(a + ib)|^2 = |a|^2 \leq |a|^2 + |b|^2 = |z|^2 \quad (30.1.26)$$

Taking the square root of both sides completes the proof.  $\square$

### Theorem 30.1.12: The Triangle Inequality

If  $z$  and  $w$  are complex numbers, then  $|z + w| \leq |z| + |w|$ . ■

*Proof.* Invoking Def. 30.1.2, Thms. 30.1.6, 30.1.8, 30.1.9, 30.1.10, and 30.1.11, we obtain:

$$|z + w|^2 = (z + w) \cdot \overline{(z + w)} \quad (30.1.27a) \qquad = |z|^2 + 2\Re(z\bar{w}) + |w|^2 \quad (30.1.27e)$$

$$= z\bar{z} + z\bar{w} + \bar{z}w + w\bar{w} \quad (30.1.27b) \qquad \leq |z|^2 + 2|z||\bar{w}| + |w|^2 \quad (30.1.27f)$$

$$= |z|^2 + z\bar{w} + \bar{z}w + |w|^2 \quad (30.1.27c) \qquad = |z|^2 + 2|z||w| + |w|^2 \quad (30.1.27g)$$

$$= |z|^2 + z\bar{w} + \bar{z}\bar{w} + |w|^2 \quad (30.1.27d) \qquad = (|z| + |w|)^2 \quad (30.1.27h)$$

Taking the square root of both sides completes the proof.  $\square$

It would be nonsensical to call something the triangle inequality if triangles weren't involved. In Euclid's *Elements* he proves that, given any triangle, the length of one side is less than the sum of the other two. This can be realized in the complex plane by thinking of  $z$ ,  $w$ , and  $z + w$  as points on a triangle (Fig. 30.3). The triangle inequality states that it is shorter to walk from the origin to the point  $z + w$ , than it is to walk from the origin to  $z$ , and then  $z$  to  $z + w$ .

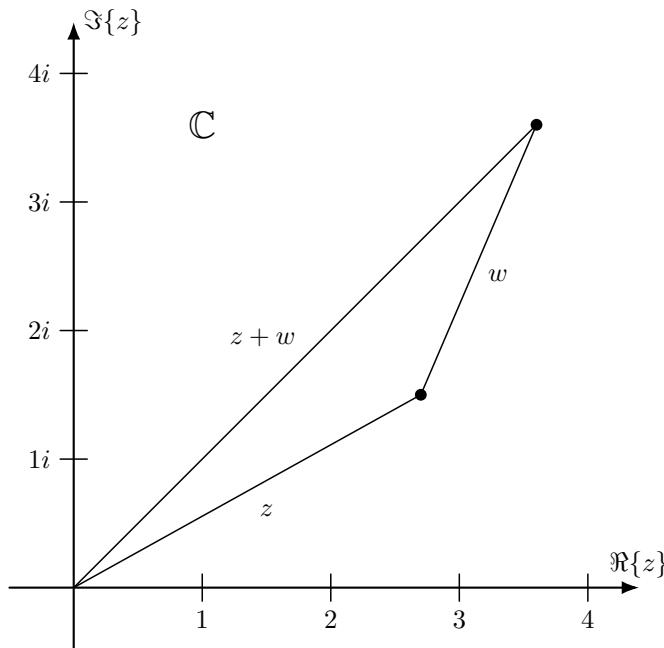


Fig. 30.3: Visual Representation of the Triangle Inequality

The complex conjugate and the modulus of a complex number can combine to form the *inverse* of a non-zero complex number. That, the complex numbers form something called a *field*. A field is a set with two operations, usually called addition and multiplication, such that the operations are commutative, associative, and such that multiplication *distributes* over addition. There is also the requirement of the existence of an *additive identity* and a *multiplicative identity*. Lastly, every number needs an *additive inverse*, and every non-zero number needs a *multiplicative inverse*. It is not difficult to see that the first eight properties are satisfied by the complex numbers, given  $z = x + iy$ ,  $-z = (-x) + i(-y)$  serves as the additive inverse. The last property is tricky, but vital for computations.

**Theorem 30.1.13.** *If  $G$  is a set, if  $*$  is an associative operation on  $G$  with an identity element  $e$ , and if  $x$  has an inverse  $x^{-1}$ , then  $x^{-1}$  is unique.*

*Proof.* For suppose  $x'^{-1}$  is a different inverse. Then:

$$x'^{-1} = x'^{-1} * e = x'^{-1} * (x * x^{-1}) = (x'^{-1} * x) * x^{-1} = e * x^{-1} = x^{-1} \quad (30.1.28)$$

And thus  $x'^{-1} = x^{-1}$ . Therefore, the inverse is unique.  $\square$

From uniqueness, once we've found a candidate for an inverse, we know that

this is indeed the inverse. We now prove that non-zero complex numbers have multiplicative inverses.

**Theorem 30.1.14.** *If  $z$  is a non-zero complex number, then there is a unique  $z^{-1}$  such that  $z \cdot z^{-1} = 1$ . The inverse of  $z = a + ib$  is:*

$$z^{-1} = \frac{a - ib}{a^2 + b^2} \quad (30.1.29)$$

*Proof.* If  $a + ib \neq 0$ , then  $a^2 + b^2 \neq 0$ , so  $(a - ib)/(a^2 + b^2)$  is well defined. But:

$$(a + ib) \cdot \frac{a - ib}{a^2 + b^2} = \frac{(a + ib)(a - ib)}{a^2 + b^2} = \frac{a^2 + b^2}{a^2 + b^2} = 1 \quad (30.1.30)$$

The uniqueness of inverses (Thm. 30.1.13) gives us our result.  $\square$

If  $|z|$  is the modulus of  $z$ , and  $\bar{z}$  is its complex conjugate,  $z^{-1}$  can be written as:

$$z^{-1} = \frac{\bar{z}}{|z|^2} \quad (30.1.31)$$

We will make use of these formulae often, so they are good to keep in mind.

### Example 30.1.1

Consider the complex number  $z = i$ . We can use Eqn. 30.1.29 to compute its multiplicative inverse, and we obtain  $i^{-1} = -i$ . We can also see this since  $i^{-1} \cdot i = 1$ , and we know that  $i^2 = -1$ . Multiplying by  $-1$ , we have  $-i^2 = i^{-1} \cdot i$ . Dividing by  $i$  obtains the result again.  $\blacksquare$

### Example 30.1.2

Now let  $z = (1+i)/F$ , where  $F$  is a non-zero real number. The multiplicative inverse of this is:

$$\left(\frac{1+i}{F}\right)^{-1} = F(1+i)^{-1} = F \frac{1-i}{2} \quad (30.1.32)$$

Invoking Pythagoras, we see that  $|z| = \sqrt{2}/F$ . Using Eqn. 30.1.31, we obtain:

$$z^{-1} = \frac{\bar{z}}{|z|^2} = \frac{\frac{1-i}{F}}{\frac{2}{F^2}} = F \frac{1-i}{2} \quad (30.1.33)$$

In agreement with our previous calculation.  $\blacksquare$

### 30.1.1 Polar Representation of Complex Numbers

In Walter Rudin's classic text on real and complex analysis, he opens with a prologue on the *exponential* function and calls it "Undoubtedly the most important function in mathematics." We take a moment to study this function.

#### Definition 30.1.3: Exponential Function

The complex exponential function is the function  $\exp : \mathbb{C} \rightarrow \mathbb{C}$  defined by:

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad (30.1.34)$$

where  $n!$  denotes the factorial of  $n$ ,  $n! = n \cdot (n-1)!$ , and where  $0! \equiv 1$ . ■

One useful result about the exponential function is that it relates multiplication and addition in a convenient way. We will need Cauchy's product theorem.

#### Theorem 30.1.15: Cauchy's Product Theorem

If  $a : \mathbb{N} \rightarrow \mathbb{C}$  and  $b : \mathbb{N} \rightarrow \mathbb{C}$  are sequences, if  $\sum a_n$  converge absolutely, and if  $\sum b_n$  converges, then:

$$\left( \sum_{j=0}^{\infty} a_j \right) \left( \sum_{k=0}^{\infty} b_k \right) = \sum_{n=0}^{\infty} \sum_{m=0}^n a_m b_{n-m} \quad (30.1.35)$$



*Proof.* Since the two sums  $\sum a_n$  and  $\sum b_n$  converge, let  $A$  and  $B$  be their limits, respectively. For all  $n \in \mathbb{N}$ , define the following partial sums:

$$A_n = \sum_{k=0}^n a_k \quad (30.1.36a) \qquad B_n = \sum_{k=0}^n b_k \quad (30.1.36b)$$

Furthermore, let  $c_n$  be the Cauchy product and  $C_n$  be the partial sums:

$$c_n = \sum_{k=0}^n a_k b_{n-k} \quad (30.1.37a) \qquad C_n = \sum_{m=0}^n c_m \quad (30.1.37b)$$

And finally, let  $\beta_n = B_n - B$ . Then, for all  $n \in \mathbb{N}$ :

$$C_n = \sum_{j=0}^n \sum_{k=0}^j a_k b_{j-k} = \sum_{j=0}^n a_j B_{n-j} = A_n B + \sum_{j=0}^n a_j \beta_{n-j} \quad (30.1.38)$$

Let  $d_n$  be the remainder term. That is:

$$d_n = \sum_{j=0}^n a_j \beta_{n-j} \quad (30.1.39)$$

Since  $\sum a_n$  is absolutely convergent, and thus  $\sum |a_n|$  converges. Let  $A'$  be the limit. Also, by the definition of  $\beta_n$ ,  $\beta_n$  converges to zero. That is, given any  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that, for all  $n > N$ , we have  $|\beta_n| < \varepsilon$ . But then:

$$|d_n| = \left| \sum_{j=0}^N a_j \beta_{n-j} + \sum_{j=N+1}^n a_j \beta_{n-j} \right| \leq \left| \sum_{j=0}^N a_j \beta_{n-j} \right| + \left| \sum_{j=N+1}^n a_j \beta_{n-j} \right| \quad (30.1.40)$$

Where this last step comes from the triangle inequality. Simplifying, we have:

$$|d_n| < \left| \sum_{j=0}^N a_j \beta_{n-j} \right| + \varepsilon A' \quad (30.1.41)$$

The first term can be made small since  $\beta_{n-j}$  is small for large  $n$  (And since  $j < N$ ), and the second term can also be made small since  $\varepsilon$  is arbitrary. So we see that  $d_n$  converges to zero. Thus:

$$\lim_{n \rightarrow \infty} C_n = \lim_{n \rightarrow \infty} A_n B + d_n = AB \quad (30.1.42)$$

This completes the proof.  $\square$

The immediate application of this is the power rule for the exponential function. We will need the *binomial theorem*. Let  $\binom{n}{k}$  (Which reads as  $n$  choose  $k$ ) denote the *binomial coefficient*:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (30.1.43)$$

The binomial theorem then says the following:

### **Theorem 30.1.16: The Binomial Theorem**

---

If  $x$  and  $y$  are real numbers, and if  $n \in \mathbb{N}$ , then:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \quad (30.1.44)$$

Where  $\binom{n}{k}$  denotes the binomial coefficient.  $\blacksquare$

*Proof.* We prove by induction. When  $n = 0$  or  $n = 1$  we can evaluate the validity of this by hand. When  $n = 2$  this is commonly known as the FOIL

rule. Suppose it is true for  $n \in \mathbb{N}$ . We must now show this implies it is true for  $n + 1$ . We have:

$$(x+y)^{n+1} = (x+y)(x+y)^n = (x+y) \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \quad (30.1.45)$$

We can further simplify, and perform a shift of index, to obtain:

$$(x+y)^{n+1} = (x+y) \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \quad (30.1.46a)$$

$$= x \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k + y \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \quad (30.1.46b)$$

$$= \sum_{k=0}^n \binom{n}{k} x^{n+1-k} y^k + \sum_{k=1}^{n+1} \binom{n}{k-1} x^{n+1-k} y^k \quad (30.1.46c)$$

$$= x^{n+1} + \sum_{k=1}^n \left[ \binom{n}{k} + \binom{n}{k-1} \right] x^{n+1-k} y^k + y^{n+1} \quad (30.1.46d)$$

The sum of these two binomial coefficients is known as Pascal's Identity. We have:

$$\binom{n}{k} + \binom{n}{k-1} = \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n+1-k)!} \quad (30.1.47a)$$

$$= n! \frac{(n+1-k)+k}{k!(n+1-k)!} \quad (30.1.47b)$$

$$= \frac{(n+1)!}{k!(n+1-k)!} \quad (30.1.47c)$$

$$= \binom{n+1}{k} \quad (30.1.47d)$$

Thus, returning to Eqn. 30.1.46d, we obtain:

$$(x+y)^{n+1} = x^{n+1} + \sum_{k=1}^n \binom{n+1}{k} x^{n+1-k} y^k + y^{n+1} \quad (30.1.48a)$$

$$= \sum_{k=0}^{n+1} \binom{n+1}{k} x^{n+1-k} y^k \quad (30.1.48b)$$

This completes the proof.  $\square$

**Theorem 30.1.17.** *If  $a$  and  $b$  are complex numbers, then:*

$$\exp(a+b) = \exp(a) \exp(b) \quad (30.1.49)$$

*Proof.* Invoking the *binomial theorem* (Thm. 30.1.16), we have:

$$\exp(a+b) = \sum_{n=0}^{\infty} \frac{(a+b)^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} a^k b^{n-k} \quad (30.1.50)$$

But by Cauchy's product theorem (Thm. 30.1.15), this last double sum can be written as the product of two sums of the form:

$$\sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{k!(n-k)!} a^k b^{n-k} = \left( \sum_{j=0}^{\infty} \frac{a^j}{j!} \right) \left( \sum_{m=0}^{\infty} \frac{b^m}{m!} \right) \quad (30.1.51)$$

But this is just the product of  $\exp(a)$  and  $\exp(b)$ . Therefore, etc.  $\square$

We next prove one of the most important theorems of complex analysis: Euler's Theorem. This is a crucial part of the theory and allows one to define the *polar representation* of a complex number. It relates the exponential function to the trigonometric functions.

**Theorem 30.1.18: Euler's Exponential Formula**

If  $\theta$  is a real number, then:

$$\exp(i\theta) = \cos(\theta) + i \sin(\theta) \quad (30.1.52)$$

*Proof.* Using the definition of the exponential function (Def. 30.1.3) and evaluating  $i\theta$  into this equation, we obtain:

$$\exp(i\theta) = \sum_{n=0}^{\infty} i^n \frac{\theta^n}{n!} \quad (30.1.53)$$

But  $i^n$  cycles between  $i$ ,  $-1$ ,  $-i$ , and  $1$ . So we may split this sum into two parts, a real part and an imaginary part, to obtain:

$$\sum_{n=0}^{\infty} \exp(i\theta) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad (30.1.54)$$

But the left sum is the Taylor expansion for  $\cos(\theta)$ , and the right sum is the Taylor expansion for  $i \sin(\theta)$ . This completes the proof.  $\square$

Euler's Theorem can also be proved by showing the two expressions satisfy the same initial value problem:  $\ddot{z} + z = 0$ ,  $z(0) = 1$ ,  $\dot{z}(0) = i$ . A corollary of this is often hailed as the most beautiful result in mathematics. This is Euler's Identity:

$$e^{i\pi} + 1 = 0 \quad (30.1.55)$$

Combining Euler's Exponential Formula and the product rule for the exponential function, we see that given any complex number  $z = a + ib$ , the following holds:

$$\exp(z) = \exp(a)(\cos(b) + i \sin(b)) \quad (30.1.56)$$

Many of the identities from trigonometry are short corollaries of this theorem, rendering memorization of these formulae redundant.

**Theorem 30.1.19: DeMoivre's Theorem**

If  $n \in \mathbb{N}$  and if  $\theta \in \mathbb{R}$ , then:

$$(\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta) \quad (30.1.57)$$



*Proof.* By Euler's formula (Thm. 30.1.18), and by Thm. 30.1.17, we have:

$$(\cos(\theta) + i \sin(\theta))^n = (\exp(i\theta))^n = \exp(in\theta) = \cos(n\theta) + i \sin(n\theta) \quad (30.1.58)$$

Therefore, etc.  $\square$

Combining the binomial theorem (Thm. 30.1.16) and DeMoivre's Theorem allows one to quickly compute any trigonometric identity one might need. Letting  $n = 2$ , we obtain the double-angle formula:

$$\cos(2\theta) + i \sin(2\theta) = \cos^2(\theta) - \sin^2(\theta) + 2i \cos(\theta) \sin(\theta) \quad (30.1.59)$$

Equating real and imaginary parts, we have:

$$\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$$

$$(30.1.60a) \quad \sin(2\theta) = 2 \cos(\theta) \sin(\theta) \quad (30.1.60b)$$

The important thing is that we can now define the polar form of a complex number.

**Theorem 30.1.20.** *If  $z$  is a complex number, then there is a unique real number  $r \geq 0$  and a real number  $\theta \in [0, 2\pi)$  such that:*

$$z = r \exp(i\theta) \quad (30.1.61)$$

*Proof.* Let  $z = x + iy$ . Define  $r$  and  $\theta$  as:

$$r = \sqrt{x^2 + y^2} \quad (30.1.62a)$$

$$\theta = \begin{cases} \arctan\left(\frac{y}{x}\right), & x > 0, y \geq 0 \\ \frac{\pi}{2} + \arctan\left(\frac{y}{|x|}\right), & x < 0, y \geq 0 \\ \pi + \arctan\left(\frac{y}{x}\right), & x < 0, y \leq 0 \\ \frac{3\pi}{2} + \arctan\left(\frac{|y|}{x}\right), & x < 0, y \geq 0 \\ \frac{\pi}{2} \operatorname{sgn}(y), & x = 0 \end{cases} \quad (30.1.62b)$$

Here  $\operatorname{sgn}(y)$  is the sign of  $y$ . Euler's Theorem completes the proof. Uniqueness of  $r$  comes from the fact that  $|\exp(i\theta)| = 1$ , so if  $z = r_1 \exp(i\theta_1)$  and  $z = r_2 \exp(i\theta_2)$ , then  $|r_1| = |r_2|$ . But  $r_1$  and  $r_2$  are non-negative, and thus  $r_1 = r_2$ .  $\square$

Eqn. 30.1.61 is the definition of the polar form of a complex number. This gives geometrical interpretations of many aspects of complex arithmetic. Multiplication can be seen as rotations and scaling in the complex plane. For if  $z = r_1 \exp(i\theta_1)$  and if  $w = r_2 \exp(i\theta_2)$ , then we have:

$$z \cdot w = r_1 r_2 \exp(i(\theta_1 + \theta_2)) \quad (30.1.63)$$

That is, multiplying  $z$  by  $w$  scales  $z$  by the magnitude of  $w$ , and rotates it in the plane by the angle  $\theta_2$ . This also allows us to define square roots. We define the  $n^{th}$  root of a complex number to be:

$$\sqrt[n]{z} = \sqrt[n]{r} \exp\left(\frac{i\theta}{n}\right) \quad (30.1.64)$$

This is well defined for all complex numbers since the  $n^{th}$  root of a non-negative real number  $r$  is well defined, and  $\exp(i\theta/n)$  is well defined for all real  $\theta$ . Thus we have avoided the messiness of square roots that occurs in the real world. By  $\sqrt[n]{r}$ , we still mean the positive root. So  $\sqrt{4} = 2$ , and not -2.

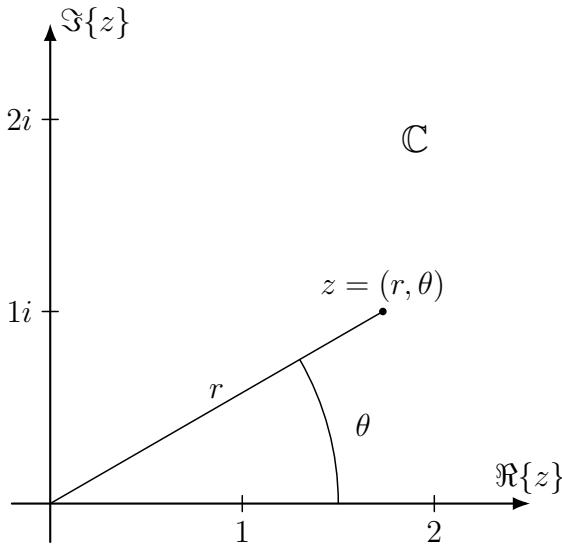


Fig. 30.4: Polar Representation of a Complex Number

### Example 30.1.3

Consider the square root of  $i$ . Using Euler's formula, we have that  $i = \exp(i\pi/2)$ . Using Eqn. 30.1.64, we obtain:

$$\sqrt{i} = \exp\left(\frac{i\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) + i \sin\left(\frac{\pi}{4}\right) = \frac{1+i}{\sqrt{2}} \quad (30.1.65)$$

We can check this solution by squaring:

$$\left(\frac{1+i}{\sqrt{2}}\right)^2 = \frac{(1-1)+i(1+1)}{2} = \frac{2i}{2} = i \quad (30.1.66)$$

in agreement with the definition of square roots. ■

We must be careful when evaluating square roots. We define the polar representation as  $z = r \exp(i\theta)$ , where  $0 \leq \theta < 2\pi$ . Problems can occur if we allow  $\theta$  to be any real number. For note that  $\exp(2\pi i) = 1 = \exp(0i)$ . Thus, we may

naively perform the following computation:

$$1 = \sqrt{1} = \exp(2\pi i/2) = \exp(\pi i) = -1 \quad (30.1.67)$$

The angle  $\theta \in [0, 2\pi)$  that we use to represent  $z$  is called the *principal value of the argument*, and is often denoted  $\text{Arg}(z)$ .

---

**Example 30.1.4**

Let  $f(z) = z^n - 1$ . This has a trivial root at  $z = 1$ , and by the Fundamental Theorem of Algebra there are at most  $n$  roots. The roots of this polynomial are called the *roots of unity*. The real solutions are 1 for odd  $n$ , and  $\pm 1$  for even  $n$ . In the complex world, there are always  $n$  solutions. Interestingly enough, these points form an  $n$ -gon around the origin of the complex plane. The solutions are:

$$z_k = \exp\left(\frac{2\pi ik}{n}\right) \quad k = 0, 1, 2, \dots, n-1 \quad (30.1.68)$$

Let's plot these solutions for various  $n$ .

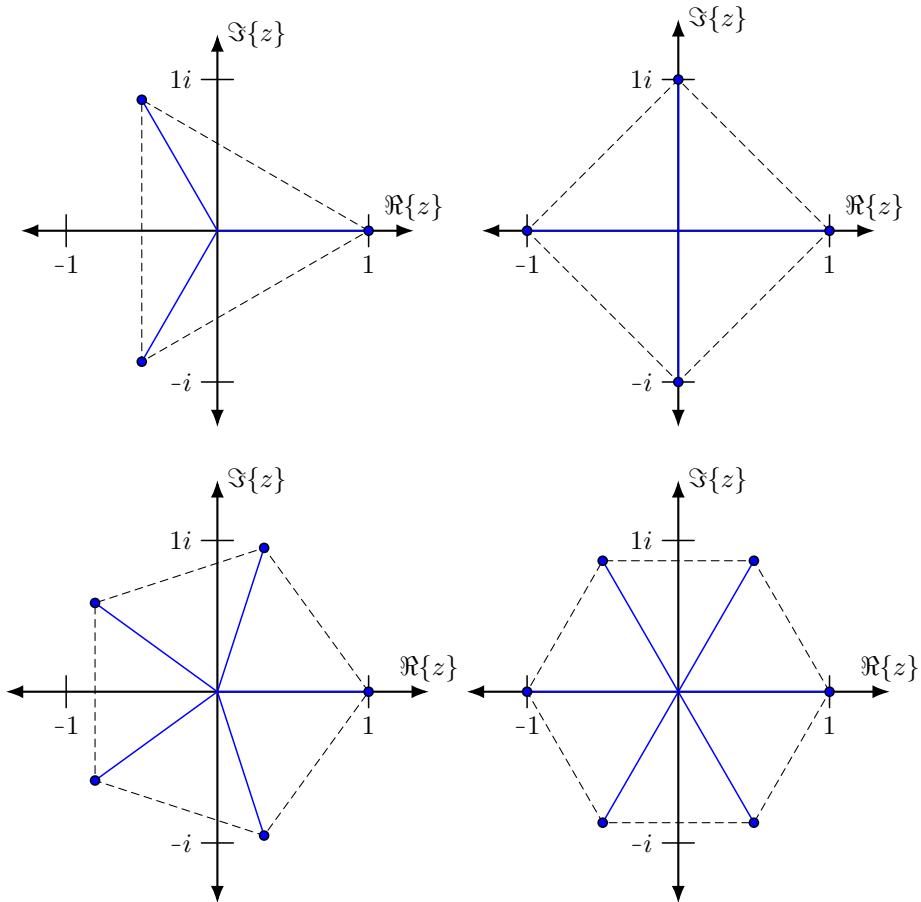


Fig. 30.5: Roots of Unity for Degrees 3 to 6.

It should be clear from the definition of  $f$  that the roots lie on the unit circle centered at the origin. While this is certainly an interesting and aesthetically appealing bit of mathematics, it also spells trouble for certain methods of numerical analysis. We'll return to this later when we discuss root finding algorithms. ■

### 30.1.2 Analytic Functions

We take a brief moment to talk about what it means to be analytic, the Cauchy-Riemann Equations, and Green's Theorem. The results here are counter-

intuitive, and it is easy to apply certain results where they do not hold.

### Definition 30.1.4: Entire Function

---

An entire function is a function  $f : \mathbb{C} \rightarrow \mathbb{C}$  such that for all  $z_0 \in \mathbb{C}$ , the following limit exists:

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0} \quad (30.1.69)$$

where  $f'$  is called the derivative of  $f$ . ■

An entire function is simply a complex function that is *differentiable* at every point in the complex plane. We can weaken this definition to include only some parts of the complex plane, and these are called *holomorphic* functions. A function  $f$  is analytic about the point  $z_0$  if its Taylor Series converges for all  $z$  to  $f(z)$  in some neighborhood of  $z_0$ :

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n \quad (30.1.70)$$

### Example 30.1.5

---

The exponential function is analytic, since we've defined it as a power series. It is indeed entire as well, since:

$$\lim_{z \rightarrow z_0} \frac{\exp(z) - \exp(z_0)}{z - z_0} = \exp(z_0) \lim_{z \rightarrow z_0} \frac{\exp(z - z_0) - 1}{z - z_0} \quad (30.1.71)$$

Letting  $w = z - z_0$ , we have:

$$\lim_{z \rightarrow z_0} \frac{\exp(z) - \exp(z_0)}{z - z_0} = \exp(z_0) \lim_{w \rightarrow 0} \frac{\exp(w) - 1}{w} \quad (30.1.72a)$$

$$= \exp(z_0) \lim_{w \rightarrow 0} \left( 1 + \sum_{n=2}^{\infty} \frac{w^{n-1}}{n!} \right) \quad (30.1.72b)$$

$$= \exp(z_0) \quad (30.1.72c)$$

This proves  $\exp$  is differentiable at every point  $z_0 \in \mathbb{C}$ , and is thus entire. ■

The remarkable fact of entire functions is that they are automatically analytic. This is certainly not true for real valued functions. One only need consider the example  $f(x) = x|x|$ . The derivative is  $f'(x) = 2|x|$ , and this has no derivative at the origin. Similarly, there are functions with two derivatives, but not three. In the real world, having  $n$  derivatives does not imply having  $n+1$  derivatives.

For complex functions, one derivative implies *all* higher derivatives exist.

### Example 30.1.6

It is often believed that *most* functions of a real variable are analytic, but the opposite is true. Real valued functions can be quite messy, and we need not construct overly pathological examples to show this. For consider the following:

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right), & x \neq 0 \\ 0, & x = 0 \end{cases} \quad (30.1.73)$$

This is a function that most students of calculus can understand, and is everywhere *smooth*: For all  $x_0 \in \mathbb{R}$ , and for all  $n \in \mathbb{N}$ , the  $n^{\text{th}}$  derivative  $f^{(n)}(x_0)$  exists.

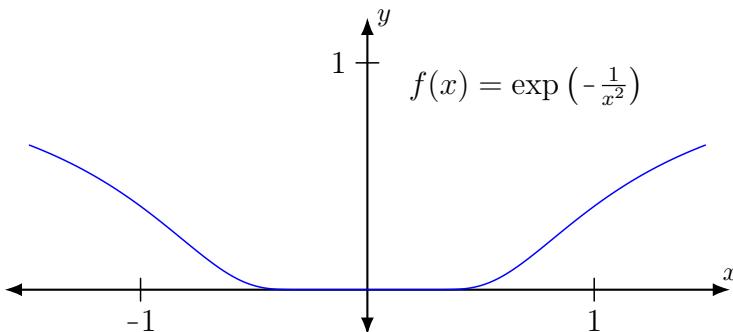


Fig. 30.6: A Smooth Function That is Not Analytic at the Origin

However, this function is not analytic at the origin. This function approaches zero so quickly at the origin that for all  $n \in \mathbb{N}$  we have:

$$\frac{d^n f}{dx^n}(0) = 0 \quad (30.1.74)$$

The Taylor expansion is thus zero, the radius of convergence is infinite, but  $f$  is not the zero function. Thus  $f$  is a function that is smooth but not analytic.

### Example 30.1.7

Further study of Ex. 30.1.6 reveals that  $f$  is analytic *everywhere else*, so one might expect smooth functions must be *somewhere* analytic, but this is false.

Consider:

$$F(x) = \sum_{k=0}^{\infty} \exp(-\sqrt{2^k}) \cos(2^k x) \quad (30.1.75)$$

Application of the  $M$  test from calculus shows that this sum converges, and that all of its derivatives exist. However, for all  $x \in \mathbb{R}$ , the Taylor series:

$$\sum_{n=0}^{\infty} F^{(n)}(x_0) \frac{(x - x_0)^n}{n!} \quad (30.1.76)$$

diverges for all  $x \neq x_0$ . So this function is smooth and *nowhere analytic*. ■

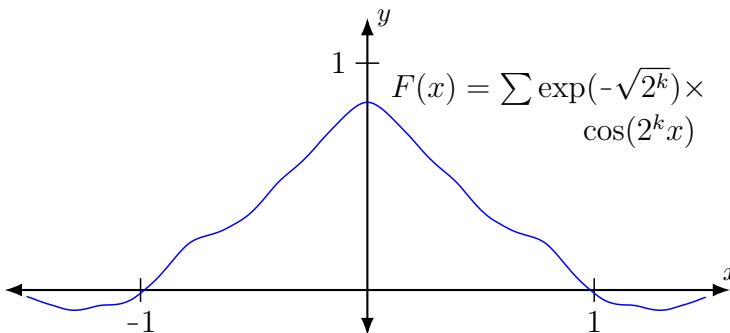


Fig. 30.7: A Smooth and Nowhere Analytic Function

The difficulties shown in Ex. 30.1.6 and Ex. 30.1.7 vanish when we study functions of a complex variable. Given a function  $f : \mathbb{C} \rightarrow \mathbb{C}$ , differentiable at  $z_0$  implies twice differentiable at  $z_0$ , which further implies smooth at  $z_0$ , and this implies analytic at  $z_0$ . There is one step that is missing here: Continuous does *not* imply differentiable. And, unfortunately, there are functions that *look* differentiable (Meaning the formula used to represent them would make us think at first glance that they are differentiable), but are not. Again, as we will see, we need not construct overly pathological examples to show this.

### Example 30.1.8

Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be defined by  $f(z) = \bar{z}$ . That is,  $f$  maps  $x + iy$  to  $x - iy$ . Then  $f$  is continuous at all  $z \in \mathbb{C}$ . For let  $\varepsilon > 0$  be given, and let  $\delta = \varepsilon/2$ . Then,

for all  $z_0$  such that  $|z - z_0| < \delta$ , we have:

$$|f(z) - f(z_0)| = |x - iy - (x_0 - iy_0)| \quad (30.1.77a)$$

$$= |(x - x_0) + i(y_0 - y)| \quad (30.1.77b)$$

$$\leq |x - x_0| + |y - y_0| \quad (30.1.77c)$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad (30.1.77d)$$

where we have applied the triangle inequality (Thm. 30.1.12) and Thm. 30.1.11 to derive these inequalities. Thus  $f$  is a continuous function. We will soon see that  $f$  is *nowhere* differentiable. ■

To reveal such functions we will need to present the *Cauchy-Riemann* equations. This set of equations provides both a *necessary* and a *sufficient* condition for a function  $f : \mathbb{C} \rightarrow \mathbb{C}$  to be entire. We will prove the easy part: Entire functions satisfy the Cauchy-Riemann equations.

**Theorem 30.1.21: Cauchy-Riemann Theorem**

If  $f : \mathbb{C} \rightarrow \mathbb{C}$  is an entire function defined by:

$$f(z) = u(x, y) + iv(x, y) \quad (30.1.78)$$

Then:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad (30.1.79a) \qquad \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \quad (30.1.79b)$$

*Proof.* As  $f$  is entire, its complex derivative exists at every point in the complex plane. Let  $z_0 = x_0 + iy_0$ , and let  $z = x + iy$ . Taking the limit, we have:

$$f'(z_0) = \lim_{x \rightarrow x_0} \frac{u(x, y_0) + iv(x, y_0) - u(x_0, y_0) - iv(x_0, y_0)}{x - x_0} \quad (30.1.80a)$$

$$= \lim_{x \rightarrow x_0} \frac{(u(x, y_0) - u(x_0, y_0)) + i(v(x, y_0) - iv(x_0, y_0))}{x - x_0} \quad (30.1.80b)$$

$$= \lim_{x \rightarrow x_0} \left( \frac{u(x, y_0) - u(x_0, y_0)}{x - x_0} \right) + i \lim_{x \rightarrow x_0} \left( \frac{v(x, y_0) - v(x_0, y_0)}{x - x_0} \right) \quad (30.1.80c)$$

Using the definition of *partial derivatives*, we obtain:

$$f'(z_0) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \quad (30.1.81)$$

Next we evaluate the limit along the path  $z = x_0 + iy$ . Since the function is complex differentiable, any path as  $z \rightarrow z_0$  will give the same value. Therefore:

$$f'(z_0) = \lim_{y \rightarrow y_0} \frac{u(x_0, y) + iv(x_0, y) - u(x_0, y_0) - iv(x_0, y_0)}{i(y - y_0)} \quad (30.1.82a)$$

$$= \lim_{y \rightarrow y_0} \frac{(u(x_0, y) - u(x_0, y_0)) + i(v(x_0, y) - iv(x_0, y_0))}{i(y - y_0)} \quad (30.1.82b)$$

$$= \frac{1}{i} \lim_{y \rightarrow y_0} \left( \frac{u(x_0, y) - u(x_0, y_0)}{i(y - y_0)} \right) + \lim_{y \rightarrow y_0} \left( \frac{v(x_0, y) - v(x_0, y_0)}{(y - y_0)} \right) \quad (30.1.82c)$$

Recalling our result from Thm. 30.1.14, the inverse of  $i$  is  $-i$ . Again using the definition of partial derivatives:

$$f'(z_0) = -i \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} \quad (30.1.83)$$

Thus, equating Eqn. 30.1.81 and Eqn. 30.1.83, we obtain:

$$\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y} \quad (30.1.84)$$

Comparing real and imaginary parts completes the proof.  $\square$

This theorem excludes many functions from being analytic.

### Example 30.1.9

Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be an analytic function, and suppose for all  $z \in \mathbb{C}$ ,  $f(z)$  is a *real* number. That is,  $f$  is a real-valued function. Then  $f$  must be a constant. For:

$$f(z) = u(x, y) + 0i \quad (30.1.85)$$

And from the Cauchy-Riemann equations, we have:

$$\frac{\partial u}{\partial x} = 0 \quad (30.1.86a) \qquad \frac{\partial u}{\partial y} = 0 \quad (30.1.86b)$$

And thus we conclude that  $u(x, y) = \text{const.}$  Given a non-constant real function  $f : \mathbb{R} \rightarrow \mathbb{R}$  that is analytic (Has a Taylor series), the complex extension  $F(x + iy) = f(x)$  is *not* analytic. Indeed, it is nowhere differentiable. Lastly, consider the function  $f(z) = z$ . This is indeed complex analytic. However:

$$\overline{f(z)} = x - iy \quad (30.1.87)$$

is *nowhere-analytic*. Indeed, it is nowhere differentiable. This is counterintuitive and reveals the bizarre nature of complex functions. It is worth recalling Ex. 30.1.8 where we showed that  $\overline{f}$  is continuous. Thus, we see that continuity does not imply differentiability, even for complex valued functions.  $\blacksquare$

We will return to this later when we discuss convolutions and the Hilbert transform. The Cauchy-Riemann equations seem to give some information for free. If we know  $f(z) = u(x, y) + iv(x, y)$  is analytic, and we know  $u(x, y)$ , then we can determine  $v(x, y)$ , up to an additive constant. In the theory of signal processing, given a real valued function  $u : \mathbb{R} \rightarrow \mathbb{R}$ , also called a *signal*, it will often be the case that we seek a real valued function  $v : \mathbb{R} \rightarrow \mathbb{R}$ , called the harmonic conjugate of  $u$ , such that  $u + iv$  is the boundary of some analytic function. Imposing certain criteria on  $u$  reveals that  $v$  is unique. Thus, given a complex signal where the imaginary part has been lost but the real part exists, we can recover the imaginary component by computing the harmonic conjugate of  $u$ . This is the *Hilbert Transform*, and we return to it in the section about Fourier Analysis.

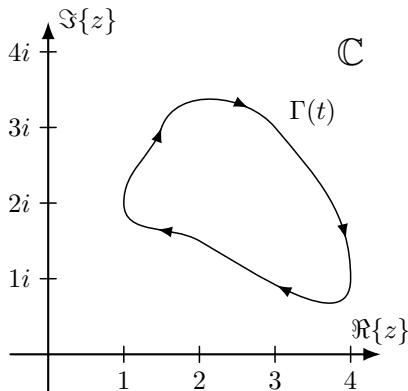
### 30.1.3 Contour Integrals

Next we introduce contour integrals. Throughout this section, integration is meant in the sense of the Riemann integral. We start by defining *Jordan Curves*.

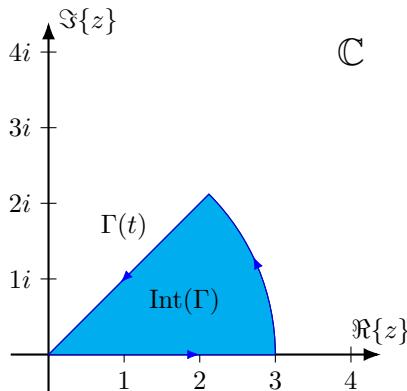
#### Definition 30.1.5: Jordan Curve

A Jordan Curve in the Complex Plane is a continuous function  $\Gamma : [0, 1] \rightarrow \mathbb{C}$  such that  $\Gamma(0) = \Gamma(1)$ , and there are no values  $0 < x_1 < x_2 < 1$  such that  $\Gamma(x_1) = \Gamma(x_2)$ . ■

A simple example of a Jordan curve is a circle. Jordan curves are *closed*, meaning they start where they end, and do not self-intersect. A Figure-8 is thus **not** a Jordan curve, but an ellipse is. An example of a Jordan curve is given below in Fig. 30.8.1. Much the way the closed unit interval  $[0, 1]$  has an ordering on it, a Jordan curve has a direction associated with it. Given a Jordan curve  $\Gamma(t)$ , one may change directions by defining  $\Gamma(t) = \Gamma(1 - t)$ . While this will plot out the same curve in the complex plane, the direction is different and thus it represents a different path. When evaluating contour integrals, the direction matters.



30.8.1: A Smooth Jordan Curve.



30.8.2: An Example of a Sector.

Fig. 30.8: Jordan Curves in the Complex Plane

For the sake of computation, we will stick to Jordan curves that are differentiable at all but finitely many points. A *sector*, which is the region contained within an arc of a circle, is an example of a Jordan curve that is differentiable at all but three points (Fig. 30.8.2). We prove Green's Theorem for such curves, particularly curves that can be broken into a *top* part and a *bottom* part. While we wish to avoid presenting theorems without proof, some results are too diffi-

cult to include. We state the *Jordan Curve Theorem*, but do not prove it. The proof can be found in a textbook on algebraic topology.

**Theorem 30.1.22.** *If  $\Gamma : \mathbb{R} \rightarrow \mathbb{R}^2$  is a Jordan curve, then  $\Gamma$  separates the plane in to two disjoint parts: The interior, denoted  $\text{Int}((\cup)\Gamma)$ , and the exterior. The interior is bounded, the exterior is unbounded, and  $\Gamma$  is their common boundary.*

A quick look at Fig. 30.8.1 can convince one of the validity of this statement. We use the fact that a Jordan curve has an interior to state Green's Theorem, which is useful for the evaluation of complex integrals. A student of electromagnetism will already understand the importance and usefulness of Green's Theorem. The Weak Green's Theorem applies to *simple* regions. There are two types of simple regions: Horizontal and Vertical.

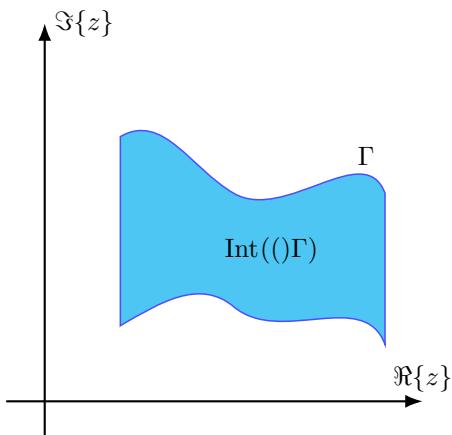
**Definition 30.1.6** A vertically simple region is a subset  $D$  of the plane  $\mathbb{R}^2$  such that there are two functions  $g_1, g_2 : [a, b] \rightarrow \mathbb{R}$  such that:

$$D = \{ (x, y) : a \leq x \leq b, g_1(x) \leq y \leq g_2(x) \} \quad (30.1.88)$$

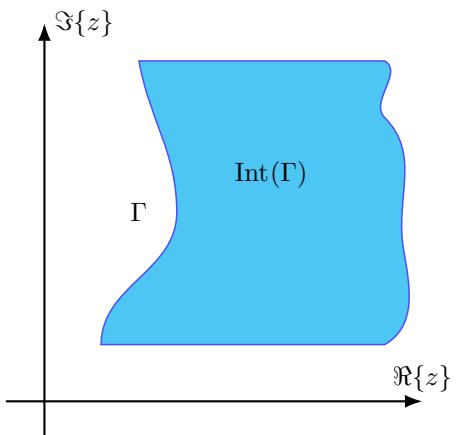
**Definition 30.1.7** A horizontally simple region is a subset  $D$  of the plane  $\mathbb{R}^2$  such that there are two functions  $g_1, g_2 : [a, b] \rightarrow \mathbb{R}$  such that:

$$D = \{ (x, y) : a \leq y \leq b, g_1(y) \leq x \leq g_2(y) \} \quad (30.1.89)$$

A vertically simple region is a subset of the plane bounded by two vertical lines, whereas a horizontally simple region is bounded by two horizontal lines (Fig. 30.9).



30.9.1: A Vertically Simple Region



30.9.2: A Horizontally Simple Region

Fig. 30.9: Examples of Simple Regions

We now prove special cases of Green's Theorem for vertically and horizontally simple regions, and then tie this together for the weak form of Green's Theorem.

**Theorem 30.1.23.** *If  $M : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a differentiable function and if  $\Gamma$  is a Jordan curve such that  $\text{Int}((\cup)\Gamma)$  is vertically simple, then:*

$$\iint_{\text{Int}((\cup)\Gamma)} \frac{\partial M}{\partial y} dA = - \oint_{\Gamma} M dx \quad (30.1.90)$$

*Proof.* Since the interior of  $\Gamma$  is a vertically simple region, there are two functions  $g_1, g_2 : [a, b] \rightarrow \mathbb{R}$  such that:

$$\Gamma = \{ (x, y) : a \leq x \leq b, g_1(x) \leq y \leq g_2(x) \} \quad (30.1.91)$$

But then:

$$\iint_{\text{Int}((\cup)\Gamma)} \frac{\partial M}{\partial y} dA = \int_a^b \int_{g_1(x)}^{g_2(x)} \frac{\partial M}{\partial y} dy dx \quad (30.1.92a)$$

$$= \int_a^b (M(x, g_2(x)) - M(x, g_1(x))) dx \quad (30.1.92b)$$

$$= - \int_a^b (M(x, g_1(x)) - M(x, g_2(x))) dx \quad (30.1.92c)$$

But since  $\text{Int}((\cup)\Gamma)$  is simple, the path at  $x = a$  is either a point or a vertical straight line. But then the integral along this portion with respect to  $x$  is zero. Similarly for  $x = b$ , and therefore:

$$\oint_{\Gamma} M dx = \int_a^b (M(x, g_1(x)) - M(x, g_2(x))) dx \quad (30.1.93)$$

Comparing Eqn. 30.1.92 and Eqn. 30.1.93 completes the proof.  $\square$

**Theorem 30.1.24.** *If  $N : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a differentiable function and if  $\Gamma$  is a Jordan curve such that  $\text{Int}((\cup)\Gamma)$  is horizontally simple, then:*

$$\iint_{\text{Int}((\cup)\Gamma)} \frac{\partial M}{\partial x} dA = \oint_{\Gamma} N dy \quad (30.1.94)$$

*Proof.* The proof is a mimicry of the proof for Thm. 30.1.23, but since the orientation of the path changes since we are now integrating with respect to  $y$ , we pick up a minus sign in the contour integral.  $\square$

**Theorem 30.1.25 (Weak Green's Theorem).** *If  $M : \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $N : \mathbb{R}^2 \rightarrow \mathbb{R}$  are differentiable functions, and if  $\Gamma$  is a Jordan curve such that the interior of  $\Gamma$  is vertically and horizontally simple (A rectangular region), then:*

$$\oint_{\Gamma} (M dx + N dy) = \iint_{\text{Int}((\cup)\Gamma)} \left( \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dA \quad (30.1.95)$$

*Proof.* Since  $\text{Int}((\cdot)\Gamma)$  is both vertically and horizontally simple, we may sum the results from Thm. 30.1.23 and Thm. 30.1.24, completing the proof.  $\square$

While this is not quite what we want, since most regions we wish to integrate over will not be simple, we can approximate the interior of a smooth Jordan curve arbitrarily well by a finite collection of simple regions. The full proof will get into the mechanics of this approximation, and show that in the *limit* we obtain the result. We'll state Green's Theorem, but neglect the full proof.

### Theorem 30.1.26: Green's Theorem

If  $M : \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $N : \mathbb{R}^2 \rightarrow \mathbb{R}$  are differentiable functions, and if  $\Gamma$  is a Jordan curve that is differentiable at all but finitely many points, then:

$$\oint_{\Gamma} (M \, dx + N \, dy) = \iint_{\text{Int}((\cdot)\Gamma)} \left( \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dA \quad (30.1.96)$$

■

We return to complex analysis and prove one of the central results of the theory: Cauchy's Integral Theorem.

**Theorem 30.1.27: Cauchy's Integral Theorem**

If  $f : \mathbb{C} \rightarrow \mathbb{C}$  is an entire function, and if  $\Gamma : [0, 1] \rightarrow \mathbb{C}$  is a Jordan curve differentiable at all but finitely many points, then:

$$\oint_{\Gamma} f(z) dz = 0 \quad (30.1.97)$$

*Proof.* For let  $f(z) = u(x, y) + iv(x, y)$ . Then:

$$\oint_{\Gamma} f(z) dz = \oint_{\Gamma} (u(x, y) + iv(x, y)) (dx + i dy) \quad (30.1.98a)$$

$$\begin{aligned} &= \oint_{\Gamma} (u(x, y) dx - v(x, y) dy) \\ &\quad + i \oint_{\Gamma} (v(x, y) dx + u(x, y) dy) \end{aligned} \quad (30.1.98b)$$

As  $f$  is entire,  $u$  and  $v$  are differentiable. Applying Green's Theorem, we obtain:

$$\oint_{\Gamma} f(z) dz = \iint_{\text{Int}(\Gamma)} \left( \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) dA + i \iint_{\text{Int}(\Gamma)} \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) dA \quad (30.1.99)$$

But since  $f$  is entire,  $u$  and  $v$  satisfy the Cauchy-Riemann equations. That is:

$$\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = 0 \quad (30.1.100a) \qquad \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = 0 \quad (30.1.100b)$$

Thus the integrands of both double integrals are zero, and hence the integrals are zero. This completes the proof.  $\square$

Finally we prove Jordan's Lemma. This is used in conjunction with Cauchy's Integral Theorem to provide a powerful means of computing the integrals of difficult functions. In particular, this is used to evaluate the limits of the *Fresnel Integrals*.

**Theorem 30.1.28** (Jordan's Inequality). *If  $x \in [0, \frac{\pi}{2}]$ , then:*

$$\frac{2}{\pi} x \leq \sin(x) \quad (30.1.101)$$

*Proof.* For let  $f : [0, \frac{\pi}{2}] \rightarrow \mathbb{R}$  be defined by  $f(x) = \frac{2}{\pi}x$ . Then  $f(0) = \sin(0)$  and  $f(\frac{\pi}{2}) = \sin(\frac{\pi}{2})$ . But since  $\sin$  is concave down on the interval  $[0, \frac{\pi}{2}]$ , it is

impossible for  $\sin(x) < f(x)$  on the open interval  $(0, \frac{\pi}{2})$ , and therefore we have that  $\sin(x) \geq f(x)$ . Therefore, etc.  $\square$

This simple theorem is best understood by graphing the two functions. We can use this to prove Jordan's Lemma, and this will conclude our discussion of complex analysis.

### Theorem 30.1.29: Jordan's Lemma

---

If  $g : \mathbb{C} \rightarrow \mathbb{C}$  is a continuous function, if  $\theta_0 \in [0, \pi]$ , if  $R$  and  $a$  are positive real numbers, and if  $\gamma_R$  is the arc from  $R$  to  $R\exp(i\theta_0)$ , then:

$$\left| \int_{\gamma_R} \exp(iaz)g(z) dz \right| \leq \frac{\pi}{a} M_R \quad (30.1.102)$$

Where  $M_R = \max \{ |g(z)| : z \in \gamma_R \}$ . ■

*Proof.* Applying the triangle inequality (Thm. 30.1.12) for integrals, Euler's Theorem (Thm. 30.1.18) and integrating in polar coordinates, we have:

$$\left| \int_{\gamma_R} \exp(iaz)g(z) dz \right| = \left| \int_{\gamma_R} \exp(iaz)g(z)iR \exp(i\theta) d\theta \right| \quad (30.1.103a)$$

$$\leq \int_{\gamma_R} |\exp(iaz)g(z)iR \exp(i\theta)| d\theta \quad (30.1.103b)$$

$$= R \int_{\gamma_R} |\exp(iaz)g(z)| d\theta \quad (30.1.103c)$$

$$= R \int_{\gamma_R} |\exp [iaR(\cos(\theta) + i \sin(\theta))] g(z)| d\theta \quad (30.1.103d)$$

$$= R \int_{\gamma_R} |\exp [aR(i \cos(\theta) - \sin(\theta))] g(z)| d\theta \quad (30.1.103e)$$

$$= R \int_{\gamma_R} |\exp (-aR \sin(\theta)) g(z)| d\theta \quad (30.1.103f)$$

$$\leq RM_R \int_{\gamma_R} \exp (-aR \sin(\theta)) d\theta \quad (30.1.103g)$$

Finally, applying Jordan's inequality (Thm. 30.1.28), we have:

$$\left| \int_{\gamma_R} \exp(iaz)g(z) dz \right| \leq RM_R \int_{\gamma_R} \exp\left(-\frac{2aR\theta}{\pi}\right) d\theta \quad (30.1.104a)$$

$$\leq \frac{\pi}{a} (1 - \exp(-aR)) M_R \quad (30.1.104b)$$

$$\leq \frac{\pi}{a} M_R \quad (30.1.104c)$$

Therefore, etc. □

## 30.2 Complex Variables

A complex function is a function whose argument is a complex variable  $z = x + iy$ , where  $i$  is the imaginary unit. Complex functions can have the problem of being multi-valued, which is a cause for caution when dealing with them. For example, in the complex realm every non-zero complex number  $z$  has two square roots  $\sqrt{z}$ . So the square root function is multi-valued. Any complex function  $f(z)$  can be written as  $f(z) = u(x, y) + iv(x, y)$ , where  $u$  and  $v$  are purely real functions. The function  $w = f(z)$  can be seen as a mapping, or transformation, of the  $z$  plane to the  $w$  plane. That is,  $f$  is a transformation of its domain onto its range, or image. A compound complex function is one of the form  $F(z) = g(f(z))$ . Since complex functions are functions of two variables, in a sense, one must be careful when considering limits of complex functions.

**Example 30.2.1** What is the limit of  $z/\bar{z}$  as  $z \rightarrow 0$ ? This is undefined. For:

$$\frac{z}{\bar{z}} = \frac{x + iy}{x - iy}$$

Letting  $x = 0$  and taking the limit on  $y$ , we get:

$$\frac{0 + iy}{0 - iy} = -1$$

Letting  $y = 0$  and taking the limit on  $x$ , we get:

$$\frac{x + 0i}{x - 0i} = 1$$

So the limit does not exist.

Continuity and the various properties of limits are defined similarly on  $\mathbb{C}$  as for  $\mathbb{R}$ , with distance between points being defined by  $d(z_1, z_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ . Differentiation is defined as:

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

**Theorem 30.2.1.** *A complex function  $f(z) = u(x, y) + iv(x, y)$  is differentiable if and only if it satisfies the Cauchy-Riemann equations:*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

**Theorem 30.2.2.** *If  $f(z) = u(x, y) + iv(x, y)$  is differentiable, then:*

$$f'(z) = u_x(x, y) + iv_y(x, y)$$

**Definition 30.2.1** A complex function  $f(z)$  is analytic, or holomorphic, at a point  $z_0$  if it is differentiable in some neighborhood of  $z_0$ .

**Definition 30.2.2** An entire function is a complex function  $f(z)$  such that  $f$  is analytic at every point  $z \in \mathbb{C}$ .

**Definition 30.2.3** A harmonic function is a function  $A(x, y)$  such that all of its second partial derivatives exists, and it satisfies the Laplace Equation:

$$\nabla^2 A = A_{xx}(x, y) + A_{yy}(x, y) = 0$$

**Theorem 30.2.3.** If  $f(z) = u(x, y) + iv(x, y)$  is differentiable on a domain  $D$ , then  $u$  and  $v$  are harmonic on the domain.

**Theorem 30.2.4.** A function  $f(z)$  is analytic if and only if its real and complex parts are harmonic conjugates of each other.

**Definition 30.2.4** A level curve of a function  $f(x, y)$  is a curve in  $\mathbb{R}^2$  such that  $f$  is constant on that curve.

One of the most basic and fundamental results from complex variables is Euler's Formula:

$$\exp(i\theta) = \cos(\theta) + i \sin(\theta)$$



Part XVII

# Calculus on Normed Spaces



# CHAPTER 31

---

## Calculus on Normed Spaces

---

31.1 Gateaux Derivative

31.2 Frechet Derivative

31.3 Malliavin Calculus



**Part XVIII**

# **Functional Analysis**



# CHAPTER 32

---

## Functional Analysis

---

### 32.1 Metric Spaces

#### 32.1.1 Basic Definitions

Functional analysis is concerned with normed spaces. This is a vector space  $V$  with a function, called a norm, from  $V$  to  $[0, \infty)$ . This is usually written  $\|\mathbf{x}\|$  for an element  $\mathbf{x} \in V$ . This norm must satisfy the following for all  $\mathbf{x}, \mathbf{y} \in V$ :

1.  $\|\mathbf{x}\| = 0$  if and only if  $\mathbf{x} = \mathbf{0}$  [Definiteness]
2.  $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$  for all  $c \in \mathbb{R}$ . [Positiveness]
3.  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$  [Triangle Inequality]

**Example 32.1.1**  $\mathbb{R}$  with  $\|\mathbf{x}\| = |x|$  and  $\mathbb{R}^n$  with  $\|\mathbf{x}\|_2$  defined by:

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{k=1}^n x_k^2}$$

are some of the most commonly used normed spaces.  $\mathbb{R}^n$  can be thought of as the set of vectors in  $n$  dimensions and the norm  $\|\mathbf{x}\|_2$  can be thought of as the length of  $\mathbf{x}$  using the Pythagorean Theorem. There are other norms one can define on  $\mathbb{R}^n$ . A common one is the  $p$  norm, defined by:

$$\|\mathbf{x}\|_p = \left( \sum_{k=1}^n x_k^p \right)^{1/p}$$

Together,  $(\mathbb{R}^n, \|\cdot\|_p)$  defines a normed space for all  $p \geq 1$ . Another type of norm on  $\mathbb{R}^n$  is the  $p$ -adic norm.

A common family of sets, which we will deal with frequently, are the  $\ell^p$  spaces.

**Definition 32.1.1**  $\ell^p$  is the set of all sequences  $x : \mathbb{R} \rightarrow \mathbb{R}$  such that the sequence of partial sums  $S : \mathbb{N} \rightarrow \mathbb{R}$  defined by  $S_N = \sum_{n=1}^N |x_n|^p$  is bounded.

**Definition 32.1.2** The  $p$  norm on  $\ell^q$  is the function  $\|\cdot\|_p : \ell^q \rightarrow \mathbb{R}$  defined by:

$$\|x\|_p = \left( \sum_{k=1}^{\infty} x_k^p \right)^{1/p}$$

**Theorem 32.1.1.** If  $p \geq 1$ , then  $(\ell^p, \|\cdot\|_p)$  is a normed space.

**Definition 32.1.3** The supremum norm on  $\mathbb{R}^n$ ,  $\|\cdot\|_\infty : \mathbb{R}^n \rightarrow \mathbb{R}$ , is the function:

$$\|\mathbf{x}\|_\infty = \max\{|x_1|, \dots, |x_n|\}$$

**Theorem 32.1.2.** If  $n \in \mathbb{N}$ , then  $(\mathbb{R}^n, \|\cdot\|_\infty)$  is a normed space.

**Definition 32.1.4**  $\ell^\infty$  is the set of sequences  $x : \mathbb{N} \rightarrow \mathbb{R}$  such that  $x$  is bounded.

**Definition 32.1.5** The supremum norm on  $\ell^\infty$  is the function  $\|\cdot\|_\infty : \ell^\infty \rightarrow \mathbb{R}$  defined by:

$$\|x\|_\infty = \sup\{|x_n| : n \in \mathbb{N}\}$$

**Theorem 32.1.3.** If  $\|\cdot\|_\infty$  is the supremum norm on  $\ell^\infty$ , then  $(\ell^\infty, \|\cdot\|_\infty)$  is a normed space.

From the fact that  $\ell^\infty$  is a normed space we have that the set of convergent sequences, again with the  $\|\cdot\|_\infty$  norm, is also a normed space. The set of null sequences, which is the set of sequences that converge to zero, is also a normed space. A stranger normed space is the set of all bounded continuous functions  $f : S \rightarrow \mathbb{R}$  with norm  $\|f\| = \sup\{|f(x)|\}$ . Furthermore, the set of all integrable functions with bounded integrals, with norm  $(\int_S |f|^p)^{1/p}$ . If you allow integral to mean Lebesgue Integrable, then this becomes a special space denoted  $L^p(S)$ .

**Definition 32.1.6** The Sobolev Space, denoted  $W^{n,p}([a, b])$  is the set of functions  $f : [a, b] \rightarrow \mathbb{R}$  such that:

$$\int_a^b \sum_{k=0}^n |f^{(k)}(x)|^p dx < \infty$$

**Definition 32.1.7** The  $p$  norm on the Sobolev space  $W^{n,q}([a, b])$  is the function  $\|\cdot\|_p : W^{n,q}([a, b]) \rightarrow \mathbb{R}$  defined by:

$$\left( \int_a^b \sum_{k=0}^n |f^{(k)}(x)|^p dx \right)^{1/p}$$

**Theorem 32.1.4.** If  $p \geq 1$ , then  $(W^{n,p}([a,b]), \|\cdot\|_p)$  is a normed space.

A lot of the things we wish to prove don't rely on the fact that all of these spaces are vector spaces. Really, we only care about the properties that the norm on the space has. What matters is that there's a set and a notion of distance on the set. This abstraction is the fundamental concept of a metric space.

**Definition 32.1.8** A metric space is a set  $S$  and a function  $d : S \times S \rightarrow [0, \infty)$  such that:

1. For all  $x, y \in S$ ,  $d(x, y) = 0$  if and only if  $x = y$ . [Definiteness]
2. For all  $x, y \in S$ ,  $d(x, y) = d(y, x)$  [Symmetry]
3. For all  $x, y, z \in S$ ,  $d(x, z) \leq d(x, y) + d(y, z)$  [Triangle Inequality]

It turns out that we can actually write the following:

**Definition 32.1.9** A metric space is a set  $X$  and a function  $d : X \times X \rightarrow \mathbb{R}$  such that:

1.  $d(x, y) = 0$  if and only if  $x = y$ .
2.  $d(x, z) \leq d(x, y) + d(z, y)$

By writing the triangle inequality in this way, symmetry comes for free (The fact that  $d(x, y) = d(y, x)$ ), as well as positivity (The fact that  $d(x, y) \geq 0$ ). Since it's easier to prove two things are true, rather than four things, it's nice to take this as the definition of a metric space, and then prove that the two definitions are equivalent. In a metric space  $(X, d)$ ,  $d$  is often called the *distance function* or *metric function*. It is meant to be an abstract mimicry of the absolute value function that is used with real numbers. Definiteness says the only point that is zero meters from a point  $x$  is  $x$  itself. Symmetry says the distance walking from  $x$  to  $y$  is the same as the distance walking from  $y$  to  $x$ . The last rule stems from Euclidean geometry. It says walking from  $x$  to  $z$  is shorter than (or equal to) walking from  $x$  to  $y$  and then  $y$  to  $z$ . In Euclidean geometry equality is achieved only when  $y$  lies between  $x$  and  $z$ . In abstract metric spaces there may be no such thing as a *line* between two points, so we need to be careful.

**Example 32.1.2**  $\mathbb{R}^n$  (for  $1 \leq p < \infty$ ):

$$d_p(\mathbf{x}, \mathbf{y}) = \left( \sum_{k=1}^n |x_k - y_k|^p \right)^{1/p} = \|\mathbf{x} - \mathbf{y}\|_p$$

**Example 32.1.3** In  $\ell^p$ , which are sequences for which  $\sum_{k=1}^{\infty} |x_k|^p < \infty$ ,  $d_p(x, y)$  forms a metric, as well as  $d_{\infty}(x, y) = \sup\{|x_k - y_k| : k \in \mathbb{N}\}$ , which is called the supremum norm.

**Example 32.1.4**  $C(S, \mathbb{R})$ , which is the set of continuous functions from  $S$  to  $\mathbb{R}$ , letting  $L^p(S)$  be the set of functions such that:

$$\int_S |x(t)|^p dt < \infty$$

Then the following is a metric:

$$d_p(x, y) = \left( \int_S |x(t) - y(t)|^p dt \right)^{1/p}$$

Also,  $d_\infty(x, y) = \sup\{|x(t) - y(t)|\}$ , which is called the supremum norm.

**Example 32.1.5** Let  $C$  be the set of sequences such that  $x_n \rightarrow 0$ . Then, with  $d_p$ , this forms a metric space. If  $C_0$  is set of sequences with only finitely many non-zero terms, then  $C_0 \subset C \subset \ell^\infty$ . Is there a sequence  $x \in C$  such that, for all  $1 \leq p < \infty$ ,  $x \notin \ell^p$ .

Since the image of the metric function lies in  $\mathbb{R}$ , we may speak of *convergence* in metric spaces.

**Definition 32.1.10** A convergent sequence in a metric space  $(X, d)$  is a sequence  $x : \mathbb{N} \rightarrow X$  such that there is an  $a \in X$  such that  $d(a, x_n) \rightarrow 0$ .

**Definition 32.1.11** A limit of a sequence  $x$  in a metric space  $(X, d)$  is an  $a \in X$  such that  $d(x_n, a) \rightarrow 0$ .

Much like convergence in real numbers, limits in metric spaces are unique.

**Theorem 32.1.5.** *If  $(X, d)$  is a metric space,  $x : \mathbb{N} \rightarrow X$  is a convergence sequence in  $X$ , and if  $a$  and  $b$  are limits of  $x$ , then  $a = b$ .*

*Proof.* For suppose not. As  $(X, d)$  is a metric space,  $d(a, b) > 0$ . Let  $\varepsilon = \frac{d(a, b)}{4}$ . Then, as  $d(a, x_n) \rightarrow 0$  and  $\varepsilon > 0$ , there is an  $N_1 \in \mathbb{N}$  such that, for all  $n > N_1$ ,  $d(a, x_n) < \varepsilon$ . But, as  $d(b, x_n) \rightarrow 0$  and  $\varepsilon > 0$ , there is an  $N_2$  such that, for all  $n > N_2$ ,  $d(b, x_n) < \varepsilon$ . Let  $n = \max\{N_1, N_2\} + 1$ . But then:

$$d(a, b) \leq d(a, x_n) + d(b, x_n) < 2\varepsilon = \frac{d(a, b)}{2}$$

A contradiction. Therefore,  $a$  is unique.  $\square$

**Theorem 32.1.6.** *If  $(X, d)$  be a metric space and if  $x, y, z \in X$ , then  $|d(x, z) - d(y, z)| \leq d(x, y)$*

*Proof.* Suppose  $d(x, z) \geq d(y, z)$ . If  $d(x, z) < d(y, z)$ , the proof is symmetric. Thus we have:

$$|d(x, z) - d(y, z)| = d(x, z) - d(y, z) \leq (d(x, y) + d(y, z)) - d(y, z) = d(x, y)$$

Therefore,  $|d(x, z) - d(y, z)| \leq d(x, y)$ .  $\square$

**Theorem 32.1.7.** If  $(X, d)$  is a metric space and  $x_n \rightarrow a$ , then for all  $b \in X$ ,  $d(x_n, b) \rightarrow d(a, b)$ .

*Proof.* For  $|d(x_n, b) - d(a, b)| \leq d(x_n, a) \rightarrow 0$ .  $\square$

**Theorem 32.1.8.** If  $(V, \|\cdot\|)$  is a normed space and if  $d : V \times V \rightarrow [0, \infty)$  is defined by  $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ , then  $(V, d)$  is a metric space.

*Proof.* In order:

1. If  $\|\mathbf{x} - \mathbf{y}\| = 0$ , then  $\mathbf{x} = \mathbf{y}$ . Similarly,  $\|\mathbf{x} - \mathbf{x}\| = \|\mathbf{0}\| = 0$ .
2.  $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \|(-1)(\mathbf{y} - \mathbf{x})\| = |-1|\|\mathbf{y} - \mathbf{x}\| = \|\mathbf{y} - \mathbf{x}\| = d(\mathbf{y}, \mathbf{x})$
3. The triangle inequality follows from the triangle inequality that norms have.

$\square$

There are metric spaces that have nothing to do with vector spaces or norms. Metric spaces are a more abstract object. Every normed space has an associated metric space since there is the “induced” metric.

**Example 32.1.6** Let  $X$  be a set and let  $d(x, y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases}$ . This is the discrete metric on  $X$ .

**Example 32.1.7** Let  $X = \{a, b, c\}$ , and  $d(a, b) = 1$ ,  $d(b, c) = 2$ . What value must  $d(a, c)$  have if  $d$  is a metric on  $X$ ? Consider the following table:

$X$	a	b	c
a	0	1	?
b	1	0	2
c	?	2	0

This obeys everything except the triangle inequality. We must pick  $d(a, c)$  such that this is upheld. So we need the following:

$$\begin{aligned} d(a, b) &\leq d(a, c) + d(c, b) & d(a, c) &\leq d(a, b) + d(b, c) & d(b, c) &\leq d(b, a) + d(a, c) \\ \Rightarrow 1 &\leq 2 + d(a, c) & \Rightarrow d(a, c) &\leq 3 & \Rightarrow 2 &\leq 1 + d(a, c) \end{aligned}$$

So we need  $1 \leq d(a, c) \leq 3$ . Pick  $d(a, c) = 2$ . This makes  $(X, d)$  a metric space.

**Example 32.1.8** Let  $X = \mathbb{R}$  and  $d(x, y) = |x - y|$ . Then  $(X, d)$  is a metric space.

**Example 32.1.9**  $\mathbb{R}$  with  $d(x, y) = |f(x) - f(y)|$ , where  $f : \mathbb{R} \rightarrow \mathbb{R}$  is injective, is a metric space. Let  $f$  be a real-valued function. Then from the triangle inequality

$$|f(x) - f(y)| \leq |f(x) - f(z)| + |f(z) - f(y)|$$

Therefore  $d$  obeys the triangle inequality. It also obeys symmetry, for:

$$|f(x) - f(y)| = |(-1)(f(y) - f(x))| = |f(y) - f(x)|$$

The absolute value function is doing most of the work. But finally we require that  $|f(x) - f(y)| = 0$  if and only if  $x = y$ . But  $|f(x) - f(y)| = 0$  if and only if  $f(x) = f(y)$ . So we require that  $f$  is injective. If  $f$  is not injective, then there exists  $x_1, x_2$  such that  $x_1 \neq x_2$  and yet  $f(x_1) = f(x_2)$ . But then  $|f(x_1) - f(x_2)| = 0$ , contradicting the fact that this is a metric. If  $f$  is injective, then this is a metric. Note injective functions need not be continuous, and can be very crazy.

**Example 32.1.10**  $\mathbb{R}$  with  $d(x, y) = |\tan^{-1}(x) - \tan^{-1}(y)|$  is a metric. Moreover,  $d(x, y) < \pi$  for all  $x, y \in \mathbb{R}$ . Thus, we have found a metric that makes  $\mathbb{R}$  a bounded set. As a fun fact,  $x_n = n$  is a Cauchy sequence in this metric space, but this sequence does not converge to anything. Thus we've found a metric on  $\mathbb{R}$  such that  $(\mathbb{R}, d)$  is not complete.

**Example 32.1.11** Can  $d(x, y) = f(x - y)$  be a metric on  $\mathbb{R}$  if  $f$  is differentiable? Not everywhere.  $f$  can not be differentiable at the origin for  $d(x, y) = f(x - y)$  to be a metric function, however  $f$  can be differentiable everywhere else. Use  $f(x) = |x|$  as an example. If  $f(x - y)$  is a metric,  $f$  must be an even function. But then  $f'(0) = 0$ . But  $f(x - y)$  also must obey the triangle inequality. Therefore:

$$f(2x) \leq f(x) + f(x) = 2f(x)$$

Define  $h(x)$  by:

$$h(x) = \begin{cases} \frac{f(x)}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Then, from the previous statement,  $h(2x) \leq h(x)$ . But then:

$$h\left(\frac{1}{2^n}\right) \leq h\left(\frac{1}{2^{n+1}}\right)$$

But from L'Hôpital's Rule,  $h(x) \rightarrow f'(0)$  as  $x \rightarrow 0$ . Therefore  $h(1) \leq f'(0)$ . But  $h(1) > 0$  since  $f(x - y)$  is a metric, a contradiction. Therefore,  $f$  can not be differentiable at the origin.

### 32.1.2 Topology

**Definition 32.1.12** The open ball of radius  $r > 0$  about a point  $x$  in a metric space  $(X, d)$  is the set  $B_r(x) = \{y \in X : d(x, y) < r\}$

The picture for this is a “circle” around the point  $x$  or radius  $r$ . However, this circle can look very strange for weird metrics.

**Example 32.1.12** If  $X$  is a set and  $d$  is the discrete metric, then  $B_r(x)$  is either the point  $x$  (If  $r \leq 1$ ), or it is the entire set  $X$ .

**Example 32.1.13** With  $X = \mathbb{R}$  and  $d$  the standard metric  $d(x, y) = |x - y|$ , we have  $B_r(x)$  is simply the open interval  $(x - r, x + r)$ .

**Example 32.1.14** Let  $X = \mathbb{R}^2$  and define  $d_p(x, y) = (|x_1 - y_1|^p + |x_2 - y_2|^p)^{1/p}$ . For  $p = 2$ , an open ball is a circle around the point  $(x, y)$  of radius  $r$ . For  $p = 1$ , we have “diamonds” around the point  $x$ . And for  $p = \infty$  we have a square around  $x$ . Let  $X = \mathbb{R}^2$  and let  $d$  be the metric such that you can only travel parallel to the  $y$  axis, or along the  $x$  axis. Consider the unit balls in  $(X, d)$  about the following points:

- a.  $(0, 0)$
- b.  $(0, 1)$
- c.  $(0, \frac{1}{2})$
- d.  $(\frac{1}{2}, \frac{1}{2})$

If  $\mathbf{x}_1 = (x_1, y_1)$  and  $\mathbf{x}_2 = (x_2, y_2)$ , then we have:

$$d(\mathbf{x}_1, \mathbf{x}_2) = \begin{cases} |y_2 - y_1|, & x_1 = x_2 \\ |x_2 - x_1| + |y_1| + |y_2|, & x_1 \neq x_2 \end{cases}$$

About the point  $(0, 0)$ , the unit ball is simply points  $(x, y)$  such that  $|x| + |y| < 1$ . This is a “diamond.” About  $(0, 1)$ , first note that to get to any point whose  $x$  coordinate is not 0, you first must travel the entirety of the  $y$  axis. Since this length is already 1, you can’t go left or right on the  $x$  axis. The unit ball is the line segment on the  $y$  axis between  $(0, 0)$  and  $(0, 2)$ . For the third one, if the  $x$  coordinate changes, we have  $0.5 + |y| + |x| < 1$ , which implies  $|y| + |x| < 0.5$ . This is again a diamond, but a smaller one. If the  $x$  coordinate does not change, we have  $|y - 0.5| < 1$ . This is another line segment. Repeat the same arguments for the fourth coordinate. The diagrams are show in Fig. 32.1.

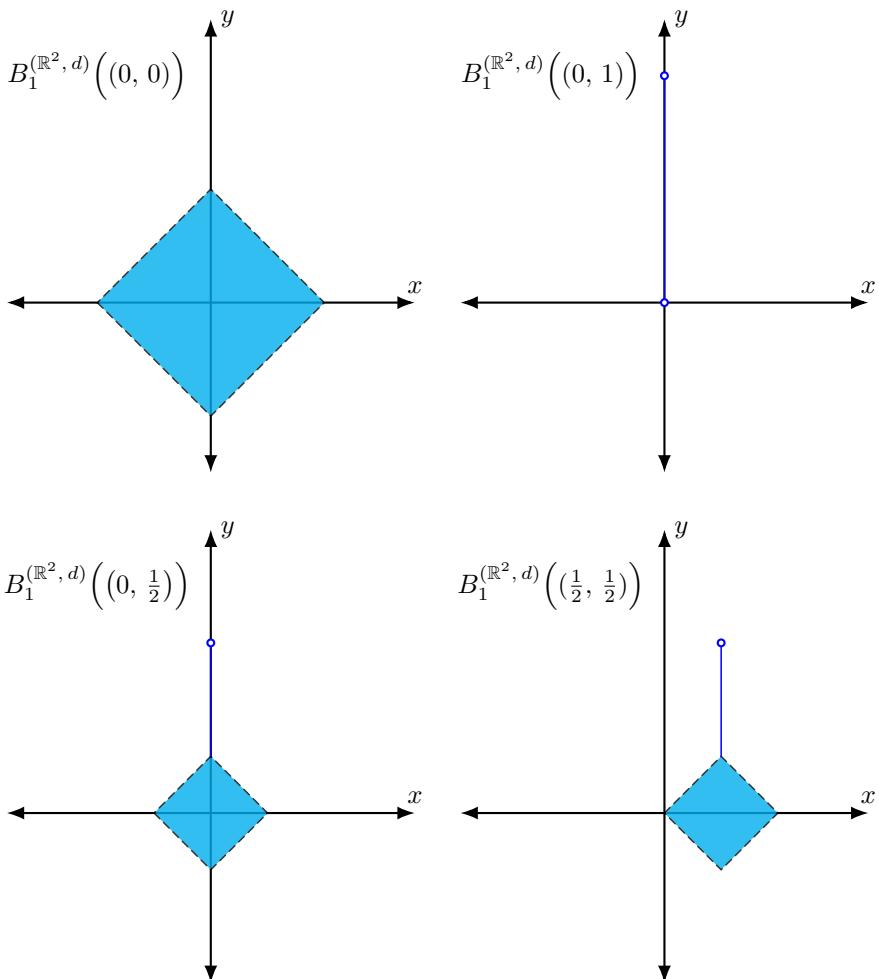


Fig. 32.1: Figures for Example ??.

If you have a vector space and a norm on it, then the open balls about a point will have the property of convexity. Convexity is a vector space property, given two points the “line” between the two remains in the set. Metric spaces have no such notion. Since the balls of  $\|\cdot\|_p$  are not convex with  $p < 1$ , we have that  $\|\cdot\|_p$  is a metric on  $\mathbb{R}^n$  if and only if  $p \geq 1$ .

**Definition 32.1.13** An open subset of a metric space  $(X, d)$  is a set  $S \subset X$  such that, for all  $x \in S$ , there is an  $r > 0$  such that  $B_r(x) \subset S$ .

**Example 32.1.15** If  $(X, d)$  is a metric space, then  $X$  is open and  $\emptyset$  is open (Vacuously true).

**Theorem 32.1.9.** *If  $(X, d)$  is a metric space,  $x \in X$ , and  $r > 0$ , then  $B_r(x)$  is an open subset of  $X$ .*

*Proof.* If  $z \in B_r(x)$ , let  $t = d(x, z)$ . Then  $0 \leq t < r$ . Let  $r' = r - t$ . But if  $y \in B_{r'}(z)$ , then  $d(x, y) \leq d(x, z) + d(y, z) < t + r' = t + r - t = r$ . Therefore  $B_{r'}(z) \subset B_r(x)$ .  $\square$

**Theorem 32.1.10.** *A finite intersection of open sets is open.*

*Proof.* If  $\mathcal{U}_1, \dots, \mathcal{U}_n$  are open and if  $x \in \cap_{k=1}^n \mathcal{U}_k$ , then there exists  $r_1, \dots, r_n$  such that  $B_{r_i}(x) \subset \mathcal{U}_i$ . Let  $r = \min\{r_1, \dots, r_n\}$ . Then  $B_r(x) \subset \cap_{k=1}^n \mathcal{U}_i$ .  $\square$

**Theorem 32.1.11.** *Arbitrary unions of open sets are open.*

Infinite intersections need not be open. The proof above would fail since the  $r_i$  can form a sequence tending to zero. But indeed, let  $X = \mathbb{R}$  and let  $d(x, y) = |x - y|$ , and take  $\mathcal{U}_n = (-\frac{1}{n}, \frac{1}{n})$ . Then all of the  $\mathcal{U}_n$  are open, yet the intersection, which is the set  $\{0\}$ , is not open. All of this mumbo-jumbo creates the more general notion of a topological space.

**Definition 32.1.14** A topological space is a set  $X$  and a subset  $\tau \subset \mathcal{P}(X)$  such that:

1.  $\emptyset, X \in \tau$
2. Finite intersections of sets in  $\tau$  are also sets in  $\tau$ .
3. Arbitrary unions of sets in  $\tau$  are also sets in  $\tau$ .

Here,  $\mathcal{P}(X)$  denotes the *power set* of  $X$ . This is the set of all subsets of  $X$ . The notion of a topological space generalizes the notion of a metric space. There is no notion of distance in such spaces, and things can be weird. There are topological spaces that have no metric associated with them.

**Definition 32.1.15** An open subset of a topological space  $(X, \tau)$  is a set  $\mathcal{U} \in \tau$ .

**Definition 32.1.16** An interior point of a subset  $S$  of a topological space  $(X, \tau)$  is a point  $x \in S$  such that there is an open subset  $\mathcal{U} \subseteq S$  such that  $x \in \mathcal{U}$ .

**Definition 32.1.17** The interior of a subset  $S$  of a topological space  $(X, \tau)$ , denoted  $\text{Int}((S))$ , is the set of all interior points of  $S$ .

**Theorem 32.1.12.** *If  $S$  is an open subset of  $(X, \tau)$ , then  $\text{Int}((S)) = S$ .*

**Definition 32.1.18** A function from a metric space  $(X, d_X)$  to a metric space  $(Y, d_Y)$  continuous at  $x \in X$  is a function  $f : X \rightarrow Y$  such that for all  $\varepsilon > 0$  there is a  $\delta > 0$  such that for all  $x_0 \in X$  such that  $d_X(x, x_0) < \delta$ , we have  $d_Y(f(x), f(x_0)) < \varepsilon$

**Theorem 32.1.13.** *If  $(X, \tau)$  is a topological space and  $S \subseteq X$ , and if  $\mathcal{O}$  is the set of all open sets  $\mathcal{U}$  such that  $\mathcal{U} \subseteq S$ , then  $\text{Int}((\cup)S) = \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}$ .*

**Definition 32.1.19** A nowhere dense subset of a topological space  $(X, \tau)$  is a subset  $S \subseteq X$  such that  $\text{Int}((\cup)S) = \emptyset$ .

**Theorem 32.1.14.** *If  $(X, d)$  is a metric space,  $y \in X$ , then  $f : X \rightarrow \mathbb{R}$  defined by  $f(x) = d(x, y)$  is uniformly continuous.*

A surprising theorem, and the entire basis of the study of topology, goes as follows:

**Theorem 32.1.15.** *If  $(X, d_x)$  and  $(Y, d_Y)$  are metric spaces, then  $f : X \rightarrow Y$  is continuous at  $x \in X$  if and only if for all open subsets of  $S \subset Y$  such that  $f(x) \in S$ ,  $f^{-1}(S)$  is an open subset of  $X$ .*

This allows us to talk about continuous functions without a notion of metric. Thus, for topological spaces, this is the *definition* of continuity. When the space we're discussing is a metric space, this theorem shows that the definition from topology and the defintition from real analysis are in fact equivalent.

**Theorem 32.1.16.** *A function  $f : X \rightarrow Y$  between metric spaces is continuous at a point  $x \in X$  if and only if for all sequences  $x_n$  such that  $d_X(x, x_n) \rightarrow 0$ , we have  $d_Y(f(x), f(x_n)) \rightarrow 0$ .*

We now have three different ways to talk about continuity. Topological spaces can be nastier, however. We saw in Thm. 32.1.5 that the limit of a convergent sequence in a metric space is unique. This is not true in a topological space and there are topological spaces with sequences which converge to every point in the space simultaneously. Indeed, it may be impossible to distinguish two points in a topological space. The ability to “Separate,” points is special. Hausdorff spaces can, but we'll save that for topology.

## Closed Sets

**Definition 32.1.20** A limit point of a subset  $S \subset X$  of a metric space  $(X, d)$  is a point  $a \in X$  such that there is a sequence  $x : \mathbb{N} \rightarrow S$  such that  $d(a, x_n) \rightarrow 0$ .

**Definition 32.1.21** A closed subset of a metric space  $(X, d)$  is a set  $S$  such that for all  $x \in X$  such that  $x$  is a limit point of  $S$ ,  $x \in S$ .

This says that if  $S$  is closed, and  $x$  is a sequence in  $S$  such that  $x_n \rightarrow a$ , then  $a \in S$ .

**Example 32.1.16** In  $\mathbb{R}$ , with the standard metric,  $(a, b)$  is open,  $\mathbb{R}$  is open (and closed),  $[a, b]$  is closed,  $[a, \infty)$  is closed,  $[a, b)$  is neither closed nor open.

**Example 32.1.17** If  $X = (0, 1)$ , and  $d(x, y) = |x - y|$ , then  $(0, 1)$  is closed. This is because there is no sequence that converges to a point in the space whose limit is not in the space. There are no sequences in  $X$  which converge to zero or one since, as far as  $X$  is concerned, neither or these points exist.

**Theorem 32.1.17.** *If  $(X, d)$  is a metric space, then a subset  $S \subset X$  is open if and only if  $X \setminus S$  is closed.*

*Proof.* Suppose  $S$  is open, and let  $x_n$  be a sequence in  $S^c$ . Suppose  $x_n \rightarrow x$  and  $x \in S$ . But  $S$  is open, and thus there is an  $\varepsilon > 0$  such that  $B_\varepsilon(x) \subset S$ . But  $x_n \rightarrow x$ , and thus this is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $d(x, x_n) < \varepsilon$ . But then for all  $n > N$ ,  $x_n \in B_\varepsilon(x)$ . But  $x_n \in S^c$ , a contradiction. Therefore,  $S^c$  is closed. On the other hand, if  $S^c$  is closed and there is an  $x \in S$  such that for all  $r > 0$ ,  $B_r(x) \cap S \neq \emptyset$ , then for all  $n \in \mathbb{N}$  there is an  $x_n \in S^c$  such that  $d(x, x_n) < \frac{1}{n}$ . But then  $x_n \rightarrow x$ , and therefore  $x \in S^c$ . But  $x \in S$ , a contradiction. Thus,  $S$  is open.  $\square$

In topology we take the definition of closed sets to be the compliment of open sets. This theorem shows that the topological definition is equivalent when we consider metric spaces.

**Definition 32.1.22** The closure of a subset  $S$  of a metric space  $(X, d)$ , denoted  $\overline{S}$ , is the set of all limit points of  $S$ .

**Theorem 32.1.18.** *If  $(X, d)$  is a metric space, if  $S \subset X$ , and if  $\Delta$  is the set of all closed subsets  $\mathcal{C} \subset X$  such that  $S \subset \mathcal{C}$ , then:  $\overline{S} = \bigcap_{\mathcal{C} \in \Delta} \mathcal{C}$*

Thus we may loosely say that the closure of a set  $S$  is the “Smallest,” closed set that contains  $S$ .

**Definition 32.1.23** The closed ball of radius  $r > 0$  about a point  $x$  in a metric space  $(X, d)$  is the set:

$$\overline{B}_r(x) = \{y \in X : d(x, y) \leq r\}$$

There exists metric spaces  $(X, d)$  such that  $\overline{B}_r(x) \neq \overline{B_r(x)}$ . For take the discrete metric,  $r = 1$ . Then the closure of  $B_1(x)$  is simply the point  $x$ . However, the closed ball  $\overline{B}_1(x)$  is the entire space. Metric spaces can be very weird like this. They have a property, that given a nested sequence of closed balls whose radius tends to zero, there is precisely one point that lies in the intersection. However, if the radius does not tend to zero it is possible that the intersection is empty. This is very counter-intuitive.

**Definition 32.1.24** A dense subset of a metric space  $(X, d)$  is a set  $S \subset X$  such that  $\overline{S} = X$ .

A subset  $S$  is dense in  $X$  if every point in  $X$  can be approximated arbitrarily well by points in  $S$ . For any point  $a \in X$  there is a sequence  $x \in S$  such that  $x_n \rightarrow a$ . The classic example is  $\mathbb{Q}$  and  $\mathbb{R}$ . Every real number can be approximated arbitrary well by a rational number. To see this, just take the continued fraction of a real number and stop once the approximation is less than  $\varepsilon$ . When we say  $\mathbb{Q}$  is dense in  $\mathbb{R}$ , we of course mean with respect to the standard metric on  $\mathbb{R}$ .  $\mathbb{Q}$  is **not** dense in  $\mathbb{R}$  with respect to the discrete metric. Indeed, if  $d$  is the discrete metric on  $X$ , then  $S \subset X$  is dense in  $X$  if and only if  $S = X$ .

**Example 32.1.18**  $\mathbb{Q}$  is dense in  $\mathbb{R}$  with respect to  $d_p$  for all  $p \geq 1$ . This includes  $d(x, y) = |x - y|$ .

**Example 32.1.19** The set of polynomials on the interval  $[a, b]$  are dense in the set of continuous functions on  $[a, b]$  with respect to the  $d_\infty$  metric. This comes from Weierstrass's Theorem.

**Example 32.1.20** The set of polynomials on  $[a, b]$  is dense in the set of continuous functions on  $[a, b]$  with respect to the  $d_p$  metric, for  $p \geq 1$ . This is because:

$$\begin{aligned} d_p(P, x) &= \left( \int_a^b |P(t) - x(t)|^p dt \right)^{1/p} &= \left( d_\infty(P, x)^p \int_a^b dt \right)^{1/p} \\ &\leq \left( \int_a^b \max\{|P(t) - x(t)|\}^p dt \right)^{1/p} &= (b-a)^{1/p} d_\infty(P, x) \end{aligned}$$

**Example 32.1.21** The continuous functions are not dense in the set of integrable functions, with respect to the supremum metric  $d_\infty$ . This is more or less because integrable functions can be discontinuous, or have jumps. This means, with respect to  $d_\infty$ , that no continuous functions could approximate such a discontinuous function arbitrary well.

**Definition 32.1.25** A separable metric space is a metric space  $(X, d)$  with a countable dense subset  $S$ .

**Example 32.1.22**  $\mathbb{R}$  is separable, with the standard metric, since  $\mathbb{Q}$  is countable and also dense in  $\mathbb{R}$ .

**Example 32.1.23** The set of continuous functions on  $[a, b]$  is separable. For take the set of polynomials with rational coefficients. This can be seen as a countable union of countably many elements. For let  $P_N$  be the set of polynomials of degree  $N$  with rational coefficients. This is countable, and the set of all polynomials with rational coefficients is simply the union of  $P_N$  over all  $N$ . This is dense in the set of polynomials, and the set of polynomials is dense in  $C[a, b]$ , and thus the set of polynomials with rational coefficients is dense in  $C[a, b]$ . Thus  $C[a, b]$  is separable.

**Example 32.1.24**  $\ell^p$  is separable with the  $d_p$  metric, simply use elements with rational entries. That is, sequences of rational numbers.

**Example 32.1.25**  $\ell^p$  with the  $d_\infty$  metric is NOT separable. Consider the real numbers in  $(0, 1)$ .

### 32.1.3 Completeness

**Definition 32.1.26** A complete metric space is a metric space  $(X, d)$  such that every Cauchy sequence  $x_n$  in  $X$  converges to a point in  $X$  with respect to  $d$ .

Recall that a sequence  $x_n$  is Cauchy if  $\forall \varepsilon > 0 \exists N \in \mathbb{N} : \forall n, m > N, d(x_n, x_m) < \varepsilon$ . Convergence with respect to  $d$  means that  $d(x, x_n) \rightarrow 0$ .

**Example 32.1.26**  $\mathbb{R}$  with the standard metric  $d(x, y) = |x - y|$  is complete.

**Example 32.1.27**  $(\mathbb{R}^n, d_p)$  is also complete for all  $n \in \mathbb{N}$ .

Completeness is both a property of the set and the metric itself. It is not a topological property.

**Example 32.1.28**  $(\mathbb{R}, d)$ , where  $d(x, y) = |\tan^{-1}(x) - \tan^{-1}(y)|$  is *not* complete. For let  $x_n = n$ . This is a Cauchy sequence, as one can see from the graph of  $\tan^{-1}(x)$ . That is, because  $\tan^{-1}(x) \rightarrow \pi/2$ ,  $x_n = n$  is a Cauchy sequence in this metric. Being even more rigorous, let  $\varepsilon > 0$  and  $N = \lceil \tan(\pi/2 - \varepsilon) \rceil$ . Then, for all  $n, m > N$ ,  $d(x_n, x_m) = |\tan^{-1}(n) - \tan^{-1}(m)| < |\pi/2 - \tan^{-1}(\min\{n, m\})| < |\pi/2 - (\pi/2 - \varepsilon)| = \varepsilon$ . But  $x_n$  does not converge. For suppose not., Suppose  $x_n = n \rightarrow x$ . Then for  $n > x + 1$ ,  $d(x_n, x) = |\tan^{-1}(n) - \tan^{-1}(x)| < |\tan^{-1}(x+1) - \tan^{-1}(x)|$ , so  $d(x_n, x) \not\rightarrow 0$ . The sequence does not converge.

Let  $X = \mathbb{R} \cup \{-\infty, \infty\}$ . Let  $d : X \times X \rightarrow \mathbb{R}$  be defined by

$$\begin{aligned} d(x, y) &= |\tan^{-1}(x) - \tan^{-1}(y)| & d(x, \infty) &= \frac{\pi}{2} - \tan^{-1}(x) \\ d(-\infty, x) &= \frac{\pi}{2} + \tan^{-1}(x) & d(-\infty, \infty) &= \pi \end{aligned}$$

Then  $d$  is a metric on  $X$ , and moreover  $(X, d)$  is complete. The counterexample we found for  $(\mathbb{R}, d)$  has been “filled,” in a sense. The hole is no longer there. The sequence  $x_n = n$  now converges to  $\infty$ . Somewhat unsurprisingly,  $\mathbb{R}$  is dense in  $X$ , with respect to  $d$ . Every element in  $X$  is the limit of a sequence of elements in  $\mathbb{R}$ .

**Definition 32.1.27** A completion of a metric space  $(X, d)$  is a complete metric space  $(\tilde{X}, \tilde{d})$  such that  $X \subset \tilde{X}$  and the restriction of  $\tilde{d}$  onto  $X$  is equal to  $d$ .

**Theorem 32.1.19.** *Every metric space has a completion.*

**Definition 32.1.28** An isometry between metric spaces  $(X, d_X)$  and  $(Y, d_Y)$  is a function  $f : X \rightarrow Y$  such that  $d_X(x, y) = d_Y(f(x), f(y))$  for all  $x, y \in X$ .

**Definition 32.1.29** Isometric metric spaces are metric spaces with an isometry between them.

**Theorem 32.1.20.** *If  $(X, d)$  is a metric space and  $(\tilde{X}_1, \tilde{d}_1)$  and  $(\tilde{X}_2, \tilde{d}_2)$  are completions of  $(X, d)$ , then  $(\tilde{X}_1, \tilde{d}_1)$  and  $(\tilde{X}_2, \tilde{d}_2)$  are isometric.*

This says the completion of a metric space is unique up to isometry. The Lebesgue space  $L^p(S)$  can be defined to be the completion of  $C(S)$  with respect to the  $d_p$  metric.

**Theorem 32.1.21.**  *$(C(S), d_\infty)$  is complete.*

*Proof.* Suppose  $x_n$  is a Cauchy sequence and let  $\varepsilon > 0$ . As  $x_n$  is Cauchy, there exists  $N \in \mathbb{N}$  such that for all  $n, m > N$ ,  $\sup |x_m(t) - x_n(t)| < \frac{\varepsilon}{3}$ . But then for all  $t \in S$ ,  $|x_m(t) - x_n(t)| < \frac{\varepsilon}{3}$ , for all  $n, m > N$ . That is, if  $x_n$  is a Cauchy sequence in  $(C(S), d_\infty)$ , then it is a Cauchy sequence in  $(\mathbb{R}, d_1)$ . But  $(\mathbb{R}, d_1)$  is complete, and therefore, for all  $t \in S$ , there is an  $x(t)$  such that  $x_n(t) \rightarrow x(t)$  with respect to the  $d_1$  metric on  $\mathbb{R}$ . We now need to show that  $x(t)$  is a continuous function. That is, that  $x(t) \in C(S)$ . Finally we need to show that  $x_n \rightarrow x$  with respect to  $d_\infty$ . We need to show that for all  $\varepsilon > 0$  and all  $t \in S$  there is a  $\delta > 0$  such that for all  $|t - t_0| < \delta$ ,  $|x(t) - x(t_0)| < \varepsilon$ . But for all  $n, m > N$ ,  $\sup |x_n(t) - x_m(t)| < \frac{\varepsilon}{3}$ . Taking the limit on  $m$ , we have  $|x(t) - x_n(t)| < \frac{\varepsilon}{2}$ . But  $x_n(t)$  is continuous, and thus there exists  $\delta > 0$  such that for all  $|t - t_0| < \delta$ ,  $|x_n(t) - x_n(t_0)| < \frac{\varepsilon}{3}$ . But  $|x(t) - x(t_0)| \leq |x(t) - x_n(t)| + |x_n(t) - x_n(t_0)| + |x_n(t_0) - x(t_0)|$ . But  $|x_n(t_0) - x(t_0)| < \sup |x_n(t) - x_n(t)| < \frac{\varepsilon}{3}$ , and therefore  $|x(t) - x(t_0)| < \varepsilon$ . So  $x(t)$  is continuous.  $\square$

The Weierstrass Approximation Theorem says that, for closed finite intervals  $S$ ,  $(C(S), d_\infty)$  is the completion of the set of polynomials with respect to the  $d_\infty$  metric. On the other hand,  $(C[0, 1], d_p)$  is not complete when  $1 \leq p < \infty$ . For define the following:

$$H(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{2} \\ 1, & \frac{1}{2} < x \leq 1 \end{cases}$$

This is discontinuous and cannot be approximated arbitrarily well by any continuous function. However, the area underneath  $H$  can be approximated arbitrarily well by continuous functions. For define:

$$x_n(t) = \begin{cases} 0, & 0 \leq t \leq \frac{1}{2} - \frac{1}{n} \\ n(t - \frac{1}{2} + \frac{1}{n}), & \frac{1}{2} - \frac{1}{n} \leq t \leq \frac{1}{2} \\ 1, & \frac{1}{2} < t \leq 1 \end{cases}$$

Then the area under  $x_n(t)$  is  $\frac{1}{2} + \frac{1}{2n}$ , and thus  $d_1(x_n(t), x_m(t)) = |\frac{1}{2m} - \frac{1}{2n}|$ , and therefore  $x_n(t)$  is a Cauchy sequence. But  $x_n(t)$  does not converge in  $(C[0, 1], d_1)$ . For suppose not, suppose  $x_n(t) \rightarrow x(t)$ , and  $x(t) \in C[0, 1]$ . If  $x(1/2) \geq 1/2$ , then, as  $x(t)$  is continuous, there is a  $\delta > 0$  such that for all  $|t - 1/2| < \delta$ ,  $x(t) > 1/4$ . But then  $d(x_n, x) = \int_0^1 |x(t) - x_n(t)| dt \geq \int_{1/2-\delta/2}^{1/2} |x(t) - x_n(t)| dt$ . But  $|x| = |(x - y) + y| \leq |x - y| + |y|$ , and thus  $|x| - |y| \leq |x - y|$ . From this we have  $d(x_n(t), x(t)) \geq \int_{1/2-\delta/2}^{1/2} (x(t) - x_n(t)) dt > \int_{1/2-\delta/2}^{1/2} \frac{1}{4} dt - \int_0^{1/2} x_n(t) dt = \frac{1}{4}\delta - \frac{1}{2n} \rightarrow \frac{1}{4}\delta$ . But then  $d(x_n(t), x(t)) \not\rightarrow 0$ . Therefore  $x_n(t)$  does not converge.

**Theorem 32.1.22.** *If  $1 \leq p < \infty$ , then  $(\ell^p, d_p)$  is complete.*

*Proof.* Let  $x_n$  be a Cauchy sequence in  $(\ell^p, d_p)$ ,  $x_n = x_n(1), x_n(2), \dots, x_n(k), \dots$ . Then, for  $n, m \in \mathbb{N}$ ,  $d_p(x_n, x_m) = (\sum_{k=0}^{\infty} |x_n(k) - x_m(k)|^p)^{1/p}$ . As  $x_n$  is Cauchy, for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that for all  $n, m > N$ ,  $d_p(x_n, x_m) < \varepsilon$ . But then, for all  $n, m > N$  and all  $k \in \mathbb{N}$ ,  $|x_n(k) - x_m(k)|^p < d_p(x_n, x_m)^p < \varepsilon^p$ . But then  $|x_n(k) - x_m(k)| < \varepsilon$ . Therefore  $x_n(k)$  is a Cauchy sequence in  $(\mathbb{R}, d)$ , and this metric space is complete. Therefore, for all  $k \in \mathbb{N}$ , there is a  $z_k$  such that  $x_n(k) \rightarrow z_k$ . We now need to show that  $z_k$  is an element of  $\ell^p$  and that  $x_n \rightarrow z_k$  with respect to the  $d_p$  metric. For let  $N \in \mathbb{N}$ . Then  $\sum_{k=0}^N |x_n(k) - x_m(k)|^p \leq \sum_{k=0}^{\infty} |x_n(k) - x_m(k)|^p < \varepsilon^p$ . Taking the limit on  $m$ , we have  $\sum_{k=0}^N |z_k - x_n(k)| < \varepsilon^p$ . The reason we have written a finite sum is to avoid getting into trouble with limits. An infinite sum is itself a limit, and taking limits of limits can get very messy very easily. For example,  $f(n, m) = \frac{m}{n+m}$ . Taking the limit on  $m$  first results in 1, whereas taking the limit on  $n$  first gives you 0. That is,  $\lim_n \lim_m f(n, m) \neq \lim_m \lim_n f(n, m)$ . You have to be careful when considering limits of limits. With this we have shown that  $z_k - x_n(k) \in \ell^p$  for all  $n \in \mathbb{N}$ . But  $x_n \in \ell^p$ , and  $\ell^p$  is closed under addition. Therefore  $z_k \in \ell^p$ . But also, for  $n > N$ , we have  $d_p(x_n, z) < \varepsilon$ . Thus,  $x_n$  converges.  $\square$

**Theorem 32.1.23.** *If  $(X, d)$  is complete and  $S$  is a closed subset of  $X$ , then  $(S, d_S)$  is complete, where  $d_S$  is the restriction of  $d$  onto  $S$ .*

*Proof.* Let  $x_n$  be a Cauchy sequence in  $S$ . Then  $x_n \rightarrow x$ ,  $x \in X$ , since  $x_n$  is Cauchy in  $X$  and  $X$  is complete. Since  $S$  is closed,  $x \in S$ . Therefore, etc.  $\square$

**Theorem 32.1.24.** *If  $(X, d)$  is complete and  $S \subset X$  is not closed, then  $(S, d_S)$  is not complete.*

*Proof.* If  $S$  is not closed then there is a convergent sequence  $x_n \in S$  whose limit it not in  $S$ . But then  $x_n$  is a Cauchy sequence in  $X$ , and therefore is also a Cauchy sequence in  $S$ , but  $x_n$  does not converge in  $S$ . Therefore  $(S, d_S)$  is not complete.  $\square$

Recall that  $c_0$  is the set of sequences which tend to zero. That is, it is the set of null sequences.

**Theorem 32.1.25.**  $c_0$  is a closed subset of  $(\ell^\infty, d_\infty)$

*proof 1.* Let  $x_n$  be a sequence in  $c_0$  that converges to  $z \in \ell^\infty$  with respect to  $d_\infty$ . Then  $\sup\{|x_n(k) - z_k|\} \rightarrow 0$ . We need to show that  $z \in c_0$ . Let  $\varepsilon > 0$ . Let  $N_1 \in \mathbb{N}$  be such that  $n > N$  implies  $\sup\{|x_n(k) - z_k|\} < \frac{\varepsilon}{2}$ . But  $x_n \in c_0$  for all  $n$ , and thus  $x_n(k) \rightarrow 0$  as  $k \rightarrow \infty$ . Thus, there is an  $N_2 \in \mathbb{N}$  such that  $n > N_2$  implies  $|x_n(k)| < \varepsilon$ . But then for  $n > \max\{N_1, N_2\}$ ,  $|z_k| \leq |z_k - x_n(k)| + |x_n(k)| < \varepsilon$ .  $\square$

*Proof 2.* We can also show that  $c_0^C$  is open. Let  $x \in c_0^C$ . Then there is an  $r > 0$  and a subsequence  $x_{k_n}$  of  $x$  such that  $x_{k_n} > r$  for all  $n$ . But then  $B_{r/2}(x)$  is an open ball contained in  $c_0^C$ . For if  $y \in B_{r/2}(x)$ , then  $d_\infty(x, y) = \sup\{|x_n - y_n|\} < r/2$ , and thus  $|y_{k_n} - x_{k_n}| < r/2$ , and therefore  $|y_{k_n}| > r/2$ . Thus,  $y$  is not a null sequence and  $c_0^C$  is open. So  $c_0$  is closed.  $\square$

Let  $X$  be the set of sequences with only finitely many nonzero terms. Then  $(X, d_\infty)$  is not complete. Let  $x_1 = (1, 0, 0, \dots)$ ,  $x_2 = (1, 1/2, 0, 0, \dots)$ ,  $x_n = (1, 1/2, \dots, 1/n, 0, 0, \dots)$ . Then  $d_\infty(x_n, x_m) = 1/\max\{n, m\} \rightarrow 0$ . But clearly  $x_n \rightarrow (1, 1/2, \dots, 1/n, \dots)$ , which is an element of  $c_0$ , but not an element of  $X$ . Thus  $X$  is not closed, and therefore is not complete. Returning to  $C[0, 1]$ , when we had that sequence of continuous functions that clearly converged to a discontinuous function, we still needed to show that there is no continuous function that the  $x_n(t)$  converged to. Here we've embedded  $X$  into a bigger space, shown that the sequence converges to something outside of  $X$ , in our case an element of  $c_0 \setminus X$ , and then used the uniqueness of limits to show that the limit does not converge in  $X$ .

### 32.1.4 Banach's Fixed Point Theorem

If  $(X, d)$  is a complete metric space, and if  $T : X \rightarrow X$  satisfies the property that, for all  $x$  and  $y$  in  $X$ ,  $d(T(x), T(y)) < kd(x, y)$  for some  $k < 1$ , then  $T$  has a unique point  $x$ , called a fixed point, such that  $T(x) = x$ .

**Definition 32.1.30** A contraction of a metric space  $(X, d)$  is a function  $T : X \rightarrow X$  such that there exists a  $k \in (0, 1)$  such that for all  $x, y \in X$ ,  $d(T(x), T(y)) < kd(x, y)$ .

**Definition 32.1.31** A fixed point of a function  $f : X \rightarrow X$  is a point  $x \in X$  such that  $f(x) = x$ .

**Theorem 32.1.26** (Banach's Fixed Point Theorem). *If  $(X, d)$  is a complete metric space and  $T : X \rightarrow X$  is a contraction, then there is a unique fixed point  $x \in X$  with respect to  $T$ .*

**Definition 32.1.32** A Lipschitz continuous function is a function  $f : [a, b] \rightarrow \mathbb{R}$  such that there is an  $L \in \mathbb{R}$  such that  $|f(x) - f(y)| < L|x - y|$  for all  $x, y \in [a, b]$ .

This says that the slopes of the secant lines of the function are bounded. The square root function  $y = \sqrt{x}$  is an example of a function that is not Lipschitz. The slopes of secant lines go to infinity as the points tend towards the origin.

**Theorem 32.1.27** (Picard's Theorem). *If  $f : [a, b] \times \mathbb{R} \rightarrow \mathbb{R}$  is Lipschitz continuous, Then there is a unique function  $x : [a, b] \rightarrow \mathbb{R}$  such that  $\frac{dx}{dt} = f(t, x(t))$  and  $x(a) = a$ .*

*Proof.* We prove Picard by using the Banach Fixed Point Theorem. First we write the problem as an integral equation. If  $\dot{x} = f(t, x(t))$ , then:

$$x(t) = \int_a^t \frac{dx}{dt} dt = x_0 + \int_a^t f(t, x(t)) dt$$

Let  $(X, d)$  be  $C[a, b]$  with the supremum norm  $d_\infty$ . Then  $(X, d)$  is a complete metric space. Let  $T : X \rightarrow X$  be defined by:

$$Tx = x_0 + \int_a^t f(t, x(t)) dt$$

All we need to do is show that  $T$  is a contraction. Applying the Banach Fixed Point theorem then shows that there is a unique fixed point of  $T$ , thus showing that there is a unique solution to our original initial value problem. If  $x, y \in X$ , then:

$$\begin{aligned} d(Tx, Ty) &= \sup\{|Tx(t) - Ty(t)|\} \\ &= \sup\{(x_0 + \int_a^t f(t, x(t)) dt) - (x_0 + \int_a^t f(t, y(t)) dt)\} \\ &= \sup\{\int_a^t f(t, x(t)) dt - \int_a^t f(t, y(t)) dt\} \\ &\leq \int_a^t |f(t, x(t)) - f(t, y(t))| dt \end{aligned}$$

But from the Lipschitz continuity of  $f$ , we have:

$$\begin{aligned} d(Tx, Ty) &\leq L \int_a^t |x(t) - y(t)| dt \\ &\leq L(t - a)d(x, y) \\ &\leq L(b - a)d(x, y) \end{aligned}$$

So  $T$  is a contraction for  $L(b - a) < 1$ . Usually we can extend this solution by taking  $b$  as the initial condition and stepping forward one interval at a time. We'll take a different approach. We have that  $d(Tx, Ty) \leq L(b - a)d(x, y)$ . From this, we obtain:

$$\begin{aligned} d(T^2x, T^2y) &\leq L \int_a^b d(Tx, Ty) dt \\ &\leq L \int_a^t L(t-a)d(x, t) dt \\ &= \frac{L^2}{2}(t-a)^2 d(x, y) \\ &\leq \frac{L^2}{2}(b-a)^2 d(x, y) \end{aligned}$$

Applying induction, we have:

$$d(T^n x, T^n y) \leq \frac{L^n}{n!} (b-a)^n$$

But this tends to zero, and thus there is an  $N$  such that, for all  $n > N$ ,  $T^n$  is a contraction. But then, by the Banach Fixed Point Theorem, there is a unique point  $x$  such that  $T^n x = x$ . But then  $Tx = T^n(Tx)$ , and thus  $Tx$  is a fixed point of  $T^n$ . But the fixed point of  $T^n$  is unique, and  $x$  is a fixed point. Therefore  $Tx = x$ . Therefore, etc.  $\square$

Without Lipschitz continuous you may lose uniqueness, but you still have existence. This is Peano's theorem. An example is  $\dot{x} = \sqrt{x}$  with  $x(0) = 0$ . This has solutions  $x(t) = 0$  and  $(t) = t^2/4$ . Now back to compactness.

### Compactness

**Definition 32.1.33** A metric space  $(X, d)$  is sequentially compact if every sequence in  $X$  has a convergent subsequence.

In topology there is a difference between sequential compactness and regular compactness, but in metric spaces they turn out to be the same. A subset of  $S$  of  $X$  is compact if every sequence in  $S$  has a subsequence which converges. That is,  $(S, d)$  is compact.

**Theorem 32.1.28.** *A subset  $S$  of a compact metric space  $(X, d)$  is compact if and only if  $S$  is closed.*

*Proof.* For let  $x_n$  be a sequence in  $S$ . Then  $x_n$  is a sequence in  $X$  and thus there is a convergent subsequence  $x_{k_n}$  with a limit  $x$ . But  $x_{k_n}$  is in  $S$  and  $S$  is closed, and therefore  $x$  is in  $S$ . Thus,  $S$  is compact. Conversely, if  $S$  is

compact, suppose it is not closed. Then there is a point  $y \in X$  such that  $y$  is a limit point of  $S$  but not contained in  $S$ . Let  $x_n$  be a sequence that converges to  $y$ . Then, as  $S$  is compact, there is a convergent subsequence. But the limit of this subsequence is  $y$ , a contradiction as  $y \notin S$ . Therefore  $S$  is closed.  $\square$

**Theorem 32.1.29.** *If  $(X, d)$  is a compact metric space, then  $(X, d)$  is complete.*

*Proof.* If  $x_n$  is Cauchy in  $X$ , then there is a convergent subsequence  $x_{k_n}$  in  $X$ . But if  $x_{k_n}$  converges to  $x$ , then  $x_n$  converges to  $x$  as well, as  $x_n$  is Cauchy. Therefore,  $(X, d)$  is complete.  $\square$

**Theorem 32.1.30** (Heine-Borel Theorem). *A subset of  $\mathbb{R}^n$  is compact if and only if it is closed and bounded.*

**Example 32.1.29** The closed unit ball of  $\ell^p$  is not compact, if  $1 \leq p \leq \infty$ . Let  $x_n(m)$  be the sequence (of sequences) such that  $x_n(m) = 1$  if  $n = m$ , and zero otherwise. Then  $d_p(x_n, x_m) = 2^{1/p}$ , so  $x_n$  has no subsequence which is Cauchy. But then there is no convergent subsequence either, and therefore  $\ell^p$  is not compact.

**Example 32.1.30** The closed unit ball in  $(C[0, 1], d_\infty)$  is not compact. For let  $x_n(t) = t^{2^n}$ . Then (Do some calculus) the maximum of  $d(x_n, x_{n+1})$  is always  $1/4$ . So this has no subsequence which is Cauchy, and thus no convergent subsequence exists.

**Definition 32.1.34** A metric space  $X$  is totally bounded if for all  $\varepsilon > 0$  there is a finite number of points  $x_n$  such that  $B_\varepsilon(x_n)$  covers the entirety of  $X$ .

**Theorem 32.1.31.** *A compact metric space is totally bounded.*

*Proof.* Suppose not. Then there is an  $\varepsilon > 0$  such that no finite collection  $B_\varepsilon(x_n)$  is a covering of  $X$ . Let  $x_1 \in X$ . Then  $B_\varepsilon(x_1)$  is not  $X$ . Thus there is an  $x_2$  such that  $x_2 \notin B_\varepsilon(x_1)$ . But also  $B_\varepsilon(x_1) \cup B_\varepsilon(x_2)$  is not the entirety of  $X$ . Continuing we have that there is a sequence  $x_n$  such that, for all  $n \neq m$ ,  $d(x_n, x_m) \geq \varepsilon$ . So there is no convergent subsequence. But  $X$  is compact, a contradiction. Therefore, etc.  $\square$

There are metric spaces that are bounded but not totally bounded. For let  $X = \mathbb{R}$  and  $d$  be the discrete metric. Then, for  $\varepsilon = 1/2$ , there is no finite covering. Every point needs its own ball, so the covering is uncountable.

**Theorem 32.1.32.** *If  $(X, d)$  is complete and totally bounded, then it is compact.*

*Proof.* Let  $x_n$  be a sequence in  $X$ . Let  $\varepsilon = 1$ . Then there are finitely many points  $y_k$  such that  $B_\varepsilon(y_k)$  covers  $X$ . Then one of these balls has infinitely

many of the  $x_n$ . Similarly, for  $\varepsilon = \frac{1}{n}$ , there is a finite number of points  $y_k$  such that  $B_{\frac{1}{n}}(y_k)$  covers  $X$ . Thus there is a point with infinitely many of the  $x_n$  in it. So, we can find a subsequence such that, for  $n, m > N$ ,  $d(x_{k_n}, x_{k_m}) < \frac{1}{N}$ . But  $(X, d)$  is complete, and therefore  $x_{k_n}$  converges. Therefore  $x_n$  has a convergent subsequence. Thus,  $(X, d)$  is compact.  $\square$

**Theorem 32.1.33.** *Compact spaces are separable.*

*Proof.* If  $X$  is compact, then it is totally bounded. But then, for  $\varepsilon = 1/n$  there is a finite covering of  $X$  with balls of radius  $\varepsilon$ . Then, taking all of the centers of all of the points for all  $n$  (Countable union of finite points is countable), we obtain a countable dense subset.  $\square$

**Example 32.1.31** There are “infinite dimension” sets that are also compact. Two in particular worth mentioning. The first is the hilbert Cube. It’s a subset of  $\ell^2$  whose elements are such that  $|x_n| < 1/n$ . That is, elements are sequences whose  $n^{th}$  elements are less than  $1/n$ . This is compact. Arzela-Ascoli. Peano.

## 32.2 Normed and Inner Product Spaces

### 32.2.1 Basic Definitions

We’re finally going to put some structure on these sets, and talk about vector spaces. In a metric space, the only thing you can really talk about is the distance between points. In a vector space we have a lot more structure. We will start off with vector spaces over the reals  $\mathbb{R}$ . The main properties are that there is a **0** element, addition is well defined and is both associative and commutative, there is a notion of scalar multiplication that is associative, and the distributive law holds.

**Example 32.2.1**  $\mathbb{R}^n$ , with its usual notion of addition, and with scalar multiplication defined over  $\mathbb{R}$ , is a vector space.

**Definition 32.2.1** A norm on a vector space  $X$  over  $\mathbb{R}$  is a function  $\|\cdot\| : X \rightarrow \mathbb{R}$  such that:

1. For all  $\mathbf{x} \in X$ ,  $\|\mathbf{x}\| \geq 0$  and  $\|\mathbf{x}\| = 0$  if and only if  $\mathbf{x} = \mathbf{0}$ . [Positive Definiteness]
2. For all  $\mathbf{x} \in X$  and  $c \in \mathbb{R}$ ,  $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$  [Homogeneity]
3. For all  $\mathbf{x}, \mathbf{y} \in X$ ,  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$  [Triangle Inequality]

We have seen before that  $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$  defines a metric, and thus  $(X, d)$  is a metric space. Thus, for every vector space there is an associated metric space, the metric  $d$  called the *induced* metric.

**Definition 32.2.2** A normed vector space is a vector space  $X$  over  $\mathbb{R}$  with a norm  $\|\cdot\|$  on  $X$ .

**Example 32.2.2**  $\mathbb{R}^n$  with  $\|\mathbf{x}\|_p$ , for  $p \geq 1$ , is a normed vector space.

**Example 32.2.3**  $\ell^p$  with  $\|x\|_p$  is also a normed vector space.

**Example 32.2.4**  $C[a, b]$  equipped with the supremum norm,  $\|x(t)\|_\infty$ , is a normed vector space.

### Inner Product Spaces

**Definition 32.2.3** An inner product on a vector space  $X$  over  $\mathbb{R}$  is a function  $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$  such that:

1. For all  $x \in X$ ,  $\langle x, x \rangle \geq 0$  and  $\langle x, x \rangle = 0$  if and only if  $x = \mathbf{0}$ . [Positive Definiteness]
2. For all  $\mathbf{x}, \mathbf{y} \in X$ ,  $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$  [Symmetry]
3. For all  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$  and all  $\alpha, \beta \in \mathbb{R}$ ,  $\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle$  [Linearity]

**Example 32.2.5**  $\mathbb{R}^2$  with  $\langle (x_1, x_2), (y_1, y_2) \rangle = x_1 y_1 + x_2 y_2$  is an inner product. Replacing this with  $\mathbb{R}^n$  and doing  $\sum_{k=1}^n x_k y_k$  is also an inner product. This is the usual dot product that one sees in a vector calculus course. In  $\ell^2$ ,  $\sum_{k=1}^\infty x_k y_k$  is an inner product as well. Note also that  $\sum |x_i y_i|$  converges since  $|x_i y_i| \leq \frac{1}{2} |x_i^2| + \frac{1}{2} |y_i|^2$ .

**Example 32.2.6** In  $C[a, b]$ , let  $\langle x(t), y(t) \rangle = \int_a^b x(t)y(t)dt$ . This defines an inner product.

**Definition 32.2.4** An inner product space is a vector space  $X$  over  $\mathbb{R}$  with an inner product  $\langle \cdot, \cdot \rangle$ .

**Theorem 32.2.1** (Cauchy-Schwarz Inequality). *If  $X$  is an inner product space and  $x, y \in X$ , then  $|\langle x, y \rangle| < \|x\| \|y\|$*

*Proof.* For all  $y \in \mathbb{R}$ ,  $\langle x + ty, x + ty \rangle = \langle x, x \rangle + 2t\langle x, y \rangle + t^2\langle y, y \rangle = \|x\|^2 + 2t\langle x, y \rangle + t^2\|y\|^2$ . Thus we have a quadratic in  $t$ . But this is always positive, and thus the discriminant must be non-positive. Therefore  $(2\langle x, y \rangle)^2 - 4\|x\|^2\|y\|^2 \leq 0$  and thus  $|\langle x, y \rangle| \leq \|x\| \|y\|$ .  $\square$

**Theorem 32.2.2.** *If  $X$  is a vector space over  $\mathbb{R}$  and  $\langle \cdot, \cdot \rangle$  is an inner product, then  $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$  is a norm on  $X$ .*

*Proof.* Positivity, homogeneity, and definiteness are pretty easy. The only tricky thing to check is the triangle inequality. We have that  $\|x + y\| = \langle x + y, x + y \rangle$ , and this simplify to  $\|x\|^2 + 2\langle x, y \rangle + \|y\|^2$ . But from the Cauchy-Schwartz inequality, we have  $\langle x, y \rangle \leq \|x\|\|y\|$ . Thus  $\|x+y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$ . Taking square roots completes the theorem.  $\square$

In  $\mathbb{R}^n$ , the Cauchy-Schwartz inequality says that the dot product of two vectors is less than or equal to the product of the magnitude of the two vectors. This is obvious from the fact that the dot product of two vector is the product of the magnitudes and the *cosine* of the angle between them. Since the cosine of a number is less than or equal to one, this would complete the theorem. In  $\ell^p$  and  $L^p$  spaces, this is the special case of the Hölder inequality for when  $p = q = 2$ .

### Convergence in Normed Spaces

In a metric space, convergence meant that  $d(x_n, x) \rightarrow 0$ . In a normed space we have the induced metric, and thus we may define convergence as  $\|x_n - x\| \rightarrow 0$ .

**Definition 32.2.5** A convergent sequence in a normed space  $X$  is a sequence  $x_n$  such that there is an  $x \in X$  such that  $\|x_n - x\| \rightarrow 0$ .

Since  $\|y\| = \|(y-x)+x\| \leq \|y-x\| + \|x\|$ , it follows that  $\|x\| - \|y\| \leq \|x-y\|$ . But then if  $x_n \rightarrow x$ , then  $\|x_n\| - \|x\| \leq \|x_n - x\|$ , and  $\|x_n - x\| \rightarrow 0$ . Therefore  $\|x_n\| \rightarrow \|x\|$ . That is, the norm function is a continuous function. Similarly, if  $x_n \rightarrow x$ , then  $\langle x_n, y \rangle \rightarrow \langle x, y \rangle$ . In fact, if  $x_n \rightarrow x$  and  $y_n \rightarrow y$ , then  $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$ . To see this, we have  $\langle x_n, y_n \rangle - \langle x, y \rangle = \langle x_n - x, y \rangle + \langle x, y - y_n \rangle$  and therefore  $|\langle x_n, y_n \rangle - \langle x, y \rangle| \leq \|x_n - x\|\|y_n\| + \|x\|\|y - y_n\|$ . But  $\|x - x_n\| \rightarrow 0$  and  $\|y - y_n\| \rightarrow 0$ . But also  $\|y_n\| = \|(y_n - y) + y\| \leq \|y_n - y\| + \|y\|$ , which is bounded. Therefore  $\langle x_n, y_n \rangle - \langle x, y \rangle \rightarrow 0$ . So inner product spaces and normed spaces are metric spaces and we can define everything we did for metric spaces and all of the previous results remain true. That is, the notions and theorems pertaining to convergence, completeness, compactness, the notion of open and closed. All of these still make sense in these new spaces.

### Banach Spaces and Hilbert Spaces

**Definition 32.2.6** A Banach Space is a normed vector space  $X$  that is complete with respect to the induced metric.

**Definition 32.2.7** A Hilbert Space is an inner product space  $X$  that is complete with respect to the induced metric.

## Linear Operators

Let  $X$  and  $Y$  be normed spaces. A mapping  $T : X \rightarrow Y$  is called a linear operator if, for all  $x, y \in X$ , and for all  $\alpha, \beta \in \mathbb{R}$ ,  $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ . Usually, with operators, we simply write  $Tx$  and  $Ty$ . Similar to how we write matrix multiplication over vectors. In  $\mathbb{R}^n$ , every  $n \times n$  matrix defines a linear operator.

**Definition 32.2.8** A linear operator from a normed vector space  $X$  to a normed vector space  $Y$  is a function  $T : X \rightarrow Y$  such that, for all  $x, y \in X$  and for all  $\alpha, \beta \in \mathbb{R}$ ,  $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$ .

**Definition 32.2.9** A bounded linear operator from a normed vector space  $X$  to a normed vector space  $Y$  is a linear operator  $T : X \rightarrow Y$  such that there is a  $K \in \mathbb{R}$  such that for all  $x \in X$ ,  $\|Tx\| \leq K\|x\|$

In a just world, “bounded” would mean  $\|Tx\| \leq K$ . However, the only linear mapping that does this is the zero mapping. For if  $\|Tx\| = 1$ , then  $\|T(2x)\| = 2$ , and so on, and thus no linear mapping is bounded (With the exception of the zero mapping). Boundedness of a norm  $T : X \rightarrow Y$  depends on the norms of the space.

**Theorem 32.2.3.** *Bounded linear operators are continuous.*

*Proof.* If  $x_n \rightarrow x$ , then  $\|Tx_n - Tx\| = \|T(x_n - x)\|$ . But  $T$  is bounded, and thus there is a  $K$  such that  $\|T(x_n - x)\| \leq K\|x_n - x\|$ . But  $\|x_n - x\| \rightarrow 0$ . Therefore, etc.  $\square$

The converse is also true.

**Theorem 32.2.4.** *If  $T$  is a continuous linear operator, than there exists a  $\delta > 0$  such that for all  $x \in B_\delta(0)$ ,  $\|Tx - T0\| < 1$ . But from linearity,  $T0 = 0$ , and thus  $\|Tx\| < 1$ . Then for any  $z \in Z$ , we have  $\|\frac{\delta}{2}\frac{z}{\|z\|}\| = \frac{\delta}{2}$ , and thus  $\|T(\frac{\delta}{2}\frac{z}{\|z\|})\| < 1$ . Letting  $K = \delta$ , we have  $\|Tx\| < K\|x\|$ . Thus,  $T$  is bounded.*

Continuity at 0 implies uniform continuity since if  $x_n - y_n \rightarrow 0$ , then  $\|Tx_n - Ty_n\| = \|T(x_n - y_n)\| \leq K\|x_n - y_n\| \rightarrow 0$ . The set of bounded linear operators form a vector space, where addition is  $(S + t)(x) = (Sx) + (Tx)$ , and scalar multiplication is defined by  $(\alpha T)(x) = \alpha(Tx)$ . We must show that when you add two bounded linear operators, the result is a bounded linear operator.

**Theorem 32.2.5.** *If  $T_1 : X \rightarrow Y$  and  $T_2 : X \rightarrow Y$  are bounded linear operators, then  $T_1 + T_2$  is a bounded linear operator.*

*Proof.* For let  $T_1$  and  $T_2$  be bounded. Then there are  $K_1, K_2$  such that, for all  $x \in X$ ,  $\|T_1x\| \leq K_1\|x\|$  and  $\|T_2x\| \leq K_2\|x\|$ . But then  $\|(T_1 + T_2)x\| = \|T_1x + T_2x\| \leq \|T_1x\| + \|T_2x\| \leq K_1\|x\| + K_2\|x\|$ . Let  $K = K_1 + K_2$ .  $\square$

**Theorem 32.2.6.** *If  $T : X \rightarrow Y$  is a bounded linear operator, and  $\alpha \in \mathbb{R}$ , then  $\alpha T$  is a bounded linear operator.*

*Proof.* For  $\|\alpha Tx\| = |\alpha| \|Tx\| \leq |\alpha| K \|x\| = K |\alpha| \|x\|$ . □

We write  $B(X, Y)$  to denote the set of bounded linear operators from  $X$  to  $Y$ . That is, linear operators  $T : X \rightarrow Y$ . We can define a norm on  $B(X, Y)$  as follows:  $\|T\|_B = \sup_{x \in X, x \neq 0} \left\{ \frac{\|Tx\|}{\|x\|} \right\}$ . This is the “Smallest  $K$ ,” used as a bound for the linear operator  $T$ . This shows that  $\|Tx\|_Y \leq \|T\|_B \|x\|_X$ .

### 32.2.2 Lecture 7: October 22, 2018

#### Bounded Linear Operators

A bounded linear operator is a function  $T : X \rightarrow Y$  between normed spaces  $X$  and  $Y$  such that  $T$  is linear, and there exists a  $K \in \mathbb{R}$  such that, for all  $x \in X$ ,  $\|Tx\|_Y \leq K \|x\|_X$ . The norm of  $T$ ,  $\|T\|$ , is then defined as the smallest such  $K$ . Equivalently:

$$\|T\| = \sup \left\{ \frac{\|Tx\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\} = \sup \{ \|Tx\|_Y : \|x\|_X = 1 \}$$

The set of all bounded linear operators from a normed space  $X$  to a normed space  $Y$  is denoted  $B(X, Y)$ . This is a vector space with addition defined as  $(T + S)x = (Tx) + (Sx)$  and  $(aT)x = a(Tx)$ .

**Theorem 32.2.7.**  *$\|T\|$  defines a norm on  $B(X, Y)$ .*

*Proof.* For  $\|T\| \geq 0$  and  $\|Tx\| = 0$  if and only if  $Tx = 0$  for all  $x \in X$ , and thus  $T$  is the zero operator. If  $\alpha \in \mathbb{R}$ , then:

$$\begin{aligned} \|\alpha T\| &= \sup \left\{ \frac{\|\alpha Tx\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\} \\ &= |\alpha| \sup \left\{ \frac{\|Tx\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\} \\ &= |\alpha| \|T\| \end{aligned}$$

Finally, if  $S, T \in B(X, Y)$ , then:

$$\begin{aligned} \|S + T\| &= \sup \left\{ \frac{\|(S + T)x\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\} \\ &= \sup \left\{ \frac{\|Sx + Tx\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\} \\ &\leq \sup \left\{ \frac{\|Sx\|_Y + \|Tx\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\} \\ &\leq \|T\| + \|S\| \end{aligned}$$

□

**Theorem 32.2.8.** *If  $Y$  is a Banach space, and if  $X$  is a normed space, then  $B(X, Y)$  is a Banach space.*

*Proof.* For let  $T_n$  be a Cauchy sequence in  $B(X, Y)$  and let  $\varepsilon > 0$ . Then there exists  $N_0 \in \mathbb{N}$  such that for all  $n, m > N_0$ ,  $\|T_n - T_m\| < \varepsilon$ . That is, for all  $n, m > N_0$ :

$$\begin{aligned} \sup \left\{ \frac{\|T_n x - T_m x\|_Y}{\|x\|_X} : x \in X, x \neq 0 \right\} &\leq \varepsilon \\ \Rightarrow \frac{\|T_n x - T_m y\|_Y}{\|x\|_X} &\leq \varepsilon \end{aligned}$$

That is,  $T_n x$  is a Cauchy sequence in  $Y$  for any fixed value  $x \in X$ . But  $Y$  is a Banach space, and is therefore complete. But then if  $T_n x$  is a Cauchy sequence in  $Y$  it has a limit  $y \in Y$ . Let  $Tx = \lim_{n \rightarrow \infty} T_n x$  for all  $x \in X$ . Then  $T \in B(X, Y)$ . For:

$$T(x + y) = \lim_{n \rightarrow \infty} T_n(x + y) = \lim_{n \rightarrow \infty} (T_n x + T_n y) = Tx + Ty$$

And similarly  $(\alpha T)x = \alpha Tx$ . Lastly,  $T$  is bounded. For all  $n, m > N$  we have  $\|T_n x - T_m x\|_Y / \|x\|_X < \varepsilon$ . Taking the limit on  $m$ , we have  $\|Tx - T_n x\|_Y / \|x\|_X \leq \varepsilon$  for all  $n > N_0$ . Thus,  $\|T_n x - Tx\|_X \leq \varepsilon \|x\|_X$ . But  $\|Tx - T_n x\|_Y = \|T_n x - (T_n - Tx)\|_Y$ , and therefore  $\|Tx\| \leq \varepsilon \|x\|_X + \|T_n x\|$ , and  $\|T_n x\| \leq \|T_n\|$ , and therefore  $\|Tx\| \leq \varepsilon \|X\|_X + \|T\| \|x\|_X$ . But then  $\|Tx\|_Y \leq (\varepsilon + \|T_n\|) \|x\|_X$ . But  $T_n$  is bounded, and therefore  $T$  is bounded. Finally, we must show that  $T_n \rightarrow T$  in  $B(X, Y)$  with respect to the norm  $\|T_n - T\|$ . That is, we must show that  $\|T - T_n\| \rightarrow 0$ . This follows since  $\|Tx - T_n x\|_Y / \|x\|_X < \varepsilon$  for  $n > N_0$ , and therefore  $\|T - T_n\| < \varepsilon$ . Therefore, etc. □

## Dual Spaces

So if  $Y$  is a Banach space, and  $X$  is any normed space, then  $B(X, Y)$  is a Banach space. One of the most important cases is  $Y = (\mathbb{R}, ||)$ , where  $||$  is the normal absolute value “norm.”  $B(X, \mathbb{R})$  is a Banach space, and it is called the continuous dual space of  $X$ , written  $X'$ . Elements of  $X'$  are called bounded linear functionals. These are bounded linear operators whose range of the operator is the real numbers. The characterization, or the representation, or realization, of these dual spaces is a major topic in functional analysis. A lot of these theorems are due to a mathematician by the name of Riesz.

**Example 32.2.7** A functional takes an element of a normed space  $X$  and spits out a real number. For example, if  $X$  is the space of continuous functions, then

the following are functionals:

$$f_1(x) = \int_0^1 x(t)t^2 dt \quad f_2(x) = x(0.5) \quad f_3(x) = 0$$

Let  $X = (\mathbb{R}^2, \ell^1)$ . What does  $X'$  look like? That is, what is the dual space of  $X$ ? let  $f : X \rightarrow \mathbb{R}$  be defined by  $f(x_1, x_2) = 2x_1 - 5x_2$ . Then  $f \in X'$  and  $\|f\| = 5$ . More generally, every element of  $\mathbb{R}^2$  defines an element of  $X'$ . Given  $(a, b) \in \mathbb{R}^2$ , we define  $f(x_1, x_2) = ax_1 + bx_2$ .  $f$  is then linear, and:

$$\begin{aligned} |f(x_1, x_2)| &= |ax_1 + bx_2| \\ &\leq |a||x_1| + |b||x_2| \\ &\leq \max\{|a|, |b|\}(|x_1| + |x_2|) \\ &= \|(a, b)\|_\infty \|(x_1, x_2)\|_{\ell^1} \end{aligned}$$

And therefore  $f$  is bounded, as  $\|(a, b)\|_\infty$  is a bound. That is,  $\|f\| \leq \|(a, b)\|_\infty$ . By choosing  $x = (x_1, x_2)$ , where  $x_1 = 1$  and  $x_2 = 0$  if  $|b| \leq |a|$ , and  $x_1 = 0$  and  $x_2 = 1$  otherwise, we get  $|f| = \max\{(a, b)\} = \|(a, b)\|_\infty$ . Therefore  $\|f\| = \|(a, b)\|_\infty$ . On the other hand, if  $f \in X'$ , let  $a = f(1, 0)$  and  $b = f(0, 1)$ . Then, for all  $(x_1, x_2) \in \mathbb{R}^2$ :

$$f(x_1, x_2) = f(x_1(1, 0) + x_2(0, 1)) = x_1 f(1, 0) + x_2 f(0, 1) = ax_1 + bx_2$$

So the dual of  $(\mathbb{R}^2, \ell^1)$  looks very much like  $(\mathbb{R}^2, \ell^\infty)$ . In fact,  $(\mathbb{R}^2, \ell^1)'$  and  $(\mathbb{R}^2, \ell^\infty)$  are isometric and isomorphic. That is, we really can't tell them apart and we can consider them as the same thing. More generally,  $(\mathbb{R}^n, \ell^n)' = (\mathbb{R}^n, \ell^\infty)$ . Even more general, if  $p$  and  $q$  are exponential conjugates of each other (That is,  $\frac{1}{q} + \frac{1}{p} = 1$ ), then  $(\mathbb{R}^n, \ell^p)' = (\mathbb{R}^n, \ell^q)$  for all  $1 \leq p \leq \infty$ . Saying  $p = \infty$  is equivalent to saying  $q = 1$ . Setting  $p = q = 2$ , we have  $(\mathbb{R}^n, \ell^2)' = (\mathbb{R}^n, \ell^2)$ . This is true of any Hilbert space: The dual of any Hilbert Space  $\mathcal{H}$  is itself. That is,  $\mathcal{H}' = \mathcal{H}$ . This is one of the Riesz Representation Theorems. In infinite dimensions,  $(\ell^p)' = \ell^q$ , where  $p$  and  $q$  are such that  $\frac{1}{p} + \frac{1}{q} = 1$ , and  $1 \leq p < \infty$ . Now, we cannot allow  $p = \infty$ . For  $(\ell^\infty)'$  is not equal to  $\ell^1$ .

**Theorem 32.2.9.** If  $1 \leq p < \infty$  and  $\frac{1}{p} + \frac{1}{q} = 1$ , then  $(\ell^p)' = \ell^q$ .

*Proof.* If  $(f_1, f_2, \dots) \in \ell^q$ , then let  $f : \ell^p \rightarrow \mathbb{R}$  be defined by  $f(x_1, x_2, \dots) = \sum_{k=1}^{\infty} x_k f_k$ . This converges from Hölder's inequality:

$$\sum_{k=1}^{\infty} |x_k f_k| \leq \left( \sum_{k=1}^{\infty} f_k^q \right)^{1/q} \left( \sum_{k=1}^{\infty} x_k^p \right)^{1/p}$$

And therefore  $|fx| = \|(f_1, f_2, \dots)\|_q \|(x_1, x_2, \dots)\|_p$ . That is, Moreover  $f$  is linear. Therefore  $f \in (\ell^p)'$  and  $\|f\| \leq \|(f_1, f_2, \dots)\|_q$ . On the other hand, let

$x_i = |f_i|^{q/p} \operatorname{sgn}(f_i)$ . Then

$$fx = \sum_{k=1}^{\infty} f_k x_k = \sum_{k=1}^{\infty} |f_k|^{q/p+1}$$

But  $\frac{1}{p} + \frac{1}{q} = 1$ , and thus  $\frac{q}{p} + 1 = q$ . Thus:

$$\begin{aligned} |fx| &= \sum_{k=1}^{\infty} |f_k|^q \\ &= \|(f_1, f_2, \dots)\|_q^q \\ &= \|(f_1, f_2, \dots)\|_q \|(f_1, f_2, \dots)\|_q^{q-1} \\ &= \|(f_1, f_2, \dots)\|_q \left( \sum_{k=1}^{\infty} |f_k^q| \right)^{\frac{q-1}{q}} &= \|(f_1, f_2, \dots)\|_q \left( \sum_{k=1}^{\infty} |x_k|^p \right)^{1/p} \\ &= \|(f_1, f_2, \dots)\|_q \|x\|_p \end{aligned}$$

Therefore,  $\|f\| = \|(f_1, f_2, \dots)\|_q$ . Thus, for all  $y \in \ell^q$  there is a bounded linear operator  $f \in \ell^p$  such that  $\|y\|_{\ell^q} = \|f\|_{(\ell^p)'}.$  That is, every  $(f_i) \in \ell^q$  defines an element of  $(\ell^p)'$  by  $fx = \sum_{k=1}^{\infty} f_k x_k$ , for any  $(x_i) \in \ell^p$ . So  $\ell^q$  can be *embedded* into  $(\ell^p)'$ . Now we need to show that this embedding is the entirety of  $(\ell^p)'$ . If  $f \in (\ell^p)'$ , let  $f_i = f(e_i)$ , where  $e_i$  is the sequence  $(0, 0, \dots, 1, 0, 0, \dots)$ , where the 1 occurs in the  $i^{\text{th}}$  spot. We need to show that  $(f_i) \in \ell^q$  and  $fx = \sum_{k=1}^{\infty} f_k x_k$  for all  $x \in \ell^p$ . If  $x \in \ell^p$ , then:

$$x = \sum_{k=1}^{\infty} x_k e_k \Rightarrow fx = f \left( \sum_{k=1}^{\infty} x_k e_k \right) = \sum_{k=1}^{\infty} f(x_k e_k) = \sum_{k=1}^{\infty} x_k f(x_k) = \sum_{k=1}^{\infty} x_k f_k$$

Choosing  $x_k = |f_k|^{q/p} \operatorname{sgn}(f_k)$  and apply Hölder.  $\square$

### 32.2.3 Lecture 8: October 29, 2018

#### Review

If  $X$  is a vector space, an inner product on  $X$  is a mapping  $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$  such that:

1.  $\langle x, x \rangle \geq 0$  and  $\langle x, x \rangle = 0$  if and only if  $x = 0$
2.  $\langle x, y \rangle = \langle y, x \rangle$
3.  $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$

Think of the dot product for vectors. This is a generalization of this concept. Every inner product on a vector space  $V$  induce a norm on  $V$ :

$$\|x\| = \sqrt{\|x, x\|}$$

An inner product that is complete with respect to the induced norm is called a Hilbert Space. A mapping  $f : X \rightarrow \mathbb{R}$  is bounded if there is a  $K \in \mathbb{R}$  such that, for all  $x \in X$ ,  $|f(x)| \leq L\|x\|$ .  $f$  is linear if  $f(ax + by) = af(x) + bf(y)$ , for all  $x, y \in X$  and all  $a, b \in \mathbb{R}$ . The smallest such  $K$  that works is called the norm of  $f$ , denoted  $\|f\|$ . For all  $x \in X$ ,  $|f(x)| \leq \|f\|\|x\|$ . The vector space of all bounded linear functionals on  $X$  is the dual space  $X'$ . This is also a Banach space with the functional norm  $\|f\|$ . One question that arises is, how do we know that there are bounded linear functionals on a space  $X$ ? In the case that  $X$  is a Hilbert space, this is rather easy, but for a more general Banach space this is not that trivial. For any normed space  $X$  we can at least one bounded linear functional because the zero mapping  $f(x) = 0$  is such a functional. The question is then does every normed space have a bounded linear functional on it? The answer is yes, and this is related to the Hahn-Banach Theorem. As said before, in the Hilbert case this is rather easy.

**Theorem 32.2.10.** *If  $X$  is a Hilbert space, then there is a non-trivial bounded linear functional  $f : X \rightarrow \mathbb{R}$ .*

*Proof.* If  $X$  is an inner product and  $z \in X$ , let  $f(x) = \langle x, z \rangle$  for all  $x \in X$ . Then  $f$  is linear since the inner product is linear. But moreover, from Cauchy-Schwarz we have:

$$|f(x)| = |\langle x, z \rangle| \leq \|x\|\|z\|$$

And thus  $\|f\| \leq \|z\|$ . But  $|f(z)| = \|z\|$ , so  $\|f\| = \|z\|$ .  $f$  is a bounded linear functional.  $\square$

Riesz's Representation theorem says that this is it. All bounded linear functionals look like this. Thus, if  $H$  is a Hilbert space, its dual  $H'$  is the space of all functions that look like  $f(x) = \langle x, y \rangle$  for some  $y \in X$ . More precisely, if  $H$  is a Hilbert space and  $f \in H'$ , then there is a  $y \in H$  such that, for all  $x \in X$ ,  $f(x) = \langle x, y \rangle$ .

### Riesz's Representation Theorem

**Theorem 32.2.11.** *If  $H$  is a Hilbert space, and  $f : H \rightarrow \mathbb{R}$  is a bounded linear functional, then there is a unique  $y \in H$  such that, for all  $x \in X$ ,  $f(x) = \langle x, y \rangle$ . Moreover,  $\|f\| = \|y\|$ .*

*Proof.* Let  $f \in H'$  and let  $N = \text{nul}(f)$ . That is,  $N$  is the null space of the functional  $f$  which is the set of all points  $x \in X$  such that  $f(x) = 0$ . The Null space actually defines a closed vector space, which is a subspace of  $H$ . If  $N = H$ , then  $f(x) = 0$ , and thus let  $y = 0$ . Otherwise, let  $z$  be a non-zero elements such that, for all  $x \in N$ ,  $\langle x, y \rangle = 0$ . For all  $x, z$ ,  $f(x)z - f(z)x \in N$ , for:

$$f(f(x)z - f(z)x) = f(f(x)z) - f(f(z)x) = f(x)f(z) - f(z)f(x) = 0$$

Therefore:

$$\begin{aligned}\langle f(x)z - f(z)x, z \rangle &= 0 \\ \Rightarrow |f(x)|\|z\|^2 - |fz|\langle x, y \rangle &= 0 \\ \Rightarrow f(x) &= \langle x, \frac{f(z)z}{\|z\|^2} \rangle\end{aligned}$$

Therefore, let  $y = \frac{f(z)}{\|z\|^2}z$ . This is unique since if for all  $x \in H$ ,  $\langle x, y_1 \rangle = \langle x, y_2 \rangle$ , then  $y_1 = y_2$ . Finally:

$$\|y\| = \frac{|f(z)|}{\|z\|^2} \|z\| = \frac{|f(z)|}{\|z\|} \leq \|f\|$$

But also:

$$|f(x)| = |\langle x, z \rangle| \leq \|x\| \|z\|$$

Thus,  $\|f\| \leq \|z\|$ . But  $\|z\| \leq \|f\|$ . Therefore,  $\|f\| = \|z\|$ .  $\square$

Much like in  $\mathbb{R}^n$ , there is a notion of orthogonality in a general inner product space.

**Definition 32.2.10** Orthogonal elements in an inner product space  $X$  are elements  $x, y \in X$  such that  $\langle x, y \rangle = 0$ .

There's also a notion of convexity for a general vector space.

**Definition 32.2.11** A convex subset of a vector space  $V$  space is a subset  $S \subset V$  such that, for all  $x, y \in V$  and for all  $\lambda \in \mathbb{R}$ ,  $\lambda x + (1 - \lambda)y \in S$ .

**Theorem 32.2.12.** *If  $S$  is a subset of  $V$ , then  $S$  is convex.*

Recall that for a general metric space  $X$ , if  $S \subset X$ , we defined  $dist(x, S) = \inf\{d(x, s) : s \in S\}$ . We proved that, if  $S$  is compact, then there is an  $s \in S$  such that  $dist(x, S) = d(x, s)$ . We showed that, without compactness, this may not be true. Indeed, even complete spaces may lack this property. If  $X$  is a Hilbert space, however, this property is guaranteed.

**Theorem 32.2.13.** *If  $H$  is a Hilbert space and if  $S \subset H$  is a closed convex subset of  $H$ , then there is a unique  $s \in S$  such that  $dist(x, S) = \|x - s\|$ .*

*Proof.* As  $dist(x, S) = \inf\{d(x, s) : s \in S\}$ , there is a sequence  $x_n \in S$  such that  $\|x - x_n\| \rightarrow dist(x, S)$ . Then, by Appolonius:

$$\begin{aligned}\|x - x_n\|^2 + \|x - x_m\|^2 &= \frac{1}{2}\|x_n - x_m\|^2 + \frac{1}{2}\left\|\frac{1}{2}(x_n + x_m) - x\right\|^2 \\ &\geq \frac{1}{2}\|x_n + x_m\|^2 + 2dist(x, S)\end{aligned}$$

But  $\|x - x_n\| \rightarrow \text{dist}(x, S)$  and  $\|x - x_m\| \rightarrow \text{dist}(x, S)$ , so:

$$\frac{1}{2}\|x_n - x_m\|^2 \leq \|x - x - m\|^2 + \|x - x_m\|^2 - 2\text{dist}(x, S)^2$$

Which can be made arbitrarily small. Therefore,  $x_n$  is Cauchy. But  $H$  is a Hilbert space, and is therefore complete, and thus  $x_n$  converges. Let  $s$  be the limit. As  $S$  is closed,  $s \in S$ . Moreover, from construction,  $\|x - s\| = \text{dist}(x, S)$ . If there is another point  $v$ , then  $\|x - s\| = \|x - v\|$ . From Appolonius:

$$\|x - s\|^2 + \|x - v\|^2 \geq \frac{1}{2}\|s - v\| + 2\text{dist}(x, S)^2$$

But  $\|x - s\| = \|x - v\| = \text{dist}(x, S)$ , and thus  $\|s - v\| = 0$ . Therefore,  $s = v$ .  $\square$

If  $S$  is a closed subspace of  $H$ , then it's automatically convex. In this case,  $x - s \perp S$ , where  $z \perp S$  means that, for all  $s \in S$ ,  $\langle s, z \rangle = 0$ . For if  $z \in S$ , then  $s + tz \in S$  for all  $t$ . Thus:

$$\|s + tz - x\| \geq \text{dist}(x, S) = \|s - x\|^2$$

And therefore:

$$\langle s - x, s - x \rangle + 2t\langle s - x, z + t^2z, z \geq s - x^2 \Rightarrow t^2\|z\|^2 + 2t\langle s - x, z \geq 0$$

Looking at the discriminant of this polynomial, we have:

$$\langle s - x, z \rangle = 0$$

Therefore,  $s - x \perp S$ . You obtain  $s \in S$  by “dropping the perpendicular of  $x$ ,” onto  $S$ . That is,  $s$  is the orthogonal projection of  $x$  onto  $S$ .  $s = P_S x$  where  $P_S : H \rightarrow H$  is the orthogonal projection. This has a few nice properties:

1. It is idempotent:  $P_S^2 = P_S$ .
2. Self adjoint:  $\langle P_S x, y \rangle = \langle x, P_S y \rangle =$
3. Linear.
4. Bounded and  $\|P_S\| = 1$ .

If  $S$  is a subset of an inner product space  $X$ , we write  $S^\perp = \{x \in X : \langle x, s \rangle = 0\}$ . This is often read aloud as “ $S$  perp” or “ $S$  perpendicular.”

**Theorem 32.2.14.** *If  $S \subset X$ , then  $S^\perp$  is a closed subspace.*

**Theorem 32.2.15.**  $S \subset (S^\perp)^\perp$

The direct sum of two subsets of a Hilbert space  $H$  is  $S_1 \oplus S_2 = \{ax + by : x \in S_1, y \in S_2\}$ .

**Theorem 32.2.16.** *If  $H$  is a Hilbert space and  $S$  is a closed subspace of  $H$ , then  $H = S \oplus S^\perp$*

*Proof.* For  $x = P_S(x) + (x - P_S(x))$ , and thus there is an element in  $S$  and an element in  $S^\perp$  such that  $x$  is the sum of those two elements. This is the only representation. For if  $x = s_1 + s_1^\perp$  and  $x = s_2 + s_2^\perp$ , then stuff.  $\square$

If  $X$  and  $Y$  are normed spaces, and if  $f \in B(X, Y)$ , then  $\{x \in X : f(x) = 0 \in Y\}$  is called the null space of  $f$ .

**Theorem 32.2.17.** *If  $X$  and  $Y$  are normed spaces, and if  $f \in B(X, Y)$ , then  $\text{nul}(f)$  is a closed linear subspace of  $X$ .*

*Proof.* Obvious since  $f$  is linear and continuous.  $\square$

In a Hilbert space  $H$ , then  $H = \text{nul}(f) \oplus \text{nul}(f)^\perp$ . Thus, if  $\text{nul}(f) \neq H$ , then  $\text{nul}(f)^\perp \neq \{0\}$ . That is, there exists a  $z \in \text{nul}(f)^\perp$  that is non-zero. This is the  $z$  we used to prove the Riesz representation theorem. Riesz's Theorem thus says that every Hilbert space is its own dual.

### 32.2.4 Lecture 9: November 5, 2018

#### Adjoint

If  $H$  is a Hilbert space and  $f \in H'$ , then there is a  $z \in H$  such that  $f(x) = \langle x, z \rangle$  for all  $x \in H$ . Moreover,  $\|f\| = \|z\|$ . The adjoint of  $T \in B(H, H)$  is an operator  $T^* : H \rightarrow H$  such that  $\langle Tx, y \rangle = \langle x, T^*y \rangle$ . There is always such an operator for any  $T \in B(H, H)$ .  $T^*$  is also bounded and linear. By Riesz there is a  $z = T^*y$  such that  $f(x) = \langle x, T^*y \rangle$ . Then  $\|T^*y\| = \|z\| = \|f\| \leq \|T\|\|y\|$ . Thus,  $\|T^*\| \leq \|T\|$ . Therefore  $T^* \in B(H, H)$ .  $T^*$  is called the adjoint of  $T$ .

**Example 32.2.8** Consider  $\mathbb{R}^n$  with the usual inner product. Let  $T$  be the matrix  $(T_{ij})$ . Then:

$$(Tx)_i = \sum_{j=1}^n T_{ij}x_j$$

and:

$$\langle Tx, y \rangle = \sum_{i=1}^n (Tx)_i y_i = \sum_{i=1}^n \sum_{j=1}^n T_{ij}x_j y_i = \sum_{j=1}^n \sum_{i=1}^n T_{ij}y_i x_j$$

If  $T^*$  is the adjoint, then:

$$\langle x, T^*y \rangle = \sum_{j=1}^n \left( \sum_{i=1}^n T_{ji}^*y_i \right) x_j$$

And thus  $T_{ji}^* = T_{ij}$ . That is, the adjoint of  $T$  is the transpose of  $T$ . If we were in  $\mathbb{C}^n$  we would use the complex conjugate of the transpose of  $T$ . In

general, if  $T = T^*$  we say that  $T$  is *self-adjoint*. This is also called symmetric or Hermitian.

**Example 32.2.9** As another example, consider  $H = \ell^2$  and let  $T(x_1, x_2, \dots) = (x_2, x_3, \dots)$ . This is linear, and:

$$\|T(x_1, x_2, \dots)\| = \|(x_2, x_3, \dots)\| = \sqrt{\sum_{n=2}^{\infty} x_n^2} \leq \sqrt{\sum_{n=1}^{\infty} x_n^2} = \|(x_1, x_2, \dots)\|$$

Therefore  $T$  is bounded and  $\|T\| \leq 1$ . But  $T(0, 1, 0, 0, \dots) = (1, 0, 0, \dots)$  showing that  $\|T\| \geq 1$ . Thus,  $\|T\| = 1$ . Then, from the definition of  $T$ :

$$\begin{aligned}\langle Tx, y \rangle &= \langle (x_2, x_3, \dots), (y_1, y_2, \dots) \rangle \\ &= x_2 y_1 + x_3 y_2 + \dots \\ &= x_1 \cdot 0 + x_2 y_1 + x_3 y_2 + \dots \\ &= \langle (x_1, x_2, \dots), (0, y_1, y_2, \dots) \rangle\end{aligned}$$

And therefore  $T^*(y_1, y_2, \dots) = (0, y_1, y_2, \dots)$ . Also  $\|T^*\| = 1$ . In general, if  $T \in B(H, H)$  then  $\|T^*\| = \|T\|$ .

**Theorem 32.2.18.** If  $T \in B(H, H)$ , then  $T^{**} = T$ .

**Theorem 32.2.19.**  $\|T\| = \|T^*\|$

*Proof.* For  $\|T\| \leq \|T^*\|$  and  $\|T^*\| \leq \|T^{**}\|$ , but  $T = T^{**}$ , and therefore  $\|T\| = \|T^*\|$ .  $\square$

**Example 32.2.10** Let  $x = C[0, 1]$  and let  $\langle x, y \rangle = \int_0^1 x(t)y(t) dt$ . Let  $K : X \times X \rightarrow \mathbb{R}$  be continuous and define  $T$  by:

$$Tx(t) = \int_0^1 K(t, s)x(s) ds$$

Then for all  $x \in X$ ,  $Tx \in X$  as well, since  $K$  is continuous. Moreover, from Cauchy-Schwarz:

$$\|Tx\|^2 = \int_0^1 \left[ \int_0^1 K(t, s)x(s) ds \right]^2 dt \leq \int_0^1 \left[ \int_0^1 K(t, s)^2 ds \int_0^1 x(s)^2 ds \right] dt$$

But  $\int_0^1 x(s)^2 ds = \|x\|^2$ . So:

$$\|Tx\|^2 \leq \|x\|^2 \int_0^1 \int_0^1 K(t, s) ds dt$$

Therefore  $T$  is bounded and:

$$\|T\| \leq \sqrt{\int_0^1 \int_0^1 K(t, s) ds dt}$$

Computing the adjoint:

$$\begin{aligned} \langle Tx, y \rangle &= \int_0^1 Tx(t)y(t) dt \\ &= \int_0^1 \left( \int_0^1 K(t, s)x(s) ds \right) y(t) dt \\ &= \int_0^1 \left( \int_0^1 K(t, s)y(t) dt \right) x(s) ds \\ &= \int_0^1 \left( \int_0^1 K(s, t)y(s) ds \right) x(t) dt \\ &= \int_0^1 Ty(s)x(s) ds \\ &= \langle Ty, x \rangle \end{aligned}$$

We may swap the order of integration since  $K$  is continuous on a compact set.

**Theorem 32.2.20.**  $\|T^*T\| = \|T\|^2$

*Proof.* For:

$$\|T^*Tx\| \leq \|T^*\| \|Tx\| \leq \|T^*\| \|T\| \|x\|$$

And therefore  $\|T^*T\| \leq \|T\|^2$ . On the other hand:  $\square$

**Theorem 32.2.21.** If  $T$  is self-adjoint, then  $\|T\| = \sup\{|\langle Tx, x \rangle| : \|x\| = 1\}$ .

*Proof.* Let  $\alpha = \sup\{\sup\{|\langle Tx, x \rangle| : \|x\| = 1\}\}$ . Then:

$$|\langle Tx, x \rangle| \leq \|Tx\| \|x\| \leq \|T\| \|x\|^2$$

Taking the supremum over  $\|x\| = 1$ , we have  $\alpha \leq \|T\|$ . But if  $\|x\| = \|y\| = 1$ , then:

$$\begin{aligned} |\langle Tx, y \rangle| &= \left| \frac{1}{4} \langle T(x+y), x+y \rangle - \frac{1}{4} \langle T(x-y), x-y \rangle \right| \\ &= \left| \frac{1}{4} \|x+y\|^2 \langle T \frac{x+y}{\|x+y\|}, \frac{x+y}{\|x+y\|} \rangle - \frac{1}{4} \|x-y\|^2 \langle T \frac{x-y}{\|x-y\|}, \frac{x-y}{\|x-y\|} \rangle \right| \end{aligned}$$

Since  $\langle Tx, y \rangle = \langle x, Ty \rangle$ , as  $T$  is self-adjoint. And from the definition of  $\alpha$ :

$$|\langle Tx, y \rangle| \leq \frac{\alpha}{4} (\|x+y\|^2 + \|x-y\|^2) \leq \frac{\alpha}{4} (2\|x\|^2 + 2\|y\|^2) = \alpha$$

Let  $y = Tx/\|Tx\|$ , we get:

$$\langle Tx, \frac{Tx}{\|Tx\|} \rangle \leq \alpha$$

And therefore  $\|T\| \leq \alpha$ . But also  $\alpha \leq \|T\|$ . Thus,  $\|T\| = \alpha$ .  $\square$

Thm. 32.2.4 can fail if  $T$  is not self-adjoint. In  $\mathbb{R}^2$ , let  $T(x_1, x_2) = (0, x_1)$ . Then:

$$\|Tx\|^2 = x_1^2 \leq x_1^2 + x_2^2 = \|x\|^2$$

And therefore  $\|T\| \leq 1$ . But  $T(1, 0) = (0, 1)$ , and thus  $\|T\| = 1$ . But if  $(x_1, x_2)$  lies on the unit circle, then  $|x_1 x_2| \leq 0.5$ . Thus:

$$|\langle Tx, x \rangle| = |\langle (x_1, x_2), (0, x_1) \rangle| = |x_1 x_2| \leq \frac{1}{2}$$

Therefore  $|\langle Tx, x \rangle| \leq 0.5 < \|T\|$  for all  $x \in \mathbb{R}^2$  such that  $\|x\| = 1$ .

## Compact Operators

Compact operators can be defined in a more general spaces than that of Hilbert or Banach spaces. They can be defined on Topological spaces, but we won't go that far. For now we will simply define them on a general metric space.

**Definition 32.2.12** A compact mapping from a metric space  $X$  to a metric space  $Y$  is a function  $T : X \rightarrow Y$  such that for all bounded subsets  $S$  of  $X$ , the image  $T(S)$  is pre-compact in  $Y$ . That is,  $\overline{T(S)}$  is compact (The closure of  $T(S)$  is compact).

**Theorem 32.2.22.** *If  $T : X \rightarrow Y$  is a linear compact operator between normed spaces  $X$  and  $Y$ , then  $T$  is continuous.*

*Proof.* For let  $S = \overline{B_1(\mathbf{0})}$ . This is bounded, so  $\overline{T(S)}$  is compact, and therefore bounded. Let  $M$  be such a bound. Thus, for all  $s \in \overline{S}$  such that  $\|s\| = 1$ ,  $\|Ts\| \leq M$ , and therefore  $\|T\| \leq M$ . Thus  $T$  is bounded and linear, and is therefore continuous.  $\square$

**Example 32.2.11** Every linear mapping  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is compact. As a another example, let  $X = C[0, 1]$  and equip this with the supremum norm. Define  $T$  as:

$$Tx(t) = \int_0^1 K(t, s)x(s) \, ds$$

Where  $K : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$  is continuous. This is a compact operator. For if  $S$  is a bounded subset then there exists an  $M$  such that for all  $x \in S$ ,  $\|x\| \leq M$ .

Thus:

$$\begin{aligned}\|Tx\| &= \sup |Tx(t)| = \sup \left| \int_0^1 K(t, s)x(s) ds \right| \\ &\leq \sup \int_0^1 |K(t, s)||x(s)| ds \\ &\leq \kappa \int_0^1 |x(s)| ds \\ &\leq \kappa \|x\|\end{aligned}$$

Where  $\kappa = \sup |K(t, s)|$ .  $\kappa$  exists since  $K(t, s)$  is continuous on a compact set and is therefore bounded. So  $T(S)$  is uniformly bounded. To apply Arzela-Ascoli we need to show that  $T(S)$  is equicontinuous. That is, for all  $\varepsilon > 0$  there is a  $\delta > 0$  such that, for all  $x \in S$ , if  $|t_2 - t_1| < \delta$  then  $|Tx(t_2) - Tx(t_1)| < \varepsilon$ . If we can show that  $T$  satisfies this, then  $\overline{T(S)}$  is compact, and thus  $T$  is compact. Let's show this. If  $x \in S$ , then:

$$\begin{aligned}|Tx(t_2) - Tx(t_1)| &= \left| \int_0^1 K(t_1, s)x(s) ds - \int_0^1 K(t_2, s)x(s) ds \right| \\ &= \left| \int_0^1 (K(t_2, s) - K(t_1, s))x(s) ds \right| \\ &\leq \int_0^1 |K(t_2, s) - K(t_1, s)||x(s)| ds \\ &\leq M \int_0^1 |K(t_2, s) - K(t_1, s)| ds\end{aligned}$$

But as  $K$  is uniformly continuous, there is a  $\delta > 0$  such that, for all  $s \in [0, 1]$ ,  $|t_2 - t_1| < \delta$  implies  $|K(t_2, s) - K(t_1, s)| < \varepsilon/M$ . Thus,  $T(S)$  is equicontinuous. We can replace the supremum norm with  $L^2$  and  $T$  is still compact. Indeed, it is true for  $L^p$  if we replace the use of Cauchy-Schwarz with the more general Hölder's Inequality. From this we have that  $T$  is a compact self-adjoint operator.

### 32.2.5 Lecture 10: November 19, 2018

#### Compact Linear Operators

A linear operator  $T : X \rightarrow Y$  is compact if  $\overline{T(S)}$  is compact for all bounded  $S \subset X$ . Example,  $Tx(t) = \int_0^1 K(t, s)x(s) ds$ , where  $K$  is continuous on  $[0, 1]^2$ , is a compact operator  $T : C[0, 1] \rightarrow C[0, 1]$ . If  $K(x, t) = K(t, x)$  for all  $(x, t) \in [0, 1]^2$ , then  $T$  is a self-adjoint operator on  $L^2[0, 1]$ .  $L^2[0, 1]$  can be seen as the completion of  $C[0, 1]$  with respect to the  $L^2$  norm.

**Theorem 32.2.23.** A linear operator  $T : X \rightarrow Y$  is compact if and only if for all bounded sequences  $x : \mathbb{N} \rightarrow X$ ,  $Tx$  has a convergent subsequence in  $Y$

We're interested mainly in the case of  $Y = X$ , and when  $T : X \rightarrow X$  is bounded and linear. This is the set of all operators  $B(X, X)$ . We rewrite this as  $B(X)$ . That is,  $B(X)$  is the set of all bounded linear operators from  $X$  to itself. Recall that if  $Y$  is a Banach space, and if  $X$  is a normed space, then  $B(X, Y)$  is a Banach space. Thus, if  $X$  is a Banach space, then  $B(X)$  is a Banach space. But we can also multiply elements in  $B(X)$  by using function composition. If  $S, T \in B(X)$ , then  $ST$  is defined by  $(ST)(x) = S(Tx)$ . But then:

$$\|(ST)x\| = \|S(Tx)\| \leq \|S\|\|Tx\| \rightarrow \|ST\| \leq \|S\|\|T\|$$

A Banach space with such a multiplication property is called a Banach Algebra. The set of compact linear operators on  $X$  is often denoted  $C(X)$ . Thus,  $C(X) \subset B(X)$ . It's a two-sided closed ideal in  $B(X)$ . That is, if  $S, T \in C(X)$ , and if  $a$  and  $b$  are scalars, then  $aS + bT \in C(X)$ ,  $ST \in C(X)$ , and  $TS \in C(X)$ . Finally, if  $F : \mathbb{N} \rightarrow C(X)$  is a sequence of compact operators, and if  $F_n \rightarrow T$ , then  $T \in C(X)$ .

**Definition 32.2.13** Orthogonal Elements in an inner product space  $(X, \langle \cdot, \cdot \rangle)$  are elements  $x, y \in X$  such that  $\langle x, y \rangle = 0$ .

**Definition 32.2.14** An orthonormal subset of an inner product space  $(X, \langle \cdot, \cdot \rangle)$  is a subset  $S \subset X$  such that for all  $x, y \in S$  such that  $x \neq y$ ,  $\langle x, y \rangle = 0$  and for all  $x \in S$ ,  $\|x\| = 1$ .

**Theorem 32.2.24.** If  $x \in X$  and  $\varphi : \mathbb{N} \rightarrow X$  is a sequence such that  $A = \{\varphi_n : n \in \mathbb{N}\}$  is an orthonormal subset of  $X$ , then for all  $x \in X$ :

$$\|x\| = \sqrt{\sum_{n=1}^N \langle x, \varphi_n \rangle^2 + \|x - \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n\|^2}$$

*Proof.* If  $m \in \mathbb{Z}_N$ , then:

$$\langle x - \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n, \varphi_m \rangle = \langle x, \varphi_m \rangle - \sum_{n=1}^N \langle x, \varphi_n \rangle \langle \varphi_n, \varphi_m \rangle$$

But  $A$  is an orthonormal subset of  $X$ , and thus if  $n \neq m$  then  $\langle \varphi_n, \varphi_m \rangle = 0$  and if  $n = m$  then  $\langle \varphi_n, \varphi_m \rangle = \|\varphi_n\| = 1$ . In terms of the Kronecker-Delta function,  $\langle \varphi_n, \varphi_m \rangle = \delta_{nm}$ . So we have:

$$\langle x - \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n, \varphi_m \rangle = 0$$

But for all  $N \in \mathbb{N}$ ,  $x = (x - \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n) + \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n$ , and these two are orthogonal. Therefore, from Pythagoras:

$$\begin{aligned}\|x\|^2 &= \|x - \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n\|^2 + \left\| \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n \right\|^2 \\ &= \sum_{n=1}^N \langle x, \varphi_n \rangle \|\varphi_n\|^2 + \left\| \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n \right\|^2\end{aligned}$$

But  $\|\varphi_n\|^2 = 1$  for all  $n$ . Therefore, etc.  $\square$

**Theorem 32.2.25** (Bessel's Inequality).  $\sum_{n=1}^N \langle x, \varphi_n \rangle \leq \|x\|^2$

**Example 32.2.12**  $A = \{e_n : n \in \mathbb{N}\}$  is an orthonormal subset of  $\ell^2$ .  $A = \{\sin(nt)/\sqrt{\pi} : n \in \mathbb{N}\}$  is an orthonormal subset of  $C[0, 1]$  with the  $L^2$  inner product.

**Definition 32.2.15** A basis of an inner product space  $X$  is an orthonormal subset  $A \subset X$  such that there is no orthonormal subset  $B \subset X$  such that  $A \subset B$ .

**Theorem 32.2.26.** If  $(X, \langle \cdot, \cdot \rangle)$  is an inner product space, then there is an  $A \subset X$  such that  $A$  is an orthonormal subset of  $X$ .

**Theorem 32.2.27.** If  $X$  is an inner product space, and if  $\varphi : \mathbb{N} \rightarrow X$  is a sequence such that  $S = \{\varphi_n : n \in \mathbb{N}\}$  is a basis of  $X$ , then for all  $x \in X$  there is a sequence  $a : \mathbb{N} \rightarrow \mathbb{R}$  such that  $x = \sum_{n=1}^{\infty} a_n \varphi_n$

### Summability

What does  $\sum_{\alpha \in A} z_{\alpha}$  if  $z_{\alpha} \in \mathbb{R}$  for all  $\alpha \in A$ ?

**Definition 32.2.16** We say  $\sum_{\alpha \in A} z_{\alpha}$  is summable to  $z \in \mathbb{R}$  if for all  $\varepsilon > 0$  there is a subset  $B \subset A$  such that, for all  $C \subset A$  such that  $B \subset C$ ,  $|\sum_{\alpha \in C} z_{\alpha} - z| < \varepsilon$ .

It turns out that, if  $\sum_{\alpha \in A} z_{\alpha}$  is summable, then only countably many  $z_{\alpha}$  are non-zero. For all  $n$ , the set  $\{z_{\alpha} : z_{\alpha} > 1/n\}$  must be finite. The set of all non-zero elements is the union over all of these  $n$ , which is the countable union of finite sets, which is thus countable.

**Theorem 32.2.28.** If  $(X, \langle \cdot, \cdot \rangle)$  is an inner product space, if  $\varphi : \mathbb{N} \rightarrow X$  is a sequence such that  $S = \{\varphi_n : n \in \mathbb{N}\}$  is a basis of  $X$ , then:

$$x = \sum_{n=1}^{\infty} \langle x, \varphi_n \rangle \varphi_n$$

**Example 32.2.13** If  $L^2[0, \pi]$ , let  $\varphi_n(t) = \sin(nt)\sqrt{2/\pi}$ . Then  $A = \{\varphi_n(t) : n \in \mathbb{N}\}$ .

### 32.2.6 Lecture 11: November 26, 2018

Let  $T$  be a linear operator on a vector space  $T$ . We say  $\lambda \in \mathbb{C}$  is an eigenvalue of  $T$  if there exists  $x \neq 0$  in  $X$  such that  $Tx = \lambda x$ . The corresponding  $x$  is called the eigenvector or eigenfunction. We're interested in the case of compact self-adjoint operators  $T$  on a Hilbert space  $\mathcal{H}$ .

**Theorem 32.2.29.** *There is a sequence of real eigenvalues  $\lambda_n$  of  $T$ , finite or infinite, such that 0 is the only possible accumulation point of  $\lambda_n$ , and corresponding basis of orthonormal eigenvectors  $x_n$ .*

*Proof.* We'll prove this in steps. First, either  $\|T\|$  or  $-\|T\|$  is an eigenvalue. This is because:

$$\|T\| = \sup_{\|x\|=1} \{|\langle Tx, x \rangle|\}$$

This comes from the fact that  $T = T^*$  for self-adjoint operators. Thus, either  $\|T\| = \langle Tx, x \rangle$  or  $-\|T\| = \langle Tx, x \rangle$ . Choose a sequence  $x_n$  such that  $\|x_n\| = 1$  and  $\langle Tx_n, x_n \rangle \rightarrow \pm \|T\|$ . By choosing a subsequence, we may assume that  $Tx_n$  converges. We can not assume that  $x_n$  converges, however. Thus,  $Tx_n \rightarrow y$ . Then:

$$\begin{aligned} \|Tx_n - \lambda x_n\|^2 &= \langle Tx_n - \lambda x_n, Tx_n - \lambda x_n \rangle \\ &= \|Tx_n\|^2 - 2\lambda \langle Tx_n, x \rangle + \lambda^2 \|x\|^2 \\ &\leq \|T\|^2 \|x\|^2 - 2\lambda \langle Tx_n, x \rangle + \lambda^2 \|x\|^2 \end{aligned}$$

And this converges to zero as  $n$  tends to infinity. Thus,  $Tx_n - \lambda x_n \rightarrow 0$ . It follows that  $\lambda x_n = Tx_n - (Tx_n - \lambda x_n)$ , and this converges to  $y$ . Therefore  $x_n \rightarrow y/\lambda$ . Note that  $\lambda$  is only equal to zero if  $T$  is the zero operator. In this case the problem is trivial. Thus we may assume  $\lambda \neq 0$ . Therefore  $Tx_n \rightarrow Ty/\lambda$ , but  $Tx_n \rightarrow y$  as well. Thus  $y = Ty/\lambda$ . Let  $\lambda_1 = \lambda$  and  $\varphi_1 = y/\|y\|$ . Then  $\|\varphi_1\| = 1$  and  $T\varphi_1 = \lambda_1\varphi_1$ . Moreover, let  $T_1 = T$  and let  $H_1 = \mathcal{H}$ . Let  $H_2 = \{x \in H_1 : x \perp \varphi_1\}$ . Define  $T_2 : H_2 \rightarrow H_1$  by  $T_2x = T_1x = Tx$ . This is the restriction of  $T$  to  $H_2$ . Then  $T_2H_2 \subseteq T_2H_2$ . For if  $x \in H_2$ , then  $\langle T_2x, \varphi_1 \rangle = \langle T_1x, \varphi_1 \rangle$ . But  $T$  is self-adjoint, and  $T_1 = T$ , and therefore  $T_1$  is self-adjoint. But then:

$$\langle T_2x, \varphi_1 \rangle = \langle T_1x, \varphi_1 \rangle = \langle x, T_1\varphi_1 \rangle = \lambda_1 \langle x, \varphi_1 \rangle$$

Therefore  $T_2x \in H_2$ , and therefore  $T_2 : H_2 \rightarrow H_2$ .  $T_2$  is self-adjoint, for:

$$\langle T_2x, y \rangle = \langle T_1x, y \rangle = \langle x, T_1y \rangle = \langle x, T_2y \rangle$$

$T_2$  is compact since if  $x_n$  is bounded in  $H_2$ , then it's bounded in  $H_1$ , and thus  $T_2x_n = T_1x_n$ , which has a convergent subsequence, and therefore  $T_2$  is

compact.  $H_2$  is a subspace of  $\mathcal{H}$  and is closed since  $x_n \in H_2$  and  $x_n \rightarrow x$  in  $\mathcal{H}$  then:

$$\langle x, \varphi_1 \rangle = \langle \lim x_n, \varphi_1 \rangle = \lim \langle x_n, \varphi_1 \rangle = 0$$

Thus  $H_2$  is closed and is therefore complete, and thus  $H_2$  is a Hilbert space. As before there is a  $\varphi_2$  and a  $\lambda_2$  such that  $\varphi_2 \in H_2$ ,  $\|\varphi_2\| = 1$ , and:

$$T_2 \varphi_2 = \lambda_2 \varphi_2$$

But then  $T\varphi_2 = \lambda_2 \varphi_2$ , where  $\lambda_2 = \pm \|T_2\|$ . Moreover,  $|\lambda_2| < |\lambda_1|$ . Continuing in this manner, let  $H_n = \{x \in \mathcal{H} : x \perp \varphi_1, \dots, \varphi_{n-1}\}$  and  $T_n : H_n \rightarrow H_n$  be the restriction of  $T$  onto  $H_n$ . We obtain a  $\varphi_n$  such that  $\|\varphi_n\| = 1$  and  $T_n \varphi_n = \lambda_n \varphi_n$ . Moreover  $|\lambda_n| \leq |\lambda_{n-1}|$ . Thus,  $|\lambda_n|$  forms a monotonically decreasing sequence and either there is an  $N \in \mathbb{N}$  such that  $\lambda_N = 0$ , in which case for all  $n > N$ ,  $\lambda_n = 0$  as well, or for all  $n \in \mathbb{N}$ ,  $|\lambda_n| > 0$ . In the first case it is clear that  $\lambda_n \rightarrow 0$ . In the second case we have  $\lambda_n \varphi_n = T_n \varphi_n$  has a convergent subsequence for  $T$  is compact. But  $|\lambda_n|$  is a monotonically decreasing sequence bounded below by zero, and therefore converges. Let  $c$  be the limit. Then  $\lambda_n x_n \rightarrow cx$ . Moreover  $\|\varphi_n - \varphi_m\|^2 = 2$  since  $\varphi_n$  and  $\varphi_m$  are orthogonal when  $n \neq m$ . Therefore  $\varphi_n$  is not a Cauchy sequence. Thus, for  $\lambda_n \varphi_n$  to converge,  $c = 0$ .  $\square$

**Theorem 32.2.30** (Hilbert-Schmidt Theorem). *If  $x \in \mathcal{H}$ , then:*

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, \varphi_n \rangle \varphi_n$$

*Proof.* For define  $y_m$  as:

$$y_m = x - \sum_{n=1}^{m-1} \langle x, \varphi_n \rangle \varphi_n$$

Then:

$$\langle y_m, \varphi_k \rangle = \langle x - \sum_{n=1}^{m-1} \langle x, \varphi_n \rangle \varphi_n, \varphi_k \rangle = \langle x, \varphi_k \rangle - \sum_{n=1}^{m-1} \langle x, \varphi_n \rangle \langle \varphi_n, \varphi_k \rangle = 0$$

If  $\lambda_N = 0$ , then  $Ty = T_n y = 0$  by setting  $m = N$ . Thus:

$$0 = Ty = Tx - \sum_{n=1}^{N-1} \langle x, \varphi_n \rangle T \varphi_n = Tx - \sum_{n=1}^{N-1} \lambda_n \langle x, \varphi_n \rangle \varphi_n$$

If  $\lambda_n \neq 0$  for all  $n \in \mathbb{N}$ , then  $y_m \in H_m$  and therefore:

$$x = (x - y_m) + y_m$$

But  $x - y_m$  is orthogonal to  $y_m$ , and therefore by Pythagoras:

$$\|x\|^2 = \|x - y_m\|^2 + \|y_m\|^2$$

Therefore  $\|y_m\| \leq \|x\|$ . Also:

$$\|Ty_m\| = \|T_m y_m\| \leq \|T_m\| \|y_m\| = |\lambda_m| \|y_m\|$$

And therefore:

$$\left\| Tx - \sum_{n=1}^{m-1} \lambda_n \langle x, \varphi_n \rangle \varphi_n \right\| \leq |\lambda_m| \|x\| \rightarrow 0$$

□

**Theorem 32.2.31.** *If  $T$  is a compact self-adjoint operator on a Hilbert Space  $\mathcal{H}$ , then there is an orthogonal basis for  $\mathcal{H}$  consisting of eigenvector of  $T$ .*

*Proof.* For any  $x \in \mathcal{H}$ :

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, \varphi_n \rangle \varphi_n$$

This sum may be infinite. Let  $\{\psi_\alpha\}_{\alpha \in A}$  be and orthogonal basis of  $\text{nul}(T)$ . Then  $T\psi_\alpha = 0$  for all  $\alpha \in A$ , and thus 0 is an eigenvalue for all  $\psi_\alpha$ . But also:

$$\lambda_n \langle \varphi_n, \psi_\alpha \rangle = \langle \lambda_n \varphi_n, \psi_\alpha \rangle = \langle T\varphi_n, \psi_\alpha \rangle = \langle \varphi_n, T\psi_\alpha \rangle = 0$$

Thus, for all  $\alpha \in A$  and all  $n \in \mathbb{N}$ ,  $\varphi_n \perp \psi_\alpha$ . Then by Hilbert-Schmidt, for every  $x \in \mathcal{H}$ :

$$T\left(x - \sum_{n=1}^{\infty} \langle x, \varphi_n \rangle \varphi_n\right) = Tx - \sum_{n=1}^{\infty} \lambda_n \langle x, \varphi_n \rangle \varphi_n = 0$$

Thus:

$$x = \sum_{n=1}^{\infty} \lambda_n \langle x, \varphi_n \rangle \varphi_n + \sum_{\alpha \in A} \langle x, \psi_\alpha \rangle \psi_\alpha$$

□

## 32.3 More Stuffs

### 32.3.1 Lecture 12: December 3, 2018

Cauchy-Schwarz says that:

$$\left( \int_a^b x(s)y(s) \, ds \right)^2 \leq \left( \int_a^b x^2(s) \, ds \right) \left( \int_a^b y^2(s) \, ds \right) \quad (32.3.1)$$

So:

$$|Tx(t)|^2 = \left( \int_0^t x(s) \, ds \right)^2 \quad (32.3.2)$$

$$\leq \left( \int_0^t 1^2 \, ds \right) \left( \int_0^t x(s)^2 \, ds \right) \quad (32.3.3)$$

$$= t \int_0^t x(s)^2 \, ds \quad (32.3.4)$$

$$\leq t \int_0^1 x(s)^2 \, ds \quad (32.3.5)$$

$$= t \|x\|_2^2 \quad (32.3.6)$$

So then:

$$\int_0^1 Tx(t)^2 \, dt \leq \int_0^1 t \|x\|_2^2 \, dt \quad (32.3.7)$$

This then implies:

$$\|Tx\|^2 \leq \frac{1}{2} \|x\|_2^2 \quad (32.3.8)$$

$$\Rightarrow \|T\| \leq \frac{1}{\sqrt{2}} \quad (32.3.9)$$

Letting  $x(t) = 1$ , we have  $Tx = t$ . Thus:

$$\|Tx\|^2 = \int_0^1 t^2 \, dt = \frac{1}{3} \quad (32.3.10)$$

And thus  $\|Tx\| = 1/\sqrt{3}$ . So  $1/\sqrt{3} \leq \|T\| \leq 1/\sqrt{2}$ . If  $x(t) = 1 - t$ , then  $Tx(t) = t - t^2/2$ . So  $\|Tx\| = \sqrt{2/15}$ . We're getting closer to the answer. Letting  $x(t) = \cos(\pi t/2)$  gives us the norm. We now have to show this. Write:

$$x(t) = \sum_{n=1}^{\infty} b_n \sin(n\pi t)$$

Then:

$$Tx(t) = \sum_{n=1}^{\infty} \frac{b_n}{n\pi} [1 - \cos(n\pi t)] = \left( \sum_{n=1}^{\infty} \frac{b_n}{n\pi} \right) - \left( \sum_{n=1}^{\infty} \frac{b_n}{n\pi} \cos(n\pi t) \right)$$

But:

$$\|x\|^2 = \sum_{n=1}^{\infty} \frac{b_n^2}{2}$$

$$\|Tx\|^2 = \left( \sum_{n=1}^{\infty} \frac{b_n}{n\pi} \right)^2 + \sum_{n=1}^{\infty} \frac{b_n^2}{2n^2\pi^2}$$

Let's maximize  $\|Tx\|^2$  subject to  $\|x\|^2$ . Using Lagrange multipliers we get:

$$\frac{2}{n\pi} \left( \sum_{k=1}^{\infty} \frac{b_k}{k\pi} \right) + \frac{b_n}{n^2\pi^2} = \lambda^2 b_n$$

Letting  $A = \sum_{k=1}^{\infty} b_k/k\pi$ , we obtain:

$$b_n = \frac{2n\pi}{\lambda^2 n^2 \pi^2 - 1} A$$

So then:

$$A = \sum_{n=1}^{\infty} \frac{2}{\lambda^2 n^2 \pi^2 - 1} A$$

And thus:

$$\sum_{n=1}^{\infty} \frac{2}{\lambda^2 n^2 \pi^2 - 1} = 1$$

And this is the expansion of cotangent:

$$1 - \frac{1}{\lambda} \cot\left(\frac{1}{\lambda}\right) = 1$$

And therefore:

$$\lambda = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \dots$$

Finally:

$$x(t) = \sum_{n=1}^{\infty} b_n \sin(n\pi t) = \sum_{n=1}^{\infty} \frac{2n\pi}{4n^2 - 1} A \sin(n\pi t) = \sqrt{2} \cos\left(\frac{\pi}{2}t\right)$$

But before we can do all of this we need to show that there is such a maximum. That is, the step that involves Lagrange multipliers is valid. We want an  $x$  such that  $\|x\| = 1$  and  $\|Tx\| = \|T\|$ . Certainly there is a sequence such that  $\|x_n\| = 1$  and  $\|Tx_n\| \rightarrow \|T\|$ . Since  $T$  is compact we may assume, taking subsequences as necessary, that  $Tx_n \rightarrow y$ . If  $T$  is self adjoint and  $x_n \rightarrow x$ , then  $Tx = y$ . It would be nice if Bolzano-Weierstrass worked and we could say  $\|x_n\|$  bounded implies that  $x_n \rightarrow x$ , but this is not always true in infinite dimensions. We'll need to weaken our notion of convergence for this. We could simply say that everything converges to everything, which is the chaotic topology, but this is not very useful for it loses uniqueness.

**Definition 32.3.1** A weakly convergent sequence in an inner product space  $X$  is a sequence  $x : \mathbb{N} \rightarrow X$  such that there is an  $a \in X$  such that for all  $z \in X$ ,  $\langle x_n, z \rangle \rightarrow \langle a, z \rangle$ . We write  $x_n \xrightarrow{w} a$

For example, if  $X$  is a Hilbert space and  $e_n$  is an orthonormal basis, then  $e_n \rightarrow 0$  since, for all  $z$ :

$$\|z\|^2 = \sum_{n=1}^{\infty} \langle e_n, z \rangle^2$$

And thus  $\langle e_n, z \rangle \rightarrow 0$ . But  $\langle 0, z \rangle = 0$ , so  $e_n \xrightarrow{w} 0$ . Normal convergence is also called strong convergence. Strong convergence implies weak convergence. For if  $x_n \rightarrow a$ , then  $\langle a - x_n, z \rangle \rightarrow 0$  for all  $z$ , and thus  $x_n \xrightarrow{w} a$ . Moreover, weak limits are unique. If  $T$  is a bounded linear operator on a Hilbert space  $H$ , and if  $x_n$  converges weakly to  $a$ , then  $Tx_n \xrightarrow{w} Ta$ . If  $x$  converges weakly to  $a$  and if  $T$  is a compact linear operator on  $H$ , then  $Tx_n$  converges strongly to  $Tx$ .

**Theorem 32.3.1.** *If  $H$  is a Hilbert space,  $T$  is a compact operator on  $H$ , and if  $x : \mathbb{N} \rightarrow H$  is a weakly convergent sequence such that  $x_n \xrightarrow{w} a$ , then the sequence  $y : \mathbb{N} \rightarrow H$  defined by  $y_n = Tx_n$  is such that  $y_n \rightarrow Tx$ . That is,  $Tx_n$  converges strongly to  $Ta$ .*

*Proof.* For suppose not. As  $T$  is compact and  $x_n$  is bounded, there is a convergent subsequence. Let  $a_1$  be the limit. If the limit is not unique then there is another convergent subsequence with a different limit  $a_2$ . But  $Tx_n$  converges weakly to  $a$ , and strong convergence implies weak convergence. Therefore  $a_1 = a_2 = a$ , a contradiction. Therefore,  $Tx_n$  converges strongly to  $Ta$ .  $\square$

**Theorem 32.3.2.** *If  $x_n \xrightarrow{w} a$  and  $\|x_n\| \rightarrow C$ , then  $\|a\| \leq C$ .*

*Proof.* For:

$$\begin{aligned} \|a\|^2 &= |\langle a, a \rangle| \\ &= \lim_{n \rightarrow \infty} |\langle x_n, a \rangle| \\ &\leq \lim_{n \rightarrow \infty} \|x_n\| \|a\| \\ &= C \|a\| \end{aligned}$$

Dividing by  $\|a\|$  gives the result.  $\square$

We now prove that there exists  $x$  such that  $\|x\| = 1$  and  $\|Tx\| = \|T\|$ . This works for any compact linear operator on a Hilbert space. We can always find a sequence, by definition, such that  $\|x_n\| = 1$  and  $\|Tx_n\| \rightarrow \|T\|$ . But since  $x_n$  is bounded by 1 there is a weakly convergent subsequence (Still to be proved). This is “Bolzano-Weierstrass,” of infinite dimensions. Then  $x_n$  converges weakly to  $x$ , and thus  $Tx_n$  converges strongly to  $Tx$ . Then  $\|Tx\| = \|T\|$ . Finally,  $\|x\| \leq \lim \|x_n\| = 1$ , and  $\|T\| = \|Tx\| \leq \|T\| \|x\|$ , so  $\|x\| \geq 1$ , and therefore  $\|x\| = 1$ . Let’s show  $T : L^2 \rightarrow L^2$  defined by  $Tx(t) = \int_0^t x(s) ds$  is compact.

Suppose  $x_n$  is bounded in  $L^2$  with bound  $M$ . That is,  $\|x_n\| \leq M$ . Then:

$$\begin{aligned} |Tx_n(t)|^2 &= \left| \int_0^t x_n(s)^2 ds \right|^2 \\ &\leq \left( \int_0^1 |x_n(s)| ds \right)^2 \\ &\leq \left( \int_0^1 ds \right) \left( \int_0^1 |x_n(s)|^2 ds \right) \\ &= \|x_n\|^2 \\ &\leq M^2 \end{aligned}$$

Taking the supremum over  $t \in [0, 1]$  gives:

$$\|Tx_n\|_\infty \leq M$$

So  $Tx_n$  is bounded in  $C[0, 1]$ . Also, if  $0 \leq t_1$  and  $t_2 \leq 1$ , then:

$$\begin{aligned} |Tx_n(t_2) - Tx_n(t_1)|^2 &\leq \left( \int_{t_1}^{t_2} |x_n(s)| ds \right)^2 \\ &\leq \left( \int_{t_1}^{t_2} ds \right) \left( \int_{t_1}^{t_2} |x_n(s)|^2 ds \right) \\ &= (t_2 - t_1) \int_{t_1}^{t_2} |x_n(s)|^2 ds \\ &\leq (t_2 - t_1) \|x_n\|^2 \\ &\leq M^2(t_2 - t_1) \end{aligned}$$

So  $|Tx_n(t_2) - Tx_n(t_1)|$  can be made arbitrarily small for  $t_2$  and  $t_1$  close enough, independent on the  $n$ . That is, a  $\delta$  may be chosen independent of  $n$ . This is the criterion for equicontinuity. The compactness of  $[0, 1]$  then gives uniform equicontinuity. Arzela-Ascoli then says there is a subsequence  $Tx_n \rightarrow y$ , with  $y \in C[0, 1]$ . That is,  $\|Tx - y\|_\infty \rightarrow 0$ . But then:

$$\begin{aligned} \|Tx_n - y\|^2 &= \int_0^1 |Tx_n(t) - y(t)|^2 dt \\ &\leq \int_0^1 \|Tx_n - y\|_\infty^2 dt \\ &= \|Tx_n - y\|_\infty^2 \end{aligned}$$

And this converges to zero. Thus,  $Tx_n \rightarrow y$ .

**Theorem 32.3.3** (Baire Category Theorem). *If  $(X, d)$  is a complete metric and  $C_n$  is a sequence of closed sets such that  $X = \cup_{n=1}^\infty C_n$ , then there is an  $N \in \mathbb{N}$  such that  $C_N$  contains an open subset.*

*Proof.* Let  $r_1 \in (0, 1)$ ,  $x_1 \in X$ . If  $B_{r_1}(X_1) \subset C_1$  then we're done. Otherwise  $B_{r_1}(x_1) \setminus C_1$  is a non-empty set, so there exists  $x_n \in B_{r_2}(x_1)$ , where  $r_2 \in (0, 1/2)$ . By induction, choose  $r_n \in (0, 1/n)$  and  $x_n$  such that  $x_n \in B_{r_{n-1}}(x_{n-1})$  and  $\overline{B_{r_n}(x_n)} \subset B_{r_{n-1}}(x_{n-1})$ . For  $n < m$ ,  $x_m \in \overline{B_{r_n}(x_n)}$ , so  $d(x_n, x_m) < 1/n$ . Then  $x_n$  is Cauchy, but  $X$  is complete so there is a limit  $x$ . Since  $\overline{B_{r_n}(x_n)}$  is closed,  $x \in \overline{B_{r_n}(x_n)}$ . But  $X = \cup_{n=1}^{\infty} C_n$  and thus there is an  $N$  such that  $x \in N$ . But then  $B_{r_N}(x) \subset C_N$ , so  $C_N$  contains an open subset.  $\square$

The Baire Category Theorem is used to prove the Uniform Boundedness Theorem, which is also called the Banach-Steinhau theorem.

**Theorem 32.3.4** (Uniform Boundedness Theorem). *If  $H$  is a Hilbert space, and if  $x_n \xrightarrow{w} a$ , then  $\|x_n\|$  is bounded.*

*Proof.* Let  $C_k = \{y \in H : |\langle x_n, y \rangle| \leq k\}$ . Then  $C_k$  is closed since  $y_j \in C_k$  and  $y_j \rightarrow y$  implies that:

$$\begin{aligned} |\langle x_n, y_j \rangle| &\leq k \\ \Rightarrow \lim_{j \rightarrow \infty} |\langle x_n, y_j \rangle| &\leq k \\ \Rightarrow |\langle x_n, y \rangle| &\leq k \end{aligned}$$

Moreover  $H = \cup_{k=1}^{\infty} C_k$ . By the Baire Category Theorem there is a  $k \in \mathbb{N}$  such that  $C_k$  contains an open subset. Let  $z_0 \in H$  and  $r \in \mathbb{R}$  be such that  $B_r(z_0) \subset C_k$ . Let  $y \in H$ ,  $y \neq 0$ , and set  $z = z_0 + \alpha y$ . where:

$$\alpha = \frac{r}{2\|y\|}$$

Then  $\|z - z_0\| < r$ , so  $z \in C_k$ . That is,  $|\langle x_n, z \rangle| \leq k$  for all  $n$ . Thus we have:

$$\begin{aligned} |\langle x_n, y \rangle| &= \left| \langle x_n, \frac{z - z_0}{\alpha} \rangle \right| \\ &\leq \frac{1}{\alpha} \left( |\langle x_n, z \rangle| + |\langle x_n, z_0 \rangle| \right) \\ &\leq \frac{1}{\alpha} (k + k) \\ &= \frac{4k}{r} \|y\| \end{aligned}$$

This is true of any  $y \in H$ . Choosing  $y = x_n$ , we get:

$$\|x_n\|^2 \leq \frac{4k}{r} \|x_n\|$$

Dividing by  $\|x_n\|$  shows boundedness.  $\square$

### 32.3.2 Lecture 13: December 10, 2018

$$Tx(t) = \int_0^{1-t} x(s) \, ds \quad (32.3.11)$$

If  $\|x\|_2 \leq M$ , then:

$$|Tx(t)|^2 = \left| \int_0^{1-t} x(s) \, ds \right|^2 \leq \left( \int_0^{1-t} 1 \, ds \right) \left( \int_0^{1-t} x(s)^2 \, ds \right) = (1-t)\|x\|_2^2 \quad (32.3.12)$$

So we have:

$$\|Tx\|^2 = \int_0^1 Tx(t)^2 \, dt \leq \int_0^1 (1-t)\|x\|_2^2 \, dt = \frac{1}{2}\|x\|_2^2 \leq \frac{1}{2}M^2 \quad (32.3.13)$$

So:

$$|Tx(t_1) - Tx(t_2)|^2 = \left| \int_0^{1-t_1} x(s) \, ds - \int_0^{1-t_2} x(s) \, ds \right|^2 \quad (32.3.14)$$

$$= \left| \int_{1-t_2}^{1-t_1} x(s) \, ds \right|^2 \quad (32.3.15)$$

$$\leq \int_{1-t_2}^{1-t_1} 1 \, ds \int_{1-t_2}^{1-t_1} x(s)^2 \, ds \quad (32.3.16)$$

$$= |t_1 - t_2| \|x\|_2^2 \quad (32.3.17)$$

$$\leq M^2 |t_1 - t_2| \quad (32.3.18)$$

This shows equicontinuity, and thus Arzela-Ascoli shows that  $T$  is compact.

**Theorem 32.3.5** (Banach-Alaoglu-Hilbert Theorem). *If  $H$  is a Hilbert space and  $x : \mathbb{N} \rightarrow H$  is a bounded sequence, then there is a weakly convergent subsequence.*

The next question would be “What about Banach Space?” If  $X$  is a normed space, we say  $x_n$  converges weakly to  $x$ , denoted  $x_n \xrightarrow{w} x$  if for all bounded linear functional  $f \in X'$ ,  $f(x_n) \rightarrow f(x)$ . For example let  $X = \ell^1$ . We have proven that the dual of  $\ell^1$  is  $\ell^\infty$  and elements of the dual take the form:

$$f(x) = \sum_{n=1}^{\infty} z_i x_i \quad (32.3.19)$$

Where  $z_i \in \ell^\infty$ . Thus,  $z_i$  is bounded and  $x_i$  is absolutely convergent, since  $x_i \in \ell^1$ , and thus the product is absolutely convergent. That is,  $x_i z_i \in \ell^1$ . The problem with this is that we don’t know that  $X' \neq \{0\}$  for a given Banach space. There are plenty these, enough to separate points, thanks to the Hahn-Banach

theorem. This says that if  $f$  is a bounded linear functional on a subspace  $M$  of  $X$ , then there exists  $F \in X'$  such that  $\|F\|_{X'} = \|f\|_M$  and  $F(x) = f(x)$  for all  $x \in M$ . So given  $m \in M$ , then  $\alpha m \in M$  for all  $\alpha \in \mathbb{R}$ . Define  $f(\alpha m) = k$ , for some  $k \in \mathbb{R}$ . Then  $f(\alpha m) = \alpha k$  and  $|f(\alpha m)| = |\alpha||f(m)| = |k|\|\alpha m\|/\|m\|$ . And  $|f(\alpha m)|/\|\alpha m\| = |k|/\|m\|$ . Thus  $f \in M'$ . Thus, Hahn-Banach can be extended to all of  $X$ . So we can for all  $x \in X$  and for all  $r \in \mathbb{R}$ , there is a bounded linear function  $f \in X'$  such that  $f(x) = r$ . If  $m_1$  and  $m_2$  are independent (That is,  $m_1 \neq \alpha m_2$  for any real number  $\alpha$ ), then let  $M = \{am_1 + bm_2 : a, b \in \mathbb{R}\}$ . Define  $f(am_1 + bm_2) = a\|m_1\|$ . Then  $f \in M'$  so this can be extended to all of  $X$  by the Hahn-Banach theorem. But  $f(m_1) = 1$  and  $f(m_2) = 0$ , so  $f$  separates points. Thus, if  $x_n$  converges weakly to  $x$  in a Banach space, and if  $x_n$  also converges weakly to  $y$ , then  $x = y$ . The uniform boundedness theorem also holds if  $X$  is complete. The Banach-Alaoglu-Hilbert theorem fails in a general Banach space. For example,  $\ell^1$ . We now talk about weak\* convergence. In  $X'$ , we say  $f_n$  converges weak\* to  $f$  if  $f_n(x) \rightarrow f(x)$  for all  $x \in X$ . Banach-Alaoglu holds if weak is replaced with weak\* and if  $X'$  is separable. The double dual,  $X''$ , is the dual of  $X'$ .  $X$  is embedded in  $X''$ . That is,  $X$  embeds naturally in  $X''$  as follows: Define  $C : X \rightarrow X''$  as follows. If  $x \in X$ ,  $Cx(f) = f(x)$  for all  $f \in X'$ . It's easy to show that  $\|Cx\| = \|x\|$ . Indeed,  $C$  is an isometry on  $X$  to  $X''$ . If  $C$  is onto, we say that  $X$  is reflexive. There are Banach spaces that are not reflexive that can be isometrically embedded into their second dual, but the canonical map is not such an embedding.

**Theorem 32.3.6.** *If  $H$  is a Hilbert space, then  $H$  is reflexive.*

**Theorem 32.3.7.** *If  $X$  is reflexive, then  $X = X''$ .*

**Theorem 32.3.8.** *If  $X$  is reflexive, weak\* convergence implies weak convergence.*

From topology, a subbasis for a topological space is a collection of sets such that every open set can be written as arbitrary unions and finite intersections of the sets. Choose as the subbasis:

$$\{f^{-1}(-\infty, a) : a \in \mathbb{R}, f \in X'\} \quad (32.3.20)$$

If  $X'$  is separable, then this space is metrizable. For let  $A$  be a countable dense subset, and define the metric  $d$  as:

$$d(f, g) = \sum_{x \in A} \frac{|f(x) - g(x)|}{1 + |f(x) - g(x)|} 2^{-n} \quad (32.3.21)$$

**Theorem 32.3.9** (Banach-Alaoglu). *If  $X$  is a normed vector space, and if  $\tau$  is the weak\* topology, then  $\overline{B}_1(0)$  is a compact subset of  $(X', \tau)$ .*

Adams Sobolev Spaces.

## 32.4 Old Notes

### 32.4.1 Summary of Lectures

The boundary of a circle in  $\mathbb{R}^2$  is nowhere dense, with respect to the metric on  $\mathbb{R}^2$ . Any open ball about any point on the circle contains points not on the circle, and thus it has empty interior. Something about  $\varepsilon$  nets.

#### Normed Spaces and Banach Spaces

There are notions of subspace, linear combination, independence, spanning, dimension, basis, Hamel basis, and *convexity*. Open and closed balls are convex. A subspace of a Banach space is complete iff closed. Schauder basis. A Schauder basis implies separable. If  $\{x_1, \dots, x_n\}$  is independent, then there exists a  $c > 0$  such that, for all  $\alpha$ ,  $|\alpha \cdot \mathbf{x}| \geq c\|\alpha\|$ . Finite dimensional subspaces are complete, as are closed subspaces. In finite dimensional normed spaces, a space is compact if and only if it is closed and bounded. Riesz's Lemma says that if  $Z$  is a subspace of a normed space  $X$ , and if  $Y$  is a proper closed subspace of  $Z$ , then there is a  $z \in Z$  such that  $\|z\| = 1$  and  $D(z, Y) \geq 1/2$ . A corollary of this is that  $B_1(0)$  is compact if and only if  $X$  is finite dimensional.

#### Linear Operators

Identity, zero, differentiation, and integration. Domain/Range of a linear operator, the null space. Inverse of a linear operator is linear.  $(ST)^{-1} = T^{-1}S^{-1}$ . In finite dimension all linear operators are continuous. An operator is bounded if and only if it is continuous. If a linear operator is continuous at some point, then it is continuous everywhere. An operator is bounded if and only if its null space is closed. There is something called the extension of a bounded linear operator.  $B(X, Y)$  is the set of bounder linear operators from  $X$  to  $Y$ . This is complete if and only if  $Y$  is complete. A functional is a mapping from a vector space  $X$  into the real numbers  $\mathbb{R}$ . For continuous linear functional, continuity at 0 implies continuity everywhere. There is something called the dual space  $X'$ , which is itself a Banach space.

#### Inner Product and Hilbert Spaces

If  $x_n \rightarrow x$  and  $y_n \rightarrow y$ , then  $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$ . There's a notion of orthogonal sets, and orthonormality. If  $(e_n)$  is orthonormal basis, then  $x = \sum \langle x, e_k \rangle e_k$  for all  $x$ . Gram-Schmidt procedure.  $\sum \alpha_k e_k$  converges if and only if  $\sum |\alpha_k|^2$  converges. A set  $M$  is total in a Hilbert space  $H$  is the span of the closure of  $M$  is equal to  $H$ . If  $M$  is complete, then it is total if and only if  $M^\perp = 0$ . Parseval's theorem. Legendre, Hermite, and Laguerre polynomials are things. Self adjoint, unitary, and normal operators.  $T^* = T$ ,  $T^* = T^{-1}$ , and  $T^*T =$

$TT^*$ . If  $X$  is a vector space over the complex numbers, and if  $T$  is self adjoint, then  $\langle Tx, x \rangle$  is a real number for all  $x$ .

## Compact Linear Operators

If  $T$  is compact and linear, then it is bounded and continuous. An operator is compact and linear if and only if for all bounded sequences  $x_n$ ,  $Tx_n$  has a convergent subsequence. Compact linear operators form a vector space. The rank of an operator is the dimension of its image. If  $T$  is linear, bounded, and of finite rank, then it is compact. If  $T_n$  is a sequence of compact linear operators, if  $Y$  is complete, and if  $\|T_n - T\| \rightarrow 0$ , then  $T$  is compact. A sequence  $x_n$  converges weakly to  $x$  if, for all  $y$ ,  $\langle x_n, y \rangle \rightarrow \langle x, y \rangle$ . If  $x_n$  converges weakly to  $x$ , then and if  $T$  is a compact linear operator, then  $Tx_n \rightarrow Tx$ . If  $H$  is a Hilbert space,  $T$  is a compact self-adjoint operator, and if  $x_n$  converges weakly to  $x$ , then  $\langle Tx_n, x_n \rangle \rightarrow \langle Tx, x \rangle$ . If  $T : H \rightarrow H$  is compact and linear, then so is its adjoint. The Hilbert-Schmidt theorem says that compact self-adjoint operators on a Hilbert space  $H$  have an orthonormal basis of eigenvectors. All of this has applications to integral operators and Sturm-Liouville Theory.

## Fundamental Theorems

Zorn's Lemma. Hahn-Banach Theorem. Sublinear functionals. If  $X$  is a normed space, and  $Z$  is a subspace, and if  $f \in Z'$ , then  $f$  be extended to  $X$  such that  $\|f\|_X = \|f\|_Z$ . This extends Hilbert spaces by Riesz. If  $X$  is a normed space and  $x \neq 0$ , then there is an  $f \in X'$  such that  $\|f\| = 1$  and  $f(x_0) = \|x_0\|$ . For all  $x$ ,  $\|x\| = \sup\{\|f(x)\|/\|f\| : f \in X', f \neq 0\}$ . There's a thing called bounded variation. If  $x \in X$  and  $g_x(f) = f(x)$  for  $f \in X'$ , then  $g_x \in X''$  and  $\|g_x\| = \|x\|$ . Reflexive implies complete. Finite and Hilbert implies reflexive.  $X'$  separable implies  $X$  is separable.  $X$  separable and reflexive implies  $X'$  is separable. Strong convergence implies weak convergence. The converse is not true. If  $X$  is finite dimensional, then weak convergence implies strong convergence. Weak convergence implies  $\|x_n\|$  is bounded. If  $x_n \rightarrow x$  weakly, and if  $\|x_n\| \rightarrow \|x\|$ , then  $x_n \rightarrow x$  strongly. Open mapping theorem. Closed graph theorem. Differentiation is a closed operator on  $C^1[a, b] \rightarrow C[a, b]$ .

## 32.5 Metric Spaces

### 32.5.1 Basic Definitions

#### Definition 32.5.1: Pseudo-Metric

A pseudo-metric on a set  $X$  is a function  $\rho : X \times X \rightarrow [0, \infty)$  such that, for all  $x, y, z \in X$ , it is true that:

$$\begin{aligned} \rho(x, y) &= \rho(y, x) && \text{(Symmetry)} \\ \rho(x, z) &\leq \rho(x, y) + \rho(y, z) && \text{(Triangle Inequality)} \end{aligned}$$



#### Definition 32.5.2: Pseudo-Metric Space

A pseudo-metric space,  $(X, \rho)$ , is a set  $X$  and a pseudo-metric  $\rho$  on  $X$ . ■

**Theorem 32.5.1.** *There exist pseudo-metric spaces  $(X, \rho)$  such that for all  $x \in X$ ,  $\rho(x, x) > 0$ .*

*Proof.* For let  $X = \mathbb{R}$  and define  $\rho : \mathbb{R}^2 \rightarrow [0, \infty)$  by:

$$\rho(x, y) = 1 + |x| + |y| \quad (32.5.1)$$

Then  $\rho$  is a pseudo-metric. For it is symmetric, since:

$$\rho(x, y) = 1 + |x| + |y| = 1 + |y| + |x| = \rho(y, x) \quad (32.5.2)$$

Moreover, it obeys the triangle inequality:

$$\rho(x, z) = 1 + |x| + |z| \quad (32.5.3a)$$

$$\leq 1 + |x| + |z| + 2|y| + 1 \quad (32.5.3b)$$

$$= (1 + |x| + |y|) + (1 + |y| + |z|) \quad (32.5.3c)$$

$$= \rho(x, y) + \rho(y, z) \quad (32.5.3d)$$

Thus,  $\rho$  is a pseudo-metric. However, for all  $x \in \mathbb{R}$ :

$$\rho(x, x) = 1 + |x| + |x| = 1 + 2|x| \geq 1 > 0 \quad (32.5.3e)$$

Thus, there are no  $x \in X$  such that  $\rho(x, x) = 0$ . Therefore, etc □

If we require  $\rho(x, x) = 0$  for all  $x \in X$ , we can still have the case where elements cannot be distinguished from. That is, there may be  $x, y \in X$  such that  $x \neq y$ , but  $\rho(x, y) = 0$ .

**Theorem 32.5.2.** *There exist pseudo-metric spaces  $(X, \rho)$  such that for all  $x \in X$ ,  $\rho(x, x) = 0$ , and there are distinct elements  $x, y \in X$  such that  $\rho(x, y) = 0$ .*

*Proof.* For let  $X$  have at least two distinct elements, and let  $\rho : X^2 \rightarrow [0, \infty)$  be defined by:

$$\rho(x, y) = 0 \quad (32.5.4)$$

Then  $\rho$  is a pseudo-metric. Symmetry and the triangle inequality are both trivial. However, since there are at least two distinct elements in  $X$ , we have unique points such that  $\rho(x, y) = 0$ . Therefore, etc.  $\square$

### Definition 32.5.3: Metric Space

---

A metric space is a pseudo-metric space  $(X, d)$  such that:

$$d(x, y) = 0 \iff x = y \quad (\text{Definiteness})$$



### Definition 32.5.4: Semi-Norm

---

A semi-norm on a vector space  $V$  over a field  $\mathbb{F} \subseteq \mathbb{C}$  is a function  $\|\cdot\| : V \rightarrow [0, \infty)$  such that, for all  $v \in V$  and  $\alpha \in \mathbb{F}$ , it is true that:

$$\|\alpha v\| = |\alpha| \|v\| \quad (\text{Homogeneity})$$

$$\|v + w\| \leq \|v\| + \|w\| \quad (\text{Triangle Inequality})$$



**Theorem 32.5.3.** *If  $V$  is a vector space over a field  $\mathbb{F} \subseteq \mathbb{C}$ , and if  $\|\cdot\|$  is semi-norm on  $V$ , then:*

$$\|\mathbf{0}\| = 0 \quad (32.5.5)$$

*Proof.* For:

$$\|\mathbf{0}\| = \|\mathbf{0}\mathbf{0}\| = |0| \|\mathbf{0}\| = 0 \quad (32.5.6)$$

Therefore, etc.  $\square$

### Definition 32.5.5: Norm

---

A norm on a vector space  $V$  over a field  $\mathbb{F} \subseteq \mathbb{C}$  is a semi-norm  $\|\cdot\|$  such that:

$$\|\mathbf{x}\| \implies \mathbf{x} = \mathbf{0} \quad (32.5.7)$$



**Example 32.5.1: S**

pose that  $\|\cdot\|_0$  is a semi-norm on a vector space  $V$ . Define the following:

$$N = \{v \in V : \|v\|_0 = 0\} \quad (32.5.8)$$

If follows from the definition of a semi-norm that  $N$  is a subspace of  $V$ . Thus we can define a function on the quotient space  $\|\cdot\| : V/N \rightarrow [0, \infty)$  by:

$$\|v + N\| = \|v\|_0 \quad (32.5.9)$$

We can then verify that this is well defined and that  $\|\cdot\|$  is a norm on  $V/N$ . ■

If  $\|\cdot\|$  is a norm on  $V$ , then we get an associated metric  $\rho$  via:

$$\rho(v, u) = \|v - u\| \quad (32.5.10)$$

**Example 32.5.2: L**

Let  $(X, \mathcal{M}, \mu)$  be a measure space. That is,  $X$  is a set,  $\mathcal{M}$  is  $\sigma$ -Algebra, and  $\mu$  is a measure on  $X$ . Let  $1 \leq p < \infty$ . Then:

$$\mathcal{L}^p(X) = \{f : X \rightarrow \mathbb{C} : f \text{ is measurable and } \int_X |f|^p d\mu < \infty\} \quad (32.5.11)$$

The set  $\mathcal{L}^p(X)$  is a vector space. We define the semi-norm on  $\mathcal{L}^p(X)$  to be:

$$\|f\|_p = \left( \int_X |f|^p d\mu \right)^{1/p} \quad (32.5.12)$$

This is not a norm, since there are many functions such that  $\|f\|_p = 0$ , yet  $f \neq 0$ . However, if  $\|f\|_p = 0$ , then  $f = 0$   $\mu$  almost-everywhere. So we create equivalence classes by comparing functions that are  $\mu$  almost-everywhere. The final thing to check is the triangle-inequality. It is not obvious and is a consequence of Minkowski's Inequality. We get a normed vector space by considering  $N$  to be the set of functions  $f$  such that  $\|f\|_p = 0$ , and we define:

$$L^p(X) = \mathcal{L}^p(X)/N \quad (32.5.13)$$

The analyst Halmos said that the only important values of  $p$  are 1, 2, and  $\infty$ . If  $f : X \rightarrow \mathbb{C}$  is measurable, then we define:

$$\|f\|_\infty = \inf\{c \geq 0 : \mu(\{x : |f(x)| > c\}) = 0\} \quad (32.5.14)$$

With the convention that  $\inf\{\emptyset\} = \infty$ . This defines a semi-norm on:

$$\mathcal{L}^\infty(X) = \{f : \|f\|_\infty < \infty\} \quad (32.5.15)$$

Homogeneity pops out rather quickly, but the triangle-inequality is still tricky.  
We call  $\|\cdot\|_\infty$  the essential supremum of  $f$ . ■

**Theorem 32.5.4.** *If  $f : X \rightarrow \mathbb{C}$  is measurable, and if:*

$$E = \{p : \|f\|_p < \infty, p \in [1, \infty)\} \quad (32.5.16)$$

*Then  $E$  is connected.*

### Example 32.5.3: L

---

Let  $X$  be a finite set, let  $\mathcal{M} = \mathcal{P}(X)$ , and let  $\mu$  be the counting measure on  $X$ . A function on  $X$  is an  $n$ -tuple  $x = (x_1, \dots, x_n)$ . Then:

$$\|x\|_p = \begin{cases} \left( \sum_{k=1}^n |x_k|^p \right)^{1/p}, & 1 \leq p < \infty \\ \max\{|x|, x \in X\} & \end{cases} \quad (32.5.17)$$

$\|\cdot\|_p$  is a norm on  $X$ . ■

### Example 32.5.4: L

---

Let  $X = \mathbb{N}$ , the set of natural numbers. Let  $\mathcal{M} = \mathcal{P}(X)$ , and let  $\mu$  be the counting measure. Then functions are sequences  $a : \mathbb{N} \rightarrow \mathbb{R}$ , or  $a : \mathbb{N} \rightarrow \mathbb{C}$ . Then:

$$\|a\|_p = \begin{cases} \left( \sum_{n=1}^{\infty} |a_n|^p \right)^{1/p}, & 1 \leq p < \infty \\ \max\{|a|, a \in X\} & \end{cases} \quad (32.5.18)$$

This defines a norm. Recall that a series is absolutely convergent if  $\sum |a_n| < \infty$ . Given an absolutely convergent series, the original series is also convergent. For this space we use the following notation:

$$\ell^p = \{a : \mathbb{R} \rightarrow \mathbb{R} : \sum_{n=1}^{\infty} |a_n|^p < \infty\} \quad (32.5.19)$$
■

In general, if  $X$  is a set then we can equip  $X$  with the counting measure and then if  $f$  is any bounded function on  $f$ , then:

$$\|f\|_\infty = \sup\{|f(x) : x \in X\} \quad (32.5.20)$$

For the space of sequences, we write  $\ell^\infty(X)$ .

**Example 32.5.5: I**

$X$  is any set, then we define the following metric:

$$\rho(x, y) = \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases} \quad (32.5.21)$$

This is often called the discrete metric. It is indeed a metric, and  $(X, \rho)$  is a metric space. ■

**Example 32.5.6: L**

Let  $(X, \rho)$  be a metric space. If  $Y \subseteq X$  is a non-empty subset of  $X$ , we can define a new metric on  $Y$  by restricting  $\rho$  to  $Y \times Y$ . We call this the metric subspace. ■

**Definition 32.5.6: Strongly Equivalent Metrics**

Strongly equivalent metrics on a set  $X$  are metrics  $\rho_1$  and  $\rho_2$  such that there exists  $c, d \in \mathbb{R}^+$  such that, for all  $x, y \in X$ :

$$c\rho_1(x, y) \leq \rho_2(x, y) \leq d\rho_1(x, y) \quad (32.5.22)$$

■

The definition of strongly equivalent metrics is indeed symmetric. For since  $c, d \in \mathbb{R}^+$ ,  $c^{-1}$  and  $d^{-1}$  are well defined and positive, and thus:

$$\frac{1}{d}\rho_2(x, y) \leq \rho_1(x, y) \leq \frac{1}{c}\rho_1(x, y) \quad (32.5.23)$$

**Theorem 32.5.5.** *If  $p, q \in [1, \infty)$ , then  $\|\cdot\|_p$  and  $\|\cdot\|_q$  are strongly equivalent.*

*Proof.* It suffices to show that for all  $p \in [1, \infty)$  there exist  $c, d \in \mathbb{R}^+$  such that:

$$c\|x\|_p \leq \|x\|_2 \leq d\|x\|_p \quad (32.5.24)$$

Also note that:

$$\partial\bar{B}_1(\mathbf{0}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1\} \quad (32.5.25)$$

Is a closed and bounded subset of  $\mathbb{R}^n$ . □

**Example 32.5.7: L**

Let  $(X, \rho)$  be a metric space on  $X$  and define:

$$d(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)} \quad (32.5.26)$$

Then  $(X, d)$  is a metric space. Definiteness and symmetry come rather immediately from the definition and the fact that  $\rho$  is a metric. The only thing to check is the triangle inequality. ■

**Definition 32.5.7: Open Ball in a Metric Space**

The open ball about a point  $x$  in a metric space  $(X, \rho)$  of radius  $r \in \mathbb{R}$  is the set:

$$B_r(x) = \{y \in X : \rho(x, y) < r\} \quad (32.5.27)$$
■

**Definition 32.5.8: Open Subsets of a Metric Space**

An open subset of a metric space  $(X, \rho)$  is a set  $\mathcal{U} \subseteq X$  such that for all  $x \in \mathcal{U}$  there is an  $\varepsilon > 0$  such that:

$$B_\varepsilon(x) \subseteq \mathcal{U} \quad (32.5.28)$$
■

**Theorem 32.5.6.** *Open balls are open.*

**Definition 32.5.9: Neighborhoods**

A neighborhood of a point  $x$  in a metric space  $(X, \rho)$  is a subset  $D \subseteq X$  such that there is an open subset  $\mathcal{U} \subseteq D$  such that  $x \in \mathcal{U}$ . ■

**Theorem 32.5.7.** *If  $(X, \rho)$  is a metric space, then  $X$  is an open subset.*

**Theorem 32.5.8.** *If  $(X, \rho)$  is a metric space, then  $\emptyset$  is an open subset.*

**Theorem 32.5.9.** *If  $\mathcal{U}_I$  is a collection of open subsets of a metric space  $(X, \rho)$ , and if  $\mathcal{U}$  is defined by:*

$$\mathcal{U} = \bigcup_{i \in I} \mathcal{U}_i \quad (32.5.29)$$

*Then  $\mathcal{U}$  is an open subset.*

**Theorem 32.5.10.** If  $(X, \rho)$  is a metric space, and if  $\mathcal{U}$  and  $\mathcal{V}$  are open subsets, then the set  $\mathcal{D}$  defined by:

$$\mathcal{D} = \mathcal{U} \cap \mathcal{V} \quad (32.5.30)$$

Is an open subset of  $X$ .

**Theorem 32.5.11.** If  $(X, \rho)$  is a metric space and  $(\mathcal{E}, \rho_{\mathcal{E}})$  is a subspace of  $(X, \rho)$ , then a set  $\mathcal{U} \subseteq \mathcal{E}$  is open in  $\mathcal{E}$  if and only if there is an open set  $\mathcal{U}$  in  $X$  such that:

$$\mathcal{V} = \mathcal{U} \cap \mathcal{E} \quad (32.5.31)$$

### Definition 32.5.10: Topology

---

A topology on a set  $X$  is a subset  $\tau \subseteq \mathcal{P}(X)$  such that:  $\emptyset \in \tau$  and  $X \in \tau$ , for any finite subset  $\mathcal{C} \subseteq \tau$ , it is true that:

$$\bigcap_{C \in \mathcal{C}} C \in \tau \quad (32.5.32a)$$

And for any subset  $\mathcal{O} \subseteq \tau$  it is true that:

$$\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \in \tau \quad (32.5.32b)$$

That is,  $\tau$  is closed to finite intersections and arbitrary unions. ■

### Example 32.5.8: I

---

$(X, \rho)$  is a metric space, then the set:

$$\tau = \{\mathcal{U} \subseteq X : \mathcal{U} \text{ is open}\} \quad (32.5.33)$$

is a topology on  $X$ . This is the *metric topology*. Not every topological space can be formed from a metric space. If  $(\mathcal{E}, \rho_{\mathcal{E}})$  is a subspace of  $(X, \rho)$ , then the *subspace topology*, or the relative topology, is the set:

$$\tau_{\mathcal{E}} = \{\mathcal{E} \cap \mathcal{U} : \mathcal{U} \in \tau\} \quad (32.5.34)$$

This is a topology on  $\mathcal{E}$ . ■

### Definition 32.5.11: Closed Subsets

---

A closed subset of a metric space  $(X, \rho)$  is a set  $\mathcal{C} \subseteq X$  such that  $X \setminus \mathcal{C}$  is an open subset of  $X$ . ■

**Theorem 32.5.12.** *If  $(X, \rho)$  is a metric space, then  $X$  is closed.*

**Theorem 32.5.13.** *If  $(X, \rho)$  is a metric space, then  $\emptyset$  is closed.*

**Theorem 32.5.14.** *If  $(X, \rho)$  is a metric space, and if  $\mathcal{C}_i$  is a finite collection of closed subsets, then the set  $\mathcal{C}$  defined by:*

$$\mathcal{C} = \bigcup_{i=1}^n \mathcal{C}_i \quad (32.5.35)$$

*is a closed subset of  $X$ .*

### Definition 32.5.12: Closure of a Set

---

If  $(X, \rho)$  is a metric space and  $\mathcal{E} \subseteq X$ , then the closure of  $\mathcal{E}$  is the set:

$$\bar{\mathcal{E}} = \bigcap \{ \mathcal{F} \subseteq X : \mathcal{E} \subseteq \mathcal{F} \wedge \mathcal{F} \text{ is closed.} \} \quad (32.5.36)$$



The closure of a set  $\mathcal{E}$  is the smallest closed set that contains  $\mathcal{E}$ .

**Theorem 32.5.15.** *If  $(X, \rho)$  is a metric space, and if  $\mathcal{E}$  is a non-empty subset of  $X$ , then  $x \in \bar{\mathcal{E}}$  if and only if for all  $\varepsilon > 0$ :*

$$B_\varepsilon(x) \cap \mathcal{E} \neq \emptyset \quad (32.5.37)$$

### Definition 32.5.13: Convergent Sequences in a Metric Space

---

A convergent sequence in a metric space  $(X, \rho)$  is a sequence  $a : \mathbb{N} \rightarrow X$  such that there is an  $x \in X$  such that for all  $\varepsilon > 0$  there exists an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  and  $n > N$ , it is true that  $d(x, a_n) < \varepsilon$ . We write  $a_n \rightarrow x$ .

### Definition 32.5.14: Limits of Convergent Sequences

---

A limit of a convergent sequence  $a : \mathbb{N} \rightarrow X$  in a metric space  $(X, \rho)$  is a point  $x \in X$  such that  $a_n \rightarrow x$ .

**Theorem 32.5.16.** *If  $(X, \rho)$  is a metric space, if  $a : \mathbb{N} \rightarrow X$  is a convergent sequence, and if  $x$  and  $y$  are limits of  $a$ , then  $x = y$ .*

### Definition 32.5.15: Equivalent Metrics

---

Equivalent metrics on a set  $X$  and metrics  $\rho$  and  $d$  such that they generate the same topology.



**Theorem 32.5.17.** If  $X$  is a set and if  $\rho$  and  $d$  are equivalent metrics on  $X$ , then for all  $x \in X$  and for all  $r \in \mathbb{R}^+$  there exists  $r_1, r_2 \in \mathbb{R}^+$  such that:

$$B_{r_1}^\rho(x) \subseteq B_r^d(x) \quad (32.5.38)$$

$$B_{r_2}^d(x) \subseteq B_r^\rho(x) \quad (32.5.39)$$

### Example 32.5.9: I

---

$(X, \rho)$  is a metric space and if  $d$  is defined by:

$$d(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)} \quad (32.5.40)$$

The  $d$  is a metric on  $X$ . Moreover,  $d$  is equivalent to  $\rho$ . This shows that the notion of boundedness is not a topological one, but a metric property. For, given any metric  $\rho$ ,  $d$  is bounded. For all  $x, y \in X$ ,  $0 \leq d(x, y) < 1$ . Letting  $\rho$  be the standard metric on  $\mathbb{R}$ ,  $\rho(x, y) = |x - y|$ , we see that the topology generated by this unbounded metric is equivalent to the topology generated by the metric:

$$d(x, y) = \frac{|x - y|}{1 + |x - y|} \quad (32.5.41)$$



**Theorem 32.5.18.** If  $(X, \rho)$  is a metric space, and if  $d : X^2 \rightarrow [0, \infty)$  is defined by:

$$d(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)} \quad (32.5.42)$$

Then  $\rho$  and  $d$  are equivalent.

*Proof.* For let  $r > 0$ . For all  $x, y \in X$ ,  $d(x, y) \leq \rho(x, y)$ , and therefore:

$$B_r^\rho(x) \subseteq B_r^d(x) \quad (32.5.43)$$

If  $\rho(x, y) \leq 1$ , then  $\rho(x, y) \leq 2d(x, y)$ . Let  $r_1 = \min\{r/2, 1\}$ , then:

$$B_{r/2}^d(x) \subseteq B_r^\rho(x) \quad (32.5.44)$$

Therefore, etc. □

**Definition 32.5.16: Continuous Functions Between Metric Spaces**

A continuous function from a metric space  $(X, \rho)$  to a metric space  $(Y, d)$  is a function  $f : X \rightarrow Y$  such that, for all  $x \in X$  and for all  $\varepsilon > 0$ , there is a  $\delta > 0$  such that:

$$f(B_\delta^\rho(x)) \subset B_\varepsilon^d(f(x)) \quad (32.5.45)$$

■

**Theorem 32.5.19.** *If  $(X, \rho)$  and  $(Y, d)$  are metric spaces, and if  $f : X \rightarrow Y$  is a function, then the following are equivalent:*

1.  *$f$  is continuous at  $x_0 \in X$ .*
2. *If  $x_n \rightarrow x_0$  then  $f(x_n) \rightarrow f(x_0)$*
3. *If  $\mathcal{V}$  is a neighborhood of  $f(x_0)$ , then  $f^{-1}(\mathcal{V})$  is a neighborhood of  $x_0$ .*

**Theorem 32.5.20.** *If  $(X, \rho)$  and  $(Y, d)$  are metric spaces, and if  $f : X \rightarrow Y$  is a function that is continuous at  $x_0 \in X$ , then for all  $\mathcal{V} \subseteq Y$  such that  $\mathcal{V}$  is open and  $f(x_0) \in \mathcal{V}$ , then  $f^{-1}(\mathcal{V})$  is an open subset of  $x_0$ .*

**Definition 32.5.17: Uniformly Continuous Functions**

A uniformly continuous function from a metric space  $(X, \rho)$  to a metric space  $(Y, d)$  is a function  $f : X \rightarrow Y$  such that for all  $\varepsilon > 0$  there exists a  $\delta > 0$  such that, for all  $x, y \in X$  such that  $\rho(x, y) < \delta$ , it is true that  $d(f(x), f(y)) < \varepsilon$ . ■

**Theorem 32.5.21.** *If  $f : X \rightarrow Y$  is uniformly continuous, then  $f$  is continuous.*

The converse is false. For define  $f(x) = x^2$ .

**32.5.2 Completeness****Definition 32.5.18: Cauchy Sequences**

A Cauchy sequence in a metric space  $(X, d)$  is a sequence  $a : \mathbb{N} \rightarrow X$  such that, for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that, for all  $n, m \in \mathbb{N}$  such that  $n, m > N$ , it is true that  $d(x_n, x_m) < \varepsilon$ . ■

**Example 32.5.10: L**

t  $X = (0, 2)$  with the usual metric, and let  $a : \mathbb{N} \rightarrow X$  be defined by:

$$a_n = \frac{1}{n} \quad (32.5.46)$$

Then  $a$  is a Cauchy sequence since:

$$|a_n - a_m| = \frac{|n - m|}{nm} < \frac{2}{\min(n, m)} \quad (32.5.47)$$

And this converges to zero. However the sequence doesn't converge, since we took zero away. █

**Example 32.5.11: L**

t  $X = C([0, 3])$  and let:

$$\|f\|_1 = \int_0^3 |f(x)| dx \quad (32.5.48)$$

Then  $\|\cdot\|_1$  is a norm on the set of continuous functions, and thus induces a metric. Let  $f_n$  be defined by:

$$f_n(x) = \begin{cases} 1, & x \leq x < 2 - \frac{1}{n} \\ Bob, & \\ 0, & x \geq 2 \end{cases} \quad (32.5.49)$$

Then  $f_n$  is Cauchy, but does not converge. █

**Definition 32.5.19: Complete Metric Spaces**

A complete metric space is a metric space  $(X, d)$  such that, for all Cauchy sequences  $a : \mathbb{N} \rightarrow X$ ,  $a$  is a convergent sequence. █

**Theorem 32.5.22.** *If  $(X, d)$  is a metric space, if  $a : \mathbb{N} \rightarrow X$  is a Cauchy sequence, and if there is a convergent subsequence of  $a$ , then  $a$  is a convergent sequence.*

**Theorem 32.5.23.** *A normed vector space  $(V, \|\cdot\|)$  is complete if and only if every absolutely convergent series converges.*

*Proof.* Suppose  $V$  is complete and let  $u_n$  be absolutely convergent. That is,

the sequence of partial sums:

$$S_N = \sum_{n=1}^N \|u_n\| \quad (32.5.50)$$

Converges in  $\mathbb{R}$ . But then:

$$\lim_{N \rightarrow \infty} \sum_{k=N}^{\infty} \|u_k\| = 0 \quad (32.5.51)$$

Define:

$$s_n = \sum_{k=1}^n u_k \quad (32.5.52)$$

But if  $m \geq n$ , then:

$$\|s_n - s_m\| \leq \sum_{k=n+1}^m \|u_k\| \leq \sum_{k=n+1}^{\infty} \|u_k\| \quad (32.5.53)$$

Thus, if  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that, for all  $n \geq N$ , it is true that:

$$\sum_{k=n+1}^{\infty} \|u_k\| < \varepsilon \quad (32.5.54)$$

Therefore  $s_n$  is a Cauchy sequence, and therefore there is an  $s \in V$  such that  $s_n \rightarrow s$ . Proving the converse, suppose  $u_n$  is a Cauchy sequence. Then there is an  $N_1 \in \mathbb{N}$  such that, for all  $n, m \geq N_1$ , we have  $\|u_n - u_m\| < 1/2$ . But then there is also an  $N_2 \in \mathbb{N}$  such that  $N_2 > N_1$ , and for all  $n, m > N_2$ ,  $\|u_n - u_m\| < 1/2$ . Continuing inductively, we find a sequence  $u_{n_k}$  such that:

$$\|u_{n_{k+1}} - u_{n_k}\| < \frac{1}{2^k} \quad (32.5.55)$$

Let  $v_k = u_{n_{k+1}} - u_{n_k}$ , and note that:

$$\sum_{n=1}^{\infty} \|v_n\| < \infty \quad (32.5.56)$$

But then there is a  $v \in V$  such that:

$$\sum_{n=1}^{\infty} v_n = v \quad (32.5.57)$$

But:

$$v = \lim_{N \rightarrow \infty} \sum_{k=1}^N v_k \quad (32.5.58)$$

$$= \lim_{N \rightarrow \infty} v_{n_{N+1}} - u_{n_1} \quad (32.5.59)$$

Therefore  $u_{n_k} \rightarrow v + u_{n_1}$ . But  $u_n$  is Cauchy and thus if there is a convergent subsequence, then it is a convergent sequence. Therefore,  $(X, \|\cdot\|)$  is complete.  $\square$

**Theorem 32.5.24.** *If  $(X, \mathcal{M}, \mu)$  is a measure space, and if  $1 \leq p \leq \infty$ .*

*Proof.* Suppose that  $f_n$  is a sequence of functions in  $L^P(X, \mathcal{M}, \mu)$  such that:

$$\sum_{k=1}^{\infty} \|f_k\|_p = B < \infty \quad (32.5.60)$$

Define  $G, G_n : X \rightarrow [0, \infty]$  be defined by:

$$G(x) = \sum_{k=1}^{\infty} |f_k(x)| \quad (32.5.61)$$

$$G_n(x) = \sum_{k=1}^n |f_k(x)| \quad (32.5.62)$$

Then, from the triangle inequality, we have that:

$$\|G_n\|_p \leq \sum_{k=1}^n \|f_k\|_p \leq B \quad (32.5.63)$$

Thus, by the monotone convergence theorem, we have:

$$\int_X G(x)^p d\mu = \lim_{n \rightarrow \infty} \int_X G_n(x)^p d\mu \leq B^p \quad (32.5.64)$$

Therefore  $G \in \mathcal{L}^p(X)$ , and thus  $G(x) < \infty$   $\mu$  almost-everywhere. But then the original series converges  $\mu$  almost everywhere. Define  $F$  be:

$$F(x) = \begin{cases} \sum_{n=1}^{\infty} f_n(x), & |\sum_{n=1}^{\infty} f_n(x)| < \infty \\ 0, & \text{Otherwise} \end{cases} \quad (32.5.65)$$

Then  $|F(x)| \leq G(x)$ , and thus  $F \in \mathcal{L}^p(X)$ . Moreover:

$$|F(x) - \sum_{k=1}^n f_k(x)|^p \leq 2^p G(x)^p \quad (32.5.66)$$

Therefore, by the Lebesgue Dominated Convergence Theorem, we have that:

$$\|F - \sum_{k=1}^n f_k(x)\|_p^p \rightarrow 0 \quad (32.5.67)$$

Therefore,  $F \in L^p(X, \mathcal{M}, \mu)$ . □

---

**Definition 32.5.20: Supremum Norm of Bounded Continuous Function**

---

The supremum norm on set  $C_b(X)$  of bounded continuous functions on a metric space  $(X, d)$  is:

$$\|f\|_\infty = \sup_{x \in X} |f(x)| \quad (32.5.68) \quad \blacksquare$$

From this, we can see that  $f_n \rightarrow f$  if and only if  $f_n \rightarrow f$  uniformly on  $X$ .

**Theorem 32.5.25.**  $C_b(X)$  is complete in the supremum norm.

*Proof.* Suppose that  $f_n$  is a Cauchy sequence in  $C_b(X)$ . For all  $f_n$  and  $x \in X$ ,  $f_n(x)$  is a Cauchy sequence in  $\mathbb{C}$ . But  $\mathbb{C}$  is complete, and thus there is a  $c_x \in \mathbb{C}$  such that  $f_n(x) \rightarrow c_x$ . Let  $f(x) = c_x$  for all  $x \in X$ . Then  $f_n \rightarrow f$ . For, let  $\varepsilon > 0$ . Then there exists  $N \in \mathbb{N}$  such that, for all  $n, m > N$  implies that:

$$|f_n(x) - f_m(x)| < \varepsilon/2 \quad (32.5.69)$$

But then:

$$|f_n(x) - f(x)| = \lim_{m \rightarrow \infty} |f_n(x) - f_m(x)| \leq \frac{\varepsilon}{2} < \varepsilon \quad (32.5.70)$$

But the uniform limit of continuous functions is continuous. Therefore, etc. □

**Theorem 32.5.26.** If  $(X, d)$  is a complete metric space, if  $(E, d')$  is a subspace of  $(X, d)$ , and if  $E$  is closed, then  $(E, d')$  is complete.

*Proof.* Suppose  $E$  is closed and suppose  $(x_n)$  is a Cauchy sequence in  $E$ . Then  $x_n$  is a Cauchy sequence in  $X$ , but  $X$  is complete. Therefore there is an  $x \in X$  such that  $x_n \rightarrow x$ . But  $E$  is closed, and therefore  $x \in E$ . Now suppose  $E$  is complete. Suppose  $x_n$  is a sequence in  $E$  and that  $x_n \rightarrow y$  in  $X$ . But convergent sequences are Cauchy sequences, and thus  $x_n$  is a Cauchy sequence. But  $E$  is complete and therefore  $y \in E$ . Therefore,  $E$  is closed. □

**Definition 32.5.21: Bounded Metric Spaces**

A bounded metric space is a metric space  $(X, d)$  such that there exists an  $x \in X$  and an  $r > 0$  such that:

$$X \subseteq B_r^{(X,d)}(x) \quad (32.5.71)$$

■

**Definition 32.5.22: Diameter of a Metric Space**

The diameter of a bounded metric space  $(X, d)$  is:

$$\text{diam}(X) = \sup_{x \in X} \{d(x, y) : x, y \in X\} \quad (32.5.72)$$

■

Every bounded metric space is contained in some open ball.

**Theorem 32.5.27.** *If  $(X, d)$  is a metric space, and then it is complete if and only if for any sequence of non-empty closed sets  $F : \mathbb{N} \rightarrow \mathcal{P}(X)$  such that  $F_{n+1} \subseteq F_n$  and  $\text{diam}(F_n) \rightarrow 0$ , there is an  $x \in X$  such that:*

$$\{x\} = \cap_{n=1}^{\infty} F_n \quad (32.5.73)$$

*Proof.* For suppose  $(X, d)$  is complete, and let  $F : \mathbb{N} \rightarrow \mathcal{P}(X)$  be a sequence of non-empty subsets of  $X$ . Then, for all  $n \in \mathbb{N}$ ,  $F_n$  is non-empty, and thus there is a sequence  $a : \mathbb{N} \rightarrow X$  such that, for all  $n \in \mathbb{N}$ ,  $x_n \in F_n$ . But then:

$$d(a_n, a_m) \leq \text{diam}(F_{\max\{n,m\}}) \quad (32.5.74)$$

But  $\text{diam}(F_n) \rightarrow 0$ , and therefore  $a$  is a Cauchy sequence. But  $(X, d)$  is complete, and therefore there is an  $x \in X$  such that  $a_n \rightarrow x$ . Moreover, there is an  $N \in \mathbb{N}$  such that  $x \in \overline{F_N}$ . But  $F_N$  is closed, and thus  $x \in F_N$ . But for all  $n > N$ ,  $F_n \subseteq F_N$ . Therefore:

$$x \in \cap_{n=1}^{\infty} F_n \quad (32.5.75)$$

If  $y \in \cap_{n=1}^{\infty} F_n$ , then  $d(x, y) \leq \text{diam}(F_n)$  for all  $n \in \mathbb{N}$ . But  $\text{diam}(F_n) \rightarrow 0$ , and thus  $d(x, y) = 0$ . Therefore,  $x = y$ . Going the other way, suppose  $X$  has the nested set property and let  $a : \mathbb{N} \rightarrow X$  be a Cauchy sequence in  $X$ . Let  $F : \mathbb{N} \rightarrow \mathcal{P}(X)$  be defined by:

$$F_n = \overline{\{a_k : k \geq n\}} \quad (32.5.76)$$

Then, for all  $n \in \mathbb{N}$ ,  $F_n$  is non-empty, and  $F_{n+1} \subseteq F_n$ . Moreover,  $\text{diam}(F_n) \rightarrow 0$ . Thus, by the nested sequence property, there is an  $x \in X$  such that  $x \in$

$\cap_{n=1}^{\infty} F_n$ . But then:

$$d(a_n, x) \leq \text{diam}(F_n) \rightarrow 0 \quad (32.5.77)$$

and therefore  $a_n \rightarrow x$ . Thus,  $a$  is a Cauchy sequence and  $(X, d)$  is complete.  $\square$

### 32.5.3 Compactness

#### Definition 32.5.23: Covers

A cover of a subset  $\mathcal{E} \subseteq X$  of a set  $X$  is a subset  $\mathcal{O} \subseteq \mathcal{P}(X)$  such that:

$$\mathcal{E} \subseteq \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \quad (32.5.78)$$



#### Definition 32.5.24: Sub-Cover

A sub-cover of a cover  $\mathcal{O}$  of a subset  $E \subseteq X$  of a set  $X$  is a subset  $\Delta \subseteq \mathcal{O}$  such that:

$$\mathcal{E} \subseteq \bigcup_{\mathcal{U} \in \Delta} \mathcal{U} \quad (32.5.79)$$



#### Definition 32.5.25: Open Covers

An open cover of a metric space  $(X, d)$  is a cover  $\mathcal{O} \subseteq \mathcal{P}(X)$  of  $X$  such that, for all  $\mathcal{U} \in \mathcal{O}$ ,  $\mathcal{U}$  is open.



#### Definition 32.5.26: Compact Sets

A compact metric space is a metric space  $(X, d)$  such that for any open cover  $\mathcal{O}$  of  $X$ , there is a finite sub-cover  $\Delta \subseteq \mathcal{O}$ .



#### Example 32.5.12: L

Let  $X = [0, 1]$  with the usual topology, and let:

$$\mathcal{U}_x = [0, x] \quad x \in (0, 1) \quad (32.5.80)$$

Then  $\mathcal{O} = \{\mathcal{U}_x : x \in (0, 1)\}$  is an open cover of  $X$ , but there is no finite sub-cover. For given any finite sub-cover, there is a greatest  $x$  such that  $\mathcal{U}_x$  is contained in the sub-cover. But then for all  $y \in (x, 1)$ ,  $y$  is not in the sub-cover. As a trivial example, any finite metric space is compact.  $\blacksquare$

**Theorem 32.5.28.** *If  $K$  is a subspace of  $X$ , then  $K$  is compact if and only if every open cover of  $K$  has a finite sub-cover.*

*Proof.* For suppose  $(K, d_K)$  is compact, and let  $\mathcal{O}$  be an open cover of  $K$ . Then:

$$\mathcal{O}_K = \{K \cup \mathcal{U} : \mathcal{U} \in \mathcal{O}\} \quad (32.5.81)$$

Is an open cover of  $K$ . But  $K$  is compact, and thus there is a finite sub-cover  $\Delta_K$ . But then:

$$\Delta = \{\mathcal{U} \in \mathcal{U} : \mathcal{U} \cap K \in \Delta_K\} \quad (32.5.82)$$

And this is a finite sub-cover.  $\square$

### Definition 32.5.27: Finite Intersection Property

A set with the finite intersection property in a metric space  $(X, d)$  is a collection of sets  $\mathcal{F} \subseteq \mathcal{P}(X)$  such that, for any sequence  $F : \mathbb{Z}_n \rightarrow \mathcal{F}$ , it is true that  $\cap_{k=1}^n F_k \neq \emptyset$ .  $\blacksquare$

**Theorem 32.5.29.** *A metric space  $(X, d)$  is compact if and only if every collection  $\mathcal{F}$  of closed sets in  $X$  with the finite intersection property is such that:*

$$\bigcap_{C \in \mathcal{F}} C \neq \emptyset \quad (32.5.83)$$

### Example 32.5.13: L

t  $F_n = [n, \infty)$ , and let  $\mathcal{F} = \{F_n : n \in \mathbb{N}\}$ . Then  $\mathcal{F}$  has the finite intersection property. However, the intersection over the entire set is empty, and hence  $\mathbb{R}$  (With the standard metric) is not compact.  $\blacksquare$

### Definition 32.5.28: Totally Bounded Metric Space

A totally bounded metric space is a metric space  $(X, d)$  such that, for all  $\varepsilon > 0$ , there exists an  $n \in \mathbb{N}$  and a sequence  $a : \mathbb{Z}_n \rightarrow X$  such that:

$$X = \bigcup_{k=1}^n B_\varepsilon^{(X, d)}(a_k) \quad (32.5.84)$$



**Definition 32.5.29:  $\varepsilon$ -Nets**

An  $\varepsilon$ -Net of a subspace  $(\mathcal{E}, d_{\mathcal{E}})$  of a metric space  $(X, d)$  is a finite collection:

$$E = \{B_{\varepsilon}^{(X,d)}(x_k) : k \in \mathbb{Z}_n\} \quad (32.5.85)$$

Such that  $E$  is an open cover of  $\mathcal{E}$ . ■

**Example 32.5.14: L**

t  $X = \ell^2$  and let:

$$e_n(x) = \begin{cases} 1, & k = n \\ 0, & k \neq n \end{cases} \quad (32.5.86)$$

Then  $e_n \in \ell^2$  and  $\|e\|_2 = 1$ , but for all  $n \neq m$ ,  $\|e_n - e_m\|_2 = \sqrt{2}$ . Let:

$$B_1 = \{f \in \ell^2 : \|f\|_2 \leq 1\} \quad (32.5.87)$$

Then  $B_1$  is bounded, but if  $\varepsilon = \sqrt{2}/2$  then no finite collection of  $\varepsilon$  balls can cover  $B_1$  since each ball can contain at most one of the  $e_n$ . Thus any cover is infinite. ■

**Theorem 32.5.30.** *A subset of  $(\mathbb{R}^n, \|\cdot\|_2)$  is totally bounded if and only if it's bounded.*

*Proof.* Totally bounded implies bounded, so it suffices to show that if  $\mathbb{R}^n$  is bounded then it is totally bounded. Let  $\mathcal{E} \subseteq \mathbb{R}^n$  be bounded. Then there is an  $r > 0$  such that:

$$\mathcal{E} \subseteq [-r, r]^n \quad (32.5.88)$$

Then, compactness, stuff like that. □

This works for any norm on  $\mathbb{R}^n$ , since all norm's on  $\mathbb{R}^n$  are strongly equivalent.

**Definition 32.5.30: Sequential Compactness**

A sequentially compact metric space is a metric space  $(X, d)$  such that, for all  $a : \mathbb{N} \rightarrow X$ , there is a convergent subsequence of  $a$ . ■

**Example 32.5.15: L**

t  $X \subseteq \mathbb{R}$  be defined by:

$$X = \left\{ \frac{1}{n} : n \in \mathbb{N} \right\} \cup \{0\} \quad (32.5.89)$$

Then  $X$  is sequentially compact, with respect to the subspace metric. ■

**Theorem 32.5.31.** IF  $(X, d)$  is a metric space, then the following are equivalent:

1.  $X$  is compact.
2.  $X$  is complete and totally bounded.
3.  $X$  is sequentially compact.

*Proof.* Suppose  $(X, d)$  is not compact, and let  $\mathcal{U}_i$  be an open cover with no finite subcover. If  $X$  is totally bounded, then there is a finite covering of  $1/2$  balls. But then at least one of these isn't covered by finitely many of the  $\mathcal{U}_i$ . Let  $F_1$  be the closure of this. Then  $F_1$  is totally bounded, and has a finite covering of  $1/4$  balls. One of these must not be covered by finitely many of the  $\mathcal{U}_i$ . Let  $F'_2$  be the closure of such a ball, and let  $F_2 = F_1 \cap F'_2$ . Then  $F_2$  is closed, non-empty, and  $\text{diam}(F_2) \leq 1/2$ . Continuing, we obtain a sequence of non-empty closed sets  $F_n$  such that for all  $n \in \mathbb{N}$ ,  $F_{n+1} \subseteq F_n$  and  $\text{diam}(F_n) < 1/2^n$ . Thus, if  $X$  is complete, there is a unique  $x$  that lies in the intersection of all of the  $F_n$ . But then there is a  $\mathcal{U}_i$  such that  $x \in \mathcal{U}_i$ , and thus eventually  $F_n \subset \mathcal{U}_i$ ; a contradiction. Thus,  $X$  is compact. Now, suppose  $X$  is compact and let  $a : \mathbb{N} \rightarrow X$  be a sequence in  $X$ . Let:

$$F_n = \overline{\{x_k : k \geq n\}} \quad \mathcal{F} = \{F_n : n \in \mathbb{N}\} \quad (32.5.90)$$

Then  $\mathcal{F}$  has the finite intersection property. Since  $X$  is compact, the intersection of the  $F_n$  is non-empty. Let  $x$  be contained in the intersection. Then:

$$B_1(x) \cap \{x_k : k \geq 1\} \neq \emptyset \quad (32.5.91)$$

Pick  $n_1$  such that  $x_{n_1} \in B_1(x)$ . Then there is an  $n_2 > n_1$  such that  $x_{n_2} \in B_{1/2}(x)$ . Continuing, we obtain a subsequence  $n_k$  such that  $x_k \in B_{1/k}(x)$ , and thus  $x_k \rightarrow x$ . Finally, we show that sequential compactness implies that  $X$  is complete and totally bounded. For suppose  $X$  is not totally bounded. Then there exists  $\varepsilon > 0$  such that  $X$  has no finite covering of  $\varepsilon$  balls. We can thus obtain a sequence  $a : \mathbb{N} \rightarrow X$  such that, for all  $n \neq m$ ,  $d(a_n, a_m) \geq \varepsilon$ . But this has no convergence subsequence, for any convergent subsequence would be a Cauchy sequence. Moreover,  $X$  is complete. For suppose not, and let  $a : \mathbb{N} \rightarrow X$  be a Cauchy sequence and suppose it does not converge. But then there is no convergent subsequence, since Cauchy sequences with convergent subsequences converge. Thus,  $X$  is complete.  $\square$

### Theorem 32.5.32: Heine-Borel Theorem

A subset  $\mathcal{E} \subseteq \mathbb{R}^n$  is compact with respect to the standard topology if and only if  $\mathcal{E}$  is closed and bounded. ■

This theorem does not generalize to other spaces. For consider  $\ell^2$  and the closed unit ball about the origin. This is closed and bounded, but it is not compact. This is simply because it is not totally bounded, nor is it sequentially compact.

### Theorem 32.5.33: Extreme Value Theorem

If  $(X, d)$  is a compact metric space and if  $f : X \rightarrow \mathbb{R}$  is continuous, then  $f$  attains its maximum and minimum. In particular, if  $f : X \rightarrow \mathbb{C}$  is continuous, then  $f$  is bounded. ■

*Proof.* Note that if  $f : X \rightarrow \mathbb{C}$  is continuous, then  $|f| : X \rightarrow \mathbb{R}$  is continuous, so we only need to prove the first statement. For if  $X$  is compact, then  $f(X)$  is compact, for  $f$  is continuous. But then  $f(X)$  is closed and bounded. Let:

$$M = \sup_{x \in X} \{f(x)\} \quad (32.5.92)$$

Then, since  $f(X)$  is bounded,  $M \in \mathbb{R}$ . But then there is a sequence  $a : \mathbb{N} \rightarrow X$  such that  $f(a_n) \rightarrow M$ . But if  $X$  is compact, then it is sequentially compact, and thus there is an  $x \in X$  and a subsequence  $a_k$  such that  $a_{k_n} \rightarrow x$ . But then  $f(x) = M$ . Similarly for the minimum value. □

### 32.5.4 Lebesgue Spaces

#### Definition 32.5.31: Lebesgue Number

A Lebesgue Number of an open cover  $\mathcal{O}$  of a metric space  $(X, d)$  is a non-zero number  $d > 0$  such that, for all  $x \in X$ , there exists a  $\mathcal{U} \in \mathcal{O}$  such that:

$$B_d^{(X,d)}(x) \subseteq \mathcal{U} \quad (32.5.93)$$
■

#### Example 32.5.16: L

t  $X = \mathbb{R}$ , and let  $d$  be the standard metric. Let  $\mathcal{O} = \{\mathcal{U}_i : i = 1, 2, 3\}$  where:

$$\mathcal{U}_1 = (-\infty, 1) \quad \mathcal{U}_2 = (0, 2) \quad \mathcal{U}_3 = (1, \infty) \quad (32.5.94)$$

Then  $d = 1/2$  is a Lebesgue number of this cover. Letting  $X = (0, 1)$  with the standard metric, for all  $x \in X$  there is a  $\delta_x > 0$  such that:

$$B_{\delta_x}^{(X,d)}(x) \subseteq X \quad (32.5.95)$$

And thus these open balls are a covering of the unit interval, but this covering has no Lebesgue number. ■

**Theorem 32.5.34: Lebesgue Covering Lemma**

If  $(X, d)$  is a compact metric space, and if  $\mathcal{O}$  is an open covering of  $X$ , then  $\mathcal{O}$  has a Lebesgue number. █

*Proof.* Suppose not. Suppose  $(X, d)$  is compact, and suppose that  $\mathcal{O}$  is a covering of  $X$  with no Lebesgue number. But then, for all  $n \in \mathbb{N}$ , there is an  $a_n$  such that, for all  $\mathcal{U} \in \mathcal{O}$ :

$$B_{1/n}^{(X,d)}(a_n) \not\subseteq \mathcal{U} \quad (32.5.96)$$

But  $X$  is compact, and thus there is a convergent subsequence such that  $a_{k_n} \rightarrow X$ . But then there is a  $\mathcal{U} \in \mathcal{O}$  such that  $x \in \mathcal{U}$ . But  $\mathcal{U}$  is open, and thus there is an  $r > 0$  such that:

$$B_r^{(X,d)}(x) \subseteq \mathcal{U} \quad (32.5.97)$$

Let  $N \in \mathbb{N}$  be such that, for all  $k_n > N$ ,  $d(x_{k_n}, x) < r/2$ . Let  $n > N$  be such that  $1/k_n < r/2$ . But then:

$$B_{1/k_n}(a_{k_n}) \subseteq \mathcal{U} \quad (32.5.98)$$

A contradiction. □

**Theorem 32.5.35.** *If  $(X, d)$  is a compact metric space, if  $(Y, \rho)$  is a metric space, and if  $f : X \rightarrow Y$  is a continuous function, then  $f$  is uniformly continuous.*

*Proof.* For let  $\varepsilon > 0$ . since  $f$  is continuous, for all  $x \in X$  there is a  $\delta_x$  such that, for all  $y \in X$  such that  $d(x, y) < \delta_x$ , it is true that  $\rho(f(x), f(y)) < \varepsilon/2$ . But then:

$$X \subseteq \bigcup_{x \in X} B_{\delta_x}^{(X,d)}(x) \quad (32.5.99)$$

But  $X$  is compact, and thus this covering has a Lebesgue number. Let  $\delta$  be such a Lebesgue number. But then if  $d(x, y) < \delta$ , then there is a  $z \in X$  such that  $x, y \in B_\delta(z)$ . But then:

$$\rho(f(x), f(y)) \leq \rho(f(x), f(z)) + \rho(f(z), f(y)) < \varepsilon \quad (32.5.100)$$

□

### 32.5.5 Equicontinuity

**Theorem 32.5.36: Arzela-Ascoli Theorem**

If  $X$  is a compact metric space, if  $F_n \in C(X)$  is a sequence of equicontinuous point-wise bounded functions, then  $F_n$  has a uniformly convergent subsequence. █

**Theorem 32.5.37.** *If  $X$  is a compact metric space and  $\mathcal{F} \subseteq C(X)$  is a closed subset with respect to the uniform norm, and if  $\mathcal{F}$  is equicontinuous on  $X$  and point-wise bounded, then  $\mathcal{F}$  is compact.*

*Proof.* It suffices to show that  $\mathcal{F}$  is sequentially compact. Let  $F_n$  be a sequence in  $\mathcal{F}$ . Then by the Arzela-Ascoli theorem, there is a uniformly convergent subsequence  $F_{k_n}$ . But  $\mathcal{F}$  is closed, and thus the limit function is contained in  $\mathcal{F}$ . Thus,  $\mathcal{F}$  is sequentially compact. But sequentially compact metric spaces are compact. Therefore, etc.  $\square$

**Theorem 32.5.38.** *If  $X$  is a compact metric space and  $\mathcal{F} \subseteq C(X)$  is a closed subset with respect to the uniform norm, and if  $\mathcal{F}$  is equicontinuous on  $X$  and point-wise bounded, then  $\mathcal{F}$  is uniformly bounded.*

*Proof.* For  $\mathcal{F}$  is compact by the previous theorem. But then  $\mathcal{F}$  is bounded with respect to  $\|\cdot\|_\infty$ . Therefore,  $\mathcal{F}$  is uniformly bounded.  $\square$

**Theorem 32.5.39.** *If  $X$  is a compact metric space and if  $\mathcal{F} \subseteq C(X)$  is closed, equicontinuous, and uniformly bounded on  $X$ , then  $\mathcal{F}$  is compact.*

*Proof.* For suppose  $\mathcal{F}$  is compact. Then  $\mathcal{F}$  is closed and uniformly bounded. Thus it suffices to show that  $\mathcal{F}$  is equicontinuous. Suppose not. Then there is a point  $x \in X$  such that  $\mathcal{F}$  is not equicontinuous at  $x$ . Thus, there exists an  $\varepsilon > 0$  such that, for all  $\delta > 0$ , there are points  $x, y$  such that  $d(x, y) < \delta$ , but  $|f(x) - f(y)| \geq \varepsilon$  for some  $f \in \mathcal{F}$ . Thus, for all  $n \in \mathbb{N}$ , there is an  $x_n \in X$  such that  $d(x, x_n) < 1/n$ , and  $|f_n(x) - f_n(x_n)| \geq \varepsilon_0$ . But if  $\mathcal{F}$  is compact, then  $f_n$  has a convergent subsequence  $f_{k_n}$ . Let  $f$  be the limit. Since  $\mathcal{F}$  is compact,  $f \in \mathcal{F}$ . But then  $f_{k_n}(x_{k_n}) \rightarrow f(x)$ . But then there is an  $N \in \mathbb{N}$  such that, for  $k_n > N$ ,  $\|f_{k_n} - f\|_\infty < \varepsilon_0/3$ . But then:

$$|f(x_{k_n}) - f(x)| = |f(x_{k_n}) - f_{k_n}(x_{k_n}) + f_{k_n}(x_{k_n}) - f_{k_n}(x) + f_{k_n}(x) - f(x)| \quad (32.5.101)$$

$$\geq |f_{k_n}(x_{k_n}) - f_{k_n}(x)| + |f(x_{k_n}) - f_{k_n}(x_{k_n}) + f_{k_n}(x) - f(x)| \quad (32.5.102)$$

$$> \varepsilon \quad (32.5.103)$$

A contradiction.  $\square$

### 32.5.6 Baire Spaces

#### Definition 32.5.32: Baire Space

A Baire space is a metric space  $(X, d)$  such that, for countable collection of open and dense sets, the intersection is also dense.  $\blacksquare$

This is a topological property, and so Baire spaces can be defined for a more general topological space. The interior of a set in a topological space is:

$$\text{Int}((\circ)A) = \bigcup\{\mathcal{U} \in \tau : \mathcal{U} \subseteq A\} \quad (32.5.104)$$

**Theorem 32.5.40.** *A metric space  $(X, d)$  is a Baire space if and only if given a countable collection  $F_n$  of closed sets such that the union over all of  $F_n$  has non-empty interior, then at least one of the  $F_n$  has non-empty interior.*

**Theorem 32.5.41.** *There exist countable Baire spaces.*

Suppose  $\mathcal{U} \subseteq X$  is open at  $x_0 \in \mathcal{U}$ . There there is a  $\delta > 0$  such  $B_\delta(x) \subseteq \mathcal{U}$ . Then:

$$\overline{B_{\delta/2}(x)} \subseteq B_\delta(x) \quad (32.5.105)$$

Thus,  $\overline{B_{\delta/2}(x)} \subseteq \mathcal{U}$  and the diameter is less than  $2\delta$ .

### Theorem 32.5.42: Baire Category Theorem

---

Every complete metric space is a Baire space. ■

*Proof.* Suppose  $\mathcal{O}_n \subseteq X$  is open and dense for all  $n \in \mathbb{N}$ . Let  $x_0 \in X$  and  $r_0 > 0$ . It will suffice to show that:

$$B_{r_0}(x_0) \cap \bigcap_{n \in \mathbb{N}} \mathcal{O}_n \neq \emptyset \quad (32.5.106)$$

Inductively, we create a sequence of points  $x_k$  and real numbers  $r_k > 0$  such that  $r_k$  is strictly monotonically decreasing, and thus that:

$$\overline{B_{r_{k+1}}(x_{k+1})} \subseteq B_{r_k}(x_k) \cap \mathcal{O}_{k+1} \quad (32.5.107)$$

□

Consider the set of all lines through the origin with rational slope. The complete of any given line is the union of two open half planes, which are open and dense subsets of  $\mathbb{R}^2$ . Since we have only a countable collection of such lines, the intersection of the complement is dense in  $\mathbb{R}^2$ . Baire's Category Theorem holds even if  $(X, d)$  is not complete, but is equivalent to a complete metric. For example, let  $X = (0, 1)$  and let  $d(x, y) = |x - y|$  be the standard metric. This is not a complete space, but is homeomorphic to  $\mathbb{R}$ , which is a complete metric space. Using this homeomorphism, we can find a metric  $\tilde{d}$  on  $(0, 1)$  that is complete and which is equivalent to the original metric. Thus,  $(0, 1)$  is a Baire space.

**Theorem 32.5.43.** *If  $V$  is a non-empty open subset of a complete metric space  $(X, d)$ , then there is a metric  $\tilde{d}$  such that  $(V, \tilde{d})$  is complete.*

Hence,  $V$  is a Baire space. Then, given a set  $F_n$  of closed subsets of  $X$  such that:

$$V = \bigcup_{n=1}^{\infty} (V \cap F_n) \quad (32.5.108)$$

Then some  $F_n \cap V$  has non-empty interior in  $V$ , and hence in  $X$ .

**Theorem 32.5.44.** *If  $X$  is a Baire space and if  $f_n$  is a sequence of continuous function in  $C(X)$  which converges point-wise to  $f : X \rightarrow \mathbb{C}$ , then the set:*

$$\{x \in X : f \text{ is continuous at } x\} \quad (32.5.109)$$

Is dense in  $X$ .

*Proof.* Let  $\varepsilon > 0$  and define:

$$A_N(\varepsilon) = \{x : |f_n(x) - f_m(x)| \leq \varepsilon, n, m \in \mathbb{N}\} \quad (32.5.110)$$

$$= \bigcap_{n, m \geq N} \{x : |f_n(x) - f_m(x)| \leq \varepsilon\} \quad (32.5.111)$$

Then  $A_N(\varepsilon)$  is closed. But also:

$$X = \bigcup_{N=1}^{\infty} A_N(\varepsilon) \quad (32.5.112)$$

Thus, by the Baire category theorem, we have:

$$\mathcal{U}(\varepsilon) = \bigcup_{N=1}^{\infty} \text{Int}((A_N(\varepsilon))) \quad (32.5.113)$$

Is non-empty and open. Moreover,  $\mathcal{U}(\varepsilon)$  is dense. But then:

$$\mathcal{C} = \bigcap_{n=1}^{\infty} \mathcal{U}\left(\frac{1}{n}\right) \quad (32.5.114)$$

Is dense in  $X$ , and  $f$  is continuous at all  $x \in \mathcal{C}$ .  $\square$

The Baire Category Theorem says that every complete metric space is a Baire space. The notion of Baire space is a topological property, and not a metric property. Thus, even if  $(X, d)$  is not complete but is equivalent to a complete metric space  $(X, \tilde{d})$ , then  $(X, d)$  is a Baire space. A topological space is called completely metrizable if there is a metric on the space that is complete and generates the topology. Given a complete metric space  $(X, d)$ , every non-empty open set  $\mathcal{V}$  has a metric  $d_{\mathcal{V}}$  such that  $(\mathcal{V}, d_{\mathcal{V}})$  is complete, and is therefore a Baire space. Thus, if:

$$\mathcal{V} = \bigcup_{n \in \mathbb{N}} (\mathcal{V} \cap F_n) \quad (32.5.115)$$

Where  $F_n$  is closed for all  $n \in \mathbb{N}$ , then for some  $N \in \mathbb{N}$ ,  $\mathcal{V} \cap F_N$  has interior.

**Theorem 32.5.45.** *If  $X$  is a Baire space, and if  $F_n$  is a sequence of continuous functions that converges point-wise to  $f : X \rightarrow \mathbb{C}$ , then the set  $\mathcal{D}$  defined by:*

$$\mathcal{D} = \{x \in X : f \text{ is continuous as } x\} \quad (32.5.116)$$

*Then  $\mathcal{D}$  is dense in  $X$ .*

*Proof.* For let  $\varepsilon > 0$ , and let:

$$A_N(\varepsilon) = \{x \in X : |f_n(x) - f_m(x)| \leq \varepsilon, n, m > N\} \quad (32.5.117)$$

Then, for all  $N \in \mathbb{N}$ ,  $A_N(\varepsilon)$  is closed. Let  $\mathcal{U}(\varepsilon)$  be defined by:

$$\mathcal{U} = \bigcup_{n \in \mathbb{N}} \text{Int}((A_N(\varepsilon))) \quad (32.5.118)$$

Then  $\mathcal{U}(\varepsilon)$  is open and dense. It is open for it is the union of open sets. For let  $\mathcal{V}$  be a non-empty subset. Then:

$$\mathcal{V} = \bigcup_{n \in \mathbb{N}} (A_n(\varepsilon) \cap \mathcal{V}) \quad (32.5.119)$$

Hence there exists an  $N \in \mathbb{N}$  such that:

$$A_N(\varepsilon) \cap \mathcal{V} \neq \emptyset \quad (32.5.120)$$

And this has interior, and therefore:

$$\text{Int}((A_N(\varepsilon)) \cap \mathcal{V}) \neq \emptyset \quad (32.5.121)$$

Therefore,  $\mathcal{V} \cap \mathcal{U}(\varepsilon) \neq \emptyset$ . Now, define:

$$\mathcal{V} = \bigcap_{n \in \mathbb{N}} \mathcal{U}\left(\frac{1}{n}\right) \quad (32.5.122)$$

And therefore  $\mathcal{C}$  is dense in  $X$ . We now want to show that  $f$  is continuous for all  $x \in \mathcal{C}$ . For let  $x_0 \in \mathcal{C}$  and let  $\varepsilon > 0$ . Let  $k \in \mathbb{N}$  be such that  $k^{-1} < \varepsilon$ . Then  $x_0 \in \mathcal{U}(k^{-1})$  and thus there is an  $N \in \mathbb{N}$  such that:

$$x_0 \in \text{Int}((A_N(k^{-1})) \quad (32.5.123)$$

But  $f_N$  is continuous, and thus there is a neighborhood  $\omega$  of  $x_0$  such that, for all  $y \in \omega$ :

$$|f_N(x_0) - f_N(y)| < \varepsilon/3 \quad (32.5.124)$$

Shrink  $\omega$  so that it resides inside of  $\text{Int}((A_N(k^{-1}))$ . Then:

$$|f_n(y) - f_N(y)| < k^{-1} \quad n \geq N \quad (32.5.125)$$

But then, use the Cauchy trick and you're down. □

## 32.6 Normed Vector Spaces

### 32.6.1 Basic Definitions

#### Definition 32.6.1: Normed Vector Spaces

A normed vector space over a field  $\mathbb{F} \subseteq \mathbb{C}$ , denoted  $(V, \|\cdot\|)$  is a vector space  $V$  over  $\mathbb{F}$  and a norm  $\|\cdot\|$  on  $V$ . ■

### 32.6.2 Banach Spaces

#### Definition 32.6.2: Banach Space

A Banach space is a normed vector space  $(V, \|\cdot\|)$  such that the metric  $d$  induced by the norm  $\|\cdot\|$  is complete on  $V$ . ■

Normed spaces are special. Give  $\mathbf{v} \in V$ , and for  $r > 0$ , we have:

$$B_r^{(V, \|\cdot\|)}(\mathbf{x}) = B_r^{(V, \|\cdot\|)}(\mathbf{0}) + \mathbf{x} \quad (32.6.1)$$

That is, open balls about arbitrary points are merely translations of an open ball about the origin.

$$\|\mathbf{v}\| - \|\mathbf{u}\| \leq \|\mathbf{v} - \mathbf{u}\| \quad (32.6.2)$$

And thus the map  $\mathbf{v} \mapsto \|\mathbf{v}\|$  is continuous. The closure of an open ball is the closed ball.

$$\overline{B_r^{(V, \|\cdot\|)}(\mathbf{x})} = \{\mathbf{y} \in V : \|\mathbf{x} - \mathbf{y}\| \leq r\} \quad (32.6.3)$$

We can also multiply open balls by constants, to get the following:

$$\varepsilon B_r^{(V, \|\cdot\|)}(\mathbf{x}) = B_{\varepsilon r}^{(V, \|\cdot\|)}(\mathbf{x}) \quad (32.6.4)$$

**Theorem 32.6.1.** Suppose  $X$  and  $Y$  are normed vector spaces over  $\mathbb{F} \subseteq \mathbb{C}$ . Let  $T : X \rightarrow Y$  be a linear transformation. Then the following are equivalent:

1.  $T$  is continuous.
2.  $T$  is continuous at some  $x_0 \in X$ .
3. There is an  $\alpha > 0$  such that  $\|Tx\| \leq \alpha \|x\|$ .

*Proof.* Suppose  $T$  is continuous at  $x_0$ . Then there is a  $\delta > 0$  such that:

$$T\left(\overline{B_\delta(x_0)}\right) \subseteq B_1(T(x_0)) \quad (32.6.5)$$

But:

$$T(\overline{B_\delta(x_0)}) = T(\overline{B_\delta(0)}) + T(x_0) \quad (32.6.6)$$

$$B_1(T(x_0)) = B_1(0) + T(x_0) \quad (32.6.7)$$

Now suppose  $z \neq 0$ . Then:

$$\|T(z)\| = \left\| \frac{\|z\|}{\delta} T\left(\frac{\delta z}{\|z\|}\right) \right\| \leq \frac{1}{\delta} \|z\| \quad (32.6.8)$$

Let  $\alpha = \delta^{-1}$ . Proving the next one:

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq \alpha \|x - y\| \quad (32.6.9)$$

And so we have continuity.  $\square$

There are linear maps that are not bounded. Let  $\ell_1^0 = \text{Span}\{e_k : x \in \mathbb{N}\}$ . Map  $e_k \rightarrow k e_k$ . Let  $\|\cdot\|_a$  and  $\|\cdot\|_b$  be norms on  $X$  that induce the same topology on  $X$ . Consider the map  $id : (X, \|\cdot\|_a) \rightarrow (X, \|\cdot\|_b)$ . Since the topologies are the same,  $id$  is continuous. Then there is a  $c \geq 0$  such that:

$$\|x\|_b \leq c \|x\|_a \quad (32.6.10)$$

We can go the other way as well, and thus we see that equivalence implies strongly equivalent. This is not true in a general metric space.

### Definition 32.6.3: Operator Norm

Let  $\mathcal{L}(X, Y)$  be the set of bounded linear transformation  $T : X \rightarrow Y$ . The operator norm on  $T$  is:

$$\|T\| = \sup\{\|T\|(x) : \|x\| \leq 1\} \quad (32.6.11)$$

$\blacksquare$

**Theorem 32.6.2.** *The operator norm on  $\mathcal{L}(X, Y)$  is a norm.*

*Proof.* For we have:

$$\|S \circ T\| \leq \|S\| \|T\| \quad (32.6.12)$$

$\square$

### Definition 32.6.4: Algebra Over a Field

An algebra over a field  $\mathbb{F}$  is a vector space  $A$  over  $\mathbb{F}$  such that  $A$  has a ring

structure  $(A, \times, +)$  such that:

$$\lambda(xy) = (\lambda x)y = x(\lambda(y)) \quad x, y \in A \quad \lambda \in \mathbb{F} \quad (32.6.13)$$

■

**Example 32.6.1**  $\mathbb{R}[x]$ ,  $\mathbb{C}[x]$ ,  $M_n(\mathbb{F})$ , and  $C_b(X)$ .

### Definition 32.6.5: Normed Algebra

A normed algebra is a normed space  $(A, \|\cdot\|)$  such that  $A$  is an algebra and such that, for all  $x, y \in A$ :

$$\|xy\| \leq \|x\|\|y\| \quad (32.6.14)$$

■

### Definition 32.6.6: Banach Algebra

A Banach Algebra is a normed algebra  $(A, \|\cdot\|)$  such that  $(A, \|\cdot\|)$  is a Banach space.

■

**Theorem 32.6.3.** *If  $Y$  is a Banach space, then  $\mathcal{L}(X, Y)$  is a Banach space.*

**Theorem 32.6.4.** *If  $X$  is a Banach Algebra, then  $\mathcal{L}(A)$  is a Banach algebra.*

*Proof.* For suppose  $T_n$  is Cauchy in  $\mathcal{L}(X, Y)$ . Then for all  $x \in X$ ,  $T_n(x)$  is Cauchy in  $Y$ , and thus  $T_n(x)$  converges to some  $y \in Y$ . Let  $T : X \rightarrow Y$  be this limit function. Then  $T : X \rightarrow Y$  is a linear map. But since  $T_n$  is Cauchy, it is uniformly bounded. But then there is an  $M \in \mathbb{R}^+$  such that:

$$\|T_n\| \leq M \quad (32.6.15)$$

And thus:

$$\|Tx\| = \lim_{n \rightarrow \infty} \|T_n x\| \leq \limsup_n \|T_n\| \|x\| \leq M \|x\| \quad (32.6.16)$$

Therefore, etc. □

Let  $X_\lambda$  be a Banach space for all  $\lambda \in \Lambda$ . Then the product is:

$$\prod_{\lambda \in \Lambda} X_\lambda = \{f : \Lambda \rightarrow \bigcup_{\lambda \in \Lambda} X_\lambda : f(\lambda) \in X_\lambda\} \quad (32.6.17)$$

Generally, we think of the indexing set to be finite,  $\Lambda = \{1, \dots, n\}$ . The product space is then:

$$\prod_{\lambda=1}^n X_\lambda = X_1 \times \dots \times X_n \quad (32.6.18)$$

Functions are therefore  $n$  tuples. There's no reason to expect that this will be a Banach space in any reasonable way. Thus we define the Banach Space Direct-Product.

### Definition 32.6.7: Banach Space Direct Product

The Banach Space Direct Product of a set of Banach spaces  $X_\lambda$  indexed over  $\Lambda$  is:

$$\prod_{\lambda \in \Lambda}^* X_\lambda = \{f \in \prod_{\lambda \in \Lambda} X_\lambda : \sup_{\lambda \in \Lambda} f(\lambda) < \infty\} \quad (32.6.19)$$

■

Then  $\|x\| = \sup_\lambda \|x_\lambda\|$  is a norm on the product space.

### Theorem 32.6.5: Open Mapping Theorem

If  $X$  and  $Y$  are Banach spaces and if  $T \in \mathcal{L}(X, Y)$  is surjective, then  $T$  is an open map. ■

*Proof.* It will suffice to find  $r > 0$  such that:

$$B_r^Y \subseteq T(B_1^X(0)) \quad (32.6.20)$$

By homogeneity,  $T(B_\delta^X(0))$  is a neighborhood of 0 for all  $\delta > 0$ . By linearity,  $T(B_\delta(x))$  is a neighborhood of  $T(x)$  for all  $x \in X$  and for all  $\delta > 0$ . But if  $V \subseteq X$  is open, and  $x \in V$ , then there is a  $\delta > 0$  such that  $B_\delta(x) \subseteq V$ . Thus,  $T(B_\delta(x))$  is a neighborhood of  $T(x)$  in  $T(V)$ . There is also an  $r > 0$  such that:

$$B_r^Y(0) \subseteq \overline{T(B_1^X(0))} \quad (32.6.21)$$

For let  $\alpha \in (0, 1)$ . Note that:

$$T(B_\alpha(x)) = \alpha T(B_1(x)) \quad (32.6.22)$$

And also:

$$B_{\alpha r}(0) \subseteq \overline{\alpha T(B_1(0))} \quad (32.6.23)$$

Let  $y \in B_r(0)$ . Then there is a  $y_1 \in T(B_1(0))$  such that:

$$\|y - y_1\| < \frac{r}{2} \quad (32.6.24)$$

But then  $y - y_1 \in B_{r/2}(0)$ . But then there is a  $y_2 \in T(B_{r/2}(0))$  such that:

$$\|y_2\| < \frac{r}{4} \quad (32.6.25)$$

That is:

$$\|y - y_1 - y_2\| < \frac{r}{2^2} \quad (32.6.26)$$

Continuing we obtain a sequence  $y_n$  such that:

$$y_n \in T(B_{1/2^n}(0)) \quad (32.6.27)$$

And such that:

$$\|y - \sum_{k=1}^n y_k\| < \frac{r}{2^n} \quad (32.6.28)$$

Note that there exists  $x_n \in X$  such that  $T(x_n) = y_n$  and:

$$\|x_n\| < \frac{1}{2^{n-1}} \quad (32.6.29)$$

But then there is an  $x \in X$  such that:

$$x = \sum_{n=1}^{\infty} x_n \quad (32.6.30)$$

Since  $X$  is complete and since this series converges. But  $T$  is continuous, and therefore  $T(x) = y$ . But:

$$\|x\| \leq \sum_{n=1}^{\infty} \|x_n\| < \sum_{n=0}^{\infty} \frac{1}{2^n} = 2 \quad (32.6.31)$$

That is,::

$$B_r(0) \subseteq T(B_2(0)) \quad (32.6.32)$$

And therefore:

$$B_{r/2}(0) \subseteq T_1(B_1(0)) \quad (32.6.33)$$

If  $T$  is surjective, we can write:

$$Y = \bigcup_{n \in \mathbb{N}} \overline{T(B_n(0))} \quad (32.6.34)$$

But  $Y$  is a Banach space, and thus by the Baire category theorem, there is an  $n \in \mathbb{N}$  such that  $\overline{T(B_n(0))}$  has interior. Therefore, etc.  $\square$

### Example 32.6.2: R

---

call that:

$$\ell_0^p = \{x \in \ell^p : \exists N \in \mathbb{N}, \forall_{n>N}, x_n = 0\} \quad (32.6.35)$$

Let  $\ell_0^p$  be defined by:

$$\ell_0^p = \text{Span}\{e_n : n \in \mathbb{N}\} \quad (32.6.36)$$

Define  $T : \ell_0^2 \rightarrow \ell_0^p$  by:

$$T(e_n) = \frac{1}{n} e_n \quad (32.6.37)$$

Then  $T$  is bounded and  $\|T\| \leq 1$ . Note that  $T$  is bijective and has an inverse:

$$T^{-1}(e_n) = ne_n \quad (32.6.38)$$

And this is not bounded, so  $T^{-1} \notin \mathcal{L}(\ell_0^p)$ . This can happen since  $\ell_0^P$  is not complete. ■

### Theorem 32.6.6: Inverse Mapping Theorem

If  $X$  and  $Y$  are Banach spaces and  $T \in \mathcal{L}(X, Y)$  is bijective, then  $T^{-1} \in \mathcal{L}(Y, X)$ . ■

*Proof.* Note that  $T^{-1}$  is linear. Since  $T$  has to be open by the open mapping theorem,  $T^{-1}$  is continuous. But continuous linear functions are bounded. Therefore,  $T^{-1}$  is bounded. □

### Theorem 32.6.7: Closed Graph Theorem

If  $X$  and  $Y$  are Banach spaces and if  $T : X \rightarrow Y$  is linear, then  $T \in \mathcal{L}(X, Y)$  if and only if the graph of  $T$  is closed in  $X \times Y$ . ■

*Proof.* Note that  $X \times Y$  is a Banach space and  $(x_n, y_n) \rightarrow (x, y)$  if and only if  $x_n \rightarrow x$  and  $y_n \rightarrow y$ . If  $T$  is bounded, then the graph of  $T$  is closed. Now, suppose that the graph of  $T$  is closed. But then the graph is a Banach space. The projection map  $P : T \rightarrow X$  given by  $P((x, T(x))) = x$  is a bounded bijection. Hence,  $P^{-1}$  is bounded. Let  $P_2$  be defined by  $P_2(x, T(x)) = T(x)$ . Then  $P_2$  is bounded. But:

$$T(x) = P_2 \circ P_1^{-1}(x) \quad (32.6.39)$$

And therefore  $T$  is bounded. □

### Example 32.6.3: S

Suppose  $T_n \in \mathcal{L}(X, Y)$  and suppose, for all  $x \in X$ :

$$T(x) = \lim_{n \rightarrow \infty} T_n(x) \quad (32.6.40)$$

Then we have that  $T$  is linear. Is  $T \in \mathcal{L}(X, Y)$ ? We have:

$$\|T\| = \left\| \lim_{n \rightarrow \infty} T_n \right\| \leq \lim_{n \rightarrow \infty} \sup_n \|T_n\| \quad (32.6.41)$$



### Theorem 32.6.8: Principle of Uniform Boundedness

If  $X$  and  $Y$  are Banach spaces, if  $T_\lambda \in \mathcal{L}(X, Y)$  for all  $\lambda \in \Lambda$ , and if for all  $x \in X$  we have:

$$\|\{\|T_\lambda(x)\| : \lambda \in \Lambda\}\| < \infty \quad (32.6.42)$$

Then  $\{\|T_\lambda\| : \lambda \in \Lambda\}$  is bounded. ■

Recall that if  $X_\lambda$  are Banach spaces for all  $\lambda \in \Lambda$ , where  $\Lambda$  is some indexing set, then we can form the Banach Space direct product:

$$\prod_{\lambda \in \Lambda}^* X_\lambda = \left\{ x \in \prod_{\lambda \in \Lambda} X_\lambda : \|x\|_\lambda < \infty \right\} \quad (32.6.43)$$

If  $\lambda_0 \in \Lambda$ , we can define:

$$P_\lambda : \prod_{\lambda \in \Lambda}^* X_\lambda \rightarrow X_{\lambda_0} \quad (32.6.44)$$

By defining  $P_{\lambda_0}(x) = x(\lambda_0)$ . Now consider the case when  $X_\lambda = Y$  for all  $\lambda \in \Lambda$ . We'll write:

$$Y_\Lambda = \prod_{\lambda \in \Lambda} Y \quad (32.6.45)$$

The principle of uniform boundedness says that if  $X$  and  $Y$  are Banach spaces and if  $T_\lambda \in \mathcal{L}(X, Y)$  for all  $\lambda \in \Lambda$ , and that for all  $x \in X$ , the set  $\{\|T_\lambda(x)\| : \lambda \in \Lambda\}$  is bounded, then  $\{\|T_\lambda\| : \lambda \in \Lambda\}$  is bounded.

### Theorem 32.6.9: Principle of Uniform Boundedness

If  $X$  and  $Y$  are Banach spaces, if  $T_\lambda \in \mathcal{L}(X, Y)$  for all  $\lambda \in \Lambda$ , and if for all  $x \in X$  we have that  $\{\|T_\lambda(x)\| : \lambda \in \Lambda\}$  is bounded, then  $\{\|T_\lambda\| : \lambda \in \Lambda\}$  is bounded. ■

*Proof.* For let:

$$Y_\Lambda = \prod_{\lambda \in \Lambda} Y \quad (32.6.46)$$

And define  $T : X \rightarrow Y_\Lambda$  by:

$$T(x) = T_\lambda(x) \quad (32.6.47)$$

Then  $T$  is well defined and linear. To see that  $T$  is bounded, use the closed graph theorem. That is, suppose  $x_n \rightarrow x$ . We need to show that  $T(x_n) \rightarrow T(x)$ . Let  $\lambda \in \Lambda$ . Then  $P_\lambda(T(x_n)) \rightarrow P_\lambda(y)$ . Then  $T_n(x_n)(\lambda) \rightarrow y(\lambda)$ . But  $T(x_n)(\lambda) = T_\lambda(x_n)$ , and this converges to  $T_\lambda(x)$ . But then  $y(\lambda) = T_\lambda$  for all  $\lambda$ , and thus  $y = T(x)$ . Therefore  $T \in \mathcal{L}(X, Y_\Lambda)$ . But if  $\|x\| \leq 1$ , then:

$$\|T_\lambda(x)\| \leq \sup_{\lambda \in \Lambda} \|T_\lambda(x)\| = \sup_{\lambda \in \Lambda} \|T(x)(\lambda)\| = \|T(x)\| \leq \|T\| \quad (32.6.48)$$

Thus,  $\|T\| \geq \|T_\lambda\|$  for all  $\lambda \in \Lambda$ . Therefore, etc.  $\square$

### Theorem 32.6.10: Banach-Stienhaus Theorem

If  $X$  and  $Y$  are Banach spaces, and if  $T_n \in \mathcal{L}(X, Y)$  converges point-wise to  $T$ , then  $T \in \mathcal{L}(X, Y)$ .  $\blacksquare$

*Proof.* By the principle of uniform boundedness,  $T_n$  is uniformly bounded. Therefore  $T$  is bounded.  $\square$

## 32.7 Zorn's Lemma

A relation on a set  $X$  is a subset  $R \subseteq X \times X$ . Given an element  $(x, y) \in R$ , we often write  $xRy$  to denote this. Here we'll write  $x \leq y$ .

### Definition 32.7.1: Ordered Sets

An ordered set is a set  $X$  and a relation  $\leq$  on  $X$ , denoted  $(X, \leq)$  such that the following are true:

1. For all  $x \in X$ ,  $x \leq x$ .
2. For all  $x, y \in X$  such that  $x \leq y$  and  $y \leq x$ , it is true that  $x = y$ .
3. For all  $x, y, z \in X$  such that  $x \leq y$  and  $y \leq z$ , it is true that  $x \leq z$ .



### Definition 32.7.2: Majorants in Ordered Sets

A majorant of a subset  $Y \subseteq X$  of an ordered set  $(X, \leq)$  is an element  $x \in X$  such that, for all  $y \in Y$ , it is true that  $y \leq x$ .  $\blacksquare$

**Example 32.7.1: L**

Let  $(X, d)$  be a metric space, and let  $x_0 \in X$ . Define the following:

$$\mathcal{N}(x_0) = \{\mathcal{V} \subseteq X : \mathcal{V} \text{ is a neighborhood of } x_0\} \quad (32.7.1)$$

We can order  $\mathcal{N}$  by reverse containment. That is, We have the following relation:

$$\leq = \{(\mathcal{U}, \mathcal{V}) \in \mathcal{N}(x_0) \times \mathcal{N}(x_0) : \mathcal{V} \subseteq \mathcal{U}\} \quad (32.7.2)$$

That is, we write  $\mathcal{U} \leq \mathcal{V}$  if  $\mathcal{V}$  is a subset of  $\mathcal{U}$ . Note that, for all  $x_0$ ,  $\mathcal{N}(x_0)$  has a least element, or a minorant, but  $X$  is such an element. But, if  $\{x_0\}$  is not open, then there is no majorant. ■

**Definition 32.7.3: Totally Ordered Sets**

A totally ordered set is an ordered set  $(X, \leq)$  such that, for all  $x, y \in X$ , either  $x \leq y$  or  $y \leq x$ . ■

**Definition 32.7.4: Maximal Element**

A maximal element of a subset  $Y \subseteq X$  of a totally ordered set  $(X, \leq)$  is an element  $y$  such that:

$$\{y' \in Y : y \leq y'\} = \{y\} \quad (32.7.3)$$

Note that  $y$  is not necessary a majorant for  $Y$  nor is  $y$  necessarily unique. ■

**Definition 32.7.5: Inductively Ordered Sets**

An inductively ordered set is an ordered set  $(X, \leq)$  such that, for all totally ordered subsets  $S \subseteq X$ , there is a majorant  $x \in X$  of  $S$ . ■

That is, there exists  $x \in X$  such that, for all  $y \in S$ ,  $y \leq x$ .

**Example 32.7.2: L**

Let  $X = \mathbb{R}$  and consider the set  $\mathcal{N}(0)$ . Then  $\mathcal{N}(0)$  is not inductively ordered. ■

**Theorem 32.7.1: Zorn's Lemma**

If  $(X, \leq)$  is an inductively ordered set, then there is a maximal element  $x \in X$ . ■

For more review, see Folland's Real Analysis and Pedersen's Analysis Now.

Zorn's lemma is used to prove that every vector space has a basis. We'll use this to discuss the notion of duals on normed vector spaces.

### Definition 32.7.6: Dual of a Normed Vector Space

The dual of a normed vector space  $(X, \|\cdot\|)$  is the set  $X^* = \mathcal{L}(X, \mathbb{F})$  with the operator norm. ■

**Theorem 32.7.2.** *If  $(X, \|\cdot\|)_X$  is a normed vector space, then  $(X^*, \|\cdot\|)$  is a Banach space.*

### Example 32.7.3: L

Let  $X = \mathbb{F}^n$  and let  $\{e_1, \dots, e_n\}$  be the standard basis. Then we can define  $e_k^* \in (\mathbb{F}^n)^*$  by:

$$e_k^*(\alpha_1 e_1 + \dots + \alpha_n e_n) = \alpha_k \quad (32.7.4)$$

Then  $\{e_1^*, \dots, e_n^*\}$  is a basis for  $(\mathbb{F})^*$ , and thus  $(\mathbb{F}^n)^* \simeq \mathbb{F}^n$ . ■

### Example 32.7.4: L

Let  $\{e_\lambda\}_{\lambda \in \Lambda}$  be a Hamel basis. Then we can define a linear basis:

$$e_\lambda^* : X \rightarrow \mathbb{F} \quad (32.7.5)$$

But also, the set  $\{\lambda \in \Lambda : e_\lambda^* \in X^*\}$  is at most finite. ■

One question that arises is, given any normed vector space  $X$ , what can we say about the dual space  $X^*$ ? We know that the zero operator is in there, but is there anything else?

### Definition 32.7.7: Minkowski Functional

A Minkowski function on a normed vector space  $(X, \|\cdot\|)$  over a field  $\mathbb{F} \subseteq \mathbb{C}$  is a function  $m : X \rightarrow \mathbb{R}$  such that the following are true:

$$m(x + y) \leq m(x) + m(y) \quad (32.7.6a)$$

$$m(tx) = tm(x) \quad t \geq 0 \quad (32.7.6b)$$



**Example 32.7.5: I**

$\|\cdot\|$  is a semi-norm on  $X$  over  $\mathbb{R}$  or  $\mathbb{C}$ , then  $m(x) = \|x\|$  is a Minkowski functional on  $X$ . If we let  $X = \ell_{\mathbb{R}}^{\infty}$ , then:

$$m(x) = \lim_{n \rightarrow \infty} \sup_{k \leq n} \{x_n\} \quad (32.7.7)$$

Is also a Minkowski functional. ■

**Theorem 32.7.3: Basic Extension Lemma**

If  $m : X \rightarrow \mathbb{R}$  is a Minkowski functional on a vector space  $X$  over  $\mathbb{R}$  if  $Y \subseteq X$  is a subspace, and if  $\varphi : Y \rightarrow \mathbb{R}$  is a linear functional such that, for all  $y \in Y$ ,  $\varphi(y) \leq m(y)$ , then there is a linear functional  $\tilde{\varphi} : X \rightarrow \mathbb{R}$  such that, for all  $x \in X$ ,  $\tilde{\varphi}(x) \leq m(x)$ , and for all  $y \in Y$ ,  $\tilde{\varphi}(y) = \varphi(y)$ . ■

If  $X$  is a normed vector space, then we can identify  $X$  with its image  $i(X)$  in  $X^{**}$ , where  $i : X \rightarrow X^{**}$  is defined by:

$$i(x)(\phi) = \phi(x) \quad (32.7.8)$$

If  $X$  is a Banach space, then  $i(X)$  is closed. Otherwise, we call  $\tilde{X} = \overline{i(X)}$  is the completion of  $X$ .

**Definition 32.7.8: Reflexive Banach Space**

A reflexive banach space is a Banach space  $X$  such that the natural map  $i : X \rightarrow X^{**}$  is surjective. ■

If  $X$  is reflexive, then  $X$  is isometrically isomorphic to  $X^{**}$ . One might suspect that the converse is true, but that's not what the definition says. Indeed, the converse is not true. There are Banach spaces  $X$  that are isometrically isomorphic to  $X^{**}$  that are not reflexive.

**Example 32.7.6: L**

Let  $X = \ell^p$ , with  $1 < p < \infty$ , and let  $q$  be the conjugate exponent of  $p$ . If  $y \in \ell^q$ , there is a linear functional  $\varphi_y^p \in (\ell^p)^*$  defined by:

$$\varphi_y^p(x) = \sum_{n=1}^{\infty} x_n y_n \quad (32.7.9)$$

From Hölder's inequality, this sum does indeed converge, and:

$$\|\varphi_y^p\| \leq \|y\|_q \quad (32.7.10)$$

Moreover, equality is obtained:

$$\|\varphi_y^p\| = \|y\|_q \quad (32.7.11)$$

and the map  $y \mapsto \varphi_y^p$  is an isometric isomorphism of  $\ell^q$  with  $(\ell^p)^*$ . Consider  $i(x) \in (\ell^p)^{**} = (\ell^q)^*$  by the identity above. Note that:

$$i(x)(\varphi_y^p) = \varphi_y^p(x) = \varphi_x^q(y) \quad (32.7.12)$$

And thus  $i : X \rightarrow X^{**}$  is surjective, so  $X$  is reflexive. ■

### Definition 32.7.9: Transpose of a Bounded Linear Operator

The transpose over a bounded linear operator  $T : X \rightarrow Y$  between normed vector spaces  $X$  and  $Y$  is the function  $T^* : Y^* \rightarrow X^*$  defined by:

$$T^*(\varphi)(x) = \varphi(T(x)) \quad (32.7.13)$$

For all  $\varphi \in Y^*$ . ■

**Theorem 32.7.4.** *If  $X$  and  $Y$  are normed vector spaced and if  $T \in \mathcal{L}(X, Y)$ , then  $T^* \in \mathcal{L}(Y^*, X^*)$  and  $\|T^*\| = \|T\|$ .*

*Proof.* For:

$$\|T^*(\varphi)\| = \sup_{\|x\|=1} |T^*(\varphi)(x)| \quad (32.7.14a)$$

$$= \sup_{\|x\|=1} |\varphi(T(x))| \leq \sup_{\|x\|=1} \|\varphi\| \|T\| \|x\| \quad (32.7.14b)$$

$$= \sup_{\|x\|=1} \|\varphi\| \|T\| \quad (32.7.14c)$$

$$\leq \|T\| \|\varphi\| \quad (32.7.14d)$$

Therefore,  $\|T^*\| \leq \|T\|$ . Let  $\varepsilon > 0$ . Then there is an  $x$  such that  $\|x\| = 1$  and  $\|T\| < \|T(x)\| + \varepsilon$ . But there exists  $\varphi \in Y^*$  such that  $\|\varphi\| = 1$  and  $\varphi(T(x)) = \|T(x)\|$ . But then:

$$\|T^*\| \geq \|T^*(\varphi)\| \geq |T^*(\varphi(x))| \quad (32.7.15a)$$

$$= \|T(x)\| \quad (32.7.15b)$$

$$> \|T\| - \varepsilon \quad (32.7.15c)$$

By letting  $\varepsilon$  tend to zero, we see that  $\|T^*\| \geq \|T\|$ . Thus,  $\|T^*\| = \|T\|$ . □

**Theorem 32.7.5.** *If  $X$  and  $Y$  are Banach Spaces, if  $T : X \rightarrow Y$  and  $S : Y^* \rightarrow X^*$  are functions such that, for all  $\varphi \in Y^*$  and for all  $x \in X$ , we have  $S(\varphi)(x) = \varphi(T(x))$ , then  $S$  and  $T$  are bounded linear operators and  $S = T^*$ .*

*Proof.* Not if  $\varphi \in Y^*$ , then  $S(\varphi) \in X^*$ , and thus:

$$\varphi(T(x + \lambda y)) = S(\varphi)(x + \lambda y) \quad (32.7.16a)$$

$$= S(\varphi)(x) + \lambda S(\varphi)(y) \quad (32.7.16b)$$

$$= \varphi(T(x)) + \lambda \varphi(T(y)) \quad (32.7.16c)$$

$$= \varphi(T(x) + \lambda T(y)) \quad (32.7.16d)$$

Since  $\varphi \in Y^*$ , we have:

$$T(x + \lambda y) = T(x) + \lambda T(y) \quad (32.7.17)$$

To see that  $T$  is bounded, we use the closed graph theorem. Suppose  $x_n \rightarrow x$  and  $T(x_n) \rightarrow y$ . Then, for all  $\varphi \in Y^*$ , we have:

$$\varphi(y) = \lim_{n \rightarrow \infty} \varphi(T(x_n)) \quad (32.7.18a)$$

$$= \lim_{n \rightarrow \infty} S(\varphi)(x_n) \quad (32.7.18b)$$

$$= S(\varphi)(x) \quad (32.7.18c)$$

$$= \varphi(T(x)) \quad (32.7.18d)$$

Thus, by the closed graph theorem,  $y = T(x)$ . □

### 32.7.1 Brushing Up on Topology

Let  $(X, \tau)$  be a topological space. Recall that elements of  $\tau$  are called open sets.

#### Definition 32.7.10: Basis of a Topology

---

A basis of a topological space  $(X, \tau)$  is a subset  $\beta \subseteq \tau$  such that, for all  $\mathcal{U} \in \tau$  and  $x \in \mathcal{U}$ , there is a  $V \in \beta$  such that  $x \in V \subset \mathcal{U}$ . ■

We say that a subset  $V \subseteq X$  is a neighborhood of  $x$  if there is an open set  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U} \subseteq V$ . We write  $\mathcal{N}(x)$  for the collection of neighborhoods of  $x$ .

#### Definition 32.7.11: Neighborhood Basis

---

A neighborhood basis of a point  $x$  in a topological space  $(X, \tau)$  is a subset  $\alpha \subseteq \mathcal{N}(x)$  such that for all  $\mathcal{U} \in \mathcal{N}(x)$  there is a  $V \in \alpha$  such that  $x \in V \subseteq \mathcal{U}$ . ■

**Example 32.7.7: I**

a metric space, the open balls form a basis for the metric topology. In  $\mathbb{R}^n$  every point has a neighborhood basis consisting of compact sets. Thus, we say that  $\mathbb{R}^n$  is locally compact. Thus it is useful to allow neighborhoods of point be more than just open sets containing the point. ■

**Theorem 32.7.6.** *If  $\alpha(x)$  is a neighborhood basis of  $x$  consisting of open sets, and:*

$$\beta = \bigcup_{x \in X} \alpha(x) \quad (32.7.19)$$

*Then  $\beta$  is a basis for  $(X, \tau)$ .*

**Theorem 32.7.7.** *If  $(X, \tau)$  a topological space, then  $\beta \subseteq \tau$  is a basis if and only if, for all  $\mathcal{U} \in \tau$ , we have:*

$$\mathcal{U} = \bigcup_{\substack{V \in \beta \\ V \subseteq \mathcal{U}}} V \quad (32.7.20)$$

**Definition 32.7.12: Separable Topological Space**

A separable topological space is a topological space  $(X, \tau)$  such that there is a countable dense subset  $\mathcal{D} \subseteq X$ . ■

**Definition 32.7.13: First Countable Topological Space**

A first countable topological space is a topological space  $(X, \tau)$  such that, for all  $x \in X$ , there is a countable neighborhood basis  $\alpha(x) \subseteq \tau$ . ■

**Definition 32.7.14: Second Countable Topological Space**

A second countable topological space is a topological space  $(X, \tau)$  such that there is a countable basis  $\beta \subseteq \tau$ . ■

**Example 32.7.8: L**

Let  $X$  be any set, and let  $\tau = \mathcal{P}(X)$ . This is called the discrete topology on  $X$  and is metrizable, for it is generated by the discrete metric on  $X$ . The topological space  $(X, \tau)$  is not second countable if  $X$  is an uncountable set. ■

**Theorem 32.7.8.** *If  $(X, \tau)$  is a metrizable metric space, then  $(X, \tau)$  is second countable if and only if  $(X, \tau)$  is separable.*

**Theorem 32.7.9.** *If  $(X, \tau)$  is metrizable, then it is first countable.*

**Theorem 32.7.10.** *If  $X$  is a set and  $S \subseteq \mathcal{P}(X)$ , then there is a smallest topology,  $\tau(S)$ , such that  $S \subseteq \tau(S)$ .*

*Proof.* Since  $\mathcal{P}(X)$  is a topology on  $X$  such that  $S \subseteq \mathcal{P}(X)$ , the set of topologies on  $X$  such that  $S$  is contained in the topology is non-empty. Let  $A$  denote this set, and define:

$$\tau = \bigcap_{\tau_A \in A} \tau_A \quad (32.7.21)$$

Then  $\tau$  is a topology, and  $S \subseteq \tau$ . Moreover, for any topology  $\tau'$  such that  $S \subseteq \tau'$ , we have  $\tau \subseteq \tau'$ . Thus,  $\tau$  is the smallest topology.  $\square$

**Theorem 32.7.11.** *If  $\beta \subseteq \mathcal{P}(X)$  has a cover of  $X$ , then  $\beta$  is a basis for  $\tau(\beta)$  if and only if, given  $\mathcal{U}, \mathcal{V}$  in  $\tau(\beta)$ , and  $x \in \mathcal{U} \cap \mathcal{V}$ , then there is an  $\omega \in \beta$  such that  $x \in \omega \subseteq \mathcal{U} \cap \mathcal{V}$ .*

**Theorem 32.7.12.** *If  $X$  is a topological space, if  $\beta \subseteq \mathcal{P}(X)$  is a cover of  $X$  such that  $\beta$  is closed under intersection, then  $\beta$  is a basis for the topological space  $(X, \tau(\beta))$ .*

**Theorem 32.7.13.** *If  $\rho \subseteq \mathcal{P}(X)$  is a cover of  $X$ , and if:*

$$\beta = \left\{ \bigcap_{i=1}^n V_i : n \in \mathbb{N}, V_i \in \rho \right\} \quad (32.7.22)$$

*Then  $\beta$  is a basis for  $\tau(\rho)$ . We call  $\rho$  a sub-basis for  $\tau(\rho)$ .*

### Definition 32.7.15: Topology Induced by Functions

The topology induced on a set  $X$  by a set  $\mathcal{F}$  of functions  $f$  from  $X$  to a topological space  $(Z_F, \tau_F)$  is the smallest topology on  $X$  such that, for all  $f \in \mathcal{F}$ ,  $f$  is continuous.  $\blacksquare$

**Theorem 32.7.14.** *If  $X$  is a set,  $\mathcal{F}$  is a set of functions from  $X$  to a topological space  $(Z_F, \tau_F)$ , if  $\tau$  is the topology induced by  $\mathcal{F}$ , then:*

$$\beta = \left\{ f^{-1}(V) : f \in \mathcal{F}, V \in \tau_F \right\} \quad (32.7.23)$$

*Is a sub-basis for  $\tau$ .*

---

**Definition 32.7.16: Weak Topology**


---

The weak topology on a normed vector space  $X$  is the topology on  $X$  generated by  $X^*$ . ■

That is, the weak topology is the smallest topology on  $X$  such that every linear functional is continuous.

## 32.8 Stuff

Last time, we described the initial topology on a space  $X$  generated by a family of functions  $\mathcal{F} : f : X \rightarrow (Z_F, \tau_F)$ .

---

**Definition 32.8.1: Initial Topology**


---

The initial topology on a normed vector space  $(X, \|\cdot\|)$  generated by a family of functions  $\mathcal{F} = \{f : X \rightarrow (Z_F, \tau_F)\}$  such that  $f$  is continuous for all  $f \in \mathcal{F}$ . ■

**Theorem 32.8.1.** *If  $\mathcal{U}$  is defined by:*

$$\mathcal{U}(\varphi, x_0, \varepsilon) = \{x \in X : |\varphi(x) - \varphi(x_0)| < \varepsilon\} \quad (32.8.1)$$

*For  $\varphi \in X^*$ ,  $x_0 \in X$ ,  $\varepsilon > 0$ , then  $\mathcal{U}$  forms a sub-basis for the weak topology on  $X$ . Moreover, the sets:*

$$\mathcal{U}(\{\varphi_1, \dots, \varphi_n\}, x_0, \varepsilon) = \{x \in X : |\varphi_k(x) - \varphi_k(x_0)| < \varepsilon, k \in \mathbb{Z}_n\} \quad (32.8.2)$$

*Form an open neighborhood basis at  $x_0$  in the weak topology.*

*Proof.* It suffices to prove the second assertion. Since open balls in  $\mathbb{F}$  form a basis for the topology, the collection:

$$\varphi^{-1}(B_r(c)) = \{x : |\varphi(x) - c| < r\} \quad (32.8.3)$$

Form a sub-basis for the weak topology. But if  $x_0 \in \{x : |\varphi(x) - c| < r\}$ , and if  $\varepsilon = r - |\varphi(x_0) - c|$ , then  $|\varphi(x) - \varphi(x_0)| < \varepsilon$ . Thus:

$$x_0 \in \{x : |\varphi(x) - \varphi(x_0)| < \varepsilon\} \subseteq \{x : |\varphi(x) - c| < r\} \quad (32.8.4)$$

Therefore, etc. □

**Theorem 32.8.2.** *If  $X$  is a normed vector space, and  $\tau$  is the normed topology and  $\tau_w$  is the weak topology, then  $\tau_w \subseteq \tau$ .*

**Theorem 32.8.3.** *If  $X$  is a normed vector space, if  $\tau$  is the normed topology, if  $\tau_w$  is the weak topology, and if  $X$  is finite dimensional, then  $\tau = \tau_w$ .*

**Theorem 32.8.4.** *If  $X$  is a normed vector space,  $\tau$  is the normed topology,  $\tau_w$  is the weak topology, and if  $X$  is infinite dimensional, then  $\tau_w \neq \tau$ .*

In any topology space  $X$ , we say that a sequence  $x_n$  converges to  $x$  if the sequence is eventually contained in any neighborhood of  $x$ .

**Theorem 32.8.5.** *If  $x_n$  is a sequence in a normed vector space, then  $x_n \rightarrow x$  weakly if and only if  $\varphi(x_n) \rightarrow \varphi(x)$  for all  $\varphi \in X^*$ .*

*Proof.* For suppose  $x_n \rightarrow x_0$  weakly. Then, for all  $\varepsilon > 0$ , let:

$$\mathcal{U}(\varphi, x_0, \varepsilon) = \{x : |\varphi(x) - \varphi(x_0)| < \varepsilon\} \quad (32.8.5)$$

This is a weak neighborhood of  $x_0$ . Hence,  $x_n$  is eventually in  $\mathcal{U}$ . Thus,  $\varphi(x_n)$  is eventually in  $B_\varepsilon(\varphi(x_0))$ , and thus  $\varphi(x_n) \rightarrow \varphi(x_0)$ . Now suppose  $\varphi(x_n) \rightarrow \varphi(x_0)$  for all  $\varphi \in X^*$ . Let  $V$  be a weak neighborhood of  $x_0$ . Then for some  $\varphi_1, \dots, \varphi_n$ , and  $\varepsilon > 0$ :

$$x_0 \in \mathcal{U}(\{\varphi_1, \dots, \varphi_n\}, x_0, \varepsilon) = \{x : |\varphi_k(x) - \varphi_k(x_0)| < \varepsilon, k \in \mathbb{Z}_n\} \subseteq V \quad (32.8.6)$$

Thus,  $x_n$  is eventually in  $V$ .  $\square$

**Theorem 32.8.6.** *Every weakly convergent sequence is bounded with respect to the normed topology.*

*Proof.* Suppose  $x_n \rightarrow x$  weakly and let  $i : X \rightarrow X^{**}$  be the canonical map. Then, for all  $\varphi \in X^*$ :

$$i(x_n(\varphi)) = \varphi(x_n) \quad (32.8.7)$$

And this is bounded, and thus  $i(x_n)$  is uniformly bounded by the Uniform Boundedness Principle. But  $i$  is isometric, and thus  $\|x_n\|$  is bounded.  $\square$

### Example 32.8.1: L

---

t  $X = \ell^2$ . Recall that if  $x \in \ell^2$ , then:

$$\|x\|_2 = \sqrt{\sum_{n=1}^{\infty} |x_n|^2} \quad (32.8.8)$$

Let  $e_n$  be the usual basis:

$$e_n(k) = \delta_{nk} = \begin{cases} 1, & n = k \\ 0, & n \neq k \end{cases} \quad (32.8.9)$$

We have seen that  $(\ell^2)^*$  is isometric to  $\ell^2$ . That is, if  $y \in \ell^2$ , then  $y$  corresponds

to  $\varphi_y$  where:

$$\varphi_y(x) = \sum_{n=1}^{\infty} x_n y_n \quad (32.8.10)$$

And every  $\varphi \in (\ell^2)^*$  can be written this way. There is a neighborhood basis for  $0 \in \ell^2$  in the weak topology given by the sets of the form:

$$\mathcal{U} = \{x \in \ell^2 : \sum_{i=1}^n |\varphi_{y_i}(x)|^2 < \varepsilon\} \quad (32.8.11)$$

Now define  $T$  as follows:

$$T = \{\sqrt{n}e_n : n \in \mathbb{N}\} \quad (32.8.12)$$

We want to find the *weak* closure of  $T$ , which is the closure of  $T$  with respect to the weak topology. That is:

$$W = \overline{T}^w = \bigcap \{F \subseteq \ell^2 : T \subseteq F, F^C \in \tau_w\} \quad (32.8.13)$$

Where  $F^C$  denotes the complement with respect to the weak topology. Let's show that  $0$  is an element of  $W$ . Suppose not. If  $0 \notin W$ , then there is a  $\mathcal{U}$  such that  $\mathcal{U} \cap T = \emptyset$ . But:

$$|\varphi_{y_i}(\sqrt{k}e_k)| = \sqrt{k}|y_i(k)| \quad (32.8.14)$$

Thus, for all  $k \in \mathbb{N}$ :

$$\sum_{i=1}^n |y_i(k)|^2 \geq \frac{\varepsilon}{k} \quad (32.8.15)$$

But then:

$$\sum_{i=1}^n \|y_i\|_2^2 = \sum_{i=1}^n \sum_{k=1}^{\infty} |y_i(k)|^2 = \sum_{k=1}^{\infty} \sum_{i=1}^n |y_i(k)|^2 \geq \sum_{k=1}^{\infty} \frac{\varepsilon}{k} \quad (32.8.16)$$

But this sum diverges, a contradiction as the functions are in  $\ell^2$ . Thus,  $0 \in W$ .



Note that any sequence  $x_n \subseteq T$  that converges to zero must be bounded. But  $\{\sqrt{n}e_n : n \leq n\}$  is weakly closed. But no sequence in  $T$  can converge weakly to zero, and thus the weak topology is not metrizable. As it turns out, the weak topology is not even first countable. We like sequences and want to continue using this notion, but cannot use this for the weak topology on infinite dimensional normed spaces. We generalize by describing nets.

---

**Definition 32.8.2: Directed Ordered Sets**


---

A directed ordered set is an ordered set  $(\Lambda, \leq)$  such that, for all  $x, y \in \Lambda$ , there is a  $z \in \Lambda$  such that  $x \leq z$  and  $y \leq z$ . ■

**Example 32.8.2: T**


---

ere are two common examples. Any totally ordered set is automatically a directed ordered set, since we can just choose the max of any two elements. Moreover, if  $(X, \tau)$  is a topological space, then  $(\tau, \subseteq)$  is a directed ordered set, since for  $\mathcal{U}$  and  $\mathcal{V}$ , then  $\mathcal{U} \cup \mathcal{V}$  is larger than both. Similarly with reverse containment, taking  $\mathcal{U} \cap \mathcal{V}$  as the *larger* set. ■

**Definition 32.8.3: Nets**


---

net on a set  $X$  is a function from a directed set  $\Lambda$  into  $X$ ,  $x : \Lambda \rightarrow X$ . ■

As with sequences, we denote the image of  $\lambda \in \Lambda$  under a net  $x$  by writing  $x(\lambda) = x_\lambda$ . We say that a net converges in a topological space if for any neighborhood  $\mathcal{U}$  of the limit point  $x_0$ , there is a  $\lambda_0 \in \Lambda$  such that, for all  $\lambda \geq \lambda_0$ ,  $x_\lambda \in \mathcal{U}$ . An accumulation point of  $x_\lambda$ . There is also a generalized notion of accumulation point for nets.

**Example 32.8.3: E**


---

ery sequence is a net. Take  $\Lambda = \mathbb{N}$ . ■

**Theorem 32.8.7.** *If  $(X, \tau)$  is a topological space, and  $\mathcal{E} \subseteq X$ , then  $x \in \overline{\mathcal{E}}$  if and only if there is a net  $a : \Lambda \rightarrow \mathcal{E}$  such that  $a_\lambda \rightarrow x$ .*

*Proof.* For suppose  $x_\lambda \rightarrow x$  with  $x_\lambda \in \mathcal{E}$ . If  $x \notin \overline{\mathcal{E}}$ , there is a  $\mathcal{U} \subseteq \mathcal{O}(x)$  such that  $\mathcal{U} \cap \mathcal{E} = \emptyset$ . But  $x_\lambda$  is eventually in  $\mathcal{U}$ , a contradiction. Thus, etc. This direction of the proof is identical for sequences, and indeed holds for sequences. Going the other way requires the use of the notion of nets. Suppose  $x \in \overline{\mathcal{U}}$ . Let  $\Lambda$  be the set of neighborhoods of  $x$ . Then this is a directed set. Pick points to get a net. □

## 32.9 More Normed Vector Space Stuff

Given a normed vector space  $X$ , the dual  $X^*$  is a Banach space. It has a normed topology and a weak topology given by  $X^{**}$ . We can also give it the weak star topology:  $\sigma(X^*, X)$ . This is the smallest (Or initial) topology determined by

the functionals in  $X$ . That is, elements of  $X^*$  are continuous. In other words, for all  $x \in X$ , the mapping  $\varphi \mapsto \varphi(x)$  is continuous. Recall that a net  $(\varphi_\lambda)$  in  $X^*$  converges to  $\varphi$  in the weak star topology if and only if it converges point-wise. That is,  $\varphi_\lambda(x) \rightarrow \varphi(x)$  for all  $x \in X$ .

### Theorem 32.9.1: Alaoglu's Theorem

If  $X$  is a normed vector space, if  $B^*$  is the closed unit disc in the dual space:

$$B^* = \{\varphi \in X^* : \|\varphi\| \leq 1\} \quad (32.9.1)$$

Then  $B^*$  is compact in the weak star topology. ■

*Proof.* Define  $D_r$  by:

$$D_r = \{z \in \mathbb{F} : |z| \leq r\} \quad (32.9.2)$$

For all  $r > 0$ . Define:

$$Z = \prod_{x \in X} D_{\|x\|} \quad (32.9.3)$$

Then, by Tychonoff's theorem,  $Z$  is compact in the product topology. But this topology is the initial topology determined by the projection maps. And a net  $(z_\lambda) \rightarrow z$  in  $Z$  if and only if  $z_\lambda(x) \rightarrow z(x)$  for all  $x \in X$ . Define  $j : B^* \rightarrow Z$  by:

$$j(\varphi)(x) = \varphi(x) \quad (32.9.4)$$

Then  $j$  is an injective map. Moreover,  $j$  has closed range in  $Z$ . For suppose  $\varphi_\lambda$  is a net and  $j(\varphi_\lambda) \rightarrow z$ , for some  $z \in Z$ . But then, for all  $x \in X$ ,  $j(\varphi_\lambda)(x) \rightarrow z(x)$ , and thus  $\varphi_\lambda(x) \rightarrow z(x)$ . Thus  $z$  is linear,  $z(x+y) = z(x) + z(y)$  and  $z(ax) = az(x)$ . Thus,  $\varphi(x) = z(x)$  and  $|\varphi(x)| \leq \|x\|$ , and therefore  $\varphi \in B^*$ . Therefore  $j(B^*)$  is compact. Moreover,  $j$  is homeomorphism between  $B^*$  and  $j(B^*)$ , and thus  $B^*$  is compact. □

## 32.10 Hilbert Spaces

### Definition 32.10.1: Sesqui-Linear Form

Sesqui-Linear form on a vector space  $V$  over a field  $\mathbb{F}$  is a function  $\langle \cdot | \cdot \rangle$  such that:

$$\langle \alpha x + y | z \rangle = \alpha \langle x | z \rangle + \langle y | z \rangle \quad x, y, z \in V \quad \alpha \in \mathbb{F} \quad (32.10.1)$$

And such that:

$$\langle x, \alpha y + z \rangle = \bar{\alpha} \langle x | y \rangle + \langle x | z \rangle \quad x, y, z \in V \quad \alpha \in \mathbb{F} \quad (32.10.2) \quad \blacksquare$$

**Definition 32.10.2: Self-Adjoint Sesqui-Linear Form**

self-adjoint sesqui-linear form on a vector space  $V$  over a field  $\mathbb{F}$  is a sesqui-linear form  $\langle \cdot | \cdot \rangle$  such that:

$$\langle x|y \rangle = \overline{\langle y|x \rangle} \quad x, y \in V \quad (32.10.3)$$

**Definition 32.10.3: Positive Sesqui-Linear Form**

positive if  $\langle x|x \rangle \geq 0$ . ■

**Theorem 32.10.1.** If  $\mathbb{F} = \mathbb{C}$  and  $\langle \cdot | \cdot \rangle$  is a sesqui-linear form, then:

$$\langle x|y \rangle = \frac{1}{4} \sum_{n=0}^3 i^n \langle x + i^n y | x + i^n y \rangle \quad (32.10.4)$$

**Theorem 32.10.2.** If  $\mathbb{F} = \mathbb{C}$ , and  $\langle \cdot | \cdot \rangle$  is a sesqui-linear form, then it is self-adjoint if and only if it is positive.

Thus, on a complex vector space, a positive sesqui-linear form is always self-adjoint.

**Definition 32.10.4: Pre-Inner Product**

pre-inner product is a positive self-adjoint sesqui-linear form. ■

**Definition 32.10.5: Inner Product**

inner product is a pre-inner product such that  $\langle x|x \rangle = 0$  implies that  $x = 0$ .

**Definition 32.10.6: Induced Semi-Norm**

iven a pre-inner product, the induced norm is:

$$\|v\| = \sqrt{\langle x|y \rangle} \quad (32.10.5)$$



If  $\mathbb{F} = \mathbb{R}$ , then:

$$\langle x|y \rangle = \|x + y\|^2 - \|x - y\|^2 \quad (32.10.6)$$

**Theorem 32.10.3: Cauchy-Schwarz Inequality**

If  $\langle \cdot | \cdot \rangle$  is a pre-inner product and  $\|\cdot\|$  is the induced semi-inner product, then:

$$|\langle x|y \rangle| \leq \|x\| \|y\| \quad (32.10.7)$$

A similar result holds for  $\mathbb{C}$ . ■

*Proof.* For:

$$0 \leq \|\alpha x + y\|^2 = \langle \alpha x + y | \alpha x + y \rangle = |\alpha|^2 \|x\|^2 + \alpha \langle x|y \rangle + \overline{\alpha \langle x|y \rangle} + \|y\|^2 \quad (32.10.8)$$

We can simplify this further to obtain:

$$0 \leq |\alpha|^2 + 2\Re(\alpha \langle x|y \rangle) + \|y\|^2 \quad (32.10.9)$$

Let  $\tau \in \mathbb{F}$  be such that  $\tau \langle x|y \rangle = |\langle x|y \rangle|$ . Let  $\alpha = t\tau$ , with  $t \in \mathbb{R}$ . Then:

$$0 \leq t^2 \|x\|^2 + 2t |\langle x|y \rangle| + \|y\|^2 \quad (32.10.10)$$

But this is a quadratic with a positive discriminant, and thus by the quadratic formula:

$$4\langle x|y \rangle^2 - 4\|x\|^2\|y\|^2 \leq 0 \quad (32.10.11)$$

Therefore, etc. Smiley face. □

**Theorem 32.10.4.** If  $\langle \cdot | \cdot \rangle$  is a pre-inner product, then the induced semi-norm is a semi-norm.

*Proof.* Apply Cauchy-Schwarz to obtain the triangle inequality. □

**Definition 32.10.7: Induced Norm**

n induced norm is a semi-induced norm induced by an inner product. ■

**Theorem 32.10.5.** Induced norms are norms.

**Definition 32.10.8: Inner Product Space**

n inner product space Is s vector space  $V$  over a field  $\mathbb{F}$  with an inner product  $\langle \cdot | \cdot \rangle$ , denoted  $(V, \langle \cdot | \cdot \rangle)$ . ■

**Definition 32.10.9: Hilbert Space**

Hilbert space is an inner product space such that the induced norm is complete. ■

**Example 32.10.1: L**

Let  $\mathcal{H} = \mathbb{F}^n$  and define:

$$\langle x|y \rangle = \sum_{k=1}^n x_k \bar{y}_k \quad (32.10.12)$$

Then  $\langle \cdot | \cdot \rangle$  is an inner product. The induced norm is  $\ell^2$ , which is complete. Thus this is a Hilbert space. Extending this to sequences:

$$\langle x|y \rangle = \sum_{k=0}^{\infty} x_k \bar{y}_k \quad (32.10.13)$$

Similarly, we can define:

$$\langle f|g \rangle = \int_X f(x) \bar{g}(x) d\mu \quad (32.10.14)$$

For  $f, g \in L^2(X, \mathcal{M}, \mu)$ . Remembering to identify functions that differ on only a set of measure zero, this is a Hilbert space. ■

**Theorem 32.10.6.** *If  $H$  is an inner product space, and if  $x, y \in H$ , then:*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad (32.10.15)$$

This theorem characterizes inner product spaces.

**Theorem 32.10.7: Jordan von-Neumann Theorem**

Let  $X$  be a complex Banach space such that:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad (32.10.16)$$

Then there is an inner product that induces the norm. ■

**Theorem 32.10.8.** *If  $H$  is an inner product space and  $x_n \rightarrow x$  and  $y_n \rightarrow y$ , then:*

$$\langle x_n|y_n \rangle \rightarrow \langle x|y \rangle \quad (32.10.17)$$

*Proof.* Recall that the norm is continuous, so  $\|x_n\| \rightarrow \|x\|$  and  $\|y_n\| \rightarrow \|y\|$ . By Cauchy-Schwarz:

$$|\langle x_n|y_n \rangle - \langle x|y \rangle| \leq |\langle x_n - x|y_n \rangle| + |\langle x|y - y_n \rangle| \leq \|x - x_n\| \|y_n\| + \|x\| \|y - y_n\| \rightarrow 0 \quad (32.10.18)$$

Therefore, etc. □

### Definition 32.10.10: Orthogonal Elements

Orthogonal elements of an inner product space  $(H, \langle \cdot | \cdot \rangle)$  are points  $x, y \in H$  such that  $\langle x | y \rangle = 0$ . █

We can define a similar notion for subsets of  $H$ .

### Theorem 32.10.9: Pythagoras' Theorem

If  $x_1, \dots, x_n$  are pairwise orthogonal elements of an inner product space  $H$ , then:

$$\sum_{k=1}^n \|x_k\|^2 = \left\| \sum_{k=1}^n x_k \right\|^2 \quad (32.10.19) \quad \text{█}$$

*Proof.* For:

$$\left\| \sum_{k=1}^n x_k \right\|^2 = \left\langle \sum_{k=1}^n x_k \mid \sum_{k=1}^n x_k \right\rangle \quad (32.10.20)$$

$$= \sum_{k=1}^n \sum_{j=1}^n \langle x_k | x_j \rangle \quad = \sum_{k=1}^n \langle x_k | x_k \rangle \quad (32.10.21)$$

$$= \sum_{k=1}^n \|x_k\|^2 \quad (32.10.22)$$

Therefore, etc. □

**Theorem 32.10.10.** If  $C$  is a closed non-empty convex subset of a Hilbert space  $H$ , then for all  $y$  there is a unique  $x \in C$  such that  $\text{dist}(y, C) = \|x - y\|$ .

Next time, on Functional analysis: Direct sum, proof of the previous theorem, the fact that there may not be a further element. Stuff.

## 32.11 Even More Stuff

**Theorem 32.11.1.** If  $C$  is a closed non-empty convex subset of a Hilbert space  $\mathcal{H}$ , and if  $h \in \mathcal{H}$ , then there is a unique  $c \in C$  such that:

$$\text{dist}(h, C) = \|h - c\| \quad (32.11.1)$$

That is, there is a unique  $x$  in  $C$  that is closest to  $h$ .

*Proof.* Translate  $C$  by  $C - h$ , so we can assume  $h = 0$ . Define the following:

$$\alpha = \inf \{ \|x\| : x \in C \} \quad (32.11.2)$$

That is,  $\alpha = \text{dist}(0, C)$ . Let  $x_n$  be a sequence in  $C$  such that  $\|x_n\| \rightarrow \alpha$ . Then, by the parallelogram law:

$$2(\|x_n\|^2 + \|x_m\|^2) = \|x_n + x_m\|^2 + \|x_n - x_m\|^2 \quad (32.11.3)$$

But then, for all  $n, m \in \mathbb{N}$ :

$$\frac{x_n + x_m}{2} \in C \quad (32.11.4)$$

Since  $C$  is convex. Therefore:

$$2(\|x_n\|^2 + \|x_m\|^2) \geq 4\alpha^2 + \|x_n - x_m\|^2 \quad (32.11.5)$$

And thus  $x_n$  is Cauchy. Then  $x_n \rightarrow x$  and  $\|x\| = \alpha$ . Moreover, from convexity,  $x$  is unique.  $\square$

### Definition 32.11.1: Orthogonal Complement

The orthogonal complement of a subset  $S \subseteq H$  of an inner product space  $H$  is the set:

$$S^\perp = \{y \in H : \forall_{x \in S}, \langle x | y \rangle = 0\} \quad (32.11.6)$$

■

**Theorem 32.11.2.** *If  $H$  is an inner product space and  $S \subseteq H$ , then  $S^\perp$  is a closed subspace.*

From linear algebra, if  $W_1$  and  $W_2$  are subspaces of a vector space  $V$  such that  $W_1 + W_2 = V$  and  $W_1 \cap W_2 = \{0\}$ , then we say  $V = W_1 \oplus W_2$ . The map  $P_1 : V \rightarrow V$  defined by taking  $v \in V$  to the unique  $w_1 \in W_1$  such that  $v = w_1 + w_2$  is called the projection of  $H$  onto  $W_1$  along  $W_2$ .

**Theorem 32.11.3.** *If  $W$  is a closed subspace of a Hilbert space  $\mathcal{H}$ , then  $\mathcal{H} = W \oplus W^\perp$ . If  $P_W : \mathcal{H} \rightarrow \mathcal{H}$  is the projection mapping of  $H$  into  $W$ , then  $P_W(h)$  is the closest element in  $W$  to  $h$ .*

*Proof.* For let  $h \in \mathcal{H}$  and let  $x$  be the closest element in  $W$  to  $h$ . Let  $x^\perp = h - x$ . Let  $w \in W$  and  $\varepsilon > 0$ . But then:

$$\|x^\perp\|^2 = \|h - x\|^2 \leq \|h - (x + \varepsilon w)\|^2 \quad (32.11.7a)$$

$$= \|h - x - \varepsilon w\|^2 \quad (32.11.7b)$$

$$= \|x^\perp - \varepsilon w\|^2 \quad (32.11.7c)$$

$$= \|x^\perp\|^2 - 2\varepsilon \Re(\langle x^\perp | w \rangle) + \varepsilon^2 \|w\|^2 \quad (32.11.7d)$$

Therefore:

$$2\varepsilon \Re(\langle x^\perp | w \rangle) = \varepsilon^2 \|w\|^2 \quad (32.11.8)$$

Thus,  $x^\perp \in W^\perp$ . But  $\mathcal{H} = W + W^\perp$  and  $W \cap W^\perp = \{0\}$ . Therefore,  $W \oplus W^\perp = \mathcal{H}$ .  $\square$

**Theorem 32.11.4.** *If  $\mathcal{H}$  is a Hilbert space and if  $S \subseteq \mathcal{H}$ , then:*

$$(S^\perp)^\perp = \text{Cl}_{\|\cdot\|}(\text{Span}(S)) \quad (32.11.9)$$

Note that given a point  $h$  in an inner product space,  $\varphi_h(x) = \langle x|v \rangle$  defines a linear function. By the Cauchy-Schwarz theorem:

$$|\varphi_h(h)| \leq \|h\| \|v\| \quad (32.11.10)$$

And thus  $\varphi_v \in H^*$  and  $\|\varphi_v\| \leq \|v\|$ . But:

$$|\varphi_v(v)| = \|v\|^2 \quad (32.11.11)$$

And thus  $\|\varphi_v\| = \|v\|$ .

### Theorem 32.11.5: Riesz's Representation Theorem

If  $\mathcal{H}$  is a Hilbert space, then the map  $\Phi : \mathcal{H} \rightarrow \mathcal{H}^*$  given by  $\Phi(v) = \varphi_v$  is a conjugate linear isometric map of  $\mathcal{H}$  onto  $\mathcal{H}^*$ . █

*Proof.* By the previous remark, it suffices to show that  $\Phi$  is surjective. Let  $\varphi \in \mathcal{H}^*$ . Let  $W = \ker(\varphi)$ . But  $\varphi$  is continuous, and thus  $W$  is a closed subspace of  $\mathcal{H}$ . If  $\varphi$  is the zero function, let  $v = 0$ . If not, then there exists a non-zero element  $v \in (\ker(\varphi))^\perp$ . Let  $y = v/\|v\|$ . □

**Theorem 32.11.6.** *If  $\mathcal{H}$  is a Hilbert space, then the bijection  $\Phi : \mathcal{H} \rightarrow \mathcal{H}^*$  mapping  $v \mapsto \varphi_v$  is a homeomorphism of  $\mathcal{H}$  with the weak topology onto  $\mathcal{H}^*$  with the weak star topology.*

*Proof.* Recall every  $\varphi \in \mathcal{H}^*$  is of the form  $\varphi_v$  for some  $v \in \mathcal{H}$ . Thus  $h_\lambda \mapsto h$  in the weak topology if and only if for all  $v \in \mathcal{H}$ :

$$\langle h_\lambda | h \rangle \rightarrow \langle h | v \rangle \quad (32.11.12a)$$

$$\implies \langle v | h_\lambda \rangle \rightarrow \langle v | h \rangle \quad (32.11.12b)$$

$$\implies \varphi_{h_\lambda}(v) \rightarrow \varphi_h(v) \quad (32.11.12c)$$

But then  $\varphi_{h_\lambda} \rightarrow \varphi_h$  in the weak star topology. □

**Theorem 32.11.7.** *If  $\mathcal{H}$  is a Hilbert space and if  $B$  is the closed unit ball, then  $B$  is weakly compact.*

### Definition 32.11.2: Orthonormal Basis

An orthonormal basis of an inner product space  $H$  is a subset  $E \subseteq H$  such that, for all  $e \in E$ ,  $\|e\| = 1$ , and for all distinct  $e_\alpha, e_\beta \in E$ ,  $\langle e_\alpha | e_\beta \rangle = 0$ . █

**Theorem 32.11.8: Bessel's Inequality**

If  $H$  is an inner product, and if  $e_n$  is an orthonormal sequence in  $H$ , then for all  $x \in H$ :

$$\sum_{n=1}^{\infty} |\langle x | e_n \rangle|^2 \leq \|x\|^2 \quad (32.11.13)$$



*Proof.* For define the following:

$$x_n = x - \sum_{k=1}^n \langle x | e_k \rangle e_k \quad (32.11.14)$$

Note that  $x_n \perp e_k$  for  $k = 1, \dots, n$ . But, by the Pythagorean theorem:

$$\|x\|^2 = \|x_n\|^2 + \sum_{k=1}^n |\langle x | e_k \rangle|^2 \quad (32.11.15)$$

Thus:

$$\|x\|^2 \geq \sup_{n \in \mathbb{N}} \sum_{k=1}^n |\langle x | e_k \rangle|^2 = \sum_{k=1}^{\infty} |\langle x | e_k \rangle|^2 \quad (32.11.16)$$

Therefore, etc. □

**Definition 32.11.3: Orthogonal Projection**

The orthogonal projection of a closed subspace  $W$  of a Hilbert space  $\mathcal{H}$  is the projection  $P_W : \mathcal{H} \rightarrow \mathcal{H}$  of  $W$  along  $W^\perp$ . ■

**Theorem 32.11.9.** If  $E$  is an orthonormal set in a Hilbert space  $\mathcal{H}$ , and if  $\mathcal{E} = \text{Cl}(\text{Span}(E))$ , then for all  $h \in H$ , the sum over  $\langle h | e_n \rangle e_n$  converges and:

$$P_{\mathcal{E}}(h) = \sum_{n=1}^{\infty} \langle h | e_n \rangle e_n \quad (32.11.17)$$

## 32.12 Even MORE Stuff!

**Theorem 32.12.1.** If  $S$  and  $T$  are self adjoint bounded operators on  $\mathcal{H}$ , and if  $S \leq T$ , then  $ASA^* \leq ATA^*$  for all  $A \in \mathcal{L}(\mathcal{H})$ .

**Theorem 32.12.2.** If  $S, T$  are self-adjoint operators on  $\mathcal{H}$ , if  $0 \leq S \leq T$ , then  $\|S\| \leq \|T\|$ .

*Proof.* For suppose  $0 \leq S \leq T$ , and let  $[x, y] = \langle Sx|y \rangle$ . Then  $[\cdot, \cdot]$  is a pre-inner product on  $\mathcal{H}$ , and thus if  $\|x\| = \|y\| = 1$ , then by Cauchy-Schwarz:

$$|\langle Sx|y \rangle|^2 = |[x, y]|^2 \leq [x, x][y, y] = \langle Sx|x \rangle \langle Sy|y \rangle \quad (32.12.1)$$

But  $S \leq T$ , and therefore:

$$\langle Sx|x \rangle \langle Sy|y \rangle \leq \langle Tx|x \rangle \langle Ty|y \rangle \leq \|T\|^2 \quad (32.12.2)$$

Therefore  $\|S\|^2 \leq \|T\|^2$ . Taking square roots completes the proof.  $\square$

**Theorem 32.12.3.** *If  $S$  is a self-adjoint bounded operators on  $\mathcal{H}$ , and if  $S \geq 0$ , then  $S \leq I$  if and only if  $\|S\| \leq 1$ .*

*Proof.* For if  $0 \leq S \leq U$ , then  $\|S\| \leq \|I\| = 1$ . But if  $0 \leq S$  and  $\|S\| \leq 1$ , then:

$$\langle Sx|x \rangle \leq \|x\|^2 = \langle x|x \rangle \quad (32.12.3)$$

Thus,  $S \leq I$ .  $\square$

**Theorem 32.12.4.** *If  $T$  is a self-adjoint bounded operator on  $\mathcal{H}$ , if  $-I \leq T \leq I$ , then  $\|T\| \leq 1$ .*

*Proof.* Suppose  $T = T^*$  and  $-I \leq T \leq I$ . Then:

$$\langle T(x+y)|x+y \rangle \leq \|x+y\|^2 \quad (32.12.4)$$

And similarly:

$$-\langle T(x-y)|x-y \rangle \leq \|x-y\|^2 \quad (32.12.5)$$

Summing, we obtain:

$$4\Re(\langle Tx|y \rangle) \leq \|x+y\|^2 + \|x-y\|^2 \quad (32.12.6)$$

By the parallelogram law:

$$4\Re(\langle Tx|y \rangle) \leq 2\|x\|^2 + 2\|y\|^2 \quad (32.12.7)$$

Therefore:

$$4|\langle Tx|y \rangle| \leq 2\|x\|^2 + 2\|y\|^2 \quad (32.12.8)$$

But, for  $\|x\| = \|y\| = 1$ ,  $\sup |\langle Tx|y \rangle| \leq 1$ , and therefore  $\|T\| \leq 1$ . On the other hand, if  $T = T^*$ , then:

$$|\langle Tx|x \rangle| \leq \|x\|^2 \quad (32.12.9)$$

But  $T = T^*$ , and thus  $\langle Tx|x \rangle$  is a real number. Thus:

$$-\langle x|x \rangle \leq \langle Tx|x \rangle \leq \langle x|x \rangle \quad (32.12.10)$$

Thus,  $-I \leq T \leq I$ .  $\square$

Let  $A \in M_n(\mathbf{F})^\dagger$ . That is,  $A = A^*$  and  $A = \mathcal{U}\mathcal{D}\mathcal{U}^*$  for some unitary  $\mathcal{U}$  and a diagonal  $\mathcal{D}$ :

$$\mathcal{D} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad (32.12.11)$$

Where  $\lambda_k \geq 0$ .

**Theorem 32.12.5.** *There is a sequence of polynomials with positive coefficients such that:*

$$\sum_{n=1}^{\infty} p_n(t) = 1 - \sqrt{1-t} \quad (32.12.12)$$

*Uniformly on  $[0, 1]$ .*

*Proof.* Let  $q_0 = 0$  and define:

$$q_{n+1}(t) = \frac{1}{2}(t + q_n(t))^2 \quad (32.12.13)$$

For all  $n \in \mathbb{N}$ . By induction we see that  $q_n$  is a sequence of polynomials with positive coefficients and such that:

$$0 \leq q_n(t) \leq 1 \quad (32.12.14)$$

For all  $t \in [0, 1]$ , and  $n \in \mathbb{N}$ . Define:

$$p_n(t) = q_n(t) - q_{n-1}(t) \quad (32.12.15)$$

For all  $n \in \mathbb{N}$ . But then:

$$2p_{n+1}(t) = q_n(t)^2 - q_{n-1}(t)^2 = (q_n(t) - q_{n-1}(t))(q_n(t) + q_{n-1}(t)) = p_n(t)(q_n(t) + q_{n-1}(t)) \quad (32.12.16)$$

Thus, each  $p_n$  has positive coefficients, and therefore:

$$q_n(t) \leq q_{n-1}(t) \quad (32.12.17)$$

But  $q_n$  is bounded by 1 and monotonic, and thus by completeness, there is a limit function. Let:

$$Q(t) = \lim_{n \rightarrow \infty} q_n(t) \quad (32.12.18)$$

But then:

$$q(t) = \frac{1}{2}(t + q(t)^2) \quad (32.12.19)$$

And therefore:

$$q(t) = 1 - \sqrt{1-t} \quad (32.12.20)$$

Moreover, by Dini's theorem, the convergence is uniform. And the  $p_n$  form a telescoping series, and therefore:

$$\sum_{n=1}^{\infty} p_n(t) = 1 - \sqrt{1-t} \quad (32.12.21)$$

Therefore, etc.  $\square$

### Example 32.12.1: D

---

Find the following  $2 \times 2$  matrices:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix} \quad (32.12.22)$$

Then  $A$  and  $B$  are positive, but:

$$AB = \begin{pmatrix} 7 & -1 \\ 2 & 1 \end{pmatrix} \quad (32.12.23)$$

And this is not symmetric, and thus not positive. The product of positive operators need not be positive.  $\blacksquare$

**Theorem 32.12.6.** *If  $S \geq 0$ , and for all  $n \in \mathbb{N}$ ,  $S^n \geq 0$ . In particular, if  $p$  is a polynomial with positive coefficients, then  $p(S) \geq 0$ .*

*Proof.* We have  $(S^n)^* = S^n$ . Thus, yeah.  $\square$

**Theorem 32.12.7.** *If  $T$  is a bounded operator,  $T \geq 0$ , then there is a unique  $A \in \mathcal{L}(\mathcal{H})$  such that  $A \geq 0$  and  $A^2 = T$ . If  $B$  commutes with  $T$ , then  $B$  commutes with  $A$ .*

*Proof.* Let  $\alpha > 0$ . If  $A^2 = \alpha T$ , then  $(\alpha^{-1/2} A)^2 = T$ . Thus we can replace  $T$  by  $\alpha T$  such that  $\|T\| \leq 1$ . Thus,  $0 \leq T \leq I$ . Then, if  $S = I - T$ , then  $0 \leq S \leq I$ . Let  $p_n$  and  $q_n$  be defined as before. Let:

$$S_n = p_n(S) \quad (32.12.24)$$

Then  $S_n \geq 0$ . Thus:

$$0 \leq \sum_{k=m}^n p_k(t) = \sum \alpha_r t^r \leq \varepsilon \quad (32.12.25)$$

For all  $t \in [0, 1]$ . Note that  $\alpha_k \geq 0$ . Hence, we have:

$$\left\| \sum_{k=m}^n S_k \right\| \leq \sum \alpha_r \|S\|^r \leq \sum \alpha r < \varepsilon \quad (32.12.26)$$

Thus,  $S_k$  forms a Cauchy sequence and therefore converges. Moreover:

$$R = \sum_{k=1}^{\infty} S_k \geq 0 \quad (32.12.27)$$

Moreover,  $0 \leq R \leq I$ . But:

$$(I - R)^2 = \left( I - \sum_{k=1}^{\infty} S_k \right)^2 \quad (32.12.28)$$

$$= \lim_{n \rightarrow \infty} (I - q_n(S))^2 \quad (32.12.29)$$

$$= \lim_{n \rightarrow \infty} (I - 2q_n(S) + q_n(S)^2) \quad (32.12.30)$$

$$= I - S \quad (32.12.31)$$

And this is equal to  $T$ . Thus,  $(I - R)^2 = T$ . Now, if  $BT = TB$ , then:

$$AB = (I - R)B = \lim_{n \rightarrow \infty} (I - q_n(I - T))B = B \lim_{n \rightarrow \infty} (I - q_n(I - T)) = BA \quad (32.12.32)$$

Lastly,  $A$  is unique. For if  $B \geq 0$  and  $B^2 = T$ , then  $B$  commutes with  $T$  and hence  $B$  commutes with  $A$ . Therefore:

$$(A - B)^2 x = (A - B)(A + B)x = (T - T)x = 0 \quad (32.12.33)$$

Hence, if  $y$  is in the range of  $A + B$ , then  $(A - B)y = 0$ . Let  $\mathcal{E}$  be the range of  $A + B$ . If we can show that  $(A - B)y = 0$  for all  $y \in \mathcal{E}^\perp$ , then we are done. But since  $A$  and  $B$  are self-adjoint,  $\mathcal{E}^\perp$  is the kernel of  $A + B$ . Thus:

$$\langle Ay|y \rangle \leq \langle (A + B)z|z \rangle = 0 \quad (32.12.34)$$

Thus, for all  $y \in \mathcal{E}^\perp$ ,  $\langle Ay|y \rangle = 0$ . But  $A \geq 0$ , hence there is a  $C$  such that  $A = C^2$  and  $C \geq 0$ . Then:

$$\langle Az|Z \rangle = \|Cz\|^2 = 0 \quad (32.12.35)$$

Thus  $Cz = 0$  and hence  $Az = C^2z = 0$ . Similarly,  $Bz = 0$ . Thus  $(A - B)z = 0$ .  $\square$

**Theorem 32.12.8.** *If  $T \geq 0$  and  $S \geq 0$ , and if  $TS = ST$ , then  $TS \geq 0$ .*

*Proof.* For:

$$TS = T(\sqrt{S})^2 = \sqrt{S}T\sqrt{S} \geq 0 \quad (32.12.36)$$

Therefore, etc.  $\square$

Here,  $A$  is called the positive square root of the operator  $T$ .



# CHAPTER 33

---

## Homeworks

---

### 33.1 Homework I

**Problem 33.1.1** Show that, for  $1 \leq p \leq q \leq \infty$ , that  $\|\cdot\|_p$  and  $\|\cdot\|_q$  are strongly equivalent.

*Solution* (1) First, if  $X$  is a set,  $d_1, d_2, d_3$  are metrics on  $X$ , if  $d_1$  is strongly equivalent to  $d_2$ , and  $d_2$  is strongly equivalent to  $d_3$ , then  $d_1$  is strongly equivalent to  $d_3$ . For if  $d_1$  is strongly equivalent to  $d_2$ , then there are  $\alpha, \beta > 0$  such that, for all  $x, y \in X$ :

$$\alpha d_1(x, y) \leq d_2(x, y) \leq \beta d_1(x, y) \quad (33.1.1)$$

But we have seen that strongly equivalent is a symmetric relation, and therefore if  $d_2$  and  $d_3$  are strongly equivalent then there are  $a, b > 0$  such that:

$$ad_3(x, y) \leq d_2(x, y) \leq bd_3(x, y) \quad (33.1.2)$$

Therefore:

$$\frac{\alpha}{b} d_1(x, y) \leq d_3(x, y) \leq \frac{\beta}{a} d_1(x, y) \quad (33.1.3)$$

Thus, strongly equivalent is a transitive relation. Let  $n \in \mathbb{N}$ ,  $p \in [1, \infty)$ , and let  $\mathbf{x} \in \mathbb{R}^n$ . Then, since  $\|\mathbf{x}\|_\infty$  is the supremum norm, it is greater than or equal to all of the components of  $\mathbf{x}$ . Thus:

$$n \|\mathbf{x}\|_\infty^p = \sum_{k=1}^n \|\mathbf{x}\|_\infty^p \geq \sum_{k=1}^n |x_k|^p = \|\mathbf{x}\|_p^p \quad (33.1.4)$$

Taking  $p^{th}$  roots, we have:

$$n^{\frac{1}{p}} \|\mathbf{x}\|_\infty \geq \|\mathbf{x}\|_p \quad (33.1.5)$$

Going the other way:

$$\|\mathbf{x}\|_\infty^p \leq \sum_{k=1}^n |x_k|^p \quad (33.1.6)$$

Taking  $p^{th}$  roots again, we obtain the following:

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_p \leq n^{\frac{1}{p}} \|\mathbf{x}\|_\infty \quad (33.1.7)$$

Thus, for all  $p \in [1, \infty)$ ,  $\|\cdot\|_p$  is strongly equivalent to  $\|\cdot\|_\infty$ . But strongly equivalent is a transitive relation, thus for all  $p, q \in [1, \infty]$ ,  $\|\cdot\|_p$  is strongly equivalent to  $\|\cdot\|_q$ .

*Solution (2)* Let  $p \in [1, \infty]$ , and let  $f : S^n \rightarrow \mathbb{R}$  be defined by:

$$f(\mathbf{x}) = \|\mathbf{x}\|_p \quad (33.1.8)$$

But  $S^n$  is a closed and bounded subset of  $\mathbb{R}^n$ , and is therefore, by the Heine-Borel theorem, it is compact. But continuous functions attain their maximum and minimum on compact sets, by the extreme value theorem. Thus, there are  $\mathbf{x}_{\min}$  and  $\mathbf{x}_{\max}$  such that, for all  $\mathbf{x} \in S^n$ :

$$\|\mathbf{x}_{\min}\|_p \leq \|\mathbf{x}\|_p \leq \|\mathbf{x}_{\max}\|_p \quad (33.1.9)$$

But also  $\|\mathbf{x}_{\min}\|_p > 0$ , for  $\mathbf{x}_{\min} \in S^n$ , and therefore  $\mathbf{x}_{\min} \neq \mathbf{0}$ . Moreover:

$$\|e_i\|_p = 1 \quad (33.1.10)$$

For all  $i \in \mathbb{Z}_n$ . But also, for all  $\mathbf{x} \in S^n$ , we have:

$$\|\mathbf{x}\|_2 = 1 \quad (33.1.11)$$

Therefore, for all  $\mathbf{x} \in S^n$ :

$$\frac{\|\mathbf{x}_{\min}\|_p}{\|\mathbf{x}_{\max}\|_p} \|\mathbf{x}\|_p \leq \|\mathbf{x}\|_2 \leq \frac{\|\mathbf{x}_{\max}\|_p}{\|\mathbf{x}_{\min}\|_p} \|\mathbf{x}\|_p \quad (33.1.12)$$

For the general  $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ , we can scale back to the unit  $n$  sphere. Thus,  $\|\cdot\|_p$  and  $\|\cdot\|_2$  are strongly equivalent for all  $p \in [1, \infty]$ . Since strongly equivalent is a transitive relation,  $\|\cdot\|_p$  and  $\|\cdot\|_q$  are strongly equivalent for all  $p, q \in [1, \infty]$ .

**Problem 33.1.2** Give an example of a metric on  $\mathbb{R}^n$  that is not strongly equivalent to  $\|\cdot\|_p$  for any  $p \in [1, \infty]$ .

*Solution* The discrete metric is not strongly equivalent to any of the metrics induced by  $\|\cdot\|_p$ . For suppose not. Then:

$$\alpha \|\mathbf{x}\|_p \leq d(\mathbf{x}, \mathbf{0}) \leq 1 \quad (33.1.13)$$

Where  $d$  is the discrete metric. But  $\|\mathbf{x}\|_p$  is unbounded, and thus if  $\alpha \|\mathbf{x}\|_p \leq 1$ , then  $\alpha = 0$ . Thus there is no  $\alpha > 0$  that satisfies the inequality.

**Problem 33.1.3** Show that, if  $a, b \geq 0$  and  $0 < \lambda < 1$ , then:

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b \quad (33.1.14)$$

*Solution* If  $a = 0$  or  $b = 0$ , then we are done. Suppose not. Define  $t = ab^{-1}$ . Then, if  $\lambda \in (0, 1)$ , and  $t \geq 1$ , we have:

$$\lambda(t^{\lambda-1} - 1) \geq 0 \quad (33.1.15a)$$

$$\Rightarrow \int_1^t \lambda(\tau^{\lambda-1} - 1) d\tau \geq 0 \quad (33.1.15b)$$

$$\Rightarrow (t^\lambda - \lambda t) - (1 - \lambda) \geq 0 \quad (33.1.15c)$$

For  $t \in (0, 1)$ , we have:

$$\lambda(t^{\lambda-1} - 1) \leq 0 \quad (33.1.16a)$$

$$\Rightarrow \int_t^1 \lambda(\tau^{\lambda-1} - 1) d\tau \leq 0 \quad (33.1.16b)$$

$$\Rightarrow (1 - \lambda) - (t^\lambda - \lambda t) \leq 0 \quad (33.1.16c)$$

Combining these two, we have for  $t \in (0, \infty)$ :

$$t^\lambda - \lambda t \leq 1 - \lambda \quad (33.1.17)$$

But  $t = ab^{-1}$ , and thus multiplying through by  $b$ :

$$a^\lambda b^{\lambda-1} \leq \lambda a + (1-\lambda)b \quad (33.1.18)$$

**Problem 33.1.4** Prove Hölder's Inequality: If  $p \in (1, \infty)$ ,  $p^{-1} + q^{-1} = 1$ , then:

$$\|fg\|_1 \leq \|f\|_p \|g\|_q \quad (33.1.19)$$

*Solution* For let:

$$\tilde{f} = \frac{f}{\|f\|_q} \quad (33.1.20a) \qquad \tilde{g} = \frac{g}{\|g\|_q} \quad (33.1.20b)$$

Then  $\|\tilde{f}\|_p = 1$  and  $\|\tilde{g}\|_q = 1$ . But if  $p \in (1, \infty)$ , then  $p^{-1} < 1$ , and thus by the previous problem, and since  $1 - p^{-1} = q^{-1}$ , we have:

$$|\tilde{f}(x)\tilde{g}(x)| \leq p^{-1}|\tilde{f}(x)|^p + q^{-1}|\tilde{g}(x)|^q \quad (33.1.21)$$

Integrating, we get:

$$\|\tilde{f}\tilde{g}\|_1 = \int_{\mathbb{R}} |\tilde{f}(x)\tilde{g}(x)| dx \quad (33.1.22a)$$

$$\leq \int_{\mathbb{R}} (p^{-1}|\tilde{f}(x)|^p + q^{-1}|\tilde{g}(x)|^q) dx \quad (33.1.22b)$$

$$= p^{-1}\|\tilde{f}\|_p^p + q^{-1}\|\tilde{g}\|_q^q \quad (33.1.22c)$$

$$= p^{-1} + q^{-1} \quad (33.1.22d)$$

But  $p^{-1} + q^{-1} = 1$ , and thus  $\|\tilde{f}\tilde{g}\|_1 = 1$ . But from the definition of  $\tilde{f}$  and  $\tilde{g}$ , we can multiply through by  $\|f\|_p\|g\|_q$  to obtain:

$$\|fg\|_1 \leq \|f\|_p\|g\|_q \quad (33.1.23)$$

**Problem 33.1.5** Prove Minkowski's Inequality

**Problem 33.1.6** A limit point of a subspace  $(E, d_E)$  of a metric space  $(X, d)$  is a point  $x \in X$  such that there exists a sequence  $a : \mathbb{N} \rightarrow E$  such that  $a_n \rightarrow x$ . Prove that  $E$  is closed if and only if it has all of its limit points.

*Solution* Suppose  $E$  is closed and let  $x$  be a limit point of  $E$ . Suppose  $x \in E^C$ . But if  $E$  is closed, then  $E^C$  is open, and thus there is an  $r > 0$  such that:

$$B_r^{(X,d)}(x) \subseteq E^C \quad (33.1.24)$$

But if  $x$  is a limit point of  $E$  then there is a sequence  $a : \mathbb{N} \rightarrow E$  such that  $a_n \rightarrow x$ . But if  $a_n \rightarrow x$  then there is an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  such that  $n > N$ , it is true that  $d(x, a_n) < r/2$ . But then, for all  $n > N$ , we have that:

$$a_n \in B_r^{(X,d)}(x) \quad (33.1.25)$$

And thus  $a_n \in E^C$ . But  $a_n \in E$ , a contradiction. Thus,  $x \in E$ . Now, suppose  $x$  has all of its limit points and suppose  $E$  is not closed. Then  $E^C$  is not open. But then there is an  $x \in E^C$  such that, for all  $\varepsilon > 0$ :

$$B_\varepsilon^{(X,d)}(x) \cap E \neq \emptyset \quad (33.1.26)$$

Define the following:

$$A_n = \left\{ y \in E : d(x, y) < \frac{1}{n} \right\} \quad (33.1.27)$$

Then for all  $n \in \mathbb{N}$ ,  $A_n$  is non-empty. By choice there is a sequence  $a : \mathbb{N} \rightarrow E$  such that, for all  $n \in \mathbb{N}$ ,  $a_n \in A_n$ . But then, for all  $n \in \mathbb{N}$ ,  $d(a_n, x) < n^{-1}$ . Thus  $a_n \rightarrow x$  and therefore  $x$  is a limit point of  $E$ . But  $x \in E^C$ , a contradiction as  $E$  contains all of its limit points. Therefore,  $E$  is closed.

**Problem 33.1.7** State and prove a result characterizing open sets in a metric space in terms of sequences, similar to the previous problem.

*Solution* A subset  $\mathcal{U} \subseteq X$  of a metric space  $(X, d)$  is open if and only if for any convergent sequence  $a : \mathbb{N} \rightarrow X$  such that the limit of  $a$  is in  $\mathcal{U}$ , there is an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  and  $n > N$ , it is true that  $a_n \in \mathcal{U}$ . For suppose  $\mathcal{U}$  is open, and let  $a : \mathbb{N} \rightarrow X$  be a convergent sequence such that there is an  $x \in \mathcal{U}$  such that  $a_n \rightarrow x$ . But if  $\mathcal{U}$  is open then there is an  $\varepsilon > 0$  such that:

$$B_\varepsilon^{(X,d)}(x) \subseteq \mathcal{U} \quad (33.1.28)$$

But if  $a_n \rightarrow x$  then there is an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  such that  $n > N$ , it is true that  $d(x, a_n) < \varepsilon$ . But then for all  $n > N$ ,  $n \in \mathbb{N}$ , we have  $a_n \in B_\varepsilon^{(X,d)}(x)$ , and thus  $a_n \in \mathcal{U}$ . Now suppose for any sequence that converges to a point in  $\mathcal{U}$ , the sequence is eventually contained within  $\mathcal{U}$ . Suppose  $\mathcal{U}$  is not open. Then there is an  $x \in \mathcal{U}$  such that, for all  $\varepsilon > 0$  there is a  $y \in \mathcal{U}^C$  such that  $d(x, y) < \varepsilon$ . Define the following:

$$A_n = \left\{ y \in \mathcal{U}^C : d(x, y) < \frac{1}{n} \right\} \quad (33.1.29)$$

Then for all  $n \in \mathbb{N}$ ,  $A_n$  is non-empty. By choice there is a sequence  $a : \mathbb{N} \rightarrow \mathcal{U}^C$  such that  $a_n \in A_n$ . But then  $a_n \rightarrow x$ . But if  $a_n \rightarrow x$ , then there is an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  such that  $n > N$  it is true that  $a_n \in \mathcal{U}$ , a contradiction. Therefore,  $\mathcal{U}$  is open.

**Problem 33.1.8** Let  $\rho$  and  $\sigma$  be metrics on  $X$  and show that  $\rho$  and  $\sigma$  are equivalent if and only if they have the same convergent sequences.

*Solution* For a sequence  $a : \mathbb{N} \rightarrow X$  converges to  $x \in X$  if and only if for all open subsets  $\mathcal{U} \subseteq X$  such that  $x \in \mathcal{U}$ , there is an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  and  $n > N$ , it is true that  $a_n \in \mathcal{U}$ . Going one way, if  $a_n \rightarrow x$  then for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  and  $n > N$ , it is true that  $d(x, a_n) < \varepsilon$ . Let  $\mathcal{U}$  be an open subset such that  $x \in \mathcal{U}$ . But then there is an  $\varepsilon > 0$  such that:

$$B_\varepsilon^{(X,d)}(x) \subseteq \mathcal{U} \quad (33.1.30)$$

But there is an  $N \in \mathbb{N}$  such that, for all  $n > N$ ,  $n \in \mathbb{N}$ , we have  $a_n \in B_\varepsilon^{(X,d)}(x)$ . Therefore  $a_n \in \mathcal{U}$ . Going the other way, let  $a : \mathbb{N} \rightarrow X$  be a sequence such that, for every open set  $\mathcal{U} \subseteq X$  such that  $x \in \mathcal{U}$ , there is an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  and  $n > N$ , it is true that  $a_n \in \mathcal{U}$ . Let:

$$A_k = B_{k^{-1}}^{(X,d)}(x) \quad (33.1.31)$$

Given  $\varepsilon > 0$  there is a  $k \in \mathbb{N}$  such that  $k^{-1} < \varepsilon$ . But  $A_k$  is open and  $x \in A_k$ , and thus there is an  $N \in \mathbb{N}$  such that for all  $n > N$  and  $n \in \mathbb{N}$ , we have that  $a_n \in A_k$ . But then  $d(x, a_n) < \varepsilon$ , so  $a_n \rightarrow x$ . This converts the notion of convergence from a metric space property to a topological property. If  $(X, \rho)$  and  $(X, \sigma)$  are equivalent then they have the same open sets, and thus convergence of sequences is preserved. For suppose  $a : \mathbb{N} \rightarrow X$  converges to  $x$  with respect to  $\rho$ . Then, for all open subsets  $\mathcal{U}$  of  $(X, d)$  such that  $x \in \mathcal{U}$ , there is an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  and  $n > N$ , it is true that  $a_n \in \mathcal{U}$ . But if  $\mathcal{U}$  is open in  $(X, \rho)$ , then it is open in  $(X, \sigma)$ , for the two metrics are equivalent. Thus  $a_n \rightarrow x$  with respect to  $\sigma$ .

**Problem 33.1.9** Let  $(X, \mathcal{M}, \mu)$  be a measure space and define  $\mathcal{U} \subseteq L^1(X)$  by:

$$\mathcal{U} = \left\{ f \in L^1(X) : \int_X \Re(f) \, d\mu < 1 \right\} \quad (33.1.32)$$

Prove that  $\mathcal{U}$  is open with respect to the metric induced by  $\|\cdot\|_1$ .

*Solution* For let  $f \in L^1(X)$  and let  $M = \|f\|_1$ . As  $f \in L^1(X)$ ,  $M < \infty$ . Define:

$$\alpha = \int_X \Re(f) d\mu \quad (33.1.33)$$

And let  $\varepsilon = \min\{1/2M, 1 - \alpha\}$ . Then if  $\|f - g\|_1 < \varepsilon$  we have:

$$\int_X \Re(g) d\mu = \int_X \Re(g - f + f) d\mu = \int_X \Re(g - f) d\mu + \int_X \Re(f) d\mu \quad (33.1.34)$$

We can simplify this further to get:

$$\int_X \Re(g) d\mu = \int_X \Re(g - f) d\mu + \alpha < \varepsilon + \alpha \leq 1 \quad (33.1.35)$$

Therefore:

$$B_\varepsilon^{(L^1(X), \|\cdot\|_1)}(f) \subseteq \mathcal{U} \quad (33.1.36)$$

And thus  $\mathcal{U}$  is open.

**Problem 33.1.10** For a metric space  $(X, d)$ , define  $\text{dist} : X \times \mathcal{P}(X) \setminus \{\emptyset\} \rightarrow [0, \infty)$  by:

$$\text{dist}(x, A) = \inf\{d(x, y) : y \in A\} \quad (33.1.37)$$

Prove that  $\text{dist}(x, A) = 0$  if and only if  $x \in \overline{A}$ . Show that, for a fixed non-empty  $A \subseteq X$ ,  $\text{dist}(x, A)$  is continuous. Prove that if  $A, B \subseteq X$  are closed disjoint non-empty subsets, then there is a continuous function  $f : X \rightarrow [0, 1]$  such that  $f(x) = 0$  if and only if  $x \in A$  and  $f(y) = 1$  if and only if  $x \in B$ .

*Solution* If  $x \in \overline{A}$ , then for all  $\varepsilon > 0$ :

$$B_\varepsilon^{(X, d)}(x) \cap A \neq \emptyset \quad (33.1.38)$$

Thus  $\text{dist}(x, A) < \varepsilon$  for all positive  $\varepsilon$ , and therefore  $\text{dist}(x, A) = 0$ . If  $\text{dist}(x, A) = 0$  then for all  $\varepsilon > 0$  there is a  $y \in A$  such that  $d(x, y) < \varepsilon$ . Therefore  $x$  is a limit point of  $A$ , and thus  $x \in \overline{A}$ .

Let  $\varepsilon > 0$  and let  $x \in X$ , and let  $\delta = \varepsilon$ . Then:

$$\text{dist}(y, A) = \inf\{d(y, z) : z \in A\} \quad (33.1.39a)$$

$$\leq \inf\{d(x, y) + d(x, z) : z \in A\} \quad (33.1.39b)$$

$$= d(x, y) + \inf\{d(x, z) : z \in A\} \quad (33.1.39c)$$

$$= d(x, y) + \text{dist}(x, A) \quad (33.1.39d)$$

Similarly:

$$\text{dist}(x, A) \leq d(x, y) + \text{dist}(y, A) \quad (33.1.40)$$

And therefore, if  $d(x, y) < \varepsilon$ :

$$|\text{dist}(x, A) - \text{dist}(y, A)| \leq d(x, y) < \varepsilon \quad (33.1.41)$$

Finally, let:

$$f(x) = \frac{\text{dist}(x, B)}{\text{dist}(x, A) + \text{dist}(x, B)} \quad (33.1.42)$$

As  $A$  and  $B$  are disjoint and closed,  $f(x)$  is well defined for all  $x \in X$ . Moreover,  $0 \leq f(x) \leq 1$ . If  $f(x) = 0$ , then  $\text{dist}(x, B) = 0$ , and thus  $x \in \overline{B}$ . But  $B$  is closed, and therefore  $x \in B$ . If  $f(x) = 1$ , then  $\text{dist}(x, A) = 0$ , and thus  $x \in \overline{A}$ . Again, as  $A$  is closed,  $x \in A$ . But  $\text{dist}(x, B)$  is continuous, and  $\text{dist}(x, A) + \text{dist}(x, B)$  is continuous, and the quotient of continuous functions is continuous. Thus,  $f$  is continuous.

**Problem 33.1.11** Show that a Cauchy sequence with a convergent subsequence is convergent.

*Solution* For let  $a : \mathbb{N} \rightarrow X$  be a Cauchy sequence and let  $k : \mathbb{N} \rightarrow \mathbb{N}$  be such that  $a \circ k$  is a convergent subsequence, and let  $x$  be the limit. That is,  $k$  is a strictly monotonically increasing sequence of natural numbers. Let  $\varepsilon > 0$ . Then there is an  $N_1 \in \mathbb{N}$  such that, for all  $k_n > N_1$ ,  $n \in \mathbb{N}$ , it is true that  $d(a_{k_n}, a) < \varepsilon/2$ . But  $a$  is Cauchy, and thus there is an  $N_2 \in \mathbb{N}$  such that, for all  $n, m \in \mathbb{N}$  such that  $n, m > N_2$ , it is true that  $d(a_n, a_m) < \varepsilon/2$ . Let  $N = \max\{N_1, N_2\}$ . Then for all  $n > N$ ,  $n \in \mathbb{N}$ ,  $k_n > N$  since  $k$  is increasing, and thus:

$$d(a_n, x) \leq d(a_n, a_{k_n}) + d(a_{k_n}, x) < \varepsilon \quad (33.1.43)$$

Therefore,  $a_n \rightarrow x$ .

**Problem 33.1.12** Let  $F : \mathbb{N} \times X \rightarrow \mathbb{C}$  be a sequence of continuous functions and let  $f : X \rightarrow \mathbb{C}$  be such that  $F_n(x) \rightarrow f$  uniformly. Show that  $f$  is continuous.

*Solution* For let  $\varepsilon > 0$ . As  $F_n \rightarrow f$  uniformly, there is an  $N \in \mathbb{N}$  such that for all  $x \in X$ , it is true that:

$$|F_N(x) - f(x)| < \frac{\varepsilon}{3} \quad (33.1.44)$$

But  $F_N(x)$  is continuous, and thus for  $x \in X$  there is a  $\delta > 0$  such that  $d(x, x_0) < \delta$  implies that:

$$|F_N(x) - F_N(x_0)| < \frac{\varepsilon}{3} \quad (33.1.45)$$

But then:

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - F_N(x)| + |F_N(x) - F_N(x_0)| + |F_N(x_0) - f(x_0)| \\ &\quad (33.1.46a) \end{aligned}$$

$$< \varepsilon \quad (33.1.46b)$$

Thus,  $f$  is continuous.

**Problem 33.1.13** Let  $X$  be a metric space. Recall that we say  $f : X \rightarrow \mathbb{C}$  is bounded if  $\|f\|_\infty < \infty$ . A sequence of functions  $f_n : X \rightarrow D$  is uniformly bounded if there is an  $M$  such that  $\|f_n\|_\infty < M$  for all  $n$ . Also,  $f_n$  is uniformly Cauchy if for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that  $n, m > N$  implies  $|f_n(x) - f_m(x)| < \varepsilon$  for all  $x \in X$ . Show that a uniformly Cauchy sequence  $f_n$  of bounded functions is uniformly bounded.

*Solution* For let  $F : \mathbb{N} \times X \rightarrow \mathbb{C}$  be a sequence of functions such that, for all  $n \in \mathbb{N}$ , there is an  $M_n \in \mathbb{R}^+$  such that:

$$\|F_n\|_\infty < M_n \quad (33.1.47)$$

And such that  $F$  is uniformly Cauchy. Let  $\varepsilon = 1$ . Then, as  $F$  is uniformly Cauchy, there is an  $N \in \mathbb{N}$  such that, for all  $n, m \in \mathbb{N}$  such that  $n, m > N$ , and for all  $x \in X$ , it is true that:

$$|F_n(x) - F_m(x)| < \varepsilon \quad (33.1.48)$$

Then, for all  $n > N$  and for all  $x \in X$ :

$$|F_n(x)| = |F_n(x) + F_{N+1}(x) - F_{N+1}(x)| \quad (33.1.49a)$$

$$\leq |F_n(x) - F_{N+1}(x)| + |F_{N+1}(x)| \quad (33.1.49b)$$

$$< \varepsilon + M_{N+1} \quad (33.1.49c)$$

$$= M_{N+1} + 1 \quad (33.1.49d)$$

Let:

$$M = \max \left( \{M_n : n \in \mathbb{Z}_N\} \cup \{M_{N+1} + 1\} \right) \quad (33.1.50)$$

Then for all  $n \in \mathbb{N}$ , and for all  $x \in X$ :

$$|F_n(x)| \leq M \quad (33.1.51)$$

Therefore,  $F$  is uniformly bounded.

**Problem 33.1.14** We say that  $D$  is dense in  $X$  if  $\overline{D} = X$ . Show that  $D$  is dense if and only if  $D$  meets every non-empty open subset.

*Solution* Suppose  $D$  is a dense subset of  $X$  and let  $\mathcal{U} \subseteq X$  be an open subset. Suppose  $\mathcal{U} \cap D = \emptyset$ . Let  $x \in \mathcal{U}$ . As  $\mathcal{U}$  is open, there is an  $r > 0$  such that:

$$B_r^{(X,d)}(x) \subseteq \mathcal{U} \quad (33.1.52)$$

But if  $D$  is dense in  $X$ , then  $x$  is a limit point of  $D$ . Thus there is a sequence  $a : \mathbb{N} \rightarrow D$  such that  $a_n \rightarrow x$ . But if  $a_n \rightarrow x$ , then there is an  $N \in \mathbb{N}$  such that, for all  $n \in \mathbb{N}$  and  $n > N$ , we have:

$$a_n \in B_r^{(X,d)}(x) \quad (33.1.53)$$

But then,  $a_n \in \mathcal{U}$ , a contradiction as as  $a_n \in D$  and  $\mathcal{U}$  and  $\mathcal{D}$  are disjoint. Therefore, etc. Now suppose  $D$  meets every open set. Suppose  $x \notin \overline{D}$ . Define the following:

$$A_n = \left\{ y \in B_{1/n}^{(X,d)}(x) : y \in D \right\} \quad (33.1.54)$$

Then  $A_n$  is non-empty for all  $n \in \mathbb{N}$ , since  $D$  meets every open set. By choice there is a sequence  $a : \mathbb{N} \rightarrow D$  such that  $a_n \in A_n$  for all  $n \in \mathbb{N}$ . But then  $x$  is a limit point of  $D$ , a contradiction. Thus,  $\overline{D} = X$ .

**Problem 33.1.15** Let  $(x_n)$  be a sequence in a complete metric space  $(X, \rho)$ . Suppose that  $\rho(x_n, x_{n+1}) < 2^{-n}$  for all  $n \in \mathbb{N}$ . Conclude that  $(x_n)$  is convergent. What if instead we have that  $\rho(x_n, x_{n+1}) < 1/n$ ?

*Solution* For let  $\varepsilon > 0$ . Let  $N \in \mathbb{N}$  such that  $2^{1-N} < \varepsilon$ . But then for  $n, m > N$ :

$$\rho(x_n, x_m) \leq \sum_{k=\min(n,m)}^{\max(n,m)} d(x_k, x_{k+1}) \leq \sum_{k=N}^{\infty} d(x_k, x_{k+1}) \leq \sum_{k=N}^{\infty} \frac{1}{2^k} \quad (33.1.55)$$

But by applying the geometric series, we have:

$$\sum_{k=N}^{\infty} \frac{1}{2^k} = 2^{1-N} < \varepsilon \quad (33.1.56)$$

Thus  $x_n$  is Cauchy, and Cauchy sequences converge in a complete metric space. Therefore, etc. If we replace  $2^{-n}$  with  $n^{-1}$ , the result may not hold. For let  $X = \mathbb{R}$ , which is complete with the standard metric, and let  $a : \mathbb{N} \rightarrow \mathbb{R}$  be defined by  $a_n = \ln(n)$ . Applying some calculus, we have:

$$d(a_{n+1}, a_n) = \ln(n+1) - \ln(n) = \int_n^{n+1} \frac{1}{x} dx < \frac{1}{n} \quad (33.1.57)$$

But  $\ln(n)$  is not a convergent sequence.

**Problem 33.1.16** A metric space is separable if it has a countable dense subset. Show that a metric space  $X$  is separable if and only if there is a countable family  $\mathcal{D}$  of open sets such that every open set in  $X$  is the union of open sets in  $\mathcal{D}$ .

*Solution* For let  $(X, d)$  be a separable metric space, and let  $A$  be a countable dense subset. Let:

$$\mathcal{D} = \bigcup_{n \in \mathbb{N}} \bigcup_{a \in A} B_{n^{-1}}^{(X,d)}(a) \quad (33.1.58)$$

Then  $\mathcal{D}$  is countable, and for all  $\mathcal{U} \in \mathcal{D}$ ,  $\mathcal{U}$  is open. Let  $\mathcal{O}$  be an open subset of  $X$ . Define:

$$\mathcal{V} = \{\mathcal{U} \in \mathcal{D} : \mathcal{U} \subseteq \mathcal{O}\} \quad (33.1.59)$$

Then:

$$\bigcup_{\mathcal{U} \in \mathcal{V}} \mathcal{U} \subseteq \mathcal{O} \quad (33.1.60)$$

Suppose the converse is false, and let  $x \in \mathcal{O}$  be such that it is not contained in this union. But  $\mathcal{O}$  is open, and thus there is an  $r > 0$  such that:

$$B_r^{(X,d)}(x) \subseteq \mathcal{O} \quad (33.1.61)$$

But by the Archimedean principle, there is an  $n \in \mathbb{N}$  such that  $n^{-1} < r/2$ . But  $A$  is dense in  $\mathcal{O}$  and thus there is a  $y \in A$  such that  $d(x, y) < n^{-1}$ . But then:

$$x \in B_{n^{-1}}^{(X,d)}(y) \subseteq B_r^{(X,d)}(x) \subseteq \mathcal{O} \quad (33.1.62)$$

And thus  $x$  is contained in this union, a contradiction. Therefore, etc. Going the other way, suppose  $(X, d)$  is a metric space such that there exists a countable set  $\mathcal{D}$  of open subsets of  $X$  such that, for all open sets  $\mathcal{O}$ ,  $\mathcal{O}$  is the union of elements of  $\mathcal{D}$ . That is, there is a sequence  $A : \mathbb{N} \rightarrow \mathcal{P}(X)$  such that:

$$\mathcal{D} = \{A_n : n \in \mathbb{N}\} \quad (33.1.63)$$

But then by choice there is a sequence:

$$a : \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} A_n \quad (33.1.64)$$

Such that, for all  $n \in \mathbb{N}$ ,  $a_n \in A_n$ . Let  $y \in X$  and let  $\varepsilon > 0$ , define:

$$\mathcal{V} = B_\varepsilon^{(X,d)}(y) \quad (33.1.65)$$

But then  $\mathcal{V}$  is an open subset of  $(X, d)$  and is therefore the union of elements of  $\mathcal{D}$ . That is, there is a sequence  $k : \mathbb{N} \rightarrow \mathbb{N}$  such that:

$$\mathcal{V} = \bigcup_{n \in \mathbb{N}} A_{k_n} \quad (33.1.66)$$

Where we write  $k_n$  to denote  $k(n)$ . But then:

$$d(y, a_{k_n}) < \varepsilon \quad (33.1.67)$$

For all  $n \in \mathbb{N}$ . Thus the set:

$$\mathcal{A} = \{a_n : n \in \mathbb{N}\} \quad (33.1.68)$$

Is a a countable dense subset, and  $(X, d)$  is separable.

## 33.2 Homework II

**Problem 33.2.1** Show that  $X$  is compact if and only if every collection of closed sets  $\mathcal{F}$  with the finite intersection property is such that:

$$\bigcap_{F \in \mathcal{F}} F \neq \emptyset \quad (33.2.1)$$

*Solution* For suppose  $X$  is compact and suppose there is a collection  $\mathcal{F}$  of closed sets with the finite intersection property such that:

$$\bigcap_{F \in \mathcal{F}} F = \emptyset \quad (33.2.2)$$

But, for all  $F \in \mathcal{F}$ ,  $F$  is closed, and therefore  $F^C$  is open. But then:

$$X = \emptyset^C = \left( \bigcap_{F \in \mathcal{F}} F \right)^C = \bigcup_{F \in \mathcal{F}} F^C \quad (33.2.3)$$

And thus:

$$\mathcal{O} = \{F^C : F \in \mathcal{F}\} \quad (33.2.4)$$

Is an open cover of  $X$ . But  $X$  is compact, and therefore there is a finite subcover  $\Delta \subseteq \mathcal{O}$ . But then:

$$\emptyset = X^C = \left( \bigcup_{U \in \Delta} U \right)^C = \bigcap_{U \in \Delta} U^C \quad (33.2.5)$$

But  $U^C \in \mathcal{F}$  for all  $U \in \Delta$ . And  $\Delta$  is finite. Thus there is a finite subset of  $\mathcal{F}$  such that the intersection is empty, a contradiction as  $\mathcal{F}$  has the finite intersection property. Therefore, etc. Now suppose  $X$  is such that every collection of closed sets  $\mathcal{F}$  with the finite intersection property is such that the intersection over all elements is non-empty. Suppose  $X$  is not compact. Then there is an open cover  $\mathcal{O}$  such that there is no finite subcover. Let:

$$\mathcal{F} = \{U^C : U \in \mathcal{O}\} \quad (33.2.6)$$

But then for any finite subset, the intersection is non-empty. For if not then  $\mathcal{O}$  has a finite subcover, which it does not. But then  $\mathcal{F}$  is a collection of closed sets with the finite intersection property, and therefore:

$$\bigcap_{F \in \mathcal{F}} F \neq \emptyset \quad (33.2.7)$$

But:

$$\emptyset = X^C = \left( \bigcup_{U \in \mathcal{O}} U \right)^C = \bigcap_{F \in \mathcal{F}} F \quad (33.2.8)$$

A contradiction. Therefore,  $X$  is compact.

**Problem 33.2.2** Show that  $E \subseteq X$  is totally bounded if and only if there is a finite  $\varepsilon$ -net for all  $\varepsilon > 0$ .

*Solution* If  $E \subseteq X$  is totally bounded, then for all  $\varepsilon > 0$  there are finitely many points  $x_k, k \in \mathbb{Z}_n$  such that:

$$E \subseteq \bigcup_{k=1}^n B_\varepsilon^{(X,d)}(x_k) \quad (33.2.9)$$

But then:

$$\mathcal{O} = \{B_\varepsilon^{(X,d)}(x_k) : k \in \mathbb{Z}_n\} \quad (33.2.10)$$

Is a finite  $\varepsilon$ -net of  $E$ . If, for all  $\varepsilon > 0$ , there is a finite  $\varepsilon$ -net of  $E$ , then there are finitely many points  $x_k, k \in \mathbb{Z}_n$  such that:

$$\mathcal{O} = \{B_\varepsilon^{(X,d)}(x_k) : k \in \mathbb{Z}_n\} \quad (33.2.11)$$

Is an open cover of  $E$ . But then for all  $\varepsilon > 0$  there are finitely many open balls that cover  $E$ , and therefore  $E$  is totally bounded.

**Problem 33.2.3** Suppose  $(X, d_X)$  is compact and that  $f : (X, d_X) \rightarrow (Y, d_Y)$  is continuous. Show that  $f(X)$  is compact.

*Solution* For let  $\mathcal{O}$  be an open cover of  $f(X)$ . But  $f$  is continuous, and thus for all  $\mathcal{U} \in \mathcal{O}$ ,  $f^{-1}(\mathcal{U})$  is an open subset of  $X$ . But then:

$$\Delta = \{f^{-1}(\mathcal{U}) : \mathcal{U} \in \mathcal{O}\} \quad (33.2.12)$$

Is an open cover of  $X$ . But  $X$  is compact, and thus there is a finite sub-cover  $\Lambda$ . But then:

$$\mathcal{O} = \{\mathcal{U} : f^{-1}(\mathcal{U}) \in \Lambda\} \quad (33.2.13)$$

Is a finite subcover of  $f(X)$ , and therefore  $f(X)$  is compact.

**Problem 33.2.4** Let  $X = (0, 1)$  and let  $\delta_x > 0$  be such that:

$$y \in B_{\delta_x}^{(X, ||)}(x) \implies \left| \frac{1}{x} - \frac{1}{y} \right| < 1 \quad (33.2.14)$$

Show that:

$$\mathcal{O} = \{B_{\delta_x}^{(X, ||)}(x) : x \in X\} \quad (33.2.15)$$

Has no Lebesgue number.

*Solution* Suppose not, and let  $d > 0$  be a Lebesgue number. Then for all  $x \in (0, 1)$ , there is a  $\mathcal{U} \in \mathcal{O}$  such that:

$$B_d^{(X, ||)}(x) \subseteq \mathcal{U} \quad (33.2.16)$$

Let  $n \in \mathbb{N}$  be such that  $n^{-1} < d$ . Let  $x = n^{-1}/2$ . Then  $x \in (0, 1)$ . But  $d > n^{-1}$ , and thus:

$$B_d^{(X, ||)}(x) = (0, x + d) \quad (33.2.17)$$

But since  $x \in (0, 1)$ , there is a  $y \in (0, 1)$  such that:

$$B_d^{(X, ||)}(x) \subseteq B_{\delta_y}^{(X, ||)}(y) \quad (33.2.18)$$

But then for all  $z \in (0, x + d)$ , we have:

$$\left| \frac{1}{z} - \frac{1}{y} \right| < 1 \quad (33.2.19)$$

Let  $N \in \mathbb{N}$  be such that  $N > x^{-1} + y^{-1} + 2$ . But then  $N^{-1} \in (0, x + d)$ , and thus:

$$\left| \frac{1}{N^{-1}} - \frac{1}{y} \right| < 1 \quad (33.2.20)$$

But:

$$\left| \frac{1}{N^{-1}} - \frac{1}{y} \right| = |N - y^{-1}| > 2 \quad (33.2.21)$$

A contradiction. Thus,  $d$  is not a Lebesgue number.

**Problem 33.2.5** Show that a compact metric space has a countable dense subset.

*Solution* For if  $(X, d)$  is compact, then it is complete and totally bounded. But if it is totally bounded, for all  $n \in \mathbb{N}$  there exists an  $N \in \mathbb{N}$  and a sequence  $a : \mathbb{Z}_N \rightarrow X$  such that:

$$X = \bigcup_{k=1}^N B_{n^{-1}}^{(X, d)}(a_k) \quad (33.2.22)$$

Define the following:

$$A_n = \bigcup_{N \in \mathbb{N}} \left\{ a : \mathbb{Z}_N \rightarrow X : X = \bigcup_{k=1}^N B_{n^{-1}}^{(X, d)}(a_k) \right\} \quad (33.2.23)$$

Then, for all  $n \in \mathbb{N}$ ,  $A_n$  is non-empty. Then by choice there is a sequence:

$$f : \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} A_n \quad (33.2.24)$$

Such that, for all  $n$ ,  $f_n \in A_n$ . Let:

$$\mathcal{D} = \bigcup_{n \in \mathbb{N}} \text{Im}(f_n) \quad (33.2.25)$$

Where  $\text{Im}$  denotes the image of  $f_n$ . From construction, for all  $n \in \mathbb{N}$ ,  $\text{Im}(f_n)$  is finite, and thus  $\mathcal{D}$  is the countable union of countable sets, and is therefore countable. Moreover,  $\overline{\mathcal{D}} = X$ . For let  $x \in X$  and let  $\varepsilon > 0$ . By the Archimedean property, there and an  $n \in \mathbb{N}$  such that  $n^{-1} < \varepsilon$ . But:

$$X = \bigcup_{y \in f_n} B_{n^{-1}}^{(X,d)}(y) \quad (33.2.26)$$

And thus there is a  $y \in f_n$  such that  $d(x, y) < n^{-1}$ . But if  $y \in f_n$ , the  $y \in \mathcal{D}$ . Thus,  $x \in \overline{\mathcal{D}}$ . Therefore  $\mathcal{D}$  is a countable dense subset.

**Problem 33.2.6** Show that the family of functions  $\mathcal{F}$  defined on  $[0, 1]$  by  $f_n(x) = x^n$ , is equicontinuous at each  $x \in [0, 1]$ .

*Solution (1)* If  $F : \mathbb{N} \times X \rightarrow Y$  is a sequence of continuous functions such that  $F_n \rightarrow f$  uniformly, then  $F$  is point-wise equicontinuous. For let  $\varepsilon > 0$  and let  $x \in X$ . But  $F_n \rightarrow f$  uniformly, and  $F_n$  is continuous for all  $n \in \mathbb{N}$ , and therefore  $f$  is continuous. But then there is a  $\delta_1 > 0$  such that, for all  $x_0 \in X$  such that  $d_X(x, x_0) < \delta_1$ , we have that:

$$d_Y(f(x), f(x_0)) < \frac{\varepsilon}{3} \quad (33.2.27)$$

But  $F_n \rightarrow f$  uniformly, and thus there is an  $N \in \mathbb{N}$  such that, for all  $n > N$  and  $n \in \mathbb{N}$ , it is true that:

$$d_Y(F_n(x), F_n(x_0)) < \frac{\varepsilon}{3} \quad (33.2.28)$$

But then, for all  $n > N$ , and for all  $x_0 \in X$  such that  $d_X(x, x_0) < \delta_1$ , we have that:

$$\begin{aligned} d_Y(F_n(x), F_n(x_0)) &\leq d_Y(F_n(x), f(x)) + \\ &d_Y(f(x), f(x_0)) + d_Y(f(x_0), F_n(x_0)) < \varepsilon \end{aligned} \quad (33.2.29)$$

But  $F$  is continuous, and thus for all  $n \in \mathbb{Z}_N$  there is a  $\delta_n$  such that  $d_X(x, x_0) < \delta_n$  implies that:

$$d_Y(F_n(x), F_n(x_0)) < \varepsilon \quad (33.2.30)$$

Let:

$$\delta = \min \left( \{\delta_0\} \cup \{\delta_n : n \in \mathbb{Z}_N\} \right) \quad (33.2.31)$$

Now, for all  $x_0 < 1$ ,  $f_n(x) = x^n$  tends to zero uniformly on  $[0, x_0]$ . Therefore, etc.

*Solution (2)* For let  $\varepsilon > 0$ , and let  $x \in [0, 1)$ . If  $x = 0$ , Let  $\delta = \varepsilon \min\{\varepsilon, \frac{1}{2}\}$ . Then, for  $0 \leq x_0 < \delta$ , and for all  $n \in \mathbb{N}$ :

$$|x_0^n| < \delta^n \leq \varepsilon \left(\frac{1}{2}\right)^n < \varepsilon \quad (33.2.32)$$

Otherwise, let  $\delta_1 = \frac{1}{2} \min\{x, 1-x\}$  and let  $y = \delta_1 + x$ . Then  $0 < y < 1$ . By the mean value theorem, for all  $x_0$  there is a  $c_{x_0}$  such that  $|x - c_{x_0}| < |x - x_0|$  and such that:

$$|x^n - x_0^n| = nc_{x_0}^{n-1} |x - x_0| \quad (33.2.33)$$

But then:

$$|x^n - x_0^n| < ny^n \delta \quad (33.2.34)$$

But, since  $0 < y < 1$ ,  $ny^n$  is bounded. For let  $f : [0, \infty) \rightarrow \mathbb{R}$  be defined by:

$$f(x) = \frac{x}{y^{1-x}} \quad (33.2.35)$$

Then by L'Hôpital, we have:

$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} \frac{x}{y^{1-x}} = \lim_{x \rightarrow \infty} \frac{-1}{y^{1-x} \ln(y)} = \lim_{x \rightarrow \infty} \frac{-y^{x-1}}{\ln(y)} \quad (33.2.36)$$

But  $0 < y < 1$ , and therefore  $y^{x-1} \rightarrow 0$  as  $x \rightarrow 0$ . Therefore:

$$\lim_{x \rightarrow \infty} f(x) = 0 \quad (33.2.37)$$

But then for any sequence  $a : \mathbb{N} \rightarrow \mathbb{R}$  such that  $a_n \rightarrow \infty$ , we have  $f(a_n) \rightarrow 0$ . Therefore,  $ny^{n-1}$  converges to zero. But convergent sequences are bounded sequences, and therefore there is an  $M \in \mathbb{R}^+$  such that, for all  $n \in \mathbb{N}$ :

$$|ny^{n-1}| \leq M \quad (33.2.38)$$

Let  $\delta = \min\{\frac{\varepsilon}{M}, \delta_1\}$ . Then for all  $x_0 \in (0, 1)$  such that  $|x - x_0| < \delta$ , we have:

$$|x^n - x_0^n| = nc_{x_0}^{n-1} |x - x_0| < ny^{n-1} \delta < \varepsilon \quad (33.2.39)$$

**Problem 33.2.7** Show that an equicontinuous family of functions on a compact metric space is uniformly equicontinuous.

*Solution* For let  $(X, d_X)$  be a compact metric space and let  $\mathcal{F}$  be a family of equicontinuous functions to a metric space  $(Y, d_Y)$ . Let  $\varepsilon > 0$ . Then, as  $\mathcal{F}$  is equicontinuous, for all  $x \in X$  there exists a  $\delta_x > 0$  such that, for all  $f \in \mathcal{F}$ , we have:

$$x_0 \in B_{\delta_x}^{(X, d_X)}(x) \implies f(x_0) \in B_{\varepsilon/2}^{(Y, d_Y)}(f(x)) \quad (33.2.40)$$

But then:

$$\mathcal{O} = \left\{ B_{\delta_x}^{(X, d_X)}(x) : x \in X \right\} \quad (33.2.41)$$

Is an open cover of  $X$ . But  $(X, d)$  is compact, and therefore this cover has a Lebesgue number  $\delta > 0$ . If  $x \in X$ , then there is a  $y \in X$  such that:

$$B_\delta^{(X, d_X)}(x) \subseteq B_{\delta_y}^{(X, d_X)}(y) \quad (33.2.42)$$

But then, if  $d_X(x, x_0) < \delta$ , then:

$$x_0 \in B_\delta^{(X, d_X)}(x) \Rightarrow x_0 \in B_{\delta_y}^{(X, d_X)}(y) \Rightarrow f(x_0) \in B_{\varepsilon/2}^{(Y, d_Y)}(f(y)) \quad (33.2.43)$$

And therefore:

$$d_Y(f(x), f(x_0)) \leq d_Y(f(x), f(y)) + d_Y(f(y), f(x_0)) < \varepsilon \quad (33.2.44)$$

Thus,  $\mathcal{F}$  is uniformly equicontinuous.

**Problem 33.2.8** Show that a subset of a compact metric space is compact if and only if it is closed.

*Solution* For let  $(X, d)$  be a compact metric space and let  $(E, d_E)$  be a compact subspace. Suppose  $E$  is not closed. Then  $E^C$  is not open, and therefore there is an  $x \in E^C$  such that, for all  $\varepsilon > 0$ :

$$B_\varepsilon^{(X, d)}(x) \cap E \neq \emptyset \quad (33.2.45)$$

Let:

$$\mathcal{O} = \left\{ \text{Cl}\left(B_\varepsilon^{(X, d)}(x)\right)^C : \varepsilon \in \mathbb{R}^+ \right\} \quad (33.2.46)$$

Where  $\text{Cl}$  denotes the closure of a set. Then  $\mathcal{O}$  is an open cover of  $E$ . But  $(E, d_E)$  is compact, and thus there is a finite subcover  $\Delta$ . But then there is at least  $r \in \mathbb{R}^+$  such that:

$$\text{Cl}\left(B_r^{(X, d)}(x)\right)^C \in \Delta \quad (33.2.47)$$

But then, for all  $0 < \varepsilon < r$ , we have:

$$B_\varepsilon^{(X, d)}(x) \cap E = \emptyset \quad (33.2.48)$$

A contradiction. Therefore,  $E$  is closed. Suppose  $(X, d)$  is compact and  $E \subseteq X$  is closed. Suppose  $(E, d_E)$  is not compact. Then there is an open cover  $\mathcal{O}_E$  of  $E$  with no finite subcover. But  $E$  is closed, and thus  $E^C$  is open. But then:

$$\mathcal{O}_X = \mathcal{O}_E \cup \{E^C\} \quad (33.2.49)$$

Is an open cover of  $X$ . But  $(X, d)$  is compact, and therefore there is a finite subcover  $\Delta_X$ . But then:

$$\Delta_E = \Delta_X \setminus \{E^C\} \quad (33.2.50)$$

Is a finite subcover of  $\mathcal{O}_E$ , a contradiction. Therefore,  $(E, d_E)$  is compact.

**Part XIX**

# **Fourier Analysis**



# CHAPTER 34

---

## Fourier Analysis

---

Fourier analysis is deeply tied to the theory of integration, and as such it is worth while to develop a few of the basic elements of this field. Many of theorems we wish to use, such as the *convolution theorem*, *Fubini's theorem*, and the *Fourier inversion theorem*, are often presented with hand-wavy “proofs,” that obscure some of the problems that arise when applying these results to real-world problems. We will not dive into the entirety of measure theory, but rather present the elementary definitions, provide examples, and move on to the more important theorems.

### 34.0.1 Basic Notions

We wish to define what it means for some function  $f : \Omega \rightarrow \mathbb{R}$  to be *integrable*, where  $\Omega$  is some space. Any attempt at defining an integral will require one to start with approximations such as the following:

$$\int_{\Omega} f(x) dx \approx \sum_n f(x_n) \mu(X_n) \tag{34.0.1}$$

Where  $X_n$  is a bunch of sets that partition  $\Omega$ ,  $x_n \in X_n$  for all  $n$ , and  $\mu(X_n)$  is the *size* or the *width* of  $X_n$ . This is precisely what is done in a calculus course where the Riemann integral is defined. To make this concrete, we'll need to decide what sets  $X_n$  are allowed to partition  $\Omega$ , and what are the properties of our *measure*  $\mu$ . Throughout,  $\emptyset$  is used to denote the empty set. This is the set with nothing in it. Our inclusion of this set in various definitions is for technical reasons that we won't often be concerned with.

---

**Definition 34.0.1:  $\sigma$ -Algebras**


---

A  $\sigma$ -Algebra on a set  $\Omega$  is a collection of subsets  $\mathcal{A}$  of  $\Omega$  such that:

1. It is true that  $\emptyset \in \mathcal{A}$  and that  $\Omega \in \mathcal{A}$ .
2. For all  $A \in \mathcal{A}$ , it is true that the complement of  $A$  is in  $\mathcal{A}$ . That is,  $A^C \in \mathcal{A}$ .
3. For any sequence  $A : \mathbb{N} \rightarrow \mathcal{A}$  of sets in  $\mathcal{A}$ , so is their intersection:

$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{A} \quad (34.0.2)$$

The elements of  $\mathcal{A}$  are called the *measurable subsets* of  $\Omega$ . Given a set  $\Omega$ , and a  $\sigma$ -Algebra  $\mathcal{A}$  on  $\Omega$ , we call the pair  $(\Omega, \mathcal{A})$  a *measure space*. ■

---

**Example 34.0.1**


---

Given a set  $\Omega$ , there are two simple  $\sigma$ -Algebras that one can define. Let  $\mathcal{A} = \{\emptyset, \Omega\}$ . This satisfies all three properties and is called the trivial  $\sigma$ -Algebra. Going in the other direction, if we let  $\mathcal{A}$  be the set of *all* subsets of  $\Omega$  (Also known as the *power set* of  $\Omega$ , denoted  $\mathcal{P}(\Omega)$ ), then this also a  $\sigma$ -Algebra. ■

The motivation for defining measurable sets in such a way is to allow one to easily describe *measures* and *probabilities* later.

---

**Definition 34.0.2: Borel  $\sigma$ -Algebra**


---

The Borel  $\sigma$ -Algebra is the *smallest*  $\sigma$ -Algebra on  $\mathbb{R}$ , denoted  $\mathcal{B}$ , such that for all  $a < b$ , the interval  $(a, b)$  is a measurable set. That is,  $(a, b) \in \mathcal{B}$ . ■

---

**Definition 34.0.3: Measures**


---

A measure on a measure space  $(\Omega, \mathcal{A})$  is a function  $\mu : \mathcal{A} \rightarrow \mathbb{R}$  such that:

1. For all  $A \in \mathcal{A}$ ,  $\mu(A) \geq 0$ .
2.  $\mu(\emptyset) = 0$ .
3. Given a *mutually disjoint* list of sets  $A_1, A_2, \dots$  that are contained in  $\mathcal{A}$ , the following is true:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \quad (34.0.3)$$

The triple  $(\Omega, \mathcal{A}, \mu)$  is called a *measurable space*. ■

It is important to remember that we are trying to model *size*, and this is what motivates our definition of measure. The first rule says that the width, or length, or size of a set is non-negative, and the second rule states that the size of *nothing* is simply zero. The last rule, which is called countable additivity, is very important but also intuitive. We may define the length of the set interval  $(0, 1)$  as 1, and the length of  $(a, b)$  to be  $b - a$  (For  $a < b$ ). What about the length of the *union* of the intervals  $(0, 1)$  and  $(3, 4)$ ? Since they have no overlap, we may as well add the lengths of the two individual intervals and claim that the length of the whole is 2. This is precisely what countable additivity tells us.

### 34.0.2 Fourier Series

### 34.0.3 The Fourier Transform

Suppose you are asked to compute  $z = x/y$  for two non-zero real numbers  $x$  and  $y$ . We could perform long-hand division, or transform the problem into subtraction by using the natural logarithm.

$$z = x/y \Rightarrow \ln(z) = \ln(x/y) \Rightarrow \ln(z) = \ln(x) - \ln(y) \quad (34.0.4)$$

Provided that  $\ln(x)$  and  $\ln(y)$  are somehow known, one can compute the difference and then compute  $z$  by exponentiating the result. In a similar manner, the Fourier transform is often introduced as a tool for converting one problem into another.

#### **Definition 34.0.4: Fourier Transform**

---

The Fourier transform of a complex valued integrable function  $f : \mathbb{R} \rightarrow \mathbb{C}$  is the function  $\mathcal{F}_\xi(f) : \mathbb{R} \rightarrow \mathbb{C}$  defined by:

$$\mathcal{F}_\xi(f) = \int_{-\infty}^{\infty} f(t) \exp(-2\pi i \xi t) dt \quad (34.0.5)$$

This is also called the *spectrum* of  $f$ . ■

The requirement that  $f$  be integrable is to avoid strange issues in mathematics. For the sake of physical application, one may assume every function is integrable. Mathematically this is far from true, but oh well. For the sake of Fourier Analysis, when we say integrable we mean Lebesgue integrable. This simply means that:

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty \quad (34.0.6)$$

**Example 34.0.2**

Consider the hat function:

$$f(t) = \begin{cases} 0, & |t| \leq 1 \\ 1, & |t| > 1 \end{cases} \quad (34.0.7)$$

We can compute the Fourier transform of the this explicitly:

$$\mathcal{F}_\xi(f) = \int_{-\infty}^{\infty} f(t) \exp(-2\pi i \xi t) dt = \int_{-1}^1 \exp(-2\pi i \xi t) dt \quad (34.0.8)$$

Here we invoke Euler's Theorem, Thm. 30.1.18, and note that the integral has symmetric limits. But sin is an *odd* function, and thus it's integral is zero, and cos is an even function. Thus, we are left with:

$$\mathcal{F}_\xi(f) = 2 \int_0^1 \cos(2\pi i \xi t) dt = \frac{\sin(2\pi \xi)}{\pi \xi} \quad (34.0.9)$$

We can be even more general, defining:

$$f(x) = \begin{cases} 1, & a \leq x \leq b \\ 0, & \text{Otherwise} \end{cases} \quad (34.0.10)$$

The Fourier transform is then:

$$\mathcal{F}_\xi(f) = \frac{i(\exp(-2\pi i \xi b) - \exp(-2\pi i \xi a))}{2\pi \xi} \quad (34.0.11)$$

Thus we see that the range of the Fourier transform generally lies in the complex plane. Only with sufficient symmetry does the problem collapse down to  $\mathbb{R}$ . ■

Recall that a complex number has a polar representation  $z = r \exp(i\theta)$ . Similarly, for a complex valued function we can write:

$$f(z) = R(r) \exp(i\Theta(\theta)) \quad (34.0.12)$$

For the Fourier transform of a function, the function  $R(r)$  is called the principle amplitude, and  $\Theta(\theta)$  is the *phase offset* from this amplitude. The Fourier transform of the hat function defined from -1 to 1 is plotted in Fig. 34.1.

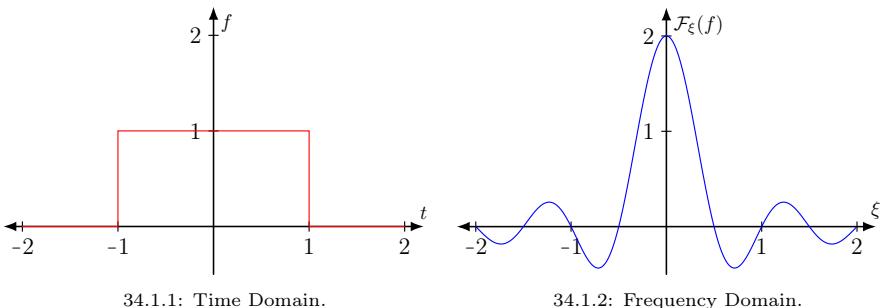


Fig. 34.1: Fourier Transform of the Hat Function.

**Example 34.0.3**

Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be defined by:

$$f(t) = \begin{cases} \beta \exp(-\alpha t), & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (34.0.13)$$

Where  $\alpha$  and  $\beta$  are positive real numbers. Since  $f(t)$  decays to zero rapidly as  $t \rightarrow \infty$ , we see that  $f$  is a Lebesgue integrable function and has a Fourier transform. We can compute this using the standard methods obtained from a course on integral calculus.

$$\mathcal{F}_\xi(f) = \int_{-\infty}^{\infty} f(t) \exp(-2\pi i \xi t) dt = \beta \int_0^{\infty} \exp(-\alpha t) \exp(-2\pi i \xi t) dt \quad (34.0.14)$$

Using the product rule for exponents, we can reduce this to the following line integral:

$$\mathcal{F}_\xi(f) = \beta \int_0^{\infty} \exp(-( \alpha + 2\pi i \xi)t) dt \quad (34.0.15)$$

Thus, we are integrating the exponential function along the line  $z = \theta t$ , where  $\theta = \tan^{-1}(2\pi\xi/\alpha)$ . Since  $\alpha$  is positive, we can use the result from Jordan's Lemma to obtain the solution:

$$\mathcal{F}_\xi(f) = \frac{\beta}{\alpha + 2\pi i \xi} \quad (34.0.16)$$

There are two ways to view the Fourier transform: The Cartesian form and the polar form. As was stated before, the polar form represents the *amplitude* and *phase offset* of the Fourier transform, whereas the Cartesian form simply represents the real and imaginary parts. To compute the Cartesian form, we simply invoke Eqn. 30.1.29 for the inverse of a complex number, and obtain:

$$\mathcal{F}_\xi(f) = \beta \frac{\alpha - 2\pi i \xi}{\alpha^2 + 4\pi^2 \xi^2} \quad (34.0.17)$$

For the polar form, we invoke Thm. 30.1.20, take the modulus of  $\mathcal{F}_\xi(f)$  and compute inverse tangents:

$$\mathcal{F}_\xi(f) = \frac{\beta}{\sqrt{\alpha^2 + 4\pi^2\xi^2}} \exp\left[i \tan^{-1}\left(\frac{-2\pi\xi}{\alpha}\right)\right] \quad (34.0.18)$$

The two are plotted below for the case of  $\alpha = \beta = 1$ . ■

### Definition 34.0.5: Inverse Fourier Transform

---

The inverse Fourier transform of a complex valued integrable function  $f : \mathbb{R} \rightarrow \mathbb{C}$  is the function  $\mathcal{F}_t^{-1}(f) : \mathbb{R} \rightarrow \mathbb{C}$  defined by:

$$\mathcal{F}_t^{-1}(f) = \int_{-\infty}^{\infty} f(\xi) \exp(2\pi i \xi t) d\xi \quad (34.0.19)$$

■

We now prove what is probably the most useful theorem in Fourier Analysis.

**Theorem 34.0.1.** *If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a continuous Lebesgue integrable function, and if its spectrum  $F$  is also continuous and Lebesgue integrable, then:*

$$f(t) = \int_{-\infty}^{\infty} F(\omega) \exp(2\pi i \omega t) d\omega \quad (34.0.20)$$

A powerful application of this is Shannon's Sampling Theorem.

**Theorem 34.0.2** (Shannon's Sampling Theorem). *If  $f(t)$  is a continuous Lebesgue integrable function such that its spectrum  $F(\omega)$  is differentiable and zero outside the interval  $[-W, W]$ , then  $f(t)$  is uniquely determined by the points  $f\left(\frac{n}{2W}\right)$ ,  $n \in \mathbb{N}$ .*

*Proof.* For let  $F$  be the spectrum of  $f$ . That is:

$$f(t) = \int_{-\infty}^{\infty} F(\omega) \exp(-2\pi i \omega t) d\omega \quad (34.0.21)$$

But  $F(\omega) = 0$  for  $|\omega| > W$ . Thus we have:

$$f(t) = \int_{-W}^{W} F(\omega) \exp(-2\pi i \omega t) d\omega \quad (34.0.22)$$

Then for  $n \in \mathbb{N}$ , we have:

$$f\left(\frac{n}{2W}\right) = \int_{-W}^{W} F(\omega) \exp\left(-2\pi i \frac{n}{2W} \omega\right) d\omega \quad (34.0.23)$$

But  $F$  is differentiable, and thus it's Fourier series converges. That is:

$$F(\omega) = \sum_{n=-\infty}^{\infty} \exp(2\pi i n \omega) \int_{-W}^W F(\tau) \exp(-2\pi i \frac{n}{2W} \tau) d\tau \quad (34.0.24a)$$

$$= \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right) e^{2\pi i n \omega} \quad (34.0.24b)$$

Therefore  $f\left(\frac{n}{2W}\right)$ ,  $n \in \mathbb{N}$  uniquely determines  $F(\omega)$ . But the spectrum  $F(\omega)$  uniquely determines  $f(t)$ . Therefore  $f(t)$  is uniquely determined and:

$$f(t) = \sum_{n=-\infty}^{\infty} \int_{-W}^W f\left(\frac{n}{2W}\right) \exp(2\pi i \omega(n+t)) d\omega \quad (34.0.25)$$

□

### 34.0.4 Convolutions

### 34.0.5 Sampling



# CHAPTER 35

## Chaos Theory

### 35.1 A Review of Differential Equations

#### 35.1.1 First Order Equations

The first differential equation that is often studied is  $\dot{x}(t) = ax(t)$ , where  $\dot{x}$  denotes the derivative of  $x$  with respect to  $t$ . In this equation  $a$  is some fixed constant parameter, and each real value  $a$  defines a different differential equation. We can solve this by integrating and invoking the fundamental theorem of calculus. In general, a differential equation is an equation that relates a differentiable functions to its derivatives. The general solution to a differential equation is the set of all functions that satisfy the differential equation.

##### Theorem 35.1.1

If  $x(t)$  is a differentiable function such that  $x(0) = x_0$  and there is a  $k \in \mathbb{R}$  such that for all  $t \in \mathbb{R}$ ,  $\dot{x}(t) = kx(t)$ , then  $x(t) = x_0 \exp(kt)$ .

*Proof.* Let  $y(t) = \exp(kt)$ . Then  $\dot{y}(t) = ky(t)$ , and  $y(t) \neq 0$  for all  $t \in \mathbb{R}$ . Let  $F(t) = x(t)/y(t)$ . Then:

$$\dot{F}(t) = \frac{y(t)\dot{x}(t) - \dot{y}(t)x(t)}{y^2(t)} = \frac{y(t)(\dot{x}(t) - kx(t))}{y^2(t)} = \frac{1}{y(t)}(\dot{x}(t) - kx(t))$$

But  $\dot{x}(t) - kx(t) = 0$ , and thus  $\dot{F}(t) = 0$  for all  $t \in \mathbb{R}$ . But then  $F(t)$  is a constant function. Thus there is a  $C \in \mathbb{R}$  such that  $x(t) = Cy(t)$ . But  $x(0) = x_0$ , and therefore  $C = x_0$ . Thus,  $x(t) = x_0 \exp(kt)$ .  $\square$

The proof shows that the solution to  $\dot{x}(t) = kx(t)$  is uniquely determined if  $x(0)$  is known. This can be replaced by knowledge of  $x(t_0)$  for any point  $t_0 \in \mathbb{R}$ . Such restraints are called initial conditions to a differential equation. Initial conditions are requirements that a solution to the differential equation must satisfy. Such requirements can be the value of  $x(t_0)$  at a certain point  $t_0$ , or a requirement on  $\dot{x}(t)$ .

### Example 35.1.1

Consider the following initial value problem:

$$\dot{x}(t) = ax(t) \quad x(t_0) = x_0 \quad (35.1.1)$$

Using the previous theorem, we can translate the problem by  $\exp(-kt_0)$  to create the initial value problem  $\dot{y}(t) = ky(t)$ ,  $y(0) = x_0$ . We know the solution to this is  $y(t) = x_0 \exp(kt)$ . Thus, the solution to the original initial value problem is:

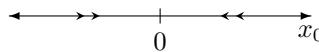
$$x(t) = x_0 \exp(k(t - t_0)) \quad (35.1.2)$$

An important class of solutions to differential equations are *equilibrium solutions*.

### Definition 35.1.1: Equilibrium Solution

An equilibrium solution to a differential equation is a solution  $x(t)$  such that  $\dot{x}(t) = 0$  for all  $t \in \mathbb{R}$ . That is, a constant solution  $x(t) = c$ .

Studying our first problem  $\dot{x}(t) = kx(t)$ , we see that the only equilibrium solution occurs when  $x_0 = 0$ . Another important concept is the limiting behavior of the solution to a differential equation. The limiting behavior is often dependent on the various parameters that may be present in a given differential equation. We can see this by further studying  $\dot{x}(t) = kx(t)$ . If  $k = 0$  we see that  $x(t)$  is a constant function. That is,  $x(t) = x_0$  for all  $t \in \mathbb{R}$ . For  $k > 0$ , solutions diverges monotonically to  $\infty$  ( $x_0 > 0$ ) or  $-\infty$  ( $x_0 < 0$ ). Finally, if  $k < 0$  then  $x(t)$  converges to 0 for all values of  $x_0$ . This notion can be represented by a *phase line*. A phase line is a one dimensional plot of the independent variable  $t$  where equilibrium solutions are marked and arrows indicating convergence or divergence to the equilibrium solutions are drawn. Consider again the example we've been studying:  $\dot{x}(t) = kx(t)$ . Suppose  $k < 0$ . We know that, for any  $x_0$ , the solution  $x(t)$  tends to zero as  $t$  tends to infinity. We can represent this with a phase line: The initial condition  $x_0 = 0$  thus gives rise to a *stable* solution to

Fig. 35.1: Phase line for  $\dot{x}(t) = kx(t)$  when  $k < 0$ .

this differential equation.

### Definition 35.1.2: Stable Equilibrium Solution

A stable equilibrium solution to a differential equation  $\dot{x}(t) = f(t, x(t))$ , is an equilibrium solution  $x(t) = x_0$  such that there exists a  $\delta > 0$  such that for all  $x_1 \in (x_0 - \delta, x_0 + \delta)$ , the solution to the initial value problem  $\dot{x}(t) = f(t, x(t)), x(t_0) = x_1$  converges to  $x_0$  as  $t \rightarrow \infty$

Being purely stable can be replaced with stable from above or stable from below. An equilibrium solution is called semi-stable if it is either stable from above or stable from below, but not totally stable. An unstable equilibrium solution is an equilibrium solution that is neither stable nor semi-stable. The stability of solutions often depends on the parameters involved in the differential equation. In the case of  $\dot{x}(t) = kx(t)$  we saw that as  $k$  cross zero the solutions drastically change. The value  $k = 0$  is said to be a *bifurcation* value of the differential equation. Bifurcations are values that a parameter can have that alter the limiting behavior of solutions.  $k = 0$  is a bifurcation since, for all  $k > 0$  we have that  $x(t)$  diverges monotonically, for all  $k < 0$  we see that  $x(t)$  converges to zero monotonically, and for  $k = 0$  all solutions are constants. Because of this, we say that there is a bifurcation at  $k = 0$ .

### The Logistics Population Model

First order differential equations have many applications in both the physical and the life sciences. Population modeling is a common such application. The growth rate of a population can be modelled based on two simple assumptions. The first is that, if the population is small, the growth rate is roughly proportional to the population. The second is that, if the population is too large, the growth rate is negative. For example, if the population is too large then there is not enough food and thus the population will start to decrease. To bridge the gap between these two assumptions we could make the rather reasonable assumption that the ratio of the population growth to the population is linear. Thus:

$$\frac{\dot{x}(t)}{x(t)} = ax(t) + b \quad (35.1.3)$$

Since the growth rate is negative for large  $x_0$  and positive for small  $x_0$ , there must be a value  $N$ , called the *ideal population*, such that  $\dot{x}(t) = 0$  for all  $t$ . If

we let  $\alpha$  be the *growth factor* of the population, Eqn. 35.1.3 becomes:

$$\dot{x}(t) = \alpha x(t) \left(1 - \frac{x(t)}{N}\right) \quad (35.1.4)$$

By hypothesis,  $\alpha$  must be positive. For if  $x(t) > N$  and if  $\alpha < 0$ , then  $\dot{x}(t) > 0$ . This would be quite strange as large populations would increase rapidly to infinity. From physical considerations we also require that  $N$  be positive. Otherwise we'd be speaking of *negative population*, which is nonsense. For the sake of easing the mathematics, we may normalize the problem so that the ideal population is  $N = 1$ . Let  $k$  be the normalized growth factor. This gives us the following normalized logistics population model:

$$\dot{x}(t) = kx(t)(1 - x(t)) \quad (35.1.5)$$

The logistics population model is another example of a first order differential equation. In general, a first order differential equation is of the form  $x(t) = f(t, x(t))$ . An  $n^{th}$  order differential equation is one of the form:

$$x^{(n)}(t) = f(t, x(t), \dot{x}(t), \ddot{x}(t), x^{(3)}(t), \dots, x^{(n-1)}(t))$$

Here we have used the notation  $x^{(n)}(t)$  to represent the  $n^{th}$  derivative of  $x$  with respect to  $t$ . A special subset of the general  $n^{th}$  order differential equation is that of the autonomous  $n^{th}$  order differential equations. These are differential equations where  $f$  is solely a function of  $x$  and its derivatives.

### Definition 35.1.3: Autonomous Differential Equation

An autonomous  $n^{th}$  order differential equation is a differential equation of the form:

$$x^{(n)}(t) = f(x(t), \dots, x^{(n-1)}(t)) \quad (35.1.6)$$

**Example 35.1.2: First Order Autonomous Differential Equations**

Consider the case of first order autonomous differential equations. Here we'd have:

$$\dot{x}(t) = f(x(t)) \quad (35.1.7)$$

If  $f$  is continuous then we can integrate this and obtain a solution for  $x$ :

$$F(x) \equiv \int \frac{1}{f(x)} dx = \int dt \quad (35.1.8)$$

If the integral on the left hand side produces an invertible function, then we can solve the differential equation and obtain  $x(t) = F^{-1}(t + C)$ , where  $C$  is a constant of integration. An equation like this is also called *separable*. That is, we can separate the  $x$  and  $t$  terms to create an expression on the left hand side which is written purely in  $x$ , and an expression on the right which is written purely in  $t$ .

The logistics population model is therefore a nonlinear first order autonomous differential equation. Nonlinear differential equations are usually very difficult to solve. When we have autonomy, however, the problem can become much easier. In the case of the logistics model, we solve for  $x(t)$  using standard techniques from Calculus.

$$\int \frac{1}{x(1-x)} dx = k \int dt \quad (35.1.9)$$

Recalling from elementary algebra the method of partial fraction decomposition, we have:

$$\frac{1}{x(1-x)} = \frac{1}{x} + \frac{1}{1-x} \quad (35.1.10)$$

This simplifies the integral, and so we obtain:

$$\int \left( \frac{1}{x} + \frac{1}{1-x} \right) dx = \ln(x) - \ln(1-x) = \ln\left(\frac{x}{1-x}\right) \quad (35.1.11)$$

Evaluating the right-hand side and exponentiating, we get:

$$\frac{x}{1-x} = A \exp(kt) \quad (35.1.12)$$

Where  $A$  is the exponential of the constant of integration. The left-hand side is an invertible function and its inverse is  $x/(1+x)$ . From this we obtain the solution to the logistics population model:

$$x(t) = \frac{1}{1 + C \exp(-kt)} \quad (35.1.13)$$

By studying Eqn. 35.1.5 we can find two equilibrium solution:  $x(t) = 0$  and  $x(t) = 1$ . This makes physical sense, for if  $x(t) = 0$  then the population is extinct and nothing more can be done, and if  $x(t) = 1$  then the population has reached its ideal value. From intuition it would seem that  $x(t) = 0$  is either unstable or semistable (But since we ignore negative populations, this would simply be unstable) and that  $x(t) = 1$  would be stable. And indeed, if  $0 < x(t) < 1$ , the from Eqn. 35.1.5 we have  $\dot{x}(t) > 0$ , so the population is increasing (To the ideal value). If  $x(t) > 1$ , then  $\dot{x}(t) < 0$  and thus the population is decreasing (Again, towards its ideal value). We can summarize this more generally for first-order autonomous differential equations. First we prove a helpful intermediate step.

**Theorem 35.1.2.** *If  $x(t)$  is a differentiable function such that  $x(t) \rightarrow a$  as  $t \rightarrow \infty$ , then there is a strictly increasing monotonic sequence  $c_n$  such that  $\dot{x}(c_n) \rightarrow 0$ .*

*Proof.* For let  $t_n = n$  for all  $n \in \mathbb{N}$ . Then  $x(t_n) \rightarrow a$  as  $t_n \rightarrow \infty$ . But convergent sequences are Cauchy sequences, and therefore  $x(t_{n+1}) - x(t_n) \rightarrow 0$ . And by the mean value theorem, for all  $n \in \mathbb{N}$  there is a  $c_n \in (t_n, t_{n+1})$  such that:

$$\dot{x}(c_n) = \frac{x(t_{n+1}) - x(t_n)}{t_{n+1} - t_n}$$

But  $t_{n+1} - t_n = 1$ , and thus  $\dot{x}(c_n) = x(t_{n+1}) - x(t_n)$ . But  $x(t_{n+1}) - x(t_n) \rightarrow 0$  and therefore  $\dot{x}(c_n) \rightarrow 0$ .  $\square$

**Theorem 35.1.3.** *If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is differentiable, and given the differential equation  $\dot{x}(t) = f(x(t))$ , if  $x_0$  is an equilibrium solution and  $f'(x_0) > 0$ , then  $x_0$  is an unstable equilibrium solution.*

*Proof.* If  $f(x_0) = 0$  and  $f'(x_0) > 0$ , then there is a  $\delta > 0$  such that, for all  $x \in (x_0 - \delta, x_0 + \delta)$ , we have:

$$\frac{f(x) - f(x_0)}{x - x_0} > 0$$

But  $f(x_0) = 0$ . Therefore, if  $x_1 > x_0$ , we have:

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} > 0 \Rightarrow f(x_1) - f(x_0) > 0 \Rightarrow f(x_1) > 0$$

Thus, for all  $x_1$  such that  $0 < x_1 - x_0 < \delta$ , the limit of the solution  $x(t)$  to the initial value problem  $\dot{x}(t) = f(x(t))$ ,  $x(0) = x_1$  can't converge to  $x_0$ . For if it did there would be a point  $t_0$  such that  $\dot{x}(t_0) < 0$  and  $0 < x(t_0) - x_0 < \delta$ , a contradiction. Similarly for if  $x_1 < x_0$ . Therefore,  $x_0$  is unstable.  $\square$

**Theorem 35.1.4.** If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is differentiable, and given the differential equation  $\dot{x}(t) = f(x(t))$ , if  $x_0$  is an equilibrium solution and  $f'(x_0) < 0$ , then  $x_0$  is a stable equilibrium solution.

*Proof.* If  $f'(x_0) < 0$  then there is a  $\delta > 0$  such that for all  $x_1 \in (x_0 - \delta, x_0 + \delta)$ ,  $f'(x_1) < 0$ . Then for all  $x_1$  such that  $0 < x_1 - x_0 < \delta$ , we have:

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} < 0 \Rightarrow f(x_1) - f(x_0) < 0 \Rightarrow f(x_1) < 0$$

□

### Notes on the Jordan Normal Form

Every square matrix  $A$  is similar to an upper triangular matrix  $J$  in *Jordan normal form* whose diagonal entries are the eigenvalues of  $A$ . That is, there exists an invertible matrix  $P$  such that  $P^{-1}AP = J$ . The trace of  $A$  is equal to the trace of  $J$ :

$$\text{Tr}(J) = \text{Tr}(P^{-1}AP) = \text{Tr}(P^{-1}PA) = \text{Tr}(IA) = \text{Tr}(A)$$

### Notes on Conjugacy

We have:

$$\begin{aligned} X'(t) &= AX(t), \quad X(0) = X_0 \\ \Rightarrow X(t) &= e^{tA}X_0 \\ \Rightarrow \phi^A(t, X_0) &= e^{tA}X_0 \end{aligned}$$

If  $B = T^{-1}AT$ , for some matrix  $T$ , then:

$$\begin{aligned} Y'(t) &= BY(t), \quad Y(0) = Y_0 \\ &= T^{-1}ATY(t) \\ \Rightarrow Y(t) &= e^{tT^{-1}AT}Y_0 \\ &= T^{-1}e^{tA}TY_0 \\ \Rightarrow \phi^B(t, Y_0) &= T^{-1}e^{tA}TY_0 \end{aligned}$$

Thus, the homeomorphism is  $h(X) = T^{-1}X$ , and we have:

$$\phi^B(t, h(X_0)) = \phi^B(t, T^{-1}X_0) = T^{-1}e^{tA}T(T^{-1}X_0) = T^{-1}e^{tA}X_0 = h(\phi^A(t, X_0))$$



**Part XX**

# **Special Functions**



# CHAPTER 36

---

## Numerical Analysis

---

### 36.1 Power Series

### 36.2 Asymptotic Expansions

### 36.3 Stationary Phase Approximation

Suppose  $g$  is an analytical function about the origin (i.e. it has a convergent MacLaurin series), and consider the integral:

$$I(k) = \int_a^b e^{ikg(x)} dx \quad (36.3.1)$$

Suppose that there is a  $c \in (a, b)$  such that  $g'(c) = 0$  and  $g''(c) \neq 0$ . Then:

$$I(k) = \exp(ikg(c)) \int_a^b \exp(ik[g(x) - g(c)]) dx \quad (36.3.2)$$

$$= \exp(ikg(c)) \int_a^b \exp\left(ik\left[\frac{g''(c)}{2}(x - c)^2 + \dots\right]\right) dx \quad (36.3.3)$$

Higher terms are extremely oscillatory, and so we neglect them. Note that higher terms can indeed cancel each other out, meaning these neglected terms may not be negligible. For example, if  $g(x) = -\sin(\pi x)$ , then  $\exp(ig(x))$  is never too oscillatory. However, so long as the interval  $[a, b]$  is small enough, the approximation is still valid. The previously mentioned  $g(x)$  is how one

approximates the  $J_0(x)$  Bessel function. Our integral then becomes:

$$I(k) \approx e^{ikg(c)} \int_a^b e^{ik\frac{g''(c)}{2}(x-c)^2} dx \quad (36.3.4)$$

$$\approx e^{ikg(c)} \int_{\infty}^{\infty} \exp\left(ik\frac{g''(c)}{2}(x-c)^2\right) dx \quad (36.3.5)$$

$$= e^{ikg(c)} \sqrt{\frac{2\pi i}{kg''(c)}} \quad (36.3.6)$$

We can use this for our double integral, and make it a single integral. The first and second integrals of  $\psi$  are nasty, however.

$$\frac{\partial \psi}{\partial \phi} = kD \left[ \frac{2D\rho \cos(B) \sin(\phi) + 2\rho\rho_0 \sin(\phi - \phi_0)}{2D^2 \sqrt{1 + 2\cos(B) \frac{\rho_0 \cos(\phi_0) - \rho \cos(\phi)}{D} + \frac{\rho^2 + \rho_0^2 - 2\rho\rho_0 \cos(\phi - \phi_0)}{D^2}}} - \frac{\rho \cos(B) \sin(\phi)}{D} \right] \quad (36.3.7)$$

The second derivative is equally bad. Solving for  $\frac{\partial \psi}{\partial \phi} = 0$  must be done iteratively by successive approximations. A further approximation can be made as  $\psi$  is analytic in  $\phi$ . Let  $\phi_s$  be the solution to  $\frac{\partial \psi}{\partial \phi}$  and let  $\phi_{s_n}$  be a sequence such that  $\phi_{s_n} \rightarrow \phi_s$ .

### 36.3.1 Root Finding

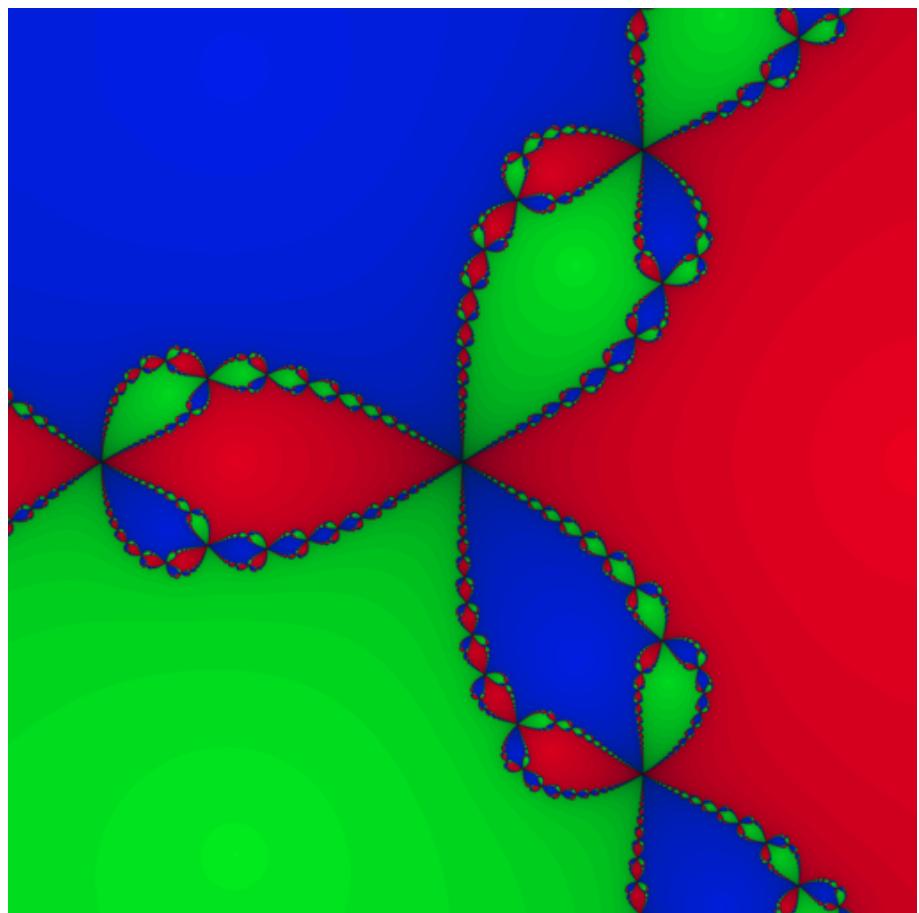


Fig. 36.1: Newton Fractal for  $z^3 - 1 = 0$

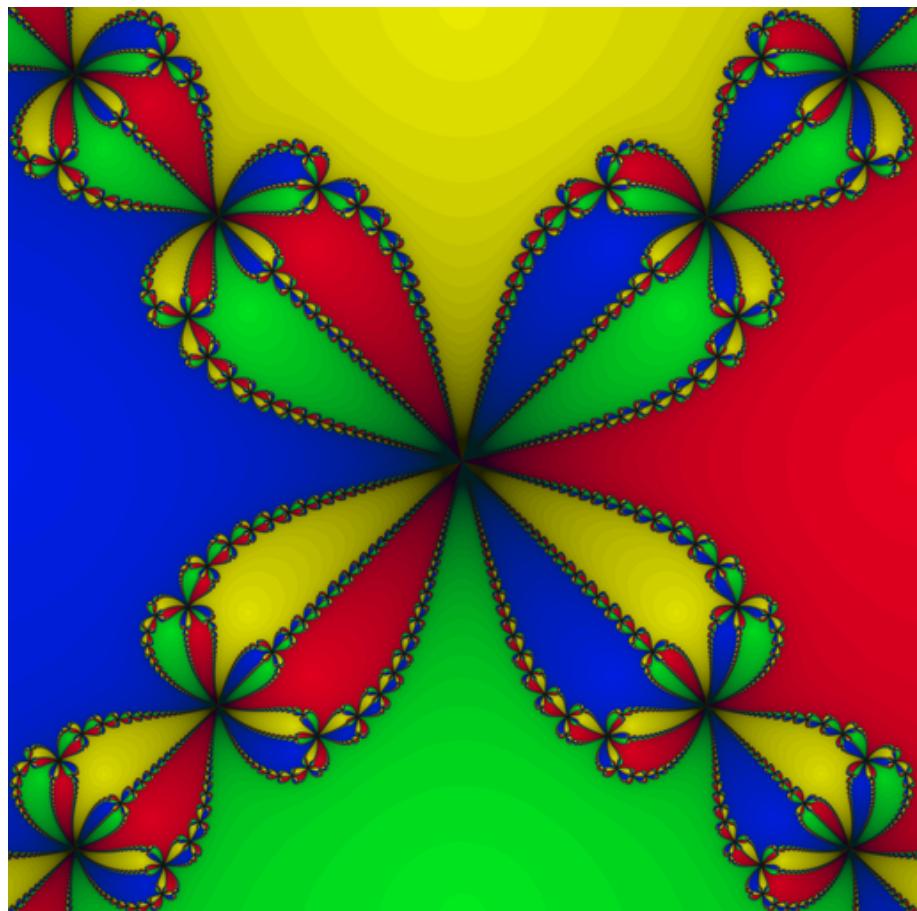


Fig. 36.2: Newton Fractal for  $z^4 - 1 = 0$

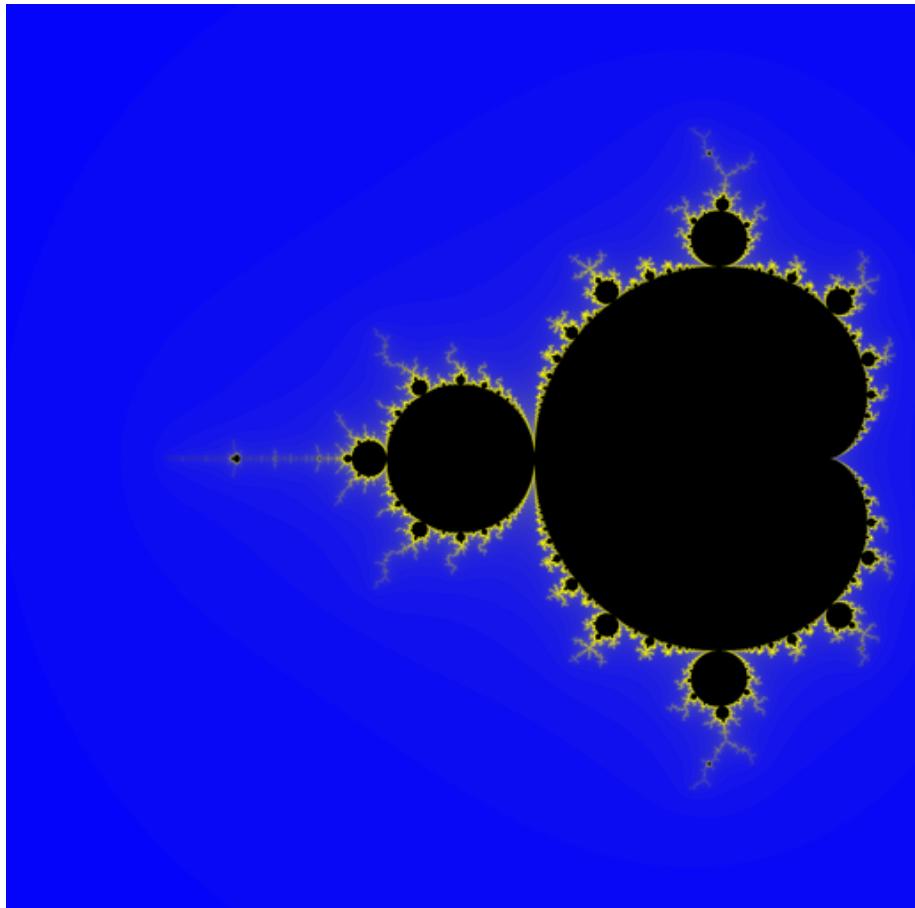


Fig. 36.3: Mandelbrot Set

## 36.4 Special Functions

There are many special functions that arise in diffraction theory. These are functions that can not be written in closed form via a combination of rational functions, trigonometric functions, logarithms, or exponentials. Usually such functions are defined as the solution to a particular differential equation, such as Bessel functions, or as the result of integrating a non-trivial function, such as Fresnel Integrals. Other times functions are defined as the inverse of a tricky algebraic equation, such that the Lambert  $W$  function. We'll discuss these three functions, numerical calculations, and their applications.

### 36.4.1 The Fresnel Integrals

The Fresnel Sine and Cosine Integrals, which are usually denoted  $S(x)$  and  $C(x)$ , respectively, occur naturally in the study of diffraction theory. By examining the *Fresnel Kernel* and using a Taylor series approximation, one comes across the following integral:

$$F(x) = \int_0^x \exp(it^2) dt \quad (36.4.1)$$

Using Euler's Theorem, we can write:

$$F(x) = \int_0^x \cos(t^2) dt + i \int_0^x \sin(t^2) dt \quad (36.4.2)$$

The Fresnel Cosine and Sine Integrals are defined as the real and imaginary parts of this equation, respectively.

#### **Definition 36.4.1: Fresnel Integrals**

he Fresnel Sine and Fresnel Cosine, denoted  $S(x)$  and  $C(x)$ , respectively, are real valued functions defined by:

$$S(x) = \int_0^x \sin(t^2) dt \quad (36.4.3a) \qquad C(x) = \int_0^x \cos(t^2) dt \quad (36.4.3b)$$

■

Graph of the Fresnel Sine and Fresnel Cosine functions are shown in Fig. 36.4.

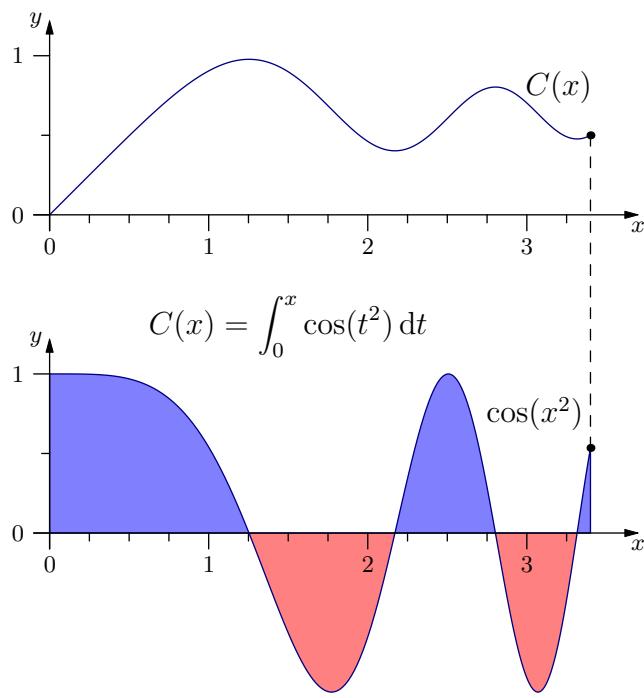


Fig. 36.4: Graph of the Fresnel Cosine Function

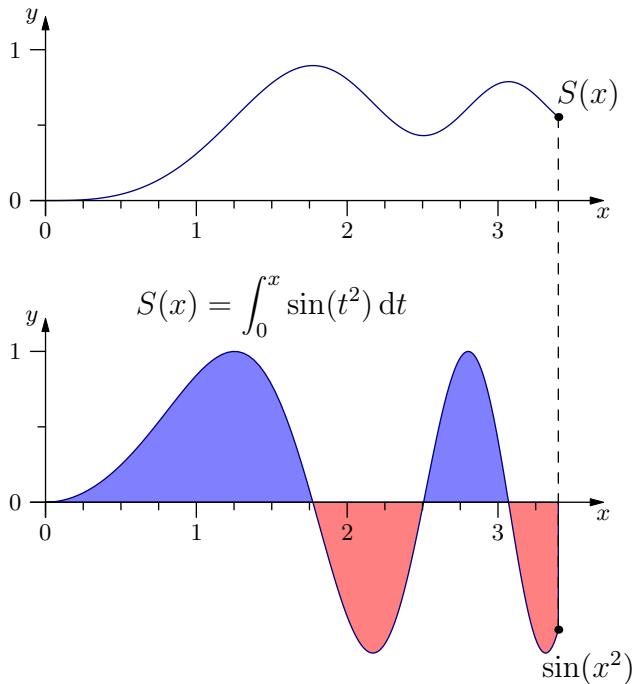


Fig. 36.5: Graph of the Fresnel Sine Function

We will be interested in functions of the form  $\exp(i\psi)$  later on. The Fresnel Approximation uses the Taylor expansion of  $\psi$  up to the quadratic term, and hence we will see something of the form  $\exp(i(a+bx+cx^2))$ . From elementary algebra we can complete the square, and do a change of variables to obtain  $\exp(i(u^2 - d^2))$ , where  $d$  is some constant. We are interested in the integral of this across the entire real line. From Euler's Formula (Thm. 30.1.18) we see that  $\exp(ix^2) = \cos(x^2) + i\sin(x^2)$ . But  $x^2$  grows rapidly, and thus  $\sin(x^2)$  and  $\cos(x^2)$  are two rapidly oscillating functions. The oscillations are so rapid that the areas cancel out, and hence  $S(x)$  and  $C(x)$  are well defined as  $x \rightarrow \infty$ . We will use Cauchy's Integral Theorem to evaluate the limits of these two functions. First, a result from Gauss.

### Theorem 36.4.1.

$$\int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi} \quad (36.4.4)$$

*Proof.* Convergence can be shown, since for all  $x \in \mathbb{R}$ :

$$0 < \exp(-x^2) \leq \frac{1}{1+x^2} \quad (36.4.5)$$

And therefore:

$$0 \leq \int_{-\infty}^{\infty} \exp(-x^2) dx \leq \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \tan^{-1}(x) \Big|_{-\infty}^{\infty} = \pi \quad (36.4.6)$$

Define the following:

$$\mathcal{I} = \int_{-\infty}^{\infty} \exp(-x^2) dx \quad (36.4.7)$$

Squaring  $\mathcal{I}$ , we obtain:

$$\mathcal{I}^2 = \left( \int_{-\infty}^{\infty} \exp(-x^2) dx \right) \left( \int_{-\infty}^{\infty} \exp(-y^2) dy \right) \quad (36.4.8a)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(-(x^2 + y^2)) dx dy \quad (36.4.8b)$$

Switching from Cartesian to Polar coordinates, we have:

$$\mathcal{I}^2 = \int_0^{2\pi} \int_0^{\infty} r \exp(-r^2) dr d\phi = 2\pi \int_0^{\infty} r \exp(-r^2) dr \quad (36.4.9)$$

This final integral can be computed from basic methods one would find in a Calculus textbook. Letting  $u = r^2$ , we have  $du = 2r dr$ , so the integral becomes:

$$\mathcal{I}^2 = \pi \int_0^{\infty} \exp(-u) du = \pi \quad (36.4.10)$$

Therefore  $\mathcal{I} = \pm\sqrt{\pi}$ . But  $\mathcal{I} > 0$ , and thus  $\mathcal{I} = \sqrt{\pi}$ .  $\square$

This result has many fundamental applications in probability theory and in statistics, where it is used to define the normal distribution. For us, we can use this to evaluate the limits of  $S(x)$  and  $C(x)$  as  $x \rightarrow \infty$ . First note, that since  $\exp(-x^2)$  is an even function, the integral on  $[0, \infty)$  is half of that of the integral on the entire real line. That is:

$$\int_0^{\infty} \exp(-t^2) dt = \frac{\sqrt{\pi}}{2} \quad (36.4.11)$$

We now evaluate the complex version of this.

**Theorem 36.4.2.**

$$\int_0^{\infty} \exp(ix^2) dx = \sqrt{\frac{\pi}{8}}(1+i) \quad (36.4.12)$$

*Proof.* For let  $C_R$  be the closed path in the complex plane defined by:

$$C_R(t) = \begin{cases} 3Rt, & 0 \leq t \leq \frac{1}{3} \\ R \exp\left(i\frac{3\pi}{4}(t - \frac{1}{3})\right), & \frac{1}{3} < t < \frac{2}{3} \\ \frac{3R}{\sqrt{2}}(1+i)(1-t), & \frac{2}{3} \leq t \leq 1 \end{cases} \quad (36.4.13)$$

Then, for all  $R > 0$ ,  $C_R$  is a Jordan Curve in the complex differentiable at all but three points. Thus, by Cauchy's Theorem, as  $\exp(iz^2)$  is an entire function:

$$\oint_{C_R} \exp(iz^2) dz = 0 \quad (36.4.14)$$

But then:

$$\begin{aligned} \int_0^R \exp(ix^2) dx + \int_{\frac{1}{3}}^{\frac{2}{3}} \exp(iz(t)^2) C'_R(t) dt \\ + \frac{1+i}{\sqrt{2}} \int_R^0 \exp(-x^2) dx = 0 \end{aligned} \quad (36.4.15)$$

But by Jordan's Lemma, this second integral tends to zero as  $R \rightarrow \infty$ . Therefore:

$$\int_0^\infty \exp(-ix^2) dx = -\frac{1+i}{\sqrt{2}} \int_\infty^0 \exp(-x^2) dx \quad (36.4.16)$$

$$= \frac{1+i}{\sqrt{2}} \int_0^\infty \exp(-x^2) dx \quad (36.4.17)$$

$$= \frac{1+i}{\sqrt{2}} \frac{\sqrt{\pi}}{2} \quad (36.4.18)$$

$$= (1+i)\sqrt{\frac{\pi}{8}} \quad (36.4.19)$$

□

**Theorem 36.4.3.** *If  $S$  and  $C$  are the Fresnel Sine and Cosine integrals, respectively, then:*

$$\lim_{x \rightarrow \infty} S(x) = \sqrt{\frac{\pi}{8}} \quad (36.4.20)$$

$$\lim_{x \rightarrow \infty} C(x) = \sqrt{\frac{\pi}{8}} \quad (36.4.21)$$

*Proof.* For:

$$\lim_{x \rightarrow \infty} (C(x) + iS(x)) = \lim_{x \rightarrow \infty} \int_0^x \exp(ix^2) dx \quad (36.4.22)$$

$$= \sqrt{\frac{\pi}{8}}(1+i) \quad (36.4.23)$$

Comparing real and imaginary parts complete the proof. □

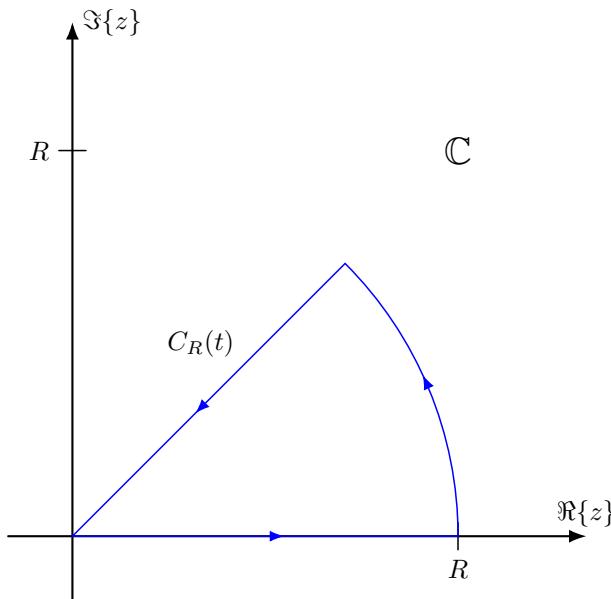


Fig. 36.6: Jordan Curve Used to Evaluate the Fresnel Integrals.

**Theorem 36.4.4.** If  $F$  is a positive real number and  $f : \mathbb{R} \rightarrow \mathbb{C}$  is defined by:

$$f(\rho) = \exp\left(i\frac{\pi}{2}\left(\frac{\rho}{F}\right)^2\right) \quad (36.4.24)$$

Then:

$$\mathcal{F}_\xi(f) = (1 + i)F \exp(-2\pi i F^2 \xi^2) \quad (36.4.25)$$

*Proof.* For:

$$\mathcal{F}_\xi(f) = \int_{-\infty}^{\infty} \exp\left(i\frac{\pi}{2}\left(\frac{\rho}{F}\right)^2\right) \exp(-2\pi i \rho \xi) d\rho \quad (36.4.26)$$

$$= \int_{-\infty}^{\infty} \exp\left(i\frac{\pi}{2}\left(\frac{\rho}{F}\right)^2 - 2\pi i \rho \xi\right) d\rho \quad (36.4.27)$$

$$= \int_{-\infty}^{\infty} \exp\left(\frac{i\pi}{2F^2} [\rho^2 - 4F^2 \rho \xi]\right) d\rho \quad (36.4.28)$$

Completing the square, we get  $(\rho - 2F^2 \xi)^2 - 4F^4 \xi^2$ . So, the integral becomes:

$$\begin{aligned} & \int_{-\infty}^{\infty} \exp\left(i\frac{\pi}{2F^2} [\rho - 2F^2 \xi]^2\right) \exp(-2\pi i F^2 \xi^2) d\rho \\ &= \exp(-2\pi i F^2 \xi^2) \int_{-\infty}^{\infty} \exp\left(i\frac{\pi}{2F^2} [\rho - 2F^2 \xi]^2\right) d\rho \end{aligned} \quad (36.4.29)$$

Let  $u = \frac{\rho - 2F^2\xi}{F}$ , so then  $F du = d\rho$ . We obtain:

$$\mathcal{F}_\xi(f) = F \exp(-2\pi i F^2 \xi^2) \int_{-\infty}^{\infty} \exp\left(i \frac{\pi}{2} s^2\right) ds \quad (36.4.30)$$

But this integral is  $1 + i$ , completing the proof.  $\square$

**Theorem 36.4.5.**  $\mathcal{F}(e^{-i\frac{\pi}{2}\left(\frac{\rho_0}{F}\right)^2}) = (1-i)Fe^{2\pi i F^2 \xi^2}$ .

*Proof.* For:

$$\mathcal{F}_\xi(f) = \int_{-\infty}^{\infty} \exp\left[-i \frac{\pi}{2} \left(\frac{\rho}{F}\right)^2\right] \exp(-2\pi i \rho \xi) d\rho \quad (36.4.31)$$

$$= \int_{-\infty}^{\infty} \exp\left(-\frac{i\pi}{2F^2} [\rho^2 + 4F^2 \rho \xi]\right) d\rho \quad (36.4.32)$$

$$= \int_{-\infty}^{\infty} \exp\left(-\frac{i\pi}{2F^2} [(\rho + 2F^2 \xi)^2 - 4F^4 \xi^2]\right) d\rho \quad (36.4.33)$$

$$= \exp(2\pi i F^2 \xi^2) \int_{-\infty}^{\infty} \exp\left(-\frac{i\pi}{2F^2} (\rho_0 + 2F^2 \xi)\right) d\rho \quad (36.4.34)$$

Let  $u = \frac{\rho_0 + 2F^2 \xi}{F}$ , then  $F du = d\rho_0$ , so we have  $Fe^{2\pi i F^2 \xi^2} \int_{-\infty}^{\infty} e^{-i\frac{\pi}{2}u^2} du$ . Let  $u = -is$ , then  $du = -ids$ , and  $u^2 = -s^2$ . So we have  $-ie^{2\pi i F^2 \xi^2} \int_{-\infty}^{\infty} e^{i\frac{\pi}{2}s^2} ds$ . But this integral is  $1+i$ . So, we have  $-iFe^{2\pi i F^2 \xi^2} (1+i) = (1-i)Fe^{2\pi i F^2 \xi^2}$ .  $\square$

**Theorem 36.4.6.** If  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $g : \mathbb{R} \rightarrow R$  are integrable, and if  $f * g$  is the convolution of  $f$  with respect to  $g$ :

$$f * g = \int_{-\infty}^{\infty} f(\tau)g(\tau-t) d\tau \quad (36.4.35)$$

Then:

$$\mathcal{F}_\xi(f * g) = \mathcal{F}_\xi(f) \cdot \mathcal{F}_\xi(g) \quad (36.4.36)$$

*Proof.* Let  $\int_{-\infty}^{\infty} |f(t)| dt = \|f\|_1$  and  $\int_{-\infty}^{\infty} |g(t)| dt = \|g\|_1$ . Then:

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(\tau)g(\tau-t)| d\tau dt &\leq \int_{-\infty}^{\infty} |f(\tau)| \int_{-\infty}^{\infty} |g(\tau-t)| d\tau dt \\ &= \int_{-\infty}^{\infty} |f(x)| \|g\|_1 dx \\ &= \|f\|_1 \|g\|_1 \end{aligned}$$

Thus,  $h(t) = f * g$  is such that  $\int_{-\infty}^{\infty} |h(t)| dt < \infty$ . Let  $H(\xi) = \mathcal{F}(h)$ . Then:

$$H(\xi) = \int_{-\infty}^{\infty} h(t) \exp(-2\pi it\xi) dt \quad (36.4.37)$$

$$= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f(\tau)g(t-\tau) d\tau \right) \exp(-2\pi it\xi) dt \quad (36.4.38)$$

But:

$$|e^{-2\pi it\xi} f(\tau)g(t-\tau)| = |f(\tau)g(t-\tau)| \quad (36.4.39)$$

But this is simply the integrand of  $h$ , and  $h$  is integrable. Thus, by Fubini's Theorem we may swap the integrals. Let  $y = t - \tau$ . Then:

$$H(\xi) = \int_{-\infty}^{\infty} f(\tau) \int_{-\infty}^{\infty} g(t-\tau) e^{-2\pi it\xi} dt d\tau \quad (36.4.40a)$$

$$= \int_{-\infty}^{\infty} f(\tau) e^{-2\pi i\tau\xi} d\tau \int_{-\infty}^{\infty} g(y) e^{-2\pi iy\xi} dy \quad (36.4.40b)$$

$$= \mathcal{F}(f) \cdot \mathcal{F}(g) \quad (36.4.40c)$$

Therefore, etc.  $\square$

**Theorem 36.4.7.** If  $T : \mathbb{R} \rightarrow \mathbb{C}$  is a Lebesgue integrable function, and if  $\hat{T} : \mathbb{R} \rightarrow \mathbb{C}$  is defined by:

$$\hat{T}(\rho_0) = \frac{1-i}{2F} \int_{-\infty}^{\infty} T(\rho) \exp\left(i\frac{\pi}{2}\left(\frac{\rho-\rho_0}{F}\right)^2\right) d\rho \quad (36.4.41)$$

then:

$$T(\rho) = \frac{1+i}{2F} \int_{-\infty}^{\infty} \hat{T}(\rho_0) \exp\left(-i\frac{\pi}{2}\left(\frac{\rho-\rho_0}{F}\right)^2\right) d\rho_0 \quad (36.4.42)$$

*Proof.* For by the definition of  $\hat{T}$ , and by the definition of convolution, we have:

$$\hat{T}(\rho) = \frac{1-i}{2F} [T * \exp\left(i\frac{\pi}{2}\left(\frac{\rho}{F}\right)^2\right)] \quad (36.4.43)$$

But  $T$  is a Lebesgue integrable function, and thus by the convolution theorem:

$$\mathcal{F}(\hat{T}) = \frac{1-i}{2F} \mathcal{F}(T) \cdot \mathcal{F}\left(\exp\left[i\frac{\pi}{2}\left(\frac{\rho_0}{F}\right)^2\right]\right) \quad (36.4.44a)$$

$$= \frac{1-i}{2F} \mathcal{F}(T) \cdot (1+i)F \left(\exp(-2\pi i F^2 \xi^2)\right) \quad (36.4.44b)$$

$$= \mathcal{F}(T) \exp(-2\pi i F^2 \xi^2) \quad (36.4.44c)$$

$$\Rightarrow \mathcal{F}(\hat{T}) \exp(2\pi i F^2 \xi^2) = \mathcal{F}(T) \quad (36.4.44d)$$

But the Fourier transform of a Gaussian is another Gaussian. That is:

$$\exp(2\pi i F^2 \xi^2) = \frac{1}{(1-i)F} \mathcal{F}\left(\exp\left[-i\frac{\pi}{2}\left(\frac{\rho_0}{F}\right)^2\right]\right) \quad (36.4.45a)$$

$$= \frac{1+i}{2F} \mathcal{F}\left(\exp\left[-i\frac{\pi}{2}\left(\frac{\rho_0}{F}\right)^2\right]\right) \quad (36.4.45b)$$

Therefore:

$$\mathcal{F}(T) = \frac{1+i}{2F} \mathcal{F}(\hat{T}) \cdot \mathcal{F}\left(e^{-i\frac{\pi}{2}\left(\frac{\rho_0}{F}\right)^2}\right) \quad (36.4.46a)$$

$$= \frac{1+i}{2F} \mathcal{F}(\hat{T} * e^{-i\frac{\pi}{2}\left(\frac{\rho_0}{F}\right)^2}) \quad (36.4.46b)$$

$$= \mathcal{F}\left(\frac{1+i}{2F} \int_{-\infty}^{\infty} \hat{T}(\rho_0) \exp\left[-i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right] d\rho_0\right) \quad (36.4.46c)$$

Therefore, by the uniqueness of the Fourier Transform:

$$T(\rho) = \frac{1+i}{2F} \int_{-\infty}^{\infty} \hat{T}(\rho_0) \exp\left[-i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right] d\rho_0 \quad (36.4.47)$$

Therefore, etc.  $\square$

**Theorem 36.4.8.** If  $T, \psi \in L^2(\mathbb{R})$ , and if  $\hat{T} : \mathbb{R} \rightarrow \mathbb{C}$  is defined by:

$$T(\rho_0) = \int_{-\infty}^{\infty} T(\rho) \exp(i\psi(\rho_0 - \rho)) d\rho_0 \quad (36.4.48)$$

then:

$$T(\rho) = \mathcal{F}_{\rho}^{-1}\left(\frac{\mathcal{F}(\hat{T})}{\mathcal{F}(\exp(i\psi))}\right) \quad (36.4.49)$$

*Proof.* For  $\hat{T}(\rho_0) = T * \exp(i\psi)$ . But then:

$$\mathcal{F}_{\xi}(\hat{T}) = \mathcal{F}_{\xi}(T * \exp(i\psi)) = \mathcal{F}_{\xi}(T) \cdot \mathcal{F}(\exp(i\psi)) \quad (36.4.50)$$

So then:

$$\mathcal{F}_{\xi}(T) = \frac{\mathcal{F}(\hat{T})}{\mathcal{F}(\exp(i\psi))} \quad (36.4.51)$$

From the uniqueness of Fourier transforms, we obtain the result.  $\square$

### 36.4.2 Bessel Functions

### 36.4.3 Lambert's W Function

### 36.4.4 Legendre Polynomials

Book Five

Geometry



**Part XXI**

# **Manifolds**



# CHAPTER 37

---

## Euclidean Spaces

---

### 37.1 Topology

#### 37.1.1 Basic Definitions

That is, topologies are closed to finite intersections, arbitrary unions, and contain both the empty set and the entire space. Elements of a topology are called the *open* subsets of  $X$ . A closed subset is the complement of an open subset. That is,  $C \subseteq X$  is closed if there exists an open set  $\mathcal{U} \in \tau$  such that  $C = X \setminus \mathcal{U}$ . [Removed chaotic and discrete topology examples, and cont funcs] These are the two extreme examples of continuous functions. From analysis one talks about continuity by means of sequences. In an ideal world these two notions would coincide, but topological spaces can be quite pathological. One of our goals is to provide any hypothesis on a space so that we can allow our familiar notions from calculus to apply. We first define sequentially continuous functions and sequential spaces. To do so needs a notion of convergence. In analysis we say that  $a_n \rightarrow x$  if for all  $\varepsilon > 0$  there is an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  with  $n > N$  it is true that  $|x - a_n| < \varepsilon$ . One can picture this by drawing an interval of radius  $\varepsilon$  about the point  $x$  and showing that eventually all of the points  $a_n$  lie within this interval. To generalize to topological space, we replace intervals with open sets.

#### Definition 37.1.1: Convergent Sequence

A convergent sequence in a topological space  $(X, \tau)$  is a sequence  $a : \mathbb{N} \rightarrow X$  such that there exists an  $x \in X$  such that for all  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  there exists an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  with  $n > N$  it is true that  $a_n \in \mathcal{U}$ . We denote this by  $a_n \rightarrow x$ .

One of the firsts things one does in analysis when discussing sequences is prove that limits are unique. In topology this is false. Consider the chaotic topology on the real numbers  $\mathbb{R}$ . Then given any sequence  $a : \mathbb{N} \rightarrow \mathbb{R}$  and any point  $x \in \mathbb{R}$  it is true that  $a_n \rightarrow x$ . We can see this since there is only one non-empty subset, and that is the entirety of  $\mathbb{R}$ . And for all  $n \in \mathbb{N}$  it is true that  $a_n \in \mathbb{R}$  by the definition of  $a$  being a sequence. Hence,  $a_n$  converges to any point  $x \in \mathbb{R}$  by the definition of convergence (Def. 37.1.1). This doesn't mean that we've stumbled upon a poor definition of convergence, but rather that the trivial topology on a set is a rather poor topological space. To rid oneself of these oddities one must require spaces to have various separation properties. The most common to impose is the Hausdorff condition, and we will discuss this after defining sequentially continuous functions.

### Definition 37.1.2: Sequentially Continuous Function

A sequentially continuous function from a topological space  $(X, \tau_X)$  to a topological space  $(Y, \tau_Y)$  is a function  $f : X \rightarrow Y$  such that for every sequence  $a : \mathbb{N} \rightarrow X$  and for every  $x \in X$  such that  $a_n \rightarrow x$ , it is true that  $f(a_n) \rightarrow f(x)$ .

We say for *every*  $x \in X$  since, as discussed before, limits may not be unique. As previously stated it is not necessarily true that sequentially continuous functions are continuous. We need to impose certain properties on the space to guarantee this. This converse, however, is true.

**Theorem 37.1.1.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, and if  $f : X \rightarrow Y$  is continuous, then it is sequentially continuous.*

*Proof.* For suppose not. Then there is a convergent sequence  $a : \mathbb{N} \rightarrow X$  and a limit  $x \in X$  of  $a$  such that  $a_n \rightarrow x$  but  $f(a_n) \not\rightarrow f(x)$  (Def. 37.1.2). But if  $f(a_n) \not\rightarrow f(x)$  then there exists an open subset  $\mathcal{V} \in \tau_Y$  such that  $f(x) \in \mathcal{V}$  and for all  $N \in \mathbb{N}$  there exists an  $n \in \mathbb{N}$  with  $n > N$  and  $f(a_n) \notin \mathcal{V}$  (Def. 37.1.1). But if  $f$  is continuous and  $\mathcal{V}$  is open, then  $f^{-1}[\mathcal{V}]$  is open (Def. ??). But since  $f(x) \in \mathcal{V}$  it is true that  $x \in f^{-1}[\mathcal{V}]$ . But  $a_n \rightarrow x$  and hence there is an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  with  $n > N$  it is true that  $a_n \in f^{-1}[\mathcal{V}]$  (Def. 37.1.1). But then for all  $n > N$  we have  $f(a_n) \in \mathcal{V}$ , a contradiction.  $\square$

It would be nice to reverse this theorem. Now we could impose something strict like requiring  $X$  to be a metric space, but should we ever encounter a non-metrizable space in the wild we'll have no tools to use. We seek a weak notion that is general enough to apply to just about every space one can imagine, but strong enough to ensure that sequentially continuous and continuous are equivalent. The concept we need is *sequential* spaces. We'll

build some point-set topology first, and then define sequential spaces. It may seem like unnecessary suspense, but sequential spaces are best defined when introducing their cousins *first countable*, *second countable*, and *separable* spaces.

### Definition 37.1.3: Interior Point

An interior point of a subset  $A \subseteq X$  of a topological space  $(X, \tau)$  is a point  $p \in A$  such that there exists an open set  $\mathcal{U} \in \tau$  such that  $\mathcal{U} \subseteq A$  and  $p \in \mathcal{U}$ .

**Example 37.1.1** If we let  $I = [0, 1]$ , considered as a subset of  $\mathbb{R}$ , then every  $x \in (0, 1)$  is an interior point of  $I$ . To see this, let  $\varepsilon > 0$  be defined by:

$$\varepsilon = \min\left\{\frac{x}{2}, \frac{1-x}{2}\right\} \quad (37.1.1)$$

since  $x \in (0, 1)$ ,  $\varepsilon > 0$ . If we define  $\mathcal{U} = (x - \varepsilon, x + \varepsilon)$ , then  $x \in \mathcal{U}$  and  $\mathcal{U} \subseteq [0, 1]$ . Moreover,  $\mathcal{U}$  is open and thus  $x$  is an interior point of  $I$ .

Fig. 37.1 shows two points  $p$  and  $q$  in a topological space  $X$  and an arbitrary subset  $A \subseteq X$ . The point  $p$  is an interior point of  $A$  since we can surround  $p$  with the blue open subset that lies entirely inside  $A$ . On the other hand,  $q$  is *not* an interior point of  $A$  since every open subset that contains  $q$  also contains points that are not in  $A$ .

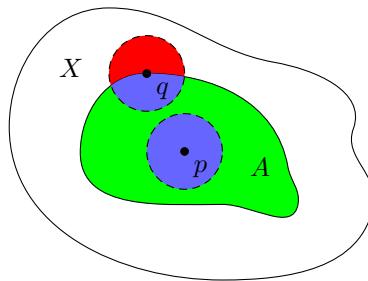


Fig. 37.1: Interior Point of a Set

**Example 37.1.2** Consider  $\mathbb{Q}$  as a subset of  $\mathbb{R}$ , where  $\mathbb{R}$  carries its usual metric topology. Then  $\mathbb{Q}$  has no interior points. For if  $q \in \mathbb{Q}$ , and if  $\mathcal{U}$  is an open subset containing  $q$ , then there is an  $\varepsilon > 0$  such that  $(q - \varepsilon, q + \varepsilon) \subseteq \mathcal{U}$ . But then  $(q - \varepsilon, q + \varepsilon) \subseteq \mathbb{Q}$ . But for all  $\varepsilon > 0$ , there is an irrational number  $r$  such that  $|q - r| < \varepsilon$ , and thus  $r \in (q - \varepsilon, q + \varepsilon)$  which is a contradiction since  $(q - \varepsilon, q + \varepsilon) \subseteq \mathbb{Q}$ . Hence,  $\mathbb{Q}$  has no interior points.

**Example 37.1.3** For the same reasons, the irrational numbers  $\mathbb{R} \setminus \mathbb{Q}$  also have no interior points.

#### Definition 37.1.4: Interior of Set

The interior of a subset  $A \subseteq X$  of a topological  $(X, \tau)$ , denoted  $\text{Int}_\tau(A)$  is the set of all interior points of  $A$ . That is:

$$\text{Int}_\tau(A) = \{ p \in X \mid p \text{ is an interior point of } A \}$$

**Example 37.1.4** The previous examples allow us to compute the interior of a few sets quickly. If  $I = [0, 1]$ , then every element of  $(0, 1)$  is an interior point. The endpoints 0 and 1 are not interior points since every  $\varepsilon$  neighborhood must contain points outside of  $[0, 1]$ . Thus,  $\text{Int}_{\mathbb{R}}(I) = (0, 1)$ . Also,  $\text{Int}_{\mathbb{R}}(\mathbb{Q}) = \emptyset$  and  $\text{Int}_{\mathbb{R}}(\mathbb{R} \setminus \mathbb{Q}) = \emptyset$ .

Borrowing from Fig. 37.1, we now present the interior of the set  $A$ ,  $\text{Int}_\tau(A)$  (Fig. 37.2).

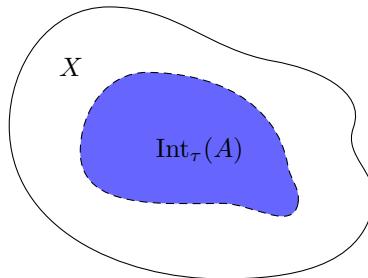


Fig. 37.2: Interior of a Set

**Theorem 37.1.2.** *If  $(X, \tau)$  is a topological space, if  $\tau_A$  is the set of all elements of  $\tau$  that are subsets of  $A$ , and if  $\text{Int}_\tau(A)$  is the interior of  $A$ , then:*

$$\text{Int}_\tau(A) = \bigcup_{\mathcal{U} \in \tau_A} \mathcal{U} \quad (37.1.2)$$

*Proof.* For if  $p \in \text{Int}_\tau(A)$ , then  $p$  is an interior point of  $A$  (Def. 37.1.4) and hence there is an open subset  $\mathcal{U} \subseteq A$  such that  $p \in \mathcal{U}$  (Def. 37.1.3). But then by hypothesis  $\mathcal{U} \in \tau_A$ , and hence  $p \in \bigcup \tau_A$ . Thus,  $\text{Int}_\tau(A) \subseteq \bigcup \tau_A$ . If  $p \in \bigcup \tau_A$ , then there is a  $\mathcal{U} \in \tau_A$  such that  $p \in \mathcal{U}$ . But by hypothesis if  $\mathcal{U} \in \tau_A$ , then  $\mathcal{U} \subseteq A$ , and hence  $p$  is an interior point of  $A$ . Thus,  $p \in \text{Int}_\tau(A)$ . Therefore,  $\text{Int}_\tau(A) = \bigcup \tau_A$ .  $\square$

Thus, an equivalent formulation of the interior of a set is the union of all open sets that are contained in  $A$ .

**Theorem 37.1.3.** *If  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then  $\text{Int}_\tau(A) \subseteq A$ .*

*Proof.* For  $\text{Int}_\tau(A) = \bigcup \tau_A$ , where  $\tau_A \subseteq A$  is the set of all element of  $\tau$  that are contained in  $A$  (Thm. 37.1.2). But then for all  $\mathcal{U} \in \tau_A$ ,  $\mathcal{U} \subseteq A$ . Thus,  $\bigcup \tau_A \subseteq A$ , and therefore  $\text{Int}_\tau(A) \subseteq A$ .  $\square$

**Theorem 37.1.4.** *If  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then  $\text{Int}_\tau(A) \in \tau$ . That is, the interior of  $A$  is open.*

*Proof.* For  $\text{Int}_\tau(A) = \bigcup \tau_A$ , where  $\tau_A \subseteq \tau$  is the set of all open subsets of  $X$  that are contained in  $A$  (Thm. 37.1.2). But then  $\text{Int}_\tau(A)$  is the union of open subsets of  $X$  and is hence open.  $\square$

**Theorem 37.1.5.** *If  $(X, \tau)$  is a topological space, and if  $\mathcal{U} \subseteq X$ , then  $\mathcal{U} \in \tau$  if and only if  $\text{Int}_\tau(\mathcal{U}) = \mathcal{U}$ .*

*Proof.* For if  $\mathcal{U} \in \tau$ , and if  $\tau_{\mathcal{U}}$  is the set of all elements of  $\tau$  that are subsets of  $\mathcal{U}$ , then  $\mathcal{U} \in \tau_{\mathcal{U}}$  since  $\mathcal{U} \subseteq \mathcal{U}$  and by hypothesis  $\mathcal{U} \in \tau$ . But  $\text{Int}_\tau(\mathcal{U}) = \bigcup \tau_{\mathcal{U}}$  (Thm. 37.1.2) and  $\mathcal{U} \subseteq \bigcup \tau_{\mathcal{U}}$ . But  $\text{Int}_\tau(\mathcal{U}) \subseteq \mathcal{U}$  (Thm. 37.1.3), and thus  $\mathcal{U} = \text{Int}_\tau(\mathcal{U})$ . In the other direction, since  $\text{Int}_\tau(\mathcal{U})$  is open (Thm. 37.1.4), if  $\mathcal{U} = \text{Int}_\tau(\mathcal{U})$ , then  $\mathcal{U}$  is open.  $\square$

**Theorem 37.1.6.** *If  $(X, \tau)$  is a topological space, if  $A, B \subseteq X$ , and if  $A \subseteq B$ , then  $\text{Int}_\tau(A) \subseteq \text{Int}_\tau(B)$ .*

*Proof.* For if  $p \in \text{Int}_\tau(A)$ , then there is an open subset  $\mathcal{U} \in \tau$  such that  $p \in \mathcal{U}$  and  $\mathcal{U} \subseteq A$  (Def. 37.1.4). But if  $\mathcal{U} \subseteq A$  and  $A \subseteq B$ , then  $\mathcal{U} \subseteq B$  and thus  $p \in \text{Int}_\tau(B)$  (Def. 37.1.4).  $\square$

**Theorem 37.1.7.** *If  $(X, \tau)$  is a topological space, if  $\mathcal{U} \in \tau$  is open, if  $A \subseteq X$ , and if  $\mathcal{U} \subseteq A$ , then  $\mathcal{U} \subseteq \text{Int}_\tau(A)$ .*

*Proof.* For if  $\mathcal{U} \subseteq A$ , then  $\text{Int}_\tau(\mathcal{U}) \subseteq \text{Int}_\tau(A)$  (Thm. 37.1.6). But  $\mathcal{U}$  is open, and therefore  $\text{Int}_\tau(\mathcal{U}) = \mathcal{U}$  (Thm. 37.1.5). Therefore  $\mathcal{U} \subseteq \text{Int}_\tau(A)$ .  $\square$

**Theorem 37.1.8.** *If  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then:*

$$\text{Int}_\tau(\text{Int}_\tau(A)) = \text{Int}_\tau(A) \tag{37.1.3}$$

*Proof.* For if  $A \subseteq X$ , then  $\text{Int}_\tau(A)$  is open (Thm. 37.1.4). But if  $\text{Int}_\tau(A)$  is open, then  $\text{Int}_\tau(\text{Int}_\tau(A)) = \text{Int}_\tau(A)$  (Thm. 37.1.5).  $\square$

**Theorem 37.1.9.** *If  $(X, \tau)$  is a topological space, then  $\text{Int}_\tau(X) = X$ .*

*Proof.* For if  $(X, \tau)$  is a topological space, then  $X \in \tau$ . But if  $X$  is open, then  $\text{Int}_\tau(X) = X$  (Thm. 37.1.5).  $\square$

**Theorem 37.1.10.** *If  $(X, \tau)$  is a topological space, and if  $A, B \subseteq X$ , then:*

$$\text{Int}_\tau(A \cap B) = \text{Int}_\tau(A) \cap \text{Int}_\tau(B) \quad (37.1.4)$$

*Proof.* For  $\text{Int}_\tau(A \cap B) = \bigcup \tau_{A \cap B}$ , where  $\tau_{A \cap B} \subseteq \tau$  is the set of all open subsets that are contained in  $A \cap B$  (Thm. 37.1.2). Similarly,  $\text{Int}_\tau(A) = \bigcup \tau_A$  and  $\text{Int}_\tau(B) = \bigcup \tau_B$ . But if  $\mathcal{U} \in \tau_{A \cap B}$ , then by definition  $\mathcal{U} \in \tau$  and  $\mathcal{U} \subseteq A \cap B$ . But  $A \cap B \subseteq A$ , and hence  $\mathcal{U} \subseteq A$ . Hence,  $\mathcal{U} \in \tau_A$  and similarly  $\mathcal{U} \in \tau_B$ . But then  $\tau_{A \cap B} \subseteq \tau_A \cap \tau_B$ , and therefore  $\bigcup \tau_{A \cap B} \subseteq \bigcup \tau_A \cap \bigcup \tau_B$  and thus  $\text{Int}_\tau(A \cap B) \subseteq \text{Int}_\tau(A) \cap \text{Int}_\tau(B)$ . But if  $x \in \text{Int}_\tau(A) \cap \text{Int}_\tau(B)$ , then  $x$  is an interior point of  $A$  and an interior point of  $B$  (Def. 37.1.4) and there exists  $\mathcal{U}_A, \mathcal{U}_B \in \tau$  such that  $x \in \mathcal{U}_A, x \in \mathcal{U}_B$ , and it is true that  $\mathcal{U}_A \subseteq A$  and  $\mathcal{U}_B \subseteq B$  (Def. 37.1.3). But then  $\mathcal{U}_A \cap \mathcal{U}_B$  is an open subset, and  $\mathcal{U}_A \cap \mathcal{U}_B \subseteq A \cap B$ . Thus,  $\mathcal{U}_A \cap \mathcal{U}_B \in \tau_{A \cap B}$ , and thus  $x \in \text{Int}_\tau(A \cap B)$ . Therefore  $\text{Int}_\tau(A \cap B) = \text{Int}_\tau(A) \cap \text{Int}_\tau(B)$ .  $\square$

The theorems proved here can completely characterize topological spaces. Kuratowski famously did this with the notion of the *closure* of a set, but interior works as well. That is, we define an *interior operator* on a set as follows:

### Definition 37.1.5: Interior Operator

An interior operator on a set  $X$  is a function  $\sigma : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$ , where  $\mathcal{P}(X)$  is the power set of  $X$ , such that:

$$\sigma(A) \subseteq A \quad (1) \quad \sigma(A \cap B) = \sigma(A) \cap \sigma(B) \quad (3)$$

$$\sigma(\sigma(A)) = \sigma(A) \quad (2) \quad \sigma(X) = X \quad (4)$$

Eqn. 3 has an equivalent formulation called *isotonicity*. We will need this later.

**Theorem 37.1.11.** *If  $X$  is a set, if  $\sigma : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  is an interior operator, if  $A, B \subseteq X$ , and if  $A \subseteq B$ , then  $\sigma(A) \subseteq \sigma(B)$ .*

*Proof.* For if  $A \subseteq B$ , then  $A \cap B = A$ , and therefore  $\sigma(A) = \sigma(A \cap B)$ . But  $\sigma(A \cap B) = \sigma(A) \cap \sigma(B)$  (Def. 37.1.5 Eqn. 3). Therefore,  $\sigma(A) \subseteq \sigma(B)$ .  $\square$

A topology gives a unique interior operator that preserves the open sets, and an interior operator gives a unique topology. We still need the notion of *morphism*, and in the ordinary study of topology we have continuous functions.

The following theorem gives an equivalent definition of continuity in terms of interior.

### Theorem 37.1.12: Kuratowski's Interior Theorem

If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, and if  $f : X \rightarrow Y$  is a function, then  $f$  is continuous if and only if for all  $B \subseteq Y$  it is true that:

$$f^{-1}[\text{Int}_Y(B)] \subseteq \text{Int}_X(f^{-1}[B])$$

*Proof.* For suppose  $f$  is continuous. But if  $B \subseteq Y$ , then  $\text{Int}_Y(B)$  is an open subset of  $Y$  (Thm. 37.1.4). But if  $\text{Int}_Y(B)$  is open and  $f$  is continuous, then  $f^{-1}[\text{Int}_Y(B)]$  is an open subset of  $X$ . Moreover,  $\text{Int}_Y(B) \subseteq B$  (Thm. 37.1.3) and therefore  $f^{-1}[\text{Int}_Y(B)] \subseteq f^{-1}[B]$ . But then  $f^{-1}[\text{Int}_Y(B)]$  is an open subset contained in  $f^{-1}[B]$ , and therefore  $f^{-1}[\text{Int}_Y(B)] \subseteq \text{Int}_X(f^{-1}[B])$  (Thm. 37.1.7). In the other direction, suppose  $f$  preserves the interior. If  $f$  is not continuous, then there is an open subset  $\mathcal{V} \in \tau_Y$  such that  $f^{-1}[\mathcal{V}]$  is not an open subset of  $X$ . But if  $\mathcal{V}$  is open, then  $\text{Int}_Y(\mathcal{V}) = \mathcal{V}$  (Thm. 37.1.5). But then by hypothesis:

$$f^{-1}[\mathcal{V}] \subseteq \text{Int}_X(f^{-1}[\mathcal{V}]) \quad (37.1.6)$$

But  $\text{Int}_X(f^{-1}[\mathcal{V}]) \subseteq f^{-1}[\mathcal{V}]$  (Thm. 37.1.3), and therefore  $f^{-1}[\mathcal{V}] = \text{Int}_X(f^{-1}[\mathcal{V}])$ . But  $\text{Int}_X(f^{-1}[\mathcal{V}])$  is open (Thm. 37.1.4), and thus  $f^{-1}[\mathcal{V}]$  is open, a contradiction. Therefore,  $f$  is continuous.  $\square$

With this, we now prove our claim that the interior operator is an equivalent formulation of topology. Def. 37.1.5 gives a complete algebraic axiomization of the notion of topologies on a set. One may be tempted to think that these four equations satisfy a *unique* interior operator (that is,  $\sigma = \text{Int}_\tau$ ), but there's a different distinct interior operator for every topology  $\tau$  on a set  $X$ . Thus, to prove any form of uniqueness we must greatly strengthen Eqn. 4 so that  $\sigma$  pertains to a single topology. We do this by requiring the  $\sigma(\mathcal{U}) = \mathcal{U}$  if and only if  $\mathcal{U}$  is open. This new requirement gives us a bijection between interior operators and topologies.

### Theorem 37.1.13: Interior Operator of a Topological Space

If  $(X, \tau)$  is a topological space, if  $\sigma : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  is an interior operator on  $X$ , and if  $\mathcal{U} \in \tau$  if and only if  $\sigma(\mathcal{U}) = \mathcal{U}$ , then  $\sigma(A) = \text{Int}_\tau(A)$ .

*Proof.* For suppose not. Then there is an  $A \in \mathcal{P}(X)$  such that  $\sigma(A) \neq \text{Int}_\tau(A)$ . But then there either  $\sigma(A) \not\subseteq \text{Int}_\tau(A)$  or  $\text{Int}_\tau(A) \not\subseteq A$ . But  $\text{Int}_\tau(A)$  is open (Thm. 37.1.4) so by hypothesis  $\sigma(\text{Int}_\tau(A)) = \text{Int}_\tau(A)$ . But  $\text{Int}_\tau(A) \subseteq A$  (Thm. 37.1.3) and therefore  $\text{Int}_\tau(A) = A \cap \text{Int}_\tau(A)$ . But  $\sigma(A \cap \text{Int}_\tau(A)) = \sigma(A) \cap \sigma(\text{Int}_\tau(A))$  (Def. 37.1.5 Eqn. 3) and  $\sigma(A) \cap \sigma(\text{Int}_\tau(A)) = \sigma(A) \cap \text{Int}_\tau(A)$ . Therefore by the transitivity of equality we obtain  $\text{Int}_\tau(A) = \sigma(A) \cap \text{Int}_\tau(A)$ , and therefore  $\text{Int}_\tau(A) \subseteq \sigma(A)$ . But  $\sigma(\sigma(A)) = \sigma(A)$  (Def. 37.1.5 Eqn. 2), and thus by hypothesis  $\sigma(A)$  is open. But since  $\sigma(A) \subseteq A$  (Def. 37.1.5 Eqn. 1), if  $\sigma(A)$  is open, then  $\sigma(A) \subseteq \text{Int}_\tau(A)$  (Thm. 37.1.7). Therefore  $\sigma(A) = \text{Int}_\tau(A)$ , a contradiction. Hence, for all  $A \subseteq X$  it is true that  $\sigma(A) = \text{Int}_\tau(A)$ .  $\square$

We now define the induced topology of an interior operator, and show that it is indeed a topology.

### Definition 37.1.6: Induced Topology of an Interior Operator

The induced topology of an interior operator  $\sigma : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  on a set  $X$  is the set of all  $\mathcal{U} \in \mathcal{P}(X)$  such that  $\sigma(\mathcal{U}) = \mathcal{U}$ . That is:

$$\tau_\sigma = \{\mathcal{U} \in \mathcal{P}(X) \mid \sigma(\mathcal{U}) = \mathcal{U}\}$$

Just because we're calling this thing a topology doesn't mean it is. We now prove that the induced topology of an interior operator is indeed a topology, and thus  $(X, \tau_\sigma)$  is a topological space.

**Theorem 37.1.14.** *If  $X$  is a set, if  $\sigma : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  is an interior operator on  $X$ , and if  $\tau_\sigma$  is the induced topology of  $\sigma$ , then  $\tau_\sigma$  is a topology on  $X$ .*

*Proof.* For if  $\sigma$  is an interior operator, then  $\sigma(X) = X$  (Def. 37.1.5 Eqn. 4) and thus  $X \in \tau_\sigma$ . Moreover, since  $\sigma(\emptyset) \subseteq \emptyset$  (Def. 37.1.5 Eqn. 1), it is therefore true that  $\emptyset \in \tau_\sigma$ . If  $\mathcal{U}, \mathcal{V} \in \tau_\sigma$ , then:

$$\begin{aligned} \sigma(\mathcal{U} \cap \mathcal{V}) &= \sigma(\mathcal{U}) \cap \sigma(\mathcal{V}) && (\text{Def. 37.1.5 Eqn. 3}) \\ &= \mathcal{U} \cap \mathcal{V} && (\text{Hypothesis}) \end{aligned}$$

and hence  $\mathcal{U} \cap \mathcal{V} \in \tau_\sigma$ . Lastly, if  $\mathcal{O} \subseteq \tau_\sigma$ , then:

$$\sigma\left(\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) \subseteq \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \quad (\text{Def. 37.1.5 Eqn. 1})$$

But for all  $\mathcal{U} \in \mathcal{O}$ ,  $\mathcal{U} \subseteq \bigcup \mathcal{O}$  and therefore  $\sigma(\mathcal{U}) \subseteq \sigma(\bigcup \mathcal{O})$  (Thm. 37.1.11). But then:

$$\bigcup_{\mathcal{U} \in \mathcal{O}} \sigma(\mathcal{U}) \subseteq \sigma\left(\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U}\right) \quad (37.1.7)$$

By hypothesis  $\sigma(\mathcal{U}) = \mathcal{U}$ , and so  $\bigcup \mathcal{O} \subseteq \sigma(\bigcup \mathcal{O})$ . But  $\sigma(\bigcup \mathcal{O}) \subseteq \bigcup \mathcal{O}$ , and therefore  $\sigma(\bigcup \mathcal{O}) = \bigcup \mathcal{O}$ . Hence,  $\bigcup \mathcal{O} \in \tau_\sigma$  and  $\tau_\sigma$  is a topology on  $X$ .  $\square$

There's a related notion called *exterior*. Like the interior, we define the exterior in terms of exterior points.

#### Definition 37.1.7: Exterior Point

An exterior point of a subset  $A \subseteq X$  of a topological space  $(X, \tau)$  is a point  $p \in X \setminus A$  such that there exists an open set  $\mathcal{U} \in \tau$  with  $\mathcal{U} \subseteq X \setminus A$  and  $p \in \mathcal{U}$ .

As one might expect from the definition, there's nothing new here and exterior points can be related to interior points by means of complement.

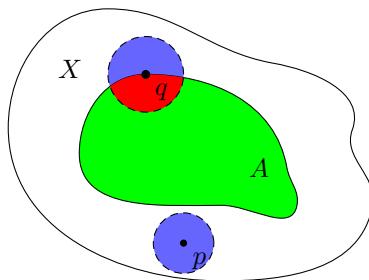


Fig. 37.3: Exterior Point of a Set

**Theorem 37.1.15.** *If  $(X, \tau)$  is a topological space, if  $A \subseteq X$ , and if  $p \in X$ , then  $p$  is an exterior point of  $A$  if and only if it is an interior point of  $X \setminus A$ .*

*Proof.* For if  $p$  is an exterior point of  $A$ , then there is an open subset  $\mathcal{U} \in \tau$  such that  $\mathcal{U} \subseteq X \setminus A$  and  $p \in \mathcal{U}$  (Def. 37.1.7). But then  $p \in X \setminus A$  and there exists an open subset  $\mathcal{U} \in \tau$  such that  $\mathcal{U} \subseteq X \setminus A$  and hence  $p$  is an interior point of  $X \setminus A$  (Def. 37.1.3). If  $p \in \text{Int}_\tau(X \setminus A)$ , then there is an open subset  $\mathcal{U} \in \tau$  such that  $p \in \mathcal{U}$  and  $\mathcal{U} \subseteq X \setminus A$ . But then  $p$  is an exterior point of  $A$  (Def. 37.1.7).  $\square$

**Example 37.1.5** Using Thm. 37.1.15 we have that every point in  $(-\infty, 0) \cup (1, \infty)$  is an exterior point of the set  $I = [0, 1]$ .

**Definition 37.1.8: Exterior of a Set**

The exterior of a subset  $A \subseteq X$  in a topological space  $(X, \tau)$ , denoted  $\text{Ext}_\tau(A)$ , is the set of all exterior points of  $A$ . That is:

$$\text{Ext}_\tau(A) = \{ p \in X \mid p \text{ is an exterior point of } A \}$$

**Theorem 37.1.16.** *If  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then:*

$$\text{Ext}_\tau(A) = \text{Int}_\tau(X \setminus A) \quad (37.1.8)$$

*Proof.* For if  $p \in \text{Ext}_\tau(A)$  if and only if  $p$  is an exterior point of  $A$  (Def. 37.1.8). But  $p$  is an exterior point of  $A$  if and only if it is an interior point of  $X \setminus A$  (Thm. 37.1.15). But  $p \in \text{Int}_\tau(X \setminus A)$  if and only if  $p \in \text{Int}_\tau(X \setminus A)$ . Therefore, we have that  $\text{Ext}_\tau(A) = \text{Int}_\tau(X \setminus A)$ .  $\square$

**Example 37.1.6** When we examined the interior of the rationals and irrationals we noted that both were empty, and hence both  $\text{Ext}_{\mathbb{R}}(\mathbb{Q})$  and  $\text{Ext}_{\mathbb{R}}(\mathbb{R} \setminus \mathbb{Q})$  are empty. That is,  $\mathbb{Q}$  has empty interior and empty exterior, even though  $\mathbb{Q}$  is dense in  $\mathbb{R}$ , which is quite paradoxical.

This example shows that the exterior of a set is not just the complement. There are three parts to a set: Interior, exterior, and boundary. This shows that, whatever boundary is defined to be, the boundary of  $\mathbb{Q}$  is the entirety of  $\mathbb{R}$ , which again may be at odds with intuition.

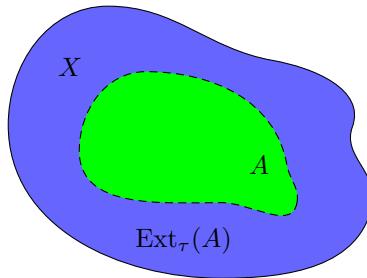


Fig. 37.4: Exterior a Set

**Theorem 37.1.17.** *If  $(X, \tau)$  is a topological space, if  $A \subseteq X$ , if  $\tau_{A^C}$  is the set of all open subsets contained in  $X \setminus A$ , then:*

$$\text{Ext}_\tau(A) = \bigcup_{U \in \tau_{A^C}} U \quad (37.1.9)$$

*Proof.* For  $\text{Ext}_\tau(A) = \text{Int}_\tau(X \setminus A)$  (Thm. 37.1.16). But  $\text{Int}_\tau(X \setminus A) = \bigcup \tau_{A^C}$  (Thm. 37.1.2) and therefore  $\text{Ext}_\tau(A) = \bigcup \tau_{A^C}$ .  $\square$

We attempt to define boundary in an intuitive manner. To do this needs the notion of *closure*, which is somewhat the *dual* concept of interior. Kuratowski's interests were in the closure axioms that could be used to define topological spaces, and indeed in his two volume treatise on topology he takes as the definition of a topological space a set  $X$  with a closure operator  $\sigma$ . Like interior, closure is defined in terms of closure points.

### Definition 37.1.9: Point of Closure

A point of closure of a subset  $A \subseteq X$  in a topological space  $(X, \tau)$  is a point  $p \in X$  such that for every open subset  $\mathcal{U} \in \tau$  such that  $p \in \mathcal{U}$  it is true that  $\mathcal{U} \cap A \neq \emptyset$ .

**Example 37.1.7** In the setting of metric spaces, points of closure of a subset  $A \subseteq X$  are points in  $X$  that can be approximated arbitrarily well by points in  $A$ . That is, the points  $x \in X$  such that for any  $\varepsilon > 0$  one can find and  $a \in A$  such that the distance between  $a$  and  $x$  is less than  $\varepsilon$ . In other words, one can find a sequence  $a : \mathbb{N} \rightarrow A$  such that  $a_n \rightarrow x$ . For example,  $\pi \in \mathbb{R}$  is a point of closure of the rationals  $\mathbb{Q}$ . We can form the sequence 3, 3.1, 3.14, 3.141, and so on. Using the actual definition of points of closure, given any open set about  $\pi$  there are rational numbers contained in this set.

### Definition 37.1.10: Closure

The closure of a subset  $A \subseteq X$  in a topological space  $(X, \tau)$  is the set:

$$\text{Cl}_\tau(X) = \{x \in X \mid x \text{ is a closure point of } A\}$$

**Example 37.1.8** Since every irrational number can be approximated arbitrarily well by rational numbers, we can see that the closure of the rationals is all of the reals. That is,  $\text{Cl}_\tau(\mathbb{Q}) = \mathbb{R}$ .

**Example 37.1.9** As a similar example the closure of the irrationals is all of  $\mathbb{R}$ .

Much like the interior of a set, many of the *dual* theorems about closures can be proved.

**Theorem 37.1.18.** *If  $(X, \tau)$  is a topological space, if  $\tau_{AC}$  is the set of all closed subsets that contain  $A$ , then:*

$$\text{Cl}_\tau(A) = \bigcap_{C \in \tau_{AC}} C \quad (37.1.10)$$

*Proof.* For suppose not. If  $x \in \text{Cl}_\tau(A)$  and if  $x \notin \bigcap \tau_{AC}$ , then there is a closed subset  $C \subseteq X$  such that  $A \subseteq C$  and  $x \notin C$ . But then  $x \in X \setminus C$  and  $X \setminus C$  is the complement of a closed set, and is thus open. But then  $X \setminus C$  is an open subset of  $X$  such that contains  $x$  and has empty intersection with  $A$ , a contradiction since  $x \in \text{Cl}_\tau(A)$  and hence  $x$  is a point of closure (Def. 37.1.10). Therefore,  $\text{Cl}_\tau(A) \subseteq \bigcap \tau_{AC}$ . Next, suppose  $x \in \bigcap \tau_{AC}$  and  $x \notin \text{Cl}_\tau(A)$ . But if  $x \notin \text{Cl}_\tau(A)$ , then  $x$  is not a point of closure (Def. 37.1.10) and hence there is an open subset  $U \in \tau$  such that  $x \in U$  and  $U \cap A = \emptyset$  (Def. 37.1.9). But then  $A \subseteq X \setminus U$  and  $x \notin X \setminus U$ . But  $X \setminus U$  is closed, and hence is an element of  $\tau_{AC}$ . But  $x \in \bigcap \tau_{AC}$ , a contradiction. Therefore, we have equality.  $\square$

**Theorem 37.1.19.** *If  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then  $A \subseteq \text{Cl}_\tau(A)$ .*

*Proof.* For  $\text{Cl}_\tau(A) = \bigcap \tau_{AC}$  (Thm. 37.1.18), where  $\tau_{AC}$  is the set of all closed subsets that contain  $A$ . But then for all  $C \in \tau_{AC}$  we have  $A \subseteq C$  and thus  $A \subseteq \bigcap \tau_{AC}$ .  $\square$

**Theorem 37.1.20.** *If  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then  $\text{Cl}_\tau(A)$  is closed.*

*Proof.* For  $\text{Cl}_\tau(A) = \bigcap \tau_{AC}$ , where  $\tau_{AC}$  is the set of all closed subsets that contain  $A$  (Thm. 37.1.18). But then  $\text{Cl}_\tau(A)$  is the intersection of closed sets, and is thus closed.  $\square$

**Theorem 37.1.21.** *If  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then  $A$  is closed if and only if  $\text{Cl}_\tau(A) = A$ .*

*Proof.* For if  $\text{Cl}_\tau(A)$  is closed (Thm. 37.1.20), and hence if  $A = \text{Cl}_\tau(A)$ , then  $A$  is closed. If  $A$  is closed, then  $A \in \tau_{AC}$  where  $\tau_{AC}$  is the set of all closed subsets of  $X$  that contain  $A$ . But  $\text{Cl}_\tau(A) = \bigcap \tau_{AC}$  (Thm. 37.1.18) and  $\bigcap \tau_{AC} \subseteq A$ , and hence  $\text{Cl}_\tau(A) \subseteq A$ . But  $A \subseteq \text{Cl}_\tau(A)$  (Thm. 37.1.19) and thus  $A = \text{Cl}_\tau(A)$ .  $\square$

**Theorem 37.1.22.** *If  $(X, \tau)$  is a topological space, if  $A, B \subseteq X$ , and if  $A \subseteq B$ , then  $\text{Cl}_\tau(A) \subseteq \text{Cl}_\tau(B)$ .*

*Proof.* For if  $x \in \text{Cl}_\tau(A)$ , then  $x$  is a point of closure of  $A$  (Def. 37.1.10). But then for all  $U \in \tau$  such that  $x \in U$ ,  $A \cap U$  is non-empty (Def. 37.1.9). But  $A \subseteq B$ , and hence  $B \cap U \neq \emptyset$ . Thus,  $x \in \text{Cl}_\tau(B)$ .  $\square$

**Theorem 37.1.23.** ??If  $(X, \tau)$  is a topological space, if  $A \subseteq X$ , if  $\mathcal{C}$  is closed, and if  $A \subseteq \mathcal{C}$ , then  $\text{Cl}_\tau(A) \subseteq \mathcal{C}$ .

*Proof.* For if  $A \subseteq \mathcal{C}$ , then  $\text{Cl}_\tau(\mathcal{C}) \subseteq \text{Cl}_\tau(A)$  (Thm. 37.1.22). But  $\mathcal{C}$  is closed, and thus  $\text{Cl}_\tau(\mathcal{C}) = \mathcal{C}$  (Thm. 37.1.21). Therefore,  $\text{Cl}_\tau(A) \subseteq \mathcal{C}$ .  $\square$

**Theorem 37.1.24.** If  $(X, \tau)$  is a topological space, and if  $A \subseteq X$ , then:

$$\text{Cl}_\tau(\text{Cl}_\tau(A)) = \text{Cl}_\tau(A) \quad (37.1.11)$$

*Proof.* For if  $A \subseteq X$ , then  $\text{Cl}_\tau(A)$  is a closed subset (Thm. 37.1.20). But if  $\text{Cl}_\tau(A)$  is closed, then  $\text{Cl}_\tau(\text{Cl}_\tau(A)) = \text{Cl}_\tau(A)$  (Thm. 37.1.21).  $\square$

**Theorem 37.1.25.** If  $(X, \tau)$  is a topological space, then  $\text{Cl}_\tau(X) = X$ .

*Proof.* For  $X$  is closed, and hence  $\text{Cl}_\tau(X) = X$  (Thm. 37.1.21).  $\square$

**Theorem 37.1.26.** If  $(X, \tau)$  is a topological space, and if  $A, B \subseteq X$ , then:

$$\text{Cl}_\tau(A \cup B) = \text{Cl}_\tau(A) \cup \text{Cl}_\tau(B) \quad (37.1.12)$$

*Proof.* For if  $x \in \text{Cl}_\tau(A \cup B)$ , then  $x$  is a point of closure of  $A \cup B$  (Def. 37.1.10) and hence for all  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$ ,  $\mathcal{U} \cap (A \cup B)$  is non-empty. But by DeMorgan's law:

$$\mathcal{U} \cap (A \cup B) = (\mathcal{U} \cap A) \cup (\mathcal{U} \cap B) \quad (37.1.13)$$

and hence either  $\mathcal{U} \cap A$  is non-empty or  $\mathcal{U} \cap B$  is non-empty. But then either  $x \in \text{Cl}_\tau(A)$  or  $x \in \text{Cl}_\tau(B)$ , and therefore  $x \in \text{Cl}_\tau(A) \cup \text{Cl}_\tau(B)$ . That is,  $\text{Cl}_\tau(A \cup B) \subseteq \text{Cl}_\tau(A) \cup \text{Cl}_\tau(B)$ . In the other direction, since  $A \subseteq A \cup B$  we have that  $\text{Cl}_\tau(A) \subseteq \text{Cl}_\tau(A \cup B)$  (Thm. 37.1.22) and similarly  $\text{Cl}_\tau(B) \subseteq \text{Cl}_\tau((A \cup B))$  and therefore  $\text{Cl}_\tau(A) \cup \text{Cl}_\tau(B) \subseteq \text{Cl}_\tau((A \cup B))$ .  $\square$

### Definition 37.1.11: Boundary of a Set

The boundary of a subset  $A \subseteq X$  in a topological space  $(X, \tau)$  is the set  $\partial A$  defined by:

$$\partial A = \text{Cl}_\tau(A) \setminus \text{Int}_\tau(A)$$

Where  $\text{Cl}_\tau(A)$  is the closure of  $A$  and  $\text{Int}_\tau(A)$  is its interior.

**Example 37.1.10** Since  $\text{Int}_\tau(\mathbb{Q}) = \emptyset$ , and since  $\text{Cl}_\tau(\mathbb{Q}) = \mathbb{R}$ , we conclude  $\partial\mathbb{Q} = \mathbb{R} \setminus \emptyset = \mathbb{R}$ . That is, the entirety of the real line is the boundary of  $\mathbb{Q}$ . In a similar manner,  $\partial(\mathbb{R} \setminus \mathbb{Q}) = \mathbb{R}$ . Both the rationals and the irrationals have the entire real line as their boundary.

**Theorem 37.1.27.** If  $(X, \tau)$  is a topological space, if  $A \subseteq X$ , then:

$$\partial A = X \setminus (\text{Int}_\tau(A) \cup \text{Ext}_\tau(A)) \quad (37.1.14)$$

### Definition 37.1.12: Isolated Point

An isolated point of a subset  $A \subseteq X$  in a topological space  $(X, \tau)$  is a point  $p \in A$  such that there exists an open set  $\mathcal{U} \in \tau$  such that  $\mathcal{U} \cap A = \{p\}$ .

### Definition 37.1.13: Limit Point

A limit point of a subset  $A \subseteq X$  in a topological space  $(X, \tau)$  is a point  $p \in X$  such that  $p \in \text{Cl}_\tau(A)$  and  $p$  is not an isolated point of  $A$ .

That is, limit points can be approximated arbitrarily well by points in  $A$ , other than the point  $p$  itself.

### Definition 37.1.14: Dense Subset

A dense subset of a topological space  $(X, \tau)$  is a subset  $A \subseteq X$  such that  $\text{Cl}_\tau(A) = X$

### Definition 37.1.15: Nowhere Dense

A nowhere dense subset of a topological space  $(X, \tau)$  is a subset  $A \subseteq X$  such that:

$$\text{Int}_\tau(\text{Cl}_\tau(A)) = \emptyset$$

**Theorem 37.1.28.** If  $(X, \tau)$  is a topological space, then  $A \subseteq X$  is nowhere dense if and only if for every non-empty open subset  $\mathcal{U} \in \tau$ ,  $A$  is not dense in the subspace topology of  $\mathcal{U}$ .

*Proof.* For if suppose not and suppose  $\mathcal{U} \in \tau$  is such that  $A$  is dense in  $\mathcal{U}$  in the subspace topology. But then:

$$\mathcal{U} = \text{Int}_\tau(\mathcal{U}) \subseteq \text{Int}_\tau(\text{Cl}_\tau(A)) \quad (37.1.15)$$

But  $A$  is nowhere dense, and hence  $\text{Int}_\tau(\text{Cl}_\tau(A)) = \emptyset$ , a contradiction. In the other direction, suppose  $\text{Int}_\tau(\text{Cl}_\tau(A)) \neq \emptyset$ . But the interior of any set is open, and hence  $A$  is dense in  $\text{Int}_\tau(\text{Cl}_\tau(A))$ , a contradiction.  $\square$

### 37.1.2 Homeomorphisms

Def continuous, def homeomorphism.

#### Definition 37.1.16: Local Homeomorphism

A local homeomorphism from a topological space  $(X, \tau_X)$  to a topological space  $(Y, \tau_Y)$  is a function  $f : X \rightarrow Y$  such that for all  $x \in X$  there exists an open set  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  and  $f|_{\mathcal{U}}$  is a continuous bijective open mapping.

Def convergent sequence.

#### Theorem 37.1.29: Kuratowski's Continuity Theorem

If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, then  $f : X \rightarrow Y$  is continuous if and only if for all  $A \subseteq X$  it is true that:

$$f[\text{Cl}_{\tau_X}(A)] \subseteq \text{Cl}_{\tau_Y}(f[A])$$

*Proof.* For suppose  $f : X \rightarrow Y$  is continuous. Since  $\text{Cl}_{\tau_Y}(f[A])$  is closed and  $f$  is continuous,  $f^{-1}[\text{Cl}_{\tau_Y}(f[A])]$  is closed. But  $f[A] \subseteq \text{Cl}_{\tau_Y}(f[A])$ , and thus  $A \subseteq f^{-1}[\text{Cl}_{\tau_Y}(f[A])]$ , and therefore:

$$\text{Cl}_{\tau_X}(A) \subseteq \text{Cl}_{\tau_X}(f^{-1}[\text{Cl}_{\tau_Y}(f[A])]) \quad (37.1.16)$$

But the closure of a closed set is the original set, and hence:

$$\text{Cl}_{\tau_X}(A) \subseteq f^{-1}[\text{Cl}_{\tau_Y}(f[A])] \quad (37.1.17)$$

But then:

$$f[\text{Cl}_{\tau_X}(A)] \subseteq f[f^{-1}[\text{Cl}_{\tau_Y}(f[A])]] \subseteq \text{Cl}_{\tau_Y}(f[A]) \quad (37.1.18)$$

In other direction, let  $C \subseteq Y$  be closed. Let  $x \in \text{Cl}_{\tau_X}(f^{-1}[C])$ . But then by hypothesis  $f(x) \in \text{Cl}_{\tau_Y}(f[f^{-1}[C]])$  and so  $f(x) \in \text{Cl}_{\tau_Y}(C)$ . But  $C$  is closed, and hence  $\text{Cl}_{\tau_Y}(C) = C$ , and therefore  $f(x) \in C$ . But then  $x \in f^{-1}[C]$ , and thus  $\text{Cl}_{\tau_X}(f^{-1}[C]) \subseteq f^{-1}[C]$ , and thus  $f^{-1}[C]$  is closed. Hence,  $f$  is continuous.  $\square$

There's a dual theorem for the interior operator.

**Theorem 37.1.30.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, then  $f : X \rightarrow Y$  is continuous if and only if for all  $B \subseteq Y$  it is true that:*

$$f^{-1}[\text{Int}_{\tau_Y}(B)] \subseteq \text{Int}_{\tau_X}(f^{-1}[B]) \quad (37.1.19)$$

Identity map is continuous, constant maps are continuous, composition is continuous.

### Definition 37.1.17: Sequentially Continuous

A sequentially continuous function from a topological space  $(X, \tau_X)$  to a topological space  $(Y, \tau_Y)$  is a function  $f : X \rightarrow Y$  such that for every convergent sequence  $a : \mathbb{N} \rightarrow X$  and for every limit  $x$  of  $a$  it is true that  $f(a_n) \rightarrow f(x)$ .

### Definition 37.1.18: Sequential Topological Space

A sequential topological space is a topological space  $(X, \tau)$  such that for every topological space  $(Y, \tau)$  and for every sequentially continuous function  $f : X \rightarrow Y$ , it is true that  $f$  is continuous.

Without sufficient separation properties (such as being Hausdorff) we don't know if the limits of sequences are unique, hence the phrasing *for every limit*.

**Theorem 37.1.31.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $f : X \rightarrow Y$  is continuous, then it is sequentially continuous.*

**Theorem 37.1.32.** *If  $(X, \tau)$  is a locally compact Lindelöf topological space, then it is  $\sigma$  compact.*

*Proof.* For every  $x \in X$  there exists compact  $K_x$  and an open  $\mathcal{U}_x$  such that  $x \in \mathcal{U}_x$  and  $\mathcal{U}_x \subseteq K_x$ . But then  $\{\mathcal{U}_x\}$  is a cover of  $X$ , and since  $(X, \tau)$  is Lindelöf there exists a countable subcover  $\mathcal{U}_n$ . Then  $\mathcal{U}_n \subseteq K_n$ , and thus  $K_n$  is a countable covering of  $X$  by compact sets, hence  $X$  is  $\sigma$  compact.  $\square$

**Theorem 37.1.33.** *Limits in Hausdorff are unique.*

*Proof.* Suppose  $a_n \rightarrow x$  and  $a_n \rightarrow y$ ,  $x \neq y$ . Then there are disjoint non-empty  $\mathcal{U}_x, \mathcal{U}_y$ . But then there exists  $N_x, N_y$  such that  $n > N_x$  implies  $a_n \in \mathcal{U}_x$  and  $n > N_y$  implies  $a_n \in \mathcal{U}_y$ . Choose  $N = \max\{N_x, N_y\}$ , a contradiction.  $\square$

**Theorem 37.1.34.** *If  $(X, \tau)$  is a Hausdorff topological space, and if  $x \in X$ , then  $\{x\}$  is a closed subset of  $X$ .*

*Proof.* For all  $y \in X$ ,  $y \neq x$ , there is a  $\mathcal{U}_y \in \tau$  such that  $y \in \mathcal{U}_y$  and  $x \notin \mathcal{U}_y$ . But then  $X \setminus \{x\} = \bigcup \mathcal{U}_y$ , and  $\{x\}$  is the complement of an open set and is therefore closed.  $\square$

**Theorem 37.1.35.** *If  $(X, \tau)$  is a Hausdorff topological space, and if  $A \subseteq X$  is finite, then  $A$  is closed.*

*Proof.* For  $A$  is the union of finitely many closed sets, and thus closed.  $\square$

### 37.1.3 Subspace Topology

Let  $X$  be a topological space with topology  $\tau$ . For any  $S \subseteq X$  let  $\tau_S$  be the subspace topology:

$$\tau_S = \{\mathcal{U} \cap S \mid \mathcal{U} \in \tau\} \quad (37.1.20)$$

Then  $\tau_S$  is a topology on  $S$ , called the subspace topology. The restriction of a continuous map  $f : X \rightarrow Y$  to  $S$  is continuous with respect to the subspace topology on  $f(S) \subseteq Y$ . If  $f|_S$  is continuous, then  $f : X \rightarrow f(X)$  is continuous with the subspace topology. Subspace of Hausdorff, first countable, second countable is still those things. Given a basis  $\mathcal{B}$  of  $X$ ,  $\mathcal{U} \cap S$  with  $\mathcal{U} \in \mathcal{B}$  forms a basis of  $S$ . Inclusion  $\iota$  is embedding. Closed in  $S$  if and only if closed in  $X$  such that  $K = S \cap C$ . Subspace topology is unique topology such that for any space  $Y$  and any function  $F : Y \rightarrow S$ ,  $F$  is continuous if and only if  $\iota \circ F : Y \rightarrow X$  is continuous. Continuity is local.

### 37.1.4 Basis for a Topology

Let  $\mathcal{B}$  be a basis for a topology on  $X$ . Then:

$$\tau(\mathcal{B}) = \{\mathcal{U} \subseteq X \mid \forall_{x \in \mathcal{U}} \exists_B \in \mathcal{B} : x \in B \subseteq \mathcal{U}\} \quad (37.1.21)$$

is a topology on  $X$ , and this is the topology generated by  $\mathcal{B}$ . If  $X_1, \dots, X_n$  are topological spaces,  $X = \prod X_k$  is the product space generated by the *open rectangles* in the  $X_i$ .

#### Definition 37.1.19: Neighborhood Basis

A neighborhood basis for a point  $x$  in a topological space  $(X, \tau)$  is a subset  $\mathcal{B}_x \subseteq \tau$  such that for all  $\mathcal{V} \in \mathcal{B}_x$  it is true that  $x \in \mathcal{V}$ , and for all  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  there exists  $\mathcal{V} \in \mathcal{B}_x$  such that  $\mathcal{V} \subseteq \mathcal{U}$ .

### 37.1.5 Continuous Maps and Products

Let  $X, Y_1, \dots, Y_n$  be topological spaces, and let  $Y = \prod_k Y_k$  be the product topological space. For each  $k$  let  $\pi_k : Y \rightarrow Y_k$  be the projection mapping sending  $\mathbf{y} \in Y$  to  $y_k$ , then  $k^{\text{th}}$  component of  $\mathbf{y}$ . Then  $f : X \rightarrow Y$  is continuous if and only if  $\pi_k \circ f$  is continuous for each  $k$ . Going the other way, if  $X = \prod_k X_k$  and  $F : X \rightarrow Y$ ,  $F = f_1 \times \dots \times f_n$ , then it is continuous if and only if  $f_k$  is continuous for all  $k$ . Let  $f : X \rightarrow Y$  be a continuous function. Note that from the definition of a function,  $f \subseteq X \times Y$ . Thus we can endow  $f$  with the subspace topology on the product topology of  $X \times Y$ . This is occasionally called the *graph* of  $f$ . If  $X$  and  $Y$  are Hausdorff and second countable, then  $X \times Y$  is, and thus any subspace of  $X \times Y$  is also Hausdorff and second countable. Hence the graph of  $f$  is a second countable Hausdorff topological space.

**Example 37.1.11** The  $n$  sphere  $S^n$  is the subset of  $\mathbb{R}^{n+1}$  such that  $\|\mathbf{x}\|_2 = 1$ , equipped with the subspace topology. We can make it into a manifold with the charts  $(\mathcal{U}_k^\pm, \phi_k^\pm)$  where  $\mathcal{U}_k^\pm$  is the  $k^{\text{th}}$  upper (or lower) hemisphere, and  $\phi_k^\pm$  is simply the projection mapping onto the hyperplane  $\mathbb{R}^n$ . We can also cover this with two simpler coordinate charts, the two stereographic projections about the north and south pole.

**Example 37.1.12** The real projective space  $\mathbb{RP}^n$  is another manifold. For any  $x, y \in \mathbb{R}^{n+1} \setminus \{0\}$ , we write  $xRy$  if there is a  $t \in \mathbb{R}$  such that  $x = ty$ . That is,  $x$  and  $y$  are equivalent if they lie on the same line through the origin. We define  $\mathbb{RP}^n$  by  $\mathbb{R}^{n+1} \setminus \{0\}/R$ , equipped with the quotient topology. That is, a subset is open if and only if  $\pi^{-1}(\mathcal{U})$  is open where  $\pi$  is the quotient map (the natural projection). For any  $k = 1, \dots, n$ , let  $\tilde{\mathcal{U}}_j$  be the  $(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \setminus \{0\}$  such that  $x_j \neq 0$ . This is an open subset and it is saturated with respect to  $\pi$ . That is, for all  $[x] \in \mathbb{RP}^n$ ,  $\pi^{-1}([x]) \cap \tilde{\mathcal{U}}_j \neq \emptyset$  if and only if  $\pi^{-1}([x]) \subseteq \tilde{\mathcal{U}}_j$ . Now let  $\mathcal{U} = \pi(\tilde{\mathcal{U}}_j)$ . Then since  $\tilde{\mathcal{U}}_j$  is saturated it is equal to  $\pi^{-1}(\mathcal{U}_j)$  and is therefore open in the quotient topology. Define  $\varphi_j$  by:

$$\varphi_j([(x_1, \dots, x_{n+1})]) = \left( \frac{x_1}{x_j}, \dots, \frac{x_{j-1}}{x_j}, \frac{x_{j+1}}{x_j}, \dots, \frac{x_{n+1}}{x_j} \right) \quad (37.1.22)$$

Then this map commutes with  $\pi$  with  $\tilde{\varphi}$ :  $\tilde{\varphi} = \varphi \circ \pi$ .

**Definition 37.1.20** A topological Lie group is a group  $(G, *)$  with a topology  $\tau$  on  $G$  such that  $\nu : G \rightarrow G$  defined by  $\nu(g) = g^{-1}$  is continuous  $* : G \times G \rightarrow G$  is continuous in the product topology, and  $(G, \tau)$  is a topological manifold.

**Theorem 37.1.36.**  $GL_n(\mathbb{R})$  is a topological Lie group of dimension  $n^2$ .

*Proof.* For we have:

$$GL_n(\mathbb{R}) = \det^{-1}(\mathbb{R} \setminus \{0\}) \quad (37.1.23)$$

and since  $\det$  is continuous (it's a polynomial), it's thus an open subset of  $\mathbb{R}^{n^2}$  (we can identify  $n \times n$  matrices with points in  $\mathbb{R}^{n^2}$ ). Moreover multiplication is continuous since it's continuous in every slot, it's a polynomial. The inverse function  $\square$

**Definition 37.1.21** A precompact subset of a topological space  $(X, \tau)$  is a subset  $\mathcal{U} \subseteq X$  such that  $\text{Cl}(\mathcal{U})$  is compact.

**Theorem 37.1.37.** *Every topological  $n$  manifold has a countable basis of precompact coordinate balls.*

*Proof.* Let  $(X, \tau)$  be a topological manifold and  $\{(\mathcal{U}_\alpha, \varphi_\alpha)\}$  a cover by coordinate charts. Since second countable spaces are Lindelof, there is a countable subcover. Since each  $\varphi_k$  is an open mapping,  $\varphi_k(\mathcal{U}_k)$  is an open subset of  $\mathbb{R}^n$ .  $\square$

**Definition 37.1.22** A locally path connected topological space is a topological space  $(X, \tau)$  such that there exists a basis of path connected open sets.

**Definition 37.1.23** A locally compact topological space is such that for all  $x \in X$  there is an open  $\mathcal{U} \subseteq X$  and a compact  $K \subseteq X$  such that  $x \in \mathcal{U}$  and  $\mathcal{U} \subseteq K$ .

**Definition 37.1.24** A connected component of  $X$  is a maximal connected subset of  $X$ .

Connected components partition the space. If two connected sets have a point in common, their union is connected.

**Definition 37.1.25** A path connected component is a maximal path component.

**Theorem 37.1.38.** *If  $(X, \tau)$  is a topological manifold, then  $M$  is locally path connected, locally compact, and the path components and connected components are identical. Moreover,  $M$  has countably many connected components, each of which is open.*

*Proof.* Locally path connected and connected implies path connected, hence connected components and path connected components are the same. Locally compact since compact coordinate balls about each point suffice. Countably many connected components since second countable.  $\square$

**Definition 37.1.26** A locally finite subset  $\mathcal{O} \subseteq \mathcal{P}(X)$  is a collection such that for all  $x \in X$  there is a neighborhood  $\mathcal{U} \in \tau$  such that  $\mathcal{U}$  intersects at most finitely many elements of  $\mathcal{O}$ .

**Definition 37.1.27** A refininement of a collection  $\mathcal{O} \subseteq \mathcal{P}(X)$  is a collection  $\mathcal{D} \subseteq \mathcal{P}(X)$  such that for all  $\mathcal{V} \in \mathcal{D}$  there is a  $\mathcal{U} \in \mathcal{O}$  such that  $\mathcal{V} \subseteq \mathcal{U}$ .

**Definition 37.1.28** A paracompact space is a space such that every open cover has a locally finite refinement.

**Theorem 37.1.39.** *If  $\mathcal{O}$  is a locally finite collection of subsets of  $X$ , then the collection of the closures is locally finite and the union of the closures is the closure of the unions.*

**Theorem 37.1.40.** *If  $(X, \tau)$  is a topological manifold, then it is paracompact.*

Thus far we have been looking at universes without walls or edges. For example, the torus, sphere, real projective spaces, and open subsets of  $\mathbb{R}^n$ . However, we can consider subsets of  $\mathbb{R}^2$  such as the closed unit disc. The interior is homeomorphic to all of  $\mathbb{R}^2$  but the bounding circle makes the entire space not open to any open subset since it is compact and no open subset of  $\mathbb{R}^n$  can be compact. However, the points on the bounding circle are homeomorphic to the closed half plane in  $\mathbb{R}^2$ . This gives rise to the notion of a smooth manifold with boundary. Let  $\mathbb{H}^n$  be the set of all  $x \in \mathbb{R}^n$  such that  $x_n \geq 0$ .

**Definition 37.1.29** A manifold with boundary is a second countable Hausdorff topological space such that every point  $x$  has an open neighborhood about it that is homeomorphic to an open subset of  $\mathbb{R}^n$  or a relatively open subset of  $\mathbb{H}^n$ .

**Definition 37.1.30** A boundary point is a point on the boundary.

**Theorem 37.1.41.** *If  $M$  is a topological manifold with boundary, then every point is either an interior point or a boundary point.*

**Definition 37.1.31** A closed manifold is a compact manifold without boundary.

**Theorem 37.1.42.** *If  $M$  is a topological manifold with boundary then the interior is an open subset of  $M$  is an open subset and the boundary is a closed subset.*

**Theorem 37.1.43.** *If  $M$  is a topological manifold with boundary, then  $M$  has a countable basis of precompact coordinate balls and half balls.  $M$  is locally compact.  $M$  is locally path connected.  $M$  has countably many components, each of which is open and connected. Moreover,  $\pi_1(M)$  is countable.*

## 37.1.6 Countability Properties

### Definition 37.1.32: First Countable Topological Space

A first countable topological space is a topological space  $(X, \tau)$  such that all  $x \in X$  there exists a countable neighborhood basis  $\mathcal{B}$  of  $x$ .

**Example 37.1.13** Any metric space.

**Definition 37.1.33: Second Countable Topological Space**

A second countable topological space is a topological space  $(X, \tau)$  such that there exists a countable basis  $\mathcal{B}$  for  $\tau$ .

**Theorem 37.1.44.** *Second countable implies first countable.*

**Example 37.1.14** This does not reverse, take the discrete metric on  $\mathbb{R}$ .

**Theorem 37.1.45.** *If  $(X, \tau)$  is a first countable topological space, if  $A \subseteq X$ , and if  $x \in X$ , then  $x \in \text{Cl}_\tau(A)$ , if and only if there is a sequence  $a : \mathbb{N} \rightarrow X$  such that  $a_n \rightarrow x$  and for all  $n \in \mathbb{N}$  it is true that  $a_n \in A$ .*

*Proof.* For suppose  $x \in \text{Cl}_\tau(A)$ . Since  $(X, \tau)$  is first countable, there exists a countable neighborhood basis  $\mathcal{B}$  of  $x$ . Since  $\mathcal{B}$  is countable, there exists a surjection  $B : \mathbb{N} \rightarrow \mathcal{B}$ . Let  $\mathcal{U} : \mathbb{N} \rightarrow \tau$  be defined by:

$$\mathcal{U}_n = \bigcap_{k \in \mathbb{Z}_{n+1}} B_k \quad (37.1.24)$$

Since  $x \in \text{Cl}_\tau(A)$ , for every open subset  $\mathcal{U} \in \tau$  that contains  $x$  there exists a point  $y \in A$  such that  $y \in \mathcal{U}$ . But then for all  $n \in \mathbb{N}$  the set  $A_n$  defined by:

$$A_n = \{y \in A \mid y \in \mathcal{U}_n\} \quad (37.1.25)$$

is non-empty. Thus by the axiom of choice there is a choice function  $a : \mathbb{N} \rightarrow X$  such that  $a_n \in A_n$ . From the definition of  $\mathcal{U}_n$ ,  $a_n \rightarrow x$ . The other direction is the definition of convergence and closure (no first countability needed).  $\square$

**Theorem 37.1.46.** *If  $(X, \tau)$  is a first countable topological space, if  $A \subseteq X$ , and if  $x \in X$ , then  $x \in \text{Int}_\tau(A)$  if and only if for every sequence  $a : \mathbb{N} \rightarrow X$  such that  $a_n \rightarrow x$  there exists an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $a_n \in A$ .*

*Proof.* For if  $x \in \text{Int}_\tau(A)$ , and if  $a : \mathbb{N} \rightarrow X$  is a sequence such that  $a_n \rightarrow x$ , then for every open subset  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  there exists an  $N \in \mathbb{N}$  such that  $n > N$  implies  $a_n \in \mathcal{U}$ . But  $\text{Int}_\tau(A)$  is open and  $\text{Int}_\tau(A) \subseteq A$ . Going the other way, suppose  $x \in A$  is such that for every sequence  $a : \mathbb{N} \rightarrow X$  such that  $a_n \rightarrow x$  there exists an  $N \in \mathbb{N}$  such that for all  $n > N$  it is true that  $a_n \in A$ . Suppose  $x \notin \text{Int}_\tau(A)$ . Since  $(X, \tau)$  is first countable, there is a countable neighborhood basis  $\mathcal{B}_x$  of  $x$ . Let  $B : \mathbb{N} \rightarrow \mathcal{B}$  be a surjection and let  $\mathcal{U}_n$  be defined by:

$$\mathcal{U}_n = \bigcap_{k \in \mathbb{Z}_{n+1}} B_k \quad (37.1.26)$$

But if  $xn \notin \text{Int}_\tau(A)$  then for all  $n \in \mathcal{N}$  there is a  $y \in \mathcal{U}_n$  such that  $y \notin A$ . By the axiom of choice there is a sequence  $y : \mathbb{N} \rightarrow X$  such that  $y_n \in \mathcal{U}_n$  and  $y_n \notin A$ . But then  $y_n \rightarrow x$  and  $y_n$  is never in  $A$ , a contradiction.  $\square$

**Theorem 37.1.47.** *If  $(X, \tau)$  is a first countable topological space, and if  $A \subseteq X$ , then  $A$  is open if and only if for every sequence  $a : \mathbb{N} \rightarrow X$  such that there exists an  $x \in A$  such that  $a_n \rightarrow x$ , then there is an  $N \in \mathbb{N}$  such that for all  $n > N$ ,  $a_n \in A$ .*

*Proof.* One direction is the definition of convergence in a topological space. Suppose for every sequence  $a : \mathbb{N} \rightarrow X$  such that there exists  $x \in A$  such that  $a_n \rightarrow x$ , it is true that there exists an  $N \in \mathbb{N}$  such that  $n > N$  implies  $a_n \in A$ . Then for every  $x \in A$  there is an open subset  $\mathcal{U}_x \subseteq A$  such that  $x \in \mathcal{U}_x$ . For suppose not, and let  $x \in A$  be such that there is not open neighborhood  $\mathcal{U}_x \in \tau$  such that  $x \in \mathcal{U}_x$  and  $\mathcal{U}_x \subseteq A$ . But  $(X, \tau)$  is first countable and thus there is a countable neighborhood basis  $\mathcal{B}_x$  about  $x$ . Let  $B : \mathbb{N} \rightarrow \mathcal{B}$  be a surjection and let  $\mathcal{U}_n$  be the intersection of  $B_0, \dots, B_n$ . Since none of these are contained in  $A$ , by hypothesis, there exists a  $y_n \in \mathcal{U}_n$  such that  $y_n \notin A$ . But then  $y_n \rightarrow x$ , a contradiction.  $\square$

**Theorem 37.1.48.** *If  $(X, \tau)$  is a first countable topological space, and if  $A \subseteq X$ , then  $A$  is closed if and only for every convergent sequence  $a : \mathbb{N} \rightarrow X$  with limit  $x \in X$  such that for all  $n \in \mathbb{N}$  it is true that  $a_n \in A$ , then  $x \in A$ .*

*Proof.* For if  $A$  is closed, and if  $a : \mathbb{N} \rightarrow A$  is such that  $a_n \rightarrow x$ , suppose  $x \notin A$ . Then since  $A$  is closed,  $X \setminus A$  is open. But if  $a_n \rightarrow x$  then for all  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$ , there is and  $N \in \mathbb{N}$  such that for all  $n > N$  it is true that  $a_n \in \mathcal{U}$ . But  $X \setminus A$  is open and  $x \in X \setminus A$ , a contradiction since  $a_n \in A$  for all  $n$ . In the other direction, a set is closed if and only if it is equal to its closure, hence apply the previous theorem.  $\square$

### Definition 37.1.34: Lindelöf Topological Space

A Lindelöf topological space is a topological  $(X, \tau)$  such that for every open cover  $\mathcal{O}$  of  $X$  there exists a countable subcover  $\Delta$ .

This is a weaker version of compactness, but it has its uses.

**Theorem 37.1.49.** *If  $(X, \tau)$  is a second countable topological space, then it is Lindelöf.*

*Proof.* For let  $\mathcal{O}$  be an open cover of  $X$ . Since  $(X, \tau)$  is second countable, there exists a countable basis  $\mathcal{B}$  for  $\tau$ . Let  $B : \mathbb{N} \rightarrow \mathcal{B}$  be a surjection. For all  $\mathcal{U} \in \mathcal{O}$ , define:

$$A_{\mathcal{U}} = \{ n \in \mathbb{N} \mid B_n \subseteq \mathcal{U} \} \quad (37.1.27)$$

Since  $\mathcal{B}$  is a basis, for all non-empty  $\mathcal{U} \in \mathcal{O}$  there is an  $n \in \mathbb{N}$  such that  $B_n \subseteq \mathcal{U}$ , and hence  $A_{\mathcal{U}}$  is non-empty. There is thus an inverse mapping (axiom of choice)  $f : \mathbb{N} \rightarrow \mathcal{O}$  such that for all  $n \in \mathbb{N}$ ,  $B_n \subseteq f_n$ , and thus  $\Delta = \{f_n\}$  is a countable subcover.  $\square$

Def subspace, def relatively open, relatively closed.

### Definition 37.1.35: Topological Embedding

A topological embedding of a topological space  $(X, \tau_X)$  into a topological space  $(Y, \tau_Y)$  is a continuous injective function  $f : X \rightarrow Y$  such that  $(X, \tau)$  is homeomorphic to  $t(f[X], \tau)f[X]$ , where  $\tau_{f[X]}$  is the subspace topology in  $Y$ .

**Theorem 37.1.50.** *If  $(X, \tau_X)$  is a topological space, if  $A \subseteq X$ , if  $\tau_A$  is the subspace topology on  $A$ , if  $(Y, \tau_Y)$  is a topological space, and if  $f : Y \rightarrow A$  is a function, then  $f$  is continuous if and only if  $\iota_A \circ f : Y \rightarrow X$  is continuous, where  $\iota_A : A \rightarrow X$  is the inclusion mapping.*

**Theorem 37.1.51.** *If  $(X, \tau_X)$  is a topological space, if  $A \subseteq X$ , if  $\tau$  is a topology on  $A$  such that for every topological space  $(Y, \tau_Y)$  and for every continuous function  $f : Y \rightarrow A$  it is true that  $\iota_A \circ f$  is continuous, then  $\tau = \tau_A$  where  $\tau_A$  is the subspace topology.*

**Theorem 37.1.52.** *Closed in subspace  $A$  if and only if  $C_A = A \cap C_X$ , where  $C_X$  is closed.*

Inclusion map is top embedding, restriction of continuous is continuous, subspace of Hausdorff is Hausdorff, same for first/second countable. If  $f : X \rightarrow Y$  is continuous, then it is continuous in its restriction to all subspaces of  $X$ . What about the converse? If  $f$  is continuous in its subspaces, is it continuous?

**Theorem 37.1.53.** *If  $(X, \tau_X)$  is a topological space, if  $(Y, \tau_Y)$  is a topological space, and if  $f : X \rightarrow Y$  is such that for every point  $x \in X$  there exists an open subset  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  and  $f|_{\mathcal{U}} : \mathcal{U} \rightarrow Y$  is continuous in the subspace topology, then  $f$  is continuous.*

This theorem lets us make continuous functions by gluing together pieces of functions that are defined on open subsets which agree on the overlap.

**Theorem 37.1.54.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $\mathcal{O}$  is an open cover of  $X$ , if  $\mathcal{F} : \mathcal{O} \rightarrow \mathcal{F}(X, Y)$  is a bijection such that for all  $\mathcal{U}, \mathcal{V} \in \mathcal{O}$  it is true that:*

$$\mathcal{F}(\mathcal{U})|_{\mathcal{U} \cap \mathcal{V}} = \mathcal{F}(\mathcal{V})|_{\mathcal{U} \cap \mathcal{V}} \quad (37.1.28)$$

and such that  $\mathcal{F}(\mathcal{U})|_{\mathcal{U}} : \mathcal{U} \rightarrow Y$  is continuous in the subspace topology, then there is a unique continuous function  $f : X \rightarrow Y$  such that for all  $\mathcal{U} \in \mathcal{F}$  it is true that  $f|_{\mathcal{U}} = \mathcal{F}(\mathcal{U})|_{\mathcal{U}}$

**Theorem 37.1.55.** If  $(X, \tau)$  is a topological space, if  $\mathcal{O}$  is a countable open cover of  $X$  such that for all  $\mathcal{U} \in \mathcal{O}$  it is true that  $(\mathcal{U}, \tau_{\mathcal{U}})$  is second countable, where  $\tau_{\mathcal{U}}$  is the subspace topology, then  $(X, \tau)$  is second countable.

*Proof.* For let  $\mathcal{B}$  be the union of all of the bases of all of the  $\mathcal{U} \in \mathcal{O}$ . This is countable since it is the countable union of countable sets. Let  $\mathcal{V} \in \tau$  be an open subset of  $X$ . But then  $\mathcal{V} \cap \mathcal{U}$  is open for all  $\mathcal{U} \in \mathcal{O}$  and it is relatively open in  $\mathcal{U}$ . But  $\mathcal{B}_{\mathcal{U}}$  is a basis for  $\mathcal{U}$ , and hence there is a subset  $\Delta_{\mathcal{V}} \subseteq \mathcal{B}_{\mathcal{U}}$  such that  $\cap \mathcal{U} = \bigcup \Delta_{\mathcal{V}}$ . But then:

$$\mathcal{V} = \mathcal{V} \cap X = \mathcal{V} \cap \left( \bigcup \mathcal{U} \right) = \bigcup (\mathcal{V} \cap \mathcal{U}) = \bigcup \left( \bigcup \Delta_{\mathcal{V}} \right) \quad (37.1.29)$$

which is the union of elements of  $\mathcal{B}$ , and hence this is a basis for  $\tau$ .  $\square$

Product topology, projection map, continuous iff components are continuous. The product topology is unique topology with this property. Projections are continuous. Product maps. Product of continuous maps is continuous. Subspace topology is the same as product topology of subspaces. That is,  $A_1 \times A_2$ , doesn't matter if we consider subspace topology of product or product topology of subspaces. Open rectangles form basis in finite products.

Def disjoint union, canonical injective map.

### Definition 37.1.36: Disjoint Union Topology

The disjoint union topology generated by the disjoint union of a set of topological spaces  $\mathcal{T}$  is subset  $\tau_T \subseteq \mathcal{P}(\coprod \mathcal{T})$  defined by declaring  $\mathcal{U}$  open if and only if for all  $\alpha$ ,  $\pi_{\alpha}(\mathcal{U})$  is open in  $X_{\alpha}$ .

Function from disjoint union to  $Y$  is continuous if and only if  $F \circ \iota_{\alpha}$  is continuous for all  $\alpha$  where  $\iota_{\alpha}$  is the canonical inclusion map. Disjoint union topology is unique topology with this trait. Canonical inclusions are embeddings. Disjoint union of Hausdorff is Hausdorff, similarly for first countable. Disjoint unions of countably many second countable spaces is second countable.

**Definition 37.1.37: Quotient Map**

A quotient map from a topological space  $(X, \tau_X)$  to a topological space  $(Y, \tau_Y)$  is a surjective continuous function  $q : X \rightarrow Y$  and such that for all  $\mathcal{V} \subseteq Y$  such that  $q^{-1}[\mathcal{V}] \in \tau_X$ , it is true that  $\mathcal{V} \in \tau_Y$

That is, the quotient map  $q$  is continuous, surjective, and uniquely defines the topology on  $Y$ . A common confusion is that a quotient map is not necessarily an open mapping. The reverse direction is a true statement.

**Theorem 37.1.56.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, and if  $f : X \rightarrow Y$  is a continuous surjective open map, then  $f$  is a quotient map.*

*Proof.* For if  $\mathcal{V} \in \tau_Y$ , then since  $f$  is continuous it is true that  $f^{-1}[\mathcal{V}] \in \tau_X$ . Moreover, if  $\mathcal{V} \subseteq Y$  is such that  $f^{-1}[\mathcal{V}] \in \tau_X$ , then since  $f$  is surjective it is true that  $f[f^{-1}[\mathcal{V}]] = \mathcal{V}$ . But by hypothesis  $f$  is an open mapping, and  $f^{-1}[\mathcal{V}]$  is an open subset of  $X$ , and therefore  $\mathcal{V}$  is an open subset of  $Y$ . Thus, for all  $\mathcal{V} \subseteq Y$  it is true that  $\mathcal{V}$  is open if and only if  $f^{-1}[\mathcal{V}]$  is open in  $X$ . Thus,  $f$  is a quotient map.  $\square$

**Definition 37.1.38: Quotient Topology**

The quotient topology on a set  $Y$  induced by a topological space  $(X, \tau_X)$  under a surjective function  $q : X \rightarrow Y$  is the set:

$$\tau_Y = \{ \mathcal{V} \in \mathcal{P}(Y) \mid q^{-1}[\mathcal{V}] \in \tau_X \}$$

The most common way of constructing quotient spaces is by means of an equivalence relation.

**Definition 37.1.39: Quotient Space of Equivalence Relation**

The quotient space induced by an equivalence relation  $R$  on a set  $X$  with respect to a topology  $\tau$  on  $X$  is the topological space  $(X/R, \tau_R)$  where  $X/R$  is the quotient set of  $X$  under  $R$  and  $\tau_R$  is the quotient topology induced by the natural projection mapping  $q : X \rightarrow X/R$  defined by  $q(x) = [x]$ .

**Definition 37.1.40: Adjunction Space**

The adjunction space of a topological space  $(X, \tau_X)$  with respect to a topological space  $(Y, \tau_Y)$  under a continuous function  $f : A \rightarrow X$  from a closed subset  $A \subseteq Y$  into  $X$  is the quotient space  $(X \sqcup_f Y, \tau_f)$  formed by the equivalence relation  $R$  on  $X \coprod Y$  generated by the relation:

$$\tilde{R} = \left\{ ((x, 0), (y, 1)) \in (X \coprod Y)^2 \mid y \in A \text{ and } x = f(y), 0 \right\}$$

Pictorially, we have a copy of  $X$  and a copy of  $Y$  in some ambient space, and we glue  $A \subseteq Y$  along  $f[A] \subseteq X$ .

**Definition 37.1.41: Saturated Subset**

A saturated subset of a set  $X$  with respect to a set  $Y$  under a function  $f : X \rightarrow Y$  is a subset  $A \subseteq X$  such that:

$$f^{-1}[f[A]] = A$$

**Theorem 37.1.57.** If  $X$  and  $Y$  are sets, if  $f : A \rightarrow Y$  is a function, and if  $A \subseteq X$ , then  $A$  is saturated in  $X$  if and only if there is a subset  $B \subseteq Y$  such that:

$$A = \bigcup_{y \in B} f^{-1}[\{y\}] \quad (37.1.30)$$

That is,  $A$  is the union of fibers.

*Proof.* For if  $A$  is the union of the fibers of  $B$ , then  $f[A] = B$ . But then  $A = f^{-1}[B] = f^{-1}[f[A]]$ , and hence  $A$  is saturated. If  $A$  is saturated, let  $B = f[A]$  and let  $y \in B$ . Then  $f^{-1}[\{y\}] \subseteq A$  since  $A$  is saturated. Thus,  $A$  is the union of fibers.  $\square$

**Theorem 37.1.58.** If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $q : X \rightarrow Y$  is a quotient map, and if  $(Z, \tau_Z)$  is a topological space, then for any continuous function  $f : Y \rightarrow Z$ ,  $f \circ q : X \rightarrow Z$  is continuous.

The quotient topology is the unique topology with this property. Closed if and only if  $q^{-1}[\mathcal{V}]$  is closed. Injective quotient map is a homeomorphism. The restriction of quotient map to saturated subset is a quotient map onto its image. Composition of quotient maps is a quotient map.

**Theorem 37.1.59.** If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $q : X \rightarrow Y$  is a continuous surjective function, then  $q$  is a quotient map if and only if for every open saturated subset  $\mathcal{U} \subseteq X$ ,  $q[\mathcal{U}]$  is open in  $Y$ .

*Proof.* For if  $q$  is a quotient map, then  $\mathcal{V} \subseteq Y$  is open if and only if  $q^{-1}[\mathcal{V}] \in \tau_X$ . But if  $\mathcal{U} \in \tau_X$  is open and saturated, then  $q^{-1}[q[\mathcal{U}]] = \mathcal{U}$ . But then  $q[\mathcal{U}]$  is such that it's pre-image under  $q$  is open, and thus  $q[\mathcal{U}]$  is open. On the other hand, suppose  $q$  maps open saturated sets to open sets and let  $\mathcal{V} \subseteq Y$  be such that  $q^{-1}[\mathcal{V}]$  is open. Let  $\mathcal{U} = q^{-1}[\mathcal{V}]$ . But then  $\mathcal{U}$  is the union of fibers and is hence saturated. But then  $\mathcal{U}$  is open and saturated, and thus  $q[\mathcal{U}]$  is open. But  $q[\mathcal{U}] = \mathcal{V}$ , and thus  $\mathcal{V}$  is open. Thus,  $q$  is a quotient map.  $\square$

Same theorem for closed subsets.

**Theorem 37.1.60.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $q : X \rightarrow Y$  is a quotient map, if  $(Z, \tau_Z)$  is a topological space, and if  $f : X \rightarrow Z$  is a continuous function such that for all  $x_1, x_2 \in X$  such that  $q(x_1) = q(x_2)$  it is true that  $f(x_1) = f(x_2)$ , then there is a unique continuous function  $\tilde{f} : Y \rightarrow Z$  such that  $f = \tilde{f} \circ q$ .*

*Proof.* For since  $q$  is a quotient map, it is surjective. Thus for all  $y \in Y$ ,  $q^{-1}[\{y\}]$  is non-empty. Thus, by the axiom of choice, there is a function  $p : Y \rightarrow X$  such that  $p(y) \in q^{-1}[\{y\}]$  for all  $y \in Y$ . That is,  $p(y)$  is a representative for the equivalence class  $[p(y)]$  of points in  $X$  that map to  $y$ . Let  $\tilde{f} : Y \rightarrow Z$  be defined by  $\tilde{f} = f \circ p$ . But by hypothesis for all  $x_1, x_2 \in q^{-1}[\{y\}]$  it is true that  $f(x_1) = f(x_2)$ . But then:

$$f(x) = f(p(y)) = \tilde{f}(y) \quad (37.1.31)$$

and thus  $f = \tilde{f} \circ q$ . But  $f$  is continuous, and thus by the characteristic property of quotient spaces,  $\tilde{f}$  is continuous.  $\square$

**Theorem 37.1.61.** *If  $(X, \tau_X)$ ,  $(Y, \tau_Y)$ , and  $(Z, \tau_Z)$  are topological spaces, if  $q_Y : X \rightarrow Y$  and  $q_Z : X \rightarrow Z$  are quotient maps, and if for all  $x_1, x_2 \in X$  it is true that  $q_Y(x_1) = q_Y(x_2)$  if and only if  $q_Z(x_1) = q_Z(x_2)$ , then there is a unique homeomorphism  $\varphi : Y \rightarrow Z$  such that  $\varphi \circ q_Y = q_Z$ .*

Projection maps are open mappings. Canonical injective mappings into disjoint union topology are closed and open mappings. Local homeo is open map. Bijective local homeo is homeo.

**Theorem 37.1.62.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $q : X \rightarrow Y$  is a quotient map, and if  $R \subseteq X \times X$  defined by:*

$$R = \{(x_1, x_2) \in X \times X \mid q(x_1) = q(x_2)\} \quad (37.1.32)$$

*is a closed subset of  $X \times X$  in the product topology, then  $Y$  is Hausdorff.*

**Theorem 37.1.63.**  *$f : X \rightarrow Y$  is an open mapping if and only if for all  $B \subseteq Y$  it is true that:*

$$\text{Int}_{\tau_X}(f^{-1}[B]) \subseteq f^{-1}[\text{Int}_{\tau_Y}(B)] \quad (37.1.33)$$

Similarly for closed mappings.

**Theorem 37.1.64.** *If  $q : X \rightarrow Y$  is a continuous surjective open function, then it is a quotient map.*

**Theorem 37.1.65.** *If  $q : X \rightarrow Y$  is a continuous injective open function, then it is a topological embedding.*

**Theorem 37.1.66.** *If  $f : X \rightarrow Y$  is a continuous bijective open function, then it is a homeomorphism.*

Def connected space, connected component, continuous image of connected is connected.

**Theorem 37.1.67.** *If  $(X, \tau)$  is a topological space, if  $A \subseteq X$  is a connected subset, then there is a unique connected component  $B$  of  $X$  such that  $A \subseteq B$ .*

*Proof.* For let  $(C, \subseteq)$  be the partial ordered set of connected subsets of  $X$  ordered by inclusion and let  $\mathcal{U}_J$  be a chain. Suppose  $\bigcup \mathcal{U}_j$  is disconnected. Then there are disjoint non-empty open sets  $\mathcal{V}_1, \mathcal{V}_2$  such that  $\mathcal{V}_1 \cup \mathcal{V}_2 = \bigcup \mathcal{U}_J$ , and such that  $\mathcal{V}_i \cap \bigcup \mathcal{U}_j \neq \emptyset$ . But then there is an  $x_1 \in \mathcal{V}_1$  such that  $x_1 \in \bigcup \mathcal{U}_J$  and an  $x_2 \in \mathcal{V}_2$  such that  $x_2 \in \bigcup \mathcal{U}_J$ . But if  $x_1 \in \mathcal{U}_J$  there is some  $\alpha \in J$  such that  $x_1 \in \mathcal{U}_\alpha$ , and similarly  $x_2 \in \mathcal{U}_\beta$ . But since the  $\mathcal{U}_J$  form a chain, either  $\mathcal{U}_\alpha \subseteq \mathcal{U}_\beta$  or  $\mathcal{U}_\beta \subseteq \mathcal{U}_\alpha$ . But then  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are not disjoint, a contradiction, and hence  $\bigcup \mathcal{U}_J$  is connected. Thus by Zorn's lemma, there is a maximal element  $B \in C$  that is comparable to  $A$ . That is,  $B$  is maximal and  $A \subseteq B$ . If  $B'$  is a different maximal element, and  $B \cup B'$  would be connected and strictly larger, contradicting maximality.  $\square$

**Theorem 37.1.68.** *The connected components of  $X$  are closed non-empty disjoint sets that partition  $X$ .*

**Theorem 37.1.69.** *Product of connected is connected (in product topology).*

The product of connected need not be connected in the box topology.

**Theorem 37.1.70.** *The quotient of connected is connected.*

### Definition 37.1.42: Locally Path Connected

A locally path connected topological space is a topological space  $(X, \tau)$  such that there exists a basis  $\mathcal{B}$  of  $\tau$  such that for all  $\mathcal{U} \in \mathcal{B}$  it is true that  $\mathcal{U}$  is path connected.

**Example 37.1.15** Paradoxically, path connected does not imply locally path connected. The Warsaw Circle, which is the topologist's sine curve wrapped around with an arc to form a path connected space, is not locally path connected.

**Theorem 37.1.71.** *If  $(X, \tau)$  is a locally path connected topological space, and if  $A \subseteq X$  is a connected component of  $X$ , then  $A$  is open.*

*Proof.* For if  $(X, \tau)$  is locally path connected, there is a basis  $\mathcal{B}$  of  $\tau$  of connected open sets. But then for all  $x \in A$ , there is an open set  $\mathcal{U}_x \in \mathcal{B}$  such that  $x \in \mathcal{U}_x$ . But by hypothesis  $\mathcal{U}_x$  is open and  $A$  is a connected component, and hence  $\mathcal{U}_x \subseteq A$ . But then  $A = \bigcup \mathcal{U}_x$ , which is open, and thus  $A$  is open.  $\square$

**Theorem 37.1.72.** *If  $(X, \tau)$  is a locally path connected topological space, and if  $A \subseteq X$  is a connected component, then it is a path connected component.*

**Theorem 37.1.73.** *If  $(X, \tau)$  is a locally path connected topological space, and if  $(X, \tau)$  is connected, then it is path connected.*

Def compact, cont image of compact is compact, EVT, finite union of compact is compact, disjoint compact subsets of a Hausdorff space can be separated by disjoint open sets. Closed subset of compact is compact. Compact Hausdorff is closed. Compact in metric if and only if complete and totally bounded. Product of compact is compact. Quotient of compact is compact. Def uni cont, Lipschitz cont, locally Lipschitz cont. Lipschitz cont  $\rightarrow$  uni cont  $\rightarrow$  cont (metric space). Cont on compact is uni cont. Locally Lipschitz cont on compact is Lipschitz cont. Implications don't reverse: Try  $\sqrt{x}$  and  $x^2$ .

**Theorem 37.1.74.** *If  $(X, \tau)$  is first countable and countably compact, then it is sequentially compact.*

#### Definition 37.1.43: Point Finite Cover

A point finite cover of a topological space  $(X, \tau)$  is a cover  $\mathcal{O}$  of  $X$  such that for all  $x \in X$  the set  $\mathcal{O}_x$  defined by:

$$\mathcal{O}_x = \{\mathcal{U} \in \mathcal{O} \mid x \in \mathcal{U}\}$$

is finite.

**Definition 37.1.44: Metacompact**

A metacompact topological space is a topological space  $(X, \tau)$  such that for every open cover  $\mathcal{O}$  of  $X$  there exists a point finite refinement  $\Delta$  of  $\mathcal{O}$ .

One useful theorem about metacompact sets is how it relates sequentially compact spaces to compact spaces. In general, these two need not be equivalent. The classic examples are the long line (which is sequentially compact but not compact) and the product space of the closed unit interval  $[0, 1]$  with itself uncountably many times (which is compact but not sequentially compact). If the space is both metacompact and countably compact, then it is compact. Since metacompact is implied by paracompactness, and since every metric space is paracompact, we can combine all of this to get a nice theorem: A metric space is compact if and only if it is sequentially compact. A more general version says that sequentially compact metacompact spaces are also compact.

**Theorem 37.1.75.** *If  $(X, \tau_X)$  is a compact topological space, if  $(Y, \tau_Y)$  is a Hausdorff topological space, and if  $f : X \rightarrow Y$  is a continuous function, then it is a closed mapping.*

*Proof.* For if  $C \subseteq X$  is closed, then it is compact since  $X$  is compact. But the continuous image of a compact set is compact, and hence  $f[C]$  is a compact subset of  $Y$ . But  $Y$  is Hausdorff, and thus compact subsets are closed. Thus, the image of closed sets is closed and hence  $f$  is a closed mapping.  $\square$

**Theorem 37.1.76.** *If  $(X, \tau_X)$  is a compact topological space, if  $(Y, \tau_Y)$  is a Hausdorff topological space, and if  $f : X \rightarrow Y$  is a continuous bijection, then it is a homeomorphism.*

**Theorem 37.1.77.** *For  $f$  is then a continuous bijective closed mapping, and hence is a homeomorphism.*

Weakening bijective to injective or surjective obtains topological embeddings and quotient maps, respectively.

**Definition 37.1.45: Proper Function**

A proper function from a topological space  $(X, \tau_X)$  to a topological space  $(Y, \tau_Y)$  is a function  $f : X \rightarrow Y$  such that for every compact subset  $C \subseteq Y$ , the pre-image  $f^{-1}[C]$  is a compact subset of  $X$

Note that the definition of a proper function does not actually require  $f$  to be continuous. Such functions play roles in the theory of groupoids, in analysis,

and in the study of smooth manifolds. If the target space  $Y$  is locally compact and Hausdorff, there's a simple theorem that can accompany this definition.

**Theorem 37.1.78.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $f : X \rightarrow Y$  is a continuous closed function, and if for all  $y \in Y$  it is true that  $f^{-1}[\{y\}]$  is a compact subset of  $X$ , then  $f$  is a proper function.*

*Proof.* For let  $C \subseteq Y$  be compact and let  $K = f^{-1}[C]$ . Suppose  $K$  is not compact. Then there is an open cover  $\mathcal{O}$  of  $K$  such that no finite subcover exists.  $\square$

There's a rewording of a previous theorem in terms of proper functions.

**Theorem 37.1.79.** *If  $(X, \tau_X)$  is a compact topological space, if  $(Y, \tau_Y)$  is a Hausdorff topological space, and if  $f : X \rightarrow Y$  is a continuous function, then it is a proper function.*

*Proof.* For if  $C \subseteq Y$  is compact, then it is closed since  $Y$  is Hausdorff. But the pre-image of a closed subset under a continuous function is closed. Thus,  $f^{-1}[C]$  is a closed subset of  $X$ . But closed subsets of compact sets are compact, and hence  $f^{-1}[C]$  is compact.  $\square$

**Theorem 37.1.80.** *If  $(X, \tau_X)$  is a topological space, if  $(Y, \tau_Y)$  is a Hausdorff topological space, if  $f : X \rightarrow Y$  is continuous, and if  $g : Y \rightarrow X$  is a continuous left inverse of  $f$ , then  $f$  is proper.*

**Theorem 37.1.81.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $f : X \rightarrow Y$  is a continuous proper function, if  $A \subseteq X$  is a saturated subset of  $X$  with respect to  $f$ , and if  $f|_A : A \rightarrow Y$  is the restriction of  $f$  to  $A$ , then  $f$  is proper in the subspace topology on  $A$ .*

### Theorem 37.1.82: Tube Lemma

---

$f : (X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $A \subseteq X$  is compact, if  $B \subseteq Y$  is compact, if  $(X \times Y, \tau_{X \times Y})$  is the product topological space, and if  $\mathcal{O} \subseteq X \times Y$  is an open subset in the product topology such that  $A \times B \subseteq \mathcal{O}$ , then there exists open subsets  $\mathcal{U} \in \tau_X$  and  $\mathcal{V} \in \tau_Y$  such that  $A \times B \subseteq \mathcal{U} \times \mathcal{V}$  and  $\mathcal{U} \times \mathcal{V} \subseteq \mathcal{O}$ .



**Theorem 37.1.83.** *If  $(X, \tau)$  is a locally compact Hausdorff space, and if  $x \in X$ , then there is a precompact subset  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$ .*

*Proof.* For if  $(X, \tau)$  is locally compact and if  $x \in X$ , then there is an open set  $\mathcal{U} \in \tau$  and a compact set  $K \subseteq X$  such that  $x \in \mathcal{U}$  and  $\mathcal{U} \subseteq K$ . But  $X$  is Hausdorff, and thus if  $K$  is compact, then  $K$  is closed. But then  $\text{Cl}_\tau(\mathcal{U}) \subseteq K$ . But then  $\text{Cl}_\tau(\mathcal{U})$  is a closed subset of a compact set, and is therefore closed. Hence,  $\mathcal{U} \in \tau$  is a precompact open set that contains  $x$ .  $\square$

**Theorem 37.1.84.** *If  $(X, \tau)$  is a locally compact Hausdorff space then there exists a basis  $\mathcal{B}$  for  $\tau$  such that for all  $\mathcal{U} \in \mathcal{B}$  it is true that  $\mathcal{U}$  is precompact.*

*Proof.* For if  $(X, \tau)$  is locally compact and Hausdorff, then for all  $x \in X$  there is a precompact open subset  $\mathcal{U}_x$  such that  $x \in \mathcal{U}_x$ . Let  $\mathcal{B}$  be defined by:

$$\mathcal{B} = \{ \mathcal{V} \in \tau \mid \exists_{x \in X} (\mathcal{V} \subseteq \mathcal{U}_x) \} \quad (37.1.34)$$

Then every element of  $\mathcal{V}$  is precompact since  $\text{Cl}_\tau(\mathcal{V}) \subseteq \text{Cl}_\tau(\mathcal{U}_x)$ , which is compact, and closed subsets of compact sets are compact. Moreover,  $\mathcal{B}$  is a basis for  $\tau$ . For if  $\mathcal{O} \in \tau$  is open, then  $\mathcal{U}_x \cap \mathcal{O}$  is open for all  $x \in X$ . And this is a subset of  $\mathcal{U}_x$  and hence contained in  $\mathcal{B}$ . But then:

$$\mathcal{O} = \mathcal{O} \cap X = \mathcal{O} \cap \left( \bigcup_{x \in X} \mathcal{U}_x \right) = \bigcup_{x \in X} (\mathcal{O} \cap \mathcal{U}_x) \quad (37.1.35)$$

Thus,  $\mathcal{B}$  is a basis.  $\square$

**Theorem 37.1.85.** *Every open or closed subspace of a locally compact Hausdorff space is again locally compact Hausdorff.*

**Theorem 37.1.86.** *If  $(X, \tau_X)$  is a topological space, if  $(Y, \tau_Y)$  is a locally compact Hausdorff space, if  $f : X \rightarrow Y$  is continuous proper function, then  $f$  is a closed map.*

*Proof.* For let  $C \subseteq X$  be closed. Suppose  $f[C]$  is not closed. Then there exists a convergent sequence  $a : \mathbb{N} \rightarrow X$  with a limit  $x \in X \setminus f[C]$  such that for all  $n \in \mathbb{N}$  it is true that  $a_n \in C$ . But  $f$  is continuous and therefore  $f(a_n) \rightarrow f(x)$ . But  $Y$  is locally compact and Hausdorff and thus there is a precompact open subset  $\mathcal{V} \in \tau_Y$  such that  $f(x) \in \mathcal{V}$ . But  $f$  is proper, and therefore  $f^{-1}[\text{Cl}_Y(\mathcal{V})]$  is a compact subset of  $X$ . But then  $C \cap f^{-1}[\mathcal{V}]$  is a closed subset of a compact set, and is therefore compact. But then  $f(x) \in f[C]$ , a contradiction. Thus,  $f[C]$  is closed.  $\square$

One of the most important theorems in the study of metric and locally compact Hausdorff spaces is the Baire category theorem. It comes in three flavors.

### Theorem 37.1.87: The First Baire Category Theorem

If  $(X, \tau)$  is a completely metrizable topological space, then it is a Baire space.

### Theorem 37.1.88: The Second Baire Category Theorem

If  $(X, \tau)$  is a locally compact Hausdorff space, then it is a Baire space.

**Theorem 37.1.89: The Third Baire Category Theorem**

If  $(X, \tau)$  is a completely metrizable topological space, if  $C : \mathbb{N} \rightarrow \mathcal{P}(X)$  is a sequence of closed sets, and if  $\bigcup C = X$ , then there exists an  $N \in \mathbb{N}$  such that  $\text{Int}_\tau(C_N) \neq \emptyset$ .

**Theorem 37.1.90.** *If  $(X, \tau)$  is a Baire space, if and if  $C \subseteq X$  is a countable closed subset of  $X$ , then there exists an isolated point  $x \in C$ .*

**Definition 37.1.46: Compact Exhaustion**

A compact exhaustion of a topological space  $(X, \tau)$  is a sequence of compact sets  $K : \mathbb{N} \rightarrow \mathcal{P}(X)$  such that for all  $n \in \mathbb{N}$  it is true that:

$$K_n \subseteq \text{Int}_\tau(K_{n+1})$$

and such that  $X = \bigcup K_n$ .

**Theorem 37.1.91: Existence of Compact Exhaustions**

If  $(X, \tau)$  is a second countable locally compact Hausdorff topological space, then there exists a compact exhaustion of  $X$ .

*Proof.* Since  $(X, \tau)$  is locally compact and Hausdorff, there is a basis of precompact open sets. But a second countable space is Lindelöf, and hence there is a countable subcover. Let  $K_n = \bigcup \text{Cl}_\tau(\mathcal{U}_k)$ .  $\square$

**37.1.7 Algebraic Topology**

Def homotopy, homotopic. Path homotopic functions in a topological space  $X$  are paths  $\gamma_1, \gamma_2 : [0, 1] \rightarrow X$  which are homotopic relative to the set  $\{0, 1\}$ . Path homotopy forms an equivalence relation on the set of all paths between two points,  $x, y \in X$ . Given two paths  $f, g : [0, 1] \rightarrow X$  such that  $f(1) = g(0)$ , their product is the new path that concatenates these two travelling as twice the speed.

$$(f * g)(t) = \begin{cases} f(2t), & 0 \leq t \leq \frac{1}{2} \\ g(2t - 1), & \frac{1}{2} < t \leq 1 \end{cases} \quad (37.1.36)$$

Since  $f(1) = g(0)$ ,  $f \star g$  is continuous at  $1/2$  and is therefore a path. If  $f, f', g, g'$  are paths such that  $f$  and  $f'$  are path homotopic, and  $g$  and  $g'$  are path homotopic, then  $f \star g$  and  $f' \star g'$  will be path homotopic as well. Multiplication of paths is *not* associative (draw a picture). Def loop ( $f : [0, 1] \rightarrow X$  vs  $f : \mathbb{S}^1 \rightarrow X$ ). Def fundamental group. Identity is constant map, inverse of  $f(t)$  is  $f(1-t)$ .

**Theorem 37.1.92.** *If  $(X, \tau)$  is a path connected topological space and if  $x, y \in X$ , then the fundamental groups  $\pi_1(X, x)$  and  $\pi_1(X, y)$  are isomorphic.*

#### Definition 37.1.47: Simply Connected

A simply connected topological space is a topological space  $(X, \tau)$  such that for all  $p \in X$  it is true that the fundamental group  $\pi_1(X, x)$  is isomorphic to the trivial group.

**Theorem 37.1.93.** *If  $(X, \tau)$  is a path connected topological space, then it is simply connected if and only if for all  $x, y \in X$  and all paths  $\gamma_1, \gamma_2 : [0, 1] \rightarrow X$  such that  $\gamma_1(0) = x$ ,  $\gamma_2(0) = x$  and such that  $\gamma_1(1) = y$  and  $\gamma_2(1) = y$ , it is true that  $\gamma_1$  and  $\gamma_2$  are path homotopic.*

*Proof.* For let  $h = \gamma_1 \star \gamma_2^{-1}$ . Since  $X$  is simply connected, there is a homotopy  $H$  between  $h$  and the path  $\gamma_2 \star \gamma_2^{-1}$ .  $\square$

**Theorem 37.1.94.** *If  $f_1, f_2 : X \rightarrow Y$ ,  $g_1, g_2 : Y \rightarrow Z$  are continuous,  $f_1$  homotopy to  $f_2$ ,  $g_1$  homotopic to  $g_2$ , then  $g_1 \circ f_1$  is homotopic to  $g_2 \circ f_2$ .*

#### Definition 37.1.48: Induced Homomorphism of $\pi_1$

The induced homomorphism from the fundamental group of a topological space  $(X, \tau_X)$  at a point  $x \in X$  into the fundamental group of a topological space  $(Y, \tau_Y)$  about a point  $y \in Y$  by a continuous function  $f : X \rightarrow Y$  such that  $f(x) = y$  is the function  $f_* : \pi_1(X, x) \rightarrow \pi_1(Y, y)$  defined by:

$$f_*[\gamma] = [f \circ \gamma]$$

Avoiding proof by

**Theorem 37.1.95.** *The induced homomorphism is a homomorphism.*

**Theorem 37.1.96.** *If  $(X, \tau_X)$ ,  $(Y, \tau_Y)$ , and  $(Z, \tau_Z)$  are topological spaces, if  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  are continuous, if  $x \in X$ , if  $y = f(x)$ , and if*

$z = g(z)$ , then:

$$(g \circ f)_* = g_* \circ f_* \quad (37.1.37)$$

where  $h_*$  denotes the induced homotopy between fundamental groups.

**Theorem 37.1.97.** If  $(X, \tau)$  is a topological space, if  $x \in X$ , and if  $\text{id}_X$  is the identity map, then  $\text{id}_{X*}$  is the unital element of  $\pi_1(X, x)$ .

Homeomorphic topological spaces have isomorphic fundamental groups. There's a weakening of the notion of convexity that is useful in the study of fundamental groups.

#### Definition 37.1.49: Star Shaped

A star shaped subset of a vector space  $(V, *_+)$  over a field  $(F, +, \cdot)$  is a subset  $\mathcal{U} \subseteq V$  such that there exists a point  $\mathbf{x} \in \mathcal{U}$  such that for all  $\mathbf{y} \in \mathcal{U}$  the straight line between  $\mathbf{x}$  and  $\mathbf{y}$  is contained in  $\mathcal{U}$ .

Thus, every convex body is star shaped but the converse is false. Indeed, a *star* is star shaped by not convex since the line between the outer vertices leaves the body.

**Theorem 37.1.98.** If  $\mathcal{U} \subseteq \mathbb{R}^n$  is star shaped, then it is simply connected.

*Proof.* Use the straight-line homotopy between the central point  $\mathbf{x}$  and any other point  $\mathbf{y}$ .  $\square$

Fundamental group of  $S^1$  is  $\mathbb{Z}$ , all higher spheres are simply connected. Fundamental group of product is product of fundamental groups. Thus, fundamental group of torus is  $\mathbb{Z}^2$ . Homotopy equivalence  $f : X \rightarrow Y$ , homotopy inverse  $g : Y \rightarrow X$ , homotopy equivalent spaces.

**Theorem 37.1.99.** If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $f : X \rightarrow Y$  is a homotopy equivalence, if  $x_0 \in X$ , and if  $y_0 = f(x_0)$ , then the induced homomorphism  $f_* : \pi_1(X, x_0) \rightarrow \pi_1(Y, y_0)$  is an isomorphism.

*Proof.* For it suffices to show that  $f_*$  is bijective. Suppose  $[\gamma_1], [\gamma_2] \in \pi_1(X, x)$  are distinct equivalence classes. Since  $f$  is a homotopy equivalence, there exists a homotopy inverse  $g : X \rightarrow Y$ . But then  $g \circ f$  is homotopic to  $\text{id}_X$ , and so there is a homotopy  $H_X$  between them. Thus, if  $f_*([\gamma_1]) = f_*([\gamma_2])$ , then there is a homotopy  $H_Y$  between them. But then  $G_X$  defined by:  $\square$

Covering maps are local homeo, open, and quotient maps. Injective covering map is a homeo. Fibers have the same cardinality, called the number of sheets of the covering. Covering map is proper if and only if it is finite sheeted. Finite product of covering maps is covering map.

**Definition 37.1.50: Lift To Covering Spaces**

lift of a continuous function  $f : X \rightarrow Y$  from a topological space  $(X, \tau_X)$  to topological space  $(Y, \tau_Y)$  to a covering space  $(E, \tau)$  of  $Y$  is a continuous function  $\tilde{f} : X \rightarrow E$  such that  $\pi \circ \tilde{f} = f$ , where  $\pi$  is the covering map  $\pi : E \rightarrow Y$ .

**Theorem 37.1.100.** *If  $B$  is connected,  $f : B \rightarrow X$  continuous, and if  $\tilde{f}_1, \tilde{f}_2$  are lifts to  $E$ , then if they are equal at one point, they are equal everywhere.*

**Theorem 37.1.101.** *If  $f : I \rightarrow X$  is a path,  $e \in E$ ,  $\pi(e) = f(0)$ , then there is a unique lift  $\tilde{f} : I \rightarrow E$  such that  $\tilde{f}(0) = e$ .*

**Theorem 37.1.102** (Monodromy Theorem). *If  $f, g : I \rightarrow X$  are path-homotopic paths, and  $\tilde{f}, \tilde{g}$  are lifts that are equal at some point  $e \in E$ , then the lifts are path homotopic.*

**Theorem 37.1.103** (Lifting Criterion). *If  $\pi : E \rightarrow X$  is a covering map, if  $Y$  is connected and locally path connected, and if  $f : Y \rightarrow X$  is continuous, and if  $y \in Y$ ,  $e \in E$  are such that  $\pi(e) = f(y)$ , then there is a lift  $\tilde{f} : Y \rightarrow E$  such that  $\tilde{f}(y) = e$  if and only if the induced homomorphism  $f_*[\pi_1(Y, y)] \subseteq \pi_*[\pi_1(E, e)]$ .*

**Theorem 37.1.104.** *If  $X$  is simply connected, and if  $\pi : E \rightarrow X$  is a covering map, then it is a homeomorphism.*

**Theorem 37.1.105.** *If  $X$  is connected, locally path connected, and semi-locally simply connected, then there is a universal cover.*

## 37.2 Homogeneous Spaces

**Definition 37.2.1: Topologically Homogeneous Space**

A topologically homogeneous space is a topological space  $(X, \tau)$  such that for all  $x, y \in X$  there exists a homeomorphism  $f : X \rightarrow X$  such that  $f(x) = y$ .

Homogeneous spaces are ones where any point  $x$  looks locally like any other point. There is a related notion called the *homogeneous spaces of a Lie group*, but the definitions are not equivalent. The naming is quite unfortunate, and we will attempt to differentiate the two clearly.

**Example 37.2.1** For any  $n \in \mathbb{N}$ , the Euclidean space  $\mathbb{R}^n$  is homogeneous. Given  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , define  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  by:

$$f(\mathbf{z}) = \mathbf{y} + (\mathbf{x} - \mathbf{z}) \quad (37.2.1)$$

Then  $f(\mathbf{x}) = \mathbf{y}$ , as desired. This is bijective and bicontinuous, and therefore a homeomorphism. That is to say, translation of  $\mathbb{R}^n$  are homeomorphisms.

**Example 37.2.2** The sphere  $\mathbb{S}^n$  is topologically homogeneous with its usual topology. If  $n \geq 1$ , given two points  $\mathbf{x}, \mathbf{y} \in \mathbb{S}^n$ , let  $R$  be the rotation mapping that takes  $\mathbf{x}$  to  $\mathbf{y}$ . This is bijective and bicontinuous, and hence a homeomorphism. For the case of  $n = 0$ ,  $\mathbb{S}^0$  is merely two points and the topology is the power set. Letting  $f$  be the function that swaps these two points shows that  $\mathbb{S}^0$  is also homogeneous.

### Definition 37.2.2: Topological Group

A topological group, denoted  $(G, *, \tau)$ , is a topological  $(G, \tau)$  with a binary operation  $*$  such that  $(G, *)$  is a group, the function  $g : G \times G \rightarrow G$  defined by  $g(a, b) = a * b$  is continuous with respect to the product topology  $\tau \times \tau$ , and the function  $\nu : G \rightarrow G$  defined by  $\nu(a) = a^{-1}$ , where  $a^{-1}$  is the inverse element of  $a$  with respect to  $*$ , is continuous.

**Example 37.2.3** The simplest example of a topological group is the Abelian group structure  $(\mathbb{R}^n, +)$  equipped with the standard topology for some  $n \in \mathbb{N}^+$ .  $\mathbf{x} + \mathbf{y}$  is continuous, as is  $\mathbf{x} \mapsto -\mathbf{x}$ , showing that  $(\mathbb{R}^n, +, \tau)$  is a topological group.

Every group  $(G, *)$  can be given a topological group structure.

**Theorem 37.2.1.** *If  $(G, *)$  is a group, and if  $\tau = \mathcal{P}(G)$ , then  $(G, *, \tau)$  is a topological group.*

*Proof.* For if  $(G, \tau)$  is the discrete topology, then for any function  $f : G \rightarrow G$  it is true that  $G$  is continuous, and thus the function  $\nu : G \rightarrow G$  defined by  $\nu(a) = a^{-1}$  is continuous. Moreover, since  $(G, \tau)$  is the discrete topology,  $\tau \times \tau$  is the discrete topology on  $G \times G$ . But then for any function  $f : G \times G \rightarrow G$  it is true that  $f$  is continuous, and thus  $g : G \times G \rightarrow G$  defined by  $g(a, b) = a * b$  is continuous. Thus,  $(G, *, \tau)$  is a topological group (Def. 37.19.11).  $\square$

Every non-trivial group can also be given a non-Hausdorff group structure.

**Theorem 37.2.2.** *If  $(G, *)$  is a group, if  $\tau = \{\emptyset, G\}$  is the trivial topology, then  $(G, *, \tau)$  is a topological group.*

*Proof.* For it  $(G, \tau)$  is a trivial topological space, and if  $f : G \rightarrow G$  is a function, then  $f$  is continuous. Thus  $\nu : G \rightarrow G$  defined by  $\nu(a) = a^{-1}$  is continuous. If  $\tau$  is trivial, then  $(G \times G, \tau \times \tau)$  is the trivial topological space on  $G \times G$ . But then for any function  $g : G \times G \rightarrow G$  it is true that  $g$  is continuous, and thus  $g(a, b) = a * b$  is continuous. Therefore,  $(G, \tau)$  is a topological group (Def. 37.19.11).  $\square$

Thus, so long as  $(G, *)$  is not the trivial group, then  $(G, *, \tau)$ , where  $\tau$  is the trivial topology, is a topological group and  $(G, *, \tau)$  is non-Hausdorff. Indeed, such spaces won't even be Kolmogorov ( $T_0$ ). It is very convenient to study  $T_0$  topological groups, since many other separation properties then come for free. That is, if  $(G, *, \tau)$  is  $T_0$ , then it is automatically accessible ( $T_1$ ), Hausdorff ( $T_2$ ), regular Hausdorff ( $T_3$ ), and completely regular Hausdorff ( $T_{3\frac{1}{2}}$ ). Quite a lot from the mere assumption of being  $T_0$ ! To prove this we must first show that topological groups are homogeneous.

**Theorem 37.2.3.** *If  $(G, *, \tau)$  is a topological group, and if  $\nu : G \rightarrow G$  is defined by  $\nu(a) = a^{-1}$ , then  $\nu$  is a homeomorphism.*

*Proof.* For since  $(G, *, \tau)$  is a topological group,  $\nu$  is continuous (Def. 37.19.11). But  $\nu$  is bijective since inverses are unique and every element of a group has an inverse. Moreover,  $\nu^{-1}(a) = (a^{-1})^{-1} = a$ , and hence  $\nu^{-1} = \text{id}_G$ , which is continuous. Thus,  $\nu$  is a homeomorphism.  $\square$

**Theorem 37.2.4.** *If  $(G, *, \tau)$  is a topological group, and if  $a \in G$ , then the function  $f : G \rightarrow G$  defined by  $f(b) = a * b$  is a homeomorphism.*

*Proof.* By Cayley's theorem,  $f$  is bijective. But the function  $g : G \times G \rightarrow G$  defined by  $g(x, y) = x * y$  is continuous since  $(G, *, \tau)$  is a topological group (Def. 37.19.11), and thus  $g|_{\{a\} \times G}$  is continuous. But  $g|_{\{a\} \times G} = f$ , and thus  $f$  is continuous. Moreover,  $f^{-1}(x) = x^{-1} * a^{-1}$ , and since  $\nu : G \rightarrow G$  defined by  $\nu(x) = x^{-1}$  is continuous (Def. 37.19.11), it then follows that  $f^{-1}$  is continuous. Hence,  $f$  is a homeomorphism.  $\square$

**Theorem 37.2.5.** *If  $(G, *, \tau)$  is a topological group, then  $(G, \tau)$  is topologically homogeneous.*

*Proof.* For let  $a, b \in G$  and let  $r = b * a^{-1}$ . Let  $f : G \rightarrow G$  be defined by  $f(x) = r * x$ . But then  $f$  is a homeomorphism (Thm. 37.2.4), and moreover:

$$f(a) = r * a = (b * a^{-1}) * a = b * (a^{-1} * a) = b * e = b \quad (37.2.2)$$

and thus  $f$  is a homeomorphism such that  $f(a) = b$ . Therefore,  $G$  is topologically homogeneous (Def. 37.2.1).  $\square$

Hence every topological group is an example of a topologically homogeneous space.

**Example 37.2.4** The circle  $\mathbb{S}^1$  with the operation  $*$  defined by:

$$\exp(i\theta_1) * \exp(i\theta_2) = \exp(i(\theta_1 + \theta_2)) \quad (37.2.3)$$

makes  $(\mathbb{S}^1, *, \tau)$  a topological group. That is,  $*$  rotates the angle  $\theta_1$  by  $\theta_2$ .

**Example 37.2.5** The general linear group  $\mathrm{GL}_n(\mathbb{R}^n)$  is a topological group, with  $\tau$  be the subspace topology given by viewing  $\mathrm{GL}_n(\mathbb{R}^n)$  as a subset of  $\mathbb{R}^{n^2}$ . In a similar vein, the orthogonal group  $\mathrm{O}_n(\mathbb{R}^n)$  is a topological group. Moreover,  $\mathrm{O}_n(\mathbb{R}^n)$  is a compact topological group since it is a closed subspace and every element is bounded by 1. Hence, by Heine-Borel, it is compact.

**Example 37.2.6** The rational numbers with the subspace topology and usual addition form a topological group.

A standard example of the quotient topology comes from considering a group  $(G, *)$  and a topological space  $(X, \tau)$ . If  $\Theta : G \times X \rightarrow X$  is a group action of  $G$  on  $X$ , then the orbits of  $X$  under  $\Theta$  form an equivalence relation on  $X$ . That is, the sets:

$$\Theta_x = \{ \Theta(x, g) \in X \mid g \in G \} \quad (37.2.4)$$

form an equivalence relation. The important case of  $(G, *)$  being a topological group gives rise to the notion of a *continuous* group action.

### Definition 37.2.3: Continuous Group Action

A continuous group action of a topological group  $(G, *, \tau_G)$  on a topological space  $(X, \tau_X)$  is a group action  $\Theta : G \times X \rightarrow X$  such that  $\Theta$  is a continuous function with respect to the product topology  $\tau_G \times \tau_X$ .

We often write  $\Theta$  as  $\cdot$ , even though  $\Theta$  is not actually a binary operation since  $G$  and  $X$  may be different sets. Recall from group theory that a group action  $\cdot$  satisfies the two following properties:

$$(g_1 * g_2) \cdot x = g_1 \cdot (g_2 \cdot x) \quad (37.2.5)$$

$$e * x = x \quad (37.2.6)$$

In the case when the quotient space is formed by a topological group, we use the following notation:

### Notation 37.2.1: Quotient by Topological Group

If  $(G, *, \tau)$  is a topological group, if  $(X, \tau)$  is a topological space, and if  $\Theta : G \times X \rightarrow X$  is a continuous group action, then we denote  $X/G$  as the quotient set formed by the equivalence relation  $R$  on  $X$  where  $xRy$  if and only if  $x$  and  $y$  are in the same orbit of  $\Theta$ .

**Example 37.2.7** Consider  $(\mathbb{Z}, +)$  with its usual group structure. The subspace topology  $\tau_{\mathbb{Z}}$  inherited from  $\mathbb{R}$  is simply the power set  $\mathcal{P}(\mathbb{Z})$ . Thus, by Thm. 37.2.1  $(\mathbb{Z}, +, \tau_{\mathbb{Z}})$  is a topological group. It acts on  $(\mathbb{R}, \tau_{\mathbb{R}})$  in the familiar way:

$$\Theta(n, x) = n + x \quad (37.2.7)$$

the orbits of  $\Theta$  on  $X$  forms the equivalence relation  $R$  where  $xRy$  if and only if  $x - y \in \mathbb{Z}$ . The quotient space is homeomorphic to the circle  $\mathbb{S}^1$  under the homeomorphism  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{S}^1$  defined by:

$$f([x]) = \exp(2\pi i x) \quad (37.2.8)$$

this is well defined since for all  $y \in [x]$  there is an  $n \in \mathbb{Z}$  with  $y = x + n$ . Hence:

$$f([y]) = \exp(2\pi i y) \quad (37.2.9a) \qquad = \exp(2\pi i x) \exp(2\pi i n) \quad (37.2.9d)$$

$$= \exp(2\pi i(x + n)) \quad (37.2.9b) \qquad = \exp(2\pi i x) \quad (37.2.9e)$$

$$= \exp(2\pi i x + 2\pi i n) \quad (37.2.9c) \qquad = f([x]) \quad (37.2.9f)$$

This is bijective and bicontinuous, and therefore a homeomorphism.

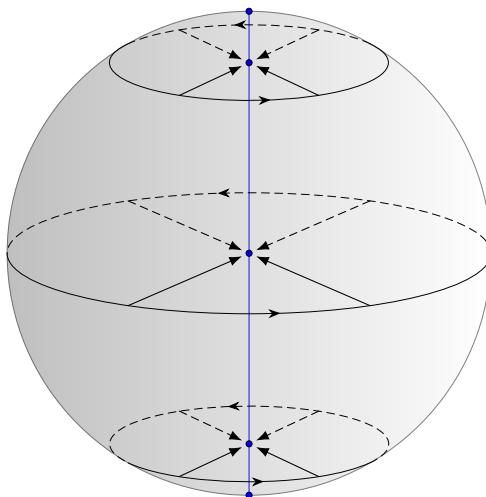
**Example 37.2.8** Consider the sphere  $\mathbb{S}^2$  and the group  $(\mathbb{S}^1, *)$  where

$$\exp(i\theta) * \exp(i\phi) = \exp(i(\theta + \phi)) \quad (37.2.10)$$

This acts on  $\mathbb{S}^2$  with  $\Theta : \mathbb{S}^1 \times \mathbb{S}^2 \rightarrow \mathbb{S}^2$  defined by:

$$\Theta(\theta, (\varphi, \psi)) = (\varphi + \theta, \psi) \quad (37.2.11)$$

that is, we rotate azimuthally by  $\theta$ . The orbits are circles concentric with the  $z$  axis, and hence  $\mathbb{S}^2/\mathbb{S}^1$  is the interval  $[-1, 1]$  (see Fig. 37.5).

Fig. 37.5: The Quotient of  $\mathbb{S}^2$  by  $\mathbb{S}^1$ 

Not every space is homogeneous.

**Example 37.2.9** The space  $X = \{0\} \cup (1, 2)$ , with the subspace topology is not homogeneous. 0 is an isolated point, and since homomorphisms preserve such a notion, there can be no homomorphism  $f : X \rightarrow X$  such that  $f(0) \in (1, 2)$ .

More interesting examples arise when we study locally Euclidean spaces, and in particular manifolds. Manifolds are locally Euclidean Hausdorff topological spaces that are second countable, and every connected manifold is homogeneous. If we rid ourselves of the Hausdorff requirement we can show that there does indeed exist non-Hausdorff locally Euclidean second countable connected spaces, and moreover we can show that they may or may not be homogeneous, paradoxically. Such examples arise from the study of the order topology.

#### Definition 37.2.4: Order Topology

The order topology from a totally ordered set  $(X, \leq)$  is the topology  $\tau$  generated by sets of the form:

$$(a, b) = \{x \in X \mid a < x \text{ and } x < b\} \quad (37.2.12)$$

together with open rays:

$$(-\infty, a) = \{x \in X \mid x < a\} \quad (37.2.13a)$$

$$(a, \infty) = \{x \in X \mid a < x\} \quad (37.2.13b)$$

**Example 37.2.10** The standard topology on  $\mathbb{R}$  is an order topology induced by the standard order. The long line is another such topology induced by the lexicographic ordering on  $[0, 1] \times \omega_1$ , where  $\omega_1$  is the first uncountable ordinal.

**Definition 37.2.5** A complete totally ordered set is a totally ordered set  $(X, \leq)$  with the least upper bound property.

**Theorem 37.2.6.** *If  $(X, \leq)$  is a totally ordered set, if  $\tau$  is the order topology, and if  $(X, \tau)$  is connected, then  $(X, \leq)$  is a complete totally ordered set.*

*Proof.* For suppose not. Then there exists a set  $\mathcal{U} \subseteq X$  that is bounded above that contains no least upper bound. Let  $\mathcal{V}_+$  be defined as follows:

$$\mathcal{V}_+ = \{x \in X \mid \forall y \in \mathcal{U} (y < x)\} \quad (37.2.14)$$

Since  $\mathcal{U}$  is bounded above,  $\mathcal{V}_+$  is non-empty. Moreover, it is open since:

$$\mathcal{V}_+ = \bigcup_{x \in \mathcal{V}_+} (x, \infty) \quad (37.2.15)$$

For if  $x_0 \in \bigcup (x, \infty)$  then there is an  $x \in \mathcal{V}_+$  such that  $x_0 \in (x, \infty)$ . But for all  $y \in \mathcal{U}$  it is true that  $y < x$ , and thus by transitivity  $y < x_0$ . Thus, by definition,  $x_0 \in \mathcal{V}_+$ . Now if  $x_0 \in \mathcal{V}_+$ , suppose it is not in  $\bigcup (x, \infty)$ . Then for all  $x \in \mathcal{V}_+$  it is true that  $x_0 < x$ , or  $x = x_0$ , and thus  $x_0$  is a greatest lower bound of  $\mathcal{V}_+$ . But by construction,  $x_0$  is then a least upper bound of  $\mathcal{U}$ , but  $\mathcal{U}$  has no such thing. Thus  $x_0 \in \bigcup (x, \infty)$ , and hence we have equality. Thus  $\mathcal{V}_+$  is open. By a similar argument,  $X \setminus \mathcal{V}_+$  is open, and hence  $X$  is disconnected, a contradiction. Therefore  $(X, \leq)$  is complete.  $\square$

The converse does not hold. Consider in  $\mathbb{R}$  the subset  $[0, 1] \cup [2, 3]$ . Since this is a closed and bounded subset, it is compact by the Heine-Borel theorem,

and compact subsets of metric spaces are complete and totally bounded. Now it may seem like we're using the word *complete* in two separate meanings: metrically complete pertains to Cauchy sequences, whereas complete ordering deal with the least upper bound principle. However, in  $\mathbb{R}$  these are the same.

**Theorem 37.2.7.** *If  $(X, \leq)$  is a totally ordered set, if  $\tau$  is the order topology, if  $\mathcal{U}$  is a connected subset of  $X$ , and if  $a, b \in \mathcal{U}$  are such that  $a < b$ , then  $(a, b) \subseteq \mathcal{U}$ .*

*Proof.* Suppose not. Then there is a  $c \in (a, b)$  such that  $c \notin \mathcal{U}$ . But then  $(c, \infty) \cap \mathcal{U}$  and  $(-\infty, c) \cap \mathcal{U}$  are disjoint non-empty subsets that are open in  $\mathcal{U}$ , and hence  $\mathcal{U}$  is disconnected, a contradiction.  $\square$

### Theorem 37.2.8: Intermediate Value Theorem

---

If  $(X, <_X)$  and  $(Y, <_Y)$  are totally ordered sets, if  $\tau_X$  and  $\tau_Y$  are the order topologies on  $X$  and  $Y$ , respectively, if  $(X, \tau_X)$  is a connected topological topological space, if  $f : X \rightarrow Y$  is a continuous function, and if  $a, b \in X$  are such that  $a <_X b$  and  $f(a) <_Y f(b)$ , then for all  $c \in (f(a), f(b))$  there is an  $x_0 \in (a, b)$  such that  $f(x_0) = c$ .  $\blacksquare$

*Proof.* For the continuous image of a connected set is connected, and thus  $f([a, b])$  is a connected subset of  $Y$ . But  $f(a) \in f([a, b])$  and  $f(b) \in f([a, b])$  and since  $f([a, b])$  is connected we have that  $(f(a), f(b)) \subseteq f([a, b])$ . But then for all  $c \in (f(a), f(b))$  it is true that  $c \in f((a, b))$ , and thus there is an  $x_0 \in (a, b)$  such that  $f(x_0) = c$ .  $\square$

**Theorem 37.2.9.** *If  $(X, \leq)$  is a totally ordered set, if  $\tau$  is the order topology, if  $(X, \tau)$  is connected, and if  $f : X \rightarrow X$  is a homeomorphism, then either  $f$  is order preserving or order reversing.*

*Proof.* For suppose not. Then there are  $a, b, c \in X$  such that  $a < b$  and  $b < c$ , yet either  $f(b) < f(a)$  and  $f(b) < f(c)$ , or  $f(a) < f(b)$  and  $f(c) < f(b)$ . Suppose  $f(a) < f(b)$  and  $f(c) < f(b)$ . Since  $a < b$  and  $b < c$ , by transitivity  $a < c$ , and thus  $a \neq c$ . Furthermore, since  $f$  is a homeomorphism it is bijective, and therefore injective, and thus  $f(a) \neq f(c)$ . But  $\leq$  is a total ordering and thus by trichotomy either  $f(a) < f(c)$  or  $f(c) < f(a)$ . Suppose  $f(a) < f(c)$ . But since  $f(a) < f(c)$  and  $f(c) < f(b)$ , and thus  $f(c) \in (f(a), f(b))$ . But then by the intermediate value theorem there is an  $x_0 \in (a, b)$  such that  $f(x_0) = f(c)$ . But if  $x_0 \in (a, b)$ , then  $x_0 < b$ . But  $b < c$ , and therefore  $x_0 < c$  and hence  $x_0 \neq c$ . But then  $f(x_0) = f(c)$  and  $x_0 \neq c$ , contradicting the fact that  $f$  is a bijection, a contradiction. Similarly, it is not true that  $f(b) < f(a)$  and  $f(b) < f(c)$ . Thus, either  $f$  is order preserving or order reversing.  $\square$

We can now prove that there are non-homogeneous spaces.

**Example 37.2.11** Let  $X = (0, 1]$ , with the standard subspace topology inherited from  $\mathbb{R}$ . This is the same as the order topology from the standard ordering. Suppose  $(0, 1]$  is homogeneous and let  $x = 1$ ,  $y = 1/2$ . Then there is a homeomorphism  $f : X \rightarrow X$  such that  $f(1) = 1/2$ . But any homeomorphism must be order preserving or order preserving, and since  $1 \geq x$  for all  $x \in (0, 1]$ , either  $f(1) \geq x$  for all  $x$ , or  $f(1) \leq x$  for all  $x$ . But if  $f(1) \leq x$  for all  $x$  then  $f(1)$  is a greatest lower bound, but  $(0, 1]$  has no such bound. Hence,  $f(1) \geq x$  for all  $x \in X$ . But then  $f(1) = 1$ , a contradiction since  $f(1) = 1/2$ . Hence,  $(0, 1]$  is not homogeneous.

The next theorem to prove is that any connected Hausdorff locally Euclidean topological space is homogeneous.

**Theorem 37.2.10.** *If  $(X, \tau)$  is a connected Hausdorff locally Euclidean topological space, then it is homogeneous.*

If the Hausdorff property is lost we may lose the homogeneity of the space as well. The fact that locally Euclidean spaces can be non-Hausdorff is one justification for including the Hausdorff property in the definition of a topological manifold, which is a locally Euclidean Hausdorff topological space that is second countable. First we'll demonstrate the existence of non-Hausdorff locally Euclidean topological spaces.

**Example 37.2.12** Let  $X$  be the disjoint union of  $\mathbb{R}$  with itself:

$$X = \mathbb{R} \sqcup \mathbb{R} \quad (37.2.16)$$

That is,  $X = (\mathbb{R} \times \{0\}) \cup (\mathbb{R} \times \{1\})$ . Furthermore, consider the relation  $R' \subseteq X \times X$  defined by:

$$R' = \left\{ \left( (x, a), (x, b) \right) \mid x \in \mathbb{R} \setminus \{0\} \text{ and } a, b \in \mathbb{Z}_2 \right\} \quad (37.2.17)$$

That is,  $(x, a)R'(y, b)$  if and only if  $a, b \in \{0, 1\}$ , and  $x = y$  are non-zero. From this we can obtain an equivalence relation  $R$  as follows:

$$R = R' \cup \{((0, 0), (0, 0))\} \cup \{((0, 1), (0, 1))\} \quad (37.2.18)$$

That is,  $R$  is the reflexive closure of  $R'$ . This is an equivalence relation. Let  $Y$  be the topological space  $(X/R, \tau_R)$ , where  $\tau_R$  is the quotient topology. This space is then locally Euclidean. For if  $x \neq 0$ , let  $r = |x|/2$ . Let  $\mathcal{U}$  be defined by:

$$\mathcal{U} = \{[(y, 0)] \in X \mid |y - x| < r\} \quad (37.2.19)$$

Then  $\mathcal{U}$  is open since for all points  $[z] \in \mathcal{U}$ , either  $z = (y, 0)$  or  $z = (y, 1)$  where  $|x - y| < r$ . By the triangle inequality we can then conclude that  $y \neq 0$ . But then:

$$\bigcup \mathcal{U} = \left( B_r^{\mathbb{R}}(x) \times \{0\} \right) \bigcup \left( B_r^{\mathbb{R}}(x) \times \{1\} \right) \quad (37.2.20)$$

And this is an open subset of  $X$  with the disjoint union topology, and therefore  $\mathcal{U}$  is open in the quotient topology. We can define a homeomorphism between  $\mathcal{U}$  and the open interval  $(x - r, x + r)$  by defining:

$$\phi([(y, 0)]) = y \quad (37.2.21)$$

Similarly, for the points  $(0, 0)$  and  $(0, 1)$  we can define  $\mathcal{U}_0 = X \setminus \{(0, 1)\}$  and  $\mathcal{U}_1 = X \setminus \{(0, 0)\}$ , both of which are homeomorphic to all of  $\mathbb{R}$ . Thus this space, which is called the *bug-eyed line*, or the *line with two origins*, is locally Euclidean but is not Hausdorff since every open set containing  $(0, 0)$  and every open set containing  $(0, 1)$  have non-empty intersection.

The bug-eyed line is shown in Fig. 37.6. The dashed line is used to denote that this final sketch is approximately what the space *looks* like. We can use this figure to show that any open set containing the upper origin must also contain the lower one (see Fig. 37.7).

**Example 37.2.13** The bug-eyed line is not homogeneous. For the points  $(0, 0)$  and  $[(1, 0)]$  can be separated by open sets, as can the points  $(0, 1)$  and  $[(1, 0)]$ . However  $(0, 0)$  can not be separated from  $(0, 1)$ , and since homeomorphisms preserve the Hausdorff property, there can be no such function  $f : X \rightarrow X$  such that  $f((0, 0)) = [(1, 0)]$ . Hence, the bug-eyed line is not homogeneous.

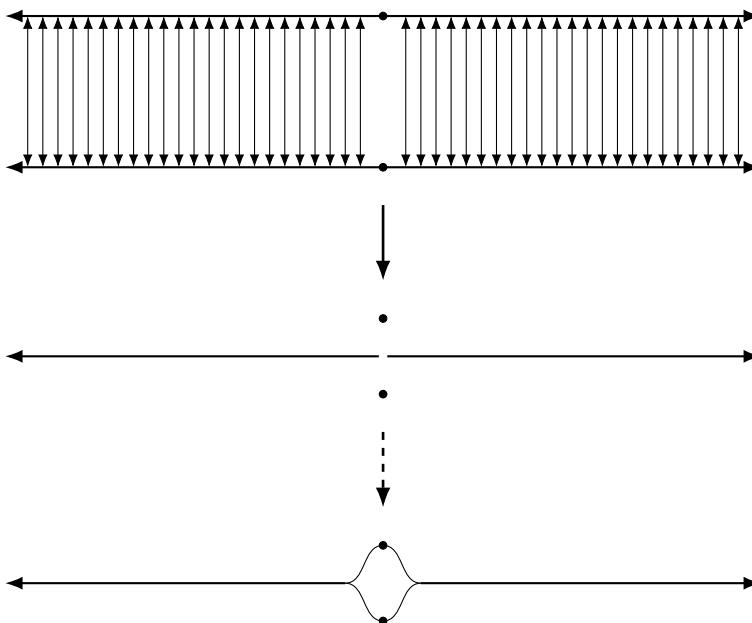


Fig. 37.6: Construction of the Bug-Eyed Line

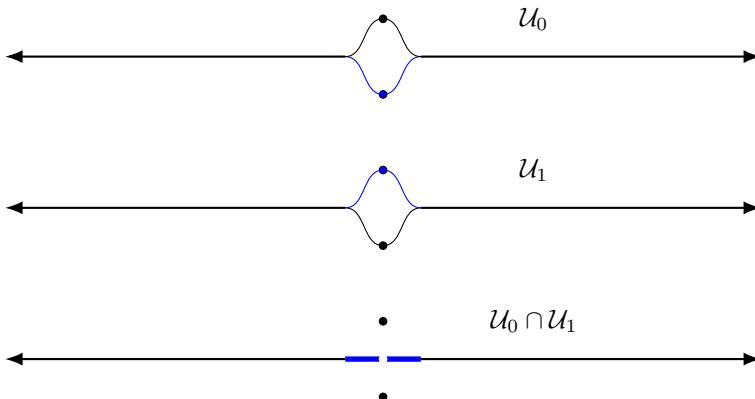


Fig. 37.7: Open Subsets of the Bug-Eyed Line

Another common example of a non-Hausdorff locally Euclidean space is the branching line. This is defined in a similar manner as the bug-eyed line, but with a slightly different equivalence relation. We define  $R$  by:

$$R = \{ ((x, a), (x, b)) \mid x < 0 \text{ and } a, b \in \mathbb{Z}_2 \} \quad (37.2.22)$$

We then look at the reflexive closure of  $R$ , adding  $((x, 0), (x, 0))$  and  $((x, 1), (x, 1))$  for all  $x \geq 0$ . This is an equivalence relation, and looking at  $\mathbb{R} \sqcup \mathbb{R}/R$  with the quotient topology gives us the branching line. The construction is shown in Fig. 37.8.

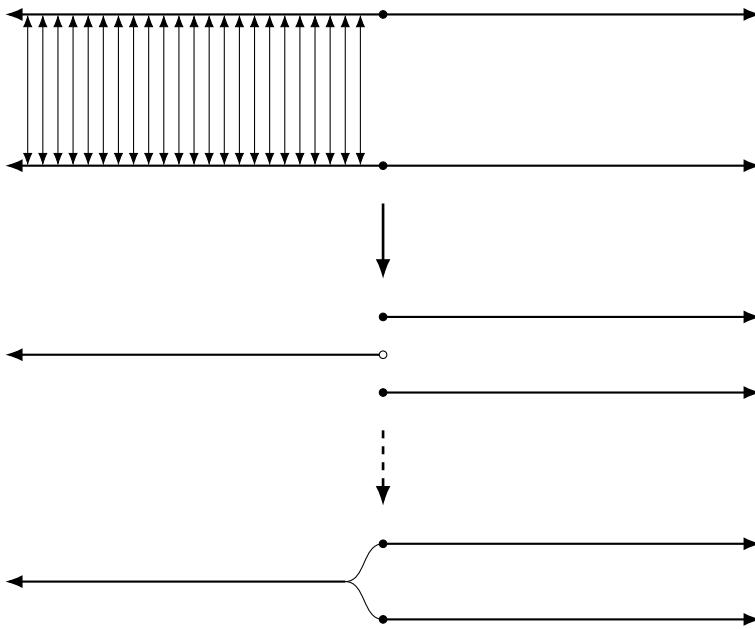


Fig. 37.8: Construction of the Branching Line

Both the bug-eyed line and the branching line are examples of non-Hausdorff, non-homogeneous, locally Euclidean second countable topological spaces. Moreover, they are both path-connected. While connected Hausdorff spaces that are locally Euclidean are indeed homogeneous, the converse is not true. That is, there are connected homogeneous locally Euclidean spaces that are not Hausdorff. If one adds the Lindelöf requirement, then it can be proved that the space is Hausdorff.

#### Definition 37.2.6: The Complete Feather

The complete feather is the set:

$$F = \left( \bigcup_{n \in \mathbb{N}^+} \{ \mathbf{x} \in \mathbb{R}^{n+1} \mid \forall i \in \mathbb{Z}_{n-1} (x_i < x_{i+1}) \wedge (x_{n-1} \leq x_n) \} \right) \bigcup \mathbb{R}$$

The complete feather is a subset of  $\bigcup \mathbb{R}^n$  with the additional requirements that  $x_0 < x_1 < \dots < x_{m-1} \leq x_m$ . No such restraint is imposed on the  $\mathbb{R}^1$  term. We place a topology by considering the following partial ordering.

**Theorem 37.2.11.** *If  $(X, \tau)$  is a Lindelöf topological space that is locally second countable, then it is second countable.*

*Proof.* For if  $(X, \tau)$  is locally second countable, for all  $x \in X$  there is an open subset  $\mathcal{U}_x \in \tau$  such that  $(\mathcal{U}_x, \tau|_{\mathcal{U}_x})$  has a countable basis, where  $\tau|_{\mathcal{U}_x}$  is the subspace topology. But then the set:

$$\mathcal{O} = \{\mathcal{U}_x \in \tau \mid x \in X\} \quad (37.2.23)$$

is an open cover of  $X$ . But  $(X, \tau)$  is Lindelöf and thus there is a countable subcover  $\Delta$ . But for all  $\mathcal{U} \in \Delta$  there is a countable basis for  $\mathcal{U}$ ,  $\mathcal{B}_{\mathcal{U}}$ . But then:

$$\mathcal{O} = \bigcup_{\mathcal{U} \in \Delta} \{\mathcal{B}_{\mathcal{U}} \mid \mathcal{U} \in \Delta\} \quad (37.2.24)$$

Is countable basis for  $(X, \tau)$ . It is countable since it is the countable union of countable sets. Moreover, let  $\mathcal{V} \in \tau$ . But since  $\Delta$  is a cover of  $X$ , we have:

$$\mathcal{V} = \mathcal{V} \cap X = \mathcal{V} \cap \left( \bigcup_{\mathcal{U} \in \Delta} \mathcal{U} \right) = \bigcup_{\mathcal{U} \in \Delta} (\mathcal{U} \cap \mathcal{V}) \quad (37.2.25)$$

But for all  $\mathcal{U} \in \Delta$  it is true that  $\mathcal{U} \cap \mathcal{V} \subseteq \mathcal{U}$ . But then there is a subset  $\Delta_{\mathcal{U}} \subseteq \mathcal{B}_{\mathcal{U}}$

$$\mathcal{U} \cap \mathcal{V} = \bigcup_{A \in \Delta_{\mathcal{U}}} A \quad (37.2.26)$$

Thus:

$$\mathcal{V} = \bigcup_{\mathcal{U} \in \Delta} \left( \bigcup_{A \in \Delta_{\mathcal{U}}} A \right) \quad (37.2.27)$$

□

### 37.3 Locally Euclidean Spaces

We wish to speak about topological spaces that are sufficiently well behaved in the sense that it allows one to do calculus. The first naive approach is to consider such spaces that can be locally approximated by Euclidean space.

#### Definition 37.3.1: Chart

A chart of dimension  $n \in \mathbb{N}$  in a topological space  $(X, \tau)$ , denoted  $(\mathcal{U}, \varphi)$ , is an open set  $\mathcal{U} \in \tau$  and a function  $\varphi : \mathcal{U} \rightarrow \mathbb{R}^n$  such that  $\varphi$  is injective, continuous, and an open mapping with respect to the standard topology on  $\mathbb{R}^n$ .

Some authors simply denote a chart by the function  $\varphi$ . Note that since the image of  $\varphi$  lies in  $\mathbb{R}^n$ , if we compose with one of the projection mappings  $\pi_k$

we obtain a function  $\pi_k \circ \phi : \mathcal{U} \rightarrow \mathbb{R}$ . These are called the coordinate functions of the chart  $(\mathcal{U}, \phi)$  and are often denoted  $x^k = \pi_k \circ \phi$ . Note that we may then write the image of  $p \in \mathcal{U}$  as:

$$\phi(p) = (x^1(p), \dots, x^n(p)) \quad (37.3.1)$$

and such notation is often useful. Avoiding the *dot dot dot* notation, we can recall that  $\mathbf{x} \in \mathbb{R}^n$  is a function  $\mathbf{x} : \mathbb{Z}_n \rightarrow \mathbb{R}$ . Thus if  $p \in \mathcal{U}$  and  $\mathbf{x} = \phi(p)$ , we can write this explicitly by:

$$(\phi(p))(k) = \mathbf{x}(k) = x_k = x^k(p) \quad (37.3.2)$$

Thus distinguishing the difference between the notation  $x^k$  and  $x_k$ . Here,  $x_k$  is simply a real number,  $x^k$  is a function  $x^k : \mathcal{U} \rightarrow \mathbb{R}$ , and  $\mathbf{x}(k)$  is the image of  $k \in \mathbb{Z}_n$  under  $\mathbf{x}$ . In other words, it is the  $k^{th}$  component of  $\mathbf{x}$ .

**Example 37.3.1** If  $X = \mathbb{R}^n$  and  $\tau$  is the standard topology, then for any open subset  $\mathcal{U} \subseteq \mathbb{R}^n$  the pair  $(\mathcal{U}, \text{id}_{\mathbb{R}^n}|_{\mathcal{U}})$  is a chart, where  $\text{id}_{\mathbb{R}^n}|_{\mathcal{U}}$  denotes the restriction of the identity map to  $\mathcal{U}$ .

**Example 37.3.2** If  $X = S^2$ , and if  $\tau$  is the subspace topology inherited from  $\mathbb{R}^3$ , then we can form a chart around the south pole  $(0, 0, -1)$  as follows. Let  $\mathcal{U} = S^2 \setminus \{(0, 0, 1)\}$ . Since  $\mathbb{R}^3$  is Hausdorff, and since  $S^2$  is a subspace of  $\mathbb{R}^3$ , it then follows that  $S^2$  is Hausdorff. But if  $S^2$  is Hausdorff, then the set  $\{(0, 0, 1)\}$  is closed since finite sets are closed in a Hausdorff space. But then  $X \setminus \{(0, 0, 1)\}$  is the complement of a closed set, and thus by definition is open. Hence,  $\mathcal{U}$  is an open subset of  $S^2$ . We can construct our open mapping  $\phi : \mathcal{U} \rightarrow \mathbb{R}^2$  by using the *stereographic projection*:

$$\phi(x, y, z) = \left( \frac{x}{1-z}, \frac{y}{1-z} \right) \quad \forall_{(x,y,z) \in S^2 \setminus \{(0,0,1)\}} \quad (37.3.3)$$

This function is continuous and bijective, and has a continuous inverse:

$$\phi^{-1}(X, Y) = \left( \frac{2X}{1+X^2+Y^2}, \frac{2Y}{1+X^2+Y^2}, \frac{-1+X^2+Y^2}{1+X^2+Y^2} \right) \quad (37.3.4)$$

Therefore  $\phi$  is a homeomorphism from  $\mathcal{U}$  to  $\mathbb{R}^2$ , and is thus necessarily a continuous injective open mapping. This chart  $(\mathcal{U}, \phi)$  is actually a chart for any point that is not the north pole. If we flip this around and do the stereographic projection about the south pole, we'll obtain a chart that contains the north pole. Thus the sphere can be covered by two charts.

**Example 37.3.3** Another example of a chart on the sphere is the *orthographic projection*. The stereographic projection is obtained by placing an observer at one of the poles, and drawing a straight line from the observer to a point on the sphere. This point then gets mapped to  $\mathbb{R}^2$  by finding where this line intersects

the  $xy$  plane. The orthographic is obtained in a similar manner by placing the observer at *infinity*. There is now the problem that such a straight line intersects the sphere twice: Once on the top and once on the bottom. That is, if  $(x, y, z)$  lies on this line, then so does  $(x, y, -z)$ . To form our chart we must consistently choose either the top or the bottom. Let  $\mathcal{U}$  be defined as follows:

$$\mathcal{U} = \{(x, y, z) \in S^2 \mid z > 0\} \quad (37.3.5)$$

Define the orthographic projection by:

$$\phi(x, y, z) = (x, y) \quad (37.3.6)$$

This is a homeomorphism between  $\mathcal{U}$  and the open unit disc  $B_1(0)$ . The inverse function is:

$$\phi^{-1}(X, Y) = (X, Y, \sqrt{1 - X^2 - Y^2}) \quad (37.3.7)$$

where we unambiguously take the positive square root. This is a chart for any point in the upper hemisphere (the equator is not included). To completely cover the sphere using orthographic projections requires 6 charts ( $\pm x, \pm y, \pm z$ , our open set  $\mathcal{U}$  corresponding to  $+z$ ).

Equivalently, one could say that  $\phi$  is a homeomorphism between  $\mathcal{U}$  and its image under  $\phi$ . That is,  $\phi(\mathcal{U}) \subseteq \mathbb{R}^n$  is homeomorphic to  $\mathcal{U}$  and  $\phi : \mathcal{U} \rightarrow \phi(\mathcal{U})$  is a homeomorphism.

**Theorem 37.3.1.** *If  $(X, \tau)$  is a topological space, if  $n \in \mathbb{N}$ , if  $(\mathcal{U}, \phi)$  is a chart in  $(X, \tau)$ , and if  $\mathcal{V} = \phi(\mathcal{U})$  is the image of  $\mathcal{U}$  under  $\phi$ , then  $\phi : \mathcal{U} \rightarrow \mathcal{V}$  is a homeomorphism.*

*Proof.* For since  $(\mathcal{U}, \phi)$  is a chart,  $\phi$  is injective, continuous, and an open mapping. But since  $\mathcal{V}$  is the image of  $\mathcal{U}$  under  $\phi$ , the restriction  $\phi : \mathcal{U} \rightarrow \mathcal{V}$  is therefore surjective. But then  $\phi : \mathcal{U} \rightarrow \mathcal{V}$  is both injective and surjective, and is therefore bijective. But since  $\phi$  is an open mapping,  $\phi^{-1}$  is continuous. But then  $\phi$  is a continuous bijection with a continuous inverse and is thus a homeomorphism.  $\square$

### Definition 37.3.2: Locally Euclidean Space

A locally Euclidean space is a topological space  $(X, \tau)$  such that for all  $x \in X$  there exists an  $n \in \mathbb{N}$  and an  $n$  dimensional chart  $(\mathcal{U}, \phi)$  of  $(X, \tau)$  such that  $x \in \mathcal{U}$ .

This is the most general setting for one to define a manifold. The dimension of a locally Euclidean space need not be constant, for we can consider the disjoint

union of a sphere with a line. Thus there will be 2 dimensional points and 1 dimensional points. Dimension is, however, a locally constant property. In particular, for any fixed point  $x \in X$  there is only one  $n \in \mathbb{N}$  such that  $x$  is locally like  $\mathbb{R}^n$ . If in addition the space  $(X, \tau)$  is connected, then there is only one unambiguous number  $n \in \mathbb{N}$  such that every point  $x \in X$  is locally like  $\mathbb{R}^n$ . This allows us to define dimension.

**Theorem 37.3.2.** *If  $X$  and  $Y$  are sets, if  $\mathcal{U} \subseteq X$ , if  $f : X \rightarrow Y$  is a function, if  $f|_{\mathcal{U}} : \mathcal{U} \rightarrow f(\mathcal{U})$  is the restriction of  $f$  to  $\mathcal{U}$ , and if  $\mathcal{V} \subseteq f(\mathcal{U})$ , then:*

$$f|_{\mathcal{U}}^{-1}(\mathcal{V}) = \mathcal{U} \cap f^{-1}(\mathcal{V}) \quad (37.3.8)$$

*Proof.* For if  $x \in f|_{\mathcal{U}}^{-1}(\mathcal{V})$ , then there is an  $x \in \mathcal{U}$  such that  $f|_{\mathcal{U}}(x) \in \mathcal{V}$ . But for all  $x \in \mathcal{U}$  it is true that  $f|_{\mathcal{U}}(x) = f(x)$ , and thus  $f(x) \in \mathcal{V}$ . But if  $f(x) \in \mathcal{V}$ , then  $x \in f^{-1}(\mathcal{V})$ . Thus  $x \in \mathcal{U}$  and  $x \in f^{-1}(\mathcal{V})$ , and therefore  $x \in \mathcal{U} \cap f^{-1}(\mathcal{V})$ . So we obtain:

$$f|_{\mathcal{U}}^{-1}(\mathcal{V}) \subseteq \mathcal{U} \cap f^{-1}(\mathcal{V}) \quad (37.3.9)$$

In the other direction, if  $x \in \mathcal{U} \cap f^{-1}(\mathcal{V})$ , then  $x \in \mathcal{U}$  and  $f(x) \in \mathcal{V}$ . But if  $x \in \mathcal{U}$ , then  $f(x) = f|_{\mathcal{U}}(x)$  and therefore  $f|_{\mathcal{U}}(x) \in \mathcal{V}$ . But then  $x \in f|_{\mathcal{U}}^{-1}(\mathcal{V})$ . That is:

$$\mathcal{U} \cap f^{-1}(\mathcal{V}) \subseteq f|_{\mathcal{U}}^{-1}(\mathcal{V}) \quad (37.3.10)$$

From the definition of equality, we are done. □

**Theorem 37.3.3.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, if  $f : X \rightarrow Y$  is a homeomorphism, and if  $\mathcal{U} \subseteq X$ , then the restriction  $f|_{\mathcal{U}} : \mathcal{U} \rightarrow f(\mathcal{U})$  is a homeomorphism between  $(\mathcal{U}, \tau_X|_{\mathcal{U}})$  and  $(f(\mathcal{U}), \tau_Y|_{f(\mathcal{U})})$ , where  $\tau_X|_{\mathcal{U}}$  and  $\tau_Y|_{f(\mathcal{U})}$  are the subspace topologies.*

*Proof.* Since  $f : X \rightarrow Y$  is a homeomorphism it is therefore bijective. But then  $f|_{\mathcal{U}} : \mathcal{U} \rightarrow f(\mathcal{U})$  is bijective. For if not, then it is either not injective or not surjective. But if it is not injective then there exists points  $x_1, x_2 \in \mathcal{U}$  such that  $x_1 \neq x_2$  and  $f|_{\mathcal{U}}(x_1) = f|_{\mathcal{U}}(x_2)$ . But since  $f|_{\mathcal{U}}$  is the restriction of  $f$  to  $\mathcal{U}$ , for all  $x \in \mathcal{U}$  it is true that  $f(x) = f|_{\mathcal{U}}(x)$ . But  $x_1, x_2 \in \mathcal{U}$  and thus  $f(x_1) = f|_{\mathcal{U}}(x_1)$  and similarly for  $x_2$ . Then by the transitivity of equality,  $f(x_1) = f(x_2)$ , a contradiction since  $f$  is injective. Therefore  $f|_{\mathcal{U}}$  is injective. It is surjective by the definition of the image of  $\mathcal{U}$  under  $f$ . Thus,  $f|_{\mathcal{U}} : \mathcal{U} \rightarrow f(\mathcal{U})$  is bijective. If  $\mathcal{V} \subseteq f(\mathcal{U})$  is open in the subspace topology then there is an open

set  $\mathcal{O} \in \tau_Y$  such that  $\mathcal{V} = \mathcal{O} \cap f(\mathcal{U})$ . But then:

$$f|_{\mathcal{U}}^{-1}(\mathcal{V}) = \mathcal{U} \cap f^{-1}(\mathcal{V}) \quad (37.3.11a)$$

$$= \mathcal{U} \cap f^{-1}(f(\mathcal{U}) \cap \mathcal{O}) \quad (37.3.11b)$$

$$= \mathcal{U} \cap (f^{-1}(f(\mathcal{U})) \cap f^{-1}(\mathcal{O})) \quad (37.3.11c)$$

$$= \mathcal{U} \cap (\mathcal{U} \cap f^{-1}(\mathcal{O})) \quad (37.3.11d)$$

$$= (\mathcal{U} \cap \mathcal{U}) \cap f^{-1}(\mathcal{O}) \quad (37.3.11e)$$

$$= \mathcal{U} \cap f^{-1}(\mathcal{O}) \quad (37.3.11f)$$

But  $f$  is continuous and therefore  $f^{-1}(\mathcal{O})$  is an open subset of  $X$ . But then  $\mathcal{U} \cap f^{-1}(\mathcal{O})$  is open in the subspace topology  $\tau_X|_{\mathcal{U}}$ , and therefore  $f|_{\mathcal{U}}$  is continuous. In a similar manner,  $f|_{\mathcal{U}}^{-1}$  is continuous and therefore  $f|_{\mathcal{U}}$  is a homeomorphism.  $\square$

While we were very liberal in our definition of a locally Euclidean space, allowing the set  $\mathcal{V} \subseteq \mathbb{R}^n$  to be any open set, we need not be. A topological space is locally Euclidean if and only if every point has an open neighborhood that is homeomorphic to all of  $\mathbb{R}^n$ .

**Theorem 37.3.4.** *If  $(X, \tau)$  is a topological space,  $x \in X$ , if  $n \in \mathbb{N}$ , and if  $(\mathcal{U}, \phi)$  is an  $n$  dimensional chart in  $(X, \tau)$  such that  $x \in \mathcal{U}$ , then there is an open subset  $\mathcal{V} \in \tau$  such that  $x \in \mathcal{V}$  and  $\mathcal{V}$  is homeomorphic to  $\mathbb{R}^n$ .*

*Proof.* For since  $\phi : \mathcal{U} \rightarrow \mathbb{R}^n$  is an open mapping,  $\phi(\mathcal{U})$  is an open subset of  $\mathbb{R}^n$ . But if  $x \in \mathcal{U}$ , then  $\phi(x) \in \phi(\mathcal{U})$ . But if  $\mathcal{U}$  is open and if  $\phi(x) \in \phi(\mathcal{U})$ , then there is an  $r > 0$  such that the open ball of radius  $r$  about  $\phi(x)$  is contained in  $\phi(\mathcal{U})$ . That is,  $B_r(\phi(x), \mathbb{R}^n) \subseteq \phi(\mathcal{U})$ . But since  $B_r(\phi(x), \mathbb{R}^n) \subseteq \phi(\mathcal{U})$ , we have:

$$\phi^{-1}(B_r(\phi(x), \mathbb{R}^n)) \subseteq \mathcal{U} \quad (37.3.12)$$

But then since  $B_r(\phi(x), \mathbb{R}^n)$  is contained in the image of  $\mathcal{U}$  under  $\phi$ , we have:

$$\phi(\phi^{-1}(B_r(\phi(x), \mathbb{R}^n))) = B_r(\phi(x), \mathbb{R}^n) \quad (37.3.13)$$

But then the restriction of  $\phi$  to  $\phi^{-1}(B_r(\phi(x), \mathbb{R}^n))$  is a homeomorphism from  $\phi^{-1}(B_r(\phi(x), \mathbb{R}^n))$  to  $B_r(\phi(x), \mathbb{R}^n)$ . But  $B_r(\phi(x), \mathbb{R}^n)$  is homeomorphic to  $\mathbb{R}^n$ , and since homeomorphic is an equivalence relation, we have that  $\phi^{-1}(B_r(\phi(x), \mathbb{R}^n))$  is homeomorphic to  $\mathbb{R}^n$ . But then  $\phi^{-1}(B_r(\phi(x), \mathbb{R}^n))$  is an open subset of  $X$  that contains  $x$  and is homeomorphic to  $\mathbb{R}^n$ , completing the proof.  $\square$

**Theorem 37.3.5.** *If  $(X, \tau)$  is a topological space, if  $Y$  is a set with at least two points, and if for all locally constant functions  $f : X \rightarrow Y$  it is true that  $f$  is constant, then  $(X, \tau)$  is connected.*

*Proof.* For suppose not. Then there are non-empty disjoint open subsets  $\mathcal{U}, \mathcal{V} \in \tau$  that partition  $X$ . That is,  $\mathcal{U} \cap \mathcal{V} = \emptyset$  and  $\mathcal{U} \cup \mathcal{V} = X$ . Since  $Y$  has at least two points, there are distinct  $y_1, y_2 \in Y$ . Let  $f : X \rightarrow Y$  be defined as follows:

$$f(x) = \begin{cases} y_1, & x \in \mathcal{U} \\ y_2, & x \in \mathcal{V} \end{cases} \quad (37.3.14)$$

Then  $f$  is locally constant. That is, since  $\mathcal{U}$  and  $\mathcal{V}$  partition  $X$ , for all  $z \in X$  it is true that either  $z \in \mathcal{U}$  or  $z \in \mathcal{V}$ , but not both. Thus suppose there is a  $z \in X$  such that for all  $\mathcal{O} \in \tau$  such that  $z \in \mathcal{O}$  it is not true that  $f$  is constant on  $\mathcal{O}$ . But either  $z \in \mathcal{U}$  or  $z \in \mathcal{V}$ . If  $z \in \mathcal{U}$ , then for all  $x \in \mathcal{U}$ ,  $f(x) = y_1$  and thus  $f$  is locally constant on  $\mathcal{U}$ . Similarly if  $z \in \mathcal{V}$ , and thus  $f$  is locally constant. It is not constant since  $\mathcal{U}$  and  $\mathcal{V}$  are non-empty, and thus  $f(X) = \{y_1, y_2\}$ , which is not a singleton. A contradiction since by hypothesis for all locally constant functions  $f : X \rightarrow Y$  it is true that  $f$  is constant. Therefore,  $(X, \tau)$  is connected.  $\square$

**Theorem 37.3.6.** *If  $(X, \tau)$  is a connected topological space, if  $Y$  is a set, and if  $f : X \rightarrow Y$  is locally constant, then  $f$  is constant.*

*Proof.* For suppose not. Then there is a function  $f : X \rightarrow Y$  that is locally constant but not constant. But if  $f$  is not constant, then there are distinct  $y_1, y_2 \in Y$  such that  $f^{-1}\{y_1\} \neq \emptyset$  and similarly for  $y_2$ . Define  $\mathcal{U}$  as follows:

$$\mathcal{U} = f^{-1}(\{y_1\}) \quad (37.3.15)$$

Since  $f$  is locally constant, for all  $x \in \mathcal{U}$  there exists an open subset  $\mathcal{V}_x \in \tau$  such that  $x \in \mathcal{V}_x$  and  $f$  is constant on  $\mathcal{V}_x$ . But since  $x \in \mathcal{U}$ , for all  $z \in \mathcal{V}_x$  it must be true that  $f(z) = y_1$ . But then:

$$\mathcal{U} = \bigcup_{x \in \mathcal{U}} \mathcal{V}_x \quad (37.3.16)$$

Since  $\mathcal{U}$  is the union of open sets, it must be open itself. In a similarly manner,  $X \setminus \mathcal{U}$  is open. But  $X \setminus \mathcal{U}$  is non-empty since  $f^{-1}(\{y_2\}) \subseteq X \setminus \mathcal{U}$  and  $f^{-1}(\{y_2\})$  is non-empty. But then  $\mathcal{U}$  and  $X \setminus \mathcal{U}$  are non-empty disjoint open sets that partition  $X$ , and therefore  $(X, \tau)$  is disconnected, a contradiction. Therefore,  $f$  is constant.  $\square$

**Theorem 37.3.7.** *If  $(X, \tau)$  is a topological space, if  $x \in X$ , if  $m, n \in \mathbb{N}$ , if  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  are  $m$  and  $n$  dimensional charts, respectively, and if  $x \in \mathcal{U}$  and  $x \in \mathcal{V}$ , then  $n = m$ .*

*Proof.* For suppose not. Since  $\mathcal{U}$  and  $\mathcal{V}$  are open,  $\mathcal{U} \cap \mathcal{V}$  is open. But  $x \in \mathcal{U}$  and  $x \in \mathcal{V}$ , and thus  $\mathcal{U} \cap \mathcal{V}$  is non-empty. But since  $\varphi : \mathcal{U} \rightarrow \mathbb{R}^m$  and  $\psi : \mathcal{V} \rightarrow \mathbb{R}^n$

are open mappings,  $\psi \circ \varphi^{-1} : \varphi(\mathcal{U} \cap \mathcal{V}) \rightarrow \mathbb{R}^n$  is an open mapping. But since  $\mathcal{U} \cap \mathcal{V}$  is open, and since  $\varphi$  is an open mapping,  $\varphi(\mathcal{U} \cap \mathcal{V})$  is an open subset of  $\mathbb{R}^m$ . But then  $\varphi(\mathcal{U} \cap \mathcal{V})$  is an open subset of  $\mathbb{R}^m$  that is homeomorphic to  $\psi(\mathcal{U} \cap \mathcal{V})$ , which is an open subset of  $\mathbb{R}^n$ . But then by invariance of domain,  $n = m$ .  $\square$

With this we can define a function  $\dim : X \rightarrow \mathbb{N}$  from any locally Euclidean topological space  $(X, \tau)$  called the *dimension* function. Our current goal is to show that this function is locally constant.

**Theorem 37.3.8.** *If  $(X, \tau)$  is a locally Euclidean topological space, if  $\dim : X \rightarrow \mathbb{N}$  is the function such that for all  $x \in X$ ,  $\dim(x)$  is the unique  $n \in \mathbb{N}$  such that there exists a chart  $(\mathcal{U}, \varphi)$  of dimension  $n$  such that  $x \in \mathcal{U}$ , then  $\dim$  is locally constant.*

*Proof.* For suppose not. Then there is a point  $x \in X$  such that for all  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$ , it is not true that  $f$  is constant on  $\mathcal{U}$ . But since  $(X, \tau)$  is locally Euclidean, there is an  $n \in \mathbb{N}$  and a chart  $(\mathcal{U}, \phi)$  of dimension  $n$  such that  $x \in \mathcal{U}$ . But for all  $z \in \mathcal{U}$ ,  $(\mathcal{U}, \phi)$  is a chart of dimension  $n$  that contains  $z$ . But the dimension is pointwise invariant, and thus the dimension is constant on  $\mathcal{U}$ , a contradiction. Thus, dimension is a locally constant function.  $\square$

**Theorem 37.3.9.** *If  $(X, \tau)$  is a connected locally Euclidean topological space, then the dimension is constant.*

*Proof.* For dimension is locally constant, and locally constant functions on a connected topological space are constant.  $\square$

## 37.4 Topological Manifolds

Manifolds are locally Euclidean spaces that have sufficiently nice structure to exclude pathological examples such as the bug-eyed line and the long line. Manifolds are required to have a second-countable topology and must also be Hausdorff.

### Definition 37.4.1: Topological Manifold

A topological manifold of dimension  $n \in \mathbb{N}$  is a Hausdorff locally Euclidean topological space of dimension  $n$  that is second countable.

**Example 37.4.1** Open subsets of  $\mathbb{R}^n$  are  $n$  dimensional manifolds. Moreover, any open subset of an  $n$  dimensional manifold, given the subspace topology, will again be an  $n$  dimensional manifold.

**Example 37.4.2** The graphs of continuous functions  $f : X \rightarrow Y$  from topological manifolds of dimension  $n$  and  $m$ , respectively, are manifolds. That is, if  $\mathcal{U} \subseteq X$  is open, and  $f|_{\mathcal{U}}$  is the restriction of  $f$ , then  $\{(p, f(p)) | p \in \mathcal{U}\}$ , equipped with the subspace topology induced by the product topology on  $X \times Y$ , is a topological manifold of dimension  $n$ . It is Hausdorff and second countable since it is the subspace of such spaces.

### Definition 37.4.2: Atlas

An atlas on a topological space  $(X, \tau)$  is a set  $\mathcal{A}$  such that for all  $z \in \mathcal{U}$  it is true that  $z$  is a chart in  $(X, \tau)$ , and such that for all  $x \in X$  there exists a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}$  such that  $x \in \mathcal{U}$ .

An atlas for a connected space necessarily has a constant and well defined dimension, whereas the by definition of a manifold, any atlas on a manifold must have the same dimension.

## 37.5 Reading From Lee (Chapter 1)

Smoothness cannot be a topological property. To see this, note that the circle and the square are homeomorphic to each other, and thus topology cannot distinguish between the two. They differ in the smooth sense since one has sharp corners (the square) and the other does not. We can be explicit, let  $\varphi : \mathbb{R}^2 \setminus \{0\} \rightarrow S^1$  be the mapping:

$$\varphi(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad (37.5.1)$$

The restriction of  $\varphi$  to the square will then be a homomorphism (Fig. 37.9).

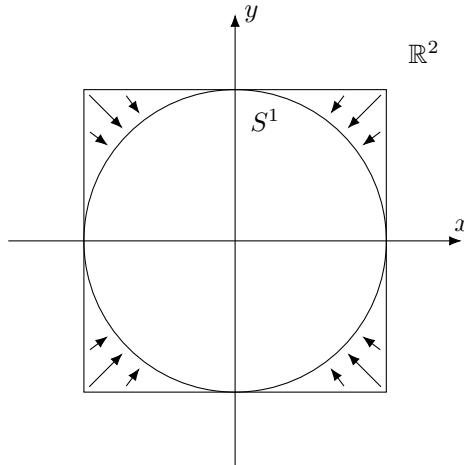


Fig. 37.9: Homemorphism from the Square to the Circle

**Theorem 37.5.1.** *A topological space  $(X, \tau)$  is a topological manifold if and only if it is Hausdorff, second countable, and for all  $p \in X$  there is an open subset  $\mathcal{U} \subseteq X$  such that  $p \in \mathcal{U}$  and  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^n$ .*

*Proof.* If every point in  $X$  is contained in an open set  $\mathcal{U}$  that is homeomorphic to all of  $\mathbb{R}^n$ , then it is a topological manifold since  $\mathbb{R}^n$  is an open subset of  $\mathbb{R}^n$ . Going the other direction, suppose  $(X, \tau)$  is a topological manifold and  $p \in X$ . Then there is an open subset  $\mathcal{U}$  such that  $p \in \mathcal{U}$  and  $\mathcal{U}$  is homeomorphic to an open subset of  $\mathbb{R}^n$ . That is, there is a  $\mathcal{V} \subseteq \mathbb{R}^n$  and a homomorphism  $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ . But since  $p \in \mathcal{U}$  and  $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ , it is then true that  $\varphi(p) \in \mathcal{V}$ . But  $\mathcal{V}$  is open, and thus there is an  $r > 0$  such that  $B_r^{\mathbb{R}^n}(\varphi(p)) \subseteq \mathcal{V}$ . That is, there is an  $r > 0$  such that the  $r$  ball centered about  $\varphi(p)$  is contained in  $\mathcal{V}$ . But  $\varphi$  is a homeomorphism, and thus  $\varphi^{-1}$  is a homeomorphism. And the restriction of a homeomorphism to a subspace is again a homeomorphism in the subspace topologies. Thus,  $\varphi^{-1}|_{B_r^{\mathbb{R}^n}(\varphi(p))}$  is a homeomorphism from  $B_r^{\mathbb{R}^n}(\varphi(p))$  to its image. But its image is simply  $\varphi^{-1}(B_r^{\mathbb{R}^n}(\varphi(p)))$ , which is open since  $\varphi$  is continuous. Thus, there is an open subset of  $(X, \tau)$  that contains  $p$  which is homeomorphic to an open ball in  $\mathbb{R}^n$ . But any open ball in  $\mathbb{R}^n$  is homeomorphic to all of  $\mathbb{R}^n$ , and since homeomorphic is an equivalence relation, there is thus an open subset containing  $p$  that is homeomorphic to all of  $\mathbb{R}^n$ .  $\square$

**Example 37.5.1** All of  $\mathbb{R}^n$  is itself a topological  $n$  dimensional manifold. It is Hausdorff since it is a metric space, given two points a distance  $d$  away one can choose open balls of radius  $d/4$  about each point as disjoint open neighborhoods, ensuring the Hausdorff property. Moreover, it is second countable since the set of all balls with rational radii centered about rational points (elements of  $\mathbb{Q}^n$ )

form a countable basis for the topology on  $\mathbb{R}^n$ . Hence it is Hausdorff and second countable. That any point  $\mathbf{x} \in \mathbb{R}^n$  is contained in an open subset homeomorphic to  $\mathbb{R}^n$  can be seen by letting  $\mathcal{U} = \mathbb{R}^n$ . That is, take as the open subset the entire space. Therefore,  $\mathbb{R}^n$  is a topological manifold.

**Example 37.5.2** In the previous example we implicitly thought of  $\mathbb{R}^n$  as being equipped with the standard Euclidean topology. This is the topology that arises from the Pythagorean distance function:

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{k \in \mathbb{Z}_n} (x_k - y_k)^2} \quad (37.5.2)$$

What if we change the metric? Any metric on  $\mathbb{R}^n$  will preserve the Hausdorff property, since all metric spaces are Hausdorff, however the discrete metric loses the second countability property, and hence cannot be considered a topological manifold (although it is still locally *zero* dimensional). What about the metrics that arise from the  $\|\cdot\|_p$  norm:

$$d_p(\mathbf{x}, \mathbf{y}) = \left( \sum_{k \in \mathbb{Z}_n} (x_k - y_k)^p \right)^{1/p} \quad (37.5.3)$$

For  $1 \leq p \leq \infty$  this will be second countable and Hausdorff and since these metrics are equivalent, they will induce the same topology, and hence will be topological manifolds of dimension  $n$ . However the case of  $n = 1$  and  $n = \infty$  are somewhat peculiar. An open ball in the  $\|\cdot\|_1$  norm looks like a diamond, whereas in  $\|\cdot\|_\infty$  the open balls are squares. So while all of these metrics give rise to the same topological structure, they create a different *smooth* structure on the space. When one considers  $\mathbb{R}^n$  it is standard to assume the Euclidean  $\|\cdot\|_2$  norm (and the structure induced by it) is being considered. That is, open balls  $B_r^{(\mathbb{R}^n, \|\cdot\|_2)}(\mathbf{x})$  are just the interiors of  $n$  dimensional spheres.

The Hausdorff condition is to exclude bizarre spaces such as the bug-eyed line and the branching line. The second countability criterion ensures the space won't be massive. For example, the uncountable disjoint union of spheres will be Hausdorff and locally Euclidean, but not second countable (but it will be first countable). This space isn't too horrible since it's metrizable (in the disjoint union topology). A more pathological example is the long line, occasionally denoted  $[0, \omega_1]$ , defined as  $[0, 1] \times \omega_1$  with the dictionary or lexicographic ordering. Here  $\omega_1$  denotes the first uncountable ordinal. This space is not metrizable, but it is Hausdorff and locally Euclidean (it is locally like  $\mathbb{R}$ ), but it is not second countable. One of the nice theorems from topology is the Smirnov metrization theorem which states that a topological space is metrizable if and only if it is locally metrizable, Hausdorff, and paracompact. Every locally Euclidean space is automatically locally metrizable since we can steal the Euclidean metric on small enough neighborhoods about every point.

The second countability condition allows one to prove paracompactness, and thus every topological manifold is automatically metrizable by the Smirnov theorem. This is perhaps one justification for the second countability condition. Note that the metric that we can place on the topological manifold  $(X, \tau)$  may not give us much geometrical information, but simply allows us to apply the results of the theory of metric spaces without much thought. For example, topological manifolds must be regular, normal, perfectly normal Hausdorff, we may apply Urysohn's lemma and Tietze extension theorem should the need arise, and more. Moreover, from an analytical point of view, we have that functions are continuous if and only if there are sequentially continuous (i.e.  $a_n \rightarrow x$  implies  $f(a_n) \rightarrow f(x)$ ) and compact if and only if sequentially compact. So while the metric does not necessarily have much to do with the structure we care about, its existence gives us this plethora of data for free.

**Example 37.5.3** Consider the sphere  $S^2$ . We can place a natural metric on this by defining the distance to be the length of the shortest curve connecting two points. One can prove the shortest curve exists and that it lies on the great circle between the two points (the circle containing the two points with the center of the sphere as the origin). This function is symmetric, positive definite, and obeys the triangle inequality and is therefore a metric. However, if we were to take two disjoint spheres, how can we metrize this? There's one way, defining:

$$d(x, y) = \begin{cases} d_1(x, y), & x, y \in S_1^2 \\ d_2(x, y), & x, y \in S_2^2 \\ 4, & \text{Otherwise} \end{cases} \quad (37.5.4)$$

where  $S_i^2$  is the  $i^{th}$  sphere and  $d_i$  is the *geodesic* metric described above on the respective spheres. For the case of when the points lie in separate sphere we've chosen the constant 4 since this is greater than the furthest two points can be on the unit sphere (which is  $\pi$ , half of the perimeter  $2\pi$ ). Doing this ensures  $d$  is a metric. So while intuitively we want to think of the disjoint union of two spheres as a two dimensional topological space, the metric that can induce the topology seems somewhat irrelevant to the picture we have in mind.

Before continuing it is perhaps important to note that paracompactness does not imply second countable. Again, the disjoint union of spheres serves as an example.

**Definition 37.5.1** A chart of dimension  $n \in \mathbb{N}$  in a topological space  $(X, \tau)$  is an open set  $\mathcal{U} \in \tau$  and a function  $\varphi : \mathcal{U} \rightarrow \mathbb{R}^n$  such that  $\varphi$  is continuous, injective, and an open mapping. A chart is denoted  $(\mathcal{U}, \varphi)$ .

Equivalently, we could say that  $\varphi$  is a homeomorphism onto its image or that  $\varphi : \mathcal{U} \rightarrow \mathcal{V}$  is a homeomorphism where  $\mathcal{V} \subseteq \mathbb{R}^n$  is open. The definition adopted

here is so that all of the details about a chart  $(\mathcal{U}, \varphi)$  are embedded in the notation. That is, we need not add a  $\mathcal{V}$  nor talk about  $\varphi^{-1}$ .

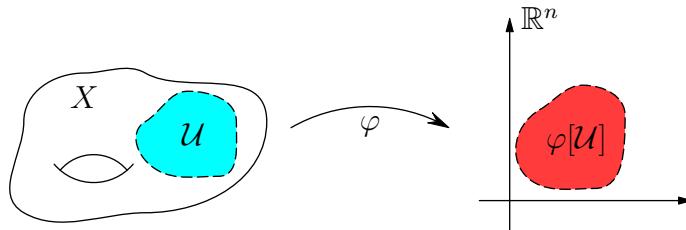


Fig. 37.10: A Chart in a Topological Space

Another equivalent definition of topological manifold is that for all  $x \in X$  there is a chart  $(\mathcal{U}, \varphi)$  such that  $x \in \mathcal{U}$ .

**Definition 37.5.2** The coordinate functions of a chart  $(\mathcal{U}, \varphi)$  of dimension  $n \in \mathbb{N}$  in a topological space  $(X, \tau)$  are the function  $x^i : X \rightarrow \mathbb{R}$  defined by  $x^i = \varphi \circ \pi_i$ , where  $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$  is the  $i^{th}$  projection mapping.

With this definition it is often common to write the image of a point  $p \in \mathcal{U}$  by:

$$\varphi(p) = (x^1(p), x^2(p), \dots, x^{n-1}(p), x^n(p)) \quad (37.5.5)$$

One should note that  $x^2$  does not denote the square of  $x$ , and this may cause confusion. It is the composition of  $\varphi$  with the map  $\pi_2 : \mathbb{R}^n \rightarrow \mathbb{R}$  which simply selects the second coordinate.

**Example 37.5.4** The graph of a continuous function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is an  $n$  dimensional manifold. That is, recalling from set theory that a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is a subset  $f \subseteq \mathbb{R}^n \times \mathbb{R}^m$ , if we endow  $f$  with the subspace topology of  $\mathbb{R}^{n+m}$ , then it will be an  $n$  dimensional manifold. To see this, for  $(\mathbf{x}, \mathbf{y}) \in f$ , let  $\pi_1 : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$  be the mapping  $\pi_1(\mathbf{x}, \mathbf{y}) = \mathbf{x}$ . This is continuous, and thus the restriction  $\pi_1|_f : f \rightarrow \mathbb{R}^n$  is continuous in the subspace topology (that is, in the graph of  $f$ ). Moreover,  $\varphi|_f$  is a homeomorphism with continuous inverse  $\varphi|_f^{-1}(\mathbf{x}) = (\mathbf{x}, f(\mathbf{x}))$ . This projection then shows that the graph of  $f$  is homeomorphic to  $\mathbb{R}^n$  itself. If we were to replace  $\mathbb{R}^n$  with an open subset  $\mathcal{U} \subseteq \mathbb{R}^n$ , the claim would still hold.

**Example 37.5.5** The most common non-trivial example of a topological manifold is the sphere  $S^n$ . This is the subset of  $\mathbb{R}^n$  such that  $\|\mathbf{x}\|_2 = 1$ , where  $\|\cdot\|_2$  is the Euclidean norm defined by the Pythagorean formula. That  $S^n$  is a topological manifold can be realized in two ways. The first is called the *orthographic projection*. We take an observer standing on the north pole and then then them off to infinity along the line through the origin and the north pole. From the

observers new perspective, only the top half of the sphere is visible. That is, this projection *at infinity* is the set:

$$\mathcal{U}_{n-1}^+ = \{\mathbf{x} \in S^2 \mid x_{n-1} > 0\} \quad (37.5.6)$$

where  $x_{n-1}$  is the last coordinate (the coordinate that corresponds to the north pole). For  $S^1$  this is just the set:

$$\mathcal{U}_y^+ = \{(x, \sqrt{1-x^2}) \in S^2 \mid x \in [-1, 1]\} \quad (37.5.7)$$

and for  $S^2$  this is:

$$\mathcal{U}_z^+ = \{(x, y, \sqrt{1-x^2-y^2}) \in S^2 \mid x^2 + y^2 \leq 1\} \quad (37.5.8)$$

Since we only see the top half, this does not cover the entire sphere. Moreover it contains the boundary, and is therefore not open. To meet the requirements of a chart we must therefore remove the equator from this projection. Thus, if we wish to cover the entire sphere with such sets we need  $2(n+1)$  charts. That these charts are homeomorphic to  $\mathbb{R}^n$  can be seen by simply using the projection map  $\pi^i : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$  which maps:

$$\pi^i(x_0, x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_{n-1}) = (x_0, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_{n-1}) \quad (37.5.9)$$

This is continuous, and thus its restriction to  $S^n$  is continuous. Moreover, its restriction is an open mapping, and hence the sets  $\text{Int}(\mathcal{U}_i^\pm)$  together with  $\pi^i$  form charts.

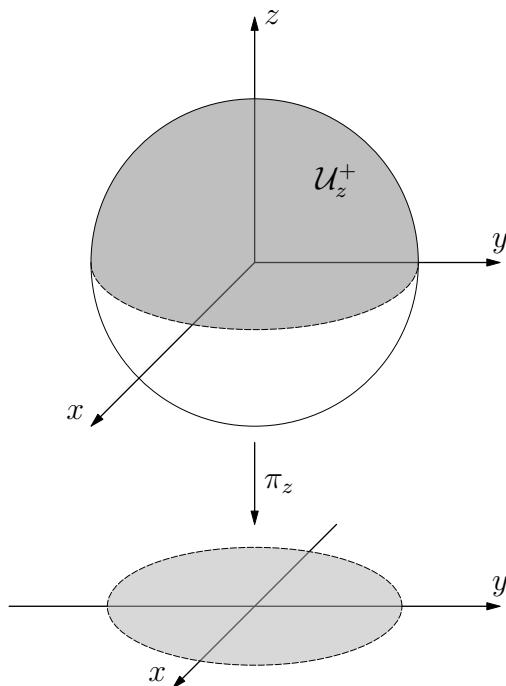


Fig. 37.11: Orthographic Chart for the Sphere

The orthographic projection is something that is somewhat approximated by satellites in space. Satellites at geosynchronous orbit (about 5 Earth radii away) have a projection that is close the orthographic projection, however these satellites can't see the equator. To realize this, note that you cannot see the entire equator. If you were to fly up a little bit, you'd see more of the horizon, and as you ascend you'd have a wider and wider cone of vision, but to see the entire equator would amount to having a *cylinder* of vision. Since a cylinder can be realized by stretching the tip of a cone out to infinity, we see that orthographic projection is well approximated by objects far from the Earth (Fig. 37.12).

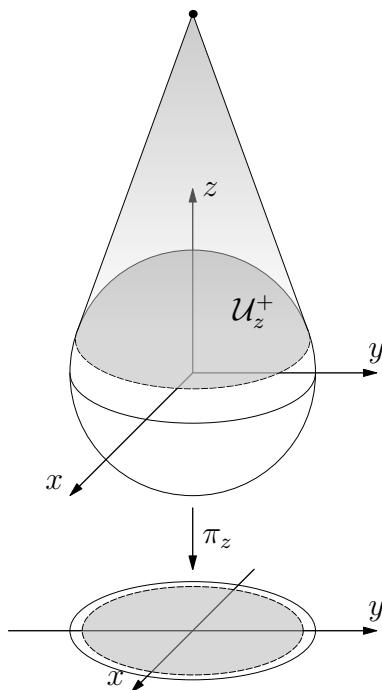


Fig. 37.12: Near-Sided Projection from Geosynchronous Orbit

This near-side projection will then cost us more than  $2(n + 1)$  charts and so we seek a simpler means of showing that  $S^n$  is a manifold. To do this we look at the far-side projections. In this method we put the observer at the exact same spot as in the near side projection, but now we consider everything we *can't* see. That is, imagine the hollow sphere and slicing off what we can see. Once this is done, the entirety of the rest of the sphere is now visible. Let's draw the picture for the case of an observer out at GEO (geosynchronous orbit).

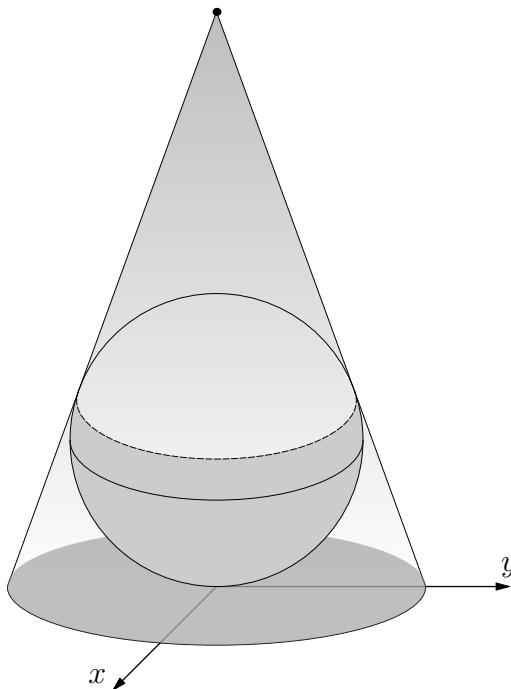


Fig. 37.13: Far-Sided Projection from Geosynchronous Orbit

We can see perhaps pictorially that we now only need two such charts to cover  $\mathbb{S}^2$ , but the formulas are somewhat messy. We've taken the orthographic projection and *fattened* it around the equator.

**Example 37.5.6** The third, perhaps easiest but less intuitive, method of showing that  $\mathbb{S}^n$  is a manifold is using *stereographic* projection. Before, we lifted our observer to infinity and then projected the sphere from this point of view. Now, we keep the observer at the north pole and imagine what they'd see if the sphere was transparent. Drawing a line from the observer to another point on the sphere will then intersect the hyperplane  $x_0 = 0$  once. That is, we use the far-sided projection of an observer standing on the Earth. The near-sided projection would simply be the point you're standing on, and hence the far-sided projection will be everything else. We can use this to describe a chart that covers almost the entirety of  $S^n$ . Note the line from the North pole to itself will be parallel to the plane, and hence will never intersect. We define the stereographic projection function  $\varphi : S^n \setminus \{(0, 0, \dots, 0, 1)\} \rightarrow \mathbb{R}^n$  as follows:

$$\varphi(\mathbf{x}) = \frac{\mathbf{x}}{1 - x_n} \quad (37.5.10)$$

Since a point is closed in a Hausdorff space, the set  $S^n$  minus the north pole

is open. Hence this open set is homeomorphic to  $\mathbb{R}^n$  and covers all but one point of the sphere. The mapping  $\varphi$  is a homemorphism since its inverse is continuous:

$$\varphi^{-1}(\mathbf{X}) = \left( \frac{2X_0}{1 + \|\mathbf{X}\|^2}, \dots, \frac{2X_{n-1}}{1 + \|\mathbf{X}\|^2}, \frac{1 - \|\mathbf{X}\|^2}{1 + \|\mathbf{X}\|^2} \right) \quad (37.5.11)$$

To make  $S^n$  a topological manifold of dimension  $n$  we need only add one more stereographic projection about any other point. The south pole is usually the most fitting choice, and so we have that  $S^n$  is a manifold that can be covered with two charts.

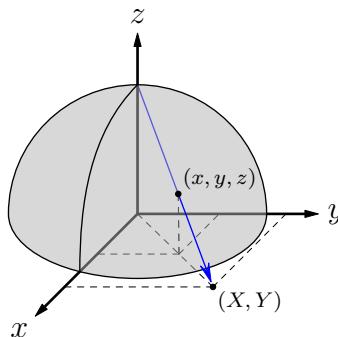


Fig. 37.14: Stereographic Projection of the Sphere

This may naturally lead one to ask the question *what's the minimum number of charts needed to cover a manifold?* Ostrand's theorem from topology tells us the answer is  $n+1$ . If the space is oriented and 2 dimensional, Morse theory can be used to reduce that to 1 or 2 (The only case where 1 is occurs is if the space is homeomorphic to  $\mathbb{R}^2$ ).

The next example to discuss are the real projective spaces, commonly denoted  $\mathbb{RP}^n$ .

**Example 37.5.7** There are three common and equivalent ways to think of  $\mathbb{RP}^2$ , the real projective plane. Much the way the torus and the Klein bottle can be realized by an equivalence relation on the unit square, so can the real projective plane. The more common method is to consider the sphere  $S^2$  and the following equivalence relation:  $\mathbf{x}R\mathbf{y}$  if and only if  $\mathbf{y} = \pm -\mathbf{x}$ . The quotient space  $S^2/R$  is then the real projective plane. Such pairing of antipodal points generalizes to all  $n$  and so we may discuss the more general real projective space  $\mathbb{RP}^n$ . The last method is by considering the set of all lines through the origin. We can develop a projection map  $\mathbb{R}^{n+1} \setminus \{0\}$  into  $\mathbb{RP}^n$  by sending  $\mathbf{x}$  to the line it spans. That is,  $\mathbf{x} \mapsto t\mathbf{x}$ , where  $t$  is a real variable. This mapping is indeed

a projection mapping and thus we can endow  $\mathbb{RP}^n$  with the quotient topology. For all  $i \in \mathbb{Z}_{n+1}$ , let  $\mathcal{U}_i$  be the subset of  $\mathbb{R}^{n+1}$  where  $x_i \neq 0$ . Since lines are closed in  $\mathbb{R}^{n+1}$ , the complement is open and hence  $\mathcal{U}_i$  is open. Moreover, it is saturated with respect to the quotient map:

$$\pi^{-1}(\pi(\mathcal{U}_i)) = \mathcal{U}_i \quad (37.5.12)$$

and hence the restriction of  $\pi$  to  $\mathcal{U}_i$  is again a quotient map. Let  $\mathcal{V}_i = \pi|_{\mathcal{U}_i}(\mathcal{U}_i)$  and define  $\varphi : \mathcal{V}_i \rightarrow \mathbb{R}^n$  by:

$$\varphi_i([\mathbf{x}]) = \frac{\pi^i(\mathbf{x})}{x_i} \quad (37.5.13)$$

This mapping is well defined, for if  $\mathbf{y} \in [\mathbf{x}]$  then there is a non-zero  $t \in \mathbb{R}$  such that  $\mathbf{y} = t\mathbf{x}$ . But then:

$$\varphi([\mathbf{y}]) = \frac{\pi^i(t\mathbf{x})}{tx_i} = \frac{t\pi^i(\mathbf{x})}{tx_i} = \frac{\pi^i(\mathbf{x})}{x_i} = \varphi([\mathbf{x}]) \quad (37.5.14)$$

Moreover, since  $x_i$  is non-zero for all  $\mathbf{x} \in \mathcal{U}_i$ , we are not dividing by zero. But  $\varphi_i \circ \pi$  is continuous, and thus  $\varphi_i$  is continuous (because  $\pi$  is a quotient mapping). Moreover, it is a homeomorphism since it's inverse:

$$\varphi_i^{-1}(\mathbf{x}) = [x_1, \dots, x_{i-1}, 1, x_i, \dots, x_n] \quad (37.5.15)$$

is continuous. Hence,  $\mathbb{RP}^n$  is locally Euclidean.

**Theorem 37.5.2.**  $\mathbb{RP}^n$  is Hausdorff.

*Proof.* For let  $[\mathbf{x}]$  and  $[\mathbf{y}]$  be distinct elements of  $\mathbb{RP}^n$ . For the case of  $n > 2$  there must be a  $\mathcal{U}_i$  such that  $\pi(\mathcal{U}_i)$  contains both points. Since  $\pi(\mathcal{U}_i)$  is locally Euclidean, it is Hausdorff, and therefore  $[\mathbf{x}]$  and  $[\mathbf{y}]$  can be separated by open subsets of  $\pi(\mathcal{U}_i)$ . For the case of  $n = 2$ , the only scenario where  $[\mathbf{x}]$  and  $[\mathbf{y}]$  do not lie in the same  $\mathcal{U}_i$  for some  $i$  is where  $[\mathbf{x}] = [(1, 0)]$  and  $[\mathbf{y}] = [(0, 1)]$ . In this case, let  $\mathcal{U}$  be the complement of the line spanned by  $(1, 1)$ . Then  $\pi(\mathcal{U})$  contains both  $[\mathbf{x}]$  and  $[\mathbf{y}]$ , and by the preceding arguments will be locally Euclidean. Being locally Euclidean, these two points can be separated by disjoint open sets. Thus,  $\mathbb{RP}^n$  is Hausdorff.  $\square$

**Theorem 37.5.3.**  $\mathbb{RP}^n$  is second countable.

*Proof.* We have shown that  $\mathbb{RP}^n$  can be covered by finitely many open subsets each of which is homeomorphic to  $\mathbb{R}^n$ . But  $\mathbb{R}^n$  is second countable, and second countability is preserved by homeomorphisms and hence each of these open subsets is second countable. But then  $\mathbb{RP}^n$  is the finite union of second countable open subspaces, and is hence second countable.  $\square$

The open subspace part is crucial. One might think that if one has a countable collection of subspaces that cover the space, each of which being second countable, then the whole space is second countable, but this is not true. For example, consider the *infinite bouquet*. Take  $\mathbb{R}$  with the equivalence relation  $R$  defined by  $xRy$  if and only if  $x = y$  or  $x$  and  $y$  are integers. The quotient space collapses all of  $\mathbb{Z}$  down to a point, and the result looks like infinitely many rings connected at the origin. The point  $[\mathbb{Z}]$  has no neighborhood that has a countable basis. Thus not only is this space not second countable, it's not even first countable (but it is separable).

**Theorem 37.5.4.**  $\mathbb{RP}^n$  is compact.

*Proof.* For the restriction  $\pi|_{S^n} : S^n \rightarrow \mathbb{RP}^n$  is surjective. Given  $[\mathbf{x}] \in \mathbb{RP}^n$ , let  $\mathbf{y} = \mathbf{x}/\|\mathbf{x}\|_2$ . Then  $\mathbf{y} \in S^n$  and  $\pi(\mathbf{y}) = [\mathbf{x}]$ . Thus  $\pi|_{S^n}$  is a continuous surjective map. But  $S^n$  is compact by the Heine-Borel theorem, and thus the image of  $S^n$  under  $\pi|_{S^n}$  is compact. Therefore,  $\mathbb{RP}^n$  is compact.  $\square$

**Theorem 37.5.5.** If  $M_1, M_2$  are topological manifolds, then  $M_1 \times M_2$  is a topological manifold.

*Proof.* For the product of Hausdorff spaces is Hausdorff, and the product of second countable spaces is second countable. Thus, we must show it is locally Euclidean. The charts of the form  $(U \times V, \varphi \times \psi)$  with  $\varphi \times \psi : U \times V \rightarrow \mathbb{R}^{n_1+n_2}$  form an atlas on  $M_1 \times M_2$ , and hence it is a topological manifold.  $\square$

### Definition 37.5.3: Hemicompact

A hemicompact topological space is a topological space  $(X, \tau)$  such that there exists a sequence  $K : \mathbb{N} \rightarrow \mathcal{P}(X)$  such that for all  $n \in \mathbb{N}$  it is true that  $K_n$  is compact, and such that for every compact subset  $C \subseteq X$  there exists an  $n \in \mathbb{N}$  such that  $C \subseteq K_n$ .

**Example 37.5.8** The real line is hemicompact. Take  $K_n = [-n, n]$  and apply the Heine-Borel theorem.

**Theorem 37.5.6.** Compact spaces are hemicompact.

*Proof.* Duh.  $\square$

**Theorem 37.5.7.** Hemicompact spaces are  $\sigma$  compact.

*Proof.* It suffices to show that the union of the  $K_n$  is the entire space. But if not then there is an  $x \in X$  not contained in  $\bigcup K_n$ . But  $\{x\}$  is a compact set, and thus there is an  $n \in \mathbb{N}$  such that  $\{x\} \subseteq K_n$ , and hence  $x \in \bigcup K_n$ . Then, the space is  $\sigma$  compact.  $\square$

**Theorem 37.5.8.** *If  $(X, \tau)$  is hemicompact and first countable, then it is locally compact.*

*Proof.* For let  $\mathcal{U}_n$  be a countable neighborhood basis, and  $K_n$  be a sequence of compact sets that contain every compact subset eventually. Let:

$$\mathcal{V}_n = \bigcap_{k \in \mathbb{Z}_n} \mathcal{U}_k \quad (37.5.16)$$

and let:

$$C_n = \bigcup_{k \in \mathbb{Z}_n} K_k \quad (37.5.17)$$

Suppose for all  $n \in \mathbb{N}$  it is true that  $\mathcal{V}_n \not\subseteq C_n$ . Then there exists  $a_n \in \mathcal{V}_n$  such that  $a_n \notin C_n$ . But then  $a_n \rightarrow x$ , and hence  $\{a_n\} \cup \{x\}$  is a compact subset of  $X$ , and hence there exists  $N \in \mathbb{N}$  such that  $K_N$  contains this. But  $K_N \subseteq C_N$ , and thus all of the  $a_n$  are contained in  $C_N$ , a contradiction. Thus, there exists  $N \in \mathbb{N}$  such that  $\mathcal{V}_N \subseteq C_N$ . But the finite intersection of open neighborhoods of  $x$  is an open neighborhood, and the finite union of compact sets is compact, and hence  $K_n$  is a compact neighborhood of  $x$ . Thus,  $X$  is locally compact.  $\square$

**Example 37.5.9** First countable and  $\sigma$  compact do not imply locally compact. The topologists sine curve serves as a counter-example. It is first countable since it is metrizable, it is  $\sigma$  compact for let:

$$K_n = \{(x, \sin(x^{-1})) \in \mathbb{R}^2 \mid \frac{1}{n} \leq x \leq 1\} \quad (37.5.18)$$

These are all compact by Heine-Borel, and the union covers all but the origin  $(0, 0)$ . Since the origin is a point, it is compact, thus adjoining  $K_0 = \{0, 0\}$  gives us a sequence of compact sets that covers the whole space. This is not a sequence that makes the space hemicompact, in indeed there is no such sequence. To see this, note that the points that evaluate to zero under  $\sin(x^{-1})$ , together with the origin  $(0, 0)$ , form a closed bounded subset and is thus compact, but no  $K_n$  covers all of this.

**Example 37.5.10** Not every  $\sigma$  compact space is hemicompact. The rationals  $\mathbb{Q}$  are  $\sigma$  compact since  $\{q_n\}$  forms a countable set of compact sets whose union is all of  $\mathbb{Q}$ . However,  $\mathbb{Q}$  is not hemicompact. For suppose  $K_n$  is a sequence that contains every compact subset of  $\mathbb{Q}$ , eventually. A subset  $C \subseteq \mathbb{Q}$  is compact if and only if it is complete and bounded. Since  $[-1/n, 1/n]$  is not complete in  $\mathbb{Q}$ , there exists  $a_n \in [-1/n, 1/n]$  such that  $a_n \notin K_n$ . Let  $C = \{a_n\} \cup \{0\}$ . This is compact, but not contained in any  $K_n$  since  $a_n \notin K_n$ . Thus,  $\mathbb{Q}$  is not hemicompact.

### 37.5.1 The Topology of Manifolds

**Theorem 37.5.9.** *If  $n \in \mathbb{N}$  and if  $(\mathbb{R}^n, \tau_{\mathbb{R}^n})$  is the usual Euclidean space, then there exists a countable basis  $\mathcal{B}$  of  $\tau_{\mathbb{R}^n}$  such that for all  $\mathcal{U} \in \mathcal{B}$ ,  $\mathcal{U}$  is diffeomorphic to  $\mathbb{R}^n$  and precompact.*

*Proof.* For let  $\mathcal{B}$  be defined by:

$$\mathcal{B} = \left\{ B_q^{\mathbb{R}^n}(\mathbf{x}) \mid \mathbf{x} \in \mathbb{Q}^n \text{ and } q \in \mathbb{Q}^+ \right\} \quad (37.5.19)$$

□

**Theorem 37.5.10.** *If  $n \in \mathbb{N}$  and if  $(X, \tau, \mathcal{A})$  is an  $n$  dimensional topological manifold, then there is a countable basis  $\mathcal{B}$  of  $\tau$  such that for all  $\mathcal{U} \in \mathcal{B}$  there is a homeomorphism  $\varphi : \mathcal{U} \rightarrow \mathbb{R}^n$ .*

*Proof.* For all  $x \in X$  there is a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}$  such that  $x \in \mathcal{U}$  and  $\varphi : \mathcal{U} \rightarrow \mathbb{R}^n$  is a homeomorphism. But then the collection of all such charts forms an open cover. But manifolds are second countable, and second countable topological spaces are Lindelöf. Thus given a cover there is a countable subcover. By the previous theorem, each of the  $\mathcal{U}_n$  has a countable basis of precompact sets by looking at  $\varphi^{-1}$ . But if  $\mathcal{U}_n$  is a countable set of open subsets that cover  $X$ , each of which has a countable basis, then the union of these basis is a basis for the entire space. But for any  $\mathcal{V} \subseteq \mathcal{U}_n$  in the basis,  $\text{Cl}_{\mathcal{U}_n}(\mathcal{V})$  is compact in  $\mathcal{U}_n$ . But then  $\text{Cl}_{\mathcal{U}_n}(\mathcal{V})$  is compact in  $X$ . But  $X$  is Hausdorff, and therefore if  $\text{Cl}_{\mathcal{U}_n}(\mathcal{V})$  is compact, then it is closed. But then  $\text{Cl}_{\mathcal{U}_n}(\mathcal{V})$  is equal to  $\text{Cl}_\tau(\mathcal{V})$ , and therefore  $\mathcal{V}$  is precompact in  $X$ . □

This can be summarized by claiming that every topological manifold has a countable basis of precompact coordinate balls.

**Theorem 37.5.11.** *If  $(X, \tau, \mathcal{A})$  is a topological manifold, then it is locally path-connected.*

Thus, a manifold is connected if and only if it is path connected.

**Theorem 37.5.12.** *If  $(X, \tau, \mathcal{A})$  is a topological manifold, then the set of connected components of  $X$  is countable.*

*Proof.* For since  $X$  is locally path connected, the connected components are open. But then the connected components form a cover of  $X$ . But manifolds are second countable, and hence Lindelöf so there is a countable subcover. But the elements of the cover are disjoint, and hence the subcover is equal to the original cover. Hence, there are only countably many connected components. □

**Theorem 37.5.13.** *If  $(X, \tau, \mathcal{A})$  is a topological manifold, then it is locally compact.*

*Proof.* For there exists a basis  $\mathcal{B}$  of  $\tau$  of precompact coordinate balls. But then for all  $x \in X$  there is a precompact  $\mathcal{U} \in \mathcal{B}$  such that  $x \in \mathcal{U}$ . But then  $\text{Cl}_\tau(\mathcal{U})$  is a compact neighborhood of  $x$ , and hence  $X$  is locally compact.  $\square$

Def paracompact, locally finite, point finite, metacompact. Def refinement.

**Theorem 37.5.14.** *If  $(X, \tau)$  is a topological space, if  $\mathcal{O} \subseteq \mathcal{P}(X)$  is locally finite, then the set  $\text{Cl}_\tau(\mathcal{O})$  defined by:*

$$\text{Cl}_\tau(\mathcal{O}) = \{ \text{Cl}_\tau(\mathcal{U}) \mid \mathcal{U} \in \mathcal{O} \} \quad (37.5.20)$$

*is locally finite as well.*

**Theorem 37.5.15.** *If  $(X, \tau)$  is a topological space, and if  $\mathcal{O} \subseteq \mathcal{P}(X)$  is locally finite, then:*

$$\text{Cl}_\tau(\bigcup \mathcal{O}) = \bigcup \text{Cl}_\tau(\mathcal{O}) \quad (37.5.21)$$

**Theorem 37.5.16.** *If  $(X, \tau)$*

### Theorem 37.5.17: Paracompactness of Manifolds

If  $(X, \tau)$  is a topological manifold, then it is paracompact. That is, if  $(X, \tau)$  is a second countable, Hausdorff, locally Euclidean topological space, then it is paracompact.

*Proof.* For if  $(X, \tau)$  is locally Euclidean, then it is locally compact. But locally compact Hausdorff spaces are regular. And if  $(X, \tau)$  is second countable, then it is locally  $\sigma$  finite. But then  $(X, \tau)$  is a regular Hausdorff space that is locally  $\sigma$  finite, and hence by the Nagata-Smirnov metrization theorem it is metrizable. But metrizable topological spaces are paracompact. Hence,  $(X, \tau)$  is paracompact.  $\square$

All of the properties were needed here, even though the bulk of the work is done by the metrization theorem. If we drop the Hausdorff property, we may consider the bug-eyed line, which is not paracompact (indeed, it's not even regular). Dropping the second countability requirement allows us to consider the open long line, which is also not paracompact. It would be slightly incorrect to think that paracompactness can replace one of these properties. If any it seems like paracompactness should be able to do away with the second countable property of manifolds, but this is not so. The disjoint union of paracompact spaces is again paracompact, and so one could consider an uncountable disjoint union of spheres. This will be paracompact, locally Euclidean, and Hausdorff, but not second countable since one would need at least

one open set per sphere, and hence any basis is uncountable. However, this is not connected. Perhaps throwing in the connectedness requirement will help, and indeed it does. First, we prove a strengthening of the claim that topological manifolds are paracompact.

**Theorem 37.5.18.** *If  $(X, \tau)$  is a topological manifold, if  $\mathcal{B}$  is a basis for  $X$ , and if  $\mathcal{O}$  is an open cover of  $X$ , then there is a locally finite refine  $\Delta$  of  $\mathcal{O}$  such that  $\Delta \subseteq \mathcal{B}$ .*

*Proof.* Since topological manifolds are locally compact, Hausdorff, and second countable, there exists a compact exhaustion  $K : \mathbb{N} \rightarrow \mathcal{P}(X)$  of  $X$ . Define  $V : \mathbb{N} \rightarrow \mathcal{P}(X)$  by:

$$V_j = K_{j+1} \setminus \text{Int}_\tau(K_j) \quad (37.5.22)$$

Then  $V_j$  is a closed subset of a compact set, and is hence compact. Define  $W : \mathbb{N} \rightarrow \mathcal{P}(X)$  by:

$$W_j = \begin{cases} K_2, & k = 0 \\ \text{Int}_\tau(K_{j+2}) \setminus K_{j-1}, & k > 0 \end{cases} \quad (37.5.23)$$

Then for all  $j \in \mathbb{N}$ ,  $W_j$  is open and  $V_j \subseteq W_j$  since  $K$  is a compact exhaustion. But  $\mathcal{O}$  is an open cover of  $X$ , and hence for all  $j \in \mathbb{N}$  and for all  $x \in W_j$ , there is an open set  $\mathcal{U}_x \in \mathcal{O}$  such that  $x \in \mathcal{U}_x$ . But  $\mathcal{B}$  is a basis, and thus there is an open set  $B_x \in \mathcal{B}$  such that  $x \in B_x$  and  $B_x \subseteq W_j \cap \mathcal{U}_x$ . The set of all such  $B_x$  for  $x \in W_j$  is thus an open cover of  $W_j$ , and hence an open cover of  $V_j$  since  $V_j \subseteq W_j$ . But  $V_j$  is compact and therefore there is a finite subcover. The union of all such finite subcovers for each  $j \in \mathbb{N}$  is thus a countable refinement of  $\mathcal{O}$  that covers  $X$  since the  $K$  is a compact exhaustion, and thus  $\bigcup K_n = X$ . Moreover, it is locally finite since  $V_j \cap V_k = \emptyset$  for all  $k > j + 2$ , and thus there are only finitely many open sets in the refinement that contain  $W_j$ .  $\square$

**Theorem 37.5.19.** *If  $(X, \tau)$  is a  $\sigma$  compact topological space, then it is a Lindelöf topological space.*

*Proof.* For if  $(X, \tau)$  is  $\sigma$  compact, then there exists a sequence  $K : \mathbb{N} \rightarrow \mathcal{P}(X)$  of compact subsets such that  $X = \bigcup K_n$ . But if  $\mathcal{O}$  is an open cover of  $X$ , then for all  $n \in \mathbb{N}$  it is true that  $\mathcal{O}$  is an open cover of  $K_n$ . But  $K_n$  is compact and hence there is a finite subcover  $\Delta_n$ . Let  $\Delta = \bigcup \Delta_n$ . Then  $\Delta$  is the countable union of finite sets, and hence hence. Moreover,  $\Delta$  is a cover of  $K_n$  for all  $n \in \mathbb{N}$ . But  $X = \bigcup K_n$ , and hence  $\Delta$  is a cover for  $X$ . Thus,  $\Delta$  is a countable subcover of  $\mathcal{O}$ .  $\square$

**Theorem 37.5.20.** *If  $(X, \tau)$  is a locally Euclidean Hausdorff topological space, then it is a topological manifold if and only if it is  $\sigma$  compact.*

*Proof.* For if  $(X, \tau)$  is a topological manifold, then it is locally compact, Hausdorff, and second countable and thus there exists a compact exhaustion. But spaces with compact exhaustions are  $\sigma$  compact, and thus  $(X, \tau)$  is  $\sigma$  compact. In the other direction, if  $(X, \tau)$  is  $\sigma$  compact, then it is Lindelöf. But if  $X$  is locally Euclidean and Hausdorff, then there is a basis  $\mathcal{B}$  of coordinate balls. But then  $\mathcal{B}$  is a cover for  $X$ , and since  $X$  is Lindelöf there exists a countable subcover  $\Delta$ . But coordinate balls are second countable, and thus  $X$  is covered by countably many open second countable subspaces and is therefore second countable. But then  $X$  is a locally Euclidean Hausdorff space that is second countable, and is therefore a topological manifold.  $\square$

**Theorem 37.5.21.** *If  $(X, \tau)$  is a connected, paracompact, locally Euclidean, Hausdorff topological space, then it is a topological manifold.*

*Proof.* The only property missing from the hypothesis is the second countability. But if  $(X, \tau)$  is locally Euclidean, then there is a basis of precompact coordinate balls  $\mathcal{B}$ . But then  $\mathcal{B}$  is an open cover, and since  $X$  is paracompact there is a locally finite refinement  $\Delta$  of  $\mathcal{B}$ . Let  $\mathcal{U}_0 \in \Delta$ . Since  $\Delta$  is a locally finite of  $\mathcal{B}$ , there is a  $\mathcal{V} \in \mathcal{B}$  such that  $\mathcal{U}_0 \subseteq \mathcal{V}$ . But then  $\text{Cl}_\tau(\mathcal{U}_0) \subseteq \text{Cl}_\tau(\mathcal{V})$ , and  $\text{Cl}_\tau(\mathcal{V})$  is compact, and therefore  $\text{Cl}_\tau(\mathcal{U}_0)$  is compact. Define  $\mathcal{U} : \mathbb{N} \rightarrow \mathcal{P}(X)$  by:

$$\mathcal{U}_n = \left\{ x \in X \mid \exists_{A: \mathbb{Z}_n \rightarrow \Delta} (\mathcal{U}_0 = A_0 \text{ and } x \in A_n \text{ and } A_i \cap A_{i+1} \neq \emptyset) \right\} \quad (37.5.24)$$

That is,  $\mathcal{U}_n$  is the set of all points  $x \in X$  that can be connected to  $\mathcal{U}_0$  be at most  $n$  sets in  $\Delta$  such that consecutive elements have non-empty overlap. Then for all  $n \in \mathbb{N}$ ,  $\text{Cl}_\tau(\mathcal{U}_n)$  is compact. For by induction, the base case is simply  $\mathcal{U}_0$ . Suppose it is true for  $n \in \mathbb{N}$ . But if  $x \in \mathcal{U}_n$ , then there is a  $\mathcal{U} \in \Delta$  such that  $\mathcal{V} \cap \text{Cl}_\tau(\mathcal{U}_{n-1}) \neq \emptyset$ . But only finitely many such sets intersect  $\text{Cl}_\tau(\mathcal{U}_{n-1})$ . For if not, then there is a sequence of points  $B : \mathbb{N} \rightarrow \Delta$  such that  $B_k \cap \text{Cl}_\tau(\mathcal{U}_{n-1}) \neq \emptyset$  and all of the  $B_k$  are distinct. But  $\text{Cl}_\tau(\mathcal{U}_{n-1})$  is compact, and  $X$  is a locally metrizable paracompact Hausdorff space, and thus by the Smirnov theorem it is metrizable. But if  $\text{Cl}_\tau(\mathcal{U}_{n-1})$  is compact and  $X$  is metrizable, then  $\text{Cl}_\tau(\mathcal{U}_{n-1})$  is sequentially compact. But then there is a convergent subsequence of  $B$ , let  $a$  be the limit. But then  $a$  is contained in infinitely many elements of  $\Delta$ , a contradiction as  $\Delta$  is a locally finite cover. Hence, only finitely many elements of  $\Delta$  intersect  $\text{Cl}_\tau(\mathcal{U}_{n-1})$ . But then  $\mathcal{U}_n$  is the union finitely many precompact sets, and is therefore precompact. But then  $\text{Cl}_\tau(\mathcal{U}_n)$  is compact. Lastly,  $\bigcup \text{Cl}_\tau(\mathcal{U}_n) = X$ . For suppose not, and let  $x \in X$  be such that  $x \notin \bigcup \text{Cl}_\tau(\mathcal{U}_n)$ . But  $X$  is locally connected and connected, and is therefore path connected. Thus, given a point  $y \in \mathcal{U}_0$ , there is a path  $\gamma : [0, 1] \rightarrow X$  such that  $\gamma(0) = y$  and  $\gamma(1) = x$ . But  $[0, 1] \subseteq \mathbb{R}$  is compact, and thus the image of  $\gamma$  is compact in  $X$ . But then there are only finitely many elements of  $\Delta$  that intersect  $\gamma$ . But then  $y$  is finitely many elements of  $\Delta$

away from  $\mathcal{U}_0$ , and hence will be contained in some  $\mathcal{U}_n$ , a contradiction. Thus,  $X$  is the union of the  $\text{Cl}_\tau(\mathcal{U}_n)$ . But then  $X$  is  $\sigma$  compact. But if  $X$  is locally Euclidean, Hausdorff, and  $\sigma$  compact, then it is second countable. Thus,  $(X, \tau)$  is a topological manifold.  $\square$

This theorem is rather subtle, as many of the counterexamples below will show.

**Example 37.5.11** Locally Euclidean, Hausdorff, and paracompactness imply that such a space is metrizable. However, connected metric spaces need not be second countable. We cannot look to the long line for a counterexample, since it is not paracompact and hence not metrizable. Such a claim can be realized by considering the Paris metric on  $\mathbb{R}^2$ . It is a path connected, but not second countable since any basis must contain an interval lying in every line through the origin. Since there are uncountably many such lines, any basis must be uncountable. This space is locally Euclidean everywhere except for the origin. If we delete the origin we get a metrizable locally Euclidean space, but it is no longer connected. Hence, the previous theorem avoids this space and shows that all of the conditions were needed.

**Example 37.5.12** More than lacking the locally Euclidean property at the origin, the Paris metric is not locally compact at the origin either. If we look at locally compact Hausdorff spaces that are connected and paracompact, then the compactification of the long line serves as an example of a space that is not second countable.

### Theorem 37.5.22: Fundamental Group of Manifolds are Countable

If  $n \in \mathbb{N}$ , if  $(X, \tau)$  is a topological manifold of dimension  $n$ , if  $x \in X$ , and if  $\pi_1(X, x)$  is the fundamental group of  $X$  at the point  $x$ , then  $\pi_1(X, x)$  is countable.

*Proof.* For if  $(X, \tau)$  is a topological manifold, then there is a basis  $\mathcal{B}$  of  $\tau$  such that for all  $\mathcal{U} \in \mathcal{B}$  it is true that  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^n$ . Moreover, since  $X$  is second countable, every subspace is second countable. But  $X$  is locally path connected, and hence its connected components are open. Thus, for any  $B, B' \in \mathcal{B}$ ,  $B \cap B'$  also has countably many connected components, each of which is therefore path connected. Come back later.  $\square$

## 37.5.2 Differentiable Manifolds

Topological manifolds are insufficient for calculus. Differentiable functions are not invariant under homeomorphism, the mapping  $x \mapsto x^{1/3}$  is such an example. The function  $f(x) = x$  is differentiable, but  $f(x^{1/3}) = x^{1/3}$  is not

differentiable at the origin. To do calculus, one needs more structure than that given by topological manifolds. Def transition maps. Draw picture. Smoothly compatible charts. There exist topological manifolds with no smooth structure (10 dimensional compact example by Kervaire, 1960).

## 37.6 Smooth Manifolds

Smooth manifolds are topological manifolds with a smooth structure on them so that one can do calculus. Such objects first require a notion of smooth functions, and we define this in terms of the familiar notions of smoothness of functions between Euclidean spaces  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ .

### Definition 37.6.1: Smooth Real-Valued Functions On $\mathbb{R}$

A smooth real-valued function on an open subset  $\mathcal{U} \subseteq \mathbb{R}^n$  is a function  $f : \mathcal{U} \rightarrow \mathbb{R}$  such that all mixed partial derivatives of all orders exist and are continuous for all  $\mathbf{x} \in \mathcal{U}$ .

$\mathbb{R}^n$  can be defined as the set of all functions  $\mathbf{x} : \mathbb{Z}_n \rightarrow \mathbb{R}$ . Given an element  $\mathbf{x} \in \mathbb{R}^n$  and  $k \in \mathbb{Z}_n$  we denote image of  $k$  as  $x_k = \mathbf{x}(k)$ . This is called the  $k^{th}$  coordinate of  $\mathbf{x}$ . The projection mapping  $\pi_k : \mathbb{R}^n \rightarrow \mathbb{R}$  for  $k \in \mathbb{Z}_n$  is the function defined by  $\pi_k(\mathbf{x}) = x_k$ . We can use this notion to define smooth functions between arbitrary Euclidean spaces by requiring the composition of  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$  with  $\pi_k$  to be smooth for all  $k$ .

### Definition 37.6.2: Smooth Euclidean Functions

A smooth function on a subset  $\mathcal{U} \subseteq \mathbb{R}^m$  to  $\mathbb{R}^n$  is a function  $f : \mathcal{U} \rightarrow \mathbb{R}^m$  such that, for all  $k \in \mathbb{Z}_n$ , the function  $\pi_k \circ f$  is a smooth real-valued function.

We can use this definition to create a notion of smoothness on a manifold by considering charts that overlap *smoothly*.

**Definition 37.6.3: Smoothly Overlapping Charts**

Smoothly overlapping charts of dimension  $n \in \mathbb{N}$  are charts  $(\mathcal{U}_1, \varphi_1)$  and  $(\mathcal{U}_2, \varphi_2)$  of dimension  $n$  on a topological space  $(X, \tau)$  such that:

$$\varphi_1 \circ \phi_2^{-1} : \varphi_2(\mathcal{U}_1 \cap \mathcal{U}_2) \rightarrow \mathbb{R}^n \quad \varphi_2 \circ \phi_1^{-1} : \varphi_1(\mathcal{U}_1 \cap \mathcal{U}_2) \rightarrow \mathbb{R}^n$$

are smooth functions, or such that  $\mathcal{U}_1 \cap \mathcal{U}_2 = \emptyset$ .

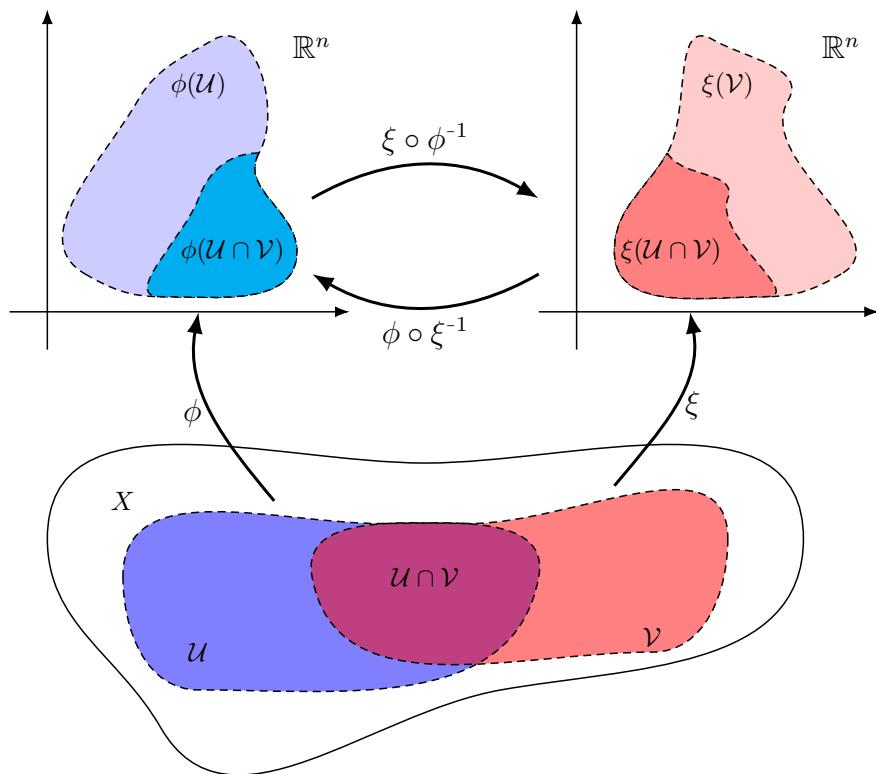


Fig. 37.15: Smoothly Overlapping Charts

**Definition 37.6.4: Transition Function**

The transition function of a chart  $(\mathcal{U}, \varphi)$  with respect a chart  $(\mathcal{V}, \psi)$  on a topological space  $(X, \tau)$  is the function  $f : \varphi(\mathcal{U} \cap \mathcal{V}) \rightarrow \psi(\mathcal{V} \cap \mathcal{V})$ :

$$f(x) = (\psi \circ \varphi^{-1})(x)$$

**Definition 37.6.5: Smooth Atlas**

A smooth atlas on a topological space  $(X, \tau)$  is an atlas  $\mathcal{A}$  such that for all charts  $(\mathcal{U}, \phi), (\mathcal{V}, \xi) \in \mathcal{A}$ , the transition function of  $(\mathcal{U}, \phi)$  with respect to  $(\mathcal{V}, \xi)$  is smooth.

**Definition 37.6.6: Maximal Smooth Atlas**

A maximal smooth atlas on a topological space  $(X, \tau)$  is a smooth atlas  $\mathcal{A}$  on  $(X, \tau)$  such that for all charts  $(\mathcal{U}, \varphi)$  of  $(X, \tau)$  such that  $(\mathcal{U}, \varphi)$  overlaps smoothly with all  $(\mathcal{V}, \psi) \in \mathcal{A}$ , it is true that  $(\mathcal{U}, \varphi) \in \mathcal{A}$ .

**Theorem 37.6.1.** *If  $(X, \tau)$  is a topological space and if  $\mathcal{A}$  is a smooth atlas of dimension  $n \in \mathbb{N}$  on  $(X, \tau)$ , then there is a unique maximal smooth atlas  $\mathcal{C}$  on  $(X, \tau)$  such that  $\mathcal{A} \subseteq \mathcal{C}$ .*

*Proof.* For let  $\mathcal{C}$  be the set of all charts on  $(X, \tau)$  that overlap smoothly with the charts in  $\mathcal{A}$ . Then since  $\mathcal{A}$  is an atlas, for all  $(\mathcal{U}, \phi) \in \mathcal{A}$  and for all  $(\mathcal{V}, \xi) \in \mathcal{A}$ , we have that  $(\mathcal{U}, \phi)$  and  $(\mathcal{V}, \xi)$  overlap smoothly, and thus  $(\mathcal{U}, \phi) \in \mathcal{C}$ . Therefore  $\mathcal{A} \subseteq \mathcal{C}$ . But  $\mathcal{A}$  is an atlas and thus for all  $x \in X$  there is a chart  $(\mathcal{U}, \phi) \in \mathcal{A}$  such that  $x \in \mathcal{U}$ . But  $\mathcal{A} \subseteq \mathcal{C}$  and thus  $(\mathcal{U}, \phi) \in \mathcal{C}$ . Thus, for all  $x \in X$  there is a chart  $(\mathcal{U}, \phi) \in \mathcal{C}$  such that  $x \in \mathcal{U}$ . Suppose  $(\mathcal{U}_1, \phi_1), (\mathcal{U}_2, \phi_2) \in \mathcal{C}$ . If  $\mathcal{U}_1$  and  $\mathcal{U}_2$  are disjoint, then these two charts overlap smoothly. Suppose it is non-empty and let  $f$  be the transition function of  $(\mathcal{U}_1, \phi_1)$  with respect to  $(\mathcal{U}_2, \phi_2)$ . Let  $p \in \phi_1(\mathcal{U}_1 \cap \mathcal{U}_2)$ . But then there is a chart  $\xi \in \mathcal{A}$  such that  $\phi_2^{-1}(p)$  is contained in the domain of  $\xi$ . From the associativity of composition, we have:

$$\phi_1 \circ \phi_2^{-1} = (\phi_1 \circ \xi^{-1}) \circ (\xi \circ \phi_2^{-1}) \quad (37.6.1)$$

But by the definition of  $\mathcal{C}$ ,  $\phi_1$  and  $\phi_2$  overlap smoothly with  $\xi$ , and thus this is the composition of smooth functions, and is therefore smooth. Therefore  $\phi_1 \circ \phi_2^{-1}$  is smooth and thus  $(\mathcal{U}_1, \phi_1)$  and  $(\mathcal{U}_2, \phi_2)$  overlap smoothly. Thus,  $\mathcal{C}$  is a smooth atlas. Moreover, it is complete from the construction. Given any

other complete atlas  $\mathcal{C}'$  that contains  $\mathcal{A}$  we would have  $\mathcal{C} \subseteq \mathcal{C}'$  and  $\mathcal{C}' \subseteq \mathcal{C}$ , and therefore  $\mathcal{C} = \mathcal{C}'$ . Thus, this completion is unique.  $\square$

### Definition 37.6.7: Smooth Manifold

A smooth manifold of dimension  $n \in \mathbb{N}$ , denoted  $(X, \tau, \mathcal{A})$  is topological manifold  $(X, \tau)$  with a maximal smooth atlas  $\mathcal{A}$  of dimension  $n$  on  $(X, \tau)$ .

Any smooth atlas  $\mathcal{A}$  on a topological space  $(X, \tau)$  defines a smooth manifold if we let  $\mathcal{C}$  be the maximal smooth atlas generated by  $\mathcal{A}$ .

**Example 37.6.1** Let  $(\mathbb{R}^n, \tau_{\mathbb{R}^n})$  be the standard  $n$  dimensional Euclidean space. We can define a trivial smooth atlas on this space by let  $\mathcal{A} = \{(\mathbb{R}^n, \text{id})\}$ , where  $\text{id}$  is the identity function. This defines a smooth atlas. By considering the unique maximal smooth atlas generated by this we obtain the standard smooth structure on  $\mathbb{R}^n$ .

**Theorem 37.6.2.** *If  $(X, \tau, \mathcal{A})$  is a smooth manifold of dimension  $n \in \mathbb{N}$ , if  $(\mathcal{U}, \varphi)$  is a chart in  $\mathcal{A}$ , if  $\mathcal{V} \in \tau$ , and if  $\varphi_{\mathcal{V}}$  denotes the restriction mapping:  $\varphi_{\mathcal{V}} : \mathcal{V} \rightarrow \mathbb{R}^n$ , then  $(\mathcal{V}, \varphi_{\mathcal{V}}) \in \mathcal{A}$ .*

*Proof.* For since  $\phi$  is a homeomorphism from  $\mathcal{U}$  to  $\phi(\mathcal{U})$ , and since  $\mathcal{V} \in \tau$ , we have that  $\phi_{\mathcal{V}}$  is a homeomorphism between  $\mathcal{V}$  and  $\phi_{\mathcal{V}}(\mathcal{V})$ , and therefore  $(\mathcal{V}, \phi_{\mathcal{V}})$  is a chart. But this chart meets  $(\mathcal{U}, \phi)$  smoothly, and  $\mathcal{A}$  is complete. Thus,  $(\mathcal{V}, \phi_{\mathcal{A}}) \in \mathcal{A}$ .  $\square$

**Theorem 37.6.3.** *If  $n \in \mathbb{N}$ , then there is a complete atlas  $\mathcal{A}$  on  $(S^n, \tau)$ , where  $\tau$  is the inherited topology from  $\mathbb{R}^{n+1}$ .*

*Proof.* For all  $k \in \mathbb{Z}_{n+1}$ , let  $\mathcal{U}_k^+$  and  $\mathcal{U}_k^-$  be defined as:

$$\mathcal{U}_k^+ = \{ \mathbf{x} \in S^n : x_k > 0 \} \quad (37.6.2) \quad \mathcal{U}_k^- = \{ \mathbf{x} \in S^n : x_k < 0 \} \quad (37.6.3)$$

Define  $\phi_{\mathcal{U}_k^+} : \mathcal{U}_k^+ \rightarrow \mathbb{R}^n$  by:

$$\phi_{\mathcal{U}_k^+}(\mathbf{x}) = (x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_{n+1}) \quad (37.6.4)$$

That is, the mapping that projects the point onto the plane defined by  $x_k = 0$ . Define  $\phi_{\mathcal{U}_k^-}$  similarly. Then all such  $\phi$  are homeomorphisms from their domain to their image. Let  $\mathcal{A}$  be defined as follows:

$$\mathcal{A} = \{ (\mathcal{U}_k^+, \phi_{\mathcal{U}_k^+}) : k \in \mathbb{Z}_n \} \bigcup \{ (\mathcal{U}_k^-, \phi_{\mathcal{U}_k^-}) : k \in \mathbb{Z}_n \} \quad (37.6.5)$$

Then  $\mathcal{A}$  is an atlas of  $(S^n, \tau)$ . For if  $\mathbf{x} \in S^n$ , then  $\|\mathbf{x}\|_2 = 1$ . But then there is a coordinate  $x_k$  of  $\mathbf{x}$  such that  $x_k \neq 0$ . But then either  $x_k > 0$  or  $x_k < 0$ , and thus either  $\mathbf{x} \in \mathcal{U}_k^+$  or  $\mathbf{x} \in \mathcal{U}_k^-$ . If  $(\mathcal{V}_2, \phi_2)$  and  $(\mathcal{V}_2, \phi_2)$  are charts, then either  $\phi_1(\mathcal{V}_1)$  and  $\phi_2(\mathcal{V}_2)$  are disjoint or they are not. If they are disjoint, then  $\phi_1$  and  $\phi_2$  overlap smoothly. If they are not disjoint, let  $\mathbf{x}$  be contained in the intersection. But then, for all  $k \in \mathbb{Z}_n$ ,  $\pi_k \circ (\phi_1 \circ \phi_2^{-1})$  is smooth, and thus  $\phi_1$  and  $\phi_2$  overlap smoothly.  $\square$

### Definition 37.6.8: Open Submanifold

An open submanifold on a manifold  $(X, \tau, \mathcal{A})$  is a an open subset  $\mathcal{U} \subseteq X$  and the collection  $\mathcal{A}_{\mathcal{U}}$  defined by:

$$\mathcal{A}_{\mathcal{U}} = \{(\mathcal{V}, \phi) \in \mathcal{A} : \mathcal{V} \subseteq \mathcal{U}\} \quad (37.6.6)$$

Together with the inherited topology  $\tau_{\mathcal{U}}$ .

**Theorem 37.6.4.** *If  $(X, \tau, \mathcal{A})$  is a smooth manifold and if  $(\mathcal{U}, \tau_{\mathcal{U}}, \mathcal{A}_{\mathcal{U}})$  is an open submanifold, then it is a smooth manifold.*

*Proof.* For by the previous theorem,  $\mathcal{A}_{\mathcal{U}}$  is a complete atlas. Moreover, a subspace of a Hausdorff topological space is also a Hausdorff topological space, and hence  $(\mathcal{U}, \tau_{\mathcal{U}})$  is a Hausdorff space. Thus,  $(\mathcal{U}, \tau_{\mathcal{U}}, \mathcal{A}_{\mathcal{U}})$  is a smooth manifold.  $\square$

### Definition 37.6.9: Product Chart

The product chart of an  $n$  dimensional chart  $(\mathcal{U}, \phi)$  on a topological space  $(X, \tau_X)$  with an  $m$  dimensional chart  $(\mathcal{V}, \xi)$  on a topological space  $(Y, \tau_Y)$  is the ordered pair  $(\mathcal{U} \times \mathcal{V}, f)$  where  $f : \mathcal{U} \times \mathcal{V} \rightarrow \mathbb{R}^{n+m}$  defined by:

$$f(p, q)_k = \begin{cases} \phi(p)_k, & k < n \\ \xi(q)_k, & n \leq k < n + m \end{cases} \quad (37.6.7)$$

Where  $\phi(p)_k$  is the  $k^{th}$  coordinate of  $\phi(p) \in \mathbb{R}^n$  and  $\xi(q)_k$  is the  $k^{th}$  coordinate of  $\xi(q) \in \mathbb{R}^m$ . We denote this by  $(\mathcal{U}, \phi) \times (\mathcal{V}, \xi)$ .

Thinking of the elements of  $\mathbb{R}^{n+m}$  as tuples of length  $n + m$ , we can write:

$$f(p, q) = (x_1(p), \dots, x_n(p), y_1(q), \dots, y_m(q)) \quad (37.6.8)$$

**Theorem 37.6.5.** *If  $(X, \tau_X, \mathcal{A}_X)$  and  $(Y, \tau_Y, \mathcal{A}_Y)$  are smooth manifolds, and if  $\mathcal{A}$  is the set of all product charts on  $X \times Y$ , then  $\mathcal{A}$  is a smooth atlas on  $(X \times Y, \tau_{X \times Y})$ , where  $\tau_{X \times Y}$  is the product topology.*

*Proof.* For if  $p \in X \times Y$  then there is an  $x \in X$  and a  $y \in Y$  such that  $p = (x, y)$ . But  $\mathcal{A}_X$  is a smooth atlas on  $(X, \tau_X)$ , and thus if  $x \in X$  then there is a  $(\mathcal{U}, \phi) \in \mathcal{A}_X$  such that  $x \in \mathcal{U}$ . Similarly, there is a  $(\mathcal{V}, \xi) \in \mathcal{A}_Y$  such that  $y \in \mathcal{V}$ . But then  $p \in \mathcal{U} \times \mathcal{V}$ , and  $\mathcal{U} \times \mathcal{V} \in \tau_{X \times Y}$ . But if  $\phi : \mathcal{U} \rightarrow \mathbb{R}^n$  is a homeomorphism between  $\mathcal{U}$  and  $\phi(\mathcal{U})$  and  $\xi : \mathcal{V} \rightarrow \mathbb{R}^m$  is a homemorphism between  $\mathcal{V}$  and  $\xi(\mathcal{V})$ , then  $f : \mathcal{U} \times \mathcal{V} \rightarrow \mathbb{R}^{n+m}$  is a homeomorphism between  $\mathcal{U} \times \mathcal{V}$  and  $f(\mathcal{U} \times \mathcal{V})$ , and thus the product chart is a chart in  $(X \times Y, \tau_{X \times Y})$ . Moreover, all of the elements of  $\mathcal{A}$  are smoothly overlapping. Thus,  $\mathcal{A}$  is an atlas on  $(X \times Y, \tau_{X \times Y})$ .  $\square$

Using the maximal smooth atlas generated by the product atlas  $\mathcal{A}$  creates the product manifold.

### 37.6.1 Smooth Mappings

#### Definition 37.6.10: Smooth Real-Valued Functions

A smooth real-valued function on a smooth manifold  $(X, \tau, \mathcal{A})$  of dimension  $n \in \mathbb{N}$  is a function  $\phi : X \rightarrow \mathbb{R}$  such that for every chart  $(\mathcal{U}, \varphi) \in \mathcal{A}$ , the function  $\phi \circ \varphi^{-1} : \phi(\mathcal{U}) \rightarrow \mathbb{R}$  is a smooth Euclidean function.

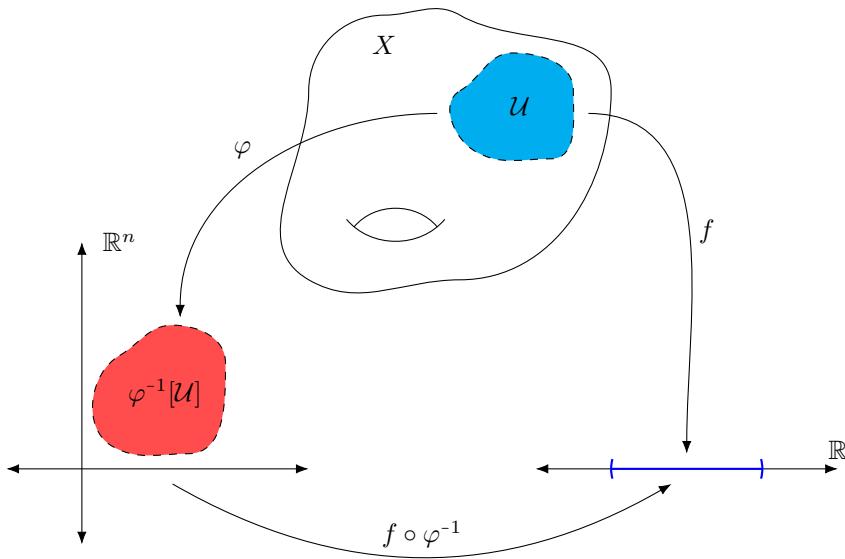


Fig. 37.16: Smooth Real-Valued Function on a Manifold

**Theorem 37.6.6.** If  $(X, \tau, \mathcal{A})$  is a manifold and if  $f, g : X \rightarrow \mathbb{R}$  are smooth real-valued functions, then  $(f + g) : X \rightarrow \mathbb{R}$  defined by:

$$(f + g)(x) = f(x) + g(x) \quad x \in X \quad (37.6.9)$$

Is a smooth real-valued function.

**Theorem 37.6.7.** If  $(X, \tau, \mathcal{A})$  is a manifold and if  $f, g : X \rightarrow \mathbb{R}$  are smooth real-valued functions, then  $(f \cdot g) : X \rightarrow \mathbb{R}$  defined by:

$$(f \cdot g)(x) = f(x) \cdot g(x) \quad x \in X \quad (37.6.10)$$

Is a smooth real-valued function.

### Definition 37.6.11: Smooth Functions Between Manifolds

A smooth function from a smooth manifold  $(X, \tau_X, \mathcal{A}_X)$  of dimension  $m \in \mathbb{N}$  to a smooth manifold  $(Y, \tau_Y, \mathcal{A}_Y)$  of dimension  $n \in \mathbb{N}$  is a function  $\phi : X \rightarrow Y$  such that for every chart  $(\mathcal{U}, \varphi) \in \mathcal{A}_X$  and for every chart  $(\mathcal{V}, \psi) \in \mathcal{A}_Y$ , the function  $\psi \circ \phi \circ \varphi^{-1} : \varphi(\mathcal{U}) \rightarrow \mathbb{R}^n$  is a smooth function.

**Theorem 37.6.8.** If  $(X, \tau_X, \mathcal{A}_X)$  and  $(Y, \tau_Y, \mathcal{A}_Y)$  are manifolds, if  $A_X \subseteq \mathcal{A}_X$  is an atlas on  $(X, \tau_X)$ , if  $A_Y \subseteq \mathcal{A}_Y$  is an atlas on  $(Y, \tau_Y)$ , and if  $f : X \rightarrow Y$  is a function such that, for all  $(U, \phi) \in A_X$  and for all  $(V, \xi) \in A_Y$  it is true that  $\xi \circ f \circ \phi^{-1} : \xi(V) \rightarrow \mathbb{R}^m$  is a smooth Euclidean function, then  $f$  is smooth.

*Proof.* Since charts in  $\mathcal{A}_X$  and  $\mathcal{A}_Y$  overlap smoothly with charts in  $A_X$  and  $A_Y$ , and since the atlases  $A_X$  and  $A_Y$  cover  $X$  and  $Y$ , respectively, we are done.  $\square$

**Theorem 37.6.9.** If  $(X, \tau_X, \mathcal{A}_X)$  is a manifold, then  $\text{id} : X \rightarrow X$  is a smooth function.

**Theorem 37.6.10.** If  $(X, \tau_X, \mathcal{A}_X)$ ,  $(Y, \tau_Y, \mathcal{A}_Y)$ , and  $(Z, \tau_Z, \mathcal{A}_Z)$  are manifolds, if  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  are smooth, then  $g \circ f : X \rightarrow Z$  is smooth.

Smoothness is a local property. A function  $\phi : M \rightarrow N$  is smooth at  $p \in M$  if there is a neighborhood  $\mathcal{U}$  of  $p$  such that the restriction of  $\phi$  to  $\mathcal{U}$  is smooth. A smooth function is thus a function that is smooth at every point.

**Theorem 37.6.11.** If  $(X, \mathcal{A}_X, \tau_X)$  and  $(Y, \mathcal{A}_Y, \tau_Y)$  are manifolds and if  $f : X \rightarrow Y$  is smooth, then  $f$  is continuous.

### Definition 37.6.12: Diffeomorphism

A diffeomorphism from a manifold  $(X, \tau_X, \mathcal{A}_X)$  to a manifold  $(Y, \tau_Y, \mathcal{A}_Y)$  is a bijective function  $f : X \rightarrow Y$  such that  $f$  and  $f^{-1}$  are smooth.

**Example 37.6.2** For any  $a, b \in \mathbb{R}$  with  $a < b$ , the interval  $(a, b)$  is diffeomorphic to the unit interval  $(0, 1)$ . For let  $\phi : (0, 1) \rightarrow (a, b)$  be defined by:

$$\phi(t) = (a - b)t + b \quad (37.6.11)$$

Then  $\phi$  is a smooth bijection and its inverse is smooth. Moreover, the unit interval is diffeomorphic to  $\mathbb{R}$ . For let  $\xi : (0, 1) \rightarrow \mathbb{R}$  be defined by:

$$\xi(t) = \frac{2t}{t(1-t)} \quad (37.6.12)$$

**Theorem 37.6.12.** If  $(X, \tau_X, \mathcal{A}_X)$  and  $(Y, \tau_Y, \mathcal{A}_Y)$  are manifolds, and if  $f : X \rightarrow Y$  is a diffeomorphism, then  $f$  is a homeomorphism from  $(X, \tau_X)$  to  $(Y, \tau_Y)$ .

*Proof.* For if  $f$  is a diffeomorphism, then it is a smooth bijection such that its inverse is smooth. But if  $f$  is smooth, then it is continuous and therefore it is a continuous bijection. But if  $f^{-1}$  is smooth, then it is continuous, and thus  $f$  is a bicontinuous bijective function, and is therefore a homeomorphism.  $\square$

A smooth homeomorphism need not be a diffeomorphism. The inverse function may not be smooth. For let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $f(x) = x^3$ . Then  $f$  is a homeomorphism and its forward direction is smooth, but  $f^{-1}$  is not smooth at the origin.

**Theorem 37.6.13.** *If  $A$  is a set, if  $(X, \tau, \mathcal{A})$  is a manifold, and if  $f : A \rightarrow X$  is an bijective function, then there exists a topology  $\tau_A$  and an atlas  $\mathcal{A}_A$  on  $X$  such that  $f$  is a diffeomorphism.*

**Definition 37.6.13** An  $n$  dimensional smooth manifold is a topological space  $(M, \tau)$  equipped with a collection  $\mathcal{A}$  or ordered pairs  $\{(\mathcal{U}_\alpha, \varphi_\alpha)\}$  such that  $\mathcal{U}_\alpha$  are open and cover the space, and such that  $\varphi_\alpha : \mathcal{U}_\alpha \rightarrow \mathcal{V}_\alpha$  is a homeomorphism to an open subset  $\mathcal{V}_\alpha$  of  $\mathbb{R}^n$  and such that for when  $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$ , the functions  $\varphi_\beta \circ \varphi_\alpha^{-1}$  are smooth.

The notion of smoothness here is simply the smoothness of functions from  $\mathbb{R}^n$  to itself. The collection  $\mathcal{A}$  is called an atlas and  $(\mathcal{U}, \varphi)$  are called charts.

**Example 37.6.3** If  $\mathcal{U}$  is an open subset of  $\mathbb{R}^n$  and  $f : \mathcal{U} \rightarrow \mathbb{R}$  is smooth, then the graph of  $f$  is a smooth manifold.

**Example 37.6.4** If  $\mathcal{U} \subseteq \mathbb{R}^n$  is open, if  $f : \mathcal{U} \rightarrow \mathbb{R}$  is smooth, if  $c$  is in the range of  $f$ , and if  $\text{grad}(f)$  is non-zero on all of  $f^{-1}(c)$ , then  $f^{-1}(c)$  is a smooth manifold.

**Example 37.6.5** Define  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  by  $f(\mathbf{x}) = \|\mathbf{x}\|_2^2$ . For all  $\mathbf{x} \in \mathbb{R}^3$  such that  $\|\mathbf{x}\| = 1$  we have  $\text{grad}(f) \neq 0$ , and thus  $f^{-1}(\{1\})$  is a smooth manifold. This is the unit sphere.

**Example 37.6.6** The orthographic projection of the sphere maps:

$$(x, y) \mapsto (x, y, \sqrt{1 - x^2 - y^2}) \quad (37.6.13)$$

The stereographic projection also exists.

**Definition 37.6.14** The product manifold of manifolds  $M_1, \dots, M_k$ , where  $M_j$  is an  $n_j$  smooth manifold then  $M = \prod_{j=1}^k M_j$  has a natural structure of an  $n = n_1 + \dots + n_k$  dimensional manifold.

**Definition 37.6.15** The quotient manifold of a smooth manifold  $M$  with an equivalence relation  $\sim$  is a manifold structure on  $M/\sim$  with the quotient topology. In particular when we have group actions.

**Example 37.6.7** Consider the sphere  $S^2$  with the equivalence relation  $R$  defined by  $pRq$  if and only if  $p = \pm q$ . The quotient space  $S^2/R$  is the real projective plane  $\mathbb{RP}^2$ . We can consider this by means of a group action on  $\mathbb{Z}_2$  on  $S^2$ . Define  $g \cdot p$  by  $1 \cdot p = p$  and  $-1 \cdot p = -p$ . This is a group action of  $\mathbb{Z}_2$  on  $S^2$ .

All of the examples thus far have been subsets of some Euclidean space  $\mathbb{R}^n$ . One natural question is whether or not there are smooth manifolds that do not live in some higher dimensional Euclidean space. That is, are there manifolds  $M$  such that  $M$  can not be embedded into some  $\mathbb{R}^m$ ?

**Definition 37.6.16** A function  $\phi : M \rightarrow N$  is smooth if for all  $p \in M$  there is a chart  $(\mathcal{U}, \varphi)$  containing  $p$  and a chart  $(\mathcal{V}, \psi)$  containing  $\phi(p)$  such that  $\psi \circ \phi \circ \varphi^{-1}$  is smooth.

Studying maps between manifolds gives us new manifolds to study. In particular, there are submersions and embeddings, and in particular embedded submanifolds.

Tangent spaces are another important topic in differential topology. The classic example is that of a sphere  $S^2$  in  $\mathbb{R}^3$ . We draw the tangent plane to a point on a sphere, which is the best linear approximation of the sphere at that point. This notion relies on an ambient space (that of  $\mathbb{R}^3$ ), but since we do not yet know if all manifolds can be embedded into such an ambient space, we need a new means of defining tangent spaces that agrees with our intuition. There are several ways of thinking of this:

- Tangent vectors are derivations on  $C^\infty(p)$ .
- Tangent vectors are equivalence classes of curves through  $p$ .

Given a function  $\phi : M \rightarrow N$  that is smooth between two manifolds, there is a function  $d\phi_o : T_p M \rightarrow T_{\phi(p)} N$  between these tangent spaces. We can further consider the collection of all tangent spaces at all points  $p$  of  $M$ , forming the tangent bundle of  $M$ , denoted  $TM$ . Set theoretically, this is the disjoint union  $TM = \coprod_p T_p M$ , but we do **not** give this the disjoint union topology. There is a natural topology that can be endowed on  $TM$ .

**Example 37.6.8** If we take  $S^1$ , the tangent bundle is  $TM = S^1 \times \mathbb{R}$ . This is an example of a trivial bundle. Most tangent bundles are not of this form. That is, for an  $m$  dimensional manifold  $M$ ,  $TM$  is usual **not** equal to  $M \times \mathbb{R}^m$ . It will always have dimension  $2m$ .

**Definition 37.6.17** A diffeomorphism from an open subset  $\mathcal{U} \subseteq \mathbb{R}^n$  to an open subset  $\mathcal{V} \subseteq \mathbb{R}^m$  is a homeomorphism  $f : \mathcal{U} \rightarrow \mathcal{V}$  such that  $f$  and  $f^{-1}$  are smooth functions.

**Definition 37.6.18** Let  $M$  an  $n$  dimensional topological manifold. Two coordinate charts  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  are said to be  $C^\infty$  compatible if  $\psi \circ \varphi^{-1}$ , defined as a function from  $\varphi(\mathcal{U} \cap \mathcal{V})$  to  $\psi(\mathcal{U} \cap \mathcal{V})$ , is a diffeomorphism.

**Definition 37.6.19** Let  $M$  be an  $n$  dimensional topological manifold. A smooth atlas on  $M$  is a collection  $\mathcal{A}$  of charts  $(\mathcal{U}, \varphi)$  that cover  $M$  and such that all charts are smoothly compatible. A smooth atlas  $\mathcal{A}$  on  $M$  is called maximal if it is not contained in any strictly larger smooth atlas. A smooth structure on  $M$  is a maximal smooth atlas.

**Definition 37.6.20** A smooth manifold, denoted  $(M, \tau, \mathcal{A})$ , is a topological manifold  $(M, \tau)$  with a smooth structure  $\mathcal{A}$ .

**Theorem 37.6.14.** *If  $M$  is an  $n$  dimensional topological manifold, and if  $\mathcal{A}$  is a smooth atlas on  $M$ , then there is a unique maximal atlas  $\mathcal{A}'$  such that  $\mathcal{A} \subseteq \mathcal{A}'$ .*

Given a smooth manifold  $(X, \tau, \mathcal{A})$ ,  $\mathcal{A}$  a maximal smooth atlas, a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}$  is called smooth,  $\mathcal{U}$  is called a coordinate domain, and  $\varphi$  is called a smooth coordinate map. This chart is called a coordinate ball if  $\varphi(\mathcal{U})$  is an open ball. Similarly, it's called a coordinate cube if  $\varphi(\mathcal{U})$  is an open cube. A chart is called regular if it is a smooth coordinate ball and there is a coordinate ball  $(\mathcal{V}, \psi)$  such that  $\varphi(\mathcal{U}) \subseteq \psi(\mathcal{V})$  with the same center and different radii.

**Example 37.6.9** A zero dimensional manifold is just a countable collection of isolated points.

**Example 37.6.10** For any  $n \in \mathbb{N}$ ,  $\mathbb{R}^n$  with the identity mapping  $\text{id}_{\mathbb{R}^n}$  forms an atlas, and thus lives inside a unique maximal smooth atlas.

**Example 37.6.11** We can put another non-compatible smooth structure on  $\mathbb{R}^n$ . Let  $\mathbb{R}$  be taken with the mapping  $f(x) = x^3$ . Then  $(\mathbb{R}, f)$  forms an atlas since  $f$  is a homomorphism. It overlaps smoothly with every other element since there's only element and  $f \circ f^{-1}$  is just the identity mapping, which is smooth and has smooth inverse. However, this atlas and the standard atlas aren't compatible since the composition is  $x^{1/3}$ , which is not differentiable at the origin. However, there is a diffeomorphism from the standard atlas to this one, and hence the two maximal atlases are essentially the same.

**Example 37.6.12** Consider the space of  $m \times n$  matrices over both  $\mathbb{R}$  and  $\mathbb{C}$ . These are smooth manifolds of dimensions  $m \cdot n$  and  $2 \cdot m \cdot n$ , respectively. From this,  $GL_n(\mathbb{R})$  is an  $n^2$  matrix since it is an open subset of  $\mathbb{R}^{n^2}$ . It is the complement of the pre-image of  $\det(0)$ . Since the determinant is continuous, and points are closed, the pre-image is closed. Hence the complement is open. Thus  $GL_n(\mathbb{R})$  is an open subset of a smooth manifold, and hence is a smooth manifold itself.

An  $n$  dimensional manifold is a topological manifold  $(M, \tau)$  with a smooth differentiable structure  $\mathcal{A}$ . In particular, for charts  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  such that  $\mathcal{U} \cap \mathcal{V} \neq \emptyset$ , then  $\varphi \circ \psi^{-1}$  is a diffeomorphism from an open subset of  $\mathbb{R}^n$  to another open subset of  $\mathbb{R}^n$ . A function  $\phi : M \rightarrow \mathbb{R}$  is smooth if for all  $p \in M$  there is a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}$  such that  $\phi \circ \varphi^{-1}$  is a smooth function from a subset  $\mathbb{R}^n$  to  $\mathbb{R}$ .

**Example 37.6.13** Matrices of full rank form a smooth manifold. Let  $m < n$  and let  $\text{Mat}_{m \times n}^F$  be the set of all  $m \times n$  matrices  $A$  such that  $A$  has full rank  $m$ . For if  $A$  has full rank, then there is an  $m \times m$  submatrix  $B$  such that  $\det(B) \neq 0$ . That is,  $A = [B|C]$ , where  $C$  is some  $m \times n - m$  matrix. Since the determinant is continuous, there is an open neighborhood about  $B$ , call it  $\mathcal{U}$ , such that none of the elements have zero determinant. Let  $\mathcal{V}$  be the set of all  $m \times (n - m)$  matrices. Then  $A$  is an element of  $\mathcal{U} \times \mathcal{V}$ , which is the product of open, and hence open in the product topology. Thus  $\text{Mat}_{n \times M}^F$  is an open subset of the entire space, and is hence a smooth manifold.

**Example 37.6.14** Given an open subset  $\mathcal{U} \subseteq \mathbb{R}^n$  and a smooth function  $f : \mathcal{U} \rightarrow \mathbb{R}$ , the graph of  $f$ , as a subset of  $\mathbb{R}^n \times \mathbb{R}$ , is a smooth manifold.

**Theorem 37.6.15.** *If  $\mathcal{U} \subseteq \mathbb{R}^n \times \mathbb{R}^k$  is open, if  $\varphi : \mathcal{U} \rightarrow \mathbb{R}^k$  is a continuously differentiable function,  $(a, b) \in \mathcal{U}$  and  $c = \varphi(a, b)$ , then the implicit function theorem.*

**Example 37.6.15** Let  $\mathcal{U} \subseteq \mathbb{R}^n$  be open, and  $\varphi : \mathcal{U} \rightarrow \mathbb{R}$  a smooth function. Level sets are manifolds when the implicit function theorem applies.

**Example 37.6.16** Let  $V$  be an  $n$  dimensional real vector space,  $k \in \mathbb{Z}_n$  such that  $k \neq 0$ , and let  $G_k(V)$  be the set of all  $k$  dimensional subspaces of  $V$ . Then this is a smooth manifold of dimension  $k(n - k)$  with a certain topology and smooth structure.

**Theorem 37.6.16** (Smooth Manifold Chart Lemma). *Let  $M$  be a set,  $\{\mathcal{U}_\alpha\}$  a collection of subsets of  $M$ , with a collection of maps  $\varphi_\alpha : M \rightarrow \mathbb{R}^n$  such that for all  $\alpha \in J$  there is injection and  $\varphi_\alpha(\mathcal{U}_\alpha)$  is open, for all  $\alpha, \beta$  such that  $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$  it is true that  $\varphi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$  is open in  $\mathbb{R}^n$  and  $\varphi_\beta \circ \varphi_\alpha^{-1}$  is smooth, countably many  $\mathcal{U}_\alpha$  cover  $M$ , and for all  $p \neq q$  then either  $p, q \in \mathcal{U}_\alpha$  for some  $\alpha$  or there are disjoint  $\mathcal{U}_\alpha, \mathcal{U}_\beta$  such that  $p \in \mathcal{U}_\alpha$  and  $q \in \mathcal{U}_\beta$ .*

**Example 37.6.17** Let  $V$  be an  $n$  dimensional real vector space over  $\mathbb{R}$ ,  $1 \leq k \leq n$  and integer, and let  $G_k(V)$  be the set of all  $k$  dimensional subspaces of  $V$ . Using the smooth manifold chart lemma  $G_k(V)$  can be given the structure of a  $C^\infty$  manifold of dimension  $k(n - k)$ . There is a natural collection  $\{(\mathcal{U}_\alpha, \varphi_\alpha)\}$  where  $\mathcal{U}_\alpha \subseteq G_k(V)$ ,  $G_k(V)$  is covered by such sets, and  $\mathcal{U}_\alpha$  is in bijection with  $\mathbb{R}^{k(n-k)}$ . Let  $P \in G_k(V)$  and  $Q$  the complementary subspace of  $P$  in  $V$ . Then  $V = P \oplus Q$  and the dimension of  $Q$  is  $n - k$ . Let  $\text{Hom}(P, Q)$  be the space of linear transformations from  $P$  to  $Q$ . Then  $\text{Hom}(P, Q)$  is equal to the space of real

$(n-k) \times k$  matrices, which is homeomorphic as a topological space to  $\mathbb{R}^{k(n-k)}$ . Let  $\mathcal{U}_Q = \{W \in G_k(V) | W \cap Q = \{0\}\}$ . Then  $P \in \mathcal{U}_Q$ . Let  $L \in \text{Hom}(P, Q)$ , then the graph of  $L$  is a  $k$  dimensional subspace of  $V$  with trivial intersection with  $Q$  and all such subspaces of  $V$  arise as some graph of  $L \in \text{Hom}(P, Q)$ . But since  $V = P \oplus Q$ , we have that  $\pi_P : V \rightarrow P$  and  $\pi_Q : V \rightarrow Q$ . Let  $W \in \mathcal{U}_Q$ . Then  $\pi_P : W \rightarrow P$  is an isomorphism and  $L = \pi_Q \circ \pi^{-1} : P \rightarrow Q$ .

## 37.7 Smooth Manifolds with Boundary

First we defined smoothness on arbitrary subsets of  $\mathbb{R}^n$ . Given  $A \subseteq \mathbb{R}^n$  and  $f : A \rightarrow \mathbb{R}^k$  we say that  $f$  is smooth if for all  $p \in A$  there is an open subset  $\mathcal{U} \subseteq \mathbb{R}^n$  that contains  $p$  and a smooth function  $F : \mathcal{U} \rightarrow \mathbb{R}^k$  such that  $F|_{\mathcal{U} \cap A} = f|_{\mathcal{U} \cap A}$ . Let  $\mathcal{U}$  be a subset of  $\mathbb{H}^n$  that is open in the subspace topology. Then  $F : \mathcal{U} \rightarrow \mathbb{R}^k$  is smooth if there is an open subset  $\tilde{\mathcal{U}} \subseteq \mathbb{R}^n$  (in the usual Euclidean topology) such that  $p \in \tilde{\mathcal{U}}$  and a smooth function  $\tilde{F} : \tilde{\mathcal{U}} \rightarrow \mathbb{R}^k$  such that  $F$  and  $\tilde{F}$  are identical on  $\mathcal{U} \cap \tilde{\mathcal{U}}$ . For such a smooth function  $F : \mathcal{U} \rightarrow \mathbb{R}^k$  we see that for the *interior* points,  $F$  is smooth in the usual sense. By the continuity of  $F$  on such points, all of the values of  $\tilde{F}$  on the boundary points are determined uniquely. That is, we can take a sequence of points  $a : \mathbb{N} \rightarrow \mathbb{U}$  that converge to a point on  $\partial \mathbb{H}^n$ , and by the continuity of all of the partial derivatives, of all orders, of  $\tilde{F}$ , we can evaluate  $\tilde{F}$  and its derivatives by examining  $\tilde{F}(a_n)$  (and similarly for partials). Thus any two extensions of  $F$  are the same.

**Example 37.7.1** Let  $B^2 \subseteq \mathbb{R}^2$  be the open unit disk and  $\mathcal{U} = B^2 \cap \mathbb{H}^2$ . Let  $f : \mathcal{U} \rightarrow \mathbb{R}$  be given by  $f(x, y) = \sqrt{1 - x^2 - y^2}$ . Then  $f$  is smooth on  $\mathcal{U}$  since  $\tilde{f} : B^2 \rightarrow \mathbb{R}$  is smooth,  $\tilde{f}(x, y) = \sqrt{1 - x^2 - y^2}$ . On the other hand, the function  $g : \mathcal{U} \rightarrow \mathbb{R}$  defined by  $g(x, y) = \sqrt{y}$  does not have a smooth extension since  $\partial g / \partial y = 1/2\sqrt{y}$ , and this tends to infinity as  $y$  tends to zero.

**Definition 37.7.1** A smooth structure on a topological manifold with boundary  $M$  is a maximal atlas  $\mathcal{A}$  consisting of charts  $(\mathcal{U}, \varphi)$  with smooth overlap in the sense discussed above.  $M$  equipped with such an atlas is called a smooth manifold with boundary. Here we've fixed the dimension.

We previously said that a product of smooth manifolds without boundary has a natural smooth structure of a manifold without boundary. It then becomes natural to ask about the Cartesian product of two manifolds with non-empty boundary.

**Theorem 37.7.1.** *If  $M_1, \dots, M_k$  are  $C^\infty$  manifolds without boundary and  $N$  is a  $C^\infty$  manifold with boundary, then the product is a  $C^\infty$  manifold with boundary and the boundary is the product of  $M_1, \dots, M_k$  with  $\partial N$ .*

**Theorem 37.7.2** (Smooth Invariance of the Boundary). *If  $M$  is a smooth manifold with boundary, if  $p \in M$ , and if  $(\mathcal{U}, \varphi)$  is a smooth structure such*

that  $\varphi[\mathcal{U}] \subseteq \mathbb{H}^n$  and  $\varphi(p) \in \partial\mathbb{H}^n$ , then the same is true for any smooth chart containing  $p$ .

*Proof.* For by the inverse function theorem, suppose  $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}^n$  are open subsets and  $f : \mathcal{U} \rightarrow \mathcal{V}$  is a smooth function. If  $p \in \mathcal{U}$  is such that  $DF_p$  is a non-singular matrix then there are neighborhoods  $\mathcal{U}_0 \subseteq \mathcal{U}$  which contain  $p$  and  $\mathcal{V}_0 \subseteq \mathcal{V}$  which contain  $F(p)$  such that  $F|_{\mathcal{U}_0}$  is a diffeomorphism from  $\mathcal{U}_0$  to  $\mathcal{V}_0$ . If this is true for every such point in  $\mathcal{U}$ , then  $f$  is an open mapping. If  $f$  is injective, then it is a diffeomorphism. Now, suppose to the contrary that  $p \in M$  is such that there exists an interior chart and a boundary chart both containing  $p$  and such that the image of  $p$  lies on the boundary of  $\mathbb{H}^n$ .  $\square$

## 37.8 Diffeomorphisms

**Definition 37.8.1** A smooth map from a manifold  $M$  to  $\mathbb{R}^k$  is a function  $\phi : M \rightarrow \mathbb{R}^k$  such that for all  $p \in M$  there exists a chart  $(\mathcal{U}, \varphi)$  such that  $p \in \mathcal{U}$  and  $f \circ \varphi^{-1}$  is smooth.

Since the charts in a smooth manifold are required to overlap smoothly, this definition of smoothness does not depend on the choice of chart.

**Theorem 37.8.1.** *If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $f : M \rightarrow \mathbb{R}^k$  is smooth, if  $p \in M$ , and if  $(\mathcal{U}, \varphi) \in \mathcal{A}$  is such that  $p \in M$ , then  $f \circ \varphi^{-1}$  is smooth.*

If  $M$  is a smooth manifold with boundary, then the definition is more or less the same except in the case when  $(\mathcal{U}, \varphi)$  is a boundary chart, we require that  $f \circ \varphi^{-1}$  to be able to extend to a smooth function.

**Theorem 37.8.2.** *If  $M$  and  $N$  are smooth manifolds,  $f : M \rightarrow N$  is smooth if and only if for all  $p \in M$  there is a chart  $(\mathcal{U}, \varphi)$  containing  $p$  and  $(\mathcal{V}, \psi)$  containing  $f(p)$  such that  $\mathcal{U} \cap f^{-1}(\mathcal{V})$  is open and  $\psi \circ f \circ \varphi^{-1}$  is smooth.*

**Theorem 37.8.3.** *Every constant map is smooth. The identity map is smooth. The inclusion map from an open subset is smooth. The composition of smooth functions is smooth.*

**Example 37.8.1** Let  $S^1$  be equipped with its usual smooth structure and  $\phi : \mathbb{R} \rightarrow S^1$  be the map  $\phi(\theta) = (\cos(\theta), \sin(\theta))$ . We can generalize this to the  $n$  torus,  $\mathbb{T}^n$ . Defined  $f : \mathbb{R}^n \rightarrow \mathbb{T}^n$  to be:

$$f(\boldsymbol{\theta}) = \exp(i2\pi\boldsymbol{\theta}) \tag{37.8.1}$$

**Example 37.8.2** The inclusion map  $\iota : S^n \rightarrow \mathbb{R}^{n+1}$  is smooth.

**Example 37.8.3** Given  $\mathbb{R}^{n+1} \setminus \{0\}$  and the quotient map  $q : \mathbb{R}^{n+1} \rightarrow \mathbb{RP}^n$ , where  $\mathbb{RP}^n$  is the real projective  $n$  space, then  $q$  is a smooth mapping.

**Definition 37.8.2** A diffeomorphism from a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  to a smooth manifold  $(N, \tau_N, \mathcal{A}_N)$  is a smooth bijective function  $\phi : M \rightarrow N$  such that  $\phi^{-1}$  is smooth.

**Example 37.8.4** The open  $n$  ball  $\mathbb{B}^n$  is diffeomorphic to the entirety of  $\mathbb{R}^n$ . Use the mapping:

$$f(\mathbf{x}) = \frac{\mathbf{x}}{\sqrt{1 - \|\mathbf{x}\|_2^2}} \quad (37.8.2)$$

This is smooth on  $\mathbb{B}^n$ , with smooth inverse:

$$f^{-1}(\mathbf{x}) = \frac{\mathbf{x}}{\sqrt{1 + \|\mathbf{x}\|_2^2}} \quad (37.8.3)$$

**Example 37.8.5** On  $\mathbb{R}$  there are two atlases that have different maximal smooth atlases. Let  $\mathcal{A}_1 = \{(\mathbb{R}, \text{id}_{\mathbb{R}})\}$  and  $\mathcal{A}_2 = \{(\mathbb{R}, x^3)\}$ . However, the mapping  $f(x) = x^{1/3}$  is a diffeomorphism.

**Theorem 37.8.4.** *If  $M$  is a topological space that is covered by countably many  $n$  dimensional coordinate charts, then  $M$  is second countable.*

**Theorem 37.8.5.** *If  $M$  is a topological space covered by  $n$  dimensional coordinate charts such that for any  $p, q \in M$  there exists  $\alpha \in J$  such that  $p, q \in U_\alpha$  or there exists  $\alpha \neq \beta \in J$  such that  $p \in U_\alpha$ ,  $q \in U_\beta$ , and  $U_\alpha \cap U_\beta = \emptyset$ , then  $M$  is Hausdorff.*

## 37.9 Partitions of Unity

**Theorem 37.9.1.** *The function  $f : \mathbb{R} \rightarrow \mathbb{R}$  defined by:*

$$f(t) = \begin{cases} \exp\left(\frac{1}{t}\right), & t > 0 \\ 0, & t \leq 0 \end{cases} \quad (37.9.1)$$

*is smooth.*

*Proof.*  $f$  is smooth at all points  $t \neq 0$  and thus the only point left to check is  $t = 0$ . Apply L'Hôpital here.  $\square$

**Theorem 37.9.2.** *If  $r_1, r_2 \in \mathbb{R}$  are real numbers such that  $0 < r_1$  and  $r_1 < r_2$ , then there is a function  $h : \mathbb{R} \rightarrow \mathbb{R}$  such that for all  $t \in \mathbb{R}$  such that  $t \leq r_1$  it is true that  $h(t) = 1$ , for all  $t \in (r_1, r_2)$  it is true that  $h(t) \in (0, 1)$ , and for all  $t \geq r_2$  it is true that  $h(t) = 0$ .*

*Proof.* Let  $f(t)$  be defined by:

$$f(t) = \begin{cases} \exp\left(\frac{1}{t}\right), & t > 0 \\ 0, & t \leq 0 \end{cases} \quad (37.9.2)$$

define  $h : \mathbb{R} \rightarrow \mathbb{R}$  by:

$$\frac{f(r_2 - t)}{f(r_2 - t) + f(t - r_1)} \quad (37.9.3)$$

□

**Theorem 37.9.3.** *If  $0 < r_1 < r_2$  there is a bump function  $H : \mathbb{R}^n \rightarrow \mathbb{R}$  such that:*

- $H = 1$  on  $\text{Cl}(B_{r_1}(0))$ .
- $H = 0$  on  $\mathbb{R}^n \setminus B_{r_2}(0)$ .
- $H$  is strictly between 0 and 1 in between these balls.

*Proof.* Let  $H(\mathbf{x}) = h(\|\mathbf{x}\|)$ . □

**Definition 37.9.1** If  $X$  is a topological space,  $f : X \rightarrow \mathbb{R}^n$ , the support is  $\text{supp}(f) = \text{Cl}_\tau(\{x \in X \mid f(x) \neq 0\})$ , compact support of  $\text{supp}(f)$  is compact.

**Definition 37.9.2** Let  $X$  be a topological space and  $\mathcal{O}$  be an open cover of  $X$ . A partition of unity subordinate to  $\mathcal{O}$  is a collection of continuous functions  $f_\alpha : X \rightarrow \mathbb{R}$  such that:

- $f_\alpha[X] \in [0, 1]$
- $\text{supp}(f_\alpha) \in \mathcal{U}_\alpha$
- $\{\text{supp}(f_\alpha)\}$  is locally finite.
- $\sum f_\alpha = 1$  for all points.

The last part of this definition is well defined since the support of all of the functions form a locally finite collection, and hence only finitely many such functions contribute to the sum at any point, and hence the sum is well defined everywhere. Smooth partition of unity is a partition of unity with smooth functions.

#### Theorem 37.9.4: Existence of Smooth Partitions of Unity

If  $M$  is a smooth manifold with boundary and if  $\mathcal{O}$  is an open cover, then there exists a smooth partition of unity subordinate to  $\mathcal{O}$ .

*Proof.* Firstly suppose the boundary is empty. For every  $\mathcal{U}_\alpha \in \mathcal{O}$  there is a basis  $B_\alpha$  consisting of regular coordinate balls. Let  $B$  be the union of all the  $B_\alpha$ . Then this is a basis for the topology on  $M$ . But  $M$  is a smooth manifold and is hence paracompact, and thus there is a locally finite refinement of  $B$ . Moreover, there's a locally finite refinement consisting of elements of  $B$ . But then the collection of the closures of the  $B_j$  will be locally finite. □

Application: We'll use this to show every manifold admits a Riemannian metric.

## 37.10 Vector Fields

If  $M$  is an  $n$  dimensional smooth manifold, then for all  $p \in M$  the space  $T_p M$  is an  $n$  dimensional vector space over  $\mathbb{R}$ . If  $\mathcal{U} \subseteq M$  is open,  $p \in \mathcal{U}$ , then  $T_p \mathcal{U}$  is isomorphic to  $T_p M$  with the inclusion map  $\iota : \mathcal{U} \rightarrow M$  giving the isomorphism  $d\iota_p$  (the differential pushforward of  $\iota$ ). Let  $(\mathcal{U}, \varphi)$  be a coordinate chart at  $p \in M$ . Then  $\varphi : \mathcal{U} \rightarrow \varphi[\mathcal{U}]$  is a diffeomorphism and hence  $d\varphi_p : T_p \mathcal{U} \rightarrow T_{\varphi(p)} \mathbb{R}^n$ . If  $M_1, \dots, M_k$  are smooth manifolds,  $\alpha : T_p \prod M_j \rightarrow \bigoplus T_{p_j} M_j$ . Let's compare  $DF_p : \mathbb{R}_p^n \rightarrow \mathbb{R}_{F(p)}^m$  with the differential pushforward  $dF_p$ . Let  $M$  be a smooth manifold and  $(\mathcal{U}, \varphi)$  a chart on  $M$ . For any  $p \in \mathcal{U}$ ,  $d\varphi_p$  gives an isomorphism between  $T_p M$  and  $\mathbb{R}^n$ .

## 37.11 Immersions, Submersions, and Embeddings

Given a smooth function  $F : M \rightarrow N$ ,  $dF_p : T_p M \rightarrow T_{F(p)} N$  represents the best linear approximation of the function  $F$  at the point  $p$ . We can try and understand  $F$  via the properties of the differential pushforward.

**Definition 37.11.1** Let  $F : M \rightarrow N$  be a smooth function. The rank of  $F$  at a point  $p \in M$  is the rank of the differential  $d_p F : T_p M \rightarrow T_{F(p)} N$ .  $F$  is said to be of full rank if for all  $p \in M$ , the rank of  $F$  at  $p$  is  $\min(m, n)$ , and  $m$  and  $n$  are the dimensions of  $M$  and  $N$ , respectively.

**Definition 37.11.2** An immersion is a smooth function  $\phi : M \rightarrow N$  such that  $d_p \phi$  is injective.

**Definition 37.11.3** A submersion is a smooth function  $\phi : M \rightarrow N$  such that  $d_p \phi$  is surjective.

**Theorem 37.11.1.** *If  $\phi : M \rightarrow N$  is smooth, if  $p \in M$ , and if  $d\phi_p : T_p M \rightarrow T_{\phi(p)} N$  is surjective, then there is a neighborhood  $\mathcal{U}$  such that  $\phi|_{\mathcal{U}}$  is a submersion.*

*Proof.* Let  $\hat{\phi} = \psi \circ \phi \circ \varphi^{-1}$  be a coordinate representation of  $\phi$  in a neighborhood of  $p \in M$ . Then the Jacobian matrix of  $d\phi_p$  is of full rank by hypothesis. Since the set of full rank matrices is open in  $\text{Mat}_{m \times n}(\mathbb{R})$ , the result follows.  $\square$

**Example 37.11.1** Projection maps are submersions.

**Example 37.11.2** A smooth curve  $\gamma : I \rightarrow M$  is an immersion if and only if  $\dot{\gamma}(t) \neq 0$  for all  $t \in I$ .

**Example 37.11.3** The natural projection map of the tangent bundle  $\pi : TM \rightarrow M$  is a submersion.

**Definition 37.11.4** A smooth mapping is a local diffeomorphism if for every  $p \in M$  there is open subset  $\mathcal{U}$  such that  $\phi|_{\mathcal{U}}$  is a diffeomorphism.

**Example 37.11.4** The quotient map of  $\mathbb{S}^n$  to  $\mathbb{RP}^n$  is a local diffeomorphism.

**Theorem 37.11.2.** If  $M$  and  $N$  are smooth manifolds of the same dimension, if  $\phi : M \rightarrow N$  is smooth, if  $p \in M$  is such that  $d_p\phi$  is invertible, then there is a connected neighborhood  $\mathcal{U}_0$  of  $p$  and  $\mathcal{V}_0$  of  $\phi(p)$  such that  $\phi|_{\mathcal{U}_0}$  is a diffeomorphism onto  $\mathcal{V}_0$ .

**Theorem 37.11.3.** Suppose  $M$  and  $N$  are smooth manifolds without boundary and  $\phi : M \rightarrow N$  a smooth function. Then  $\phi$  is a local diffeomorphism if and only if it is a smooth immersion and a smooth submersion.

**Theorem 37.11.4** (Theorem on Rank). If  $M$  and  $N$  are smooth manifolds,  $\phi : M \rightarrow N$  is smooth with constant rank  $r$ , and if  $p \in M$ , then there are charts  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  with  $p \in \mathcal{U}$  and  $\phi(p) \in \mathcal{V}$  such that

*Proof.* It do be like that some times though. □

## 37.12 Notes from O'Neill (Chapter 1)

A smooth function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is one such that  $\pi_k \circ f : \mathbb{R}^n \rightarrow \mathbb{R}$  is smooth for all  $k \in \mathbb{Z}_m$ . A chart  $(\mathcal{U}, \varphi)$  in a topological space  $(X, \tau)$  is an open set  $\mathcal{U} \subseteq X$  and an injective continuous open mapping  $\varphi : \mathcal{U} \rightarrow \mathbb{R}^n$ . The coordinate functions of the chart  $(\mathcal{U}, \varphi)$  are the functions  $x^k : \mathcal{U} \rightarrow \mathbb{R}$  defined by  $x^k = \pi_k \circ \varphi$  for all  $k \in \mathbb{Z}_n$ . Smoothly overlapping charts are charts  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  such that either  $\mathcal{U} \cap \mathcal{V} = \emptyset$  or both  $\varphi \circ \psi^{-1}$  and  $\psi \circ \varphi^{-1}$  are smooth functions (in the Euclidean sense). The domain of  $\varphi \circ \psi^{-1}$  is  $\psi(\mathcal{V})$ , which is an open subset of  $\mathbb{R}^n$ , and hence this composition is a function from an open subset of  $\mathbb{R}^n$  to  $\mathbb{R}^n$  and thus asking if it is smooth is a valid well defined question. Similarly for  $\psi \circ \varphi^{-1}$ . A smooth atlas is a collection of charts  $(\mathcal{U}_\alpha, \varphi_\alpha)$  such that the  $\mathcal{U}_\alpha$  cover  $X$  and each chart overlaps smoothly. A maximal smooth atlas is a smooth atlas  $\mathcal{A}$  on  $(X, \tau)$  such that for all smooth atlases  $\mathcal{A}'$  such that  $\mathcal{A} \subseteq \mathcal{A}'$  it is true that  $\mathcal{A} = \mathcal{A}'$ .

**Theorem 37.12.1.** If  $(X, \tau)$  is a topological space, if  $\mathcal{A}$  and  $\mathcal{A}'$  are smooth atlases on  $(X, \tau)$ , then they are compatible if and only if  $\mathcal{A} \cup \mathcal{A}'$  is a smooth atlas on  $(X, \tau)$ .

*Proof.* If they are compatible, then every chart in  $\mathcal{A}$  overlaps smoothly with every chart in  $\mathcal{A}'$ , and vice versa. But  $\mathcal{A}$  is a smooth atlas and thus every chart

in  $\mathcal{A}$  overlaps smoothly with every other chart in  $\mathcal{A}$ . Thus, every chart in  $\mathcal{A}$  overlaps smoothly with every chart in  $\mathcal{A} \cup \mathcal{A}'$ , and similarly for  $\mathcal{A}'$ . Hence if  $\mathcal{A}$  and  $\mathcal{A}'$  are compatible, then  $\mathcal{A} \cup \mathcal{A}'$  is a smooth atlas. Moreover, if  $\mathcal{A} \cup \mathcal{A}'$  is a smooth atlas, then every element of  $\mathcal{A} \cup \mathcal{A}'$  overlaps smoothly with every other element of  $\mathcal{A} \cup \mathcal{A}'$ . In particular, every element of  $\mathcal{A}$  overlaps smoothly with every element of  $\mathcal{A}'$ , and vice-versa. Hence,  $\mathcal{A}$  and  $\mathcal{A}'$  are compatible.  $\square$

**Theorem 37.12.2.** *If  $\mathcal{A}$  is a smooth atlas on  $(X, \tau)$ , then there is a unique maximal smooth atlas  $\mathcal{C}$  such that  $\mathcal{A} \subseteq \mathcal{C}$ .*

*Proof.* The set of all atlases on  $(X, \tau)$  forms a partially ordered set by inclusion. Given any chain of atlases, by the previous theorem the union is then again an atlas compatible with all atlases in the chain. That is, every chain has an upper bound. Hence by Zorn's lemma every Atlas has a maximal atlas. If there is another maximal atlas, the union will again be a smooth atlas, contradicting maximality, and hence there is a unique maximal smooth atlas.  $\square$

A smooth manifold is a second countable Hausdorff topological space with a maximal smooth atlas.

**Theorem 37.12.3.** *If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $(\mathcal{U}, \varphi) \in \mathcal{A}$ , and if  $\mathcal{V}$  is an open subset such that  $\mathcal{U} \cap \mathcal{V} \neq \emptyset$ , then  $(\mathcal{U} \cap \mathcal{V}, \varphi|_{\mathcal{U} \cap \mathcal{V}}) \in \mathcal{A}$ .*

*Proof.* For since  $\mathcal{U}$  and  $\mathcal{V}$  are open,  $\mathcal{U} \cap \mathcal{V}$  is open. But then  $\varphi|_{\mathcal{U} \cap \mathcal{V}}$  is a continuous and injective open mapping. If  $(\mathcal{O}, \psi) \in \mathcal{A}$ , then it overlaps smoothly with  $(\mathcal{U}, \varphi)$ . But smoothness is a local property, and hence it overlaps smoothly with  $(\mathcal{U} \cap \mathcal{V}, \varphi|_{\mathcal{U} \cap \mathcal{V}})$  and since  $\mathcal{A}$  is maximal, it contains this chart.  $\square$

**Theorem 37.12.4.** *If  $(X, \tau, \mathcal{A})$  is a smooth manifold and  $A \in \tau$  is an open subset of  $X$ , then  $(A, \tau|_A, \mathcal{A}|_A)$  is a smooth manifold where  $\tau|_A$  is the subspace topology and is the subatlas defined by:*

$$\mathcal{A}|_A = \{(\mathcal{U}, \varphi) \in \mathcal{A} \mid \mathcal{U} \subseteq A\} \quad (37.12.1)$$

**Definition 37.12.1** The product function of a function  $\varphi : \mathcal{U} \rightarrow X$  and a function  $\psi : \mathcal{V} \rightarrow Y$  is the function  $\varphi \times \psi : \mathcal{U} \times \mathcal{V} \rightarrow X \times Y$  defined by:

$$\varphi \times \psi(u, v) = (\varphi(u), \psi(v))$$

**Theorem 37.12.5.** *If  $(X, \tau_X, \mathcal{A}_X)$  and  $(Y, \tau_Y, \mathcal{A}_Y)$  smooth manifolds of dimension  $n$  and  $m$ , respectively, then  $(X \times Y, \tau_{X \times Y}, \mathcal{A}_{X \times Y})$  is a smooth manifold of dimension  $n + m$ .*

The natural smooth structure on  $\mathbb{R}^n$  can be seen as the product manifold of  $n$  copies of  $\mathbb{R}$ .

### 37.12.1 Smooth Functions

**Definition 37.12.2** A smooth function from a manifold  $(M, \tau, \mathcal{A})$  into  $\mathbb{R}$  is a function  $f : M \rightarrow \mathbb{R}$  such that for all  $(\mathcal{U}, \varphi) \in \mathcal{A}$  the function  $f \circ \varphi^{-1}$  is Euclidean smooth.

**Theorem 37.12.6.** *A function  $f : M \rightarrow \mathbb{R}$  is smooth if and only if there is a cover  $\mathcal{C}$  or charts such that for all  $(\mathcal{U}, \varphi) \in \mathcal{C}$  the function  $f \circ \varphi^{-1}$  is Euclidean smooth.*

*Proof.* Since the atlas of a smooth manifold is maximal,  $\mathcal{C}$  is contained in it. But charts in a smooth atlas overlap smoothly, and thus if  $(\mathcal{V}, \psi)$  is a different chart then we have:

$$f \circ \psi^{-1} = f \circ (\varphi^{-1} \circ \varphi) \circ \psi^{-1} = (f \circ \varphi^{-1}) \circ (\varphi \circ \psi^{-1}) \quad (37.12.2)$$

The composition of smooth functions, and hence smooth.  $\square$

**Definition 37.12.3** A smooth function from a manifold  $M$  to  $\mathbb{R}^m$  is a function  $f : M \rightarrow \mathbb{R}^m$  such that  $f \circ \pi_k$  is smooth for all  $k \in \mathbb{Z}_m$ .

**Definition 37.12.4** The set of all smooth functions from a smooth manifold  $(M, \tau, \mathcal{A})$  to  $\mathbb{R}$  is denoted  $C^\infty(M, \mathbb{R})$ .

**Theorem 37.12.7.** *If  $f, g \in C^\infty(M, \mathbb{R})$ , then  $f + g \in C^\infty(M, \mathbb{R})$  and  $fg \in C^\infty(M, \mathbb{R})$ .*

The space  $C^\infty(M, \mathbb{R})$  is a vector space over  $\mathbb{R}$ . Moreover, it is an algebra over  $\mathbb{R}$  since multiplication of vectors makes sense.  $fg$  is simply the function  $(fg)(p) = f(p)g(p)$ . This operation is distributive and associative, and hence  $C^\infty(M, \mathbb{R})$  is an associative algebra over the field of real numbers  $\mathbb{R}$ . Moreover, there is a unital element since the constant mapping  $f(p) = 1$  is such that  $fg = g$  for all  $g \in C^\infty(M, \mathbb{R})$ . Hence, this space is also a unital algebra. Another algebraic structure that  $C^\infty(M, \mathbb{R})$  has is that of a commutative ring. We can now extend the notion of smoothness to functions between arbitrary manifolds.

**Definition 37.12.5** A smooth function from a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  to a smooth manifold  $(N, \tau_N, \mathcal{A}_N)$  is a function  $f : M \rightarrow N$  such that for all  $(\mathcal{U}, \varphi) \in \mathcal{A}_M$  and for all  $(\mathcal{V}, \psi) \in \mathcal{A}_N$ , the function  $\psi \circ f \circ \varphi^{-1}$  is Euclidean smooth.

This function  $\psi \circ f \circ \varphi^{-1}$  is defined on the overlaps of the domains in question. That is, the domain is  $\varphi(\mathcal{U} \cap f^{-1}(\mathcal{V}))$ , which is an open subset of  $\mathbb{R}^n$  since  $\varphi$  is an open mapping,  $\mathcal{U}$  is open, and  $f^{-1}(\mathcal{V})$  is open since  $f$  is smooth, and hence continuous. The range is some subset of  $\mathbb{R}^m$ . Specifically, the range is the set  $\psi[\mathcal{V} \cap \phi[\mathcal{U}]]$ , so we defined the function:

$$\psi \circ \phi \circ \varphi^{-1} : \varphi[\mathcal{U} \cap \phi^{-1}(\mathcal{V})] \rightarrow \psi[\mathcal{V} \cap \phi[\mathcal{U}]] \quad (37.12.3)$$

The schematic for this can be seen in Fig. 37.17.

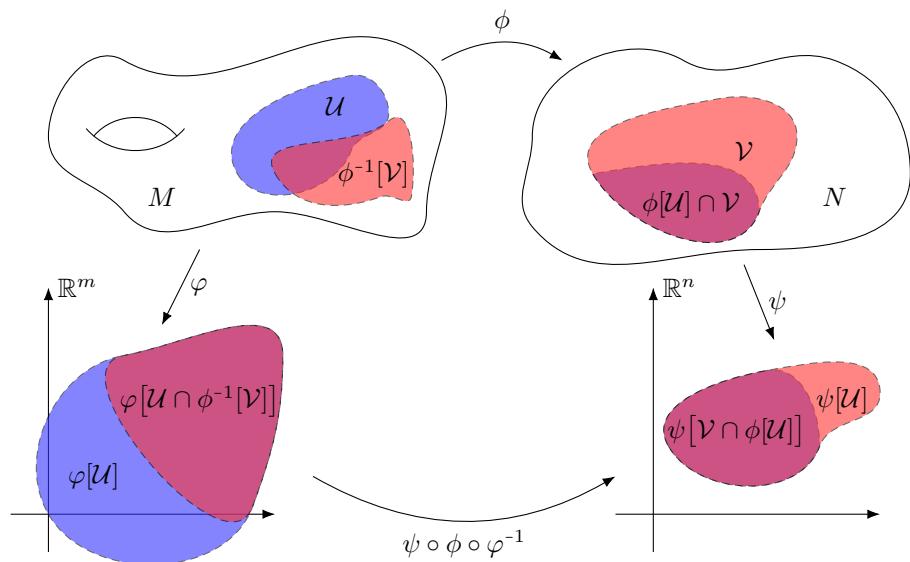


Fig. 37.17: Smooth Function Between Manifolds

**Theorem 37.12.8.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold, then the identity mapping  $\text{id}_X : X \rightarrow X$  is a smooth function.

**Theorem 37.12.9.** If  $(X, \tau_X, \mathcal{A}_X)$ ,  $(Y, \tau_Y, \mathcal{A}_Y)$ , and  $(Z, \tau_Z, \mathcal{A}_Z)$  are smooth manifolds, if  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  are smooth functions, then  $g \circ f$  is a smooth function.

**Theorem 37.12.10.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold,  $(U, \varphi) \in \mathcal{A}$ , and if  $(U, \tau_U, \mathcal{A}_U)$  is the open submanifold, then  $\varphi : U \rightarrow \mathbb{R}^n$  is a smooth function.

From this we have that all of the coordinate functions are smooth, since  $\varphi \circ \pi_k$  is the composition of smooth functions, which is smooth. Smoothness is a local property. A function is smooth if and only if it is locally smooth at every point.

**Theorem 37.12.11.** If  $(X, \tau_X, \mathcal{A}_X)$  and  $(Y, \tau_Y, \mathcal{A}_Y)$  are smooth manifolds, and if  $f : X \rightarrow Y$  is smooth, then it is continuous.

Here, smoothness is a property of the atlases  $\mathcal{A}_X$  and  $\mathcal{A}_Y$ , whereas continuity is a property of the topologies  $\tau_X$  and  $\tau_Y$ . This theorem shows that topological information is encoded into the structure of atlases.

**Theorem 37.12.12.** *If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $\mathcal{O}$  is an open cover, if  $\mathcal{F}$  is a collection of smooth functions  $\varphi_\alpha : \mathcal{U}_\alpha \rightarrow N$  for each  $\mathcal{U}_\alpha \in \mathcal{O}$ , then there is a unique function  $\varphi|_{\mathcal{U} \cap \mathcal{V}} = \psi|_{\mathcal{U} \cap \mathcal{V}}$   $\varphi : \bigcup \mathcal{O} \rightarrow N$  such that  $\varphi|_{\mathcal{U}_\alpha} = \varphi_\alpha$ .*

This stems from the local property of smoothness. If we can define a bunch of functions on various patches of a manifold, and if these functions agree on the overlap, then we can define a single function on the union of all such sets.

**Definition 37.12.6** A diffeomorphism from a manifold  $(X, \tau_X, \mathcal{A}_X)$  to a manifold  $(Y, \tau_Y, \mathcal{A}_Y)$  is a smooth bijection  $f : X \rightarrow Y$  such that  $f^{-1} : Y \rightarrow X$  is smooth.

By the previous theorem, any diffeomorphism is automatically a homeomorphism since smooth functions are continuous.

**Theorem 37.12.13.** *If  $(X, \tau, \mathcal{A})$  is a smooth manifold, then the identity map  $\text{id}_X : X \rightarrow X$  is a diffeomorphism.*

**Theorem 37.12.14.** *The composition of diffeomorphisms is a diffeomorphism.*

**Theorem 37.12.15.** *The inverse of a diffeomorphism is a diffeomorphism.*

**Theorem 37.12.16.** *The restriction of a diffeomorphism to an open subset of a smooth manifold is a diffeomorphism onto its image.*

**Example 37.12.1** The interval  $(a, b)$  is diffeomorphic to  $(0, 1)$ , which is diffeomorphic to all of  $\mathbb{R}$ .

**Theorem 37.12.17.** *If  $(X, \tau_X, \mathcal{A}_X)$  is a manifold, if  $Y$  is a set, and if  $f : X \rightarrow Y$  is a bijective function, then there is a unique topology  $\tau_Y$  and a unique maximal atlas  $\mathcal{A}_Y$  such that  $f$  is a diffeomorphism.*

*Proof.* That there is a unique topology comes from point-set topology. Define:

$$\tau_Y = \{ f[\mathcal{U}] \mid \mathcal{U} \in \tau_X \} \quad (37.12.4)$$

This is a topology since  $f$  is bijective. That is,  $Y = f[X]$ ,  $\emptyset = f[\emptyset]$ , and the forward image preserves unions and intersections:

$$f \left[ \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \right] = \bigcup_{\mathcal{U} \in \mathcal{O}} f[\mathcal{U}] \quad (37.12.5)$$

$$f[\mathcal{U} \cap \mathcal{V}] = f[\mathcal{U}] \cap f[\mathcal{V}] \quad (37.12.6)$$

Hence,  $\tau_Y$  will be closed to finite intersections and arbitrary unions. It is an open mapping by definition, and it is continuous since  $f$  is a bijection and hence the pre-image of open subsets of  $Y$  will be open subsets of  $X$  by definition. Thus,  $\tau_Y$  is a topology that makes  $f$  a continuous open mapping, and hence a homeomorphism. For the atlas, define:

$$\mathcal{A}_Y = \{(f[\mathcal{U}], f^{-1} \circ \varphi) \mid (\mathcal{U}, \varphi) \in \mathcal{A}_X\} \quad (37.12.7)$$

Thus  $(Y, \tau_Y, \mathcal{A}_Y)$  a smooth manifold that is diffeomorphic to  $(X, \tau_X, \mathcal{A}_X)$ .  $\square$

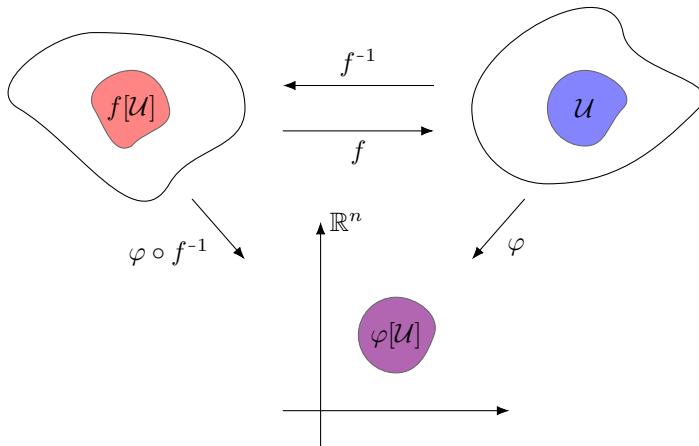


Fig. 37.18: Creating a Manifold Structure on an Arbitrary Set

Since there is a unique topology such that  $f$  is a homeomorphism, and since diffeomorphisms are automatically homeomorphisms, the topology  $\tau_Y$  part is done. To find  $\mathcal{A}_Y$  we simply look at the forward image of all of the charts in  $\mathcal{A}_X$  by the bijection  $f$ .

**Example 37.12.2** Much the way a continuous bijection need not be a homeomorphism, a smooth homeomorphism need not be a diffeomorphism. The classic example is the mapping  $f(x) = x^3$ . This is indeed a homeomorphism, and is smooth, but  $f^{-1}(x) = x^{1/3}$ , which is not differentiable at the origin (let alone smooth).

Every chart is a diffeomorphism from its domain to an open subset of  $\mathbb{R}^n$ .

**Definition 37.12.7** A bump function on an open subset  $\mathcal{U}$  in a topological space  $(X, \tau)$  about a point  $x \in X$  is a function  $f : X \rightarrow \mathbb{R}$  such that  $f[X] = [0, 1]$ ,  $\text{supp}(f) \subseteq \mathcal{U}$  and such that there exists an open subset  $\mathcal{V} \in \tau$  such that  $f[\mathcal{V}] = \{1\}$ .

**Theorem 37.12.18.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $x \in X$ , if  $\mathcal{U} \in \tau$  is such that  $x \in \mathcal{U}$ , then there is a smooth bump function  $f : X \rightarrow \mathbb{R}$  about  $x$  contained in  $\mathcal{U}$ .

*Proof.* For given  $\varepsilon > 0$ , let  $f : \mathbb{R} \rightarrow [0, 1]$  be defined by:

$$f(t) = \begin{cases} \exp\left(-\frac{(2\varepsilon-x)^2}{4\varepsilon^2(x-\varepsilon)^2}\right), & \varepsilon \leq x \leq 2\varepsilon \\ 0, & x < \varepsilon \\ 1, & 2\varepsilon < x \end{cases} \quad (37.12.8)$$

Then  $f(\varepsilon) = 0$ ,  $f(2\varepsilon) = 1$ , and all derivatives evaluate to zero at both points, and hence this function is smooth. Since  $(X, \tau, \mathcal{A})$  is a manifold, there is a chart  $(\mathcal{V}, \varphi)$  such that  $x \in \mathcal{V}$ . But  $x \in \mathcal{U}$ , and  $\mathcal{U} \in \tau$  is open, and hence  $\mathcal{U} \cap \mathcal{V}$  is a non-empty open subset. But  $\mathcal{A}$  is maximal, and so  $(\mathcal{U} \cap \mathcal{V}, \varphi|_{\mathcal{U} \cap \mathcal{V}}) \in \mathcal{A}$ . Relabel this as  $(\mathcal{O}, \psi)$ . Then  $\psi[\mathcal{O}]$  is open and  $\psi(x) \in \psi[\mathcal{O}]$  and thus there is an  $\varepsilon > 0$  such that the ball of radius  $\sqrt{\varepsilon}$  centered about  $\psi(x)$  is contained in  $\psi[\mathcal{O}]$ . Define  $h : M \rightarrow \mathbb{R}$  to be:

$$h(p) = \begin{cases} f(\|\psi(p)\|_2^2), & \|\psi(x) - \psi(p)\|_2^2 < \varepsilon \\ 0, & \text{otherwise} \end{cases} \quad (37.12.9)$$

□

### 37.12.2 Tangent Vectors

We now wish to generalize the concept of the tangent plane found in calculus. To do this needs a generalization of the directional derivative that is used to define such spaces. The key properties are linearity and the Liebniz rule.

**Definition 37.12.8** A tangent vector at a point  $p$  in a manifold  $(X, \tau, \mathcal{A})$  is a linear functional  $v : C^\infty(M, \mathbb{R}) \rightarrow \mathbb{R}$  that is Liebnizian. That is, for all  $a, b \in \mathbb{R}$  and for all  $f, g \in C^\infty(M, \mathbb{R})$ :

$$v(af + bg) = av(f) + bv(g) \quad (\text{Linearity})$$

$$v(fg) = v(f)g(p) + f(p)v(g) \quad (\text{Liebnizian})$$

**Definition 37.12.9** The tangent space at a point  $p$  in a manifold  $(X, \tau, \mathcal{A})$  is the set  $T_p X$  of all tangent vectors to  $p$ .

If we define function addition and scalar multiplication in the usual way, then  $T_p X$  is a vector space. That is:

$$(v + w)(f) = v(f) + w(f) \quad (37.12.10)$$

$$(av)(f) = a(v(f)) \quad (37.12.11)$$

**Definition 37.12.10** The partial derivative of a smooth function  $f \in C^\infty(M, \mathbb{R})$  on a smooth manifold  $(X, \tau, \mathcal{A})$  at a point  $x \in X$  with respect to a chart  $(\mathcal{U}, \varphi)$  in the  $k^{th}$  direction, with  $k \in \mathbb{Z}_n$ , is the value:

$$\partial_k f(x) = \frac{\partial}{\partial x_k} (f \circ \varphi^{-1}) \quad (37.12.12)$$

This is well defined since  $f \circ \varphi^{-1}$  is a smooth function from an open subset of  $\mathbb{R}^n$  into  $\mathbb{R}$ , and thus from analysis we may take derivatives of arbitrary order in any of the components.

**Theorem 37.12.19.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $x \in X$ , if  $\mathcal{U}, \varphi \in \mathcal{A}$ , if  $x \in \mathcal{U}$ , and if  $\partial_k|_x : C^\infty(M, \mathbb{R}) \rightarrow \mathbb{R}$  is defined by  $\partial_k|_x(f) = \partial_k(f)(x)$ , then  $\partial_k|_x \in T_x(X)$ .

Tangent spaces are the best linear approximation to the manifold  $(X, \tau, \mathcal{A})$  about the point  $x$ . This is made clear by the following theorems.

**Theorem 37.12.20.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $x \in X$ , if  $\mathbf{0} : X \rightarrow \mathbb{R}$  is the zero function, and if  $v \in T_x(X)$ , then  $v(\mathbf{0}) = 0$ .

*Proof.* For:

$$v(\mathbf{0}) = v(\mathbf{0} - \mathbf{0}) = v((0)) - v((0)) = 0 \quad (37.12.13)$$

□

**Theorem 37.12.21.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $x \in X$ , if  $\mathbf{1} : X \rightarrow \mathbb{R}$  is the function such that  $f(p) = 1$  for all  $p \in X$ , and if  $v \in T_x(X)$ , then  $v(\mathbf{1}) = 0$ .

*Proof.* For:

$$v(\mathbf{1}) = v((1) \cdot (1)) = v(\mathbf{1}) \cdot \mathbf{1}(x) + \mathbf{1}(x) \cdot v(\mathbf{1}) = v(\mathbf{1}) + v(\mathbf{1}) \quad (37.12.14)$$

Thus, from the cancellation law,  $v(\mathbf{1}) = 0$ .

□

**Theorem 37.12.22.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $\mathcal{U} \in \tau$ , if  $x \in \mathcal{U}$ , if  $f, g \in C^\infty(M, \mathbb{R})$ , if  $f|_{\mathcal{U}} = g|_{\mathcal{U}}$ , and if  $v \in T_x(X)$ , then  $v(f) = v(g)$ .

*Proof.* For there exists a bump function  $\alpha$  about the point  $x$  in the open set  $\mathcal{U}$ . Let  $h = f - g$ . But then:

$$v(\alpha \cdot h)v(\alpha) \cdot h(x) + v(h) \cdot \alpha(x) \quad (37.12.15)$$

But  $\alpha(x) = 1$  and  $h(x) = 0$  since  $h(x) = f(x) - g(x)$ , and  $f|_{\mathcal{U}} = g|_{\mathcal{U}}$  and the point  $x$  is contained in  $\mathcal{U}$ . But  $\alpha \cdot h$  is the zero function. For if  $p \in \mathcal{U}$  then  $h(p) = 0$ , and if  $p \notin \mathcal{U}$  then  $\alpha(p) = 0$ . But then  $v(\alpha \cdot h) = v(\mathbf{0}) = 0$  by the previous theorem. Therefore  $v(h) = 0$ , and  $v(f) = v(g)$ . □

**Theorem 37.12.23.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $x \in X$ , if  $\mathcal{U} \in \tau$  is an open neighborhood of  $x$ , if  $f \in C^\infty(X, \mathbb{R})$  is a constant mapping, and if  $v \in T_x(X)$ , then  $v(f) = 0$ .

*Proof.* For if  $f$  is a constant, then  $f = c \cdot \mathbf{1}$  for some  $c \in \mathbb{R}$ . But then:

$$v(f) = v(c \cdot \mathbf{1}) = cv(\mathbf{1}) = c \cdot 0 = 0 \quad (37.12.16)$$

□

**Theorem 37.12.24.** If  $(X, \tau, \mathcal{A})$  is a smooth manifold, if  $x \in X$ , if  $(\mathcal{U}, \tau) \in \mathcal{A}$  is such that  $x \in \mathcal{U}$ , if  $(\mathcal{U}, \tau_{\mathcal{U}}, \mathcal{A}_{\mathcal{U}})$  is the open submanifold structure of  $\mathcal{U}$ , then there is an isomorphism  $\phi : T_x(\mathcal{U}) \rightarrow T_x(X)$ .

*Proof.* For given  $v \in T_x(\mathcal{U})$ , let  $\phi(v) \in T_x(X)$  be such that  $\phi(v)(f) = v(f|_{\mathcal{U}})$  for all  $f \in C^\infty(X, \mathbb{R})$ . By the previous theorems, this is a bijection and is also linear, and hence an isomorphism.  $\square$

### Theorem 37.12.25: Tangent Space Basis Theorem

If  $(X, \tau, \mathcal{A})$  is a smooth manifold of dimension  $n \in \mathbb{N}$ , if  $p \in X$ , and if  $(\mathcal{U}, \varphi) \in \mathcal{A}$  is a chart such that  $x \in \mathcal{U}$ , then the set  $\{\partial_k|_x | k \in \mathbb{Z}_n\}$  is a basis for  $T_x(X)$ , where  $\partial_k|_x$  is the function that maps  $f \in C^\infty(X, \mathbb{R})$  to  $\partial/\partial x_k(f \circ \varphi^{-1})|_x$ . Moreover, for all  $v \in T_x(X)$ , the following is true:

$$v = \sum_{k \in \mathbb{Z}_n} v(x^k) \partial_k|_p \quad (37.12.17)$$

### 37.12.3 The Differential Pushforward

**Definition 37.12.11** The pushforward of a tangent vector  $v$  in a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  at a point  $p \in M$  by a smooth function  $\phi : M \rightarrow N$  into a smooth manifold  $(N, \tau_N, \mathcal{A}_N)$  is the function  $d\phi_p(v) : C^\infty(N, \mathbb{R}) \rightarrow \mathbb{R}$  defined by:

$$d\phi_p(v)(f) = v(f \circ \phi) \quad (37.12.18)$$

The most important aspect of the pushforward is that it maps tangent vectors in  $M$  to tangent vectors in  $N$ .

**Theorem 37.12.26.** If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if  $\phi : M \rightarrow N$  is a smooth function, if  $p \in M$ , if  $v \in T_p M$ , and if  $d\phi_p$  is the differential pushforward of  $\phi$ , then  $d\phi_p(v) \in T_{\phi(p)} N$ .

*Proof.* For if  $a, b \in \mathbb{R}$ ,  $f, g \in C^\infty(M, \mathbb{R})$ , then:

$$d\phi_p(v)(af + bg) = v((af + bg) \circ \phi) \quad (37.12.19a)$$

$$= v(a(f \circ \phi) + b(g \circ \phi)) \quad (37.12.19b)$$

$$= av(f \circ \phi) + bv(g \circ \phi) \quad (37.12.19c)$$

$$= a d\phi_p(v)(f) + b d\phi_p(v)(g) \quad (37.12.19d)$$

Finally, if  $f, g \in C^\infty(N, \mathbb{R})$ , then:

$$d\phi_p(v)(fg) = v((fg) \circ \phi) \quad (37.12.20a)$$

$$= v((f \circ \phi)(g \circ \phi)) \quad (37.12.20b)$$

$$= v(f \circ \phi)g(\phi(p)) \quad (37.12.20c)$$

$$= v(\phi(p))v(g \circ \phi) \quad (37.12.20d)$$

$$= v(f \circ \phi)g(\phi(p)) + f(\phi(p))v(g \circ \phi) \quad (37.12.20e)$$

$$= d\phi_p(v)(f)g(\phi(p)) + f(\phi(p))d\phi_p(v)(g) \quad (37.12.20f)$$

Thus,  $d\phi_p(v)$  is a tangent vector to  $\phi(p)$ . Hence,  $d\phi_p(v) \in T_{\phi(p)}N$ .  $\square$

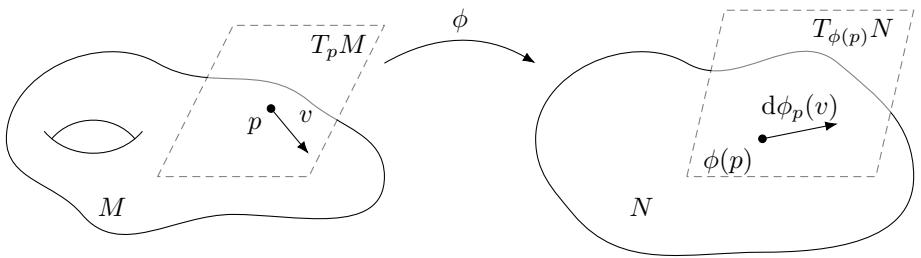


Fig. 37.19: The Pushforward of a Tangent Vector

**Theorem 37.12.27.** If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if  $\phi : M \rightarrow N$  is a smooth function, if  $p \in M$ , if  $(U, \varphi) \in \mathcal{A}_M$  and  $(V, \psi) \in \mathcal{A}_N$  are charts such that  $p \in U$  and  $\phi(p) \in V$ , then:

$$d\phi_p(\partial_{\varphi^k}|_p) = \sum_{k \in \mathbb{Z}_n} \frac{\partial(\psi^k \circ \phi)}{\partial \varphi^k}(p) \partial_{\psi^k}|_{\phi(p)} \quad (37.12.21)$$

*Proof.* For by the basis theorem, we have:

$$d\phi_p(\partial_{\varphi^k}|_p) = \sum_{k \in \mathbb{Z}_n} d\phi_p(\partial_{\varphi^k}|_p)(\psi^k) \partial_{\psi^k}|_{\phi(p)} = \sum_{k \in \mathbb{Z}_n} \frac{\partial(\psi^k \circ \phi)}{\partial \varphi^k}(p) \partial_{\psi^k}|_{\phi(p)} \quad (37.12.22)$$

Completing the proof.  $\square$

The Jacobian matrix of  $\phi$  about the point  $p \in M$  with respect to the charts  $(U, \varphi) \in \mathcal{A}_M$  and  $(V, \psi) \in \mathcal{A}_n$  is the matrix:

$$I_{ij} = \frac{\partial(\psi^i \circ \phi)}{\partial \varphi^k}(p) \quad (37.12.23)$$

**Theorem 37.12.28.** If  $(M, \tau_M, \mathcal{A}_M)$ ,  $(N, \tau_N, \mathcal{A}_N)$ , and  $(P, \tau_P, \mathcal{A}_P)$  are smooth manifolds, if  $\phi : M \rightarrow N$  and  $\xi : N \rightarrow P$  are smooth functions, then:

$$d(\xi \circ \phi)_p = d\xi_{\phi(p)} \circ d\phi_p \quad (37.12.24)$$

*Proof.* For if  $p \in M$ ,  $v \in T_p M$ , and  $f \in C^\infty(P, \mathbb{R})$ , then:

$$d(\xi \circ \phi)_p(v)(f) = v(f \circ \xi \circ \phi) \quad (37.12.25)$$

$$= v((f \circ \xi) \circ \phi) \quad (37.12.26)$$

$$= d\phi_p(v)(f \circ \xi) \quad (37.12.27)$$

$$= d\xi_{\phi(p)}(d\phi_p(v)(f)) \quad (37.12.28)$$

$$= (d\xi_{\phi(p)} \circ d\phi_p)(v)(g) \quad (37.12.29)$$

Hence,  $d(\xi \circ \phi)_p = d\xi_{\phi(p)} \circ d\phi_p$ .  $\square$

**Theorem 37.12.29.** If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if  $\phi : M \rightarrow N$  is a smooth function, and if for every  $p \in M$  the differential pushforward  $d\phi_p$  is a isomorphism between  $T_p M$  and  $T_{\phi(p)} N$ , then  $\phi$  is a local diffeomorphism.

**Theorem 37.12.30.** If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, and if  $\phi : M \rightarrow N$  is a local diffeomorphism, then for all  $p \in M$ , the differential pushforward  $d\phi_p$  is an isomorphism between  $T_p M$  and  $T_{\phi(p)} N$ .

*Proof.* For if  $\phi$  is a local diffeomorphism, for all  $p \in M$  there is a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}_M$  and a chart  $(\mathcal{V}, \psi) \in \mathcal{A}_N$  such that  $p \in \mathcal{U}$  and  $\mathcal{U}$  and  $\mathcal{V}$  are diffeomorphic. Since  $d\phi_p$  is linear, it suffices to show that it is bijective. Suppose  $v, u \in T_p M$  are distinct but such that  $d\phi_p(v) = d\phi_p(u)$ . Then for all  $f \in C^\infty(N, \mathbb{R})$  we have:

$$d\phi_p(v - u)(f) = (v - u)(f \circ \phi) = v(f \circ \phi) - u(f \circ \phi) = 0 \quad (37.12.30)$$

But  $\phi$  is a local diffeomorphism, and thus for all  $f \in C^\infty(N, \mathbb{R})$  there exists a  $g \in C^\infty(M, \mathbb{R})$  such that  $f|_{\mathcal{V}} = g|_{\mathcal{U}} \circ \phi^{-1}$ . Since tangent vectors are local objects, if  $u(g) = v(g)$  for all  $g \in C^\infty(\mathcal{U}, \mathbb{R})$ , then  $u = v$ , a contradiction. Hence,  $d\phi_p$  is injective. By a reverse argument, it is surjective.  $\square$

### 37.12.4 Curves

**Definition 37.12.12** A curve in a smooth manifold is a smooth mapping  $\alpha : I \rightarrow M$ , where  $I$  is an open interval in  $\mathbb{R}$  with the subspace smooth structure.

**Definition 37.12.13** The velocity vector of a curve  $\alpha : I \rightarrow M$  at a point  $t \in I$  is the tangent vector:

$$\dot{\alpha}(t) = d\alpha\left(\frac{d}{du}\Big|_t\right) \quad (37.12.31)$$

The directional derivative of a function  $f \in C^\infty(M, \mathbb{R})$  along a curve  $\alpha : I \rightarrow M$  can be computed from the definition of the velocity vector. We have:

$$\dot{\alpha}(t)f = \frac{d}{du}(f \circ \alpha)|_{\alpha(t)} \quad (37.12.32)$$

We can apply the basis theorem as well, given a point  $t \in I$  and a chart  $(U, \varphi)$  that contains  $\alpha(t)$ , we have:

$$\dot{\alpha}(t) = \sum_{k \in \mathbb{Z}_n} \frac{d}{du}(\varphi^k \circ \alpha) \partial_{\varphi^k}|_{\alpha(t)} \quad (37.12.33)$$

This is reminiscent of the chain rule from multivariable calculus. If  $h : J \rightarrow I$  is a smooth function from the open interval  $J$  to the open interval  $I$ , and if  $\beta : J \rightarrow M$  is the composition of  $h$  and  $\alpha : I \rightarrow M$ , then:

$$\dot{\beta}(s) = \frac{dh}{du}(s)\dot{\alpha}(h(s)) \quad (37.12.34)$$

for all  $s \in J$ . Lastly, the differential pushforward preserves velocities. If  $\alpha : I \rightarrow M$  is a curve, and if  $\phi : M \rightarrow N$  is smooth, then:

$$d\phi_{\alpha(t)}(\dot{\alpha}(t)) = (\phi \circ \alpha)'(t) \quad (37.12.35)$$

This is because:

$$d\phi_{\alpha(t)}(\dot{\alpha}(t)) = d\phi_{\alpha(t)}\left(d\alpha_t\left(\frac{d}{du}\right)|_t\right) \quad (37.12.36)$$

$$= (d\phi_{\alpha(t)} \circ d\alpha_t)\left(\frac{d}{du}|_t\right) \quad (37.12.37)$$

$$= d(\phi \circ \alpha)_t\left(\frac{d}{du}|_t\right) \quad (37.12.38)$$

$$= (\phi \circ \alpha)'(t) \quad (37.12.39)$$

**Definition 37.12.14** A regular curve is a curve  $\alpha : I \rightarrow M$  such that for all  $t \in I$ ,  $\dot{\alpha}(t) \neq 0$ .

**Definition 37.12.15** A curve segment is a function  $\alpha : [a, b] \rightarrow M$  such that there exists an open interval  $I$  such that  $[a, b] \subseteq I$  and there exists a smooth extension  $\tilde{\alpha} : I \rightarrow M$ .

### 37.12.5 Vector Fields

**Definition 37.12.16** A vector field on a smooth manifold  $(M, \tau, \mathcal{A})$  is a function  $V : M \rightarrow \mathcal{F}(C^\infty(M, \mathbb{R}), \mathbb{R})$  such that for all  $p \in M$ ,  $V(p) \in T_p M$ .

That is, to every point in  $M$ ,  $V$  assigns a tangent vector to that point. There's a way to define a function  $Vf$  given a vector field  $V$  and a function  $f \in C^\infty(M, \mathbb{R})$ . We let:

$$(Vf)(p) = V_p(f) \quad (37.12.40)$$

This is now a function in  $\mathcal{F}(M, \mathbb{R})$ . To see this, note that for all  $p \in M$ ,  $V_p$  is a tangent vector, and thus  $V_p(f)$  is a real number since by definition a tangent vector outputs a real number. For  $Vf : M \rightarrow \mathbb{R}$  is a function. We can thus ask if it is continuous, differentiable, or smooth. This gives rise to the definition of a smooth vector field.

**Definition 37.12.17** A smooth vector field on a manifold  $(M, \tau, \mathcal{A})$  is a vector field  $V$  such that for all  $f \in C^\infty(M, \mathbb{R})$ , the function  $Vf : M \rightarrow \mathbb{R}$  defined by:

$$(Vf)(p) = V_p(f) \quad (37.12.41)$$

is a smooth function. That is,  $Vf \in C^\infty(M, \mathbb{R})$ .

We can multiply vector fields by elements of  $C^\infty(M, \mathbb{R})$ , as well as add two vector fields together in the following way:

$$(fV)(p) = f(p)V_p \quad (37.12.42)$$

$$(V + W)_p = V_p + W_p \quad (37.12.43)$$

### Notation 37.12.1: Smooth Vector Fields on a Manifold

The set of all smooth vector fields on a manifold  $(M, \tau, \mathcal{A})$  is denoted  $\mathfrak{X}(M)$ .

This notation, while standard and can be found in a plethora of textbooks on differential geometry and topology, is somewhat strange and the connection between it and vector fields is not obvious. An attempt will be made to remind the reader of this notation when context seems unclear. By the definition of multiplication by elements of  $C^\infty(M, \mathbb{R})$  and addition, we have that  $\mathfrak{X}(M)$  is a module over  $C^\infty(M, \mathbb{R})$ .

**Example 37.12.3** Given a smooth manifold  $(M, \tau, \mathcal{A})$  and a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}_M$ , there coordinate vector fields are the elements  $\partial_{\varphi^k} \in \mathfrak{X}(\mathcal{U})$  defined by:

$$\partial_{\varphi^k}(p) = \partial_{\varphi^k}|_p \quad (37.12.44)$$

By the basis theorem, for any vector field  $V$ , we have:

$$V = \sum_{k \in \mathbb{Z}_n} V \varphi^k \partial_{\varphi^k} \quad (37.12.45)$$

Where  $V\varphi^k$  is the function from  $\mathcal{U}$  into  $\mathbb{R}$  defined by:

$$(V\varphi^k)(p) = V_p(\varphi^k) \quad (37.12.46)$$

Where  $\varphi^k$  is the  $k^{th}$  coordinate function:

$$\varphi^k = \pi^k \circ \varphi \quad (37.12.47)$$

$\pi^k$  be the projection mapping from  $\mathbb{R}^n$  to  $\mathbb{R}$ .

We now look at some differentiable algebra to show that vector fields and derivations on an algebra  $C^\infty(M, \mathbb{R})$  over the field  $\mathbb{R}$  are actually one in the same.

**Definition 37.12.18** A derivation on an algebra  $\mathcal{A}$  over a ring  $R$  is a function  $D : \mathcal{A} \rightarrow \mathcal{A}$  such that for all  $a, b \in R$  and for all  $\mathbf{x}, \mathbf{y} \in \mathcal{A}$ , the following is true:

$$D(ax + by) = aD(\mathbf{x}) + bD(\mathbf{y}) \quad (37.12.48)$$

$$D(\mathbf{x} \times \mathbf{y}) = D(\mathbf{x}) \times \mathbf{y} + \mathbf{x} \times D(\mathbf{y}) \quad (37.12.49)$$

**Theorem 37.12.31.** *If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $D$  is a derivation on the algebra  $C^\infty(M, \mathbb{R})$  over  $\mathbb{R}$ , then there is a vector field  $V \in \mathfrak{X}(M)$  such that, for all  $f \in C^\infty(M, \mathbb{R})$ ,  $D(f) = Vf$ .*

*Proof.* For let  $V : M \rightarrow \mathcal{F}(M, \mathbb{R})$  be the function such that, for all  $f \in C^\infty(M, \mathbb{R})$ , we have

$$V(p)(f) = D(f)(p) \quad (37.12.50)$$

Then for all  $p$ ,  $V_p \in T_p M$ . For since  $D$  is a derivation, it is linear and Liebnizian, and hence  $V_p$  is a tangent vector to  $p$ .  $\square$

The converse is also true: Every vector field gives rise to a derivation.

**Theorem 37.12.32.** *If  $R$  is a ring and  $\mathcal{A}$  is an algebra over  $R$ , and if  $D_1, D_2$  are derivations on  $\mathcal{A}$ , then  $D_1 \circ D_2 - D_2 \circ D_1$  is a derivation on  $\mathcal{A}$ .*

*Proof.* For if  $a, b \in R$  and  $\mathbf{x}, \mathbf{y} \in \mathcal{A}$ , then:

$$(D_1 \circ D_2 - D_2 \circ D_1)(a\mathbf{x} + b\mathbf{y}) = D_1(D_2(a\mathbf{x} + b\mathbf{y})) - D_2(D_1(a\mathbf{x} + b\mathbf{y})) \quad (37.12.51)$$

$$= aD_1(D_2(\mathbf{x})) + bD_1(D_2(\mathbf{y})) \\ - aD_2(D_1(\mathbf{x})) - bD_2(D_1(\mathbf{y})) \quad (37.12.52)$$

$$= a(D_1(D_2(\mathbf{x})) - D_2(D_1(\mathbf{x}))) \\ + b(D_1(D_2(\mathbf{y})) - D_2(D_1(\mathbf{y}))) \quad (37.12.53)$$

$$= a(D_1 \circ D_2 - D_2 \circ D_1)(x) \\ + b(D_1 \circ D_2 - D_2 \circ D_1)(y) \quad (37.12.54)$$

And therefore  $D_1 \circ D_2 - D_2 \circ D_1$  is linear. It is also Liebnizian, since we have:

$$(D_1 \circ D_2 - D_2 \circ D_1)(\mathbf{x} \times \mathbf{y}) = D_1(D_2(\mathbf{x} \times \mathbf{y})) - D_2(D_1(\mathbf{x} \times \mathbf{y})) \quad (37.12.55)$$

$$= D_1(D_2(\mathbf{x}) \times \mathbf{y} + \mathbf{x} \times D_2(\mathbf{y})) \\ - D_2(D_1(\mathbf{x}) \times \mathbf{y} + \mathbf{x} \times D_1(\mathbf{y})) \quad (37.12.56)$$

Applying linearity and to this next sum, we can simplify:

$$D_1(D_2(\mathbf{x}) \times \mathbf{y} + \mathbf{x} \times D_2(\mathbf{y})) - D_1(D_2(\mathbf{x}) \times \mathbf{y} + \mathbf{x} \times D_2(\mathbf{y})) \\ = D_1(D_2(\mathbf{x}) \times \mathbf{y}) + D_1(\mathbf{x} \times D_2(\mathbf{y})) \quad (37.12.57)$$

Fill this in later, taking too long.  $\square$

**Definition 37.12.19** The Lie bracket on a smooth manifold  $(M, \tau, \mathcal{A})$  is the function  $[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$  defined by

$$[V, W] = VW - WV \quad (37.12.58)$$

That is,  $[V, W]$  is the function  $[V, W] : M \rightarrow \mathcal{F}(C^\infty(M, \mathbb{R}), \mathbb{R})$  such that, for all  $p \in M$  and for all  $f \in C^\infty(M, \mathbb{R})$  we have

$$[V, W]_p = V_p(Wf) - W_p(Vf) \quad (37.12.59)$$

By the previous theorem, the Lie bracket of two vector fields is again a vector fields since the bracket of two derivations is again a derivations, and the only derivations on a smooth manifold are precisely the vector fields.

**Theorem 37.12.33.** *The following are true of any Lie bracket:*

$$[aV + bW, X] = a[V, W] + b[W, X] \quad (37.12.60)$$

$$[V, W] = -[W, V] \quad (37.12.61)$$

$$[X, [Y, Z]] + [Y, [X, Z]] + [Z, [X, Y]] = \mathbf{0} \quad (37.12.62)$$

**Example 37.12.4** Consider the usual smooth manifold structure on  $\mathbb{R}^2$  with the vector fields  $V = y\partial_y$  and  $W = x\partial_y$ . Computing the Lie Bracket, we have:

$$[V, W](f) = y\partial_y(x\partial_y(f)) - x\partial_y(\partial_y(f)) \quad (37.12.63)$$

$$= xy\partial_y^2(f) - x\partial_y(f) - xy\partial_y^2(f) \quad (37.12.64)$$

$$= -x\partial_y(f) \quad (37.12.65)$$

$$= -Wf \quad (37.12.66)$$

The Lie bracket, while being linear in  $\mathbb{R}$ , is not linear over  $C^\infty(M, \mathbb{R})$ .

**Theorem 37.12.34.** *If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $[\cdot, \cdot]$  is the Lie bracket on  $M$ , if  $f, g \in C^\infty(M, \mathbb{R})$ , and if  $V, W \in \mathfrak{X}(M)$ , then:*

$$[fV, gW] = fg[V, W] + f(Vg)W - g(Wf)V \quad (37.12.67)$$

**Definition 37.12.20** A vector field  $V$  on a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  that is related to a vector field  $W$  on a smooth manifold  $(N, \tau_N, \mathcal{A}_N)$  is a vector field  $V \in \mathfrak{X}(M)$  such that there exists a smooth function  $\phi : M \rightarrow N$  such that:

$$d\phi(V_p) = W_{\phi(p)} \quad (37.12.68)$$

**Theorem 37.12.35.** *If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if  $V \in \mathfrak{X}(M)$  and  $W \in \mathfrak{X}(N)$ , then  $V$  is related to  $W$  if and only if there exists a smooth function  $\phi : M \rightarrow N$  such that for all  $f \in C^\infty(N, \mathbb{R})$  the following is true:*

$$V(f \circ \phi) = Wf \circ \phi \quad (37.12.69)$$

*Proof.* For:

$$\Leftrightarrow d\phi_p(V_p)(f) = W_{\phi(p)}(f) \quad (37.12.70)$$

$$\Leftrightarrow V_p(f \circ \phi) = W_{\phi(p)}(f) \quad (37.12.71)$$

$$\Leftrightarrow V(f \circ \phi)(p) = W(f)(\phi(p)) \quad (37.12.72)$$

$$\Leftrightarrow V(f \circ \phi) = Wf \quad (37.12.73)$$

□

**Theorem 37.12.36.** *If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if the function  $\phi : M \rightarrow N$  is smooth, if  $V_1, V_2 \in \mathfrak{X}(M)$ , if  $W_1, W_2 \in \mathfrak{X}(N)$ , if  $\phi$  relates  $V_1$  to  $W_1$ , and if  $\phi$  relates  $V_2$  to  $W_2$ , then  $\phi$  relates  $[V_1, V_2]$  to  $[W_1, W_2]$ .*

**Theorem 37.12.37.** *If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if the function  $\phi : M \rightarrow N$  is a diffeomorphism, and if  $V \in \mathfrak{X}(M)$ , then there is a unique vector field  $W \in \mathfrak{X}(N)$  such that  $\phi$  relates  $V$  to  $W$ .*

### 37.12.6 One Forms

Whenever we have the notion of linear functionals on vector spaces, we ultimately end up talking about the dual space. We have seen that tangent vectors can be considered as linear functions on  $C^\infty(M, \mathbb{R})$  with the additional property that they are Liebnizian. A vector field is a function on a manifold  $M$  that assigns to every point  $p \in M$  a tangent vector  $V_p$  at  $p$ . The dual notion of this is called a one form.

**Definition 37.12.21** The cotangent space on a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  at a point  $p \in M$  is the dual space of  $T_p M$ . That is,  $T_p^* M$  is the vector space of linear functionals on  $T_p M$ .

**Definition 37.12.22** A one form on a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  is a function  $\Theta : M \rightarrow \mathcal{F}(C^\infty(M, \mathbb{R})^*, \mathbb{R})$  such that for all  $p \in M$ ,  $\Theta_p \in T_p^* M$ .

Since dual spaces can be confusing, it is perhaps worthwhile to spell out the inputs and outputs of  $\Theta$ . Given a point  $p \in M$ ,  $\Theta_p$  is a cotangent vector living in the space  $T_p^* M$ . That is,  $\Theta_p$  takes in tangent vectors in  $T_p M$ , and outputs a real numbers. Moreover, it does this in a linear fashion. If  $p \in M$ , and if  $v, u \in T_p M$ , then:

$$\Theta_p(v + u) = \Theta_p(v) + \Theta_p(u) \quad (37.12.74)$$

Much the way  $\mathfrak{X}(M)$  was given a module structure, we can do this for one forms. Given a vector field  $V \in \mathfrak{X}(M)$  and a one form  $\Theta$ , we define  $\Theta V : M \rightarrow \mathbb{R}$  to be the function:

$$(\Theta v)(p) = \Theta_p(V_p) \quad (37.12.75)$$

That is,  $\Theta_p$  is the cotangent vector that  $\Theta$  assigns to  $p$ , and  $V_p$  is the tangent vector that  $V$  assigns to  $p$ . Since cotangent vectors take in as inputs tangent vectors and output real number,  $\Theta V$  is a well defined function from  $M$  to  $\mathbb{R}$ . Thus we can ask if it is continuous, differentiable, or smooth.

**Definition 37.12.23** A smooth one form on a manifold  $(M, \tau_M, \mathcal{A}_M)$  is a one form  $\Theta$  on  $M$  such that for every smooth vector field  $V \in \mathfrak{X}(M)$ ,  $\Theta V : M \rightarrow \mathbb{R}$  is a smooth function.

#### Notation 37.12.2: Smooth One Forms on a Manifold

The set of all smooth one forms on a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  is denoted  $\mathfrak{X}^*(M)$ .

Like  $\mathfrak{X}(M)$ ,  $\mathfrak{X}^*(M)$  is also a module over  $C^\infty(M, \mathbb{R})$ . To see this we need to define multiplication and addition, and we do this in the same way as with

vector fields:

$$(f\Theta)(p) = f(p)\Theta_p \quad (37.12.76)$$

$$(\Theta + \Xi)(p) = \Theta_p + \Xi_p \quad (37.12.77)$$

With this,  $\mathfrak{X}^*(M)$  is a module over  $C^\infty(M, \mathbb{R})$ . Given any function  $f \in C^\infty(M, \mathbb{R})$  on a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$ , we can always associate to this a smooth one form. This is called the *differential* of  $f$ .

**Definition 37.12.24** The differential of a smooth function  $f \in C^\infty(M, \mathbb{R})$  on a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  is the one form  $df \in \mathfrak{X}^*(M)$  defined by:

$$df(v) = v(f) \quad (37.12.78)$$

**Theorem 37.12.38.** If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $\Theta$  is a one form on  $M$ , and if  $(\mathcal{U}, \varphi) \in \mathcal{A}$  is a chart, then for all  $p \in \mathcal{U}$  the following is true:

$$\Theta_p = \sum_{k \in \mathbb{Z}_n} \Theta_p(\partial_{\varphi^l}|_p) d\varphi^k \quad (37.12.79)$$

**Theorem 37.12.39.** If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $(\mathcal{U}, \varphi) \in \mathcal{A}$  is a chart, and if  $f \in C^\infty(M, \mathbb{R})$ , then for all  $p \in \mathcal{U}$ :

$$df = \sum_{k \in \mathbb{Z}_n} \frac{\partial f}{\partial \varphi^k} d\varphi^k \quad (37.12.80)$$

**Theorem 37.12.40.** If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $a, b \in \mathbb{R}$ , and if  $f, g \in C^\infty(M, \mathbb{R})$ , then:

$$d(af + bg) = a df + b dg \quad (37.12.81)$$

**Theorem 37.12.41.** If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $f, g \in C^\infty(M, \mathbb{R})$ , then:

$$d(fg) = g d(f) + f dg \quad (37.12.82)$$

**Theorem 37.12.42.** If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $f \in C^\infty(M, \mathbb{R})$ , and if  $\alpha \in C^\infty(\mathbb{R}, \mathbb{R})$ , then:

$$d(\alpha \circ f) = \dot{\alpha}(f) df \quad (37.12.83)$$

### 37.12.7 Submanifolds

**Definition 37.12.25** A submanifold of a manifold  $(M, \tau, \mathcal{A})$  is a manifold  $(P, \tau_P, \mathcal{A}_P)$  such that  $(P, \tau_P)$  is a topological subspace of  $(M, \tau)$ , and such that the inclusion map  $\iota : P \rightarrow M$  is smooth and the differential  $d\iota$  is injective.

**Theorem 37.12.43.** *If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if  $\phi : M \rightarrow N$  is a smooth function, and if  $(P, \tau_P, \mathcal{A}_P)$  is a submanifold of  $M$ , then  $\phi|_P : P \rightarrow N$  is a smooth function.*

*Proof.* For  $\phi|_P = \phi \circ \iota$ , where  $\iota$  is the inclusion. Since  $(P, \tau_P, \mathcal{A}_P)$  is a submanifold,  $\iota$  is a smooth, and hence  $\phi|_P$  is the composition of smooth functions, which is smooth.  $\square$

The fact that  $d\iota$  is injective means that for every point  $p \in P$ ,  $T_p P$  is a vector subspace of  $T_p M$ .

**Example 37.12.5** Any open subset of a smooth manifold  $(M, \tau, \mathcal{A})$  is automatically a submanifold. Non-trivial examples include the sphere  $S^n$  which is a submanifold of  $\mathbb{R}^{n+1}$ . Hyperplanes in  $\mathbb{R}^n$  obtained by holding any  $k < n$  components constant are also submanifolds.

**Definition 37.12.26** A chart that is adapted to a subset  $P$  of a manifold  $(M, \tau, \mathcal{A})$  is a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}$  such that there exists an  $m < n$  such that:

$$P \cap \mathcal{U} = \{p \in \mathcal{U} \mid \varphi^j(p) = 0, j > n - m\} \quad (37.12.84)$$

That is,  $P$  looks roughly like a hyperplane. Since  $\varphi$  might be a curvilinear mapping, this *hyperplane* may not be a plane at all, but rather some  $m$  dimensional submanifold of  $\mathbb{R}^n$ . As it turns out, all submanifolds of a manifold  $(M, \tau, \mathcal{A})$  look like this, at least locally.

**Theorem 37.12.44.** *If  $(M, \tau, \mathcal{A})$  is a smooth manifold of dimension  $n \in \mathbb{N}$ , and if  $(P, \tau_P, \mathcal{A}_P)$  is a submanifold of dimension  $m \in \mathbb{N}$ , then there is a chart  $(\mathcal{U}, \varphi)$  that is adapted to  $P$  at each point of  $p$ .*

*Proof.* For let  $(\mathcal{U}, \varphi) \in \mathcal{A}$  be such that  $p \in \mathcal{U}$  and let  $(\mathcal{V}, \psi) \in \mathcal{A}_P$  be also such that  $p \in \mathcal{U}_P$ . But since the differential of the inclusion map is injective, the Jacobian matrix has rank  $m$  and thus we may rearrange it so that the first  $m$  rows are linearly independent. But then  $\varphi^k|_P$  for  $k \in \mathbb{Z}_m$  forms a chart for  $P$  on a neighborhood  $\mathcal{W}$  of  $p$ . I don't know, some exercise that is reference in O'Neill, fuck this. Come back later.  $\square$

**Theorem 37.12.45.** *If  $P$  is a submanifold of  $M$ , if  $\phi : N \rightarrow M$  is smooth, and if  $\phi[N] \subseteq M$ , then the induced map  $\phi : N \rightarrow P$  is smooth.*

**Theorem 37.12.46.** *If  $(M, \tau, \mathcal{A})$  is a smooth manifold, and  $P \subseteq M$ , then either there is no smooth structure on  $P$  that makes  $P$  a submanifold of  $M$ , or this is a unique smooth structure.*

*Proof.* For if there are two different smooth structures, by the previous theorem the identity mapping  $\text{id}_P : P \rightarrow P$  is a diffeomorphism from the first structure to the second, a contradiction.  $\square$

By this theorem it then makes sense to ask whether or not a subset  $P$  of a smooth manifold  $(M, \tau, \mathcal{A})$  is a submanifold or not, since there is either a unique smooth atlas on  $P$  that makes it a submanifold, or there is no such structure.

**Theorem 37.12.47.** *A subset  $P$  of a manifold  $(M, \tau, \mathcal{A})$  is a submanifold if and only if for all  $p \in P$  there exists a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}$  such that  $p \in \mathcal{U}$  and  $(\mathcal{U}, \varphi)$  is adapted to  $P$ .*

*Proof.* For if  $(\mathcal{U}, \varphi)$  is such a chart, then the restriction of  $\varphi$  to  $\mathcal{U} \cap P$  is a homeomorphism from an open subset of  $P$  (with the subspace topology) to a subset of  $\mathbb{R}^n$ . But by the definition of this chart,  $n - m$  of the coordinates will be constant, and hence  $\varphi|_P$  is a homeomorphism from an open subset of  $P$  to an open subset of  $\mathbb{R}^m$ . The collection of all such charts form an atlas for  $P$ . For the cover  $P$ , and if two charts  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  overlap then they overlap smoothly because of reasons. The inclusion map is also smooth because reasons, and  $d\iota$  is injective or whatever.  $\square$

**Definition 37.12.27** A tangent vector yield to a submanifold  $(P, \tau_P, \mathcal{A}_P)$  of a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  is a vector field  $V \in \mathfrak{X}(M)$  such that for all  $p \in P$  it is true that  $V_p \in T_p P$

**Theorem 37.12.48.** *If  $(M, \tau_M, \mathcal{A}_M)$  is a smooth manifold, if  $(P, \tau_P, \mathcal{A}_P)$  is a submanifold of  $M$ , and if  $V \in \mathfrak{X}(M)$  is a tangent vector to  $P$ , then  $V|_P \in \mathfrak{X}(P)$ .*

**Theorem 37.12.49.** *If  $(M, \tau_M, \mathcal{A}_M)$  is a smooth manifold, if  $(P, \tau_P, \mathcal{A}_P)$  is a submanifold of  $M$ , if  $V, W \in \mathfrak{X}(M)$  are tangent vector fields to  $P$ , if  ${}^M[\cdot, \cdot]$  is the Lie bracket on  $M$ , and if  ${}^P[\cdot, \cdot]$  is the Lie bracket on  $P$ , then:*

$${}^M[V, W]|_P = {}^P[V|_P, W|_P] \quad (37.12.85)$$

### 37.12.8 Immersions and Submersions

**Theorem 37.12.50.** *If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if the function  $\phi : M \rightarrow N$  is smooth, then for all  $p \in M$  the function  $d\phi_p : T_p M \rightarrow T_{\phi(p)} N$  is injective if and only if the Jacobian matrix has rank  $m$  for every chart  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  containing  $p$  and  $\phi(p)$ , respectively.*

**Definition 37.12.28** An immersion from a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  to another smooth manifold  $(N, \tau_N, \mathcal{A}_N)$  is a smooth function  $\phi : M \rightarrow N$  such that for all  $p \in M$  the differential pushforward  $d\phi_p$  is injective.

**Theorem 37.12.51.** *If  $\alpha : I \rightarrow M$  is a regular curve ( $\dot{\alpha}(t) \neq 0$ ), then  $\alpha$  is an immersion.*

**Definition 37.12.29** An embedding is an injective immersion  $\phi : M \rightarrow N$  such that  $\phi$  is a homeomorphism from  $M$  onto its image.

**Theorem 37.12.52.** *If  $(M, \tau, \mathcal{A})$  is a smooth manifold and if  $(P, \tau_P, \mathcal{A}_P)$  is a submanifold of  $M$ , then the inclusion map  $\iota : P \rightarrow M$  is an embedding.*

**Theorem 37.12.53.** *If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if  $\phi : M \rightarrow N$  is an embedding, then  $\phi[M]$  is a submanifold of  $N$ .*

**Definition 37.12.30** An immersed submanifold of a smooth manifold  $(M, \tau, \mathcal{A})$  is a subset  $P$  such that the inclusion map  $\iota : P \rightarrow M$  is an immersion.

By the previous theorem, every submanifold is an immersed submanifold but the converse is not true. The standard picture of the Klein bottle represents an immersed submanifold that is not a submanifold since its image in  $\mathbb{R}^3$  is not homeomorphic to the original Klein bottle.

**Theorem 37.12.54.** *If  $(M, \tau_M, \mathcal{A}_M)$ ,  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds,  $\phi : M \rightarrow N$  a smooth function, and if for all  $p \in M$  it is true that the differential pushforward  $d\phi_p$  is surjective, then for every chart  $(U, \varphi) \in \mathcal{A}_M$  and  $(V, \psi) \in \mathcal{A}_N$ , the Jacobian matrix has rank  $n$ .*

**Definition 37.12.31** A regular value of a smooth function  $\phi : M \rightarrow N$  from a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  to a manifold  $(N, \tau_N, \mathcal{A}_N)$  is a point  $q \in N$  such that for all  $p \in \phi^{-1}\{q\}$  the differential pushforward  $d\phi_p$  is surjective.

**Theorem 37.12.55.** *If  $(M, \tau_M, \mathcal{A}_M)$ ,  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds,  $\phi : M \rightarrow N$  a smooth function, and if  $q \in N$  is a regular point of  $\phi$ , then the dimension of  $M$  is equal to the dimension of  $N$  plus the dimension of  $\phi^{-1}\{q\}$ .*

**Definition 37.12.32** The codimension of a submanifold  $P$  of a smooth manifold  $(M, \tau, \mathcal{A})$  is the dimension of  $M$  minus the dimension of  $P$ .

**Definition 37.12.33** A hypersurface in a smooth manifold  $(M, \tau, \mathcal{A})$  is a submanifold of codimension 1.

**Theorem 37.12.56.** *If  $(M, \tau, \mathcal{A})$  is a smooth manifold, if  $f \in C^\infty(M, \mathbb{R})$ , and if  $c \in \mathbb{R}$  is a regular point of  $f$ , then  $f^{-1}\{c\}$  is a hypersurface in  $M$ .*

**Example 37.12.6** This is one way of showing that the sphere  $S^n$  is a manifold since it can be realized as a hypersurface in  $\mathbb{R}^{n+1}$ . Let  $f(\mathbf{x}) = \|\mathbf{x}\|_2^2$  for all  $\mathbf{x} \in \mathbb{R}^{n+1}$ . Then  $1 \in \mathbb{R}$  is a regular point, and hence  $f^{-1}\{1\}$  is a hypersurface (and hence a submanifold, and therefore a manifold) but this is simply  $S^n$ .

**Definition 37.12.34** A submersion from a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  to another smooth manifold  $(N, \tau_N, \mathcal{A}_N)$  is a function  $\phi : M \rightarrow N$  such that for all  $p \in M$  the differential pushforward  $d\phi_p$  is surjective.

Since  $\phi^{-1}\{q\}$  is a submanifold for every  $q \in N$  this then forms a partition of  $M$ . Thus, submersions can be seen as ways of partitioning a manifold  $M$  by the level-sets or fibers of another manifold.

### 37.12.9 Partitions of Unity

If  $f_\alpha \in C^\infty(M, \mathbb{R})$  is a collection of functions such that  $\text{supp}\{f_\alpha\}$  forms a locally finite collection, then  $\sum_\alpha f_\alpha$  is well defined since at every point there are only finitely many functions  $f_\alpha$  that contribute to the sum, and hence this sum converges everywhere.

**Definition 37.12.35** A smooth partition of unity is of a smooth manifold  $(M, \tau, \mathcal{A})$  is a subset  $\mathcal{F} \subseteq C^\infty(M, \mathbb{R})$  such that for all  $f \in \mathcal{F}$ ,  $f[M] \subseteq [0, 1]$ , and such that the set:

$$\mathcal{O} = \{\text{supp } f \mid f \in \mathcal{F}\} \quad (37.12.86)$$

is locally finite, and  $\sum_\alpha f_\alpha = 1$ .

**Definition 37.12.36** A subordinate partition of unity on a manifold  $(M, \tau, \mathcal{A})$  with respect to an open cover  $\mathcal{O}$  of  $M$  is a partition of unity  $\mathcal{F}$  such that the supports of  $\mathcal{F}$  form a refinement of  $\mathcal{O}$ .

#### Theorem 37.12.57: Existence of Partitions of Unity on Manifolds

If  $(M, \tau, \mathcal{A})$  is a smooth manifold, and if  $\mathcal{O}$  is an open cover of  $M$ , then there exists a partition of unity  $\mathcal{F}$  that is subordinate to  $\mathcal{O}$ .

The second countability of smooth manifolds is required for this theorem to hold since the proof relies on paracompactness. Hence, large manifolds like the long line may fail to have such partitions. This theorem is actually a significantly weaker version of a purely topological result:

#### Theorem 37.12.58: Partitions of Unity Theorem

If  $(X, \tau)$  is an accessible topological space ( $T_1$ ), then it is paracompact and Hausdorff if and only if every open cover has a subordinate partition of unity.

Since, by definition, manifolds are Hausdorff, and since the second countability criterion can prove paracompactness, the existence of subordinate partitions of unity falls out immediately. Furthermore, if somehow we have a connected locally Euclidean space that always has partitions of unity, then we know it is a manifold. That is, since locally Euclidean spaces are always  $T_1$ , this theorem then proves the space is Hausdorff and paracompact. Connectedness then gives us that the space is second countable, and hence a manifold. We can trivially weaken this to having countably many connected components.

### 37.12.10 Orientability

**Definition 37.12.37** A smooth oriented manifold is a manifold  $(M, \tau, \mathcal{A})$  such there exists an atlas  $\mathcal{O} \subseteq \mathcal{A}$  such that for all  $(\mathcal{U}, \varphi), (\mathcal{V}, \psi) \in \mathcal{O}$ , the determinant of the Jacobian matrix is positive.

**Example 37.12.7** The torii, spheres, and Euclidean spaces  $\mathbb{T}^n$ ,  $\mathbb{S}^n$ , and  $\mathbb{R}^n$ , respectively, are all orientable. The Möbius strip and Klein bottle are not.

**Theorem 37.12.59.** *If  $A$  is a set,  $\mathcal{A}$  is a topological atlas on  $A$  such that all charts in  $\mathcal{A}$  overlap smoothly, and if for all  $x, y \in A$  it is true that either there exists a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}$  such that  $x, y \in \mathcal{U}$  or there exist charts  $(\mathcal{U}, \varphi), (\mathcal{V}, \psi) \in \mathcal{A}$  such that  $x \in \mathcal{U}$ ,  $y \in \mathcal{V}$ , and  $\mathcal{U} \cap \mathcal{V}$  are disjoint, then there is a unique topology  $\tau$  on  $A$  such that  $(A, \tau)$  is a Hausdorff topological space such that  $\mathcal{A}$  is a smooth atlas on  $A$ .*

We can see what each condition in the statement of this theorem is doing. the first criterion, then  $\mathcal{A}$  be an atlas, simply requires that the domains of the charts in  $\mathcal{A}$  cover the entire set  $A$ . The smoothly overlapping part is so that we can ensure that  $\mathcal{A}$  can be a smooth atlas, and the last criterion gives us the Hausdorff condition. If  $x, y$  lie in the same chart, since Euclidean space is Hausdorff we can thus separate  $x$  and  $y$ . If they lie in distinct disjoint charts, then again they can be separated. Note the the disjoint condition is crucial. The bug-eyed line forms a countable example. We can cover the bug-eyed line with charts, but the two origins cannot be separated by disjoint charts, and hence is a non-Hausdorff manifold.

### 37.12.11 Special Manifolds

The projection mappings  $\pi : M \times N \rightarrow M$  and  $\sigma : M \times N \rightarrow N$  are smooth mappings, and are submersions. A map  $\phi : P \rightarrow M \times N$  is smooth if and only if  $\pi \circ \phi$  and  $\sigma \circ \phi$  are smooth. The subspace  $M \times q$  and  $p \times N$  are submanifolds of  $M \times N$ .

**Theorem 37.12.60.** *If  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  are smooth manifolds, if  $(M \times N, \tau_{M \times N}, \mathcal{A}_{M \times N})$  is the product manifold, if  $(p, q) \in M \times N$ , then:*

$$T_{(p,q)}(M \times N) = T_{(p,q)}(M \times \{q\}) \oplus T_{(p,q)}(\{p\} \times N) \quad (37.12.87)$$

Where  $\oplus$  is the vector space direct sum.

**Definition 37.12.38** The lift of a smooth function  $f \in C^\infty(M, \mathbb{R})$  on a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  with respect to the product manifold of  $M$  with  $(N, \tau_N, \mathcal{A}_N)$  is the function  $\tilde{f} \in C^\infty(M \times N, \mathbb{R})$  defined by  $\tilde{f} = f \circ \pi_M$ .

By the previous theorem, for any point  $p \in M$  and for any tangent vector  $v \in T_p M$  there is a unique tangent vector  $\tilde{v} \in T_{(p,q)}(M \times N)$  such that  $d\pi_M(\tilde{v}) = v$

and we can similarly define this as the lift of  $v$ . Extending this further, given a smooth vector field  $V \in \mathfrak{X}(M)$  we can lift this to  $M \times N$  by defining  $\tilde{V}_{(p,q)}$  to be the unique lift of  $V_p$ . Lifts of vector fields are not invariant under multiplication by elements of  $C^\infty((M \times N), \mathbb{R})$ . For example, in  $\mathbb{R}^2$ , we have that  $d/dx$  can be lifted to  $\partial_x$ , but  $y\partial_x$  is not a lift. A horizontal lift is one such that  $\pi_N(V)$  is simply the zero vector field.

**Theorem 37.12.61.** *If  $V, W \in \mathfrak{X}(M)$ , then  $[\tilde{V}, \tilde{W}] = [\tilde{V}, \tilde{W}]$ .*

### 37.12.12 Vector Spaces as Manifolds

If  $V$  is an  $n$  dimensional vector space over  $\mathbb{R}$  and if  $\varphi, \psi$  are linear isomorphisms from  $V$  to  $\mathbb{R}^n$ , then  $\varphi \circ \psi^{-1}$  is a linear isomorphism from  $\mathbb{R}^n$  to itself, which is therefore a diffeomorphism. By a previous theorem there is a unique Hausdorff topology on  $V$  and a unique maximal atlas that makes  $V$  into an  $n$  dimensional smooth manifold such that all of the linear isomorphisms form charts.

**Theorem 37.12.62.** *If  $V$  is an  $n$  dimensional vector space, if  $p, v \in V$ , if  $\alpha : I \rightarrow V$  is the curve  $\alpha(t) = p + tv$ , is  $v_p = \dot{\alpha}(0)$ , and if  $(U, \varphi)$  is a linear chart containing  $p$ , then:*

$$v_p = \sum_{k \in \mathbb{Z}_n} \varphi^k(v) \partial_{\varphi^k}|_p \quad (37.12.88)$$

*Proof.* For by linearity:

$$\varphi^k(\alpha(t)) = \varphi^k(p) + t\varphi^k(v) \quad (37.12.89)$$

But then:

$$v_p = \dot{\alpha}(0) = \sum_{k \in \mathbb{Z}_n} \frac{d}{dt} (\varphi^k \circ \alpha) \partial_{\varphi^k}|_p = \sum_{k \in \mathbb{Z}_n} \varphi^k(v) \partial_{\varphi^k}|_p \quad (37.12.90)$$

□

The function  $v \mapsto v_p$  is a linear isomorphism from  $V$  to  $T_p V$ , similarly for  $v_q \mapsto v_p$ .

### 37.12.13 The Tangent Bundle

**Definition 37.12.39** The tangent bundle of a smooth manifold  $(M, \tau, \mathcal{A})$  is the set:

$$TM = \coprod_{p \in M} T_p M = \bigcup_{p \in M} \{p\} \times T_p M \quad (37.12.91)$$

We can topologize  $TM$  in a way that makes this a manifold of dimension  $2n$  and this is **not** the disjoint union topology. Indeed, for any uncountable manifold

$M$ , the disjoint union topology will not be second countable, but instead will be a locally Euclidean Hausdorff topological space of the same dimension as  $M$ . That is, a *large*  $n$  dimensional manifold. To topologize  $TM$ , we consider the projection mapping  $\pi$  sending  $(p, v) \in TM$  to  $p \in M$ . Given a chart  $(\mathcal{U}, \varphi)$  in  $M$ , we look at  $\pi^{-1}(\mathcal{U})$ , which is a subset of  $TM$ . Since the  $\partial_{\varphi^k}|_p$  uniquely determine any tangent vector at any point  $p \in \mathcal{U}$ , we can define a new chart  $(\pi^{-1}(\mathcal{U}), \tilde{\varphi})$  defined by:

$$\tilde{\varphi}(p, v) = (\varphi(p), \dot{\varphi}(v)) \quad (37.12.92)$$

Where  $\dot{\varphi}(p)$  is the function such that the  $k^{th}$  component is  $\dot{\varphi}(v)_k = v(\varphi^k)$ . We do this for every chart, showing that we now have a  $2n$  dimensional atlas that gives rise to a Hausdorff topology. To complete the claim that  $TM$  is a smooth manifold we need to show that this atlas has smoothly overlapping charts. Suppose  $(\pi^{-1}(\mathcal{U}, \tilde{\varphi}))$  and  $(\pi^{-1}(\mathcal{V}, \tilde{\psi}))$  are two such charts with non-empty overlap. Then of  $k \in \mathbb{Z}_n$ , we have:

$$\pi^k \circ (\tilde{\varphi} \circ \tilde{\psi}^{-1}) = (\pi^k \circ \tilde{\varphi}) \circ \tilde{\psi}^{-1} = (\varphi^k \circ \pi) \circ \tilde{\psi}^{-1} = \varphi^k \circ \psi^{-1} \quad (37.12.93)$$

Which is smooth. Similarly, for  $k \in \mathbb{Z}_n$ , we have:

$$\pi^{n+k} \circ (\tilde{\varphi} \circ \tilde{\psi}^{-1}) = (\pi^{n+k} \circ \tilde{\varphi}) \circ \tilde{\psi}^{-1} \quad (37.12.94)$$

And that's nice.

**Definition 37.12.40** A smooth section in the tangent bundle  $TM$  of a smooth manifold  $(M, \tau, \mathcal{A})$  is a smooth function  $f : M \rightarrow TM$  such that  $\pi \circ f = \text{id}_M$ .

Every vector field is thus a smooth section in the tangent bundle. This gives us the generalization to vector fields on a smooth map. A vector field on a smooth map  $\phi : M \rightarrow N$  is a smooth function  $V : M \rightarrow TN$  such that  $\pi \circ V = \phi$ . The velocity of a curve is thus an example of a smooth vector field on the function  $\alpha : I \rightarrow M$ .

### 37.12.14 Integral Curves

**Definition 37.12.41** An integral curve of a vector field  $V \in \mathfrak{X}(M)$  on a smooth manifold  $(M, \tau, \mathcal{A})$  is a curve  $\alpha : I \rightarrow M$  such that for all  $t \in I$ ,  $\dot{\alpha}(t) = V_{\alpha(t)}$ .

## 37.13 O'Neill Problems Chapter 7

**Theorem 37.13.1.** *Every Lie group is parallelizable.*

*Proof.* Parallel transport of non-zero tangent vector at tangent space of unital element by left multiplication (differential).  $\square$

Non-orientable space is not simply connected since it is not its own universal cover since double cover is connected.

## 37.14 Manifolds Review

Since  $\varphi$  is an injective open mapping, it is automatically a homeomorphism onto its image. Indeed, it is common to define locally Euclidean in terms of homeomorphisms. The dimension can vary in this definition and we can consider the disjoint union of a sphere and a line. However, dimension is locally constant. This requires Brouwer's invariance of domain.

### Theorem 37.14.1: Invariance of Domain

If  $\mathcal{U} \subseteq \mathbb{R}^n$  is open, and if  $f : \mathcal{U} \rightarrow \mathbb{R}^n$  is a continuous injective function, then  $f(\mathcal{U})$  is open.

*| Proof.* Difficult. □

We can now prove that dimension is locally constant.

### Theorem 37.14.2: Invariance of Dimension

If  $\mathcal{U} \subseteq \mathbb{R}^n$  is open and  $\mathcal{V} \subseteq \mathbb{R}^m$  is open, and if  $f : \mathcal{U} \rightarrow \mathcal{V}$  is a homeomorphism, then  $n = m$ . ■

*Proof.* For suppose not and suppose  $n > m$ . Let  $\tilde{\mathcal{V}}$  be the extension of  $\mathcal{V}$  into  $\mathbb{R}^n$  where the last  $n - m$  coordinates are simply 0. But  $f$  is a homeomorphism, and thus the extension  $\tilde{f} : \mathcal{U} \rightarrow \mathbb{R}^n$  is an injective continuous function, and hence an open mapping. But  $f(\mathcal{U})$  is not open since for any point and for any  $\varepsilon > 0$  the  $\varepsilon$  ball must leak into the last  $n - m$  coordinates, a contradiction. Thus  $n = m$ . □

**Theorem 37.14.3.** *If  $(X, \tau)$  is a locally Euclidean topological space, if  $x \in X$ , and if  $n, m \in \mathbb{N}$  are such that there exists open sets  $\mathcal{U}_n$  and  $\mathcal{V}_m$  such that there are injective open mapping  $\varphi_n : \mathcal{U}_n \rightarrow \mathbb{R}^n$  and  $\varphi_m : \mathcal{V}_m \rightarrow \mathbb{R}^m$ , then  $m = n$ .*

*Proof.* For  $\mathcal{U}_n \cap \mathcal{V}_m$  is non-empty since  $x$  is in the intersection, and the intersection of open is open. Thus the resection  $\varphi_n|_{\mathcal{U}_n \cap \mathcal{V}_m}$  and  $\varphi_m|_{\mathcal{U}_n \cap \mathcal{V}_m}$  are open mappings into  $\mathbb{R}^n$  and  $\mathbb{R}^m$ , respectively. But then these are homeomorphisms onto their images. But by composing  $\varphi_n \circ \varphi_m^{-1}$  we obtain a homeomorphism from an open subset of  $\mathbb{R}^m$  to  $\mathbb{R}^n$ . But then  $n = m$ , by the previous theorem. □

We have thus shown that there is a well defined dimension function  $\dim : X \rightarrow \mathbb{N}$  that assigns to every  $x \in X$  the unique dimension of  $x$ . That is, the space  $\mathbb{R}^n$  that  $x$  locally looks like.

**Definition 37.14.1** A locally constant function on from a topological space  $(X, \tau)$  to a set  $Y$  is a function  $f : X \rightarrow Y$  such that for all  $x \in X$  there exists an open subset  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  and  $f|_{\mathcal{U}}$  is a constant mapping.

Note that there need not be any topology on  $Y$ .

**Theorem 37.14.4.** *If  $(X, \tau)$  is a connected topological space, if  $Y$  is a set with at least two distinct points, and if  $f : X \rightarrow Y$  is a locally constant function, then it is a constant function.*

*Proof.* For suppose not. Then there are  $c_1, c_2 \in Y$  such that  $f^{-1}[\{c_1\}]$  and  $f^{-1}[\{c_2\}]$  are non-empty. Let  $x \in f^{-1}[\{c_1\}]$ . Since  $f$  is a locally constant function, there exists an open subset  $\mathcal{U}_x \subseteq X$  such that  $x \in \mathcal{U}_x$  and  $f|_{\mathcal{U}_x}$  is a constant mapping. But this is true of all  $x \in f^{-1}[\{c_1\}]$ , and hence:

$$f^{-1}[\{c_1\}] = \bigcup_{x \in f^{-1}[\{c_1\}]} \mathcal{U}_x \quad (37.14.1)$$

which is the union of open, and hence open. Similarly, the complement can be written as:

$$X \setminus f^{-1}[\{c_1\}] = \bigcup_{x \notin f^{-1}[\{c_1\}]} \mathcal{U}_x \quad (37.14.2)$$

which is the union of open, and hence open. Moreover it is non-empty since  $f^{-1}[\{c_2\}]$  is a subset, and this set is non-empty. But then  $X$  is the union of two disjoint non-empty open subsets and is hence disconnected, a contradiction. Thus,  $f$  is constant.  $\square$

**Theorem 37.14.5.** *If  $(X, \tau)$  is a locally Euclidean topological space, then the dimension function  $\dim : X \rightarrow \mathbb{N}$  is locally constant.*

*Proof.* For let  $x \in X$ . Then there is an open subset  $\mathcal{U} \subseteq X$  such that  $x \in \mathcal{U}$  and there exists a continuous injective open mapping  $f : \mathcal{U} \rightarrow \mathbb{R}^n$ . But then for all  $y \in \mathcal{U}$ ,  $\mathcal{U}$  is an open set such that  $y \in \mathcal{U}$  and  $f : \mathcal{U} \rightarrow \mathbb{R}^n$  is a continuous injective open mapping, and hence  $\dim(y) = \dim(x)$ . Thus,  $\dim$  is locally constant.  $\square$

**Theorem 37.14.6.** *If  $(X, \tau)$  is a connected locally Euclidean topological space, then there is a unique dimension.*

*Proof.* For  $\dim$  is locally constant, and a locally constant function on a connected space is constant.  $\square$

**Definition 37.14.2** Compatible charts on a topological space  $(X, \tau)$  are charts  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  such that either  $\mathcal{U} \cap \mathcal{V}$  are empty, or  $\varphi \circ \psi^{-1} : \psi(\mathcal{V}) \rightarrow \mathbb{R}^n$  and  $\psi \circ \varphi^{-1} : \varphi(\mathcal{U}) \rightarrow \mathbb{R}^n$  are smooth.

**Definition 37.14.3** Compatible atlases on a topological manifold  $(X, \tau)$  are atlases  $\mathcal{A}$  and  $\mathcal{A}'$  such that for every chart in  $\mathcal{A}$  is compatible with every chart in  $\mathcal{A}'$ , and vice-versa.

**Theorem 37.14.7.** If  $(X, \tau)$  is a topological manifold and if  $\mathcal{A}$  and  $\mathcal{A}'$  are smooth atlases, then they are compatible if and only if  $\mathcal{A} \cup \mathcal{A}'$  is a smooth atlas.

*Proof.* For if  $\mathcal{A}$  and  $\mathcal{A}'$  are compatible, then every element of  $\mathcal{A} \cup \mathcal{A}'$  is compatible with every other element, and moreover the domains cover  $X$ . Hence, the union is a smooth atlas. If the union is a smooth atlas, then every element of  $\mathcal{A}$  is compatible with every element of  $\mathcal{A} \cup \mathcal{A}'$ , and in particular it is compatible with every element of  $\mathcal{A}'$ . Similarly, every element of  $\mathcal{A}'$  is compatible with  $\mathcal{A}$ , and hence they are compatible atlases.  $\square$

### Theorem 37.14.8: Existence and Uniqueness of Maximal Smooth Atlases

---

If  $(X, \tau)$  is a topological manifold and if  $\mathcal{A}$  is a smooth atlas, then there is a maximal smooth atlas  $\mathcal{C}$  such that  $\mathcal{A} \subseteq \mathcal{C}$ .  $\blacksquare$

*Proof.* For consider the set of all atlases on  $(X, \tau)$ . This forms a partially ordered set by inclusion. Given a chain of atlases, the union over the entire chain is again an atlas by the previous theorem. But then every chain is bounded, and thus by Zorn's lemma for every  $\mathcal{A}$  there is a maximal element  $\mathcal{C}$  that bounds  $\mathcal{A}$ . If  $\mathcal{C}'$  is another, then they are compatible since they are both compatible with  $\mathcal{A}$ , and thus  $\mathcal{C} \cup \mathcal{C}'$  is a smooth atlas, contradicting maximality. Hence,  $\mathcal{C}$  is unique.  $\square$

**Definition 37.14.4** A smooth function from a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  to a smooth manifold  $(N, \tau_N, \mathcal{A}_N)$  is a function  $\phi : M \rightarrow N$  such that for all  $p \in M$  there is a chart  $(\mathcal{U}, \varphi) \in \mathcal{A}_M$  and a chart  $(\mathcal{V}, \psi) \in \mathcal{A}_N$  such that  $p \in \mathcal{U}$ ,  $\phi(p) \in \mathcal{V}$ , and the mapping  $\psi \circ \phi \circ \varphi^{-1} : \varphi(\mathcal{U} \cap \phi^{-1}(\mathcal{V})) \rightarrow \mathbb{R}^n$  is a smooth function.

Note that since charts in a smooth atlas overlap smoothly, this definition does not actually depend on the choice of charts.

**Definition 37.14.5** A diffeomorphism from a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  to a smooth manifold  $(N, \tau_N, \mathcal{A}_N)$  is a homeomorphism  $\phi : M \rightarrow N$  such that  $\phi$  and  $\phi^{-1}$  are smooth.

In particular, we can consider the set of all smooth functions from the manifold  $M$  into  $\mathbb{R}$ , where  $\mathbb{R}$  has its usual smooth structure. This structure  $C^\infty(M, \mathbb{R})$  forms a commutative ring.

**Definition 37.14.6** A tangent vector at a point  $p$  in a smooth manifold  $(M, \tau, \mathcal{A})$  is a function  $v : C^\infty(M, \mathbb{R}) \rightarrow \mathbb{R}$  that is linear and Liebnizian. That is:

$$\begin{aligned} v(af + bg) &= av(f) + bv(g) && \text{(Linearity)} \\ v(fg) &= v(f)g(p) + f(p)v(g) && \text{(Liebnizian)} \end{aligned}$$

**Definition 37.14.7** The tangent space at a point  $p$  in a smooth manifold  $(M, \tau, \mathcal{A})$  is the set  $T_p M$  of all tangent vectors at  $p$ .

This may seem abstract, but it is a direct generalization of the directional derivative one studies in vector calculus.

**Theorem 37.14.9.** If  $\phi : M \rightarrow N$  is a smooth function, if  $p \in M$ , and if  $v$  is a tangent vector at  $p$ , then the function  $d_p \phi : C^\infty(N, \mathbb{R}) \rightarrow \mathbb{R}$  defined by  $d_p \phi(v)(f) = v(f \circ \phi)$  is an element of  $T_{\phi(p)} N$ .

*Proof.* For:

$$v((af + bg) \circ \phi) = v(a(f \circ \phi) + b(g \circ \phi)) = av(f \circ \phi) + bv(g \circ \phi) \quad (37.14.3)$$

and:

$$v((fg) \circ \phi) = v((f \circ \phi)(g \circ \phi)) = v(f \circ \phi)(g \circ \phi)(p) + (f \circ \phi)(p)v(g \circ \phi) \quad (37.14.4)$$

□

**Definition 37.14.8** The differential pushforward of a smooth function  $\phi : M \rightarrow N$  from a smooth manifold  $(M, \tau_M, \mathcal{A}_M)$  to a smooth manifold  $(N, \tau_N, \mathcal{A}_M)$  at a point  $p \in M$  is the function  $d_p \phi : T_p M \rightarrow T_{\phi(p)} N$  defined by  $d_p \phi(v)(f) = v(f \circ \phi)$  for all  $f \in C^\infty(N, \mathbb{R})$ .

**Definition 37.14.9** An open curve in a topological space  $(X, \tau)$  is a continuous function  $\alpha : I \rightarrow X$ , where  $I$  is the open unit interval.

**Definition 37.14.10** A smooth curve in a smooth manifold  $(M, \tau, \mathcal{A})$  is a curve  $\alpha : I \rightarrow M$  such that  $\alpha$  is a smooth function with respect to the standard smooth structure on  $I$ .

**Definition 37.14.11** The velocity of a smooth curve  $\alpha : I \rightarrow M$  at a point  $t \in I$  differential pushforward evaluated at the tangent vector  $d/dx|_{x=t}$ . That is:

$$\dot{\alpha}(t) = d_t \alpha \left( \frac{d}{dx} \Big|_{x=t} \right) \quad (37.14.5)$$

**Definition 37.14.12** The tangent bundle of a smooth manifold  $(M, \tau, \mathcal{A})$  is the set  $TM = \coprod_{p \in M} T_p M$ . That is, the disjoint union of all tangent spaces.

There is a topology that one can place on the tangent bundle that makes it a  $2n$  dimensional manifold (with a natural smooth atlas) and this is not the disjoint union topology. Indeed for any manifold of non-zero dimension the disjoint union topology is not second countable, and hence not a manifold (but it will be an  $n$  dimensional locally Euclidean Hausdorff space).

**Definition 37.14.13** A vector field is a smooth function  $V : M \rightarrow TM$  such that for all  $p \in M$ ,  $V(p) \in T_p M$ .

The collection  $\mathfrak{X}(M)$  of all smooth vector fields on a smooth manifold  $M$  has a module structure over  $C^\infty(M, \mathbb{R})$  if we define addition and scalar multiplication as follows:

$$(fV)_p = f(p)V_p \quad (37.14.6)$$

$$(V + W)_p = V_p + W_p \quad (37.14.7)$$

We can also evaluate vector fields at functions in  $C^\infty(M, \mathbb{R})$  as follows:

$$(Vf)(p) = V_p(f) \quad (37.14.8)$$

This is well defined since for all  $p \in M$ ,  $V_p$  is a tangent vector in  $T_p M$ , which is a linear functional. Thus  $Vf$  is a function from  $C^\infty(M, \mathbb{R})$  to itself. The structure of  $C^\infty(M, \mathbb{R})$  can also be seen as an algebra over the field  $\mathbb{R}$ . As such we can define derivations.

**Definition 37.14.14** A derivation on a  $C^\infty(M, \mathbb{R})$  is a function  $D : C^\infty(M, \mathbb{R}) \rightarrow C^\infty(M, \mathbb{R})$  that is  $\mathbb{R}$  linear and Liebnizian:

$$D(af + bg) = aD(f) + bD(g) \quad (37.14.9)$$

$$D(fg) = D(f)g + fD(g) \quad (37.14.10)$$

This is slightly redundant since every derivation comes from a vector field. Given  $V \in \mathfrak{X}(M)$ , the function  $D : C^\infty(M, \mathbb{R}) \rightarrow C^\infty(M, \mathbb{R})$  defined by  $D(v) = Vf$  is a derivation, and moreover every derivation comes from a vector field. Simply let  $V_p$  be the tangent vector such that  $V_p(f) = D(f)(p)$  for all  $p \in M$  and  $f \in C^\infty(M, \mathbb{R})$ .

We can also evaluate vector fields on other vector fields. Given  $V, W \in \mathfrak{X}(M)$ , we denote  $V(W)$  to be the function  $V(W) : C^\infty(M, \mathbb{R}) \rightarrow C^\infty(M, \mathbb{R})$  defined by:

$$V(W)(f) = V(Wf) \quad (37.14.11)$$

More explicitly, if  $p \in M$ , then:

$$\left( (V(W))(f) \right)(p) = V_p(Wf) \quad (37.14.12)$$

Remember that  $Wf$  is a function from  $C^\infty(M, \mathbb{R})$  into itself, where as  $V_p$  is a tangent vector and hence a linear functional from  $C^\infty(M, \mathbb{R})$  into  $\mathbb{R}$ . Thus, while strange, this definition is well posed. What's important is that now we can define the Lie bracket of two vector fields.

**Definition 37.14.15** The Lie Bracket is the function  $[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$  defined by:

$$[V, W] = V(W) - W(V) \quad (37.14.13)$$

More explicitly, given  $f \in C^\infty(M, \mathbb{R})$  and  $p \in M$ , we have:

$$[V, W]_p(f) = V_p(Wf) - W_p(Vf) \quad (37.14.14)$$

Before moving on to connections, we define integral curves.

**Definition 37.14.16** An integral curve of a vector field  $V \in \mathfrak{X}(M)$  is a curve  $\alpha : I \rightarrow M$  such that for all  $t \in I$ ,  $\dot{\alpha}(t) = V_{\alpha(t)}$ .

**Definition 37.14.17** A bilinear form on a vector field  $V$  over a field  $F$  is a function  $g : V \times V \rightarrow F$  such that:

$$g(a\mathbf{x} + b\mathbf{y}, \mathbf{z}) = ag(\mathbf{x}, \mathbf{z}) + bg(\mathbf{y}, \mathbf{z}) \quad (37.14.15)$$

$$g(\mathbf{x}, a\mathbf{y} + b\mathbf{z}) = ag(\mathbf{x}, \mathbf{y}) + bg(\mathbf{x}, \mathbf{z}) \quad (37.14.16)$$

**Definition 37.14.18** A symmetric form on a vector space  $V$  over a field  $F$  is a function  $g : V \times V \rightarrow F$  such that for all  $\mathbf{x}, \mathbf{y} \in V$ ,  $g(\mathbf{x}, \mathbf{y}) = g(\mathbf{y}, \mathbf{x})$ .

**Definition 37.14.19** A non-degenerate form on a vector space  $V$  over a field  $F$  is a function  $g : V \times V \rightarrow F$  such that for all  $\mathbf{x} \in V$  such that  $\mathbf{x} \neq \mathbf{0}$  there exists a  $\mathbf{y} \in V$  such that  $g(\mathbf{x}, \mathbf{y}) \neq 0$ .

In other words, if  $g(\mathbf{x}, \mathbf{y}) = 0$  for all  $\mathbf{y}$ , then  $\mathbf{x} = \mathbf{0}$ . Lastly, positive-definiteness.

**Definition 37.14.20** A positive-definite form on a vector space  $V$  over an ordered field  $F$  is a function  $g : V \times V \rightarrow F$  such that for all  $\mathbf{x} \in V$  it is true that  $g(\mathbf{x}, \mathbf{x}) \geq 0$  and  $g(\mathbf{x}, \mathbf{x}) = 0$  implies that  $\mathbf{x} = \mathbf{0}$ .

Semi-Riemannian and Riemannian geometry are concerned with smooth manifolds that have symmetric bilinear forms that are non-degenerate (Semi-Riemannian) or positive-definite (Riemannian). The vector spaces are simply the tangent spaces, and we need a function  $g$  that takes in a point  $p \in M$  and two elements of  $T_p M$  and returns a real number. We further want this function to be a symmetric non-degenerate bilinear form for every  $p$ , and we would also like this function to vary smoothly between points. So in essence, we want a function  $g : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow C^\infty(M, \mathbb{R})$ . Given a point  $p \in M$ , and two vector fields  $V, W \in \mathfrak{X}(M)$ , we want  $g(V, W)(p) = g_p(V_p, W_p)$  to vary smoothly with  $p$ . We also want symmetry, bilinearity, and non-degeneracy.

**Definition 37.14.21** A metric tensor on a smooth manifold  $(M, \tau, \mathcal{A})$  is a smooth non-degeneracy symmetric bilinear form  $g : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow C^\infty(M, \mathbb{R})$ .

**Definition 37.14.22** A semi-Riemannian manifold is a smooth manifold  $(M, \tau, \mathcal{A})$  with a metric tensor  $g$ .

## 37.15 Connections

We now want to talk about transporting data along a manifold in a parallel manner. For Euclidean space  $\mathbb{R}^n$  there is an intuitive manner to do this: Simply translate your vector from point  $a$  to point  $b$ . On a sphere there's a slightly intuitive manner. Translations may leave the sphere and so we need a new method. Given two points we can simply rotate a tangent vector at  $a$  to  $b$ , but the resultant vector depends on how one rotated the sphere. That is, along which path did one rotate. This is simply a consequence of the curvature of the sphere. For the problem of generalizing to arbitrary manifolds one might think coordinates suffice to perform parallel transport, but this is not true. Even on the sphere, using the two stereographic projections about the north and south pole, one runs into incompatibility issues. So the idea is to find a way of describing the rate of change of one vector field with respect to another. Given a vector field  $V$  and a tangent vector  $v \in T_p M$  for some point  $p \in M$  such that  $V_p = v$ , and given another vector field  $W$ , we can think of parallel transport as transporting  $v$  along an integral curve of  $W$ .

**Definition 37.15.1** A connection  $\nabla$  on a smooth manifold  $(M, \tau, \mathcal{A})$  is a function  $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$  such that:

$$\nabla(fV_1 + gV_2, W) = f\nabla(V_1, W) + g\nabla(V_2, W) \quad (37.15.1)$$

$$\nabla(V, aW_1 + bW_2) = a\nabla(V, W_1) + b\nabla(V, W_2) \quad (37.15.2)$$

$$\nabla(V, fW) = (Vf)W + f\nabla(V, W) \quad (37.15.3)$$

**Definition 37.15.2** A Levi-Civita connection on a semi-Riemannian manifold  $(M, g)$  is a connection  $\nabla$  that is torsion free (preserves the Lie bracket):

$$\nabla(V, W) - \nabla(W, V) = [V, W] \quad (37.15.4)$$

and preserves the metric tensor:

$$Xg(V, W) = g(\nabla(X, V), W) + g(V, \nabla(X, W)) \quad (37.15.5)$$

**Theorem 37.15.1: Fundamental Theorem of Riemannian Geometry**

If  $(M, g)$  is a semi-Riemannian manifold, then there is a unique Levi-Civita connection  $\nabla$  on  $(M, g)$ .

*Proof.* Just use the Koszul Formula:

$$\begin{aligned} 2g(\nabla(X, V), W) &= \partial_X(g(V, W)) + \partial_Y(g(X, W)) \\ &\quad - \partial_Z(g(X, V)) + g([X, V], W) \\ &\quad - g([X, W], V) - g([V, W], X) \end{aligned} \tag{37.15.6}$$

□

## 37.16 Pendulums

We can apply our new found Levi-Civita connection to a classic problem. Consider the sphere with it's standard embedding into  $\mathbb{R}^3$ . There is a natural metric tensor we can associate too it, combined with the standard smooth structure induced by the orthographic projections, that make  $S^2$  a Riemannian manifold. That is, there is a natural metric tensor that is also positive-definite. Take the dot product  $\langle \cdot | \cdot \rangle : \mathbb{R}^3 \rightarrow \mathbb{R}$  apply this to tangent vectors in  $T_p S^2$ . Since the tangent spaces are isometric to subspaces of  $\mathbb{R}^3$ , there's a nice way to do this. The dot product varies smoothly as one moves smoothly from one point to another, and so this creates a metric tensor  $g$  and  $(M, g)$  is a Riemannian manifold.

Consider a pendulum in Paris oscillating in the direction of north to south. We can attach a tangent vector to this pendulum pointing towards the north pole. What happens as the Earth rotates around it's axis after a full day? The resulting parallel transport is no long pointing towards the north pole, but rather has tipped 270 degrees and oscillates east to west. This is in agreement with the parallel transport we get from our Levi-Civita connection on a sphere. This connection gives rotations as the means of parallel transport, in correspondence with our intuition. The change in angle is a function of the area of the curve enclosed.

## 37.17 Spivak Calculus on Manifolds

Def  $\mathbb{R}^n$ ,  $\|\cdot\|_2$  is positive definite, Cauchy-Schwarz inequality, triangle inequality. That is,  $\|\cdot\|_2$  is a norm.  $\langle \cdot | \cdot \rangle$  on  $\mathbb{R}^n$  is inner product. Polarization identity:

$$\langle \mathbf{x} | \mathbf{y} \rangle = \frac{\|\mathbf{x} + \mathbf{y}\| + \|\mathbf{x} - \mathbf{y}\|}{4} \tag{37.17.1}$$

Parallelogram Law (for inner product space):

$$\frac{\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2}{2} = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \tag{37.17.2}$$

Apollonius's Identity:

$$\|\mathbf{z} - \mathbf{x}\|^2 + \|\mathbf{z} - \mathbf{y}\|^2 = \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|^2 + \left\|\mathbf{z} - \frac{\mathbf{x} + \mathbf{y}}{2}\right\|^2 \quad (37.17.3)$$

**Theorem 37.17.1.**

$$\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad (37.17.4)$$

*Proof.* For:

$$\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} + (-\mathbf{y})\| \leq \|\mathbf{x}\| + \|-\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\| \quad (37.17.5)$$

□

**Theorem 37.17.2.**

$$\|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| \quad (37.17.6)$$

*Proof.* For:

$$\|\mathbf{x}\| = \|\mathbf{x} - \mathbf{y} + \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\| \quad (37.17.7)$$

And therefore:

$$\|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} + \mathbf{y}\| \quad (37.17.8)$$

and similarly:

$$\|\mathbf{y}\| - \|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\| \quad (37.17.9)$$

and hence the inequality is proved. □

$L^p(\mathbb{R}, \mu)$  is an inner product space.

**Theorem 37.17.3.**  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  is a linear isometry if and only if it preserves inner products.

*Proof.* For if  $\|T\mathbf{x}\| = \|\mathbf{x}\|$ , then:

$$\langle T\mathbf{x}|T\mathbf{y} \rangle = \frac{\|T\mathbf{x} + T\mathbf{y}\| + \|T\mathbf{x} - T\mathbf{y}\|}{4} \quad (37.17.10a)$$

$$= \frac{\|T(\mathbf{x} + \mathbf{y})\| + \|T(\mathbf{x} - \mathbf{y})\|}{4} \quad (37.17.10b)$$

$$= \frac{\|\mathbf{x} + \mathbf{y} + \|\mathbf{x} - \mathbf{y}\|\|}{4} \quad (37.17.10c)$$

$$= \langle \mathbf{x} | \mathbf{y} \rangle \quad (37.17.10d)$$

If  $T$  preserves inner products, then:

$$\|T\mathbf{x}\| = \sqrt{\langle T\mathbf{x}|T\mathbf{x} \rangle} = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle} = \|\mathbf{x}\| \quad (37.17.11)$$

completing the proof. □

**Definition 37.17.1** The angle between two non-zero vectors  $\mathbf{x}, \mathbf{y}$  in an inner product space  $V$  is:

$$\angle(\mathbf{x}, \mathbf{y}) = \arccos\left(\frac{\langle \mathbf{x} | \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}\right) \quad (37.17.12)$$

**Theorem 37.17.4.** If  $T$  is a linear isometry,  $\mathbf{x}, \mathbf{y}$  non-zero, then:

$$\angle(T\mathbf{x}, T\mathbf{y}) = \angle(\mathbf{x}, \mathbf{y}) \quad (37.17.13)$$

*Proof.* For if  $T$  is a linear isometry, then:

$$\angle(T\mathbf{x}, T\mathbf{y}) = \arccos\left(\frac{\langle T\mathbf{x} | T\mathbf{y} \rangle}{\|T\mathbf{x}\| \|T\mathbf{y}\|}\right) = \arccos\left(\frac{\langle \mathbf{x} | \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}\right) = \angle(\mathbf{x}, \mathbf{y}) \quad (37.17.14)$$

□

The converse need not hold. Scaling  $\mathbf{x} \mapsto a\mathbf{x}$  is angle preserving but is not a linear isometry.

**Theorem 37.17.5.** If  $T$  is a linear isometry, and if  $\lambda$  is an eigenvalue, then  $|\lambda| = 1$ .

*Proof.* For if  $\mathbf{x}$  is an eigenvector for  $\lambda$  and  $\mathbf{y}$  is an eigenvector of  $\mu$ , then:

$$\|\mathbf{x}\| = \|T\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda| \|\mathbf{x}\| \quad (37.17.15)$$

but  $\mathbf{x}$  is an eigenvector and is thus non-zero, and hence  $\|\mathbf{x}\| \neq 0$ . Thus  $|\lambda| = 1$ . □

**Theorem 37.17.6.** If  $(V, \langle \cdot | \cdot \rangle)$  is a finite dimensional inner product space of dimension  $n \in \mathbb{N}$  over an absolute value field  $(F, |\cdot|)$ , if  $\mathbf{e} : \mathbb{Z}_n \rightarrow V$  is an orthogonal basis, if  $T : V \rightarrow V$  is a linear function such that there exists a sequence  $\lambda : \mathbb{Z}_n \rightarrow F$  such that for all  $k \in \mathbb{Z}_n$  it is true that  $T(\mathbf{e}_k) = \lambda_k \mathbf{e}_k$ , and if for all  $j \in \mathbb{Z}_n$  it is true that  $|\lambda_j| = |\lambda_0|$ , then  $T$  is an angle preserving linear transformation.

*Proof.* For since  $\mathbf{e}$  is an basis, for all  $\mathbf{x}, \mathbf{y} \in V$  there exists sequences  $a, b : \mathbb{Z}_n \rightarrow F$  such that:

$$\mathbf{x} = \sum_{k \in \mathbb{Z}_n} a_k \mathbf{e}_k \quad (37.17.16a) \qquad \mathbf{y} = \sum_{k \in \mathbb{Z}_n} b_k \mathbf{e}_k \quad (37.17.16b)$$

But since  $\mathbf{e}$  is an orthogonal basis, it is true that:

$$\left\langle \sum_{k \in \mathbb{Z}_n} a_k \mathbf{e}_k \mid \sum_{k \in \mathbb{Z}_n} b_k \mathbf{e}_k \right\rangle = \sum_{k \in \mathbb{Z}_n} a_k b_k \|\mathbf{e}_k\|^2 \quad (37.17.17)$$

but then:

$$\angle(T\mathbf{x}, T\mathbf{y}) = \arccos \left( \frac{\sum_{k \in \mathbb{Z}_n} a_k b_k \|T(\mathbf{e}_k)\|^2}{\sqrt{\sum_{k \in \mathbb{Z}_n} a_k^2 \|T(\mathbf{e}_k)\|^2} \sqrt{\sum_{k \in \mathbb{Z}_n} b_k^2 \|T(\mathbf{e}_k)\|^2}} \right) \quad (37.17.18a)$$

$$= \arccos \left( \frac{\sum_{k \in \mathbb{Z}_n} a_k b_k \|\lambda_k \mathbf{e}_k\|^2}{\sqrt{\sum_{k \in \mathbb{Z}_n} a_k^2 \|\lambda_k \mathbf{e}_k\|^2} \sqrt{\sum_{k \in \mathbb{Z}_n} b_k^2 \|\lambda_k \mathbf{e}_k\|^2}} \right) \quad (37.17.18b)$$

$$= \arccos \left( \frac{\sum_{k \in \mathbb{Z}_n} a_k b_k |\lambda_k|^2 \|\mathbf{e}_k\|^2}{\sqrt{\sum_{k \in \mathbb{Z}_n} a_k^2 |\lambda_k|^2 \|\mathbf{e}_k\|^2} \sqrt{\sum_{k \in \mathbb{Z}_n} b_k^2 |\lambda_k|^2 \|\mathbf{e}_k\|^2}} \right) \quad (37.17.18c)$$

$$= \arccos \left( \frac{|\lambda_0|^2 \sum_{k \in \mathbb{Z}_n} a_k b_k \|\mathbf{e}_k\|^2}{\sqrt{|\lambda_0|^2 \sum_{k \in \mathbb{Z}_n} a_k^2 \|\mathbf{e}_k\|^2} \sqrt{|\lambda_0|^2 \sum_{k \in \mathbb{Z}_n} b_k^2 \|\mathbf{e}_k\|^2}} \right) \quad (37.17.18d)$$

$$= \arccos \left( \frac{|\lambda_0|^2 \sum_{k \in \mathbb{Z}_n} a_k b_k \|\mathbf{e}_k\|^2}{|\lambda_0| \sqrt{\sum_{k \in \mathbb{Z}_n} a_k^2 \|\mathbf{e}_k\|^2} |\lambda_0| \sqrt{\sum_{k \in \mathbb{Z}_n} b_k^2 \|\mathbf{e}_k\|^2}} \right) \quad (37.17.18e)$$

$$= \arccos \left( \frac{\sum_{k \in \mathbb{Z}_n} a_k b_k \|\mathbf{e}_k\|^2}{\sqrt{\sum_{k \in \mathbb{Z}_n} a_k^2 \|\mathbf{e}_k\|^2} \sqrt{\sum_{k \in \mathbb{Z}_n} b_k^2 \|\mathbf{e}_k\|^2}} \right) \quad (37.17.18f)$$

$$= \angle(\mathbf{x}, \mathbf{y}) \quad (37.17.18g)$$

completing the proof.  $\square$

The hypothesis involved cannot be relaxed. For if the  $\mathbf{e}_k$  are not orthogonal, we can consider  $\mathbb{R}^2$  with  $\mathbf{e}_0 = (1, 0)$  and  $\mathbf{e}_1 = (1, 1)$ , and taking  $\lambda_0 = 1$  and  $\lambda_1 = -1$ , and let  $T$  be the linear transformation defined by  $T\mathbf{e}_i = \lambda_i \mathbf{e}_i$ . Then all of the conditions are satisfied with exception of the basis being orthogonal. However:

$$\angle((0, 1), (1, 0)) = \frac{\pi}{2} \neq \arccos \left( -\frac{2}{\sqrt{5}} \right) = \angle(T(0, 1), T(1, 0)) \quad (37.17.19)$$

and thus  $T$  is not angle preserving.

**Example 37.17.1** The rotation matrix in  $\mathbb{R}^2$  is an angle preserving linear transformation. Let  $\theta \in [0, 2\pi)$  and  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be defined by the matrix:

$$T = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \quad (37.17.20)$$

$T$  is angle preserving since it is a linear isometry. If  $\mathbf{x} \in \mathbb{R}^2$ , then:

$$\|T(\mathbf{x})\| = \|(\cos(\theta)x_0 + \sin(\theta)x_1, -\sin(\theta)x_0 + \cos(\theta)x_1)\| \quad (37.17.21a)$$

$$= \sqrt{(\cos(\theta)x_0 + \sin(\theta)x_1)^2 + (-\sin(\theta)x_0 + \cos(\theta)x_1)^2} \quad (37.17.21b)$$

$$= \sqrt{x_0^2 + x_1^2} \quad (37.17.21c)$$

and this is just the Euclidean norm of  $\mathbf{x}$ . This example also shows that angle function  $\angle$  is a good definition of angle, since:

$$\angle(\mathbf{x}, T\mathbf{x}) = \arccos\left(\frac{\langle \mathbf{x}|T(\mathbf{x}) \rangle}{\|\mathbf{x}\|\|T(\mathbf{x})\|}\right) = \arccos\left(\frac{\cos(\theta)\|\mathbf{x}\|^2}{\|\mathbf{x}\|^2}\right) = \theta \quad (37.17.22)$$

That is, the angle between  $\mathbf{x}$  and the rotation of  $\mathbf{x}$  by  $\theta$  is simply  $\theta$ , as hoped.

Pythagoras theorem:  $\mathbf{x}, \mathbf{y}$  orthogonal:

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \quad (37.17.23)$$

**Theorem 37.17.7.** *If  $B \subseteq \mathbb{R}^n$  is compact, if  $x \in \mathbb{R}$ , and if  $\mathcal{O}$  is an open cover of  $\{x\} \times B$ , then there is an open subset  $\mathcal{U} \subseteq \mathbb{R}$  such that  $x \in \mathcal{U}$  and such that there exists a finite subset  $\Delta \subseteq \mathcal{O}$  such that  $\Delta$  covers  $\mathcal{U} \times B$ .*

*Proof.* Since  $\mathcal{O}$  is a cover, for all  $\mathbf{y} \in B$  there is a  $\mathcal{V} \in \mathcal{O}$  such that  $(x, \mathbf{y}) \in \mathcal{V}$ . Thus, these  $\mathcal{V}$  cover  $\{x\} \times B$ , and since  $\{x\} \times B$  is compact, there is a finite subcover  $\Delta$ . Each element of  $\mathcal{V}$  is of the form  $\mathcal{U}_k \times \mathcal{W}_k$  with  $\mathcal{U}_k \subseteq \mathbb{R}$  open. Let  $\mathcal{U} = \bigcap \mathcal{U}_k$ .  $\square$

Spivak 1-17

**Theorem 37.17.8.** *If  $A, B \subseteq \mathbb{R}^n$ ,  $A$  closed,  $B$  compact, then  $\text{dist}(A, B) > 0$ .*

This may fail for two closed subsets. Consider  $\mathbb{R}^2$  and the graphs of  $y_1 = x^{-1}$  and  $y_2 = -x^{-1}$ . Both of these graphs are closed subsets of  $\mathbb{R}^2$ , but the distance between them is asymptotic to zero, though zero is never attained.

**Theorem 37.17.9.** *If  $U$  open  $C \subseteq U$  compact, there exists  $D$  compact such that  $C \subseteq \text{Int}_\tau(D)$  and  $D \subseteq U$ .*

Def projections, product functions.

**Definition 37.17.2: Oscillation of a Function**

The oscillation of a bounded function  $f : X \rightarrow \mathbb{R}$  from a metric space  $(X, d)$  into  $\mathbb{R}$  at a point  $x_0 \in X$  is the real number:

$$o(f, x_0) = \lim_{\delta \rightarrow 0^+} (M(x_0, f, \delta) - m(x_0, f, \delta))$$

where:

$$M(x_0, f, \delta) = \sup \{ f(x) \in \mathbb{R} \mid x \in X \text{ and } d(x, x_0) < \delta \}$$

and:

$$m(x_0, f, \delta) = \inf \{ f(x) \in \mathbb{R} \mid x \in X \text{ and } d(x, x_0) < \delta \}$$

**Theorem 37.17.10.** *If  $(X, d)$  is a metric space, if  $f : X \rightarrow \mathbb{R}$  is bounded, and if  $x_0 \in X$ , then  $f$  is continuous at  $x_0$  if and only if the oscillation of  $f$  at  $x_0$  is zero.*

**Theorem 37.17.11.** *If  $A \subseteq \mathbb{R}^n$  is a closed subset, if  $f : A \rightarrow \mathbb{R}$  is bounded, and if  $\varepsilon > 0$ , then the set  $B \subseteq \mathbb{R}^n$  defined by:*

$$B = \{ x \in A \mid o(f, x) \geq \varepsilon \} \quad (37.17.24)$$

*is a closed subset of  $\mathbb{R}^n$ .*

*Proof.* For it suffices to show that  $\mathbb{R}^n \setminus B$  is open. But  $B \subseteq A$ , and thus:

$$\mathbb{R}^n \setminus A \subseteq \mathbb{R}^n \setminus B \quad (37.17.25)$$

Suppose  $x \in \mathbb{R}^n \setminus B$ . Then either  $x \notin A$ , or  $o(x, f) < \varepsilon$ . If  $x \notin A$ , then  $x \in \mathbb{R}^n \setminus A$ , and since  $A$  is closed,  $\mathbb{R}^n \setminus A$  is open, and thus there is a ball about  $x$  that is contained in  $\mathbb{R}^n \setminus A$ . If  $x \in A$ , then there is a  $\delta > 0$  such that:

$$M(x, f, \delta) - m(x, f, x) < \varepsilon \quad (37.17.26)$$

But then for all  $y$  such that  $d(x, y) < \delta/2$  we have that:

$$M(x, f, \delta/2) - m(x, f, x/2) < \varepsilon \quad (37.17.27)$$

and thus there is a ball about  $x$  contained in  $\mathbb{R}^n \setminus B$ . Therefore  $\mathbb{R}^n \setminus B$  is open, and hence  $B$  is closed.  $\square$

If  $A \subseteq \mathbb{R}^2$  is defined by:

$$A = \{ (x, y) \in \mathbb{R}^2 \mid x > 0 \text{ and } 0 < y < x^2 \} \quad (37.17.28)$$

then every line through the origin contains an integral about the origin such that the line does not intersect  $A$  in this region. For the equation of such a line is  $y = mx$ . If  $m \leq 0$  then the line never intersects  $A$ . If  $m > 0$ , then for all  $0 < x < m$ ,  $(x, mx)$  is not contained in  $A$ . Define  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  by  $f(\mathbf{x}) = 0$  if  $\mathbf{x} \notin A$  and 1 if  $\mathbf{x} \in A$ . Let  $g_m : \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $g_m(t) = f((t, mt))$ . Then for all  $m$ ,  $g_m(t)$  is a constant function and hence continuous. However,  $f$  is not continuous. In particular,  $g_m$  is continuous at the origin for all  $m$  and  $f$  is not. This shows that to check for continuity it is necessary to look at every sequence, not just linear ones. Spivak 1.30

## 37.18 Spivak (Chapter 2)

Def differentiable  $f : \mathbb{R} \rightarrow \mathbb{R}$ .

## 37.19 Homework I

### 37.19.1 Sequential Spaces

#### Definition 37.19.1: Sequentially Continuous

A sequentially continuous function from a topological space  $(X, \tau_X)$  to a topological space  $(Y, \tau_Y)$  is a function  $f : X \rightarrow Y$  such that for every convergent sequence  $a : \mathbb{N} \rightarrow X$  and for every  $x \in X$  such that  $a_n \rightarrow x$ , it is true that  $f(a_n) \rightarrow f(x)$ .

We say *for every*  $x$  since limits need not be unique in spaces that aren't at least  $T_1$  (an *accessible* space).

**Theorem 37.19.1.** *If  $(X, \tau_X)$  and  $(Y, \tau_Y)$  are topological spaces, and if  $f : X \rightarrow Y$  is a continuous function, then it is sequentially continuous.*

*Proof.* For suppose not. Then there is a sequence  $a : \mathbb{N} \rightarrow X$  and a point  $x \in X$  such that  $a_n \rightarrow x$ , but  $f(a_n) \not\rightarrow f(x)$  (Def. 37.19.1). But then there exists an open subset  $\mathcal{V} \in \tau_Y$  such that  $f(x) \in \mathcal{V}$ , and for all  $N \in \mathbb{N}$  there exists an  $n \in \mathbb{N}$  such that  $n > N$  and  $f(a_n) \notin \mathcal{V}$ . But  $f$  is continuous, and therefore  $f^{-1}[\mathcal{V}]$  is an open subset of  $X$ . And since  $f(x) \in \mathcal{V}$  it is true that  $x \in f^{-1}[\mathcal{V}]$ . But  $a_n \rightarrow x$ , and thus there is an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  with  $n > N$  it is true that  $a_n \in f^{-1}[\mathcal{V}]$ . But then for all  $n > N$  it is true that  $f(a_n) \in \mathcal{V}$ , a contradiction. Therefore,  $f$  is sequentially continuous.  $\square$

This theorem does not reverse.

**Example 37.19.1** Let  $X$  be the order topology on the ordinal  $\omega_1 + 1$ , where  $\omega_1$  is the first uncountable ordinal. Define  $f : X \rightarrow \mathbb{R}$  by:

$$f(x) = \begin{cases} 0, & x \neq \omega_1 \\ 1, & x = \omega_1 \end{cases} \quad (37.19.1)$$

then  $f$  is sequentially continuous, but not continuous. If  $a : \mathbb{N} \rightarrow X$  is any sequence, either it converges to  $\omega_1$  or it is not. In the latter case the limit of  $f(a_n)$  will be zero, and in the former case there is an  $N \in \mathbb{N}$  such that for all  $n > N$  it is true that  $a_n = \omega_1$ , and hence  $f(a_n) \rightarrow 1$ . In both scenarios,  $f$  is sequentially continuous. It is not continuous since  $\{\omega_1\}$  is not an isolated point of  $X$ , yet  $f^{-1}[(0, \infty)] = \{\omega_1\}$ .

The problem with this space is that it is not first countable. First countability implies that sequentially continuous and continuous are identical concepts. There's a weaker such notion, called sequential spaces.

### Definition 37.19.2: Sequentially Open

A sequentially open subset of a topological space  $(X, \tau)$  is a subset  $\mathcal{U} \subseteq X$  such that for every sequence  $a : \mathbb{N} \rightarrow X$  such that  $a$  converges to an element  $x \in \mathcal{U}$  there exists an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  with  $n > N$ , it is true that  $a_n \in \mathcal{U}$ .

**Theorem 37.19.2.** *If  $(X, \tau)$  is a topological space, if  $\mathcal{U} \in \tau$  is an open subset of  $X$ , then  $\mathcal{U}$  is sequentially open.*

*Proof.* For suppose not. Then there is a sequence  $a : \mathbb{N} \rightarrow X$  such that  $a$  converges to a point  $x \in \mathcal{U}$ , but for all  $N \in \mathbb{N}$  there is an  $n \in \mathbb{N}$  such that  $n > N$  and  $a_n \notin \mathcal{U}$  (Def. 37.19.2). But if  $\mathcal{U}$  is an open subset containing  $x$ , and if  $a_n \rightarrow x$ , then there is an  $N \in \mathbb{N}$  such that for all  $n > N$  it is true that  $a_n \in \mathcal{U}$ , a contradiction.  $\square$

### Definition 37.19.3: Sequential Space

A sequential topological space is a topological space  $(X, \tau)$  such that for every sequentially open subset  $\mathcal{U} \subseteq X$  it is true that  $\mathcal{U} \in \tau$ .

As the name suggests, the usefulness of sequential spaces stems from the fact that continuity and sequential continuity are equivalent. One direction is true of any topological space (Thm. 37.19.1). We now prove the other direction in the setting of a sequential space.

**Theorem 37.19.3.** *If  $(X, \tau_X)$  is a sequential topological space, if  $(Y, \tau_Y)$  is a topological space, and if  $f : X \rightarrow Y$  is a sequentially continuous function, then  $f$  is continuous.*

*Proof.* For suppose not. Then there is a open subset  $\mathcal{V} \in \tau_Y$  such that  $f^{-1}[\mathcal{V}]$  is not open. But  $(X, \tau_X)$  is sequential, and thus if  $f^{-1}[\mathcal{V}]$  is not open, then it is not sequentially open (Def. 37.19.3). But then there is a sequence  $a : \mathbb{N} \rightarrow X$  and a point  $x \in f^{-1}[\mathcal{V}]$  such that  $a_n \rightarrow x$  but for all  $N \in \mathbb{N}$  there is an  $n \in \mathbb{N}$  with  $n > N$  such that  $a_n \notin f^{-1}[\mathcal{V}]$  (Def. 37.19.2). But  $f$  is sequentially continuous, and thus if  $a_n \rightarrow x$ , then  $f(a_n) \rightarrow f(x)$  (Def. 37.19.1). But  $\mathcal{V}$  is open and  $f(x) \in \mathcal{V}$ . Therefore if  $f(a_n) \rightarrow f(x)$ , then there is an  $N \in \mathbb{N}$  such that for all  $n \in \mathbb{N}$  with  $n > N$ , it is true that  $f(a_n) \in \mathcal{V}$ . But then for all  $n > N$  it is true that  $a_n \in f^{-1}[\mathcal{V}]$ , a contradiction. Therefore,  $f$  is continuous.  $\square$

Two amazing equivalent definitions of sequential spaces that we will not prove since these will not be used.

- $(X, \tau)$  is sequential if and only if it is the quotient of a first countable space.
- $(X, \tau)$  is sequential if and only if it is the quotient of a metrizable space.

In just about every setting we will be working with first countable spaces. It would be nice to know that first countable spaces are sequential, and indeed they are.

**Theorem 37.19.4.** *If  $(X, \tau)$  is a first countable topological space, then it is a sequential topological space.*

*Proof.* For suppose not. Then there is a sequentially open subset  $\mathcal{U} \subseteq X$  such that  $\mathcal{U}$  is not open (Def. 37.19.3). But since  $(X, \tau)$  is first countable, for all  $x \in \mathcal{U}$  there is a countable neighborhood basis. By the axiom of choice there is a function  $\mathcal{B} : \mathcal{U} \rightarrow \mathcal{P}(\tau)$  such that for all  $x \in \mathcal{U}$ ,  $\mathcal{B}_x$  is a countable neighborhood basis of  $x$ . But if for all  $x \in \mathcal{U}$  it is true that  $\mathcal{B}_x$  is countable, then there is surjective function  $\mathcal{V}_x : \mathbb{N} \rightarrow \mathcal{B}_x$ . Then there exists an  $N \in \mathbb{N}$  such that  $\mathcal{V}_{x,N} \subseteq \mathcal{U}$ . For suppose not and let  $B : \mathbb{N} \rightarrow X$  be defined by:

$$B_n = \bigcap_{k \in \mathbb{Z}_n} \mathcal{V}_{x,k} \tag{37.19.2}$$

Since  $B_n$  is the intersection of finitely many open sets, it is open. Moreover it is non-empty since  $x \in B_n$  for all  $n$ . But then  $B_n$  is an open neighborhood about  $x$ , and since  $\mathcal{B}_x$  is a neighborhood basis there is an element  $V \in \mathcal{B}_x$  such that  $V \subseteq B_n$ . But  $\mathcal{V} : \mathbb{N} \rightarrow \mathcal{B}_x$  is a surjection, and hence there is an  $N \in \mathbb{N}$  such that  $\mathcal{V}_{x,N} = V$ . But by hypothesis,  $\mathcal{V}_{x,N} \not\subseteq \mathcal{U}$  and hence there is an element

$y \in \mathcal{V}_{x,N}$  such that  $y \notin \mathcal{U}$ . Therefore we have shown that the sequence  $A_n$  defined by:

$$A_n = \{y \in B_n \mid y \notin \mathcal{U}\} \quad (37.19.3)$$

is non-empty for all  $n \in \mathbb{N}$ , and hence by the axiom of (countable) choice there is a sequence  $y : \mathbb{N} \rightarrow X$  such that  $y_n \in A_n$  for all  $n \in \mathbb{N}$ . That is, for all  $n \in \mathbb{N}$ ,  $y_n \in B_n$  and  $y_n \notin \mathcal{U}$ . But if  $y_n \in B_n$  for all  $n \in \mathbb{N}$ , then  $y_n \rightarrow x$ . For if not, then there is an open set  $U \in \tau$  such that  $x \in U$  and for all  $N \in \mathbb{N}$  there exists an  $n > N$  such that  $y_n \notin U$ . But  $\mathcal{B}_x$  is a neighborhood basis of  $x$ , and hence there is a  $V \in \mathcal{B}_x$  such that  $V \subseteq U$ . But  $\mathcal{V}_x : \mathbb{N} \rightarrow \mathcal{B}_x$  is a surjection, and hence there is an  $N \in \mathbb{N}$  such that  $\mathcal{V}_{x,N} = V$ . But then for all  $n > N$ ,  $B_n \subseteq \mathcal{V}_{x,N}$  and thus  $B_n \subseteq U$ . But then for all  $n > N$ ,  $y_n \in U$ , a contradiction. Hence,  $y_n \rightarrow x$ . But for all  $n \in \mathbb{N}$ ,  $y_n \notin \mathcal{U}$ . Thus  $y_n$  is a sequence such that  $y_n \rightarrow x$ , but for all  $N \in \mathbb{N}$  there is an  $n \in \mathbb{N}$  with  $n > N$  such that  $y_n \notin \mathcal{U}$ , a contradiction since  $x \in \mathcal{U}$  and  $\mathcal{U}$  is sequentially open (Def. 37.19.2). Thus, for all  $x \in \mathcal{U}$  there is an open subset  $\mathcal{V}_{x,N}$  such that  $x \in \mathcal{V}_{x,N}$  and  $\mathcal{V}_{x,N} \subseteq \mathcal{U}$ . But then  $\mathcal{U}$  is simply the union over all of these open sets, and is thus open, a contradiction. Hence,  $(X, \tau)$  is sequential.  $\square$

And now, our useful corollary.

**Theorem 37.19.5.** *If  $(X, \tau_X)$  is a first countable topological space, if  $(Y, \tau_Y)$  is a topological space, and if  $f : X \rightarrow Y$  is sequential continuous, then  $f$  is continuous.*

*Proof.* For since  $(X, \tau_X)$  is first countable, it is sequential (Thm. 37.19.4). But if  $(X, \tau_X)$  is a sequential topological space and if  $f$  is sequentially continuous, then  $f$  is continuous (Thm. 37.19.3). Therefore,  $f$  is continuous.  $\square$

It's laborious to show a certain space is first countable, especially in the case of manifold theory, and so we rely on the following theorems.

**Theorem 37.19.6.** *If  $(X, \tau)$  is a locally metrizable topological space, then it is first countable.*

*Proof.* For suppose not. Then there is a point  $x \in X$  with no countable neighborhood basis. But if  $(X, \tau)$  is locally metrizable, then there is an open subset  $\mathcal{U}$  such that  $x \in \mathcal{U}$  and  $(\mathcal{U}, \tau_{\mathcal{U}})$  is metrizable, where  $\tau_{\mathcal{U}}$  is the subspace topology. But then there is a metric  $d : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}$  such that  $d$  induces  $\tau_{\mathcal{U}}$ . Let  $\mathcal{B} \subseteq \tau$  be defined by:

$$\mathcal{B} = \{\mathbb{B}_{n^{-1}}^{(\mathcal{U}, d)}(x) \mid n \in \mathbb{N}^+\} \quad (37.19.4)$$

That is, the set of all open balls about  $x$  of radius  $1/n$ . Suppose  $\mathcal{V} \in \tau$  is such that  $x \in \mathcal{V}$ . But then  $x \in \mathcal{V} \cap \mathcal{U}$ . But  $\mathcal{V} \cap \mathcal{U}$  is an open subset of  $\mathcal{U}$ , and thus there is an  $r > 0$  such that:

$$\mathbb{B}_r^{(\mathcal{U}, d)}(x) \subseteq \mathcal{V} \quad (37.19.5)$$

but by Archimedean's Theorem, there is an  $N \in \mathbb{N}$  such that  $N > r$ . But it is true that  $\mathbb{B}_{N-1}^{(U,d)}(x) \in \mathcal{B}$ , and hence  $\mathcal{B}$  is a countable neighborhood basis of  $x$ , a contradiction. Thus,  $(X, \tau)$  is first countable.  $\square$

### 37.19.2 Various Types of Compactness

#### Definition 37.19.4: $\sigma$ Compact

A  $\sigma$  compact topological space is a topological space  $(X, \tau)$  such that there exists a sequence  $K : \mathbb{N} \rightarrow \mathcal{P}(X)$  of compact subsets of  $X$  such that:

$$X = \bigcup_{n \in \mathbb{N}} K_n$$

This definition gives the following triviality.

**Theorem 37.19.7.** *If  $(X, \tau)$  is a compact topological space, then it is  $\sigma$  compact.*

*Proof.* For let  $K : \mathbb{N} \rightarrow \mathcal{P}(X)$  be defined by  $K_n = X$ . Then  $K_n$  is compact for all  $n$  and  $\bigcup K_n = X$ . Thus,  $(X, \tau)$  is  $\sigma$  compact (Def. 37.19.4).  $\square$

#### Definition 37.19.5: Lindelöf Topological Space

A Lindelöf topological space is a topological space  $(X, \tau)$  such that for every open cover  $\mathcal{O}$  of  $X$  there exists a countable subcover.

**Theorem 37.19.8.** *If  $(X, \tau)$  is a second countable topological space, then it is Lindelöf.*

*Proof.* For suppose not. Then there exists an open cover  $\mathcal{O}$  with no countable subcover (Def. 37.19.5). But if  $(X, \tau)$  is second countable, then there is a countable basis  $\mathcal{B}$  of  $\tau$ . But if  $\mathcal{B}$  is a basis, and since  $\mathcal{O}$  is an open cover, for all  $\mathcal{V} \in \mathcal{O}$  it is true that  $\mathcal{V}$  is open and thus there is an element  $\mathcal{U} \in \mathcal{B}$  such that  $\mathcal{U} \subseteq \mathcal{V}$ . That is, the function  $F : \mathcal{O} \rightarrow \mathcal{P}(\mathcal{B})$  defined by:

$$F(\mathcal{V}) = \{\mathcal{U} \in \mathcal{B} \mid \mathcal{U} \subseteq \mathcal{V}\} \quad (37.19.6)$$

is such that  $F(\mathcal{V}) \neq \emptyset$  for all  $\mathcal{V} \in \mathcal{O}$ . But then  $F[\mathcal{O}] \subseteq \mathcal{P}(\mathcal{B})$  is a non-empty subset of  $\mathcal{P}(\mathcal{B})$  such that  $\emptyset \notin F[\mathcal{O}]$ . Hence by the axiom of choice there is a function  $G : F[\mathcal{O}] \rightarrow \mathcal{B}$  such that for all  $\Delta \in F[\mathcal{O}]$ ,  $G(\Delta) \in \Delta$ .

Let  $B = G[F[\mathcal{O}]]$ . But  $\mathcal{B}$  is countable and  $B \subseteq \mathcal{B}$ , and thus  $B$  is countable. Moreover,  $F \circ G : \mathcal{O} \rightarrow B$  is surjective by the definition of  $B$  and hence there exists a right inverse  $H : B \rightarrow \mathcal{O}$ . And since  $B$  is countable,  $H[B]$  is a countable subset of  $\mathcal{O}$ . But  $\mathcal{O}$  has no countable subcover, and hence there is an  $x \in X$  such that  $x \notin \bigcup H[B]$ . But  $\mathcal{O}$  is a cover of  $X$ , and thus there is a  $\mathcal{V} \in \mathcal{O}$  such that  $x \in \mathcal{V}$ . And since  $\mathcal{B}$  is a basis, there is an element  $\mathcal{U}_1 \in \mathcal{B}$  such that  $x \in \mathcal{U}_1$ . But then  $x \in \mathcal{U}_1 \cap \mathcal{V}$  and since  $\mathcal{B}$  is a basis, there is an element  $\mathcal{U}_2 \in \mathcal{B}$  such that  $x \in \mathcal{U}_2$  and  $\mathcal{U}_2 \subseteq \mathcal{U}_1 \cap \mathcal{V}$ . But then  $H(\mathcal{U}_2)$  is an element of  $H[B]$  such that  $\mathcal{U}_2 \subseteq H(\mathcal{U}_2)$ , and hence  $x \in H(\mathcal{U}_2)$ . A contradiction, since  $x \notin \bigcup H[B]$ . Therefore,  $(X, \tau)$  is Lindelöf.  $\square$

**Theorem 37.19.9.** *If  $(X, \tau)$  is a  $\sigma$  compact topological space, then it is Lindelöf.*

*Proof.* For suppose not. Then there exists an open cover  $\mathcal{O}$  of  $X$  with no countable subcover. But  $X$  is  $\sigma$  compact, and hence there is a sequence  $K : \mathbb{N} \rightarrow \mathcal{P}(X)$  of compact sets such that  $X = \bigcup K_n$  (Def. 37.19.4). But then for all  $n \in \mathbb{N}$ ,  $K_n$  is a subset of  $X$ , and hence  $\mathcal{O}$  is an open cover of  $K_n$ . But by hypothesis,  $K_n$  is compact and hence there is a finite subcover  $\mathcal{D} \subseteq \mathcal{O}$  of  $K_n$ . That is, if we define the sequence  $A_n : \mathbb{N} \rightarrow \mathcal{P}(\mathcal{O})$  by:

$$A_n = \{\mathcal{D} \in \mathcal{P}(\mathcal{O}) \mid \mathcal{D} \text{ is finite and } K_n \subseteq \bigcup \mathcal{D}\} \quad (37.19.7)$$

then for all  $n \in \mathbb{N}$ ,  $A_n$  is non-empty. Hence by the axiom of (countable) choice, there is a choice function  $\Delta : \mathbb{N} \rightarrow \mathcal{P}(\mathcal{O})$  such that for all  $n \in \mathbb{N}$ ,  $\Delta_n$  is a finite open subcover of  $K_n$ . But then the union of all  $\Delta_n$  is the countable union of finite collections, and is hence countable. But then this collection covers  $K_n$  for all  $n \in \mathbb{N}$ , and hence covers  $\bigcup K_n$ . But  $\bigcup K_n = X$ , a contradiction since  $\mathcal{O}$  has no countable subcover. Thus,  $(X, \tau)$  is Lindelöf.  $\square$

This theorem reverses if we add locally compact. The requirement of local compactness can be seen by examining the irrational numbers with the subspace topology. This is Lindelöf (since it is second countable), but it is not  $\sigma$  compact. For any compact subset of the irrationals must have empty interior, and by the Baire category theorem the irrationals cannot be written as the countable union of nowhere dense subsets. Hence they cannot possibly  $\sigma$  compact.

### Definition 37.19.6: Locally Compact Topological Space

A locally compact topological space is a topological space  $(X, \tau)$  such that for all  $x \in X$  there exists a compact subset  $K \subseteq X$  and an open set  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  and  $\mathcal{U} \subseteq K$ .

**Theorem 37.19.10.** *If  $(X, \tau)$  is a locally compact Lindelöf space, then it is  $\sigma$  compact.*

*Proof.* For if  $(X, \tau)$  is locally compact, then for all  $x \in X$  there is a compact set  $K \subseteq X$  and an open set  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  and  $\mathcal{U} \subseteq K$  (Def. 37.19.6). Invoking the axiom of choice, there is a function  $A : X \rightarrow \tau \times \mathcal{P}(X)$  such that for all  $x \in X$   $A_x = (\mathcal{U}, K)$  where  $x \in \mathcal{U}$  and  $\mathcal{U} \subseteq K$ , where  $K$  is compact. But then the collection of all such  $\mathcal{U}$  is an open cover of  $X$ . But  $X$  is Lindelöf and hence there is a countable subcover. But then the subcollection of all  $K$  form a countable collection of compact sets that cover  $X$ . Hence,  $(X, \tau)$  is  $\sigma$  compact (Def. 37.19.4).  $\square$

### Definition 37.19.7: Sequentially Compact

A sequentially compact topological space is a topological space  $(X, \tau)$  such that for every sequence  $a : \mathbb{N} \rightarrow X$  there exists a strictly increasing sequence  $k : \mathbb{N} \rightarrow \mathbb{N}$  such that  $a \circ k : \mathbb{N} \rightarrow X$  converges.

Since sequences are easier to work with than large open covers, we seek a simple constraint that forces compact spaces to be sequentially compact. As it turns out, first countability suffices.

**Theorem 37.19.11.** *If  $(X, \tau)$  is first countable and compact, then it is sequentially compact.*

*Proof.* For suppose not. Then there is a sequence  $a : \mathbb{N} \rightarrow X$  with no convergent subsequence. But since  $X$  is compact, there must be a limit point. For if not, then for all  $x \in X$  there is an open set  $\mathcal{U}_x$  such that only finitely many  $n \in \mathbb{N}$  are such that  $a_n \in \mathcal{U}_x$ . But these  $\mathcal{U}_x$  cover  $X$ , and since  $X$  is compact there is a finite subcover. But then finitely many such sets contain every  $a_n$ , a contradiction since  $\mathbb{N}$  is not finite. Hence, there is a limit point. But  $(X, \tau)$  is first countable and hence there is a neighborhood basis  $\mathcal{B}$  of  $x$ . Let  $B : \mathbb{N} \rightarrow \mathcal{B}$  be a surjection and let  $U_n = \bigcap B_k$  be the intersection of the first  $n$  elements. Then  $U_n$  is an open subset that contains  $x$ , and hence there is a  $j \in \mathbb{N}$  such that  $a_j \in U_n$ , giving us a convergent subsequence. A contradiction. Hence,  $(X, \tau)$  is sequentially compact.  $\square$

Lastly, before moving away from compactness issues, the separation axioms that manifolds satisfy are often useful and so we prove them.

**Theorem 37.19.12.** *If  $(X, \tau)$  is a compact Hausdorff topological space, then it is regular.*

*Proof.* For suppose  $C \subseteq X$  is closed, and  $x \in X \setminus C$ . But closed subsets of compact spaces are compact, and thus  $C$  is compact. And since  $X$  is Hausdorff, for all  $y \in C$  there are open subsets  $\mathcal{U}_y, \mathcal{V}_y$  such that  $x \in \mathcal{U}_y$ ,  $y \in \mathcal{V}_y$ , and  $\mathcal{U}_y \cap \mathcal{V}_y = \emptyset$ . But the  $\mathcal{V}_y$  cover  $C$ , and hence there is a finite subcover  $\mathcal{V}_n$ . Let  $\mathcal{U} = \bigcap \mathcal{U}_n$  and  $\mathcal{V} = \bigcup \mathcal{V}_n$ . Then since  $\mathcal{U}$  is the finite intersection of open subsets, it is open. And since  $\mathcal{V}$  is the union of open subsets, it is open. But  $x \in \mathcal{U}$ ,  $C \subseteq \mathcal{V}$ , and  $\mathcal{U} \cap \mathcal{V} = \emptyset$ . Hence,  $(X, \tau)$  is regular.  $\square$

**Theorem 37.19.13.** *If  $(X, \tau)$  is a compact Hausdorff space, then it is normal.*

*Proof.* For suppose not. Then there are two disjoint closed sets  $C_1$  and  $C_2$  that cannot be separated by disjoint open subsets. But if  $(X, \tau)$  is compact and Hausdorff, then it is regular (Thm. 37.19.12). Thus for all  $x \in C_1$  there is an open set  $\mathcal{U}_x$  and an open set  $\mathcal{V}_x$  such that  $x \in \mathcal{U}_x$ ,  $C_2 \subseteq \mathcal{V}_x$ , and  $\mathcal{U}_x \cap \mathcal{V}_x = \emptyset$ . But the  $\mathcal{V}_x$  cover  $C_1$ , and since  $C_1$  is compact, finitely many cover it. Let  $U_k$  and  $V_k$  be the corresponding finite collections. But  $C_2 \subseteq \bigcap V_k$ , and  $\bigcap V_k$  is the finite intersection of open subsets, which is therefore open. Moreover,  $(\bigcap V_k) \cap (\bigcup U_k) = \emptyset$  since  $V_k \cap U_k = \emptyset$  for all  $k$ . But then  $\bigcap V_k$  and  $\bigcup U_k$  are open subsets that are disjoint and contain  $C_1$  and  $C_2$ , which is a contradiction. Therefore,  $(X, \tau)$  is normal.  $\square$

**Theorem 37.19.14.** *If  $(X, \tau)$  is locally compact and Hausdorff, then it is regular.*

*Proof.* For let  $C \subseteq X$  be closed and  $x \in X \setminus C$ . Since  $X$  is locally compact there is an open subset  $\mathcal{U} \in \tau$  and a compact subset  $K \subseteq X$  such that  $x \in \mathcal{U}$  and  $\mathcal{U} \subseteq K$  (Def. 37.19.6). But since  $K$  is compact and  $X$  is Hausdorff,  $K$  is closed. But  $C$  is closed, and therefore  $C \cap K$  is closed. If  $C \cap K = \emptyset$ , let  $\mathcal{V} = X \setminus K$ . Since  $K$  is closed,  $\mathcal{V}$  is therefore open. But  $C \subseteq \mathcal{V}$  and  $\mathcal{U} \cap \mathcal{V} = \emptyset$ , thus separating  $x$  and  $C$ . If  $C \cap K \neq \emptyset$ , then since  $X$  is Hausdorff,  $\{x\}$  is closed. But then  $(C \cap K) \cup \{x\}$  is a closed subset of a compact set, and is hence compact. But compact Hausdorff spaces are regular (Thm. 37.19.12). Thus there are open subsets  $\mathcal{U}_0, \mathcal{V}_0$  such that  $x \in \mathcal{U}_0$ ,  $C \cap K \subseteq \mathcal{V}_0$  and  $\mathcal{U}_0 \cap \mathcal{V}_0 = \emptyset$ . Let  $\mathcal{V} = \mathcal{V}_0 \cup (X \setminus K)$ . Then  $C \subseteq \mathcal{V}$ , and  $\mathcal{V} \cap \mathcal{U}_0 = \emptyset$ , hence separating  $x$  and  $C$ .  $\square$

**Theorem 37.19.15.** *If  $(X, \tau)$  is a regular Lindelöf topological space, then it is normal.*

*Proof.* For let  $C_1, C_2$  be disjoint closed sets. Since  $X$  is regular, for all  $x \in C_1$  there are open subsets  $\mathcal{U}_x, \mathcal{V}_x$  that have disjoint closures and such that  $x \in \mathcal{U}_x$  and  $C_2 \subseteq \mathcal{V}_x$ . But then the collection of all  $\mathcal{U}_x$  together with  $X \setminus C_1$  form an open cover of  $X$ . But  $X$  is Lindelöf and hence there is a countable subcover

(Def. 37.19.5). Let  $\mathcal{U}_k$  and  $\mathcal{V}_k$  be the corresponding sets, and define:

$$U_n = \mathcal{U}_n \setminus \left( \bigcup_{k \in \mathbb{Z}_n} \text{Cl}_\tau(V_k) \right) \quad (37.19.8)$$

$$V_n = \mathcal{V}_n \setminus \left( \bigcup_{k \in \mathbb{Z}_n} \text{Cl}_\tau(U_k) \right) \quad (37.19.9)$$

Let  $U = \bigcup U_n$  and  $V = \bigcup V_n$ . Then by construction,  $U \cap V = \emptyset$  and  $U$  and  $V$  contain  $C_1$  and  $C_2$ , respectively. Hence,  $(X, \tau)$  is normal.  $\square$

**Theorem 37.19.16.** *If  $(X, \tau)$  is a locally compact Hausdorff topological space that is  $\sigma$  compact, then it is normal.*

*Proof.* For if  $(X, \tau)$  is locally compact and Hausdorff, then it is regular (Thm. 37.19.14). But if  $(X, \tau)$  is  $\sigma$  compact, then it is Lindelöf (Thm. 37.19.9). But regular Lindelöf spaces are normal (Thm. 37.19.15), completing the proof.  $\square$

### Definition 37.19.8: Paracompact

A paracompact topological space is a topological space  $(X, \tau)$  such that for every open cover  $\mathcal{O}$  of  $X$ , there exists a locally finite refinement  $\Delta$  of  $\mathcal{O}$ .

### Theorem 37.19.17: Michael's Theorem

If  $(X, \tau)$  is regular and such that every cover  $\mathcal{O}$  has a refinement  $\Delta = \bigcup \mathcal{O}_k$  where each  $\mathcal{O}_k$  is locally finite, then  $X$  is paracompact.

**Theorem 37.19.18.** *If  $(X, \tau)$  is regular and Lindelöf, then it is paracompact.*

*Proof.* For suppose not. But if  $X$  is regular and not paracompact, then by Michael's theorem there is a cover  $\mathcal{O}$  such that there is no countably locally finite refinement (Thm. 37.19.17). But  $X$  is Lindelöf and hence there is a countable subcover  $\Delta$ . But then  $\Delta = \bigcup \{B_n\}$ , where  $B_n \in \Delta$ . And  $\{B_n\}$  is finite, and hence locally finite, which is a contradiction. Therefore,  $X$  is paracompact.  $\square$

**Theorem 37.19.19.** *If  $(X, \tau)$  is a topological space, if  $\mathcal{O} \subseteq \tau$  is a countable subset of open sets such that for all  $\mathcal{U} \in \mathcal{O}$ ,  $(\mathcal{U}, \tau_{\mathcal{U}})$  is second countable, where  $\tau_{\mathcal{U}}$  is the subspace topology, then  $(X, \tau)$  is second countable.*

*Proof.* For let  $\mathcal{B}$  be the collection of all of the bases for all of the  $\mathcal{U} \in \mathcal{O}$ . Since it is the countable union of countable sets, it is countable. But since  $\mathcal{U}$  is open for all  $\mathcal{U} \in \mathcal{O}$ , this is a countable open cover of  $X$ . It suffices to show that it is a basis. Let  $\mathcal{V} \in \tau$  be an open subset. But then:

$$\mathcal{V} = \mathcal{V} \cap X \tag{37.19.10}$$

$$= \mathcal{V} \cap \left( \bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} \right) \tag{37.19.11}$$

$$= \bigcup_{\mathcal{U} \in \mathcal{O}} (\mathcal{V} \cap \mathcal{U}) \tag{37.19.12}$$

But  $\mathcal{V} \cap \mathcal{U}$  is an open subset of the subspace  $(\mathcal{U}, \tau_{\mathcal{U}})$ , and hence there is a subset of  $\Delta_{\mathcal{V}} \subseteq \mathcal{B}_{\mathcal{U}}$  such that  $\mathcal{V} \cap \mathcal{U} = \bigcup \Delta_{\mathcal{V}}$ . But then the entire of  $\mathcal{V}$  is the union of all such collections, each of which is contained in  $\mathcal{B}$ , and hence  $\mathcal{V}$  can be written as the union of elements of  $\mathcal{B}$ . Thus,  $\mathcal{B}$  is a basis.  $\square$

The requirement that the covering collection be open subspaces is crucial. The quotient space  $\mathbb{R}/R$ , where  $R$  is the equivalence relation generated by  $nRm$  for all  $n, m \in \mathbb{Z}$ , can be thought of as a countable collection of rings all glued together at the origin. Hence, it can be covered by countably many closed subspaces, each of which is homeomorphic to  $\mathbb{S}^1$  in the subspace topology, and hence each of which is second countable. However, this space  $\mathbb{R}/R$  is not even first countable, let alone second countable. The point  $[0] \in \mathbb{R}/R$  has no countable neighborhood basis.

### 37.19.3 Connectedness

**Theorem 37.19.20.** *If  $(X, \tau)$  is locally path connected, if  $x \in X$ , and if  $\mathcal{U} \subseteq X$  is a path connected component of  $X$ , then  $\mathcal{U}$  is open.*

*Proof.* For if  $(X, \tau)$  is locally path connected, there is a basis  $\mathcal{B}$  of open and path connected subsets of  $X$ . But if  $\mathcal{U} \subseteq X$  is a path connected component, then for all  $x, y \in \mathcal{U}$  there is a path  $\gamma : [0, 1] \rightarrow \mathcal{U}$  connecting  $x$  and  $y$ , and for all  $z \in X$  such that  $z \notin \mathcal{U}$ , there is no path between  $x$  and  $z$ . But  $\mathcal{B}$  is a basis, and hence for all  $x \in \mathcal{U}$  there is a  $B \in \mathcal{B}$  such that  $x \in B$ . By choice, we get a function  $B : \mathcal{U} \rightarrow \mathcal{B}$ . Moreover, since  $B \in \mathcal{B}$ , it is path connected. But then for all  $x \in \mathcal{U}$ ,  $x \in B_x$ , and since  $B_x$  is path connected it is true that  $B_x \subseteq \mathcal{U}$  since  $\mathcal{U}$  is a path connected component. But then  $\mathcal{U} = \bigcup B_x$ , which is the union of open sets, and hence  $\mathcal{U}$  is open.  $\square$

**Theorem 37.19.21.** *If  $(X, \tau)$  is locally path connected and connected, then it is path connected.*

*Proof.* For if not then there are two points  $x, y \in X$  with no path between them. But the since  $(X, \tau)$  is locally path connected, the path connected components

of  $x$  and  $y$  are open (Thm. 37.19.20). But let  $\mathcal{U}$  be the path connected component containing  $x$ , and let  $\mathcal{V}$  be the union of all other path connected components. Then  $\mathcal{V}$  is non-empty since  $y \in \mathcal{V}$ , and hence  $\mathcal{U}$  and  $\mathcal{V}$  are non-empty disjoint open subsets that cover  $X$ , a contradiction since  $X$  is connected.  $\square$

### 37.19.4 Topological Manifolds

#### Definition 37.19.9: Locally Euclidean

A locally Euclidean topological space is a topological space  $(X, \tau)$  such that for all  $x \in X$  there exists an open subset  $\mathcal{U} \in \tau$  and an  $n \in \mathbb{N}$  such that  $x \in \mathcal{U}$  and  $\mathcal{U}$  is homeomorphic to an open subset of  $\mathbb{R}^n$ .

**Theorem 37.19.22.** *If  $(X, \tau)$  is locally Euclidean, then for all  $x \in X$  there is an open subset  $\mathcal{U} \in \tau$  and an  $n \in \mathbb{N}$  such that  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^n$ .*

*Proof.* For if  $(X, \tau)$  is locally Euclidean, then for all  $x \in X$  there is an open subset  $\mathcal{V} \in \tau$  and an  $n \in \mathbb{N}$  such that  $x \in \mathcal{V}$  and  $\mathcal{V}$  is homeomorphic to an open subset of  $\mathbb{R}^n$ . But then there is an injective continuous open mapping  $\varphi : \mathcal{V} \rightarrow \mathbb{R}^n$ . Let  $\mathbf{y} = \varphi(x)$ . But since  $\varphi$  is an open mapping, and since  $\mathcal{V}$  is open,  $\varphi[\mathcal{V}]$  is an open subset of  $\mathbb{R}^n$ . But  $\mathbf{y} = \varphi(x)$ , and hence  $\mathbf{y} \in \varphi[\mathcal{V}]$ . But if  $\varphi[\mathcal{V}]$  is open and  $\mathbf{y} \in \varphi[\mathcal{V}]$ , then there is an  $r > 0$  such that the open ball  $B$  defined by:

$$B = \mathbb{B}_r^{(\mathbb{R}^n, \|\cdot\|_2)}(\mathbf{y}) \quad (37.19.13)$$

is contained in  $\varphi[\mathcal{V}]$ . But  $\varphi$  is continuous, and open balls are open, and hence  $\varphi^{-1}[B]$  is an open subset of  $\mathcal{V}$ . Moreover, since  $\mathbf{y} \in B$ ,  $x \in \varphi^{-1}[B]$ . But  $\varphi$  is an injective continuous open mapping, and thus the restriction of  $\varphi$  to an open subset is an injective continuous open mapping. Hence,  $\varphi|_{\varphi^{-1}[B]}$  is a homeomorphism onto its image, which is  $B$ . And since open balls in  $\mathbb{R}^n$  are homeomorphic to  $\mathbb{R}^n$ ,  $\varphi^{-1}[B]$  is homeomorphic to  $\mathbb{R}^n$ . Thus, there is an open subset  $\varphi^{-1}[B]$  containing  $x$  and an  $n \in \mathbb{N}$  such that  $\varphi^{-1}[B]$  is homeomorphic to  $\mathbb{R}^n$ .  $\square$

**Theorem 37.19.23.** *If  $(X, \tau)$  is a locally Euclidean topological space, then there exists a basis  $\mathcal{B}$  such that for all  $\mathcal{U} \in \mathcal{B}$  there is an  $n \in \mathbb{N}$  such that  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^n$ .*

*Proof.* For if  $(X, \tau)$  is locally Euclidean, then for all  $x \in X$  there is an open subset  $\mathcal{U}_x \in \tau$  and an  $n \in \mathbb{N}$  such that  $\mathcal{U}_x$  is homeomorphic to  $\mathbb{R}^n$  (Thm. 37.19.22). Let  $\varphi_x : \mathcal{U}_x \rightarrow \mathbb{R}^n$  be such a homeomorphism, and let  $\mathcal{B}_x$  be the set:

$$\mathcal{B}_x = \{ \varphi_x^{-1} [\mathbb{B}_r^{(\mathbb{R}^n, \|\cdot\|_2)}(\mathbf{y})] \mid r > 0, \mathbf{y} \in \mathbb{R}^n \} \quad (37.19.14)$$

Then by construction, every element of  $\mathcal{B}_x$  is homeomorphic to  $\mathbb{R}^n$ . Let  $\mathcal{B}$  be the collection of all such sets for all  $x \in X$ . If  $\mathcal{U}, \mathcal{V}$  are elements of  $\mathcal{B}$ , then there is an  $x \in X$  such that  $\mathcal{U} \in \mathcal{B}_x$ . Let  $y \in \mathcal{U} \cap \mathcal{V}$  and let  $\mathbf{y}$  be the image of  $y$  under  $\varphi_x$ . But  $\mathcal{U} \cap \mathcal{V}$  is open, and hence there is an  $r > 0$  such that the ball about  $\mathbf{y}$  is contained in the image of  $\varphi_x[\mathcal{U} \cap \mathcal{V}]$ . But this  $r$  ball is contained in  $\mathcal{B}$ . Hence,  $\mathcal{B}$  is a basis.  $\square$

**Theorem 37.19.24.** *If  $(X, \tau)$  is a locally Euclidean topological space, then there is a basis  $\mathcal{B}$  such that for all  $\mathcal{U} \in \mathcal{B}$  it is true that  $\mathcal{U}$  is precompact in  $\tau$  and such that there exists an  $n \in \mathbb{N}$  such that  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^n$ .*

*Proof.* For by Thm. 37.19.23, there is a basis  $\mathcal{B}$  of  $\tau$  such that for all  $\mathcal{U} \in \mathcal{B}$  there is an  $n \in \mathbb{N}$  such that  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^n$ . Let  $\varphi_{\mathcal{U}} : \mathcal{U} \rightarrow \mathbb{R}^n$  be such a homeomorphism. Define  $\mathcal{B}_x$  by:

$$\mathcal{B}_x = \{ \varphi_{\mathcal{U}}^{-1} [\mathbb{B}_r^{(\mathbb{R}^n, \|\cdot\|_2)}(\mathbf{y})] \mid r > 0, \mathbf{y} \in \mathbb{R}^n \} \quad (37.19.15)$$

But by the Heine-Borel theorem, for all  $\mathcal{V} \in \mathcal{B}_x$ ,  $\text{Cl}_{\mathbb{R}^n}(\varphi_{\mathcal{U}}[\mathcal{V}])$  is compact in  $\mathbb{R}^n$ , and since  $\varphi_{\mathcal{U}}$  is a homeomorphism,  $\text{Cl}_{\tau}(\mathcal{V})$  is compact in  $\mathcal{U}$ . But then  $\text{Cl}_{\tau}(\mathcal{V})$  is compact in  $X$ . The collection of all such  $\mathcal{B}_x$  is thus a basis of precompact subsets that are homeomorphic to open balls in  $\mathbb{R}^n$ , which are homeomorphic to  $\mathbb{R}^n$ .  $\square$

**Theorem 37.19.25.** *If  $(X, \tau)$  is locally Euclidean, then it is locally compact.*

*Proof.* For if  $(X, \tau)$  is locally Euclidean, then there exists a basis  $\mathcal{B}$  of precompact coordinate balls (Thm. 37.19.24). But then for all  $x \in X$  there is a  $\mathcal{U} \in \mathcal{B}$  such that  $x \in \mathcal{U}$  and  $\text{Cl}_{\tau}(\mathcal{U})$  is compact. Thus,  $(X, \tau)$  is locally compact (Def. 37.19.6).  $\square$

**Theorem 37.19.26.** *If  $(X, \tau)$  is locally Euclidean, then it is locally metrizable.*

*Proof.* For if  $(X, \tau)$  is locally Euclidean, then for all  $x \in X$  there is an  $n \in \mathbb{N}$  and a  $\mathcal{U} \in \tau$  such that  $x \in \mathcal{U}$  and  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^n$  (Thm. 37.19.22). But  $\mathbb{R}^n$  is metrizable, and thus  $\mathcal{U}$  is metrizable. Hence,  $X$  is locally metrizable.  $\square$

**Theorem 37.19.27.** *If  $(X, \tau)$  is locally Euclidean, then it is first countable.*

*Proof.* For if  $X$  is locally Euclidean, then it is locally metrizable (Thm. 37.19.26). But locally metrizable spaces are first countable (Thm. 37.19.6).  $\square$

**Theorem 37.19.28.** *If  $(X, \tau)$  is locally Euclidean and Hausdorff, then it is regular.*

*Proof.* For if  $(X, \tau)$  is locally Euclidean, then it is locally compact (Thm. 37.19.25). But locally compact Hausdorff spaces are regular (Thm. 37.19.14). Thus,  $X$  is regular.  $\square$

**Theorem 37.19.29.** *If  $(X, \tau)$  is locally Euclidean, and if  $C \subseteq X$  is compact, then it is sequentially compact.*

*Proof.* For locally Euclidean spaces are first countable (Thm. 37.19.27) and every subspace of a first countable space is first countable. But then  $C$  is a compact first countable space, and is therefore sequentially compact (Thm. 37.19.11).  $\square$

### Definition 37.19.10: Topological Manifold

A topological manifold is a locally Euclidean, Hausdorff, second countable topological space of constant dimension.

**Theorem 37.19.30.** *If  $(X, \tau)$  is a topological manifold, then it is Lindelöf.*

*Proof.* For if  $X$  is a topological manifold, then it is second countable (Def. 37.19.10). But second countable topological spaces are Lindelöf (Thm. 37.19.8). Thus,  $(X, \tau)$  is Lindelöf.  $\square$

**Theorem 37.19.31.** *If  $(X, \tau)$  is a topological manifold, then it is  $\sigma$  compact.*

*Proof.* For if  $X$  is a topological manifold, then it is Lindelöf (Thm. 37.19.30). But topological manifolds are locally Euclidean (Def. 37.19.9) and locally Euclidean spaces are locally compact (Thm. 37.19.25). But if  $(X, \tau)$  is locally compact and Lindelöf, then it is  $\sigma$  compact (Thm. 37.19.10).  $\square$

**Theorem 37.19.32.** *If  $(X, \tau)$  is a locally Euclidean Hausdorff topological space that is  $\sigma$  compact, then it is a topological manifold.*

*Proof.* For if  $(X, \tau)$  is locally Euclidean, then there is a basis  $\mathcal{B}$  of precompact coordinate balls (Thm. 37.19.24). But if  $X$  is  $\sigma$  compact, then it is Lindelöf (Thm. 37.19.9). But  $\mathcal{B}$  is an open cover of  $X$ , and hence there is a countable subcover  $\Delta$ . But then  $\Delta$  is a countable collection of open subspaces of  $X$ , each of which is homeomorphic to  $\mathbb{R}^n$  and hence second countable. Hence  $X$  is second countable (Thm. 37.19.19). But if  $X$  is locally Euclidean, Hausdorff, and second countable, then it is a topological manifold (Def. 37.19.10).  $\square$

**Theorem 37.19.33.** *If  $(X, \tau)$  is a locally Euclidean Hausdorff topological space that is Lindelöf, then  $X$  is a topological manifold.*

*Proof.* For if  $(X, \tau)$  is locally Euclidean, then it is locally compact (Thm. 37.19.25). But locally compact Lindelöf spaces are  $\sigma$  compact (Thm. 37.19.10). But locally Euclidean Hausdorff topological spaces that are  $\sigma$  compact are topological manifolds (Thm. 37.19.32). Therefore,  $(X, \tau)$  is a topological manifold.  $\square$

**Theorem 37.19.34.** *If  $(X, \tau)$  is a topological manifold, then it is paracompact.*

*Proof.* For if  $(X, \tau)$  is a topological manifold, then it is locally Euclidean and Hausdorff (Def. 37.19.10). But locally Euclidean Hausdorff spaces are regular (Thm. 37.19.28). But manifolds are second countable (Def. 37.19.10) and second countable spaces are Lindelöf (Thm. 37.19.8). But regular Lindelöf spaces are paracompact (Thm. 37.19.18). Hence,  $X$  is paracompact.  $\square$

### Definition 37.19.11: Topological Group

A topological group, denoted  $(G, *, \tau)$ , is a topological space  $(G, \tau)$  and a binary operation  $*$  such that  $(G, *)$  is a group, and such that  $g : G \times G \rightarrow G$  defined by  $g(a, b) = a * b$  is continuous with respect to the product topology on  $G$ , and such that  $\nu : G \rightarrow G$  defined by  $\nu(a) = a^{-1}$ , where  $a^{-1}$  is the inverse element of  $a$  under  $*$ , is continuous.

### Definition 37.19.12: Continuous Group Action

A continuous group action of a topological group  $(G, *, \tau)$  on a topological space  $(X, \tau_X)$  is a function  $\Theta : G \times X \rightarrow X$  such that for all  $x \in X$  and for all  $a, b \in G$ , the following are true:

$$\begin{aligned}\Theta(e, x) &= x \\ \Theta(a, \Theta(b, x)) &= \Theta(a * b, x)\end{aligned}$$

where  $e$  is the unital element of  $G$ , and such that  $\Theta$  is continuous with respect to the product topology.

**Theorem 37.19.35.** *If  $(X, \tau_X)$  is a Hausdorff topological space, if  $(G, *, \tau)$  is a compact Hausdorff topological group, if  $\Theta$  is a continuous group action of  $G$  on  $X$ , and if  $(X/G, \tau_q)$  is the quotient topology formed by the orbits of  $\Theta$ , then  $X/G$  is Hausdorff.*

*Proof.* Need to fill in.  $\square$

**Theorem 37.19.36.** *If  $(X, \tau_X)$  is a second countable topological space, if  $(G, *, \tau)$  is a compact Hausdorff topological group, if  $\Theta$  is a continuous group action of  $G$  on  $X$ , and if  $(X/G, \tau_q)$  is the quotient topology formed by the orbits of  $\Theta$ , then  $X/G$  is second countable.*

*Proof.* Also need to fill in.  $\square$

### 37.19.5 Problems: Part A

**Problem 37.19.1** (Lee Exercise 1.6) Show that  $\mathbb{R}\mathbb{P}^n$  is Hausdorff and second countable.

*Solution* For  $\mathbb{R}\mathbb{P}^n$  is homeomorphic to the quotient space of  $\mathbb{S}^n$  by the multiplicative group  $G = \{-1, 1\}$  with the group action  $\Theta : G \times \mathbb{S}^n \rightarrow \mathbb{S}^n$  defined by  $\Theta(n, \mathbf{s}) = n \cdot \mathbf{s}$  (this is well defined since  $n = \pm 1$ , and hence  $\|n \cdot \mathbf{s}\| = 1$  and thus  $n \cdot \mathbf{s}$  still lies in  $\mathbb{S}^n$ ). Equipping  $G$  with the discrete topology makes  $(G, *, \tau)$  a topological group, and  $\Theta$  a continuous group action on  $\mathbb{S}^n$ . But  $G$  is finite, and hence compact, and moreover it is Hausdorff since the discrete topology is always Hausdorff (it is metrizable, in fact). Thus  $\mathbb{S}^n/G$  is the quotient of a second countable Hausdorff space by a compact Hausdorff group, and is therefore Hausdorff (Thm. 37.19.35) and second countable (Thm. 37.19.36). Thus,  $\mathbb{R}\mathbb{P}^n$  is Hausdorff and second countable.

**Problem 37.19.2** (Lee Problem 1.5) Show that if  $(X, \tau)$  is a locally Euclidean connected Hausdorff space, then it is a manifold if and only if it is paracompact.

*Solution* For topological manifolds are paracompact (Thm. 37.19.34). Going the other way, if  $X$  is locally Euclidean, then there is a basis of precompact open subsets  $\mathcal{B}$ , each of which is homeomorphic to  $\mathbb{R}^n$  (Thm. 37.19.24). But  $X$  is paracompact, and hence there is a locally finite refinement  $\Delta$  of  $\mathcal{B}$  (Def. 37.19.8). Let  $\mathcal{U}_0$  be an element of  $\Delta$ . Since  $\Delta$  is a refinement of  $\mathcal{B}$ , there is an element  $\mathcal{V} \in \mathcal{B}$  such that  $\mathcal{U}_0 \subseteq \mathcal{V}$ . But then  $\text{Cl}_\tau(\mathcal{U}_0) \subseteq \text{Cl}_\tau(\mathcal{V})$ , and  $\mathcal{V}$  is precompact, and thus  $\text{Cl}_\tau(\mathcal{V})$  is compact. But then  $\text{Cl}_\tau(\mathcal{U}_0)$  is a closed subset of a compact set, and is hence compact. For all  $n \in \mathbb{N}$ , define  $\mathcal{U}_n$  by:

$$\mathcal{U}_n = \left\{ x \in X \mid \exists_{A: \mathbb{Z}_n \rightarrow \Delta} (\mathcal{U}_0 = A_0, x \in A_{n-1} \text{ and } \forall_{i < n-1} (A_i \cap A_{i+1} \neq \emptyset)) \right\} \quad (37.19.16)$$

That is, the set of all points that are separated from  $\mathcal{U}_0$  by at most  $n$  consecutive elements of  $\Delta$ . Then  $\mathcal{U}_n$  is precompact. We prove by induction. The base case of  $\mathcal{U}_0$  is true from the previous paragraph. Suppose it is true for  $n \in \mathbb{N}$ . Since  $X$  is locally Euclidean, if  $\mathcal{C} \subseteq X$  is compact, then it is sequentially compact (Thm. 37.19.29). But  $\mathcal{U}_{n+1}$  is the union of  $\text{Cl}_\tau(\mathcal{U}_n)$  and elements of  $\mathcal{V} \in \Delta$  such that  $\mathcal{V} \cap \mathcal{U}_n \neq \emptyset$ . But every element of  $\Delta$  is precompact, and since  $\text{Cl}_\tau(\mathcal{U}_{n+1})$  is not compact, there must be infinitely many such  $\mathcal{V}$ . But  $\mathcal{V} \cap \mathcal{U}_n$  is non empty for all such  $\mathcal{V}$ , and hence by the axiom of choice there is a sequence  $a : \mathbb{N} \rightarrow \mathcal{U}_n$  such that  $a_j$  lies in a distinct  $\mathcal{V}$  for all  $j \in \mathbb{N}$ . But  $\text{Cl}_\tau(\mathcal{U}_n)$  is compact, and hence sequentially compact, and thus there is a convergent subsequence  $a_k : \mathbb{N} \rightarrow \text{Cl}_\tau(\mathcal{U}_n)$  with a limit  $x$ . But  $x \in \text{Cl}_\tau(\mathcal{U}_n)$  and hence there is an open set  $V \in \tau$  that has non-empty intersection with only finitely many elements of  $\Delta$ , since  $\Delta$  is a locally finite refinement. But since  $a_k$  converges to an element of  $V$ , there is an  $N \in \mathbb{N}$  such that for all  $j > N$ ,  $a_{k_j} \in V$ . But each

$a_{k_j}$  lies in a different  $\mathcal{V} \in \Delta$ , and hence infinitely many elements of  $\Delta$  have non-empty intersection with  $V$ , a contradiction. Hence,  $\mathcal{U}_{n+1}$  is covered by finitely many elements of  $\Delta$  and is therefore precompact. Moreover,  $\bigcup \text{Cl}_\tau(\mathcal{U}_n) = X$ . For if  $y \in X$ , let  $x \in \mathcal{U}_0$  be any point. Then since  $X$  is locally path connected and connected, it is path connected (Thm. 37.19.21). Let  $\gamma : [0, 1] \rightarrow X$  be a path from  $x$  to  $y$ . But  $[0, 1]$  is compact and hence  $\gamma[[0, 1]]$  is a compact subset of  $X$ . But then it is covered by only finitely many elements of  $\Delta$ , and hence  $y$  is contained in one of the  $\mathcal{U}_n$ . Thus  $X$  is  $\sigma$  compact (Def. 37.19.4). But locally Euclidean  $\sigma$  compact Hausdorff spaces are topological manifolds (Thm. 37.19.32).

**Problem 37.19.3** Lee Problem 1-7.

*Solution* Let  $\mathbf{x} \in \mathbb{S}^n \setminus N$ , where  $N$  is the north pole. Let  $\Gamma$  be the line through  $N$  and  $\mathbf{x}$ . We can parameterize this as follows:

$$\Gamma(t) = tN + (1 - t)\mathbf{x} \quad (37.19.17)$$

To find when this lies on the  $\mathbb{R}^n$  hyperplane of  $\mathbb{R}^{n+1}$  defined by  $(x_1, \dots, x_n, 0)$  we simply need to find  $t$  such that  $\Gamma(t) = (x_1, \dots, x_n, 0)$ . Solving for the last coordinate, we get (Since  $N = (0, \dots, 0, 1)$ ):

$$t + (1 - t)x_{n+1} = 0 \implies t(1 - x_{n+1}) = -x_{n+1} \implies t = \frac{x_{n+1}}{x_{n+1} - 1} \quad (37.19.18)$$

evaluating  $\Gamma$  at this particular time obtains:

$$\Gamma\left(\frac{x_{n+1}}{x_{n+1} - 1}\right) = \frac{1}{1 - x_{n+1}}(x_1, \dots, x_n, 0) \quad (37.19.19)$$

in agreement with  $\sigma$ . For the south pole we repeat the argument, obtaining the curve  $\Lambda$  defined by:

$$\Lambda(t) = tS + (1 - t)\mathbf{x} \quad (37.19.20)$$

Solving for  $\Lambda(t) = (x_1, \dots, x_n, 0)$ , we look at the last component and set this equal to zero:

$$-t + (1 - t)x_{n+1} = 0 \implies -t(1 + x_{n+1}) = -x_{n+1} \implies t = \frac{x_{n+1}}{1 + x_{n+1}} \quad (37.19.21)$$

Evaluating this  $t$  into  $\Lambda$  gives:

$$\begin{aligned} \Lambda\left(\frac{x_{n+1}}{1 + x_{n+1}}\right) &= \left(0, \dots, 0, \frac{-x_{n+1}}{1 + x_{n+1}}\right) + \left(1 - \frac{x_{n+1}}{1 + x_{n+1}}\right)(x_1, \dots, x_n, x_{n+1}) \\ &= \left(0, \dots, 0, \frac{-x_{n+1}}{1 + x_{n+1}}\right) + \left(\frac{1}{1 + x_{n+1}}\right)(x_1, \dots, x_n, x_{n+1}) \\ &= \frac{(x_1, \dots, x_n, 0)}{1 + x_{n+1}} \end{aligned}$$

To show that  $\sigma$  is bijective, it suffices to show that  $\sigma \circ \sigma^{-1}$  and  $\sigma^{-1} \circ \sigma$  are the identity mappings, where:

$$\sigma^{-1}(X_1, \dots, X_n) = \frac{(2X_1, \dots, 2X_n, \|\mathbf{X}\|^2 - 1)}{\|\mathbf{X}\|^2 + 1} \quad (37.19.22)$$

But:

$$\begin{aligned} (\sigma \circ \sigma^{-1})(\mathbf{X}) &= \sigma\left(\frac{(2X_1, \dots, 2X_n, \|\mathbf{X}\|^2 - 1)}{\|\mathbf{X}\|^2 + 1}\right) \\ &= \frac{1}{1 - \frac{\|\mathbf{X}\|^2 - 1}{\|\mathbf{X}\|^2 + 1}} \left( \frac{2X_1}{\|\mathbf{X}\|^2 + 1}, \dots, \frac{2X_n}{\|\mathbf{X}\|^2 + 1} \right) \\ &= \frac{1}{\|\mathbf{X}^2\| + 1} \cdot \frac{1}{1 - \frac{\|\mathbf{X}\|^2 - 1}{\|\mathbf{X}\|^2 + 1}} (2X_1, \dots, 2X_n) \\ &= \frac{1}{\|\mathbf{X}\|^2 + 1 - \|\mathbf{X}\|^2 + 1} (2X_1, \dots, 2X_n) \\ &= \frac{1}{2} (2X_1, \dots, 2X_n) \\ &= (X_1, \dots, X_n) \\ &= \mathbf{X} \end{aligned}$$

and therefore  $\sigma \circ \sigma^{-1}$  is the identity. Going the other way, for  $\mathbf{x} \in \mathbb{S}^n \setminus \{N\}$  we have:

$$\begin{aligned} (\sigma^{-1} \circ \sigma)(\mathbf{x}) &= \sigma^{-1}\left(\frac{(x_1, \dots, x_n)}{1 - x_{n+1}}\right) \\ &= \frac{1}{\|\frac{(x_1, \dots, x_n)}{1 - x_{n+1}}\|^2 + 1} \left( \frac{2x_1}{1 - x_{n+1}}, \dots, \frac{2x_n}{1 - x_{n+1}}, \frac{\|(x_1, \dots, x_n)\|^2}{(1 - x_{n+1})^2} - 1 \right) \end{aligned}$$

But since  $\mathbf{x} \in \mathbb{S}^n \setminus \{N\}$ ,  $\|\mathbf{x}\| = 1$ . And:

$$\begin{aligned} \frac{1}{\|\frac{(x_1, \dots, x_n)}{1 - x_{n+1}}\|^2 + 1} &= \frac{(1 - x_{n+1})^2}{\|(x_1, \dots, x_n)\|^2 + (1 - x_{n+1})^2} \\ &= \frac{(1 - x_{n+1})^2}{\|\mathbf{x}\|^2 + 1 - 2x_{n+1}} \\ &= \frac{(1 - x_{n+1})^2}{1 + 1 - 2x_{n+1}} \\ &= \frac{(1 - x_{n+1})^2}{2(1 - x_{n+1})} \\ &= \frac{1 - x_{n+1}}{2} \end{aligned}$$

returning to the problem, we then get:

$$\begin{aligned}
 (\sigma^{-1} \circ \sigma)(\mathbf{x}) &= \frac{1-x_{n+1}}{2} \left( \frac{2x_1}{1-x_{n+1}}, \dots, \frac{2x_n}{1-x_{n+1}}, \frac{\|(x_1, \dots, x_n)\|^2}{(1-x_{n+1})^2} - 1 \right) \\
 &= \frac{1-x_{n+1}}{2} \left( \frac{2x_1}{1-x_{n+1}}, \dots, \frac{2x_n}{1-x_{n+1}}, \frac{1-x_{n+1}^2}{(1-x_{n+1})^2} - 1 \right) \\
 &= \frac{1-x_{n+1}}{2} \left( \frac{2x_1}{1-x_{n+1}}, \dots, \frac{2x_n}{1-x_{n+1}}, \frac{1+x_{n+1}}{1-x_{n+1}} - 1 \right) \\
 &= \frac{1-x_{n+1}}{2} \left( \frac{2x_1}{1-x_{n+1}}, \dots, \frac{2x_n}{1-x_{n+1}}, \frac{2x_{n+1}}{1-x_{n+1}} \right) \\
 &= (x_1, \dots, x_n, x_{n+1})
 \end{aligned}$$

and therefore,  $\sigma^{-1}\sigma$  is the identity. Thus,  $\sigma$  and  $\sigma^{-1}$  are bijections and inverses of each other. The transition map  $\tilde{\sigma} \circ \sigma^{-1}$ , where  $\tilde{\sigma}$  is the stereographic graphic projection from the south pole, can be computed as follows:

$$\begin{aligned}
 (\tilde{\sigma} \circ \sigma^{-1})(\mathbf{X}) &= \tilde{\sigma} \left( \frac{(2X_1, \dots, 2X_n, \|\mathbf{X}\|^2 - 1)}{\|\mathbf{X}\|^2 + 1} \right) \\
 &= \frac{1}{1 + \frac{\|\mathbf{X}\|^2 - 1}{\|\mathbf{X}\|^2 + 1}} \left( \frac{2X_1}{\|\mathbf{X}\|^2 + 1}, \dots, \frac{2X_n}{\|\mathbf{X}\|^2 + 1} \right) \\
 &= \frac{1}{\|\mathbf{X}\|^2} (X_1, \dots, X_n) \\
 &= \frac{\mathbf{X}}{\|\mathbf{X}\|^2}
 \end{aligned}$$

Since  $N$  and  $S$  are not included,  $\mathbf{0}$  is not in the domain of  $\tilde{\sigma} \circ \sigma^{-1}$ , and hence this is well defined everywhere and smooth. By symmetry,  $\sigma \circ \tilde{\sigma}^{-1}$  is smooth. Hence, these two charts constitute a smooth atlas on  $\mathbb{S}^n$  since they cover the set and overlap smoothly. Lastly, show that the standard smooth structure on  $\mathbb{S}^n$  and the stereographic one are compatible. For  $i \in \mathbb{Z}_n$ , let  $\varphi_i$  be the projection mapping of  $\mathcal{U}_i^+$  down to  $\mathbb{S}^n$ . We have:

$$\begin{aligned}
 (\varphi_i \circ \sigma^{-1})(\mathbf{X}) &= \varphi_i \left( \frac{(2X_1, \dots, 2X_n, \|\mathbf{X}\|^2 - 1)}{\|\mathbf{X}\|^2 + 1} \right) \\
 &= \frac{(2X_1, \dots, 2X_{i-1}, 2X_{i+1}, \dots, \|\mathbf{X}\|^2 - 1)}{\|\mathbf{X}\|^2 + 1}
 \end{aligned}$$

which is smooth. Similarly:

$$\begin{aligned}
 (\sigma \circ \varphi_i^{-1})(\mathbf{X}) &= \sigma \left( (x_1, \dots, x_{i-1}, \sqrt{1 - \|\mathbf{X}\|^2}, x_i, \dots, x_n) \right) \\
 &= \frac{(x_1, \dots, x_{i-1}, \sqrt{1 - \|\mathbf{X}\|^2}, x_{i+1}, \dots, x_{n-1})}{1 - x_n}
 \end{aligned}$$

which, since the domain of  $\sigma$  does not include  $x_n = 1$ , this is well defined and smooth. Hence, the standard smooth structure on  $\mathbb{S}^n$  is compatible with the stereographic one.

**Problem 37.19.4** (Lee Problem 1-9) Show that  $\mathbb{CP}^n$  is a smooth  $2n$  dimensional manifold.

*Solution* Given a  $\mathbf{z}$ , let  $\hat{\mathbf{z}} = \mathbf{z}/\|\mathbf{z}\|$ . Then  $[\mathbf{z}]$  is the same as the orbit of  $\hat{\mathbf{z}}$  by the rotation group  $U(1)$  acting on  $\mathbb{S}^{2n+1}$ , and similarly the quotient topologies are the same. But  $U(1)$  is a compact topological group. It is compact since it is a closed and bounded subset of Euclidean space, since all of the elements of  $U(1)$  have norm 1. But then  $\mathbb{S}^{2n+1}/U(1)$  is Hausdorff (Thm. 37.19.35) and second countable (Thm. 37.19.36). Moreover, since it is the quotient of a compact topological space ( $\mathbb{S}^{2n+1}$  is compact),  $\mathbb{CP}^n$  is compact as well. Lastly, we must show that  $\mathbb{CP}^n$  can be given a smooth structure compatible with this quotient topology. For all  $i \in \mathbb{Z}_{n+1}$ , let  $\mathcal{U}_i$  be the set of all  $\mathbf{z} \in \mathbb{C}^{n+1}$  such that the  $i^{th}$  component  $z_i$  is non-zero. Let  $\mathcal{V}_i = q[\mathcal{U}_i]$ , where  $q$  is the canonical projection of  $\mathbb{C}^{n+1}$  down to  $\mathbb{CP}^n$  sending  $\mathbf{z} \mapsto [\mathbf{z}]$ . Since  $\mathcal{U}_i$  is a saturated set,  $\mathcal{V}_i$  is therefore open and the restriction of  $q$  to  $\mathcal{U}_i$  is a quotient map. Let  $\varphi_i : \mathcal{V}_i \rightarrow \mathbb{C}^n$  be defined by:

$$\varphi_i[\mathbf{z}] = \left( \frac{z_1}{z_i}, \dots, \frac{z_{i-1}}{z_i}, \frac{z_{i+1}}{z_i}, \dots, \frac{z_{n+1}}{z_i} \right) \quad (37.19.23)$$

this is well defined, for if  $\mathbf{w} \in [\mathbf{z}]$ , then there is a  $c \in \mathbb{C}$  such that  $\mathbf{w} = c\mathbf{z}$ . But then:

$$\begin{aligned} \varphi_i[\mathbf{w}] &= \left( \frac{w_1}{w_i}, \dots, \frac{w_{i-1}}{w_i}, \frac{w_{i+1}}{w_i}, \dots, \frac{w_{n+1}}{w_i} \right) \\ &= \left( \frac{cz_1}{cz_i}, \dots, \frac{cz_{i-1}}{cz_i}, \frac{cz_{i+1}}{cz_i}, \dots, \frac{cz_{n+1}}{cz_i} \right) \\ &= \left( \frac{z_1}{z_i}, \dots, \frac{z_{i-1}}{z_i}, \frac{z_{i+1}}{z_i}, \dots, \frac{z_{n+1}}{z_i} \right) \\ &= \varphi_i[\mathbf{z}] \end{aligned}$$

Thus,  $\varphi_i$  is continuous. Moreover, it is a homeomorphism onto its image, with inverse:

$$\varphi_i^{-1}(z_1, \dots, z_n) = [z_1, \dots, z_{i-1}, 1, z_{i+1}, \dots, z_n] \quad (37.19.24)$$

which is continuous. Thus,  $\mathcal{V}_i$  is homeomorphic to an open subset of  $\mathbb{C}^n$ . But  $\mathbb{C}^n$  is homeomorphic to  $\mathbb{R}^{2n}$  (Indeed, set theoretically it simply *is*  $\mathbb{R}^{2n}$ ). This shows that  $\mathbb{CP}^n$  is locally Euclidean, and since it is also Hausdorff and second countable, it is therefore a topological manifold (Def. 37.19.10). Lastly, we must show that the  $\varphi_i$  overlap smoothly. But:

$$\begin{aligned} \varphi_i \circ \varphi_j^{-1}(z_1, \dots, z_n) &= \varphi_i([z_1, \dots, z_{j-1}, 1, z_{j+1}, \dots, z_n]) \\ &= \left( \frac{z_1}{z_i}, \dots, \frac{z_{j-1}}{z_i}, \frac{1}{z_i}, \frac{z_{j+1}}{z_i}, \dots, \frac{z_{i-1}}{z_i}, \frac{z_{i+1}}{z_i}, \dots, \frac{z_n}{z_i} \right) \end{aligned}$$

which is smooth.

**Problem 37.19.5** (Lee Problem 1-12) Prove that the product of smooth manifolds together with a manifold with boundary is a smooth manifold with boundary.

*Solution* We prove by induction on the number of manifolds without boundary. In the base case, let  $(M, \tau_M, \mathcal{A}_M)$  be a smooth manifold without boundary and  $(N, \tau_N, \mathcal{A}_N)$  a smooth manifold with boundary. Let  $(M \times N, \tau_M \times \tau_N)$  be the product topological space. Since  $M$  and  $N$  are second countable and Hausdorff,  $M \times N$  is. We must now show that it is locally Euclidean with boundary. For all  $q \in \text{Int}(N)$  there is an open subset  $\mathcal{V} \in \tau_N$  such that  $q \in \mathcal{V}$  and  $\mathcal{V}$  is homeomorphic to  $\mathbb{R}^n$ . For all  $p \in M$  there is an open subset  $\mathcal{U} \in \tau_M$  such that  $p \in \mathcal{U}$  and  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^m$ . But then  $(p, q) \in \mathcal{U} \times \mathcal{V}$ , and by the definition of the product topology,  $\mathcal{U} \times \mathcal{V} \in \tau_M \times \tau_N$ . Moreover, since  $\mathcal{U}$  is homeomorphic to  $\mathbb{R}^m$  and  $\mathcal{V}$  is homeomorphic to  $\mathbb{R}^n$ ,  $\mathcal{U} \times \mathcal{V}$  is homeomorphic to  $\mathbb{R}^m \times \mathbb{R}^n$ , which is homeomorphic to  $\mathbb{R}^{m+n}$ . Thus for all  $(p, q) \in M \times \text{Int}(N)$  there is an open subset containing  $(p, q)$  that is homeomorphic to Euclidean space. If  $q \in \partial N$ , then there is an open subset  $\mathcal{V}$  such that  $q \in \mathcal{V}$  and  $\mathcal{V}$  is homeomorphic to  $\mathbb{H}^n$ . But then  $\mathcal{U} \times \mathcal{V}$  is homeomorphic to  $\mathbb{R}^m \times \mathbb{H}^n$ , which is homeomorphic to  $\mathbb{H}^{m+n}$ . Thus, every point  $(p, q) \in M \times N$  is locally either  $\mathbb{R}^{m+n}$  or  $\mathbb{H}^{m+n}$ . That is,  $(M \times N, \tau_M \times \tau_N)$  is a topological manifold with boundary. Moreover, this shows that  $(p, q)$  lies on the boundary of  $M \times N$  if and only if  $q \in \partial N$ . Hence,  $\partial(M \times N) = M \times \partial N$ . We now how to show that this is a smooth manifold. Give  $p \in M$  and  $q \in \text{Int}(N)$ , there are smooth charts  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  such that  $p \in \mathcal{U}$ ,  $q \in \mathcal{V}$ ,  $\varphi : \mathcal{U} \rightarrow \mathbb{R}^m$  is a diffeomorphism, and  $\psi : \mathcal{V} \rightarrow \mathbb{R}^n$  is a diffeomorphism. But then the product function  $\varphi \times \psi$  is a diffeomorphism from  $\mathcal{U} \times \mathcal{V}$  to  $\mathbb{R}^m \times \mathbb{R}^n$ . If  $(U, \phi)$  and  $(V, \xi)$  are different charts, then  $\phi \times \xi$  overlaps smoothly since:

$$(\varphi \times \psi) \circ (\phi \times \xi)^{-1} = (\varphi \circ \phi^{-1}) \times (\psi \circ \xi^{-1}) \quad (37.19.25)$$

and since  $M$  and  $N$  are smooth manifolds,  $\varphi$  and  $\phi$  overlap smoothly, and similarly for  $\psi$  and  $\xi$ . But then this is the product of smooth functions, and hence smooth. Now if  $q \in \partial N$ , then there is a smooth boundary chart  $(\mathcal{V}, \psi)$  such that  $\psi : \mathcal{V} \rightarrow \mathbb{H}^n$  is a diffeomorphism. But the  $\varphi \times \psi$  is a diffeomorphism from  $\mathcal{U} \times \mathcal{V}$  to  $\mathbb{R}^m \times \mathbb{H}^n$ , which is diffeomorphic to  $\mathbb{H}^{m+n}$ . Therefore,  $M \times N$  is a smooth manifold with boundary. For the general case, if  $M_1, \dots, M_n$  are manifolds with boundary, and if  $N$  is a manifold with boundary, let  $M = \prod M_i$ . Then the product of the  $M_k$  with  $N$  is homeomorphic to  $M \times N$ , which reduces us to the previous problem. Thus,  $M_1 \times \dots \times M_n \times N$  is a smooth manifold and  $\partial(M \times N) = M \times \partial N$ .

### 37.19.6 Problems: Part B

**Problem 37.19.6** Show that the group action  $\Theta : \text{Aut}(V) \times V \rightarrow V$  given by  $\Theta(T, v) = Tv$  is continuous, where  $V$  is an  $n$  dimensional vector space over  $\mathbb{R}$ , and  $\text{Aut}(V)$  carries the subspace topology of  $\mathbb{R}^{n^2}$ .

*Solution* Both spaces are second countable, and hence first countable, and thus to check continuity it suffices to show that  $\Theta$  is sequentially continuous (Thm. 37.19.5). Let  $(T_n, v_n)$  be a sequence such that  $T_n \rightarrow T$  and  $v_n \rightarrow v$ . But then:

$$\|T_nv_n - Tv\| = \|T_nv_n - T_nv + T_nv - Tv\| \quad (37.19.26a)$$

$$\leq \|T_nv_n - T_nv\| + \|T_nv - Tv\| \quad (37.19.26b)$$

$$\leq \|T_n\| \|v_n - v\| + \|T_n - T\| \|v\| \quad (37.19.26c)$$

and since  $T_n \rightarrow T$  and  $v_n \rightarrow v$ , this converges to zero. Hence,  $T_nv_n \rightarrow Tv$ , and thus  $\Theta$  is sequentially continuous, and therefore continuous. Moreover, it is a group action:

$$\Theta(\text{id}_V, v) = \text{id}_V(v) = v \quad (37.19.27)$$

and for  $T, S \in \text{Aut}(V)$  we have:

$$\Theta(T, \theta(S, v)) = \Theta(T, Sv) = T(Sv) = (TS)v = \Theta(TS, v) \quad (37.19.28)$$

and thus  $\Theta$  is a continuous group action (Def. 37.19.12).

**Problem 37.19.7** Let  $G$  be a topological Lie group and  $V$  a finite dimensional real vector space with the usual topology. A real representation of  $G$  on  $V$  is a continuous group homeomorphism  $\rho : G \rightarrow \text{Aut}(V)$ . Any real representation  $(V, \rho)$  or  $G$  on  $V$  defines a group action  $\Theta : G \times V \rightarrow V$  given by  $\Theta(g, v) = \rho(g)v$ . Show that  $\rho$  is a representation if and only if  $\Theta$  is a continuous group action such that  $\Theta_g = \Theta(g, \cdot) \in \text{Aut}(V)$  for all  $g \in G$ .

*Proof.* Suppose  $\rho$  is a real representation of  $G$  on  $V$ . Since  $G$  and  $V$  are manifolds, so is  $G \times V$ . And manifolds are locally metrizable, and hence first countable (Thm. 37.19.6), and therefore sequential (Thm. 37.19.4). Thus, it suffices to show that  $\Theta$  is a group action and that it is sequentially continuous (Thm. 37.19.3). It is indeed a group action, for:

$$\Theta(e, v) = \rho(e)v = \text{id}_v(v) = v \quad (37.19.29)$$

since  $\rho$  is a group homeomorphism, and thus takes the identity of  $G$  to the identity of  $\text{Aut}(V)$ . Moreover, if  $g_1, g_2 \in G$ ,  $v \in V$ , then since  $\rho$  is a group

homeomorphism we obtain::

$$\Theta(g_1, \Theta(g_2, v)) = \Theta(g_1, \rho(g_2)v) \quad (37.19.30a)$$

$$= \rho(g_1)(\rho(g_2)v) \quad (37.19.30b)$$

$$= (\rho(g_1)\rho(g_2))v \quad (37.19.30c)$$

$$= \rho(g_1 * g_2)v \quad (37.19.30d)$$

and thus  $\Theta$  is a group action. It is continuous, for let  $(g_n, v_n) \rightarrow (g, v)$ . Then:

$$\Theta(g_n, v_n) = \rho(g_n)v_n \quad (37.19.31)$$

But  $\rho$  is continuous, and is therefore sequentially continuous (Thm. 37.19.1). Thus if  $g_n \rightarrow g$ , then  $\rho(g_n) \rightarrow \rho(g)$  (Def. 37.19.1). But then:

$$\|\Theta(g_n, v_n) - \Theta(g, v)\| = \|\rho(g_n)v_n - \rho(g)v\| \quad (37.19.32a)$$

$$= \|\rho(g_n)v_n - \rho(g)v_n + \rho(g)v_n - \rho(g)v\| \quad (37.19.32b)$$

$$\leq \|\rho(g_n)v_n - \rho(g)v_n\| + \|\rho(g)v_n - \rho(g)v\| \quad (37.19.32c)$$

$$\leq \|\rho(g_n) - \rho(g)\| \|v_n\| + \|\rho(g)\| \|v_n - v\| \quad (37.19.32d)$$

But  $v_n$  is a convergent sequence, and is therefore bounded, and hence  $\|v_n\|$  is bounded. Since  $v_n \rightarrow v$  and  $\rho(g_n) \rightarrow \rho(g)$ , this whole things tends to zero. Therefore,  $\Theta$  is sequentially continuous and therefore continuous. Moreover, if  $g \in G$ , then  $\Theta(g, \cdot) \in \text{Aut}(V)$ . For  $\Theta(g, \cdot) = \rho(g)(\cdot)$ , and  $\rho(g) \in \text{Aut}(V)$  by hypothesis. In the other direction, suppose  $\Theta$  is a continuous group action such that  $\Theta(g, \cdot) \in \text{Aut}(V)$ . Again, since  $G$  is a manifold, it suffices to show that  $\rho$  is sequentially continuous. If  $g_n \rightarrow g$ , then for all  $v \in V$ ,  $\Theta(g_n, v) = \rho(g_n)v \rightarrow \rho(g)v$  since  $\Theta$  is continuous and therefore  $\Theta(g_n, v) \rightarrow \Theta(g, v)$ . But then:

$$\|\rho(g_n)v - \rho(g)v\| \leq \|\rho(g_n) - \rho(g)\| \|v\| \quad (37.19.33)$$

and this tends to zero, showing that  $\rho(g_n) \rightarrow \rho(g)$ . Hence,  $\rho$  is continuous. Moreover, it is a group homeomorphism. For if  $e \in G$  is the identity, then for all  $v \in V$  it is true that:

$$\rho(e)v = \Theta(e, v) = v \quad (37.19.34)$$

since  $\Theta$  is a group action. Hence,  $\rho(e)$  is the identity mapping. If  $g_1, g_2 \in G$ , then:

$$\rho(g_1 * g_2)(\cdot) = \Theta(g_1 * g_2, \cdot) = \Theta(g_1, \Theta(g_2, \cdot)) = (\rho(g_1)\rho(g_2))(\cdot) \quad (37.19.35)$$

and thus  $\rho(g_1 * g_2) = \rho(g_1)\rho(g_2)$ . Therefore,  $\rho$  is a group homeomorphism.  $\square$

## 37.20 Homework II

**Problem 37.20.1** (Lee Problem 2-2) Prove that if  $M_1, \dots, M_k$  are smooth manifolds with or without boundary such that at most one of them has non-empty boundary, if  $\mathcal{M}$  is the product manifold, if  $\pi_i$  is the  $i^{th}$  projection mapping, then  $F : N \rightarrow \mathcal{M}$  is smooth if and only if  $\pi_i \circ F$  is smooth for all  $i \in \mathbb{Z}_k$ .

*Solution* First we show that  $\pi_i$  is smooth. Let  $\mathcal{M} = M_1 \times M_2$  and let  $\pi_1$  be the projection map  $(p, q) \mapsto p$ . Let  $(p, q) \in \mathcal{M}$ . Since  $M_1$  and  $M_2$  are manifolds, there are charts  $(\mathcal{U}_1, \varphi_1)$  and  $(\mathcal{U}_2, \varphi_2)$  such that  $p \in \mathcal{U}_1$  and  $q \in \mathcal{U}_2$  but by the definition of the product manifold,  $\mathcal{U}_1 \times \mathcal{U}_2$  is open in  $\mathcal{M}$  and  $(p, q) \in \mathcal{U}_1 \times \mathcal{U}_2$ . Moreover, the product function  $\varphi_1 \times \varphi_2$  is smooth. But then for all  $(x, y) \in \mathcal{U}_1 \times \mathcal{U}_2$  we have:

$$\varphi_1 \circ \pi_1 \circ (\varphi_1 \times \varphi_2)^{-1}(x, y) = \varphi_1 \circ (\pi_1 \circ (\varphi_1 \times \varphi_2)^{-1}(x, y)) \quad (37.20.1)$$

$$= \varphi_1 \circ (\pi_1(\varphi_1^{-1}(x), \varphi_2^{-1}(y))) \quad (37.20.2)$$

$$= \varphi_1 \circ (\varphi_1^{-1}(x)) \quad (37.20.3)$$

$$= x \quad (37.20.4)$$

And thus we have the projection map from  $\mathbb{R}^{d_1+d_2}$  to  $\mathbb{R}^{d_1}$ , which is smooth. Hence for all  $(p, q) \in \mathcal{M}$  there is a chart  $(\mathcal{U}, \varphi)$  containing  $(p, q)$  and a chart  $(\mathcal{V}, \psi)$  containing  $\pi_1(p, q)$  such that  $\psi \circ \pi_1 \circ \varphi^{-1}$  is smooth, and thus  $\pi_1$  is smooth. Similarly,  $\pi_2$  is smooth. For the general case we proceed by induction and write:

$$\mathcal{M} = \prod_{k=1}^{n+1} M_k = \left( \prod_{k=1}^n M_k \right) \times M_{n+1} \equiv \widehat{\mathcal{M}} \times M_{n+1} \quad (37.20.5)$$

which is the product of two manifolds, one of dimension  $d_1 + d_2 + \dots + d_n$  and the other of dimension  $d_{n+1}$ , and thus by the previous argument the projection maps are smooth. Hence,  $\pi_{n+1}$  is smooth. But by the induction hypothesis, all of the  $\pi_i$  are smooth for  $i = 1, \dots, n$ , and thus all  $\pi_i$  are smooth for  $i = 1, \dots, n, n+1$ . Now, suppose  $f : N \rightarrow \mathcal{M}$  is smooth. Then  $F_i = \pi_i \circ f$  is the composition of smooth function and hence by theorem 2.10 (d) in Lee's text,  $F_i$  is smooth. In the other direction, we again start with the case that  $\mathcal{M} = M_1 \times M_2$ . Suppose  $F : N \rightarrow \mathcal{M}$  is such that  $\pi_1 \circ F$  and  $\pi_2 \circ F$  are smooth. Let  $v \in N$  and let  $(\mathcal{V}, \psi)$  be a chart containing  $v$  and let  $(\widehat{\mathcal{U}}, \widehat{\varphi})$  be a chart containing  $F(v)$ . Let  $\mathcal{U}_i = \pi_i[\widehat{\mathcal{U}}]$  and let  $\varphi_i = \pi_i \circ \widehat{\varphi}$ . Then the  $\varphi_i$  are the composition of smooth functions, and hence are smooth, and the  $\mathcal{U}_i$  are the projections of open sets and are hence open. Moreover,  $(\mathcal{U}_1 \times \mathcal{U}_2, \varphi_1 \times \varphi_2)$  is a chart containing  $F(v)$ . But

$$\varphi_1 \times \varphi_2 \circ (F \circ \psi^{-1})(x) = (\varphi_1 \circ \pi_1 \circ F \circ \psi^{-1}), \varphi_2 \circ \pi_2(F) \circ \psi^{-1}) \quad (37.20.6)$$

and by hypothesis,  $F \circ \pi_i$  is smooth. But then this is the composition of smooth functions in each component, and is therefore smooth in the each component. But a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is smooth if and only if it is smooth in each component, and thus this transition map is smooth. Thus,  $F$  is smooth.

**Problem 37.20.2** (Lee 2-3) Show that  $z^n : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  is smooth, the antipodal map  $\mathbf{x} \mapsto -\mathbf{x}$ , and the function  $f : \mathbb{S}^3 \rightarrow \mathbb{S}^2$  defined by:

$$f(w, z) = (z\bar{w} + w\bar{z}, iw\bar{z} - iz\bar{w}, z\bar{z} - w\bar{w}) \quad (37.20.7)$$

*Solution* Since  $\mathbb{S}^1$  can be covered by two charts, we simply need to check that this function is smooth on these charts. Let  $\mathcal{U}^- = \mathbb{S}^1 \setminus \{(0, 1)\}$  and similarly for  $\mathcal{U}^+$ . The stereographic projection onto  $\mathbb{R}$  is just:

$$\varphi_-(x, y) = \frac{x}{1-y} \quad (37.20.8a) \qquad \varphi_+(x, y) = \frac{x}{1+y} \quad (37.20.8b)$$

The inverse functions are:

$$\varphi_-^{-1}(X) = \left( \frac{X}{1+X^2}, \frac{X^2-1}{X^2+1} \right) \quad (37.20.9)$$

and similarly for  $\varphi_+$ . To show that  $p_n$  is smooth it suffices to show that  $\varphi_- \circ p_n \circ \varphi_-^{-1}$  is smooth. We have:

$$(\varphi_- \circ p_n \circ \varphi_-^{-1})(X) = (\varphi_- \circ p_n) \left( \frac{2X + i(X^2 - 1)}{1 + X^2} \right) \quad (37.20.10)$$

$$= \varphi_- \left( \frac{(2X + i(X^2 - 1)^2)^n}{(1 + X^2)^n} \right) \quad (37.20.11)$$

Thus, the  $x$  and  $y$  components will both be rational functions in  $X$ , and therefore  $\varphi_{-1}(x, y)$  will be a rational function in  $X$ , which is smooth. In a similarly manner,  $\varphi_+ \circ p_n \circ \varphi_+^{-1}$  is smooth. Thus for every point  $\mathbf{x} \in \mathbb{S}^1$  there is a chart  $(\mathcal{U}, \varphi)$  such that  $\mathbf{x} \in \mathcal{U}$  and  $\varphi \circ p_n \circ \varphi^{-1}$  is smooth. Therefore,  $p_n$  is smooth. For the antipodal map on  $\mathbb{S}^n$  we can use the orthographic projections. That is,  $(\mathcal{U}_j^+, \varphi_+)$  is the chart where  $\mathcal{U}_j^+$  is the  $j^{th}$  upper hemisphere and  $\varphi$  is the mapping that projects  $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$  down to the  $\mathbb{R}^n$  hyper plane obtained by fixed all but the  $j^{th}$  coordinate, and similarly let  $(\mathcal{U}^-, \varphi_-)$  be the opposite hemisphere. Given  $\mathbf{s} \in \mathbb{S}^n$ ,  $\mathbf{s}$  is contained in one of these, and  $-\mathbf{s}$  will be contained in the opposite. We thus need to show that, if  $f$  is the antipodal map, then  $\varphi_- \circ f \circ \varphi_+^{-1}$  is smooth. We have:

$$\varphi_- \circ f \circ \varphi_+^{-1}(\mathbf{x}) = \varphi_- \circ f(x_0, \dots, x_{j-1}, \sqrt{1 - \|\mathbf{x}\|^2}, x_{j+1}, \dots, x_n) \quad (37.20.12)$$

$$= \varphi_-(-x_0, \dots, -x_{j-1}, -\sqrt{1 - \|\mathbf{x}\|^2}, -x_{j+1}, \dots, -x_n) \quad (37.20.13)$$

$$= (-x_0, \dots, -x_{j-1}, -x_{j+1}, \dots, -x_n) \quad (37.20.14)$$

$$= -\mathbf{x} \quad (37.20.15)$$

Hence, this is a smooth mapping since multiplying by a constant is a smooth mapping from  $\mathbb{R}^n$  to itself. Lastly, we show that the *Hopf Fibration* is smooth. Again, sticking to orthographic projections, let  $(\mathcal{U}, \varphi)$  and  $(\mathcal{V}, \psi)$  be orthographic charts in  $\mathbb{S}^3$  and  $\mathbb{S}^2$ , respectively (Suppose both in the  $x$  axis). Let  $f : \mathbb{S}^3 \rightarrow \mathbb{S}^2$  be the Hopf fibration. Then for  $\mathbf{x} \in \mathbb{B}^3$ , we have:

$$\psi \circ f \circ \varphi^{-1}(\mathbf{x}) = \psi \circ f \circ (\sqrt{1 - \|\mathbf{x}\|^2}, \mathbf{x}) = \psi \circ f \left( (\sqrt{1 - \|\mathbf{x}\|^2} + ix_1), (x_2 + ix_3) \right) \quad (37.20.16)$$

Using the definition of  $f$  and  $\psi$  (which simply projection down to the  $yz$  plane), we have:

$$\psi \circ f \circ \varphi^{-1}(\mathbf{x}) = (2\sqrt{1 - \|\mathbf{x}\|^2}x_1 + 2x_2x_3, 1 - x_1^2) \quad (37.20.17)$$

which is smooth in both components, and hence is a smooth function from an open subset of  $\mathbb{R}^3$  to an open subset of  $\mathbb{S}^2$  and is therefore smooth. Similarly for the other hemispheres.

**Problem 37.20.3** (Lee 2-6) Show that if  $f : \mathbb{R}^{n+1} \setminus \{0\} \rightarrow \mathbb{R}^{k+1} \setminus \{0\}$  is a smooth homogeneous function of degree  $d \in \mathbb{Z}$ , then the induced maps  $\tilde{f} : \mathbb{R}P^n \rightarrow \mathbb{R}P^k$  are well defined and smooth.

*Solution* It is indeed well defined, for if  $\mathbf{x}, \mathbf{y}$  have the same equivalence class:  $[\mathbf{x}] = [\mathbf{y}]$ , then there is a  $\lambda \in \mathbb{R} \setminus \{0\}$  such that  $\mathbf{y} = \lambda\mathbf{x}$ . But then:

$$\tilde{f}([\mathbf{y}]) = [f(\mathbf{y})] = [f(\lambda\mathbf{x})] = [\lambda^d f(\mathbf{x})] = [f(\mathbf{x})] = \tilde{f}([\mathbf{x}]) \quad (37.20.18)$$

and therefore elements of the same equivalence class map to the same points.

**Problem 37.20.4** Lee 2-10.

*Solution* Let  $(M, \tau_M, \mathcal{A}_M)$  and  $(N, \tau_N, \mathcal{A}_N)$  be manifolds,  $F : M \rightarrow N$  continuous.  $F^*$  is linear. For if  $a, b \in \mathbb{R}$ ,  $f, g \in C(M, \mathbb{R})$ , then:

$$F^*(af + bg) = (af + bg) \circ F \quad (37.20.19a)$$

$$= ((af) \circ F) + ((bg) \circ F) \quad (37.20.19b)$$

$$= a(f \circ F) + b(g \circ F) \quad (37.20.19c)$$

$$= aF^*(f) + bF^*(g) \quad (37.20.19d)$$

If  $F : M \rightarrow N$  is smooth, then for all  $f \in C^\infty(N, \mathbb{R})$  we have  $F^*(f) = f \circ F$ , which is the composition of smooth functions and is hence smooth. Therefore, we obtain  $F^*[C^\infty(N, \mathbb{R})] \subseteq C^\infty(M, \mathbb{R})$ . Conversely, suppose  $F^*[C^\infty(N, \mathbb{R})] \subseteq C^\infty(M, \mathbb{R})$  and suppose that  $F$  is not smooth. Then there is a point  $p \in M$  such that for every chart  $(\mathcal{U}, \varphi) \in \mathcal{A}_M$  that contains  $p$  and for every chart  $(\mathcal{V}, \psi) \in \mathcal{A}_N$  that contains  $F(p)$ , the function  $\psi \circ F \circ \varphi^{-1}$  is not smooth. But since  $N$  is a manifold, there is a precompact chart  $(\mathcal{V}, \psi)$  that contains  $F(p)$ . But  $F$

is continuous, and hence  $F^{-1}[\mathcal{V}]$  is an open subset of  $M$  that contains  $p$ . Let  $(\tilde{\mathcal{U}}, \tilde{\varphi})$  be a precompact chart containing  $p$ , and let  $\mathcal{U} = \tilde{\mathcal{U}} \cap F^{-1}[\mathcal{V}]$  and  $\varphi = \tilde{\varphi}|_{\mathcal{U}}$ . Then, since  $\mathcal{U}$  is the non-empty intersection of two open subsets, it will open and non-empty, and hence  $(\mathcal{U}, \varphi)$  is a chart in  $M$ . Since  $\mathcal{A}_M$  is maximal,  $(\mathcal{U}, \varphi) \in \mathcal{A}_M$ . By hypothesis,  $\psi \circ F \circ \varphi^{-1}$  is not smooth, and hence there is a component  $k$  such that composing with the projection map  $\pi_k : \mathbb{R}^m \rightarrow \mathbb{R}$  is not smooth. But we may use the bump function to extend  $\pi_k \circ \psi$  to all of  $N$ , obtaining a smooth function in  $C^\infty(N, \mathbb{R})$ . But then by hypothesis,  $\pi_k \circ \psi \circ F$  is smooth. And  $\varphi^{-1}$  is smooth, so  $\pi_k \circ \psi \circ F \circ \varphi^{-1}$  is smooth, a contradiction. Hence,  $F$  is smooth. Lastly, show that if  $F$  is a homeomorphism, then it is a diffeomorphism if and only if  $F^*$  is an isomorphism. Since  $F$  and  $F^{-1}$  are continuous, they are both smooth if and only if  $F^*[C^\infty(N, \mathbb{R})] \subseteq C^\infty(M, \mathbb{R})$  and  $F^{-1*}[C^\infty(M, \mathbb{R})] \subseteq C^\infty(N, \mathbb{R})$ . That is, they are smooth if and only if  $F^*|_{C^\infty(N, \mathbb{R})} : C^\infty(N, \mathbb{R}) \rightarrow C^\infty(M, \mathbb{R})$  is a bijective function. Since  $F^*$  is linear,  $F$  and  $F^{-1}$  are smooth if and only if  $F^*$  is an isomorphism.

**Problem 37.20.5** (Lee 2-13) Show that paracompactness and subordinate partitions of unity are equivalent.

*Solution* That paracompactness implies subordinate partitions of unity was proved both in class and in Lee. Going the other way, let  $(X, \tau)$  be a topological space and suppose every open cover has a subordinate partition of unity. Let  $\mathcal{O}$  be an open cover and let  $\mathcal{F}$  be a subordinate partition of unity. Let  $\Delta$  be defined by:

$$\Delta = \{\mathcal{U} \mid \exists_{f \in \mathcal{F}} (\mathcal{U} = f^{-1}[\mathbb{R} \setminus \{0\}])\} \quad (37.20.20)$$

Then, since all  $f$  are continuous, all of the  $\mathcal{U}$  are the pre-images of open sets and are therefore open. Since  $\mathcal{F}$  is a partition of unity, for every point  $x \in X$  there is an  $f \in \mathcal{F}$  such that  $f(x) \neq 0$ , hence  $\Delta$  is an open subcover of  $X$ . Moreover, every  $\mathcal{U}$  is contained in the support of some function, and since  $\mathcal{F}$  is a partition of unity, these form a locally finite cover. Therefore  $\Delta$  is a locally finite refinement of  $\mathcal{O}$ . Thus,  $(X, \tau)$  is paracompact.

**Problem 37.20.6** (Lee 2-14) Show that disjoint closed subsets can be separated by smooth functions.

*Solution* For manifolds are Hausdorff and second countable (Def. 37.19.10), and are regular (Thm. 37.19.28) and Lindelöf (Thm. 37.19.30). But regular Lindelöf spaces are normal (Thm. 37.19.15), and thus if  $C_1$  and  $C_2$  are disjoint closed sets then there are disjoint open sets  $\mathcal{U}_1, \mathcal{U}_2$  containing  $C_1$  and  $C_2$ . Furthermore, we can shrink  $\mathcal{U}_1$  and  $\mathcal{U}_2$  so that they have disjoint closures. From theorem 2.29 we can find functions  $f, g : M \rightarrow \mathbb{R}$  such that  $f^{-1}[\{0\}] = C_1$  and  $g^{-1}[\{1\}] = C_2$ . But since  $\text{Cl}_\tau(\mathcal{U}_1)$  is closed, we can find a bump function for  $\text{Cl}_\tau(\mathcal{U}_1)$  supported on  $\mathcal{U}$  that evaluates to 1 on  $C_1$ , and similarly for  $C_2$ . Denote these bump functions  $h_1, h_2$ . Let  $F : M \rightarrow \mathbb{R}$  be defined by:

$$F(p) = g_1(p)f(p) + g_2(p)g(p) \quad (37.20.21)$$

## 37.21 Homework III

### 37.21.1 Weierstrass Approximation Theorem

**Theorem 37.21.1:** Weierstrass Approximation Theorem

If  $f : [0, 1] \rightarrow \mathbb{R}$  is a continuous function and if  $\varepsilon > 0$ , then there is a polynomial  $P : [0, 1] \rightarrow \mathbb{R}$  such that  $\|P - f\|_\infty < \infty$ . That is:

$$\sup \left\{ |f(x) - P(x)| \mid x \in [0, 1] \right\} < \varepsilon$$

**Theorem 37.21.2.** If  $f : [0, 1] \rightarrow \mathbb{R}$  is a continuous function, if  $f(0) = 0$ ,  $f(1) = 0$ , and if  $\varepsilon > 0$ , then there is a polynomial  $P : [0, 1] \rightarrow \mathbb{R}$  such that  $P(0) = f(0)$ ,  $P(1) = f(1)$ , and  $\|P - f\|_\infty < \varepsilon$ .

*Proof.* For by the Weierstrass approximation theorem there exists a polynomial  $Q : [0, 1] \rightarrow \mathbb{R}$  such that  $\|f - Q\| < \varepsilon/4$  (Thm. 37.21.1). Define  $P$  by:

$$P(x) = Q(x) - Q(0) - (Q(1) - Q(0))x \quad (37.21.1)$$

Then:

$$\|P - f\|_\infty \leq \|Q - f\|_\infty + 2|Q(0)| + |Q(1)| < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon \quad (37.21.2)$$

and moreover,  $P(0) = 0$  and  $P(1) = 0$ .  $\square$

**Theorem 37.21.3.** If  $f : [0, 1] \rightarrow \mathbb{R}$  is a continuous function, and if  $\varepsilon > 0$ , then there is a polynomial  $P : [0, 1] \rightarrow \mathbb{R}$  such that  $P(0) = f(0)$ ,  $P(1) = f(1)$ , and  $\|P - f\|_\infty < \varepsilon$ .

*Proof.* For let  $g(x) = f(x) - f(b)x - (1-x)f(a)$ . Then  $g$  is continuous,  $g(0) = 0$ ,  $g(1) = 0$ , and thus there is a polynomial  $Q : [0, 1] \rightarrow \mathbb{R}$  such that  $Q(0) = g(0)$ ,  $Q(1) = f(1)$ , and  $\|Q - g\|_\infty < \varepsilon$  (Thm. 37.21.2). Let  $P = Q + f(b)x + (1-x)f(a)$ . Then  $P(0) = f(0)$ ,  $P(1) = f(1)$ , and  $\|P - f\|_\infty < \varepsilon$ .  $\square$

**Theorem 37.21.4.** If  $f : [0, 1] \rightarrow \mathbb{R}^n$  is a smooth function, and if  $\varepsilon > 0$ , then there is a smooth function  $\Gamma : [0, 1] \rightarrow \mathbb{R}^n$  such that  $\Gamma(0) = f(0)$ ,  $\Gamma(1) = f(1)$ , and  $\|\Gamma - f\|_\infty < \varepsilon$ .

*Proof.* For let  $g_k = f \circ \pi_k$ , where  $\pi_k : \mathbb{R}^n \rightarrow \mathbb{R}$  is the projection mapping. But the projection of a continuous function is continuous, and thus  $g_k : [0, 1] \rightarrow \mathbb{R}^n$  is continuous. By the Weierstrass approximation theorem, there is a polynomial  $P_k : [0, 1] \rightarrow \mathbb{R}$  such that  $\|g_k - P_k\|_\infty < \varepsilon/n$ , and such that  $P_k(0) = g_k(0)$  and  $P_k(1) = g_k(1)$  (Thm. 37.21.4). Let  $\Gamma : [0, 1] \rightarrow \mathbb{R}^n$  be the function such that

$\Gamma \circ \pi_k = P_k$ . Since polynomials are smooth, and since  $\Gamma$  is smooth in every projection, it is smooth. But then:

$$\|f - \Gamma\|_\infty \leq \sum_{k \in \mathbb{Z}_n} \|g_k - P_k\|_\infty < \sum_{k \in \mathbb{Z}_n} \frac{\varepsilon}{n} = \varepsilon \quad (37.21.3)$$

completing the proof.  $\square$

**Theorem 37.21.5.** *If  $(M, \tau, \mathcal{A})$  is a smooth connected manifold, and if  $x, y \in M$ , then there is a smooth function  $\Gamma : [0, 1] \rightarrow M$  such that  $\Gamma(0) = x$  and  $\Gamma(1) = y$ .*

*Proof.* For suppose not. Then there are points  $x, y \in M$  with no smooth path between them. But manifolds are locally path connected, and locally path connected topological spaces that are connected are also path connected (Thm. 37.19.21). Hence, there is a path  $\gamma : [0, 1] \rightarrow M$  such that  $\gamma(0) = x$  and  $\gamma(1) = y$ . Let  $A \subseteq [0, 1]$  be the set of elements  $t$  such that there is a smooth curve from  $\gamma(0)$  to  $\gamma(t)$ . Then  $A$  is non-empty since  $0 \in A$ . That is, the constant function is a smooth function from  $\gamma(0)$  to  $\gamma(0)$ . Moreover  $A$  is bounded by 1, and hence there is a least upper bound  $r$ . Suppose  $r < 1$ . Since  $M$  is a manifold, there is a chart  $(U, \varphi)$  such that  $\gamma(r) \in U$  and  $U$  is homeomorphic to  $\mathbb{R}^n$ . Since  $U$  is open,  $\gamma^{-1}[U]$  is open. Hence there exists  $t_0 < r < t_1$  such that  $\gamma[[t_0, t_1]] \subseteq U$ . But then  $\gamma|_{[t_0, t_1]} \circ \varphi$  is the composition of continuous function, and is thus a continuous function from  $[t_0, t_1]$  to  $\mathbb{R}^n$ . And since  $[t_0, t_1]$  is compact, its image is compact. Hence there is an  $\varepsilon > 0$  such that the image sits inside that  $\varepsilon$  ball about the origin. But then there is a smooth function  $P$  such that  $P(t_0) = \varphi(\gamma(t_0))$ ,  $P(t_1) = \varphi(\gamma(t_1))$ , and  $\|P - \varphi\| \circ \gamma|_{[t_0, t_1]} < \varepsilon/2$  and hence the image of  $P$  sits inside the  $\varepsilon$  ball. But then the pull back of  $P$  by  $\varphi$  is a smooth curve connecting  $\gamma(t_0)$  and  $\gamma(t_1)$ . But  $t_0 < r$  and hence there is a smooth curve  $\Gamma_0 : [0, 1] \rightarrow M$  connecting  $\gamma(0)$  and  $\gamma(t_0)$ . Let  $\Gamma_1$  be the concatenation of  $\Gamma_0$  and the pull back of  $P$  by  $\varphi$ , smoothed out at  $\gamma(t_0)$  by any smoothing function ( $\exp(-1/t^2)$  will do). But then  $\Gamma$  is a smooth curve connecting  $\gamma(0)$  and  $\gamma(t_1)$ , a contradiction since  $r < t_1$  and  $r$  is the least upper bound of such  $t$ .  $\square$

## 37.21.2 Homework III

**Problem 37.21.1** Lee 3-1.

*Solution* It suffices to look at a single connected component of  $M$ . Suppose  $F$  is not constant. Then there are  $x, y \in M$  such that  $F(x) \neq F(y)$ . But since  $(M, \tau, \mathcal{A})$  is a smooth connected manifold, there exists a smooth path  $\Gamma : [0, 1] \rightarrow M$  such that  $\Gamma(0) = x$  and  $\Gamma(1) = y$  (Thm. 37.21.5). But for all  $p \in M$  it is true that  $d_p F$  is the zero map, and thus for all  $t \in [0, 1]$ ,  $d_{\Gamma(t)} F = 0$ . Let  $\gamma : [0, 1] \rightarrow N$  be defined by  $\gamma = F \circ \Gamma$ . Then since  $\gamma$  is the composition of

smooth functions, it is smooth, and moreover  $\gamma(0) = F(x)$  and  $\gamma(1) = F(y)$ . But for all  $t \in [0, 1]$ ,  $\dot{\gamma}(t) = 0$ . Thus, for all  $f \in C^\infty(N, \mathbb{R})$  it is true that  $\dot{\gamma}(t)(f) = 0$ . But  $f \circ \gamma$  is a function from  $[0, 1]$  to  $\mathbb{R}^n$ , and if  $\dot{\gamma}(t)(f) = 0$  then  $f \circ \gamma$  is constant. Then in particular, for any chart  $(\mathcal{V}, \psi)$  in  $N$  that overlaps with  $\gamma[[0, 1]]$ , the projection maps  $\pi_k : \mathcal{V} \rightarrow \mathbb{R}$  are such that  $\gamma$  is constant on  $\mathcal{V}$ . But then  $\gamma$  is a locally constant function and since  $\gamma[[0, 1]]$  is connected,  $\gamma$  is thus constant. But then  $\gamma(0) = \gamma(1)$  and hence  $F(x) = F(y)$ , a contradiction. In the other direction, if  $F$  is constant, let  $p \in M$  and  $v \in T_p M$ . Then for all  $f \in C^\infty(N, \mathbb{R})$  we have:

$$d_p F(v)(f) = v(f \circ F) \quad (37.21.4)$$

But  $f \circ F$  is a constant function, and a derivation of a constant is zero since:

$$v(1) = v(1 \cdot 1) = v(1) \cdot 1 + 1 \cdot v(1) = v(1) + v(1) \quad (37.21.5)$$

and hence  $v(1) = 0$ . Then for any  $c$  such that  $v(c) = cv(1) = 0$ . But then  $d_p F(v)(f) = 0$  for all  $f \in C^\infty(N, \mathbb{R})$  and thus  $d_p F$  is the zero mapping for all  $p \in M$ .

**Problem 37.21.2** If  $M_1, \dots, M_n$  are smooth manifolds, if  $\mathcal{M} = \prod M_k$ , and if  $\alpha : T_p \mathcal{M} \rightarrow \bigoplus T_{p_k} M_k$  is defined by;

$$\alpha(v) = (d\pi_{1p}(v), \dots, d\pi_{np}(v)) \quad (37.21.6)$$

then  $\alpha$  is an isomorphism. The same is true if one of the  $M_k$  has boundary.

*Solution* We must show that  $\alpha$  is linear and bijective. It is linear since:

$$\alpha(av + bw) = (d\pi_{1p}(av + bw), \dots, d\pi_{np}(av + bw)) \quad (37.21.7)$$

$$= (a d\pi_{1p}(v) + b d\pi_{1p}(w), \dots, a d\pi_{np}(v) + b d\pi_{np}(w)) \quad (37.21.8)$$

$$= (a d\pi_{1p}(v), \dots, a d\pi_{np}(v)) + (b d\pi_{1p}(w), \dots, b d\pi_{np}(w)) \quad (37.21.9)$$

$$= a(d\pi_{1p}(v), \dots, d\pi_{np}(v)) + b(d\pi_{1p}(w), \dots, d\pi_{np}(w)) \quad (37.21.10)$$

$$= a\alpha(v) + b\alpha(w) \quad (37.21.11)$$

and thus,  $\alpha$  is linear. It is injective, for if  $v, w \in T_p \mathcal{M}$  and if  $\alpha(v) = \alpha(w)$ , then the projections of  $v$  and  $w$  agree in every component, and hence  $v = w$ . Lastly, it is surjective. For if  $v_k \in T_{p_k} M_k$ , let  $v = v_1 \times \dots \times v_n$ . Then the  $k^{th}$  component of  $\alpha(v)$  is:

$$d\pi_{kp}(v)(f) = v(f \circ \pi_k) = v_k(f) \quad (37.21.12)$$

and thus  $d\pi_{kp}(v) = v_k$  so  $\alpha$  is surjective.

**Problem 37.21.3** Show that  $T\mathbb{S}^1$  is diffeomorphic to  $\mathbb{S}^1 \times \mathbb{R}$ .

*Solution* For every  $s \in \mathbb{S}^1$  and for every  $v \in T_s \mathbb{S}^1$ , by the basis theorem there is a  $c \in \mathbb{R}$  such that  $v = c d/d\theta$ . Let  $f : T\mathbb{S}^1 \rightarrow \mathbb{S}^1 \times \mathbb{R}$  be defined by  $f(s, v) = (s, r)$ . We must show that this is a diffeomorphism. It is injective, for if  $f(s_1, v_1) = f(s_2, v_2)$ , then  $s_1 = s_2$ , and hence  $v_1, v_2 \in T_{s_1} \mathbb{S}^1$ . But  $v_1 = c d/d\theta$  and  $v_2 = c d/d\theta$ , and hence  $v_1 = v_2$ . Thus,  $f$  is injective. It is surjective by the basis theorem. That is, every element of  $T\mathbb{S}^1$  is of the form  $(s, c d/d\theta)$  for some  $c \in \mathbb{R}$ . But this is smooth in both components, and hence it is a smooth function. The inverse function  $f^{-1}(s, c) = (s, c d/d\theta)$  is also smooth in both components, and hence  $f$  is a diffeomorphism.

**Problem 37.21.4** Let  $M$  be a smooth manifold with or without boundary and  $p \in M$ . Let  $C_p^\infty(M)$  denote the algebra of germs of smooth real-valued functions at  $p$  and  $\mathcal{D}_p^\infty(M)$  the vector space of derivations of  $C_p^\infty(M)$ . Show that  $\Phi : \mathcal{D}_p(M) \rightarrow T_p M$  defined by:

$$\Phi(v)(f) = v([f]_p) \quad (37.21.13)$$

is an isomorphism.

*Solution* We have to show that  $\Phi$  is linear and bijective. It is linear since:

$$\Phi(av + bw)(f) = (av + bw)([f]_p) = av([f]_p) + bw([f]_p) = a\Phi(v) + b\Phi(w) \quad (37.21.14)$$

it is injective since if  $\Phi(u) = \Phi(v)$ , then for all  $f$  we have that  $v([f]_p) = w([f]_p)$  and thus  $v$  and  $w$  are the same derivation, and hence  $v = w$ . It is surjective, for if  $\tilde{v}$  is a tangent vector, let  $v \in \mathcal{D}_p^\infty$  simply be the element such that  $v([f]) = \tilde{v}(f)$ . This is well defined since if  $g \in [f]$ , then  $g$  agrees with  $f$  on some neighborhood of  $p$  and thus  $\tilde{v}(f) = \tilde{v}(g)$ . But then  $\Phi(v) = \tilde{v}$ , and therefore  $\Phi$  is surjective. Thus,  $\Phi$  is an isomorphism.

## 37.22 Homework 4

**Problem 37.22.1** Lee 4-1.

*Solution* By Theorem 3.11 of the text, for any point  $p \in \mathbb{H}^n$  the inclusion map  $\iota : \mathbb{H}^n \rightarrow \mathbb{R}^n$  is such that the differential  $d_p \iota$  is a isomorphism, and hence the differential is invertible. But  $\iota$  cannot be a local homeomorphism for any point  $p \in \partial \mathbb{H}^n$  into  $\mathbb{R}^n$  since  $\mathbb{R}^n$  has no boundary, whereas any open subset of  $\mathbb{H}^n$  that contains a point  $p \in \partial \mathbb{H}^n$  will have boundary. Since there is no local homeomorphism, there can be no local diffeomorphism either. Hence, the theorem does not extend to manifolds with non-empty boundary.

**Problem 37.22.2** Lee 4-4.

*Solution* It suffices to show that  $A = \{(n, 2n\alpha) \bmod 1\}$  is dense in the unit square. If this set is dense in the rationals  $Q^2$  of the unit square, then it will be dense in the entire unit square since:

$$\text{Cl}_\tau(A) = \text{Cl}_\tau(\text{Cl}_\tau(A)) = \text{Cl}_\tau(Q^2) = I^2 \quad (37.22.1)$$

Let  $(p_1/p_2, q_1/q_2)$  be rational points in the square. Then by Dirichlet's approximation theorem for all  $\varepsilon > 0$  there exists  $n, m \in \mathbb{N}$  such that:

$$\left| n - \alpha \frac{q_1}{q_2} p_2 \right| m < \varepsilon \quad (37.22.2)$$

Moving the  $p_2/p_1$  to the other side, and choosing  $\varepsilon$  sufficiently small, shows that any rational point can be approximated arbitrarily well.

### Problem 37.22.3 Lee 4-6

*Solution* For suppose not, and let  $\phi : M \rightarrow \mathbb{R}^n$  be a smooth submersion. Since  $M$  is compact and  $\phi$  is continuous,  $\phi[M]$  is a compact subset of  $\mathbb{R}^n$  and hence by the Heine-Borel theorem it is closed and bounded. If  $\phi$  is a submersion, then it is an open mapping and hence  $\phi[M]$  is an open subset of  $\mathbb{R}^n$ . But  $\mathbb{R}^n$  is connected, and hence if  $\phi[M]$  is both closed and open, then it is either empty or all of  $\mathbb{R}^n$ . But  $M$  is non-empty, and hence  $\phi[M]$  is non-empty. But  $\phi[M]$  is bounded and  $\mathbb{R}^n$  is not bounded, a contradiction. Hence there is no submersion.

### Problem 37.22.4 Lee 4-9.

*Solution* For suppose  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are different smooth atlases and let  $p \in E$ . Given smooth charts  $(\mathcal{U}, \varphi) \in \mathcal{A}_1$  and  $(\mathcal{V}, \psi) \in \mathcal{A}_2$  containing  $p$ , we have that there exist smooth sections  $\sigma_1, \sigma_2 : M \rightarrow E$  whose image contain  $p$ . But then restricting the charts to be contained in these images we find that:

$$\varphi \circ \psi^{-1} = (\varphi \circ \sigma_1) \circ (\pi \circ \psi^{-1}) \quad (37.22.3)$$

which is smooth, and similarly the other way around. Hence  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are compatible.

### Problem 37.22.5 Lee 4-10.

*Solution* The map  $q$  gives a 2 to 1 covering of  $\mathbb{RP}^n$  by  $\mathbb{S}^n$ . Given a point  $s$  on the sphere, choose a neighborhood about  $s$  small enough to lie in a single hemisphere. Then this neighborhood  $\mathcal{U}$  and its reflection  $-\mathcal{U}$  have empty intersection and map to the same open subset of  $\mathbb{RP}^n$ ,  $\tilde{\mathcal{U}}$ . Hence the pre-image of  $\tilde{\mathcal{U}}$  is  $\mathcal{U} \cap -\mathcal{U}$ , both of which are homeomorphic to  $\tilde{\mathcal{U}}$ .

## 37.23 Homework 5

**Problem 37.23.1** Lee 8-1.

*Proof.* For given a closed subset  $A \subseteq M$  and an open subset  $\mathcal{U} \subseteq M$  there is a bump function  $f : M \rightarrow \mathbb{R}$  such that  $\text{supp}(f) \subseteq \mathcal{U}$  and  $f|_A = 1$ . Let  $V : M \rightarrow TM$  be defined by:

$$V_p = \begin{cases} f(p)X_p, & p \in \mathcal{U} \\ 0, & p \notin \mathcal{U} \end{cases} \quad (37.23.1)$$

Then  $V$  is a smooth vector field. Since smoothness is a local property we need only check this pointwise. For the points inside of  $\mathcal{U}$   $V$  is the product of a smooth function with a smooth vector field, and hence  $V$  is smooth. For points in the exterior  $V$  is zero, which is smooth. And moreover  $V$  is smooth on the boundary. Hence,  $V$  is a smooth vector field on  $M$ . By definition,  $V|_A = X$  since  $V_p = f(p)X_p = X_p$  since  $f(p) = 1$  for all  $p \in A$ .  $\square$

**Problem 37.23.2** Lee 8-16.

*Solution* Give  $X = y\partial_z - 2xy^2\partial_y$  and  $Y = \partial_y$  we have:

$$[X, Y] = XY - YX \quad (37.23.2)$$

$$= y\partial_{yz} - 2xy^2\partial_{yy} - \partial_z - y\partial_{yz} + 4xy\partial_y + 2xy^2\partial_{yy} \quad (37.23.3)$$

$$= -\partial_z + 4xy\partial_y \quad (37.23.4)$$

For  $X = x\partial_y - y\partial_x$  and  $Y = y\partial_z - z\partial_y$  we compute:

$$\begin{aligned} [X, Y] &= x\partial_z + xy\partial_{yz} - xz\partial_{yy} - y^2\partial_{xz} + yz\partial_{xy} \\ &\quad - \left( xy\partial_{yz} - y^2\partial_{xz} - xz\partial_{yy} + z\partial_x + yz\partial_{xy} \right) \end{aligned} \quad (37.23.5)$$

$$= x\partial_z - z\partial_x \quad (37.23.6)$$

Lastly, given  $X = x\partial_y - y\partial_x$  and  $Y = x\partial_y + y\partial_x$  we know all of the cross terms cancel, and so we are left with:

$$[X, Y] = x\partial_x - y\partial_y - (-x\partial_x + y\partial_y) \quad (37.23.7)$$

$$= 2x\partial_x - 2y\partial_y \quad (37.23.8)$$

**Problem 37.23.3** Lee 8-17.

*Solution* To show that  $X \oplus Y$  is smooth it suffices to show that this takes smooth functions to smooth functions. Thus, given  $(f, g) \in C^\infty(M \times N, \mathbb{R})$  we have:

$$(X \oplus Y)(f, g) = (X(f), Y(g)) \quad (37.23.9)$$

and since  $X$  and  $Y$  are smooth vector fields, this function is smooth in all coordinates and is hence smooth. Next we compute:

$$[X_1 \oplus Y_1, X_2 \oplus Y_2] = X_1 \oplus Y_1(X_2 \oplus Y_2) - X_2 \oplus Y_2(X_1 \oplus Y_1) \quad (37.23.10)$$

$$= X_1 X_2 \oplus Y_1 Y_2 - X_2 X_1 \oplus Y_2 Y_1 \quad (37.23.11)$$

$$= (X_1 X_2 - X_2 X_1) \oplus (Y_1 Y_2 - Y_2 Y_1) \quad (37.23.12)$$

$$= [X_1, X_2] \oplus [Y_1, Y_2] \quad (37.23.13)$$

**Problem 37.23.4** Lee 9-3.

*Solution* We need to solve  $\dot{\gamma}(t) = (\dot{x}(t), \dot{y}(t)) = V_{\gamma(t)}$ . We get  $\dot{x}(t) = y(t)$  and  $\dot{y}(t) = 1$ . This yields  $y(t) = t + c_1$ , and hence  $\dot{x}(t) = t + c_1$  giving  $x(t) = t^2/2 + c_1 t + c_0$ . Hence, the flow is  $\theta : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}$ :

$$\theta(t, (x, y)) = \left( \frac{t^2}{2} + yt + x, t + y \right) \quad (37.23.14)$$

For  $V = x\partial_x + 2y\partial_y$  we need  $\dot{x}(t) = x(t)$  and  $\dot{y}(t) = 2y(t)$ , so  $x(t) = c_0 \exp(t)$  and  $y(t) = c_1 \exp(2t)$ . The flow is then:

$$\theta(t, (x, y)) = (x \exp(t), y \exp(2t)) \quad (37.23.15)$$

**Problem 37.23.5** Lee 9-4.

*Solution* The infinitesimal generator of  $\theta(t, z) = z \exp(it)$  is the vector field  $V_z = \dot{\theta}^{(z)}(0)$ . We compute this and get  $V_z = iz$ . This is smooth, since the product of a smooth function by a constant is smooth, and  $f(z) = z$  is smooth. Moreover it is nonvanishing on  $\mathbb{S}^{2n-1}$  since  $\|iz\| = |i|\|z\| = \|z\| = 1$  for all  $z \in \mathbb{S}^{2n-1}$ . Hence,  $V$  is a non-vanishing vector field on  $\mathbb{S}^{2n-1}$ .

**Problem 37.23.6** Lee 9-6.

*Solution* For suppose not, and let  $C$  be a compact set that contains the image of  $\gamma[[t_0, b]]$ . Let  $y_n$  be the sequence  $\gamma(b - 1/n)$  starting with  $n$  large enough so that  $b - 1/n$  falls inside the interval  $[t_0, b]$ . Since  $C$  is compact, and since  $y_n \in C$  for all  $n$ , there is a convergent subsequence. Let  $\Lambda$  be the set of all limit points of this sequence. By the previous remark it is non-empty. Suppose there are at least two elements  $z_0, z_1$ . But  $\gamma$  is continuous, and hence if a subsequence of  $\gamma(b - 1/n)$  converges, then since  $b - 1/n$  is a Cauchy sequence, the entirety of  $\gamma(b - 1/n)$  must converge. But then  $y_n \rightarrow z_0$  and  $y_n \rightarrow z_1$  and therefore  $z_0 = z_1$ , a contradiction. Thus there is a unique limit point and we may define  $\gamma(b)$  to be this point. But by the uniqueness and existence theorem there is an integral curve of  $V$  through  $\gamma(b)$  for some interval  $(-\varepsilon, \varepsilon)$ , and hence  $\gamma$  may be extended to  $b + \varepsilon$ , a contradiction since  $b$  is the least upper bound of such values.

**Problem 37.23.7** Let  $p \in M$  and let  $\Lambda$  be the set of all points  $q \in M$  such that there is a diffeomorphism  $\phi : M \rightarrow M$  such that  $\phi(p) = q$ .  $\Lambda$  is non-empty since  $p \in \Lambda$  since the identity mapping is a diffeomorphism. Moreover,  $\Lambda$  is open. For if  $q \in \Lambda$ , let  $\mathcal{U}$  be an open subset of  $M$  that contains  $q$  such that  $\mathcal{U}$  is diffeomorphic to  $\mathbb{B}^n$ . Given any point  $x \in \mathcal{U}$ , we construct a diffeomorphism  $\phi : M \rightarrow M$  taking  $q$  to  $x$  as follows. Let  $r \in (0, 1)$  be such that  $\varphi(x)$  and  $\varphi(q)$  are both contained in the closed  $r$  ball centered about the origin, where  $\varphi$  is the coordinate chart for  $\mathcal{U}$ . Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a bump function that is 1 on this closed  $r$  ball and supported in  $\mathbb{B}^n$ . Let  $g(y) = f(y)(\varphi(q) - \varphi(x))$ . Let  $V$  be the vector field  $V(y) = g(y)$  and let  $\theta$  be the associated flow. The  $\theta(1, \varphi(q)) = \varphi(x)$ , and  $\theta$  is the identity map outside of  $\mathbb{B}^n$ . Moreover,  $\theta(1, \cdot)$  is a diffeomorphism. Hence, pulling back by  $\varphi$  and writing:

$$\phi(y) = \begin{cases} y, & y \notin \mathcal{U} \\ \varphi^{-1}(\theta(1, \varphi(y))), & y \in \mathcal{U} \end{cases} \quad (37.23.16)$$

gives us our diffeomorphism. Hence  $\mathcal{U} \subseteq \Lambda$  showing that  $\Lambda$  is open. Moreover,  $\Lambda$  is closed for if  $q_n$  is a convergent sequence with limit  $q$ , there is an open coordinate ball  $\mathcal{U}$  containing  $q$ . Since  $q_n \rightarrow q$  eventually the sequence is contained in  $\mathcal{U}$ . But then by the previous argument, eventually there is an  $N$  such that  $q_N$  can be swapped to  $q$  by diffeomorphism, and hence  $q \in \Lambda$ . But then  $\Lambda$  is open and closed, and since  $M$  is connected, either  $\Lambda = \emptyset$  or  $\Lambda = M$ . But  $p \in \Lambda$ , and hence  $\Lambda = M$ .

**Problem 37.23.8** Lee 9-10.

*Solution* At the point  $(1, 0)$  the flow is  $(0, 1)$ , which has non-zero  $y$  component and hence we choose that  $x$  axis. We write  $\Psi(s, t) = (t^2/2 + s, t)$ . The inverse of this is  $(s, t) = \Psi^{-1}(x, y) = (x - y^2/2, y)$ , which gives us our chart that turns this vector field into a coordinate vector field. For the next one we have  $(1, 0)$  which has non-zero  $x$  component and so we choose the  $y$  axis with parameterization  $X(s) = (0, s)$ . We have  $\Psi(s, t) = \theta_t(s, 0) = (s \exp(t), 0)$ . To get the coordinate vector field we invert this.



**Part XXII**

# **Euclidean Geometry**



# CHAPTER 38

---

## Convex Geometry

---

Very old notes that need a lot of touching up.

### 38.1 Convexity: Part I

#### 38.1.1 Convex Sets

**Definition 38.1.1** A subset  $K$  of a Banach Space  $X$  is convex if and only if  $[(x, y \in K) \wedge (\lambda \in [0, 1])] \Rightarrow [(1 - \lambda)x + \lambda y \in K]$ .

**Theorem 38.1.1.** *Convex spaces are path-connected.*

*Proof.* Let  $x, y \in V$  and define  $f : [0, 1] \rightarrow V$  by  $f(\lambda) = (1 - \lambda)x + \lambda y$ . Therefore, etc.  $\square$

**Theorem 38.1.2.** *Convex sets are connected.*

*Proof.* As convex sets are path-connected, they are connected.  $\square$

**Theorem 38.1.3.** *Any Banach Space  $X$  is convex.*

*Proof.* Let  $x, y \in V$ . As  $V$  is a vector space, for all  $\lambda \in \mathbb{R}$ ,  $(1 - \lambda)x + \lambda y \in V$ . Therefore, etc.  $\square$

The set of all convex subsets of  $\mathbb{R}^n$  is denoted  $\mathcal{K}_n$ .

**Theorem 38.1.4.** *If  $X$  is a Banach Space and  $A, B \subset X$  are convex sets, then  $A \cap B$  is convex.*

*Proof.*  $[x, y \in A \cap B] \Rightarrow [[\lambda \in [0, 1]] \Rightarrow [(1 - \lambda)x + \lambda y \in A] \wedge [(1 - \lambda)x + \lambda y \in B]] \Rightarrow [\forall \lambda \in [0, 1] : (1 - \lambda)x + \lambda y \in A \cap B]$ .  $\square$

**Theorem 38.1.5.** *In a Banach Space  $X$ ,  $\forall \xi \in X, \forall r > 0$ ,  $B_r(\xi)$  is convex.*

*Proof.*  $[x, y \in B_r(\xi)] \Rightarrow [[\lambda \in [0, 1]] \Rightarrow [|(1 - \lambda)x + \lambda y - \xi| \leq \frac{(1-\lambda)}{2} \|x - \xi\| + \frac{\lambda}{2} \|y - \xi\| < r]] \Rightarrow [(1 - \lambda)x + \lambda y \in B_r(\xi)].$   $\square$

**Theorem 38.1.6.** *There exist convex sets  $A$  and  $B$  such that  $A \cap B \neq \emptyset$ , yet  $A \cup B$  is not convex.*

*Proof.* For let  $A = \{(x, y) \in \mathbb{R}^2 : (x - 1)^2 + y^2 \leq 1\}$ , and  $B = \{(x, y) \in \mathbb{R}^2 : (x + 1)^2 + y^2 \leq 1\}$ . Both of these sets are discs, and thus convex, their intersection is the point  $(0, 0)$ , however their union is not convex.  $\square$

**Theorem 38.1.7.** *If  $K \subset X$  is convex and  $\psi : X \rightarrow X$  is a linear transformation, then  $\psi(K)$  is convex.*

*Proof.*  $[X, Y \in \psi(K)] \Rightarrow [\exists x, y \in K : \psi(x) = X \wedge \psi(y) = Y].$   $[\lambda \in [0, 1]] \Rightarrow [(1 - \lambda)x + \lambda y \in K] \Rightarrow [\psi((1 - \lambda)x + \lambda y) = (1 - \lambda)\psi(x) + \lambda\psi(y) = (1 - \lambda)X + \lambda Y \in \psi(K)]$   $\square$

**Theorem 38.1.8.** *There exist non-convex, compact subsets  $K$  of a Banach Space  $X$  :  $K_\xi$  is convex for every subspace  $\xi$ .*

*Proof.* Let  $r > 0$ ,  $\mathcal{M} = \{B_r(\mathbf{0}) \setminus B_{r/2}(\mathbf{0})\}$ . This is not convex, but for any subspace  $\xi$ ,  $\mathcal{M}_\xi = B_r(\mathbf{0}) \cap \xi$ , which is convex.  $\square$

**Theorem 38.1.9.** *If  $X$  is a Banach Space,  $K \subset X$ , and for every affine subspace  $\xi$ ,  $K \cap \xi$  is convex, then  $K$  is convex.*

*Proof.* Let  $x, y \in K$ . Let  $W$  be a subspace containing  $y$ , and let  $\xi$  be the affine subspace  $\{x - y + v : v \in W\}$ . Then  $\xi \cap K$  is convex, and thus  $(1 - \lambda)x + \lambda y \in \xi \cap K \Rightarrow (1 - \lambda)x + \lambda y \in K$ .  $\square$

**Theorem 38.1.10.** *If  $X$  is a Banach Space,  $K, L \subset X$  are convex, then  $K + L$  is convex.*

*Proof.* For let  $\chi$  and  $\zeta$  be elements of  $K + L$ . Then there are points  $x_1, x_2$  and  $y_1, y_2$  such that  $\chi = x_1 + x_2$  and  $\zeta = y_1 + y_2$  with  $x_1, y_1 \in K$  and  $x_2, y_2 \in L$ . If  $\lambda \in [0, 1]$ , then  $(1 - \lambda)\chi + \lambda\zeta = (1 - \lambda)(x_1 + x_2) + \lambda(y_1 + y_2) = (1 - \lambda)x_1 + \lambda y_1 + (1 - \lambda)x_2 + \lambda y_2$ . But  $K$  and  $L$  are convex, and thus  $(1 - \lambda)x_1 + \lambda y_1 \in K$  and  $(1 - \lambda)x_2 + \lambda y_2 \in L$ . But then  $(1 - \lambda)x_1 + \lambda y_1 + (1 - \lambda)x_2 + \lambda y_2 = (1 - \lambda)\chi + \lambda\zeta \in K + L$ . As  $\lambda$  is arbitrary in  $[0, 1]$ ,  $K + L$  is convex.  $\square$

**Theorem 38.1.11.** *If  $V$  is a Banach Space,  $K \subset V$  is convex, and  $\alpha \in \mathbb{R}$ , then  $\alpha K$  is convex.*

*Proof.* For let  $X, Y \in \alpha K$ . Then  $X = \alpha x$  and  $Y = \alpha y$  for  $x, y \in K$ . Let  $\lambda \in (0, 1)$  be arbitrary. Then  $(1 - \lambda)X + \lambda Y = (1 - \lambda)(\alpha x) + \lambda(\alpha y) = \alpha[(1 - \lambda)x + \lambda y]$ . As  $K$  is convex,  $(1 - \lambda)x + \lambda y \in K$ . Thus  $\alpha[(1 - \lambda)x + \lambda y] \in \alpha K$ , and  $\alpha K$  is convex.  $\square$

**Theorem 38.1.12.** *If  $K$  and  $L$  are compact and convex, and  $K \cup L$  is convex, then  $K \cap L \neq \emptyset$ .*

*Proof.* For if not, then  $K \cup L$  is the union of two closed, disjoint subsets, and is thus disconnected. A contradiction.  $\square$

**Theorem 38.1.13.** *Any set is a subset of its convex hull.*

*Proof.* For  $x \in K \Rightarrow 1 \cdot x = x \in \text{conv}(K)$ . Therefore, etc.  $\square$

**Theorem 38.1.14.** *If  $x \in \text{conv}(K)$ , then there are two points in  $\text{conv}(K)$  such that  $x = \lambda v_1 + (1 - \lambda)v_2$ ,  $0 \leq \lambda \leq 1$ .*

*Proof.* Let  $x = \sum_{k=1}^{n+1} |\lambda_k| x_k$ . If  $x = |\lambda_{n+1}| x_{n+1}$ , we are done. If not, then let  $v_2 = \frac{1}{1-|\lambda_{n+1}|} \sum_{k=1}^n |\lambda_k| x_k$ . As  $\frac{1}{1-|\lambda_{n+1}|} \sum_{k=1}^n |\lambda_k| = 1$ ,  $v_1 \in \text{conv}(K)$ . Thus,  $x = \lambda_{n+1} x_{n+1} + (1 - \lambda_{n+1})v_2$ . Let  $x_{n+1} = v_1$  and  $\lambda = \lambda_{n+1}$ . Therefore, etc.  $\square$

**Theorem 38.1.15.** *The convex hull of a convex set is itself.*

*Proof.*  $[x \in \text{conv}(K)] \Rightarrow [\exists v_1, v_2 \in K \wedge \lambda \in [0, 1] : x = \lambda v_1 + (1 - \lambda)v_2] \Rightarrow [x \in K]$ . The last step is from convexity.  $\square$

**Theorem 38.1.16.** *If  $K$  is compact and convex,  $\lambda_n \in [0, 1]$ ,  $v_n \in K$ , and  $\sum_{k=1}^n \lambda_k \rightarrow 1$ , then  $\sum_{k=1}^n \lambda_k v_k$  converges in  $K$ .*

*Proof.* Let  $x \in K$  be arbitrary,  $\varepsilon > 0$ ,  $s_n = \sum_{k=1}^n \lambda_k v_k$ , and  $\Lambda_n = \sum_{k=1}^n \lambda_k$ . Let  $S_n = s_n + (1 - \Lambda_n)x$ . For each  $n \in \mathbb{N}$ ,  $S_n \in K$ . It now suffices to show  $S_n$  converges. As  $K$  is compact,  $\{\|v\| : v \in K\}$  is bounded, let  $M$  be a bound. As  $\Lambda_n \rightarrow 1$ ,  $\exists N \in \mathbb{N} : n, m > N \Rightarrow |\Lambda_n - \Lambda_m| < \frac{\varepsilon}{2M}$ . But then for  $n \geq m > N$ ,  $d(S_n, S_m) = \left\| \sum_{k=m}^n \lambda_k v_k + x(\Lambda_n - \Lambda_m) \right\| \leq \sum_{k=m}^n |\lambda_k| \|v_k\| + \|x\|(|\Lambda_n - \Lambda_m|) \leq 2M|\Lambda_n - \Lambda_m| < \varepsilon$ . Thus,  $S_n$  is Cauchy. As  $K$  is compact, it is complete, and thus  $S_n$  converges in  $K$ . But  $d(S_n, s_n) \leq |1 - \Lambda_n|M \rightarrow 0$ . Thus,  $s_n$  converges to the same limit. Therefore, etc.  $\square$

**Theorem 38.1.17.** *The convex hull of a compact set is compact.*

*Proof.* Yeah  $\square$

**Theorem 38.1.18.** *The convex hull of an open set is open.*

*Proof.* Let  $x \in \text{conv}(K)$ . Then  $x = \sum_{k=1}^n |\lambda_k| x_k$ , where  $\sum_{k=1}^n |\lambda_k| = 1$ . At least one  $\lambda_k$  must be non-zero, suppose  $\lambda_1$  is. Define the function  $f : K \rightarrow K$  by  $f(z) = \frac{1}{|\lambda_1|}(z - \sum_{k=2}^n |\lambda_k| v_k)$ . Then  $f$  is continuous, injective, and thus  $f^{-1}$  exists and is equal to  $f^{-1}(z) = \lambda_1 z + \sum_{k=2}^n \lambda_k v_k$ . Thus,  $x \in f^{-1}(K) \subset \text{conv}(K)$ . Therefore, etc.  $\square$

**Theorem 38.1.19.** *There exist closed sets whose convex hull is open (And not closed).*

*Proof.* Let  $K = \{(x, y) \in \mathbb{R}^2 : \frac{1}{1+x^2} \leq y\}$ . The complement is open, and thus this is closed. But  $\text{conv}(K) = \{(x, y) \in \mathbb{R}^2 : y > 0\}$ , which is open. As  $\mathbb{R}^2$  is connected, this set is not closed.  $\square$

**Theorem 38.1.20** (Carathéodory's Theorem). *If  $S$  is an  $n$  dimensional subset of a Banach Space, there exists  $n + 1$  points  $x_k$  in  $S$  such that  $\text{conv}(S) = \text{conv}(x_1, \dots, x_{n+1})$ .*

*Proof.* Let  $y \in \text{conv}(S)$ . If  $y \notin S$ , then there are points  $x_1, \dots, x_m$  and positive real numbers  $\lambda_1, \dots, \lambda_m$  such that  $y = \sum_{k=1}^m \lambda_k x_m$  and  $\sum_{k=1}^m \lambda_k = 1$ . Let  $m$  be the minimal number of points needed. Suppose the  $x_i$  are affinely dependent. Then there are real numbers  $\alpha_i$  such that  $\sum_{k=1}^m \alpha_k x_k = \mathbf{0}$  and  $\sum_{k=1}^m \alpha_k = 1$ . But then  $y = y + t\mathbf{0}$  for all  $t \in \mathbb{R}$ , and thus  $y = \sum_{k=1}^m \lambda_k x_k + t \sum_{k=1}^m \alpha_k x_k = \sum_{k=1}^m (\lambda_k + t\alpha_k) x_k$  and  $\sum_{k=1}^m (\lambda_k + t\alpha_k) = 1$ . Let  $|t|$  be the smallest value such that  $\lambda_j + t\alpha_j = 0$  for some  $j$ . Then, for all other values,  $\lambda_k + t\alpha_k \in [0, 1]$ . Thus,  $y$  is represented by a convex combination of  $m - 1$  points, a contradiction as  $m$  is minimal. Thus, the  $x_k$  are affinely independent. But then  $m \leq n + 1$ . Thus, etc.  $\square$

**Definition 38.1.2** A set  $P$  is said to be a convex polytope if and only if it is the convex hull of finitely many points.

**Definition 38.1.3** A polytope in  $\mathbb{R}^2$  is called a polygon.

The set of all polytopes in  $\mathbb{R}^n$  is denoted  $\mathcal{P}_n$ .

**Definition 38.1.4** A polytope  $P$  is said to be a simplex if and only if then generating points are affinely independent.

**Theorem 38.1.21.** *An  $n$  dimensional simplex has  $n + 1$  vertices.*

**Theorem 38.1.22.** *If  $P$  is a polytope in some Banach Space  $V$ , and  $\xi$  is a subspace, then  $P_\xi$  is a polytope.*

**Theorem 38.1.23.** *If  $K$  is a compact set of finite dimension  $n \geq 3$ , and if for every  $n - 1$  dimensional subspace  $\xi$ ,  $K_\xi$  is a convex polytope, then  $K$  is a convex polytope. (Page 23)*

**Theorem 38.1.24** (Radon's Theorem). *If  $S$  Is a finite dimensional affinely dependent subset of a Banach Space  $V$ , then there are sets  $S_1, S_2 \in V$  such that  $S_1 \cap S_2 = \emptyset$ ,  $S = S_1 \cup S_2$ , and  $\text{conv}(S_1) \cap \text{conv}(S_2) \neq \emptyset$ .*

*Proof.* As  $S$  is affinely dependent, there are distinct points  $x_1, \dots, x_n$  and non-zero real numbers  $\lambda_1, \dots, \lambda_n$  such that  $\sum_{k=1}^n \lambda_k x_k = \mathbf{0}$  and  $\sum_{k=1}^n \lambda_k = 0$ . As none of the values of  $\lambda_k$  are zero, some must be positive and some negative. Let  $\lambda_1, \dots, \lambda_j$  be positive and  $\lambda_{j+1}, \dots, \lambda_n$  are negative. Then  $\sum_{k=1}^j \lambda_k = -\lambda_{j+1} \dots \lambda_n$ . Let this sum be  $c$ . Let  $T_1 = \{x_1, \dots, x_j\}$  and  $T_2 =$

$\{x_{j+1}, \dots, x_n\}$ . As the  $x_k$  are distinct,  $T_1 \cap T_2 = \emptyset$ . However,  $\frac{1}{c} \sum_{k=1}^j \lambda_k x_k = \frac{-1}{c} \sum_{k=j+1}^n \lambda_k x_k \in \text{conv}(T_1) \cap \text{conv}(T_2)$ . Let  $S_1 = S \setminus T_2$ , and  $S_2 = T_2$ . Therefore, etc.  $\square$

**Theorem 38.1.25.** *If  $K$  is compact, convex, and a subset of a two dimensional Banach Space, then if  $A = \{x \in K : B_1(x) \subset K\}$ , then  $\bigcup_{x \in A} B_1(x)$  is convex.*

**Theorem 38.1.26.** *If  $K$  is a compact, convex subset of a two dimensional Banach Space, then  $\bigcup_{x \in K} B_1(x)$  is convex.*

### 38.1.2 Convexity in the Euclidean Plane

**Definition 38.1.5**  $A(K)$  is the area of a compact set  $K$ .

**Definition 38.1.6** The perimeter  $P(K)$  is the length of the boundary curve,  $\mu(\partial K)$ .

**Definition 38.1.7** The section  $X_\ell(K)$  is defined for any given straight line  $\ell$  to be the length  $\mu(K \cap \ell)$ .

**Definition 38.1.8** Given a straight line  $\ell$ , the width  $W_\ell(K)$  is  $\mu(\ell_K)$ .

**Definition 38.1.9** The mean width  $W(K)$  is the average of the width of all  $W_\ell(K)$ .

**Definition 38.1.10** The minimum width  $w(K)$  is the minimum of  $W_\ell(K)$  over all possible directions of the line  $\ell$ .

**Definition 38.1.11** The maximum width  $\check{W}$  is the maximum of  $W_\ell(K)$ .

**Definition 38.1.12** The diameter  $D(K)$  is the maximum distance between any two points inside  $K$ .

**Definition 38.1.13** The inradius  $r_K$  is the maximum radius of any circle inside  $K$ .

**Definition 38.1.14** The circumradius  $R_K$  is the minimum radius of any circle containing  $K$ .

For bounded sets  $K \subset \mathbb{R}^2$ , define  $R_K(\theta)$  as the projection of  $K$  onto the line  $\sin(\theta)y = \cos(\theta)x$ .

**Theorem 38.1.27.** *If  $K$  is a compact subset of  $\mathbb{R}^n$ , then there are points  $x, y \in K$  such that  $d(x, y) = D(K)$ .*

*Proof.* As  $K$  is compact, it is bounded. Let  $M$  be such a bound. Then the set  $D = \{r \in \mathbb{R}^+ | d(x, y) = r, x, y \in K\}$  is bounded by  $M$ . Let  $d$  be the least upper bound of this set. Then we can find a sequence of points  $x_n, y_n \in K$

such that  $d(x_n, y_n) \rightarrow d$ . As  $K$  is compact, there are convergent subsequences  $x_{n_k}$  and  $y_{n_k}$  with limits in  $K$ , call them  $x$  and  $y$ . Then  $d(x, y) = d$ . Thus, there are points  $x, y \in K$  such that  $d(x, y) \geq d(a, b)$  for all  $a, b \in K$ . Therefore  $D(K)$  is equal to  $d$ .  $\square$

**Theorem 38.1.28.** *There exist bounded sets  $K$  such that  $R_K(\theta)$  is not continuous.*

*Proof.* For define:

$$\mathcal{M} = \{(x, y) \in [-1, 1]^2 : x \in \mathbb{Q}, y \in \mathbb{R}\} \quad (38.1.1)$$

Also, let  $K = \mathcal{M} \cap \overline{D^2}$ , where  $\overline{D^2}$  denotes the closed unit disc in  $\mathbb{R}^2$ . Then  $f_K(\theta)$  is discontinuous at  $\theta = \frac{\pi}{2}$ . For,  $f_K(\frac{\pi}{2}) = \mu(\mathbb{Q} \cap [-1, 1]) = 0$ . But, for any  $0 < \delta < \pi$ ,  $f_K(\frac{\pi}{2} \pm \delta) = 2$ . Thus:

$$f_K(\theta) = \begin{cases} 0, & \theta = \pm\frac{\pi}{2} \\ 2, & \theta \neq \pm\frac{\pi}{2} \end{cases} \quad (38.1.2)$$

And this is discontinuous.  $\square$

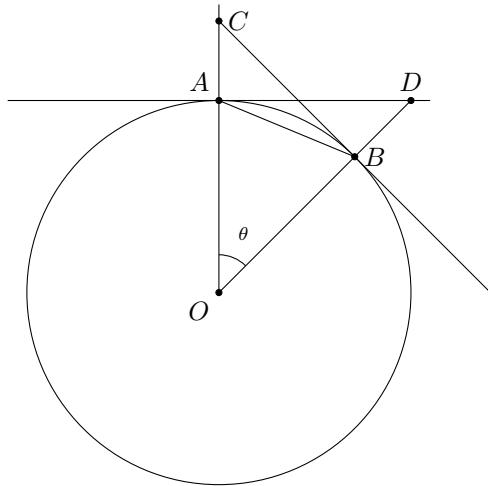


Fig. 38.1: Drawing for Thm. 38.1.28.

**Theorem 38.1.29.** *If  $K \subset \mathbb{R}^2$  is compact and convex, then  $R_K(\theta)$  is continuous.*

*Proof.* For let  $K$  be compact and convex, and without loss of generality suppose it contained within the unit disc and contains the origin. Let  $\theta \in (0, 2\pi)$

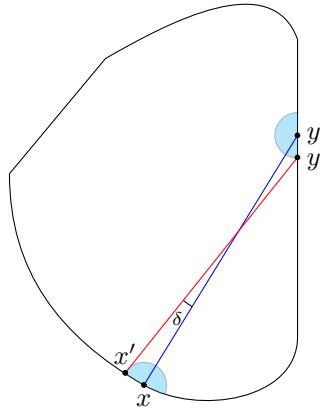


Fig. 38.2: Drawing for Thm. 38.1.29.

be given and let  $\ell_\theta$  be a line through the origin making an angle  $\theta$  with the horizontal. Let  $x = (x_1, x_2), y = (y_1, y_2) \in K$  be such that  $W_{\ell_\theta}(K) = d(x, y)$ . Such points exist as  $K$  is compact. Let  $\varepsilon > 0$  be given. About  $x$ , consider  $B_{\frac{\varepsilon}{2}}(x)$ , and similarly  $B_{\frac{\varepsilon}{2}}(y)$ . Neither of these are empty, as  $K$  is convex. Let  $x', y' \in K \cap (B_{\frac{\varepsilon}{2}}(x) \cup B_{\frac{\varepsilon}{2}}(y))$  be such that  $d(x', y')$  is maximized. As this set is compact, such points exist. Let  $\delta$  be defined by:

$$\delta = \min\left\{ \left| \theta - \frac{x'_2}{\sqrt{x'^2_1 + x'^2_2}} \right|, \left| \theta - \frac{y'_2}{\sqrt{y'^2_1 + y'^2_2}} \right| \right\} \quad (38.1.3)$$

Then, since  $|\theta - \theta_0| < \delta$ :

$$d(x, y) - \varepsilon \leq W_{\ell_{\theta_0}}(K) \leq d(x, y) + \varepsilon \quad (38.1.4)$$

and thus  $|W_{\ell_\theta}(K) - W_{\ell_{\theta_0}}(K)| < \varepsilon$ .  $\square$

**Theorem 38.1.30.** *If  $K$  is a compact subset of  $\mathbb{R}^2$  and  $x, y \in K$  such that  $d(x, y) = D(K)$ , then the lines perpendicular to  $\overline{xy}$  at  $x$  and  $y$  contain all of  $K$  in between.*

*Proof.* Suppose not. Let  $\overline{X}$  be the line perpendicular to  $\overline{xy}$  containing point  $x$ , and similarly define  $\overline{Y}$ . Suppose there is a point  $z \in K$  that falls on the exterior of the region:

$$\mathcal{U} = \{(x, y) \in \mathbb{R}^2 : (x, y) \text{ Lies Between } \overline{X} \text{ and } \overline{Y}\} \quad (38.1.5)$$

Note that  $d(x, z) \neq d(y, z)$ , as then  $z$  would lie between these two lines. Suppose  $d(y, z) < d(x, z)$ . Where the line  $\overline{xz}$  cuts  $\overline{Y}$  denote as  $A$ . But then:

$$d(x, z) > d(x, A) = \sqrt{d(y, A)^2 + d(x, y)^2} \geq d(x, y) \quad (38.1.6)$$

A contradiction as  $d(x, y) = D(K)$ . Thus  $z \in \mathcal{U}$ .  $\square$

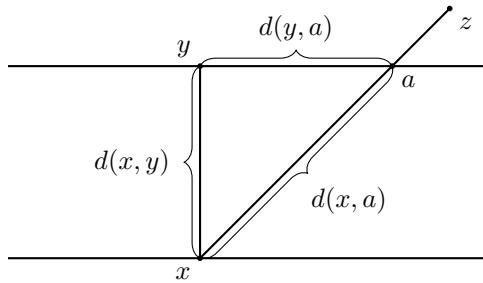


Fig. 38.3: Drawing for Thm. 38.1.30.

**Theorem 38.1.31.** If  $K \subset \mathbb{R}^2$  is compact and convex, then there are points  $x, y \in K$  such that  $\check{W}(K) = d(x, y)$ .

*Proof.* As  $f_K(\theta)$  is continuous for convex compact set, and as it is continuous on a compact set  $[0, 2\pi]$ , it attains its maximum. Let  $\theta$  be such a maximum. Let  $\ell_\theta$  be the line through the origin which makes an angle  $\theta$  with the horizontal axis and consider set  $K_{\ell_\theta}$ . As  $K$  is compact and  $\ell_\theta$  is closed,  $K_{\ell_\theta}$  is compact. Then  $W = \{d(x, y) : x, y \in K_{\ell_\theta}\}$  is bounded, has a least upper bound, and therefore there are points  $x, y \in K$  such that  $\check{W}(K) = d(x, y)$ .  $\square$

**Theorem 38.1.32.** If  $K$  is a compact convex set of  $\mathbb{R}^2$ , then  $D(K) = \check{W}(K)$ .

*Proof.* As  $K$  is compact,  $D(K)$  and  $\check{W}(K)$  exists and there are points  $x, y$  such that  $d(x, y) = D(K)$  and points  $x', y'$  such that  $d(x', y') = \check{W}(K)$ . Suppose  $d(x', y') > d(x, y)$ . A contradiction, as  $d(x, y)$  is the diameter of  $K$ . So  $d(x, y) \geq d(x', y')$ . Now suppose  $d(x, y) > d(x', y')$ . But as  $d(x', y') = \check{W}(K)$ ,  $d(x', y')$  is the greatest length of any line segment that terminates in  $K$  and such that perpendiculars at these terminating points contain all of  $K$ . But as  $d(x, y) = D(K)$ , the lines perpendicular to  $\overline{xy}$  at  $x$  and  $y$  contain all of  $K$ , a contradiction. Thus  $d(x', y') \geq d(x, y)$ . But it was just showed that  $d(x, y) \geq d(x', y')$ . Thus,  $d(x, y) = d(x', y')$ .  $D(K) = \check{W}(K)$ .  $\square$

**Definition 38.1.15** If  $Q$  is a convex polygon with interior (That is, positive area) in  $\mathbb{R}^2$ , then the perimeter of  $Q$  is the sum of the lengths of its edges.

**Definition 38.1.16** The perimeter of a line segment  $e$  is  $P(e) = 2|e|$ .

This is for the sake of continuity. If we take a rectangle of length  $|e|$  and width  $\frac{1}{n}$ , then the perimeter is  $2|e| + \frac{2}{n} \rightarrow 2|e|$  as  $n \rightarrow \infty$ . Thus, for the purpose of continuity we define the perimeter of line segments to be twice their length.

**Theorem 38.1.33** (Cauchy's Perimeter Theorem). *If  $K$  is a compact convex subset of the plane, then  $P(K) = \pi W(K)$ .*

*Proof.* Suppose that  $Q$  is a convex polygon with edges  $e_1, \dots, e_n$ . At each  $e_i$ , let  $\theta_i$  be the angle made with  $e_i$  and the horizontal axis of  $\mathbb{R}^2$ . The mean width is:

$$W(Q) = \frac{1}{2\pi} \int_0^{2\pi} W_{\ell_\theta}(Q) d\theta \quad (38.1.7a)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2} \sum_{i=1}^n |e_i| |\cos(\theta - \theta_i)| d\theta \quad (38.1.7b)$$

$$= \frac{1}{4\pi} \sum_{i=1}^n |e_i| \int_0^{2\pi} |\cos(\theta - \theta_i)| d\theta \quad (38.1.7c)$$

$$= \frac{1}{\pi} \sum_{i=1}^n |e_i| \quad (38.1.7d)$$

$$= \frac{1}{\pi} P(Q) \quad (38.1.7e)$$

Thus,  $P(Q) = \pi W(Q)$ . For any convex compact subset  $K \subset \mathbb{R}^2$ , we may find a polygon  $Q$  that approximates the boundary with a perimeter  $P(Q)$  that is as close to  $P(K)$  and a width  $W(Q)$  as close to  $W(K)$  as desired. That is, for all  $n \in \mathbb{N}$ , we can obtain a polynomial  $Q_n$  such that:

$$\max\{|W(Q_n) - W(K)|, |P(K) - P(Q_n)|\} < \frac{1}{n} \quad (38.1.8)$$

But then:

$$|P(K) - \pi W(K)| = |P(K) - P(Q_n) + P(Q_n) - \pi W(Q_n) + \pi W(Q_n) - \pi W(K)| \quad (38.1.9a)$$

$$\leq |P(K) - P(Q_n)| + |P(Q_n) - \pi W(Q_n)| + \pi |W(Q_n) - W(K)| \quad (38.1.9b)$$

$$< \frac{1}{n} + 0 + \frac{\pi}{n} \quad (38.1.9c)$$

$$= \frac{1 + \pi}{n} \quad (38.1.9d)$$

And this tends to zero as  $n$  tends to infinity. Thus,  $P(K) = \pi W(K)$ . □

**Theorem 38.1.34.** *There exist compact path-connected sets  $K \subset \mathbb{R}^2$  such that  $P(K) \neq \pi W(K)$ .*

*Proof.* Consider the set  $K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1, y \geq 0\}$ . Then  $P(K) = 2\pi$ , but  $W(K) = \pi(\pi + 2)$ . To see this, consider the set  $\mathcal{K} = \{(x, y) \in$

$\mathbb{R}^2 : x^2 + y^2 \leq 1, y \geq 0\}$ . This is convex and has perimeter  $\pi + 2$  and therefore  $W(\mathcal{K}) = \pi(\pi + 2)$ . But, as the image shows,  $W(K) = W(\mathcal{K})$ . That is,  $W_{\ell_\theta}(K)$  is the length of the line segment  $\overline{AB}$ , as is  $W_{\ell_\theta}(\mathcal{K})$ . Therefore the averages  $W(K)$  and  $W(\mathcal{K})$  are the same. Thus,  $P(K) \neq \pi W(K)$ .  $\square$

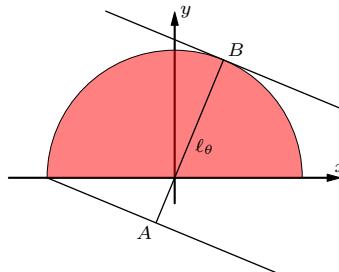


Fig. 38.4: Caption

**Definition 38.1.17** A functional  $f$  with respect to subset inclusion is said to be monotonic on a family of sets  $\mathcal{P}$  if and only if  $f(K) \leq f(L)$  for all  $K, L \in \mathcal{P}$  such that  $K \subset L$ .

**Theorem 38.1.35.** If  $K \subset L \in \mathcal{K}_2$ , then  $W_\ell(K) \leq W_\ell(L)$ .

*Proof.* For let  $\pi_\ell : \mathbb{R}^2 \rightarrow \ell$  be the orthogonal projection map of  $\mathbb{R}^2$  to  $\ell$ . Then  $\pi_\ell(K) \leq \pi_\ell(L)$  as  $K \subset L$ , and thus  $W_\ell(K) \leq W_\ell(L)$ .  $\square$

**Theorem 38.1.36.** If  $K \subset L \in \mathcal{K}_2$ , then  $W(K) \leq W(L)$ .

*Proof.* For  $W_\ell(K) \leq W_\ell(L)$ , and thus  $W(K) = \frac{1}{2\pi} \int_0^{2\pi} W_{\ell_\theta}(K) d\theta \leq \frac{1}{2\pi} \int_0^{2\pi} W_{\ell_\theta}(L) d\theta = W(L)$ .  $\square$

**Theorem 38.1.37.** If  $K \subset L \in \mathcal{K}_2$ , then  $X_\ell(K) \leq X_\ell(L)$ .

*Proof.* For if  $x \in \ell \cap K$ , then  $x \in \ell \cap L$ , and thus  $X_\ell(K) = \mu(\ell \cap K) \leq \mu(\ell \cap L) = X_\ell(L)$ .  $\square$

**Theorem 38.1.38.** If  $K \subset L \subset \mathcal{K}_2$ , then  $D(K) \leq D(L)$ .

*Proof.* For suppose not. Suppose  $D(K) > D(L)$ . Then, there are points  $x, y \in K$  such that  $d(x, y) > \sup\{d(x', y') : x', y' \in L\}$ . But as  $K \subset L$ ,  $x, y \in L$  and thus  $d(x, y) \leq \sup\{d(x', y') : x', y' \in L\}$ . Thus,  $D(K) \leq D(L)$ .  $\square$

**Theorem 38.1.39.** If  $K \subset L \subset \mathcal{K}_2$ , then  $R_K \leq R_L$ .

*Proof.* For suppose not. Suppose  $R_K > R_L$ . But as  $K \subset L$ , either this circle contains all of  $L$  as well or it doesn't. But then  $R_L \not\leq R_K$ . Thus,  $R_L \geq R_K$ .  $\square$

**Theorem 38.1.40.** If  $K \subset L \in \mathcal{K}_2$ , then  $r_K \leq r_L$ .

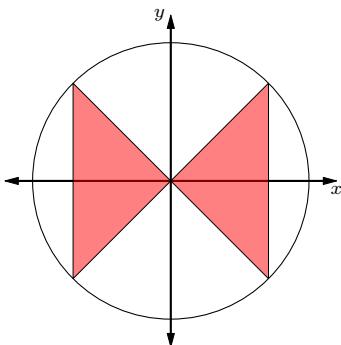
*Proof.* For suppose not. Suppose  $r_K > r_L$ . But as the inscribed circle of radius  $r_K$  fits entirely in  $K$ , and  $K \subset L$ , then it fits inside of  $L$ . But then  $r_K > r_L$ . Thus,  $r_L \geq r_K$ .  $\square$

**Theorem 38.1.41.** If  $K, L \in \mathcal{K}_2$  and  $K \subset L$ , then  $P(K) \leq P(L)$ .

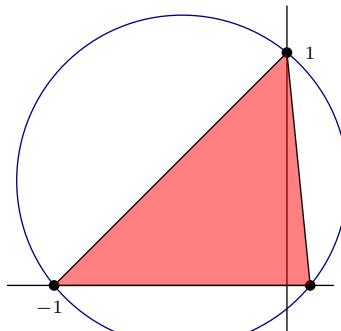
*Proof.* As  $K \subset L \in \mathcal{K}_2$ ,  $W(K) \leq W(L)$ . As  $K$  and  $L$  are convex,  $P(K) = \pi W(K)$  and  $P(L) = \pi W(L)$ . Thus,  $P(K) \leq \pi W(L) = P(L)$ . Therefore,  $P(K) \leq P(L)$ .  $\square$

**Theorem 38.1.42.** There exists sets  $K, L \subset \mathbb{R}^2$  such that  $L$  is convex,  $K \subset L$ , yet  $P(K) > P(L)$ .

*Proof.* For let  $L = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ . Then  $P(L) = 2\pi$ . Let  $K = \{(x, y) \in \mathbb{R}^2 : -\sqrt{2}x \leq y \leq \sqrt{2}x, \frac{-1}{\sqrt{3}} \leq x \leq \frac{1}{\sqrt{3}} \vee \sqrt{2}x \leq y \leq -\sqrt{2}x, \frac{-1}{\sqrt{3}} \leq x \leq \frac{1}{\sqrt{3}}\}$ .  $P(K) = 4(1 + \sqrt{\frac{2}{3}}) \approx 7.26 > 2\pi$ .  $\square$



38.5.1: Drawing for Theorem 3.2.15



38.5.2: Drawing for Theorem 3.2.16

**Theorem 38.1.43.** There exist compact convex sets in  $\mathbb{R}^2$  such that  $D(K) < 2R_K$ .

*Proof.* For consider the set  $K = \{(x, y) \in \mathbb{R}^2 : [0 \leq y \leq x+1 \wedge x \geq 0] \vee [0 \leq y \leq -10x+1 \wedge x \geq 0]\}$ . From Euclid, the smallest circle containing this triangle is the one defined by the three vertices (Three points define a triangle). This circle has the formula  $(x + 0.45)^2 + (y - 0.45)^2 = (0.45)^2 + (0.45 + \frac{1}{10})^2$ . Thus,  $2R_K = \sqrt{0.45^2 + (0.45 + \frac{1}{10})^2} > \sqrt{2} = D(K)$ .  $\square$

**Theorem 38.1.44.** If  $K$  is a compact convex set in  $\mathbb{R}^2$ , then  $D(K) \leq 2R_K$ .

*Proof.* For suppose not. Suppose  $D(K) > 2R_K$ . As  $K$  is compact, there are points  $x$  and  $y$  in  $K$  such that  $d(x, y) = D(K)$ . But then  $d(x, y) > 2R_K$ , and thus at least one of  $x$  or  $y$  is not contained in the circle. A contradiction. Thus  $D(K) \leq 2R_K$ .  $\square$

**Theorem 38.1.45.** *If  $K$  is a compact convex set in  $\mathbb{R}^2$ , then  $2\pi r_K \leq P(K)$ .*

*Proof.* From Calculus of Variations we know that the circle maximizes the area contained with a set of perimeter  $p$ . As the inscribed circle has perimeter  $2\pi r_K$  and as the circle is a subset of  $K$ , it is true that  $A(K)$  is greater than or equal the area of the circle. Thus,  $P(K) \geq 2\pi r_K$ .  $\square$

**Theorem 38.1.46.** *For a compact convex set of  $\mathbb{R}^2$ ,  $P(K) \leq 2\pi R_K$ .*

*Proof.* As  $K$  is convex,  $P(K) = \pi W(K) \leq \pi \check{W}(K) = \pi D(K) \leq 2\pi R_K$ .  $\square$

**Theorem 38.1.47.** *If  $K$  is a compact and convex subset of  $\mathbb{R}^2$ , then  $2D(K) \leq P(K)$ .*

*Proof.* Suppose not. Suppose  $2D(K) > P(K)$ . As  $K$  is compact and convex, there exist points  $x, y \in K$  such that  $d(x, y) = D(K)$ . But as  $K$  is convex, the line contained between  $x, y$  is contained in  $K$ . Thus,  $P(\overline{xy}) = 2D(K)$ . But as  $\overline{xy} \subset K$ ,  $P(K) > P(\overline{xy})$ , a contradiction. Thus,  $2D(K) \leq P(K)$ .  $\square$

**Theorem 38.1.48.** *There exist compact convex subsets of  $\mathbb{R}^2$  such that  $2D(K) = P(K)$ .*

*Proof.* For take a straight line segment of length  $\ell$ . Then  $D(K) = \ell$ ,  $P(K) = 2\ell$ , and thus  $2D(K) = P(K)$ .  $\square$

**Theorem 38.1.49.** *If  $K$  is a compact convex subset of  $\mathbb{R}^2$ , then  $P(K) \leq \pi D(K)$ .*

*Proof.* From Cauchy's Perimeter theorem,  $P(K) = \pi W(K) \leq \pi \check{W}(K) = \pi D(K)$ .  $\square$

**Theorem 38.1.50.** *For a compact convex subset  $K$  of  $\mathbb{R}^2$ ,  $P(K) = \pi D(K)$  if and only if  $W_\ell(K)$  is a constant.*

*Proof.* For then  $\frac{1}{2\pi} \int_0^{2\pi} W_{\ell_\theta}(K) d\theta = \check{W}(K)$ . But as  $W_{\ell_\theta}(K)$  is continuous for compact convex bodies, it must be true that  $W_{\ell_\theta}(K) = \check{W}(K)$  for all  $\ell_\theta$ . Thus,  $K$  is of constant width.  $\square$

There are many types of shapes that have constant width besides discs. The Reuleaux Triangle is such an example. Triangles are the simplest convex bodies in the plane other than points and lines. Any convex polygon can be written as the union of triangles with disjoint interiors.

**Theorem 38.1.51.** *If  $\Delta_s$  is an equilateral triangle with edge length  $s$ , then  $\Delta_s$  has the following properties:*

$$1. A(\Delta_s) = \frac{\sqrt{3}}{4}s^2 \quad \Rightarrow. P(\Delta_s) = 3s. W(\Delta_s) = \frac{3s}{\pi}4. R_{\Delta_s} = \frac{1}{\sqrt{3}}s. r_{\Delta_s} = \frac{1}{2\sqrt{3}}s$$

*Proof.* In order:

1. From Pythagoras,  $A(\Delta_s) = 2 \times \left[ \frac{1}{2} \left( \frac{1}{2}s \right) \left( \frac{\sqrt{3}}{2}s \right) \right] = \frac{\sqrt{3}}{4}s^2$
2. There are three edges, each of length  $s$ , and thus  $P(\Delta_s) = 3s$ .
3.  $W(\Delta_s) = \frac{1}{\pi}P(\Delta_s) = \frac{3s}{\pi}$
4. The circumcircle gives the following equations:

$$(a) R_{\Delta_s}^2 = \frac{s^2}{4} + h^2$$

$$(b) h + R_{\Delta_s} = \frac{\sqrt{3}}{2}s$$

This has solution  $R_{\Delta_s} = \frac{1}{\sqrt{3}}s$

$$5. r_{\Delta_s} = \frac{R_{\Delta_s}}{2} = \frac{1}{2\sqrt{3}}s$$

□

**Theorem 38.1.52.** *If  $T$  is a triangle in the plane, then there is a linear transformation  $\psi$  such that  $\psi T$  is equilateral.*

*Proof.* For let  $T$  be a triangle with vertices  $a = (x_1, y_1)$ ,  $b = (x_2, y_2)$ ,  $c = (x_3, y_3)$ . Let  $A = d(b, c)$ ,  $B = d(a, c)$ , and  $C = d(a, b)$ . Suppose If  $A = B = C$ , we are done. Thus, suppose  $C \geq B > A$ . At point  $a$  and with radius  $C$ , construct the circle  $b, c', d$ , and point  $b$  and with radius  $C$ , construct the circle  $a, c', e$ . If we can shift  $c$  to  $c'$  in a linear fashion, we are done. Let  $\psi =$  □

**Theorem 38.1.53.** *If  $T$  is a triangle with edges  $a, b, c$  and opposite angles  $\alpha, \beta, \gamma$ , respectively, then  $A(T) = \frac{\sin(\alpha)}{2}bc = \frac{\sin(\beta)}{2}ac = \frac{\sin(\gamma)}{2}ab$*

*Proof.* Suppose the triangle is acute. The proof is symmetric for all sides, so we prove it for just  $\alpha$ . Note that the perpendicular  $h$  dropped from the vertex of  $a$  onto  $b$  satisfies  $h^2 + \ell_1^2 = c^2$  and  $h^2 + \ell_2^2 = a^2$ , where  $\ell_1 + \ell_2 = b$ . Then  $\sin(\alpha) = \frac{h}{c}$  and  $A(T) = \frac{1}{2}h\ell_1 + \frac{1}{2}h\ell_2 = \frac{h}{2}(\ell_1 + \ell_2)h = \frac{1}{2}bh = \frac{1}{2}bc\sin(\alpha)$ . An identical argument works for obtuse triangles.

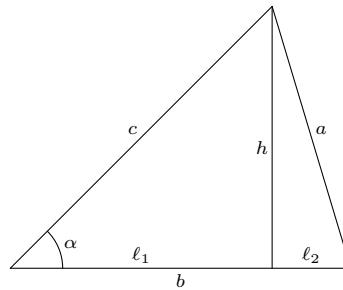


Fig. 38.6: Triangle

□

**Theorem 38.1.54** (The Law of Sines). *For a triangle with edges  $a, b, c$  and opposite angles  $\alpha, \beta, \gamma$ ,  $\frac{\sin(\alpha)}{a} = \frac{\sin(\beta)}{b} = \frac{\sin(\gamma)}{c}$*

*Proof.* From the previous theorem, divide by  $\frac{abc}{2}$  to obtain the result. □

**Theorem 38.1.55** (The Law of Cosines). *Given a triangle with lengths  $a, b, c$  and opposite edges  $\alpha, \beta, \gamma$ ,  $c^2 = a^2 + b^2 - 2ab \cos(\gamma)$ .*

*Proof.* If  $\gamma = \frac{\pi}{2}$ , this is Pythagoras' Theorem. Thus, suppose  $0 < \gamma < \frac{\pi}{2}$ . Where  $c$  and  $a$  meet, drop a perpendicular onto  $c$  and call this  $h$ .  $h$  satisfies  $h^2 + \ell_1^2 = c^2$  and  $h^2 + \ell_2^2 = a^2$  from Pythagoras' Theorem, where  $\ell_1 + \ell_2 = b$ . But  $\ell_2 = a \cos(\gamma)$ , so  $\ell_1 = b - a \cos(\gamma)$ . Thus  $c^2 = h^2 + b^2 + a^2 \cos^2(\gamma) - 2ab \cos(\gamma)$ . But  $h = a \sin(\gamma)$ . Thus,  $c^2 = b^2 + a^2 \cos^2(\gamma) + \sin^2(\gamma) - 2ab \cos(\gamma) = b^2 + a^2(\sin^2(\gamma) + \cos^2(\gamma)) - 2ab \cos(\gamma) = a^2 + b^2 - 2ab \cos(\gamma)$ . A similar construction is done if  $\frac{\pi}{2} < \gamma < \pi$ . □

**Theorem 38.1.56.** *Given a triangle  $\Delta$  with lengths  $a \leq b \leq c$ ,  $D(\Delta) = c$ .*

*Proof.* At the midpoint of  $c$ , and with radius  $c$ , construct a circle. This circle contains the entirety of  $\Delta$ , and thus for all points  $x, y \in \Delta$ ,  $d(x, y) \leq c$ . Thus,  $c = D(\Delta)$ . □

**Theorem 38.1.57** (Heron's Formula). *For a triangle with lengths  $a, b, c > 0$ ,  $A(T)^2 = \frac{1}{16}(a+b+c)(-a+b+c)(a-b+c)(a+b-c)$ .*

*Proof.* For  $A(T) = \frac{1}{2}ab \sin(\gamma) = \frac{1}{2}ab\sqrt{1 - \cos^2(\gamma)} = \frac{1}{4}\sqrt{4a^2b^2 - (a+b-c)^2} = \frac{1}{4}\sqrt{(2ab - (a^2 + b^2 - c^2))(2ab + (a^2 + b^2 + c^2))} = \frac{1}{4}\sqrt{(c^2 - (a-b)^2)((a+b)^2 - c^2)} = \frac{1}{4}\sqrt{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}$ . Squaring this gives the result. □

**Theorem 38.1.58.** *For any triangle  $\Delta$ ,  $r_\Delta P(\Delta) = 2A(\Delta)$ .*

*Proof.* For let  $\Delta$  has sides  $a, b, c$ , with opposite angles  $\alpha, \beta, \gamma$ , respectively. Then  $A(\Delta) = \frac{ab}{2} \sin(\gamma)$  and  $P(\Delta) = a + b + c$ . Thus,  $\frac{2A(\Delta)}{P(\Delta)} = \frac{ab \sin(\gamma)}{a+b+c}$ . But this is the radius of the incircle of  $\Delta$ . Therefore, etc.  $\square$

**Theorem 38.1.59** (Viviani's Theorem: Page 17).

**Theorem 38.1.60.** *There exist convex polygon's such that the inradius is not unique.*

*Proof.* For consider the rectangle  $[0, 2] \times [0, 1]$ . The diameter of any circle that sits inside this body must be at most 1, and thus the radius is at most  $\frac{1}{2}$ . However, there are multiple circles that achieve this. For example  $(x - \frac{1}{2})^2 + (y - \frac{1}{2})^2 = \frac{1}{2}$  and  $(x - \frac{3}{2})^2 + (y - \frac{3}{2})^2 = \frac{1}{2}$ .  $\square$

**Theorem 38.1.61.** *For an acute triangle,  $\frac{2(A)}{abc} = \frac{1}{R_T}$ .*

**Theorem 38.1.62.** *If  $\Delta$  is a triangle with vertices  $A, B, C$  and sides  $a, b, c$ , then given a circle that contains  $A, B, C$ , the radius of this circle  $R$  satisfies  $\frac{1}{R} = \frac{2A(\Delta)}{abc}$*

**Definition 38.1.18** If  $K \in \mathcal{K}_2$ , then  $-K = \{-x : x \in K\}$ .

**Theorem 38.1.63.** *If  $K \in \mathcal{K}_2$ , then  $W_\ell(K) = W_\ell(-K)$ .*

*Proof.* For let  $x, y \in K$  be such that  $d(x, y) = W_\ell(K)$ . Then  $d(-x, -y) = W_\ell(-K) = d(x, y)$ . Therefore, etc.  $\square$

**Theorem 38.1.64.** *There exist sets  $K$  and  $L$  such that  $W_\ell(K) < W_\ell(L)$  for all  $\ell$  yet  $K \not\subset L$  for any translation or rotation of  $K$ .*

**Definition 38.1.19** For  $K \in \mathcal{K}$  and unit vector  $u$ ,  $\bar{X}_u(K)$  is the mean value of  $X_\ell(K)$  over all lines  $\ell$  parallel to  $u$  such that  $X_\ell(K) > 0$ .

**Theorem 38.1.65.** *For  $K \in \mathcal{K}_2$ ,  $\bar{X}_u(K) = W_\ell(K) = A(K)$ .*

## 38.2 On Uniform Convergence

**Definition 38.2.1** A sequence of functions  $f_n$  is said to converge point-wise on a set  $A$  to a function  $f$ , if  $\forall \varepsilon > 0$  and  $\forall x \in A$ , there is an  $N \in \mathbb{N}$  such that  $n > N \Rightarrow |f(x) - f_n(x)| < \varepsilon$ .

**Definition 38.2.2** A sequence of functions  $f_n$  converge uniformly on a set  $A$  to  $f$  if and only if  $\forall \varepsilon > 0 \exists N \in \mathbb{N}$  such that  $\forall x \in A$  and  $n > N$ ,  $|f(x) - f_n(x)| < \varepsilon$ .

**Definition 38.2.3** A sequence of functions  $f_n$  are point-wise equicontinuous on a set  $A$  if and only if  $\forall \varepsilon > 0 \forall x \in A \exists \delta > 0 \forall n \in \mathbb{N} : |x - x_0| < \delta \Rightarrow |f_n(x) - f_n(x_0)| < \varepsilon$

**Definition 38.2.4** A sequence of functions  $f_n$  are uniformly equicontinuous on a set  $A$  if and only if  $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in A \forall n \in \mathbb{N} : |x - x_0| < \delta \Rightarrow |f_n(x) - f_n(x_0)| < \varepsilon$ .

**Definition 38.2.5** A subset  $A$  of the real line is open if  $\forall x \in A \exists r > 0 : \forall y \in (x-r, x+r), y \in A$ .

**Definition 38.2.6** An open cover  $\Delta$  of a set  $A \subset S$  is a set of open subsets  $A_k \subset S$ , such that  $A \subset \bigcup_{k \in I} A_k$ , where  $I$  is some index, countable or uncountable.

**Definition 38.2.7** A set  $A$  is said to be compact if and only if for all open coverings  $\Delta$  there is a finite sub-cover  $\Delta_0 \subset \Delta$ , such that  $A \subset \bigcup_{A_k \in \Delta_0} A_k$ .

**Theorem 38.2.1** (Heine-Borel Theorem). *Any closed-bounded subset of the real line is compact.*

*Proof.* For let  $A$  be a closed and bounded subset of  $\mathbb{R}$  with least upper bound  $b$  and greatest lower bound  $a$ . Let  $\Delta$  be an open covering, and let  $X$  be the set of points  $y \in A$  such that for all  $s < y$  such that  $s \in A$ , there is a finite refinement of  $\Delta$  which covers these points.  $X$  is non-empty, as  $a \in X$ . The set  $X$  is bounded, as for all points  $y \in X$  we have that  $a \leq y \leq b$ . As bounded sets have a least upper bound, let  $x$  be the least upper bound. Suppose  $x < b$ . As  $x \in [a, b]$ , there exists an element  $A_k$  of  $\Delta$  such that  $x \in A_k$ . But as  $A_k$  is open and therefore there is an  $r > 0$  such that  $y \in (x - r, x + r) \Rightarrow y \in A_k$ . But then  $y = x + \frac{r}{2} > x$  and  $y \in A_k$ . Therefore  $x$  is not the least upper bound as we have found an element in  $X$  greater than  $x$ . Therefore  $x \not< b$ . And thus  $x = b$ .  $\square$

**Theorem 38.2.2.** *If a sequence of functions are point-wise equicontinuous on a closed and bounded set, then they are uniformly equicontinuous.*

*Proof.* For let  $A$  be a closed bounded subset of  $\mathbb{R}$ , and let  $f_n(x)$  be a sequence of point-wise equicontinuous functions on  $A$ . As the set is closed and bounded, it is compact by the Heine-Borel theorem. Let  $\varepsilon > 0$  be given. For  $x \in A$ , define the function  $\delta(x) = \min\{\sup\{\delta > 0 : |x - x_0| < \delta, x_0 \in A \Rightarrow |f_n(x) - f_n(x_0)| < \frac{\varepsilon}{2}, \forall n \in \mathbb{N}\}, b - a\}$ . Construct the open covering  $\mathcal{U}$  as follows:  $\mathcal{U} = \{(x - \delta(x), x + \delta(x)) : x \in A\}$ . This is an open covering, as every set in  $\mathcal{U}$  is open, and for all  $x \in A$ ,  $x \in (x - \delta(x), x + \delta(x)) \in \mathcal{U}$ . But as  $A$  is compact, there is a finite sub-cover. Let  $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$  be the centers of the remaining open sets in the sub-cover. Further refine this sub-covering as follows: If  $(x_j - \delta(x_j), x_j + \delta(x_j)) \subset (x_k - \delta(x_k), x_k + \delta(x_k))$  for  $j \neq k$ , then remove it from the sub-cover as it is superfluous. We now have a set of points  $a = z_0 < z_1 < \dots < z_{N-1} < z_N = b$  such that  $A \subset \bigcup_{i=0}^N (z_i - \delta(z_i), z_i + \delta(z_i))$ . Let  $\delta = \min\{\delta(z_0), \dots, \delta(z_N), \delta(b), \frac{(a+\delta(a))-(z_1-\delta(z_1))}{2}, \dots, \frac{(z_{N-1}+\delta(z_{N-1}))-(b-\delta(b))}{2}\}$ . That is,  $\delta$  is the smallest of the  $\delta(z_i)$ , or half of the smallest intersection of two consecutive intervals. Let  $x \in A$  be arbitrary. If  $(x - \delta, x + \delta)$  is contained

entirely in one of the  $(z_i - \delta(z_i), z_i + \delta(z_i))$  sets, then we have that  $|x - x_0| < \delta \Rightarrow |x - x_0| < \delta(z_i) \Rightarrow |f_n(x) - f_n(x_0)| < \frac{\varepsilon}{2}$  for all  $n \in \mathbb{N}$ . Suppose that  $(x - \delta, x + \delta)$  is contained in two of the  $(x - \delta(z_i), x + \delta(z_i))$  sets. Note, it cannot be in three or more as we have refined the sub-cover in such a manner as to prevent this. Let  $y$  be the center of the intersection of these two sets. Then we have that for  $|x - x_0| < \delta$ , then  $|f_n(x) - f_n(x_0)| = |f_n(x) - f_n(y) + f_n(y) - f_n(x_0)| \leq |f_n(x) - f_n(y)| + |f_n(y) - f_n(x_0)|$ . But  $|x - y|$  and  $|x_0 - y|$  are less than  $\frac{(z_i+\delta(z_i))-(z_{i+1}-\delta(z_{i+1}))}{2}$  apart, and therefore  $|f_n(x) - f_n(y)| < \frac{\varepsilon}{2}$ , and  $|f_n(y) - f_n(x_0)| < \frac{\varepsilon}{2}$ . Therefore,  $|f_n(x) - f_n(x_0)| < \varepsilon$ . And as  $x$  is arbitrary,  $f_n(x)$  is uniformly equicontinuous.  $\square$

**Theorem 38.2.3.** *If a sequence of point-wise equicontinuous functions converge, then the limit is point-wise continuous.*

*Proof.* For let  $f_n : A \rightarrow \mathbb{R}$  be equicontinuous,  $\varepsilon > 0$  and  $x \in A$  be given. Choose  $\delta > 0$  to satisfy the criterion of equicontinuity at  $x$ . Let  $x_0$  be an arbitrary point in  $(x - \delta, x + \delta) \cap A$ . It suffices to show that  $|f(x) - f(x_0)| < \varepsilon$ . As  $f_n \rightarrow f$  we have that  $\exists N_1 \in \mathbb{N}$  such that  $n > N_1 \Rightarrow |f(x) - f_n(x)| < \varepsilon$ . We also have that  $\exists N_2 \in \mathbb{N}$  such that  $n > N_2 \Rightarrow |f(x_0) - f_n(x_0)| < \varepsilon$ . Let  $N = \max\{N_1, N_2\} + 1$ . But we have that  $|f(x) - f(x_0)| = |f(x) - f_N(x) + f_N(x) - f_N(x_0) + f_N(x_0) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < 3\varepsilon$ .  $f$  is continuous.  $\square$

**Theorem 38.2.4.** *If  $f_n \rightarrow f$  on a closed bounded subset of  $\mathbb{R}$ , and if  $f_n$  is equicontinuous, then the convergence is uniform.*

*Proof.* Let  $A$  be a closed bounded subset of  $\mathbb{R}$ ,  $f_n(x)$  a sequence of equicontinuous functions, and let  $\varepsilon > 0$  be given. As  $f_n(x)$  is equicontinuous on a closed bounded set, it is uniformly equicontinuous. But the limit of equicontinuous functions is continuous. Let  $\delta > 0$  be such that,  $\forall x \in A, \forall n \in \mathbb{N}, |x - x_0| < \delta, x_0 \in A \Rightarrow |f_n(x) - f_n(x_0)| < \frac{\varepsilon}{3}$  and  $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \frac{\varepsilon}{3}$ . Let  $\mathcal{U} = \{(x - \frac{\delta}{2}, x + \frac{\delta}{2}) : x \in A\}$ . This is an open cover of  $A$  and thus there is a finite subcover. Let  $x_0 < x_1 < \dots < x_n$  be the centers of the finitely many sets  $(x_k - \frac{\delta}{2}, x_k + \frac{\delta}{2})$  that cover  $A$ . There is thus another finite sequence of positive integers,  $N_0, N_1, \dots, N_n$ , such that  $n > N_k \Rightarrow |f(x_k) - f_n(x_k)| < \frac{\varepsilon}{3}$ , for  $k = 0, 1, 2, \dots, n$ . Let  $N = \max\{N_0, N_1, \dots, N_n\}$ . It suffices to show that, for any point  $x_0 \in A$ , for all  $n > N$ ,  $|f(x_0) - f_n(x_0)| < \varepsilon$ . Let  $x_0$  be arbitrary and let  $x_k$  be the nearest point to  $x_0$  in the above sequence (If there are two nearest points, pick your favorite). Then we have that, for  $n > N$ ,  $|f(x_0) - f_n(x_0)| = |f(x_0) - f(x_k) + f(x_k) - f_n(x_k) + f_n(x_k) - f_n(x_0)| \leq |f(x_k) - f(x_0)| + |f(x_k) - f_n(x_k)| + |f_n(x_k) - f_n(x_0)| < \varepsilon$ . The convergence is uniform.  $\square$

**Theorem 38.2.5** (Integration of a Uniformly Convergent Sequence of Functions). *If  $f_n \rightarrow f$  uniformly on a closed bounded set  $A$  with  $g.u.b(A) = a$ , then  $\int_a^x f_n \rightarrow \int_a^x f$  uniformly on  $A$ .*

*Proof.* Let  $\varepsilon > 0$  be given, let  $b = l.u.b.(A)$ , and choose  $N \in \mathbb{N}$  such that  $n > N \Rightarrow |f(x) - f_n(x)| < \frac{\varepsilon}{b-a}$ . Then we have  $|\int_a^x f_n - \int_a^x f| = |\int_a^x (f_n - f)| \leq \int_a^x |f_n - f| < \int_a^x \frac{\varepsilon}{b-a} = \frac{\varepsilon}{b-a}(x - a) \leq \varepsilon$ .  $\square$

**Theorem 38.2.6** (Differentiation of a Uniformly Convergent Sequence of Functions). *If  $f'_n \rightarrow g$  uniformly on a closed bounded set  $A$ , and if  $f_n \rightarrow f$  on  $A$ , then  $f' = g$ .*

*Proof.* Let  $a = g.u.b.(A)$  and  $b = l.u.b.(A)$ . We have that  $f_n(x) - f_n(a) = \int_a^x f'_n \rightarrow \int_a^x g$  uniformly. But  $f_n(x) - f_n(a) \rightarrow f(x) - f(a)$ . Therefore  $f'(x) = \frac{d}{dx}(f(x) - f(a)) = \frac{d}{dx} \int_a^x g = g(x)$ .  $f' = g$ .  $\square$

**Theorem 38.2.7** (The Product of a Uniformly Convergent Sequence and a Bounded Function). *If  $f_n \rightarrow f$  uniformly, and if  $g$  is a bounded function, then  $f_n g \rightarrow fg$  uniformly.*

*Proof.* For let  $\varepsilon > 0$  and  $x$  be given, and let  $g$  be a bounded function with bound  $M$ , and choose  $N \in \mathbb{N}$  such that  $n > N \Rightarrow |f(x) - f_n(x)| < \frac{\varepsilon}{M}$ . Then we have that  $|f(x)g(x) - f_n(x)g(x)| = |g(x)||f(x) - f_n(x)| < M|f(x) - f_n(x)| < \varepsilon$ .  $\square$

**Theorem 38.2.8.** *If  $f$  is continuous on a compact set  $A$ , then it is uniformly continuous.*

*Proof.* For let  $\varepsilon > 0$  be given, let  $a = g.u.b.(A)$ ,  $b = l.u.b.(A)$ , and for  $x \in A$  define  $\delta(x) = \min\{\sup\{\delta > 0 : |x-x_0| < \delta, x_0 \in A \Rightarrow |f(x)-f(x_0)| < \frac{\varepsilon}{2}\}, b-a\}$ . Let  $\Delta = \{(x - \delta(x), x + \delta(x)) : x \in A\}$ . Then  $\Delta$  is an open cover of  $A$  and therefore there is an open subscover. Let  $x_k$  be the centers of the finitely many sets  $(x_k - \delta(x_k), x_k + \delta(x_k))$  that cover  $A$ . Further refine this by removing overlaps. That is, if  $(x_i - \delta(x_i), x_i + \delta(x_i)) \subset (x_j - \delta(x_j), x_j + \delta(x_k))$  for  $i \neq j$ , then remove it for it is superfluous. We thus obtain a new sequence  $z_1, \dots, z_N$  such that the intervals  $(z_k - \delta(z_k), z_k + \delta(z_k))$  cover  $A$ . Define  $\delta = \min\{\delta(z_1), \dots, \delta(z_N), \frac{(z_0 + \delta(z_0)) - (z_1 - \delta(z_1))}{2}, \dots, \frac{(z_{N-1} + \delta(z_{N-1})) - (z_N - \delta(z_N))}{2}\}$ . Let  $x, x_0 \in A$  such that  $|x - x_0| < \delta$ . Let  $x_k$  be the closest point in the sequence to  $x$  (If there are two such points, pick your favorite). Then  $|f(x) - f(x_0)| = |f(x) - f(x_k) + f(x_k) - f(x_0)| \leq |f(x) - f(x_k)| + |f(x_k) - f(x_0)| < \varepsilon$ .  $\square$

The proof of this is a mimicry of the proof that equicontinuity on a compact set implies uniform equicontinuity.

**Definition 38.2.8** A set  $A$  is called sequentially compact if given a sequence  $x_n \in A$ , there is a convergent subsequence  $x_{n_k}$ .

**Theorem 38.2.9.** *Compact sets of  $\mathbb{R}$  are sequentially compact.*

*Proof.* Let  $A$  be a compact set in  $\mathbb{R}$ , and let  $x_n$  be a sequence in  $A$ . A point  $x \in A$  is the limit of a subsequence of  $x_n$  if for every  $\varepsilon > 0$  there are infinitely many of the  $x_n$  such that  $|x - x_n| < \varepsilon$ . Suppose there is no such point.

That is, for each  $x \in A$  only finitely many of the  $x_n$  lie within sufficiently small  $\varepsilon$ -neighborhoods. Let  $\varepsilon(x) = \sup\{\varepsilon > 0 : \text{Only finitely many } x_n \text{ lie within } \varepsilon \text{ of } x\}$ . Define  $E = \{(x - \varepsilon(x), x + \varepsilon(x)) : x \in A\}$ . This is an open cover of  $A$ , and therefore there is a finite subcover. Thus, at least one of the finitely many intervals  $(x - \varepsilon(x), x + \varepsilon(x))$  must contain infinitely many of the  $x_n$ , a contradiction. Thus there is a convergent subsequence.  $\square$

**Theorem 38.2.10.** *Continuous functions on compact sets are bounded.*

*Proof.* For suppose not. Let  $f : A \rightarrow \mathbb{R}$  be a continuous function on a compact set  $A$ , and let  $x_n$  be a sequence of points in  $A$  such that  $f(x_n) > n$ . Such a sequence must exist as  $f$  is not bounded. As  $A$  is compact, there must a point  $x \in A$  such that some subsequence  $x_{n_k}$  that converges to  $x$ . Let  $\varepsilon > 0$ . Then, as  $f$  is continuous, there is a  $\delta > 0$  such that  $|x - x_0| < \delta$ ,  $x_0 \in A \Rightarrow |f(x) - f(x_0)| < \varepsilon$ . But then for all points  $x_{n_k}$  such that  $|x - x_{n_k}| < \delta$ ,  $-\varepsilon < f(x_{n_k}) - f(x) < \varepsilon \Rightarrow f(x) - \varepsilon < f(x_{n_k}) < f(x) + \varepsilon$ . A contradiction as  $f(x_{n_k})$  is unbounded. Thus,  $f$  is bounded.  $\square$

**Theorem 38.2.11.** *Continuous functions on compact sets attain their maximum and minimum.*

*Proof.* For let  $f : A \rightarrow \mathbb{R}$  be a continuous function on a compact set  $A$ . Let  $f(A) = \{y \in \mathbb{R} : \exists x \in A \mid f(x) = y\}$ . (This is called the image of  $A$  under  $f$ ). As  $f$  is continuous, it is bounded, and thus the set  $f(A)$  is bounded. But bounded sets have an l.u.b. and a g.u.b. Therefore, etc.  $\square$

**Theorem 38.2.12** (Uniform Limit Theorem). *If  $f_n \rightarrow f$  uniformly, and if the  $f_n$  are continuous, then  $f$  is continuous.*

*Proof.* For let  $\varepsilon > 0$  be given and let  $x \in A$ . Let  $N \in \mathbb{N}$  such that  $n > N$  implies  $|f(\chi) - f_n(\chi)| < \frac{\varepsilon}{3}$  for all  $\chi \in A$ . Let  $\delta > 0$  be chosen such that  $|x - x_0| < \delta$ ,  $x_0 \in A \Rightarrow |f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3}$ . Then  $|f(x) - f(x_0)| = |f(x) - f_N(x) + f_N(x) - f_N(x_0) + f_N(x_0) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f(x_0) - f_N(x_0)| < \varepsilon$ .  $\square$

**Theorem 38.2.13.** *If  $f_n g \rightarrow fg$  uniformly on a compact set  $A$ , and if  $g$  is continuous and positive, then  $f_n \rightarrow f$  uniformly.*

*Proof.* As  $g$  is positive on a compact set, its minimum is also positive and is attained on  $A$ . Let  $x_{min} \in A$  be such a minimum of  $g$ . Let  $\varepsilon > 0$  be given and let  $N \in \mathbb{N}$  be such that for  $n > N$ ,  $|f_n g - fg| < \varepsilon \cdot g(x_{min})$ . Then,  $|f_n g - fg| = |g||f_n - f| \leq |g(x_{min})||f_n - f| < \varepsilon \cdot g(x_{min}) \Rightarrow |f_n - f| < \varepsilon$ .  $\square$

**Theorem 38.2.14.** *If  $f'_n$  is uniformly bounded, then  $f_n$  is equicontinuous.*

*Proof.* For let  $M$  be such a bound for  $f'_n$  and let  $\varepsilon > 0$  be given. Choose  $\delta = \frac{\varepsilon}{M}$ . Then for  $x, x_0 \in A$  and  $|x - x_0| < \delta$ ,  $|\int_{x_0}^x f'_n| = |f_n(x) - f_n(x_0)| \leq \int_{x_0}^x |f'_n| \leq (x - x_0)M < \varepsilon$ .  $\square$

**Theorem 38.2.15.** *If  $f'_n$  is uniformly bounded, and if  $f_n \rightarrow f$  on a closed and bounded subset of  $\mathbb{R}$ , then the convergence is uniform.*

*Proof.* From the previous lemma,  $f_n$  is equicontinuous. But a sequence of equicontinuous functions on a compact set is uniformly equicontinuous. And a sequence of uniformly equicontinuous functions that converge does so uniformly. Therefore, etc.  $\square$

**Theorem 38.2.16.** *If  $f_n \rightarrow f$ ,  $f'_n \rightarrow g$  and if  $f''_n - f'_n$  is uniformly bounded on a closed bounded set, then the convergences are uniform and  $f' = g$ .*

*Proof.* Let  $A$  be the closed bounded set under consideration. First note that as  $f''_n - f'_n$  is uniformly bounded,  $f'_n - f_n$  is equicontinuous. But as  $f'_n$  and  $f_n$  converge to  $g$  and  $f$ , respectively, then  $f'_n - f_n$  converges to  $g - f$  uniformly. Let  $M$  be a bounded for  $f''_n - f'_n$ . Let  $a$  be the greatest lower bound and  $b$  be the least upper bound of  $A$ . We then have that  $-Me^{-a} \leq e^{-x}[f''_n(x) - f'_n(x)] = \frac{d}{dx}[e^{-x}f'_n(x)] \leq Me^{-a}$ . That is,  $\frac{d}{dx}[e^{-x}f'_n(x)]$  is uniformly bounded, and therefore  $e^{-x}f'_n(x)$  is equicontinuous. But equicontinuity on a compact set implies uniform equicontinuity. As  $f'_n \rightarrow g$ , and  $e^{-x}$  is bounded on  $A$ ,  $e^{-x}f'_n \rightarrow e^{-x}g$ . But a convergent uniformly equicontinuous sequence of functions converges uniformly. Thus,  $e^{-x}f'_n(x) \rightarrow e^{-x}g(x)$  uniformly, and therefore, as  $e^{-x}$  is continuous and positive on  $A$ ,  $f'_n(x) \rightarrow g(x)$  uniformly. But also  $f'_n - f_n \rightarrow g - f$  uniformly, and therefore  $f_n \rightarrow f$  uniformly. Thus,  $f' = g$ .  $\square$

**Theorem 38.2.17.** *If  $f'_n - f_n$  is uniformly bounded and if  $f_n \rightarrow f$  on a closed and bounded set  $A$ , then the convergence is uniform.*

*Proof.* Using the inequality from the previous theorem, let  $M$  be a bound for  $f'_n - f_n$  and let  $a$  be the least upper bound of  $A$ . Then  $-Me^{-a} \leq \frac{d}{dx}[e^{-x}f_n] \leq Me^{-a}$ . Thus  $e^{-x}f_n$  is uniformly equicontinuous and therefore  $e^{-x}f_n \rightarrow e^{-x}f$  uniformly, and thus  $f_n \rightarrow f$  uniformly.  $\square$

**Theorem 38.2.18.** *If  $f_n^{(N+1)} - f_n^{(N)}$  is bounded and a compact set, and if  $f_n^{(k)} \rightarrow f_k$  for  $k = 0, 1, \dots, N$ , then the convergence is uniform and  $f'_k = f_{k+1}$  for  $k = 0, 1, \dots, N - 1$ .*

*Proof.* A simple induction and application of the previous theorem proves this.  $\square$

### 38.3 On Analyticity

We deal with functions on intervals for simplicity.

**Definition 38.3.1** A real-valued function  $f$  is said to be smooth, denoted  $f \in C^\infty$  if, for all  $k$ ,  $\frac{d^k}{dx^k} f(x) \equiv f^{(k)}(x)$  exists.

**Theorem 38.3.1** (Taylor's Theorem). *If  $f \in C^\infty$ , on some interval  $[a, b]$ , and if  $x_0 \in (a, b)$ , then  $f(x) - \sum_{k=0}^n f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} = \int_{x_0}^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$*

*Proof.* We prove by induction. The base case says  $f(x) - f(x_0) = \int_{x_0}^x f'(t) dt$ , which is true. Suppose it holds for some  $n \in \mathbb{N}$ . Then  $f(x) - \sum_{k=0}^{n+1} f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} = f(x) - \sum_{k=0}^n f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} - f^{(n+1)}(x) \frac{(x-x_0)^{n+1}}{(n+1)!} = \int_{x_0}^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt - f^{(n+1)}(x) \frac{(x-x_0)^{n+1}}{(n+1)!}$ . But  $\int_{x_0}^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt = \int_{x_0}^x f^{(n+2)}(t) \frac{(x-t)^{n+1}}{(n+1)!} dt + f^{(n+1)}(x) \frac{(x-x_0)^{n+1}}{(n+1)!}$  from integration by parts. Thus,  $f(x) - \sum_{k=0}^{n+1} f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} = \int_{x_0}^x f^{(n+2)}(t) \frac{(x-t)^{n+1}}{(n+1)!} dt$   $\square$

**Theorem 38.3.2.** *If  $f \in C^\infty$  and  $f^{(n)}(x) \rightarrow 0$  (Point-wise) on  $[a, b]$ , and if  $F(x) \equiv f(x) - \sum_{k=0}^{\infty} f^{(k)}(x_0) \frac{(x-x_0)^k}{k!}$ , where  $x_0 \in [a, b]$  is fixed, then  $\int_{x_0}^x F^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$  converges.*

*Proof.* For let  $x, x_0 \in [a, b]$  fixed. We will show that  $\int_{x_0}^x F^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$  is Cauchy. Let  $\varepsilon > 0$ ,  $N_0 = 1$ , and let  $n > m > N_0$  be arbitrary. We have that  $F(x) = \left( f(x) - \sum_{k=0}^N f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} \right) - \left( g(x) - \sum_{k=0}^N f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} \right)$ , where  $N \in \mathbb{N}$  is arbitrary. From Taylor's Theorem we thus have  $F(x) = \int_{x_0}^x F^{N+1}(t) \frac{(x-t)^N}{N!} dt$ . Then  $|\int_{x_0}^x F^{n+1}(t) \frac{(x-t)^n}{n!} dt - \int_{x_0}^x F^{m+1}(t) \frac{(x-t)^m}{m!} dt| = |F(x) - F(x)| = 0 < \varepsilon$ .  $\square$

**Theorem 38.3.3.** *If  $f \in C^\infty$  and  $f^{(n)}(x) \rightarrow 0$  (Point-wise) on some interval  $[a, b]$ , then  $f^{(n)}(x)$  is uniformly bounded.*

*Proof.* For let  $x_0 \in (a, b)$  be arbitrary. As  $f^{(n)}(x_0) \rightarrow 0$ ,  $\sum_{k=0}^{\infty} f^{(k)}(x_0) \frac{(x-x_0)^k}{k!}$  converges everywhere. Let  $g(x) \equiv \sum_{k=0}^{\infty} f^{(k)}(x_0) \frac{(x-x_0)^k}{k!}$ . Define  $F(x) = f(x) - g(x)$ . Then:

$$F^{(n)}(x) = f^{(n)}(x) - g^{(n)}(x)$$

$$= \left( f^{(n)}(x) - \sum_{k=n}^N f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} \right) - \left( g^{(n)}(x) - \sum_{k=n}^N f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} \right)$$

From Taylor's theorem, this is equal to:

$$\begin{aligned} \int_{x_0}^x f^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt - \int_{x_0}^x g^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt \\ = \int_{x_0}^x F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt \end{aligned}$$

That is, for all  $N > n$ ,  $F^{(n)}(x) = \int_{x_0}^x F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt$ . But for all  $x_1 \in (a, b)$ :

$$F^{(n)}(x) - \sum_{k=n}^N F^{(k)}(x_1) \frac{(x-x_1)^k}{k!} = \int_{x_1}^x F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt$$

Now, suppose  $f^{(n)}(x)$  is not uniformly bounded.  $g^{(n)}(x)$  is uniformly bounded by its definition, and thus  $F^{(n)}(x)$  is not uniformly bounded. Let  $k_n$  be a subsequence of  $n$  such that  $F^{(k_n)}(x_{k_n}) > n$ . Such a sequence exists as  $F^{(n)}(x)$  is not uniformly bounded. As  $[a, b]$  is closed and bounded, it is compact. Thus  $x_{k_n}$  has a convergent subsequence  $\varphi(x_{k_n})$  (We use this notation so as to avoid writing  $x_{k_m n}$ ). Let  $x_1$  be the limit of this subsequence. As  $F^{(n)}(x_1) \rightarrow 0$ ,  $\sum_{k=n}^N F^{(k)}(x_1) \frac{(x-x_1)^k}{k!}$  converges. Let  $M$  be a bound for  $F^{(k)}(x_1)$ . Such a bound exists as this sequence converges. As  $F^{(n)}(x) = \int_{x_0}^x F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt$ , we have that:

$$\sum_{k=n}^N F^{(k)}(x_1) \frac{(x-x_1)^k}{k!} = - \int_{x_0}^{x_1} F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt$$

Thus, for all  $n$  and  $N$ :

$$\left| \int_{x_0}^{x_1} F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt \right| \leq M e^{b-a}$$

Thus, we have that:

$$\begin{aligned} |F^{(n)}(x)| &= \left| \int_{x_0}^x F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt \right| \\ &= \left| \int_{x_0}^{x_1} F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt + \int_{x_1}^x F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt \right| \\ &\leq M e^{b-a} + \left| \int_{x_1}^x F^{(N+n+1)}(t) \frac{(x-t)^{N+n}}{(N+n)!} dt \right| \end{aligned}$$

But as  $N$  is arbitrary, we may take it to be large enough to make the latter term close to a fixed finite value for each point. Thus  $F^{(n)}(\varphi(x_{k_n})) \not\rightarrow \infty$  and therefore  $F^{(n)}(x)$  is not unbounded, and is therefore uniformly bounded. Thus  $f^{(n)}(x)$  is uniformly bounded.  $\square$

**Definition 38.3.2** An analytic function about a point  $x_0$  is a function  $f : \mathcal{U} \rightarrow \mathbb{R}$  such that  $f(x) = \sum_{n=0}^{\infty} f^n(x_0) \frac{(x-x_0)^n}{n!}$  for all  $x \in \mathcal{U}$ .

**Theorem 38.3.4** (Lagrange's Remainder Theorem). *A function  $f(x)$  is analytic if and only if  $\int_{x_0}^x f^{n+1}(t) \frac{(x-t)^n}{n!} dt \rightarrow 0$ .*

*Proof.* For if  $f(x)$  is analytic, then  $f(x) - \sum_{k=0}^n f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} = \int_{x_0}^x f^{n+1}(t) \frac{(x-t)^n}{n!} dt \rightarrow 0$ . If  $\int_{x_0}^x f^{n+1}(t) \frac{(x-t)^n}{n!} dt \rightarrow 0$ , then  $f(x) - \sum_{k=0}^n f^{(k)} \frac{(x-x_0)^k}{k!} \rightarrow 0$ , and thus  $f(x)$  is analytic.  $\square$

**Theorem 38.3.5.** *If  $f \in C^\infty$  and  $f^{(n)}$  is uniformly bounded, then it is analytic.*

*Proof.* For  $|\int_{x_0}^x f^{n+1}(t) \frac{(x-t)^n}{n!} dt| \leq \int_{x_0}^x |f^{n+1}(t)| \frac{(x-t)^n}{n!} dt$ . As  $f^{(n)}(x)$  is uniformly bounded, and for all  $x \frac{(x-x_0)^n}{n!} \rightarrow 0$ , we have that  $\int_{x_0}^x f^{n+1}(t) \frac{(x-t)^n}{n!} dt \rightarrow 0$ .  $\square$

**Theorem 38.3.6.** *If  $f^{(n)}(x) \rightarrow 0$ , then  $f$  is analytic.*

*Proof.* For  $f^{(n)}(x)$  is thus uniformly bounded, and therefore analytic.  $\square$

## 38.4 On Infinite Order O.D.E.'s

**Definition 38.4.1** An infinite order O.D.E. is a differential equation with no largest order of derivative.

An infinite order O.D.E. then necessarily has an infinite number of terms.

**Definition 38.4.2** A linear infinite order O.D.E. is a differential equation of the form  $\sum_{n=0}^{\infty} a_n(x) \frac{d^n f}{dx^n} = F(x)$ .

Unlike normal differential equation of order  $n \in \mathbb{N}$ , infinite order differential equations have the problem of convergence. That is,  $\sum_{n=0}^{\infty} a_n(x) \frac{d^n f}{dx^n} = F(x)$  may have a different solution set if point-wise convergence is considered rather than uniform. We now consider the main topic of the paper.

**Theorem 38.4.1.** *Consider the following differential equation on some interval  $(a, b)$ :*

$$\sum_{n=0}^{\infty} \frac{d^n f}{dx^n} = 0$$

*Be the convergence uniform or point-wise, the only solution is  $f(x) = 0$*

We will prove this via the tools we have developed in the previous sections. First, some preliminary results.

**Theorem 38.4.2.** *If, for some open set  $A$ ,  $f : A \rightarrow \mathbb{R}$  is continuous and positive at some point  $x_0$ , then there exists and open interval  $(a, b)$  that contains  $x_0$  such that  $f(x) > 0$  on this interval.*

*Proof.* For let  $A$  be open, let  $f : A \rightarrow \mathbb{R}$  be continuous, and let  $x_0 \in A$  be such that  $f(x_0) > 0$ . Let  $\varepsilon = f(x_0) > 0$ . As  $f$  is continuous, there is a  $\delta > 0$  such that  $|x - x_0| < \delta$  and  $x \in A$  implies  $|f(x_0) - f(x)| < \varepsilon = f(x_0)$ . As  $A$  is open and  $x_0 \in A$  there is an  $r > 0$  such that  $(x_0 - r, x_0 + r) \in A$ . Then  $(x_0 - r, x_0 + r) \cap (x_0 - \delta, x_0 + \delta)$  is an open interval in  $A$  such that  $0 < f(x) < 2f(x_0)$ .  $\square$

**Theorem 38.4.3** (The Fundamental Theorem of the Calculus of Variations). *If  $f$  is a continuous function on  $(a, b)$ , and if for all  $\alpha, \beta \in (a, b)$   $\int_{\alpha}^{\beta} f = 0$ , then  $f = 0$ .*

*Proof.* For suppose not. Let  $f$  be positive at some point  $x$ . Then, as  $f$  is continuous, there is a  $\delta > 0$  such that for all  $x_0 \in (x - \delta, x + \delta) \cap (a, b)$ ,  $f(x_0) > 0$ . But then the integral on this subinterval is positive, a contradiction. Thus  $f = 0$ .  $\square$

**Theorem 38.4.4** (Cauchy Criterion). *If  $\sum a_n$  converges, then  $a_n \rightarrow 0$ .*

*Proof.* For let  $s_n$  be the  $n^{th}$  partial sum. As convergent sequence are Cauchy sequences,  $s_{n+1} - s_n \rightarrow 0 \Rightarrow a_{n+1} \rightarrow 0$ .  $\square$

**Theorem 38.4.5.** *If  $\sum_{n=0}^N \frac{d^n f}{dx^n} \rightarrow 0$  uniformly on some interval  $(a, b)$ , then  $f = 0$ .*

*Proof.* For any  $\alpha, \beta \in (a, b)$ ,  $\int_{\alpha}^{\beta} \sum_{n=0}^N \frac{d^n f}{dx^n} \rightarrow \int_{\alpha}^{\beta} 0 = 0$ . Thus,  $\int_{\alpha}^{\beta} f + \sum_{n=0}^{N-1} \frac{d^n f}{dx^n} \Big|_{\alpha}^{\beta} \rightarrow 0$ . As the latter term tends to 0,  $\int_{\alpha}^{\beta} f = 0$ . As  $\alpha$  and  $\beta$  are arbitrary,  $f = 0$ .  $\square$

**Theorem 38.4.6.** *If  $\sum_{n=0}^N \frac{d^n f}{dx^n} \rightarrow 0$  point-wise on some interval  $(a, b)$ , then  $f = 0$ .*

*Proof.* Suppose not. Let  $x \in (a, b)$  be such that  $f(x) \neq 0$ . Consider the interval  $[\frac{a+x}{2}, \frac{x+b}{2}] = [\alpha, \beta]$  and let  $S_N = \sum_{n=0}^N \frac{d^n f}{dx^n}$ . Note that  $S'_N = S_{N+1} - f$ . So  $S'_N - S_N = f^{(n+1)} - f$ , and thus  $|S'_N - S_N| = |f^{(n+1)} - f|$ . As  $\sum_{n=0}^N \frac{d^n f}{dx^n}$  converges,  $\frac{d^n f}{dx^n} \rightarrow 0$ . But then  $f^{(n)}(x)$  is uniformly bounded on  $[\alpha, \beta]$ . Let  $M_1$  be such a bound. As  $f$  is continuous on  $[\alpha, \beta]$  it is bounded. Let  $M_2$  be such a bound. Let  $M = M_1 + M_2$ . Then  $|S'_N - S_N| = |f - f^{(N+1)}| \leq M$ . That is,  $|S'_N - S_N|$  is uniformly bounded. Therefore  $S_N$  converges uniformly to zero. But if the convergence is uniform, then  $f = 0$ . A contradiction. Thus  $f$  is not nonzero anywhere, and therefore  $f = 0$ .  $\square$

$a$  and  $b$  need not be finite. The theorem holds on all of  $\mathbb{R}$ .

## 38.5 Other Results

**Theorem 38.5.1.** *A sum of  $K$  continuous functions is continuous.*

*Proof.* For let  $f_n$ ,  $n = 1, 2, \dots, K$  be continuous, let  $x$  be a point in their domains, and let  $\varepsilon > 0$  be given. Then, there is a  $\delta_n$  such that  $|x - x_0| < \delta_n$ , with  $x_0$  also in the domain, implies  $|f_n(x) - f_n(x_0)| < \frac{\varepsilon}{K}$ . Let  $\delta = \min\{\delta_1, \dots, \delta_K\}$ . Then  $|\sum_{n=1}^K [f_n(x) - f_n(x_0)]| \leq \sum_{n=1}^K |f_n(x) - f_n(x_0)| < \sum_{n=1}^K \frac{\varepsilon}{K} = \varepsilon$ .  $\square$

**Theorem 38.5.2.** *The set of rational numbers  $\frac{p}{q}$  where  $p$  and  $q$  are prime is dense in  $\mathbb{R}^+$ .*

*Proof.* If  $x = 0$ , from Euclid we have  $\frac{1}{p_n} \rightarrow 0$ , where  $p_n$  is the  $n^{th}$  prime. Let  $x \in \mathbb{R}^+$  be given. From the Prime Number Theorem,  $\frac{p_n}{n \ln(n)} \rightarrow 1$ . Let  $p_{\lceil nx \rceil}$  be the  $\lceil nx \rceil^{th}$  prime. Then  $\frac{p_{\lceil nx \rceil}}{p_n} \frac{n \ln(n)}{nx \ln(nx)} \rightarrow 1$ . But  $\frac{n \ln(x)}{nx \ln(nx)} \rightarrow \frac{1}{x}$ . Therefore  $\frac{p_{\lceil nx \rceil}}{p_n} \rightarrow x$ .  $\square$

**Theorem 38.5.3.** *If  $p$  is a positive integer, then  $e^p$  is irrational.*

*Proof.* For let  $m$  and  $n$  be positive integers, and let:

$$I_n = \frac{1}{n!} \int_0^\infty [x(x-p)]^n e^{-x} dx \quad J_n = \frac{1}{n!} \int_0^\infty [x(x+p)]^n e^{-x} dx$$

By induction, we have that  $I_n$  and  $J_n$  are integers for all integer  $p$ . But now:

$$\begin{aligned} me^p I_n &= \frac{me^p}{n!} \int_0^\infty [x(x-p)]^m e^{-x} dx \\ &= \frac{me^p}{n!} \int_0^p [x(x-p)]^n e^{-x} dx + \frac{m}{n!} \int_p^\infty [x(x-p)]^n e^{-(x-p)} dx \\ &= \frac{me^p}{n!} \int_0^p [x(x-p)]^n e^{-x} dx + \frac{m}{n!} \int_0^\infty [(u+p)u]^n e^{-u} du \\ &= \frac{me^p}{n!} \int_0^p [x(x-p)]^n e^{-x} dx + mJ_n \end{aligned}$$

But for  $x \in [0, p]$ ,  $|x(x-p)| \leq p^2/4$  and  $0 < e^{-x} \leq 1$ , and therefore:

$$\left| \frac{me^p}{n!} \int_0^p |x(x-p)|^n e^{-x} dx \right| \leq \frac{me^p p^{2n}}{4^n n!}$$

Let  $N$  be such that  $N! > me^p (p^2/4)^N$ , we have:

$$\left| \frac{me^p}{n!} \int_0^p [x(x-p)]^n e^{-x} dx \right| < 1$$

But moreover, this integral is non-zero since the integrand is positive on the interval  $(0, p)$ . So we have:

$$0 < me^p I_n - mJ_n < 1$$

Therefore  $me^p$  cannot be an integer, and therefore  $e^p$  is irrational.  $\square$

**Theorem 38.5.4.** *If  $p$  and  $q$  are positive integers, then  $e^{p/q}$  is irrational.*

*Proof.* For suppose not. Then  $(e^{p/q})^p = e^p$  is rational, a contradiction. Therefore, etc.  $\square$

**Theorem 38.5.5** (Kronecker's Theorem). *If  $\alpha$  is irrational, and if  $g$  is a continuous function, then:*

$$\frac{1}{2\pi} \int_0^{2\pi} g(e^{i\theta}) d\theta = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{n=0}^N g(e^{ink\alpha})$$

*Proof.* Let  $I$  be the functional  $I(g) = \frac{1}{2\pi} \int_0^{2\pi} g(\exp(i\theta)) d\theta$ . For  $N \in \mathbb{N}$ , let  $I_N(g)$  be the functional  $I_N(g) = \frac{1}{N+1} \sum_{n=0}^N g(\exp(ik\alpha))$ . Then  $I$  and  $I_N$  are linear functionals. Let  $\|g\|$  be the supremum norm,  $\|g\| = \sup\{g(\exp(i\theta))\}$ . Then, from the definition of  $I$  and  $I_N$ ,  $|I(g)| \leq \|g\|$  and  $|I_N(g)| \leq \|g\|$ . If  $g$  is the constant mapping  $g(\theta) = 1$ , then  $I_N(g) = I(g) = 1$ . If  $g(\exp(i\theta)) = \exp(in\theta)$  for  $n \in \mathbb{N}$ , then:

$$I(g) = \frac{1}{2\pi} \int_0^{2\pi} e^{in\theta} d\theta = 0$$

Let  $r = e^{in\alpha}$ . Then we have:

$$I_N(g) = \frac{1}{N+1} \sum_{n=0}^N r^n = \frac{1-r^{N+1}}{1-r}$$

But  $\alpha$  is irrational, and thus for all  $N \in \mathbb{N}$ ,  $1-r \neq 0$ . But then:

$$|I_N(g)| = \frac{1}{N+1} \left| \frac{1-r^{N+1}}{1-r} \right| \leq \frac{1}{N+1} \frac{2}{|1-r|} \rightarrow 0$$

If  $g(\exp(i\theta)) = \sum_{n=0}^N a_n \exp(in\theta)$ , then the result holds by induction. If  $g$  is continuous, and  $\varepsilon > 0$ , then there is a polynomial  $P$  such that  $\sup\{|P(x) - g(x)|\} < \varepsilon/3$ . But then:

$$|I(g) - I_N(g)| \leq |I(g) - I(P)| + |I(P) - I_N(P)| + |I_N(P) - I_N(g)|$$

But  $|I(g) - I(P)| < \varepsilon/3$ , and from before we have that there is an  $N \in \mathbb{N}$  such that  $|I(P) - I_N(P)| < \varepsilon/3$ . Finally,  $|I_N(P) - I_N(g)| = |I_N(P-g)| < \|P-g\| < \varepsilon$ . Therefore  $|I_N(g) - I(g)| < \varepsilon$ .  $\square$

## 38.6 An Almost Group

**Definition 38.6.1** A group is a set  $G$  with an operation  $*$  satisfying the following:

1.  $a * (b * c) = (a * b) * c$  for all  $a, b, c \in G$
2. There is an  $e \in G$  such that  $a * e = e * a = a$  for all  $a \in G$
3. For all  $a \in G$  there is an  $a^{-1} \in G$  such that  $a * a^{-1} = a^{-1} * a = e$

**Theorem 38.6.1.** *The identity of a group is unique.*

*Proof.* Suppose not, and let  $e'$  be a different identity. But  $e' = e' * e = e$ . Thus  $e$  is unique.  $\square$

**Definition 38.6.2** A quasigroup is a group but the operation need not be associative.

**Definition 38.6.3** An Abelian Quasigroup is a quasigroup with a commutative operation.

An interesting thing to note is that  $e$  is an identity for *all* elements of  $G$ . There are, however, groups with elements  $a, b$  such that  $a * b = b * a = a$ , and yet  $b \neq e$ . The key difference is that  $a * b$  does not necessarily equal  $a$  for *all*  $a \in G$ .

**Theorem 38.6.2.** *There exist abelian quasigroups  $\langle G, *\rangle$  with elements  $a, b \in G$  such that  $a * b = b * a = a$ , yet  $b \neq e$ .*

*Proof.* In a pathological construction, let  $G = \mathbb{R}$ . Consider the following operation:

$$x * y = \begin{cases} (x + y)^2, & x, y \neq 0 \\ x, & y = 0, x \neq 0 \\ y, & x = 0, y \neq 0 \\ 0, & x, y = 0 \end{cases} \quad (38.6.1)$$

The identity is zero. For  $0 * 0 = 0$ , and if  $x \neq 0$ , then  $x * 0 = 0 * x = x$ . The inverse is  $-x$ . For if  $x \neq 0$ , then  $x * (-x) = (x - x) = 0$ . The operation is not associative, for, in general:

$$x * (y * z) = (x + (y + z)^2)^2 \neq ((x + y)^2 + z)^2 \quad (38.6.2)$$

For take  $x = 2$ ,  $y = 1$ , and  $z = 1$ . Then  $x * (y * z) = 36$ , but  $(x * y) * z = 100$ . It is, however, commutative. For if  $x, y \neq 0$ , then:

$$x * y = (x + y)^2 = (y + x)^2 = y * x \quad (38.6.3)$$

The case of either element being zero is identity, and thus commutative. Let  $x = 4$  and  $y = -2$ . Then  $x * y = (4 - 2)^2 = 4 = x$ , and  $y * x = (-2 + 4)^2 = 4 = x$ . Also,  $4 * (-6) = (-6) * 4 = (4 - 2)^2 = (-2)^2 = 4$ . Thus, 4 has three *identities*, that is  $0, -2, -6$ . 4 is not the only element, for let  $x \neq 0$ . Then  $y = x - \sqrt{x}$  and  $y = -x - \sqrt{x}$  are also identities for  $x$ . Thus, with the exception of 0 and 1, every positive element has three identities. Note that  $-2$  is only an identity for the elements 4 and 1. Thus, for any other elements  $x * (-2) \neq -2$ . Thus,  $-2$  is not a true identity.  $\square$

## 38.7 On Sequences

### Some Fun Stuff

**Theorem 38.7.1.** *Given an enumeration  $\{x_n\}_{n=1}^{\infty}$  of the rationals  $\mathbb{Q} \cap [0, 1]$ , for all  $\varepsilon > 0$  there is a  $k \in \mathbb{N}$  such that  $|x_{k+1} - x_k| < \varepsilon$ .*

*Proof.* For let  $x_n$  be such an enumeration. Then, for all  $n \in \mathbb{N}$ ,  $0 \leq x_n \leq 1$ .  $\square$

**Definition 38.7.1** The Fibonacci Numbers are formed by the sequence  $F_{n+2} = F_{n+1} + F_n$ , with  $F_0 = F_1 = 1$ .

**Definition 38.7.2** Two positive integers are said to be coprime if they share no common factors.

**Theorem 38.7.2.** *Any two consecutive Fibonacci numbers are coprime.*

*Proof.* We have that  $F_0 = F_1 = 1$  and thus  $F_2 = 2$ , and also  $F_3 = 3$ . Suppose there is some integer  $N \in \mathbb{N}$  such that  $F_{N+2}$  and  $F_{N+1}$  are not coprime. Then there is a least integer  $n \in \mathbb{N}$  such that  $F_{n+2}$  and  $F_{n+1}$  are not coprime. That is, there are integers  $a, b, c \in \mathbb{N}$  such that  $F_{n+2} = ab$  and  $F_{n+1} = ac$  where  $b > c$ . But then  $F_n = F_{n+2} - F_{n+1} = a(b - c)$ . Let  $\alpha = b - c \in \mathbb{N}$ . Then  $F_n$  and  $F_{n+1}$  are also not coprime. But this is impossible as  $n$  is the least integer such that  $F_{n+2}$  and  $F_{n+1}$  are coprime, and  $n - 1 < n$ , a contradiction. Therefore there is no  $N$  such that  $F_{N+2}$  and  $F_{N+1}$  are coprime. Consecutive Fibonacci numbers are coprime.  $\square$

**Theorem 38.7.3.** *For all  $N \in \mathbb{N}$ ,  $\sum_{n=1}^N n \cdot n! = (N + 1)! - 1$ .*

*Proof.* For  $n \cdot n! = n \cdot n! + n! - n! = n!(n + 1) - n! = (n + 1)! - n!$ . Thus,  $\sum_{n=1}^N n \cdot n! = \sum_{n=1}^N (n + 1)! - n! = (N + 1)! - 1$ , as this is a telescoping series.  $\square$

**Theorem 38.7.4.** *If  $f(x)$  is an increasing function on  $[1, N+1]$ , then  $\sum_{n=2}^{N+1} f(n) \leq \int_1^{N+1} f(x) dx \leq \sum_{n=1}^N f(n)$ .*

*Proof.* For  $x \in [n, n+1]$ ,  $f(n+1) \leq f(x) \leq f(n)$ , as  $f$  is decreasing. Thus  $\int_n^{n+1} f(n+1) dx \leq \int_n^{n+1} f(x) dx \leq \int_n^{n+1} f(n) dx \Rightarrow f(n+1) \leq \int_n^{n+1} f(x) dx \leq f(n)$ . Summing over this, we obtain  $\sum_{n=1}^N f(n+1) \leq \int_1^{N+1} f(x) dx \leq \sum_{n=1}^N f(n)$ . Finally, applying a shift of index to the leftmost term,  $\sum_{n=2}^{N+1} \leq \int_1^{N+1} f(x) dx \leq \sum_{n=1}^N f(n)$ .  $\square$

**Theorem 38.7.5.** If  $f$  is decreasing, then  $\int_1^{n+1} f(x) dx \leq \sum_{k=1}^{n+1} f(k) \leq \int_1^{n+1} f(x) dx + f(1)$

*Proof.* For  $\int_1^{n+1} f(x) dx \leq \sum_{k=1}^n f(k) \leq \sum_{k=1}^{n+1}$ . But  $\sum_{k=2}^{N+1} f(k) \leq \int_1^{N+1} f(x) dx$  so  $\sum_{k=1}^{n+1} f(k) \leq \int_1^{n+1} f(x) dx + f(1)$ . Combining these together gives the result.  $\square$

**Theorem 38.7.6.**  $\lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n+k} = \ln(2)$ .

*Proof.* From the previous theorem,  $\int_1^n \frac{1}{n+x} dx \leq \sum_{k=1}^n \frac{1}{n+k} \leq \frac{1}{n+1} + \int_1^n \frac{1}{n+x} dx$ , and thus  $\ln(n+x)|_1^{n+1} \leq \sum_{k=1}^n \frac{1}{n+k} \leq \frac{1}{n+1} + \ln(n+x)|_1^{n+1} \Rightarrow \ln(\frac{2n+1}{n+1}) \leq \sum_{k=1}^n \frac{1}{n+k} \leq \ln(\frac{2n+1}{n+1}) + \frac{1}{n+1}$ . As  $\frac{2n+1}{n+1} \rightarrow 2$  and as  $\ln(x)$  is continuous,  $\ln(\frac{2n+1}{n+1}) \rightarrow \ln(2)$ . But also  $\frac{1}{n+1} \rightarrow 0$ . Thus, by the squeeze theorem,  $\sum_{k=1}^n \frac{1}{n+k} \rightarrow \ln(2)$ .  $\square$

**Theorem 38.7.7.**  $\sum_{k=1}^n \frac{1}{\sqrt{k}} < 2\sqrt{n}$ .

*Proof.* From the theorem we have that  $\sum_{k=1}^n \frac{1}{\sqrt{k}} \leq \int_1^n \frac{1}{\sqrt{x}} dx + 1 < \int_1^n \frac{1}{\sqrt{x}} dx + 2 = 2\sqrt{n} - 2 + 2 = 2\sqrt{n}$ .  $\square$

**Theorem 38.7.8.** If  $x \bmod 1 < \frac{1}{2}$ , then  $2\lfloor x \rfloor = \lfloor 2x \rfloor$ .

*Proof.* Let  $0 \leq x \bmod 1 \leq 0.5$ . Then  $0 \leq x - \lfloor x \rfloor < 0.5 \Rightarrow 2x - 2\lfloor x \rfloor < 1$  and thus  $2\lfloor x \rfloor \leq \lfloor 2x \rfloor \leq 1 + 2\lfloor x \rfloor$ . But then we have that  $0 \leq \lfloor 2x \rfloor - 2\lfloor x \rfloor < 1$ . But this is the difference of two integers, and is thus an integer. But there are no integers between 0 and 1, and therefore  $\lfloor 2x \rfloor - 2\lfloor x \rfloor = 0$ . Thus,  $\lfloor 2x \rfloor = 2\lfloor x \rfloor$ .  $\square$

## A Peculiar Family of Sequences and their Averages

Consider the sequence  $1, 2, 1, 1, 3, 1, 1, 1, 4, 1, 1, 1, 1, 5, \dots, n, \dots (n \text{ } 1's) \dots, n+1$  and also the generalization  $1^k, 2^k, \dots (2^k \text{ } 1's), \dots, 3^k, \dots (3^k \text{ } 1's), \dots, n^k, \dots (n^k \text{ } 1's) \dots, (n^k)^k$

**Theorem 38.7.9.** If  $a_n, b_n$  are sequences,  $a_n \rightarrow A$  and  $a_n - b_n \rightarrow 0$ , then  $b_n \rightarrow A$ .

*Proof.* For  $|A - b_n| \leq |A - a_n| + |a_n - b_n| \rightarrow 0$ , thus  $|A - b_n| \rightarrow 0$  and therefore  $b_n \rightarrow A$ .  $\square$

**Theorem 38.7.10.** Let  $a_n$  be a sequence and  $f, g$  be strictly increasing integer valued functions such that for all  $m < f(n)$ ,  $a_{f(n)} > a_m$  and for all  $m > g(n)$ ,  $a_{g(n)} < a_m$ . If  $a_{f(n)} \rightarrow A$  and  $a_{f(n)} - a_{g(n)} \rightarrow 0$ , then  $a_n \rightarrow A$ .

*Proof.* Let  $\varepsilon > 0$  be given. We have that  $a_{g(n)} \rightarrow A$  as well from the previous lemma. Thus, there is an  $N_1 \in \mathbb{N}$  such that for all  $n > N_1$ ,  $|A - a_{g(n)}| < \varepsilon$ . Thus, for  $n > N_1$ ,  $A - \varepsilon < a_{g(n)} < A + \varepsilon$ . But for all integers  $n > g(N_1)$ ,  $a_n > a_{g(N_1)}$ , and thus  $A - \varepsilon < a_n$  for all  $n > g(N_1)$ . As  $a_{f(n)} \rightarrow A$ , there is an  $N_2$  such that for all  $n > N_2$ ,  $|A - a_{f(n)}| < \varepsilon$ . Thus, for  $n > N_2$ ,  $A - \varepsilon < a_{f(n)} < A + \varepsilon$ . As  $f$  is a monotonically increasing function on the integers,  $f(n) \geq n$ . Thus,  $a_{f(n)} > a_n$  for all  $n$ . But then for  $n > \max\{g(N_1), N_2\}$ ,  $A - \varepsilon < a_{g(n)} < a_n < a_{f(n)} < A + \varepsilon$ . Thus,  $a_n \rightarrow A$ .  $\square$

**Theorem 38.7.11.** If  $f$  and  $g$  are continuous functions defined on  $\mathbb{R}^+$ , and if  $\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) = A$ , and if  $S = \{(x, y) : x \in \mathbb{R}^+, \min\{f(x), g(x)\} \leq y \leq \max\{f(x), g(x)\}\}$ , and if  $a_n$  is any sequence such that  $(n, a_n) \in S$  for all  $n \in \mathbb{N}$ , then  $a_n \rightarrow A$ .

*Proof.* As  $(n, a_n) \in S$ :

$$\begin{aligned} \min\{f(n), g(n)\} &\leq a_n \leq \max\{f(n), g(n)\} \\ &\Rightarrow 0 \leq a_n - \min\{f(n), g(n)\} \leq \max\{f(n), g(n)\} - \min\{f(n), g(n)\} \end{aligned}$$

But  $\max\{f(n), g(n)\} - \min\{f(n), g(n)\} \rightarrow 0$ , and thus  $a_n - \min\{f(n), g(n)\} \rightarrow 0$ . From the lemma,  $a_n \rightarrow A$ .  $\square$

**Theorem 38.7.12.** If  $P(x)$  and  $Q(x)$  are polynomials of degree  $n$ , with leading coefficients  $a_n$  and  $b_n$ , respectively, then  $\lim_{x \rightarrow \infty} \frac{P(x)}{Q(x)} = \frac{a_n}{b_n}$ .

*Proof.* From repeated application of L'Hôpital's Rule:

$$\lim_{x \rightarrow \infty} \frac{P(x)}{Q(x)} = \lim_{x \rightarrow \infty} \frac{a_n x^n + \dots + a_0}{b_n x^n + \dots + b_0} = \lim_{x \rightarrow \infty} \frac{n! a_n}{n! b_n} = \frac{a_n}{b_n}$$

$\square$

**Theorem 38.7.13.** The average of the family of sequences we were considering is 2. That is, let  $a_n(k)$  be the  $n^{\text{th}}$  term in the sequence  $1^k, 2^k, \dots, (2^k 1's), \dots, 3^k, \dots$ , then the average  $\frac{\sum_{n=1}^N a_n(k)}{N}$  converges to 2 for all  $k \geq 1$ .

## 38.8 A Class of Differentiability

**Definition 38.8.1** A function  $f : (a, \infty) \rightarrow \mathbb{R}$ ,  $a > 0$ , is said to be Kiwi Continuous if  $f(x) - xf'(x)$  is bounded.

A function is Kiwi Continuous if the set of  $y$ -intercepts of the tangent lines of  $f(x)$  is bounded.

**Theorem 38.8.1.** *If  $f : [a, \infty) \rightarrow \mathbb{R}$  is Kiwi Continuous, then  $f'$  is bounded.*

*Proof.* By the definition,  $-m \leq f(x) - xf'(x) \leq m$ . Therefore  $-\frac{m}{x^2} \leq \frac{f(x)}{x^2} - \frac{f'(x)}{x} \leq \frac{m}{x^2}$ . But  $\frac{f(x)}{x^2} - \frac{f'(x)}{x} = -\frac{d}{dx}\left(\frac{f(x)}{x}\right)$ . So  $-\frac{m}{x^2} \leq \frac{d}{dx}\left(\frac{f(x)}{x}\right) \leq \frac{m}{x^2}$ . Let  $x_0 \in (a, \infty)$ . Then  $-\int_{x_0}^x \frac{m}{\tau^2} d\tau = -\left[-\frac{m}{\tau} + \frac{m}{x_0}\right] = \frac{m}{x} - \frac{m}{x_0} \leq \int_{x_0}^x \frac{d}{d\tau}\left(\frac{f(\tau)}{\tau}\right) d\tau = \frac{f(x)}{x} - \frac{f(x_0)}{x_0} \leq \int_{x_0}^x \frac{m}{\tau^2} d\tau = \frac{m}{x_0} - \frac{m}{x}$ . So  $\left|\frac{f(x)}{x}\right| \leq m\left|\frac{1}{x} - \frac{1}{x_0}\right| \leq m\left|\frac{2}{a}\right|$ . Therefore  $|f(x)| \leq 2\frac{m}{a}x$ . But  $|f(x) - xf'(x)| \leq m$ . Thus  $|f(x) - xf'(x)| \geq |f(x)| - x|f'(x)|$ , and therefore  $|f'(x)| \leq \frac{m+|f(x)|}{x} \leq \frac{m+\frac{2m}{a}x}{x} \leq \frac{m}{a} + \frac{2m}{a} = \frac{3m}{a}$ . Therefore,  $|f'(x)|$  is bounded.  $\square$

## 38.9 Degenerate Fredholm Equations of the First Kind

**Definition 38.9.1** A Fredholm Equation of the first kind is an equation of the form:

$$f(x) = \int_a^b g(x_0)K(x, x_0)dx_0$$

**Definition 38.9.2** A degenerate Fredholm of the First Kind is an equation of the form:

$$f(x) = \int_a^b g(x_0)K_1(x)K_2(x_0)dx_0$$

**Theorem 38.9.1.** *If  $f(x) = \int_a^b g(x_0)K_1(x)K_2(x_0)dx_0$ ,  $f$  and  $K_1$  are non-zero, and if  $K_2$  is continuous and non-zero at some point  $\xi \in (a, b)$ , then there exists two solutions  $g_1(x_0)$  and  $g_2(x_0)$ .*

*Proof.* If  $f$  and  $K_1$  are non-zero, then:

$$f(x) = \int_a^b g(x_0)K_1(x)K_2(x_0)dx_0 \tag{38.9.1a}$$

$$= K_1(x) \int_a^b g(x_0)K_2(x_0)dx_0 \tag{38.9.1b}$$

$$\implies \frac{f(x)}{K_1(x)} = \int_a^b g(x_0)K_2(x_0)dx_0 \tag{38.9.1c}$$

But  $\int_a^b g(x_0)K_2(x_0)dx_0$  is a number  $c \in \mathbb{R}$ . Since  $K_2$  is continuous and positive at a point  $\xi \in (a, b)$ , there is an  $\varepsilon > 0$  such that  $\forall_{x \in B_\varepsilon(\xi)}$ ,  $K_2(x) > \frac{K_2(\xi)}{2}$ . Let

$G_r(x)$  be defined as follows:

$$G_r(x) = \begin{cases} 0, & x \notin (\xi - \epsilon, \xi) \\ \frac{2r}{\varepsilon}(x - (\xi - \varepsilon)), & x \in (\xi - \epsilon, \xi - \frac{\varepsilon}{2}) \\ \frac{2r}{\varepsilon}(x - \xi), & x \in (\xi - \frac{\varepsilon}{2}, \xi) \end{cases} \quad (38.9.2)$$

Let  $F(r) = \int_a^b G_r(x) K_2(x) dx$ . Then we have:

$$F(r) = \int_a^b G_r(x) K_1(x) dx \quad (38.9.3a)$$

$$= \int_{\xi-\epsilon}^{\xi} G_r(x) K_1(x) dx \quad (38.9.3b)$$

$$\geq \frac{K_1(\xi)}{2} \int_{\xi-\varepsilon}^{\xi} G_r(x) dx \quad (38.9.3c)$$

$$= \frac{K_1(\xi)}{2} \frac{\varepsilon r}{2} \quad (38.9.3d)$$

Therefore,  $F(r) \rightarrow \infty$  as  $r \rightarrow \infty$ . Furthermore,  $F(0) = 0$ . Suppose  $c > 0$ . Let  $M = \{r \in \mathbb{R} : c < F(r)\}$ .  $M$  is bounded below, for 0 is such a bound. Then there exists a Greatest Lower Bound  $\alpha$ . From the continuity of  $F$ ,  $\lim_{r \rightarrow \alpha} F(r) = F(\alpha)$ , and  $F(\alpha) = c$ . Therefore  $G_\alpha(x)$  is a function such that:

$$\int_a^b G_\alpha(x) K_1(x) dx = c$$

For the second function, repeat the argument on the interval  $(\xi, \xi + \varepsilon)$   $\square$

**Theorem 38.9.2.** *There infinitely many solutions to degenerate Fredholm Equations of the First Kind.*

*Proof.* By the previous theorem, there are at least two. Let  $g_1$  and  $g_2$  be such solutions. Then, for all  $\lambda \in \mathbb{R}$ , define  $G_\lambda$  by:  $G_\lambda(x) = \lambda g_1(x) + (1 - \lambda)g_2(x)$ . For all  $\lambda \in \mathbb{R}$ ,  $G_\lambda$  is a solution.  $\square$

Part XXIII

# Riemannian Geometry



# CHAPTER 39

---

## Semi-Riemannian Geometry

---

### 39.1 Definitions

Book Six

Geometric Topology





**Part XXIV**

# **Surgery Theory**



# CHAPTER 40

---

## Surgery Theory

---

### 40.1 Lecture 2: Surgery Structure Sets

Let  $X$ ,  $M_1$ , and  $M_2$  be closed, compact  $n$  dimensional manifolds without boundary. Two homotopy equivalences  $f_i : M_i \rightarrow X$  are called equivalent if there exists a cobordism  $(W; M_1, M_2)$  and a map  $(F; f_1, f_2) : (W; M_1, M_2) \rightarrow (X \times [0, 1]; X \times \{0\}, X \times \{1\})$  such that  $F, f_1, f_2$  are homotopy equivalences. The structure set  $S(X)$  is the set of equivalence classes of homotopy equivalences  $f : M \rightarrow X$  from closed manifolds of dimension  $n$  to  $X$ .

$$\begin{array}{ccc} A & \xrightarrow{\alpha} & B \\ \gamma \downarrow & & \downarrow \beta \\ C & \xrightarrow{\delta} & D \end{array}$$

Fig. 40.1: Example of a Commutative Diagram.

---

#### Definition 40.1.1: Surgery Structure Set

---

The surgery structure set of a closed (without boundary) compact manifold  $M$  is  $S(M) = \{f : N^n \rightarrow M^n | f \text{ is a Homotopy Equivalence}\}$  ■

---

#### Definition 40.1.2: Base Point of a Surgery Structure Set

---

The base point of a surgery structure set is the map  $id_X : X \rightarrow X$ . ■

Let  $N_1$  and  $N_2$  be two manifold structures. And let  $f_1 : N_1^n \rightarrow M^n$  and  $f_2 : N_2^n \rightarrow M^n$  be two homotopy equivalences. We call  $g : N_1 \rightarrow N_2$  a cat-

homeomorphism if  $g$ , together with  $f_1$  and  $f_2$ , form the commutative diagram in Fig. 40.1. That is,  $g$  is a cat-homeomorphism if it homotopy commutes.

**Example 40.1.1** Some examples of surgery structure sets:

$$1. \ S^{Top}(S^n) = \{S^n\} \quad 2. \ S^{PL}(S^n) = \{S^n\} \quad 3. \ S^{Diff}(S^7) = \mathbb{Z}_{28}$$

### Orientable and Non-Orientable

Stiefel-Whitney classes  $w_1, \dots, w_n$  are cohomological classes. Orientable means that  $w_1 = 0$ .

**Example 40.1.2**

- |                                     |  |
|-------------------------------------|--|
| 1. $\mathbb{RP}^2$ - Non-Orientable | 5. $\mathbb{RP}^3$ - Orientable                            |
| 2. $\mathbb{RP}^4$ - Non-Orientable | 6. $\mathbb{RP}^5$ - Orientable                            |
| 3. $\mathbb{RP}^6$ - Non-Orientable | 7. $\mathbb{RP}^7$ - Orientable                            |
| 4. $\mathbb{RP}^8$ - Non-Orientable | 8. $\mathbb{CP}^n$ - Orientable for all $n \in \mathbb{N}$ |

Returning to surgery exact sequences, the goal is to compute  $S^{Cat}(\mathcal{M}^n)$ , where  $n$  is the dimension of  $\mathcal{M}^n$ . The notion of a surgery helps solve this question. Let  $X = \mathbb{S}^2 \setminus \{(a_1, b_1, c_1), (a_2, b_2, c_2)\}$ . That is, the sphere with two points removed. Stretch these two points out to create a sphere with two holes removed. One could imagine taking a hollow cylinder and stretching it to connect the two holes in the sphere. The result is a spherical coffee cup, see Fig. 40.2. This can be continuously deformed into a torus.

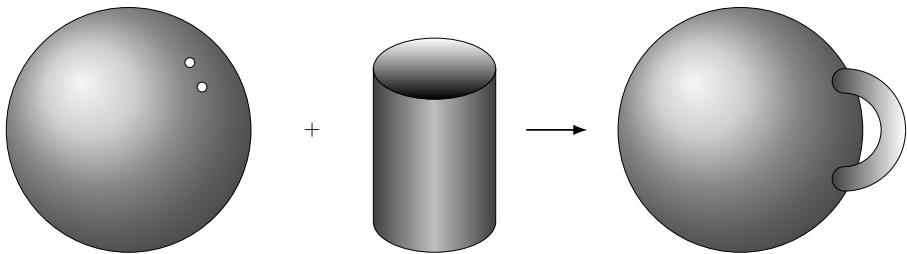


Fig. 40.2: Simple Surgery Example.

Recall that  $S^0$  is two points, and that  $D^2$  is the open unit disc. Then  $S^0 \times D^2 = D^2 \coprod D^2$  is simply two disjoint open unit discs. This is a good representation of the idea of the disjoint union, denoted  $X \coprod Y$ . We have:

$$S^0 \times D^2 = D^2 \coprod D^2$$

We can also represent a cylinder as the closed  $S^1 \times \overline{D}^1$ . The codimension of a surgery is the dimension of the object minus the dimension of a surgery. So, for the surgery in Fig. 40.2, the dimension of the entire thing is 2, the dimension of the surgery is 2, so the codimension is 0. This is called a Zero-Surgery. A zero-surgery takes out 2 holes and connects them with a tube.

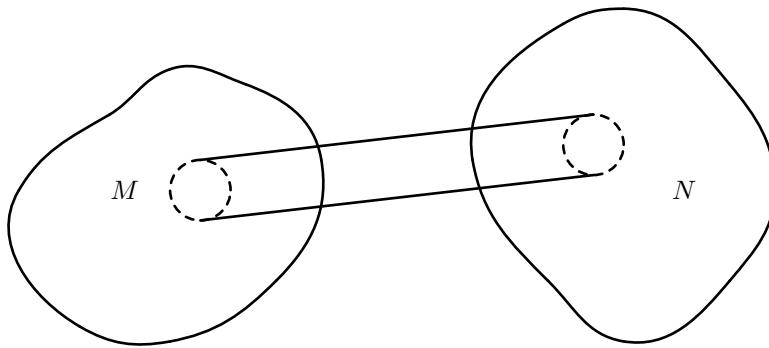


Fig. 40.3: Example of a Zero Surgery.

Let  $\mathcal{M}^n$  be an  $n$  dimensional manifold. Embed  $S^k \times D^{n-k}$  into  $\mathcal{M}^n$ . Let  $\partial(X)$  be the boundary of  $X$ . Then we have:

$$\partial(S^k \times D^{n-k}) = S^k \times S^{n-k-1} \quad \dim(S^{k+1} \times D^{n-k-1}) = \dim(S^k \times D^{n-k}) = n$$

Remove  $\partial(S^k \times D^{n-k})$  and glue on  $S^{k+1} \times D^{n-k-1}$ . We also have that  $\partial(D^{k+1} \times S^{n-k-1}) = S^k \times S^{n-k-1}$ . Glue  $\mathcal{M}^n \cup (D^{k+1} \times S^{n-k-1})$  along  $\partial(S^k \times S^{n-k-1})$ .

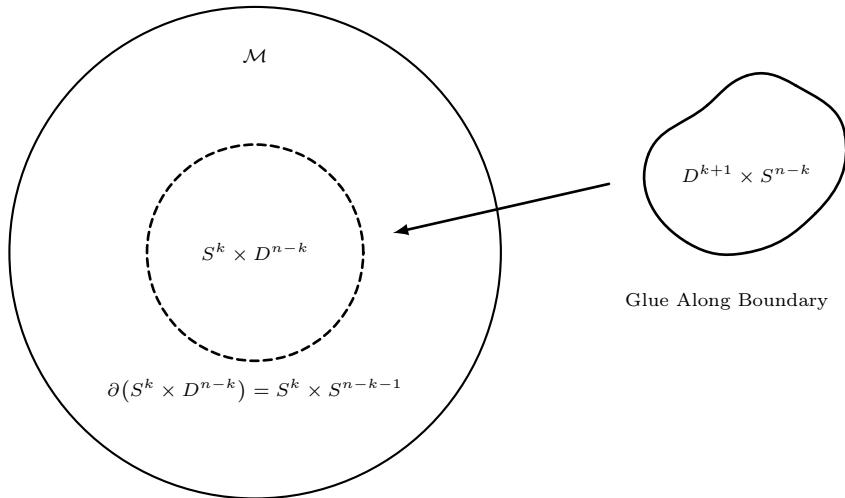


Fig. 40.4: Gluing  $D^{k+1} \times S^{n-k}$  along  $\partial(S^k \times D^{n-k})$ . The new manifold is  $\mathcal{M}^n \setminus (S^k \times D^{n-k} \coprod (D^{k+1} \times S^{n-k-1}))$

We now consider  $k$  surgeries  $\mathcal{M} \xrightarrow{k\text{-surgery}} \mathcal{N}$ . We have seen  $S^2 \xrightarrow{0\text{-surgery}} T^2$ . Note:  $\pi_1(S^2)$  is trivial, and  $\pi_1(T^2) = \mathbb{Z}^2$ . This happens because  $n < 5$ . When  $n \geq 5$ , we have the following result.

**Theorem 40.1.1.** *If  $\mathcal{M}$  is an  $n$  dimensional manifold,  $n \geq 5$ , and if  $\mathcal{N}$  is the result of a  $k$  surgery on  $\mathcal{M}$ , then  $\pi_1(\mathcal{M}) = \pi_1(\mathcal{N})$ .*

### More On Surgery Exact Sequences

Recall that a surgery exact sequence looks like the following:

$$\underbrace{L_{n+1}(\mathbb{Z}\pi_1\mathcal{M})}_{\text{Group}} \rightarrow \cdots \rightarrow \underbrace{S^{Cat}(\mathcal{M}^n)}_{\text{Not a Group}} \rightarrow \underbrace{[M, G/0]}_{\text{Group}} \rightarrow \underbrace{L_n(\mathbb{Z}\pi_1(\mathcal{M}))}_{\text{Group}}$$

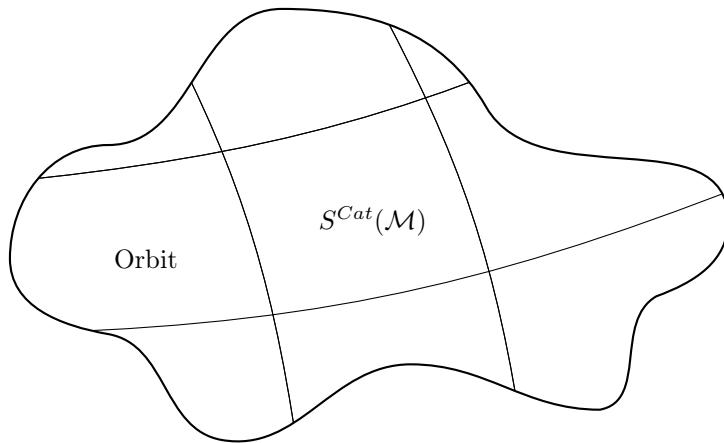
An exact sequence of groups is of then form  $G_{n+1} \xrightarrow{g_n} G_n \rightarrow \dots$ , where  $\text{Im}(g_n) = \ker(g_{n-1})$ . We refine our notion of a surgery exact sequence:

$$\cdots \rightarrow L_{n+1}(\mathbb{Z}\pi_1(\mathcal{M})) \dashrightarrow S^{Cat}(\mathcal{M}) \xrightarrow{g} [M, G/o] \xrightarrow{\sigma} L_n(\mathbb{Z}\pi_1(\mathcal{M}))$$

The dotted line means  $L_{n+1}(\mathbb{Z}\pi_1(\mathcal{M}))$  acts on  $S^{Cat}(\mathcal{M})$ . Exact means  $\text{Im}(g) = \ker(\sigma)$ . Each element  $f \in [M, G/o]$  either pulls back to  $\emptyset$  or something non-empty. If the pullback is non-empty, you get a blob in  $S^{Cat}(\mathcal{M})$ :  $f^{-1}(\{x\})$ . But:

$$\bigcup_{f \in [M, G/o]} g^{-1}(\{f\}) = S^{Cat}(\mathcal{M}) \quad (40.1.1)$$

This process creates a partition of  $S^{Cat}(\mathcal{M})$ . Now,  $L_{n+1}(\mathbb{Z}(\pi_1\mathcal{M}))$  acts on  $S^{Cat}(\mathcal{M})$  in some fashion. Partition the space into orbits. Exactness here means that partitioning by point inverses is the same as partitioning by orbits. That is, the two partitions are identical. See Fig. 40.5 for a partitioning into orbits.

Fig. 40.5: Partition of  $S^{Cat}(\mathcal{M})$ .

The next object to talk about is  $L_n(\mathbb{Z}\pi_1(\mathcal{M}))$ . These are called Wall groups. They are difficult to compute, but there are some facts that are known about them:

- Wall groups only have 2-torsion.
  - 2-torsion means that elements of finite order have order 2.
  - This implies the groups are Abelian.
- They can be orientable or not.
  - $L_n(\mathbb{Z}\pi_1(\mathcal{M})^\pm)$  indicates orientable or not.

### 40.1.1 Lecture 3: Vector Bundles

#### Group Rings

**Definition 40.1.3** If  $G$  is a group and  $R$  is a ring, then the group ring  $RG$  is the collection of all finite linear combinations (Formal Sums):  $r_1g_1 + \dots + r_ng_n$ , where  $r_k \in R$  and  $g_k \in G$ .

**Example 40.1.3** If  $G$  is a group, and  $\mathbb{Z}G = \{\sum_{k=0}^n n_k g_k : n_k \in \mathbb{Z}, g_k \in G\}$ , then  $\mathbb{Z}G$  is a group ring. This is a special group ring, denoted  $SP_{\mathbb{Z}}(G)$ .

**Theorem 40.1.2.** *If  $R$  is a ring and  $G$  is a group, then the group ring  $RG$  is a ring.*

From the previous lecture we saw that  $L_{n+1}(\mathbb{Z}\pi_1(\mathcal{M}))$  is a group. But from the previous theorem, we have that  $\mathbb{Z}\pi_1(\mathcal{M})$  is a ring. So, we may think of the  $L_n$  as a *Functor*:  $L_n : \text{Rings} \rightarrow \text{Groups}$ . To recap the notation,  $S(\mathcal{M})$  is the Surgery Structure Set on the manifold  $\mathcal{M}$ , and  $L_n(\mathbb{Z}\pi_1(\mathcal{M}))$  is a Wall Group.

## Matrices and Vector Bundles

The next monster we need to understand in the Surgery Exact Sequence is the  $[\mathcal{M}, G/o]$  that keep appearing. First, a quick recap on some notions in linear algebra.

**Definition 40.1.4** An orthogonal matrix is an invertible square matrix  $A$  such that  $A^T = A^{-1}$

Let  $\mathcal{O}(n)$  be the group of  $n \times n$  orthogonal matrices. There is a simple map then from  $\mathcal{O}(n)$  to  $\mathcal{O}(n+1)$ ,  $\psi_n : \mathcal{O}(n) \rightarrow \mathcal{O}(n+1)$ , defined by:

$$\psi_n(A) = \left[ \begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array} \right]$$

We can also define a map  $\varphi_{nm} : \mathcal{O}(n) \times \mathcal{O}(m) \rightarrow \mathcal{O}(n+m)$  defined by:

$$\varphi_{nm}(A, B) = \left[ \begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array} \right]$$

This is, in general, not a bijection. From this we can create a sequence:

$$\mathcal{O}(1) \xrightarrow{\psi_1} \mathcal{O}(2) \xrightarrow{\psi_2} \mathcal{O}(3) \xrightarrow{\psi_3} \mathcal{O}(4) \xrightarrow{\psi_4} \cdots \mathcal{O}(n) \xrightarrow{\psi_n} \cdots$$

We can then define  $\mathcal{O}$  as the *Direct Limit* of this sequence:

$$\mathcal{O} = \lim_{n \rightarrow \infty} \mathcal{O}(n)$$

Now, let  $\mathcal{M}$  be a manifold. An  $n$  dimensional vector bundle is a map  $P : E \rightarrow \mathcal{M}$  such that, for each point  $x \in \mathcal{M}$ , the *fiber* of  $x$ , the pre-image  $p^{-1}(x)$ , is homeomorphic to  $\mathbb{R}^n$ .

**Definition 40.1.5** The fiber of a point  $y$  in a set  $Y$  under the map  $f : X \rightarrow Y$  is the pre-image  $f^{-1}(y) \subset X$ .

**Definition 40.1.6** A real  $n$  dimensional vector bundle on a manifold  $\mathcal{M}$  is a manifold  $E$  and a continuous map  $p : E \rightarrow \mathcal{M}$  such that, for all  $x \in \mathcal{M}$ , the fiber of  $x$  is homeomorphic to  $\mathbb{R}^n$  and there exists an open set  $\mathcal{U}$  such that  $x \in \mathcal{U}$  and  $p^{-1}(\mathcal{U})$  is homeomorphic to  $\mathcal{U} \times \mathbb{R}^n$

The requirement that there is an open neighborhood  $\mathcal{U}_x$  for all  $x$  such that  $p^{-1}(\mathcal{U})$  is homeomorphic to  $\mathcal{U}_x \times \mathbb{R}^n$  is called *local triviality*. There is another notion called *global triviality*.

**Example 40.1.4** A classic example is a cylinder with a disk (Or the boundary of a cylinder with the circle). Given a point  $(x, y, z)$  in the cylinder, collapse this (Or project it) down onto the  $xy$  plane by the map  $p(x, y, z) = (x, y)$ . This

is continuous, and is an example of a vector bundle:  $(D^1, D^1 \times \mathbb{R}, p)$ . The pre-image, or fiber, of any point in  $D^1$  is a line, which is certainly homeomorphic to  $\mathbb{R}$ . Again, taking any point  $x$  and looking at an open ball about that point that is entirely contained within  $D^1$ , the pre-image  $p^{-1}(B_r(x))$  is another cylinder, which is homeomorphic to  $\mathbb{R}^3$ , which is itself homeomorphic to  $B_r(x) \times \mathbb{R}^1$ . The fibers of  $x$  and  $\mathcal{U}$  are shown in Fig. 40.6

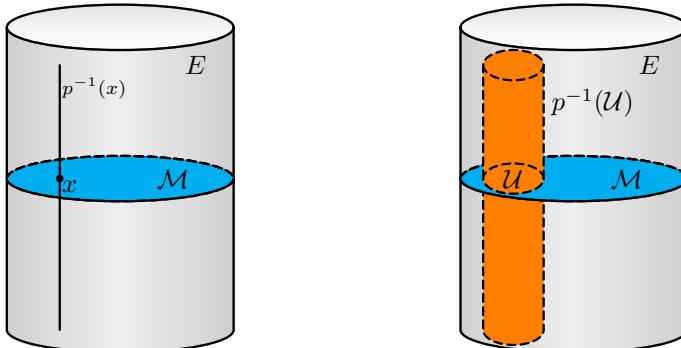


Fig. 40.6: Example of a Vector Bundle:  $(D^1, D^1 \times \mathbb{R}, p)$ .

**Example 40.1.5** If  $\mathcal{M}$  is a manifold,  $E = \mathcal{M} \times \mathbb{R}^n$ , and if  $p : E \rightarrow \mathcal{M}$  is defined by  $p(x, \mathbf{y}) = x$  for all  $(x, \mathbf{y}) \in \mathcal{M} \times \mathbb{R}^n$ , then  $(E, \mathcal{M}, p)$  is a vector bundle. This is called the trivial  $n$  dimensional vector bundle of  $\mathbb{R}^n$ . The fibers of points  $x \in \mathcal{M}$  are  $\mathbb{R}^n$ , which is homeomorphic to  $\mathbb{R}^n$ . Given any open set  $\mathcal{U}$  containing  $x$ , the pre-image is  $\mathcal{U} \times \mathbb{R}^n$ .

**Example 40.1.6** The Möbius strip can be seen as a vector bundle  $S^1 \times [0, 1] \rightarrow [0, 1]$  where the map  $(x, t) \rightarrow x$  is by a “twist.” This is a non-orientable bundle which is non-trivial. It has local triviality, but no global triviality.

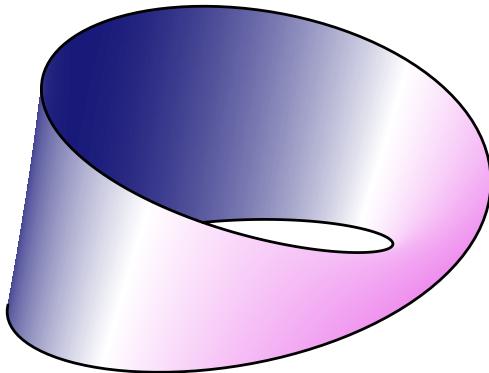


Fig. 40.7: Möbius Strip.

Returning to our discussion of orthogonal matrices,  $\mathcal{O}(n)$  is a group under matrix multiplication. The matrix  $I_n$  is orthogonal, and if  $A$  is orthogonal, then  $(A^{-1})^T = (A^T)^{-1} = A$ . But  $(A^{-1})^{-1} = A$ , and thus  $A^{-1}$  is orthogonal as well. Thus we have an identity, associativity, and closure of inverses. Therefore  $\mathcal{O}(n)$  is a group. This group can act on the set of  $n$  dimensional vectors in  $\mathbb{R}^n$  by the map  $(A, \mathbf{v}) \rightarrow A\mathbf{v}$ , for all  $A \in \mathcal{O}(n), \mathbf{v} \in \mathbb{R}^n$ . Thus, we have the  $\mathcal{O}(n)$  acts over the fibers of an  $n$  dimensional real vector bundle  $(E, \mathcal{M}, p)$ .  $\mathcal{O}(1)$  is the identity. That is, the “Do nothing,” action on a 1 dimensional vector bundle.  $\mathcal{O}(2)$  can perform *reflections* and *rotations* on the fibers. Endowed with this action, any real vector bundle of dimension  $n$  is an example of a *principal  $\mathcal{O}(n)$  bundle*. If  $g \in \mathcal{O}(n)$  and  $v \in E$ , then  $p(gv) = g(p(v))$ .

## Principal G-Bundles

If  $G$  is a group, and if  $X$  is a topological space, then there is a structure/notion of a *principal  $G$ -Bundle* on  $X$ . That is,  $X$  has some bundle over it (The space  $E$  from our previous discussion), and  $G$  acts on the fibers of  $X$ . This is denoted  $\text{Prin}_G(X)$ .

Construction by John Milnor, Classifying space. No idea why I wrote this...

If  $G$  is a group, there is a complex (space)  $BG$  such that we may form the set:

$$[\mathcal{M}, BG] = \{f : \mathcal{M} \rightarrow BG\}/\text{Homotopy}$$

That is, the set of continuous maps from  $\mathcal{M}$  to  $BG$  modded out by homotopy. Two maps are equivalent if they are homotopic.

**Theorem 40.1.3.** *There is a continuous surjective function  $f : \text{Prin}_G(\mathcal{M}) \rightarrow [\mathcal{M}, BG]$ .*

Elaborating more on the  $BG$ ,  $B$  is a functor  $B : \text{Groups} \rightarrow \text{Spaces}$ . If  $G$  is a finitely presented group, then  $\pi_1(BG) = G$ , and more over, for all  $n \geq 2$ ,  $\pi_n(BG) = 0$ . That is,  $\pi_n(BG)$  is the trivial group for all  $n \geq 2$ .  $\pi_1(X)$  can be seen as the *homotopy class* of  $[S^1, X]$ .  $\pi_n$  the homotopy class for  $[S^n, X]$ .

**Example 40.1.7**  $\pi_1(B\mathbb{Z}) = \mathbb{Z}$ . For all  $n \geq 2$ ,  $\pi_n(B\mathbb{Z}) = 0$ . Thus  $B\mathbb{Z}$  is homotopy equivalent to  $S^1$ .

## Covering Spaces

**Definition 40.1.7** A covering space of a topological space  $X$  is a space  $E$  such that there exists a continuous surjection  $p : E \rightarrow X$  such that for all  $x \in X$ , there is an open set  $\mathcal{U}$  such that  $x \in \mathcal{U}$  such that there exists a set of disjoint open sets  $E_r \subset E$  where  $p^{-1}(\mathcal{U}) = \cup_r E_r$  and for all  $r$ ,  $p$  is a homeomorphism between  $E_r$  and  $\mathcal{U}$ .

**Example 40.1.8** The first example is  $S^1$  and  $\mathbb{R}$ . Define the map  $p : \mathbb{R} \rightarrow S^1$  by  $p(x) = \exp(2\pi ix)$ . This “wraps,” the real line around the circle over and over again. Given a point  $y \in S^1$ , the pre-image, or fiber, of  $y$  with respect to  $p$  is  $\{x + n : n \in \mathbb{Z}\}$  for some  $x \in [0, 1)$ . Given a small enough neighborhood around  $y$ , the pre-image is of the form  $\{x + n - \varepsilon, x + n + \varepsilon : n \in \mathbb{Z}\}$ , which is a bunch of copies of  $(0, 1)$ , or a bunch of copies of the neighborhood around  $y$ .

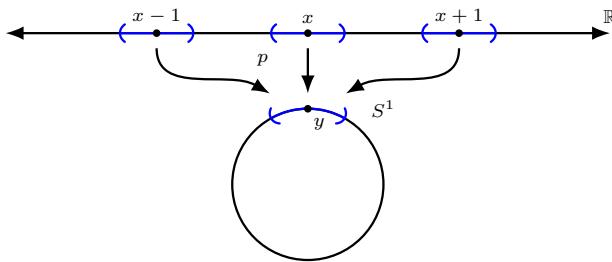


Fig. 40.8:  $\mathbb{R}$  is the Universal Cover of  $S^1$ .

**Definition 40.1.8** A universal covering space of a topological space  $X$  is a covering space  $E$  of  $X$  such that  $E$  is simply connected.

That is, if  $E$  is a covering space of  $X$ , then we say that  $E$  is a universal covering space if  $\pi_1(E) = 0$ . In the previous example we saw that  $\mathbb{R}$  is a covering space of  $S^1$ . But  $\mathbb{R}$  is simply connected. That is,  $\pi_1(\mathbb{R}) = 0$ . Therefore  $\mathbb{R}$  is a universal covering space of  $S^1$ . Up to homotopy equivalence,  $B\mathbb{Z}^n = T^n$ , the  $n$  torus. This is because  $\pi_1(B\mathbb{Z}^n) = \mathbb{Z}^n$ , and  $\pi_n(B\mathbb{Z}^n) = 0$  for  $n \geq 2$ . For  $S^n$ , if  $n \geq 2$  then  $S^n$  is simply connected,  $\pi_1(S^n) = 0$ . But then the identity map makes  $S^n$  a covering space for itself. That is,  $id_{S^n}$  is a covering map. But

since  $S^n$  is simply connected ( $n \geq 2$ ), we have that  $S^n$  is a universal covering of itself. Moreover, it can be shown that  $S^n$  is a universal covering space of  $\mathbb{R}P^n$  for  $n \geq 2$ . All universal covering spaces are homotopy equivalent to each other.

## Eilenberg-MacLane Spaces

**Definition 40.1.9** An Eilenberg-MacLane space is a topological space  $X$  such that there exists a non-trivial group  $G$  and an  $n \in \mathbb{N}$  such that  $\pi_n(X) = G$  and, for all  $m \neq n$ ,  $\pi_m(X) = 0$ .

This  $B$  functor takes a group  $G$  and spits out an Eilenberg-MacLane space. That is,  $\pi_1(BG) = G$ , and  $\pi_n(BG) = 0$  for all  $n \geq 2$ . Eilenberg-MacLane spaces are analogous to prime numbers in Number Theory, but for the study of topological spaces. These have a special notation:

### Notation 40.1.1

An Eilenberg-MacLane space  $X$  is of the type  $K(G, n)$  if  $\pi_n(X) = G$  and, for all  $m \neq n$ ,  $\pi_m(X) = 0$ . We write  $X \in K(G, n)$ .

Every principal  $G$  bundle over  $\mathcal{M}$  can be imagined as  $[\mathcal{M}, BG]$ . A principal  $\mathcal{O}(n)$  bundle can be identified with a map  $f : \mathcal{M} \rightarrow B\mathcal{O}(n)$ . For example,  $f$  be the constant map. Constant maps are homotopic to each other. This is the easiest bundle.

## 40.1.2 Lecture 4: Principal G-Bundles

A brief discussion on complexes. A simplex is a generalization of the notation of a triangle. A triangle can be considered as the convex-hull of 3 non-coplanar points. This is called a 2-simplex. A 0-simplex is thus a point, and a 1-simplex is a line. This can be generalized to higher dimensions. A 3-simplex is a tetrahedron, and an  $n$ -simplex is an  $n$  dimensional triangle, defined on  $n + 1$  non-hyper-coplanar points.

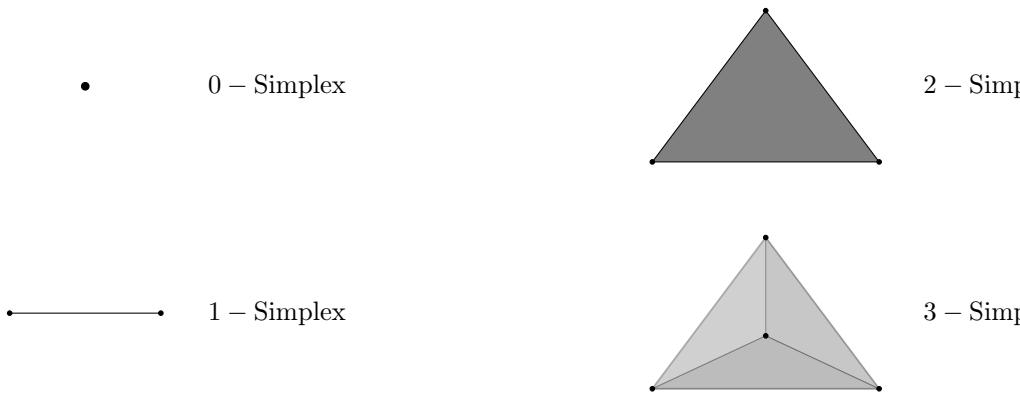


Fig. 40.9: Examples of Simplices.

A simplicial complex is a set of simplices  $\mathcal{H}$  such that the face of any element of  $\mathcal{H}$  is also contained in  $\mathcal{H}$ , and the intersection of two simplices  $\sigma_1, \sigma_2 \in \mathcal{K}$  is a face of both  $\sigma_1$  and  $\sigma_2$ . We return to studying surgery exact sequences for  $n \geq 5$ . Let  $\mathcal{M}$  be an  $n$  dimensional manifold, and let  $G = \pi_1(\mathcal{M})$ . In our surgery exact sequence we still have this mysterious object  $[\mathcal{M}, G/Cat]$ . Let Cat be either PL or Top. The generalized Poincare Conjecture says that, for  $n \geq 5$ ,  $S^{PL}(S^n) = S^{Top}(S^n) = \{S^n\}$ . Let  $\mathcal{M} = S^n$ . Then  $G = \pi_1(\mathcal{M}) = \{e\}$ . Then we have the following:

$$\begin{array}{ccccccc}
 L_6(\mathbb{Z}) & \longrightarrow & S^{Cat}(S^5) & \longrightarrow & [S^5, G/Cat] & \longrightarrow & L_5(\mathbb{Z}) \\
 & & & & \downarrow & & \\
 & & & & [S^5, G/Cat] & \longrightarrow & 0 \\
 & & & & \downarrow & & \\
 & & & & 0 & \longrightarrow & \pi_5(G/Cat) \longrightarrow 0
 \end{array}$$

Fig. 40.10: Diagram for the Surgery Exact Sequence of  $S^5$ .

So,  $\pi_5(G/Cat) = \{e\}$ . This gives us:

$$\begin{aligned}
 \cdots &\rightarrow S^{Cat}(S^6) \rightarrow [S^6, G/Cat] \rightarrow L_6(\mathbb{Z}) \rightarrow \cdots \\
 &\cdots \rightarrow 0 \rightarrow \pi_6(G/Cat) \rightarrow \mathbb{Z}_2 \rightarrow 0
 \end{aligned}$$

So, we have  $\pi_6(G/Cat) \cong \mathbb{Z}_2$ . In general,  $\pi_n(G/o) \cong L_n(\mathbb{Z})$ .

**Theorem 40.1.4** (Wall's Theorem).

$$L_n(\mathbb{Z}) = \begin{cases} \mathbb{Z}, & n \equiv 0 \pmod{4} \\ 0, & n \equiv 1 \pmod{4} \\ \mathbb{Z}_2, & n \equiv 2 \pmod{4} \\ 0, & n \equiv 3 \pmod{4} \end{cases}$$

All  $L$  groups are periodic, and never have odd torsion. That is, there is never  $\mathbb{Z}_3, \mathbb{Z}_5$ , etc. Wall groups are hard to compute. Whatever  $G/Cat$  is, its homotopy groups for  $n \geq 5$  are known.

## Principle G-Bundles

A few things are needed:

- Map  $p : E \rightarrow X$ , where  $E$  is a total space and  $X$  is a base space.
- The inverse-image  $E_x = p^{-1}(\{x\})$  is called the fiber over  $x$ .
- $G$  (Group) acts on each  $E_x$  freely and transitively.
- $G$  has to act ‘continuously.’ Nearby points are taken to nearby points.

Then  $p : E \rightarrow X$  is a  $G$ -principle bundle. Freely means the only element that fixes everything is the identity.

**Example 40.1.9** Take a sphere  $S^n$  and a projection  $p : S^n \rightarrow \mathbb{RP}^n$ .  $\mathbb{RP}^n$  is created by glueing antipodal points together. If  $x \in \mathbb{RP}^n$ , then  $p^{-1}(\{x\})$  consists of 2 antipodal points in  $S^n$ . Now  $\mathbb{Z}_2 = \{0, 1\}$  can act on a sphere. 0 maps  $x \mapsto x$  and 1 maps  $x \mapsto -x$ . Note that  $1 + 1 = 0$ , as in  $\mathbb{Z}_2$ . Given any point, you can get to another point in the fiber. This is trivial in this example as there are only two points in the fiber. Also only the identity maps a point back to itself. This action is free and transitive, so  $p : S^n \rightarrow \mathbb{RP}^n$  is a  $\mathbb{Z}_2$ -principle bundle.

Let  $M$  be a manifold (Or a space) with dimension  $n$  and fundamental group  $G$ . A universal cover  $\tilde{M}$  of  $M$  includes a map  $p : \tilde{M} \rightarrow M$  such that  $\tilde{M}$  is simply connected of dimension  $n$ , i.e.  $\pi_1(\tilde{M}) = e$ , and  $\forall_{x \in M}, p^{-1}(x)$  is a collection of discrete points.  $\pi_1(\mathbb{RP}^n) = \mathbb{Z}_2$  and  $S^n$  is a universal cover of  $\mathbb{RP}^n$ . This might come from a general theory.

**Example 40.1.10** Take the circle  $S^1$ .  $\pi_1(S^1) = \mathbb{Z}$ . There's a map  $p(x) = e^{2\pi ix}$  of modulus 1. Note that  $p^{-1}(0) = \mathbb{Z}$ . So  $p^{-1}(x)$  is just a shift of  $\mathbb{Z}$  to  $\mathbb{Z} + r$ . Note that  $\pi_1(\mathbb{R}) = e$ . So  $\mathbb{R}$  is a universal cover of  $S^1$ .

**Example 40.1.11** We may think  $\mathbb{R}^2$  is a universal cover of  $S^2$ , but  $S^2$  is already simply connected. So  $p$  is the identity map, and the universal cover of  $S^2$  is  $S^2$ . All universal covers are homotopy equivalent.

Let  $x \in \mathcal{M}$ . Then, for all  $z \in p^{-1}(\{x\})$ , and for all  $g \in \pi_1(M)$ , there is an action  $gx \in p^{-1}(x)$ . This uses the homotopy lifting property. There is an action  $G$  on  $\tilde{\mathcal{M}}$  which preserves the fiber (Takes every element of a fiber to the same fiber. It does not mix fibers), is transitive, and is free. The map  $p : \tilde{\mathcal{M}} \rightarrow \mathcal{M}$  is a  $\pi_1(M)$  Principal Bundle.

## Functors

Let  $F$  be a functor  $F : Space \rightarrow Groups$ . So for all spaces  $X$ , we have a group  $F(X)$ . There are many such examples:

- Cohomology
- Homology
- K-Theory
- Other Stuff

**Homology:** Take  $M$  and triangulate. Take maps from the simplicial complex of  $M$  to  $G$  (Group) (Certain conditions). There's an equivalence relation on these maps. That set after taking the equivalence relations is the homology:  $H_n(M, G)$ .  $n$  describes the type of simplices. If  $n > \dim(M)$ , then  $H_n(G, M) = 0$ .  $H_n(M, G) = \{f : \Delta^n \rightarrow G\}$ . Cohomology is the set  $H^n(M, G) = [H_n(M, G), G]$ , that is, the *dual*. We want to talk about cohomology. Under special conditions there is something called the Brown Representation Theorem. Consider Cohomology  $H^n(M, G)$ , with coefficients in  $G$ . Cohomology is Homotopy invariant, that is if  $M \cong N$ , then  $H^n(M, G) \cong H^n(N, G)$ . The Brown-Representation Theorem says that there is a classifying space  $BG$  such that, for all spaces  $M$ , there is a one-to-one correspondence between  $H^n(X, G) \leftrightarrow [X, BG]$ . In general, if  $F$  is a functor, then the Brown-Representation Theorem says that there is a classifying space  $Y$  such that  $F(X)$  has a one-to-one correspondence with the homotopy classes of maps,  $[X, Y]$ .  $F(x) \leftrightarrow [X, Y]$ .

**Example 40.1.12** The Eilenberg-MacLane Space  $K(G, n)$  has the property that  $\forall_{j \neq n}, \pi_n(K(G, n)) = G$ , and  $\pi_j(K(G, n)) = 0$ .  $K(G, n)$  is the classifying space for cohomology.

**Theorem 40.1.5.**  $K(G, n)$  is the classifying space for cohomology. That is, up to homotopy,  $H^n(X, G) \leftrightarrow [X, K(G, n)]$ .

Let  $\text{Prin}_G(X)$  be the collection of  $G$ -principal bundles on  $X$ . With a certain equivalence relation, it turns out the  $\text{Prin}_G(X)$  is a group. So  $\text{Prin}_G : Spaces \rightarrow Groups$  is a functor. The Brown-Representation Theorem implies that there is a classifying space  $BG$   $\text{Prin}_G \leftrightarrow [X, BG]$ .

**Theorem 40.1.6.** If  $p : E \rightarrow X$  and  $p' : E' \rightarrow X$  are both bundles over  $X$ , then there exists  $p \oplus p' : E \oplus E' \rightarrow X$  COME BACK TO LATER

## Grothendieque Groupification of Semigroup

**Definition 40.1.10** A semi-group is a group without the requirement for inverses.

**Example 40.1.13**  $\{0, 1, 2, \dots\}$  is a semi-group under addition.

Let  $G$  be a semi-group. Constraint  $G \times G / \sim$ .  $(a, b) \sim (c, d)$  if  $a + d = b + c$ . So,  $(2, 3) \sim (4, 5) \sim (7, 8) \sim (-1, 0) \equiv -1$ . The equivalence class of all of these things is called  $-1$ . We still have all of the positive integers,  $(4, 2) \sim (5, 3) \sim (6, 4) \equiv 2$ . This process adds all of the negatives. This process, called Grothendieque Construction on a Semi-group creates a group out of a semi-group. It is, in a way, the 'smallest' group containing the semi-group. The groupification of  $\{0, 1, 2, 3, \dots\}$  will be  $\mathbb{Z}$ .

**Example 40.1.14** What are the vector bundles over a dot? There is  $\mathbb{R}^0$  (A dot),  $\mathbb{R}^n, \dots, \mathbb{R}^n, \dots$  There is an operation on this set  $\{\mathbb{R}^n : n \geq 0\}$ . This makes a semi-group, and there is a Grothendieque Groupification  $G_r(\mathbb{Z}_{\geq 0}, +) = \mathbb{Z}$

Suppose  $M$  is a monoid/semigroup. Not required to have an inverse but should have an identity. For example,  $(\mathbb{Z}_{\geq 0}, +)$  is a monoid. Has identity, but no inverse. Construct  $M \times M = \{(a, b) : a, b \in M\}$  with the operation  $(a, b) + (c, d) = (a + c, b + d)$ . Think of  $(a, b)$  as  $a - b$ . Note that in regular math  $'3 - 1' = '4 - 2'$ , so we want  $(3, 1)$  to equal  $(4, 2)$ . We do this with the equivalence relation  $(a, b) \sim (c, d)$  if and only if  $a + d = b + c$ . Let  $M \times M / \sim$  be called  $M_G$ .

**Theorem 40.1.7.**  $M_G$  is a group.

**Theorem 40.1.8.** There is an injection  $i : M \rightarrow M_G$  with the following property:

1.  $i(a) \sim (a, 0) \sim (a + 1, 1) \sim (a + 2, 2) \sim \dots$

This construction is functorial, so if there are monoids  $M, N$  with a semi-homomorphism  $\phi : M \rightarrow N$  ( $\phi(a * b) = \phi(a) * \phi(b)$ ) (Homomorphism for a semi-group), then there is a HM  $\phi_G : M_G \rightarrow N_G$  so  $G(Monoids, Semihomomorphism) \rightarrow (Groups, Homomorphism)$  is a functor.

## Suspension

Let  $X$  and  $Y$  be disjoint topological spaces. The wedge product  $X \vee Y$  is the one-point union of  $X$  and  $Y$ . Take  $X$ , take  $Y$ , and glue one point together.

**Theorem 40.1.9.** If  $X$  and  $Y$  are disjoint topological spaces, then  $\pi_1(X \vee Y) = \pi_1(X) \oplus \pi_1(Y)$ .

In the same context, the smash product of  $X$  and  $Y$  is  $X \wedge Y = X \times Y / (X \vee Y)$ . Picture  $X = (0, 1]$  in the  $x$  axis and  $Y = (0, 1]$  in the  $y$  axis. Then  $X \times Y$  is a

square in the  $xy$  plane, and  $X \vee Y$  is the  $x$  and  $y$  axes from 0 to 1. So  $X \wedge Y$  takes all of the points on the two axes between 0 and 1 and smashes them down to the origin.

**Example 40.1.15**  $S^1 \times [0, 1]$  is the hollow cylinder, and  $S^1 \vee [0, 1]$  is the boundary of the edge of the cylinder (The lid) and the line going down the cylinder parallel with the  $z$ -axis (the spine). So  $X \wedge Y$  smashes down to a cone. This is then homeomorphic to  $D^2$ .

**Example 40.1.16** The torus can be visualized by the diagram shown in Fig. ???. So  $T^2 = S^1 \times S^1$ . Using the diagram, we can see that the smash product  $S^1 \wedge S^1$  is homotopy equivalent to a sphere.

**Definition 40.1.11** Let  $X$  be a topological space. Then the suspension of  $X$ , denoted  $\Sigma X$ , is  $S^1 \wedge X$ .

So  $\Sigma S^n = S^{n+1}$ . The usefulness of smash has to do with  $\pi_k(\Sigma X) = \pi_{k-1}(X)$ . There's another thing called the Freudenthal suspension theorem.

## Higher Homotopy

The fundamental group, which is the first homotopy group, is  $\pi_1(X)$ . Ingredients needed:

1. Topological Space  $X$
2. A basepoint  $x_0$

**Definition 40.1.12** The fundamental group of a topological space  $X$  about a base point  $x_0$  is the set:

$$\pi_1(X) = [(S^1, \star), (X, \star)] = \text{Hom}((S^1, \star), (X, \star)) = \{f : S^1 \rightarrow X : f(x) = \star\}/\text{Homotopy} \quad (40.1.2)$$

$\pi_1(X)$  is a group using concatenation. Higher homotopy groups:

**Definition 40.1.13**  $\pi_n(X) = [(S^n, \star), (X, \star)]$

It turns out that  $\pi_n(X)$  has a certain operation, for  $n \geq 2$ , such that it is an Abelian group. However,  $\pi_1(X)$  need not be an Abelian group.

**Example 40.1.17** The Klein bottle is an example of a space such that  $\pi_1(X)$  is not an Abelian group.

The question becomes 'What are the Homotopy groups of sphere?' That is, what is  $\pi_m(S^n)$ ? Recall stereographic projection from before. Take  $S^n$  and remove the north pole (The point  $(0, 0, 1)$ ). This can be projected down to  $\mathbb{R}^2$ . This can be generalized to  $n$  dimensions, and in general  $S^n \setminus \{\text{North Pole}\}$  is homeomorphic to  $\mathbb{R}^n$ . Now  $\mathbb{R}^n$  has 0 homotopy groups because it is contractible (Can be smushed down to a point). If  $m < n$ , then we are mapping a small 'sphere' into a big 'sphere'.

**Theorem 40.1.10.** *If  $n \neq m$ , then there is no continuous function  $f$  such that  $f : S^n \rightarrow S^m$  is surjective.*

Now, if  $m < n$ , then we map  $S^m$  into  $S^n$ . But since there is no surjective continuous function we can remove a point from  $S^n$ , map it down to  $\mathbb{R}^n$  and then contract. So,  $\pi_m(S^n) = 0$  for all  $m < n$ . The next case is when  $m = n$ . There are three obvious maps: The constant map, the identity map, and the antipodal map. It can be shown that there are, for all  $n$ , countably many maps. So  $\pi_n(S^n) = \mathbb{Z}$ . Another neat little fun fact is that  $\pi_3(S^2) = \mathbb{Z}$ . (Related to Hopf fibration). Now, the suspension theorem says that  $\pi_{n+1}(\Sigma X) = \pi_n(X)$ . So  $\mathbb{Z} = \pi_3(S^2) = \pi_4(\Sigma S^4) = \pi_4(S^3) = \pi_5(\Sigma S^3) = \pi_5(S^4) = \dots$  so, if  $m - n = 1$ , then  $\pi_m(S^n) = \mathbb{Z}$ .

**Theorem 40.1.11.**  $\pi_3(S^2) = \mathbb{Z}$

**Theorem 40.1.12.** *If  $m - n = 1$ , and  $n \geq 2$ , then  $\pi_m(S^n) = \mathbb{Z}$*

These are examples of stability theorems, or stability results.

## Fibrations

**Definition 40.1.14** A fibration is a map between topological spaces that has the homotopy lifting property for every space  $X$ .

A fibration gives rise to a long exact sequence of homotopy groups

$$\begin{aligned} \pi_3(S^1) \rightarrow \pi_3(S^3) \rightarrow \pi_3(S^2) \rightarrow \pi_2(S^1) \rightarrow \pi_2(S^3) \rightarrow \pi_2(S^1) \rightarrow \dots \\ \dots \rightarrow \pi_2(S^3) \rightarrow \pi_2(S^2) \rightarrow \pi_1(S^1) \rightarrow \pi_1(S^3) \rightarrow \pi_1 \end{aligned}$$

We need to know that  $\pi_n(S^1) = 0$  if  $n \geq 2$ . An element of  $\pi_n(S^1)$  is  $f : S^n \rightarrow S^1$ . Stanley owe's me an explanation. This becomes:

$$0 \rightarrow \mathbb{Z} \rightarrow A \rightarrow 0 \rightarrow 0 \rightarrow \mathbb{Z} \rightarrow \mathbb{Z} \rightarrow 0 \rightarrow 0$$

## Classifying Space

If  $G$  is a group (discrete or not), then there is a classifying space (Topological space)  $BG$  (unique up to homotopy) such that :

1.  $\pi_1(BG) = G$  and  $\pi_n(BG) = 0$  for all  $n \geq 2$ .
2. There is a contractible space  $EG$  that is a principle  $G$  bundle with a  $G$  action such that  $BG \simeq EG/G$ .  $EG \rightarrow BG$ .
3. For all spaces  $X$  with a continuous map  $f : X \rightarrow BG$ , there is a pullback diagram

4. The correspondence  $(X \rightarrow BG) \rightarrow (Y_f \downarrow X)$  has the property that, if  $f \simeq g$ , then  $(Y_f \downarrow X) \xrightarrow{\text{Principle G-Bundle}} (Y_g \downarrow X)$ . This gives a map  $[X, BG] \rightarrow \text{Prin}_G(X)$  which is a bijection.

**Definition 40.1.15** Let  $V$  be a finite dimensional vector space over  $\mathbb{F}$ . We say that  $B : V \times V \rightarrow \mathbb{F}$  is symmetric bilinear if:

1.  $B(v, v') = B(v', v)$
2.  $B(v + w, v') = B(v, v') + B(w, v')$
3.  $B(\lambda v, w) = \lambda B(v, w)$

**Example 40.1.18** Let  $A$  be a symmetric real  $n \times n$  matrix. Then  $B : \mathbb{R}^n \rightarrow \mathbb{R}^n \rightarrow \mathbb{R}$  given by  $B(X, Y) = X^T A Y$ . This  $B$  gives rise to a map  $Q : V \rightarrow \mathbb{F}$  given by  $Q(x) = B(x, x)$ .

$$Q(X) = [x, y] \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} [x \quad y] = 2x^2 + y^2$$

Which is a quadratic form.

### 40.1.3 Lecture 5: The Wall L-Groups

The Wall L-Groups are defined on all commutative rings. In fact, there is a functor  $L_n$  which takes commutative rings to groups. Some facts about this:

1. L-Groups are 4 periodic. For all commutative rings  $R$ , we have  $L_n(R) \simeq L_{n+4}(R)$ . This is hard to prove.
2. Surgery theory requires for  $R = \mathbb{Z}G$ , where  $G$  is the fundamental group of the manifold in question, and  $\mathbb{Z}G$  is all finite linear combinations of the elements in  $G$ .
3. There is a whole theory of computing  $L_n(R)$  when  $R$  is a field, usually denoted  $\mathbb{F}$ .
4. Potentially true statement: L-groups only have 2 or 4 torsion, if they have torsion at all. This is hard to prove, as well.

5. For  $G$  equal to the trivial group,  $L_n(\mathbb{Z}[e])$  we have  $L_n(\mathbb{Z}[e]) = \begin{cases} \mathbb{Z}, & n \cong 0(4) \\ 0, & n \cong 1(4) \\ \mathbb{Z}_2, & n \cong 2(4) \\ 0, & n \cong 3(4) \end{cases}$

Suppose  $\mathcal{M}^n$  is a closed manifold and  $n \geq 5$ , and  $\pi_1(\mathcal{M}) = e$ . Suppose  $n = 5$ . Then:

$$\begin{aligned} L_6(\mathbb{Z}[e]) &\longrightarrow [\mathcal{M}, G/Cat] \longrightarrow S^{Cat}(\mathcal{M}) \longrightarrow L_5(\mathbb{Z}[e]) \\ \mathbb{Z}_2 &\longrightarrow [\mathcal{M}, G/Cat] \xrightarrow{f} S^{Cat}(\mathcal{M}) \longrightarrow 0 \end{aligned}$$

If  $n = 6$ , we have:

$$\begin{aligned} L_7(\mathbb{Z}[e]) &\rightarrow [M, G/Cat] \rightarrow S^{Cat}(\mathcal{M}) \rightarrow L_6(\mathbb{Z}[e]) \\ 0 &\rightarrow [M, G/Cat] \xrightarrow{g} S^{Cat}(\mathcal{M}) \xrightarrow{\text{?}} \mathbb{Z}_2 \end{aligned}$$

In the case of  $n = 4$ , there are these things called 'Good' groups in which some of these results still hold. The dimensions can be broken up like this:

- 2 Completely solved.
- 3 This is Knot Theory.
- 4 Very hard.
- $\geq 5$  Surgery Theory.

How do you classify manifolds?

- In two dimensions the genus (number of wholes) and the orientation (Euler characteristic) gives you everything.
- In three dimensions, Thurston and Perelman did the classification of this.
- Four is a big vacuum of unsolved stuff. 'Good' groups come up here.
- For every group  $G$  there is a manifold  $\mathcal{M}$  of dimensions 5 or greater such that  $\pi_1(\mathcal{M}) = G$

Let's study  $L_n(\mathbb{Z}[e])$ . This is surprisingly hard enough to study. The computation of this was known by Brouder, but the use of the surgery exact sequences wasn't done until Wall (Hence, Wall groups).

### The Witt Group

$L_0(\mathbb{Z}[e])$  is equal to something called the Witt group  $Witt(\mathbb{Z})$ . First let's talk about the Witt group of fields. The Witt group of a field  $\mathbb{F}$  is the set of symmetric bilinear forms  $B : V \times V \rightarrow \mathbb{F}$  of finite dimensional vector spaces  $V$  over  $\mathbb{F}$ , modulo some equivalence relation. So  $B$  can be represented by some

symmetric matrix in  $M_n(\mathbb{F})$  with respect to some basis  $\{\beta\}$ . So if  $B : V \times V \rightarrow \mathbb{F}$  has matrix  $A_\beta$  and if  $D : V \times V \rightarrow \mathbb{F}$  has matrix  $A'_\delta$ , then construct the matrix:

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & A' \end{pmatrix}_{\beta, \delta}$$

Then one can get a Bilinear form  $B \perp D : V \oplus V \rightarrow V \oplus V$  using this matrix. This is called the orthogonal sum. The equivalence relation on these Bilinear forms is a bit complicated. Consider the matrix:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix} = x^2 - y^2$$

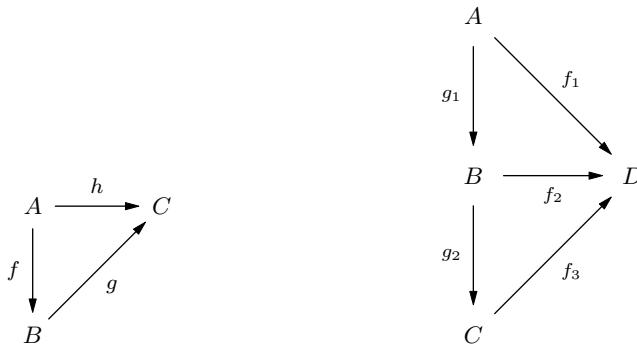
This is called a Hyperbolic form  $H_2(\mathbb{F})$

1. Two forms  $B_1 : V \times V \rightarrow \mathbb{F}$  and  $B_2 : W \times W \rightarrow \mathbb{F}$ , with  $\dim(V) = \dim(W)$ . If  $A^T[B_1]A = [B_2]$ , then  $B_1 \sim B_2$ .
2. We can also write  $B_1 \sim B_2$  is  $B_1 = B_2 \perp_k H_2(\mathbb{F})$ . So  $H_2(\mathbb{F})$  is the 0 element in  $Witt(\mathbb{F})$ . Note that, since  $H_2(\mathbb{F})$  has dimension 2,  $\dim(B_1) = \dim(B_2) \pmod{2}$ .
3. What is the inverse of  $B_1$ ? It is a form  $B_2$  for which  $B_1 \perp B_2 \simeq H_2(\mathbb{F})$

There is a map, called the signature map, of a matrix  $W(\mathbb{F}) \rightarrow L_0(\mathbb{Z}[e]) \simeq \mathbb{Z}$ . It is defined for matrices with real eigenvalues. It is the number of positive eigenvalues minus the number of negative eigenvalues.

## Manifold Structures

Let  $X$  be a closed manifold. Then a homotopy equivalence  $f : \mathcal{N} \rightarrow X$  is called a manifold structure on  $X$ . Two manifold structures,  $f_1 : \mathcal{N}_1 \rightarrow \mathcal{M}$  and  $f_2 : \mathcal{N}_2 \rightarrow \mathcal{M}$ , are called equivalent on  $S(X)$  if there is a homeomorphism  $g : \mathcal{M} \rightarrow \mathcal{N}$  that Fig. 40.11.1 is a commutative diagram. Since the composition of homeomorphisms is a homeomorphism, if  $f_1 : \mathcal{L} \rightarrow X$  and  $f_2 : \mathcal{M} \rightarrow X$  are equivalent manifold structures on  $X$ , and if  $f_2 : \mathcal{M} \rightarrow X$  and  $f_3 : \mathcal{N} \rightarrow X$  are equivalent manifold structures, then  $f_1 : \mathcal{L} \rightarrow X$  and  $f_3 : \mathcal{N} \rightarrow X$  are equivalent manifold structures. That is, the diagram shown in Fig. 40.11.1 is a commutative diagram.



40.11.1: Commutative Diagram for Manifold Structures  
40.11.2: Equivalent Manifolds form an Equivalence Relation.

#### 40.1.4 Lecture 6: The Brown Representation Theorem

A functor  $f : \text{Spaces} \rightarrow \text{Groups}$  takes a topological space and returns a group. There are many examples, such as homology, cohomology, and K-Theory. Under certain circumstances there is a space  $B_{\circ f}$  such that there is a one-to-one functor  $f(\mathcal{M}) \leftrightarrow [\mathcal{M}, B_{\circ f}]$ , where  $\mathcal{M}$  is a manifold.

**Example 40.1.19** Let  $G$  be a group, and  $X \in K(G, n)$  an Eilenberg-MacLane space. This is not usually a manifold, but may be a complex, for example. As  $X \in K(G, n)$ , we have that  $\pi_n(X) = G$  and, for all  $m \neq n$ ,  $\pi_m(X) = 0$ . Then  $K(G, n)$  is the  $B_{\circ f}$ , where the  $\circ f$  is cohomology with coefficients in  $G$ . That is, we have  $H^n(\mathcal{M}; G) \leftrightarrow [\mathcal{M}, K(G, n)]$  is a one-to-one mapping.

Consider a manifold  $\mathcal{M}$  and the semi-group of vector bundles  $V(\mathcal{M})$  on  $\mathcal{M}$  with  $\oplus$  give by the *Whitney Sum* (More on that later). The Grothendieck construction gives us a group  $E(\mathcal{M})$  where the elements are vector bundles and “negative,” vector bundles (Virtual bundles).  $E$  can then be thought of as a functor from spaces to groups:  $E : \text{Spaces} \rightarrow \text{Groups}$ . There is some sloppiness ahead that will be clarified later. There is a space  $BO$  such that  $E(\mathcal{M}) \leftrightarrow [\mathcal{M}, BO]$  is a one-to-one mapping. Note that the  $B_{\circ f}$  are classifying spaces.  $BO$  is also a classifying space. Let  $\mathcal{O}(n)$  be the set of orthogonal matrices, as defined in a previous lecture.  $n \times n$  matrices such that  $A^T = A^{-1}$ . We saw before that there is a natural mapping  $\psi_n$  of  $\mathcal{O}(n)$  into  $\mathcal{O}(n+1)$ . We can then form the sequence:

$$\mathcal{O}(1) \xrightarrow{\psi_1} \mathcal{O}(2) \xrightarrow{\psi_2} \mathcal{O}(3) \xrightarrow{\psi_3} \mathcal{O}(4) \xrightarrow{\psi_4} \cdots \mathcal{O}(n) \xrightarrow{\psi_n} \cdots$$

And define  $\mathcal{O}$  to be the *direct limit* of this sequence. This is a subset of “infinite” dimensional matrices. Orthogonal matrices act on vector bundles, there are “rotations,” of the fibers in various dimensions. Let  $H$  be a vector bundle over  $\mathcal{M}$ . “Compactify,” the fibers, which are homeomorphic to  $\mathbb{R}^n$ , making them now

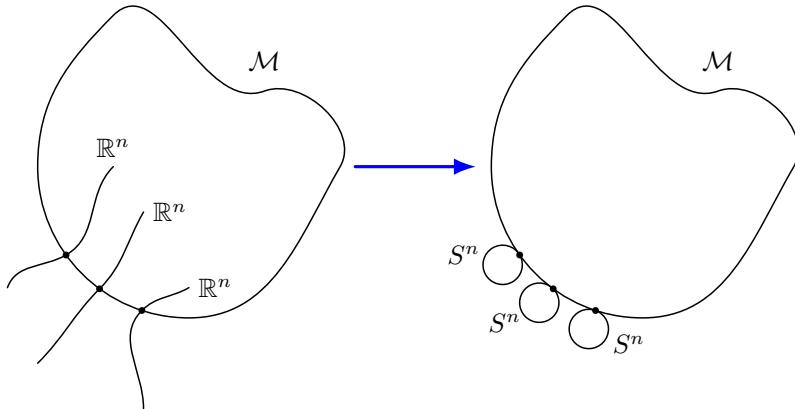


Fig. 40.12: Turning a Vector Bundle into a Sphere Bundle.

homeomorphic to  $S^n$ . This is, in a way, adding a point “at infinity,” or performing the one point compactification of  $\mathbb{R}^n$ . An example is the stereographic projection of the sphere onto the plane. The compactification of this is adding the “North Pole,” which was previously ignored as it is projected “to infinity.” The stereographic projection gives a bijection  $f : S^n \setminus \{\text{North Pole}\} \rightarrow \mathbb{R}^n$ . We now have the mapping  $f : \mathbb{R}^n \cup \{\infty\} \rightarrow (S^n \setminus \{\text{North Pole}\}) \cup \{\text{North Pole}\} = S^n$ . How do we make  $\mathbb{R}^n \cup \{\infty\}$  into a topological space? If  $\mathcal{U}$  is open in  $\mathbb{R}^n$ , we say that it is still open. Moreover, we say that  $[-\infty, a) = (-\infty, a) \cup \{\infty\}$  and  $(a, \infty] = (a, \infty) \cup \{\infty\}$  are also open sets. The topology on  $\mathbb{R}^n$  is then the topology generated by the three types of sets. Compactifying the fibers of a vector bundle creates something called a sphere bundle. An example is shown in Fig. 40.12. A spherical fibration is a space  $\mathcal{M}$  where at each point  $x$ , the fiber of  $x$  is equivalent to a sphere  $S^n$  and all *transition maps* are homotopy equivalent, rather than homeomorphic. The collection of all spherical fibrations on a manifold  $\mathcal{M}$  is a semigroup. The Grothendieck group associated with this is denoted  $S(\mathcal{M})$ . This group is a classifying space. So we have that a vector bundle on  $\mathcal{M}$  gives rise to a spherical fibration on  $\mathcal{M}$ .

$$\begin{array}{ccc} [\mathcal{M}, BO] & \longrightarrow & [\mathcal{M}, BG] \\ \text{Vector Bundle} & & \text{Spherical Fibration} \end{array}$$

We have the following diagrams to consider: Given  $\varphi$  and  $f$ , we can form  $\hat{\varphi} : \mathcal{M} \rightarrow BO$  by taking the composition,  $\hat{\varphi}(x) = (f \circ \varphi)(x)$ . The lifting problem is, given the central diagram, can we find a continuous map  $\hat{\varphi} : \mathcal{M} \rightarrow BG$  such that the diagram becomes commutative? The answer is not always. The final diagram gives rise to all “kernels,” like in exactness:  $BG/BO \dashrightarrow BO \rightarrow BG$ . We thus define  $G/O \equiv BG/BO$ . The homotopy collection  $[\mathcal{M}, G/O]$

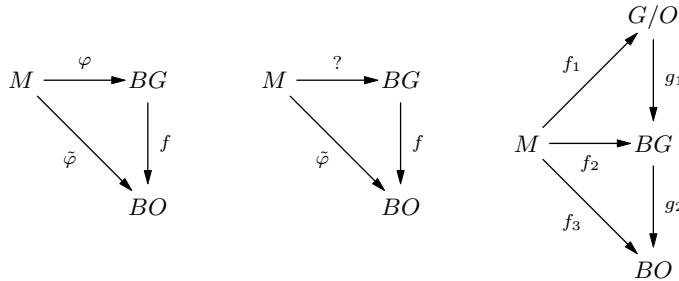


Fig. 40.13: Diagrams for the Lifting Property.

“computes,” the extent of which a map  $\psi : \mathcal{M} \rightarrow B$  can be “lifted,” to  $\hat{\varphi} : \mathcal{M} \rightarrow BO$ . Loops on  $\mathbb{R}^n$ , that is, continuous functions  $f : S^1 \rightarrow \mathbb{R}^n$ , can be contracted to a point. To put another way, no loops can wrap around “holes” in  $\mathbb{R}^n$ . Therefore,  $[\mathbb{R}^n, G/O]$  is a point. For more examples, we have  $[S^1, G/O] = \pi_1(G/O)$ , and  $[S^n, G/P] = \pi_n(G/O)$ . In our discussions  $O$  gives rise to a differentiable manifold  $BO$ . We can replace this with piece-wise linear or topological and obtain spaces like  $BPL$  or  $BT$ , respectively. From the Poincare Conjecture we have that  $S^{Top}(S^n) = S^{PL}(S^n) = \{e\}$ . That is, the trivial group. Our Surgery Exact Sequence now becomes:

$$\cdots \rightarrow L_{n+1}(\mathbb{Z}\pi) \rightarrow 0 \rightarrow [\Sigma \mathcal{M}, G/Top] \rightarrow L_n(\mathbb{Z}\pi) \rightarrow 0 \rightarrow \\ \rightarrow [\mathcal{M}, G/Top] \rightarrow L_{n-1}(\mathbb{Z}\pi) \rightarrow \cdots$$

So  $L_{n+1}(\mathbb{Z} \simeq [S^{n+1}, G/Top] = \pi_n(G/Top)$ , therefore  $L_{n+1}(G/Top) \simeq L_n(\mathbb{Z}\pi)$ .

**Learn about Suspending Space, Wall Groups, and Grothendieque complete of semigroup**

#### 40.1.5 Lecture 1: Singular and Simplicial Homology

The fundamental group  $\pi_1(X)$  is useful for studying low dimensional spaces. However, it is poor for studying higher dimensional spaces since, for example, it is unable to distinguish spheres of dimensions  $n \geq 2$ . The first solution to this is to study the homotopy groups  $\pi_n(X)$ . For this, we have that  $\pi_i(X) = 0$  for  $i < n$ , and  $\mathbb{Z}$  for  $i = n$ . A drawback is that homotopy groups are difficult to compute. The problem of  $\pi_i(S^n)$  is very difficult for when  $i > n$ . Homology groups,  $H_n(X)$  are one such solution to this difficulty. Homology groups share some characteristics with homotopy groups. If  $X$  is a CW complex, then  $H_n(X)$  depends only on the  $(n + 1)$ -skeleton of  $X$ . Also,  $H_i(S^n)$  and  $\pi_i(S^n)$  are isomorphic for  $1 \leq i \leq n$ . One benefit is that  $H_i(S^n) = 0$  for  $i > n$ .

## Simplicial Homology

The torus, projective plane, and Klein bottle can be created from a square by identifying opposite edges in certain ways. We can divide the square into two triangles, meaning these surfaces can be built from two triangles and then identifying edges. This can be done for all closed polygons as well. We can generalize this to  $n$  dimensions by considering the  $n$ –simplex, which is the convex hull of a set of  $n + 1$  points in  $\mathbb{R}^n$  that do not lie in an  $n$ –dimensional hyperplane.  $n$ –Simplexes are denoted  $\Delta^n$ . The interior of  $\Delta^n$  is denoted  $\overset{\circ}{\Delta}{}^n$ . A face of an  $n$ –simplex is a simplex formed by removing one of the vertices from  $\Delta^n$ .

**Definition 40.1.16** A  $\Delta$ –Complex on a space  $X$  is a set of maps  $\sigma_\alpha : \Delta^n \rightarrow X$  such that:

1. The restriction  $\sigma_\alpha|_{\overset{\circ}{\Delta}{}^n}$  is injected. Each point  $x \in X$  is in the image of only one such  $\overset{\circ}{\Delta}{}^n$ .
2. Each restriction of  $\sigma_\alpha$  to a face is equal to one of the maps  $\sigma_\beta : \Delta^{n-1} \rightarrow X$ .
3. A set  $A \subset X$  is open if and only if  $\sigma_\alpha^{-1}$  is open in  $\Delta^n$  for all  $\sigma_\alpha$ .

Let  $X$  be a  $\Delta$ –Complex, and let  $\Delta_n(X)$  be the free abelian group whose basis is the open  $n$ –simplices of  $X$ . Elements of  $\Delta_n(X)$  are called  $n$ –chains and can be written as  $\sum_\alpha n_\alpha e_\alpha^n$ , where  $n_\alpha$  is an integer.

**Definition 40.1.17** The Boundary Homomorphism  $\partial_n : \Delta_n(X) \rightarrow \Delta_{n-1}(X)$  is the map:

$$\partial_n(\sigma_\alpha) = \sum_i (-1)^i \sigma_\alpha |(v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$$

**Theorem 40.1.13.** *The composition of  $\Delta_n(X) \xrightarrow{\partial_n} \Delta_{n-1}(X) \xrightarrow{\partial_{n-1}} \Delta_{n-2}(X)$  is zero.*

**Definition 40.1.18** The  $n^{th}$  simplicial homology group of  $X$ , denoted  $H_n^\Delta(X)$ , is the quotient group  $\ker(\partial_n) / \text{Im}(\partial_{n+1})$  formed from the chain complex  $\cdots \rightarrow \Delta_n(X) \xrightarrow{\partial_n} \Delta_{n-1}(X) \xrightarrow{\partial_{n-1}} \Delta_{n-2}(X) \rightarrow \cdots$

The triangle on vertices  $a, b, c$  can be defined by  $[a, b, c]$ . We have:  $\partial_2([a, b, c]) = [b, c] - [a, c] + [a, b]$ . The negative sign on  $[a, c]$  is used to preserve orientation. That is, if you start at  $b$ , travel to  $c$ , then to  $a$ , and finally loop back to  $b$ , this should be the “same” as going from  $b$  to  $c$ , then “negative”  $a$  to  $c$ , and finally  $a$

to  $b$ . Note that then:

$$\partial_2 \partial_1([a, b, c]) = \partial_1([b, c] - [a, c] + [a, b]) \quad (40.1.3a)$$

$$= \partial_1([b, c]) - \partial_1([a, c]) + \partial_1([a, b]) \quad (40.1.3b)$$

$$= (b - c) + (c - a) + (a - b) \quad (40.1.3c)$$

$$= 0 \quad (40.1.3d)$$

We can perform the same calculation for the tetrahedron:

$$\begin{aligned} \partial_3 \partial_2([a, b, c, d]) &= \partial_2([b, c, d]) - \partial_2([a, c, d]) + \partial_2([a, b, d]) - \partial_2([a, b, c]) \\ &= ([c, d] - [b, d] + [b, c]) - ([c, d] - [a, d] + [a, c]) + \\ &\quad ([b, d] - [a, d] + [a, b]) - ([b, c] - [a, c] + [a, b]) \\ &= 0 \end{aligned}$$

Given a complex  $W$ , we look at the following chain:

$$C_{n+1} \xrightarrow{\partial_n} C_n \xrightarrow{\partial_{n-1}} C_{n-1} \xrightarrow{\partial_{n-2}} \cdots \longrightarrow C_2 \xrightarrow{\partial_1} C_1$$

Where  $\text{Im}(\partial_{n+1}) \subset \ker(\partial_n)$  and  $H_r^\Delta(W) = \ker(\partial_n)/\text{Im}(\partial_{n+1})$ . We call the elements of  $\text{Im}(\partial_{n+1})$  *boundaries*, and the elements of  $\ker(\partial_n)$  *cycles*.

**Example 40.1.20** Let  $W = [a, b]$ . That is, the line connecting  $a$  and  $b$ . Then  $C_2 = 0$  and  $C_1 = \mathbb{Z}\{[a, b]\} \simeq \mathbb{Z}$ . But also  $C_0 = \mathbb{Z}\{[a], [b]\} \simeq \mathbb{Z}^2$ . So we have  $0 \rightarrow C_1 \rightarrow C_0 \rightarrow 0$ . Now  $\text{Im}(\partial_1) \subset \ker(\partial_0)$ , and  $\partial_0$  is a mapping from  $\mathbb{Z}^2$  into 0, and thus the kernel of  $\partial_0$  is all of  $\mathbb{Z}^2$ . Moreover, since  $\partial_1$  is a homomorphism, either the image of  $\partial_1$  is 0 or  $\partial_1$  is injective. But  $\partial_1([a, b]) = b - a \neq 0$ , and therefore  $\partial_1$  is injective. So we have  $H_0^\Delta(W) = \mathbb{Z}^2/\mathbb{Z} = \mathbb{Z}$ . For  $n > 0$ ,  $H_n^\Delta(W) = 0$ .

**Example 40.1.21** Let  $W = [a, b, c]$ , a triangle including its interior. Then  $C_3 = 0$  and  $C_2 = \mathbb{Z}\{[a, b, c]\} \simeq \mathbb{Z}$ . Also  $C_1 = \mathbb{Z}\{b - a, c - a, c - b\} = \mathbb{Z}^3$ .  $\partial_2$  is not the zero mapping, and thus  $\text{Im}(\partial_2) = \mathbb{Z}$ . Also  $\ker(\partial_2) = 0$ , and thus  $H_2^\Delta(W) = 0/0 = 0$ . For  $\partial_1$  we have  $\partial_1([b, c]) = \partial_1([a, c]) - \partial_1([a, b])$ , and thus  $\text{Im}(\partial_1) = \mathbb{Z}^2$ . But then  $\ker(\partial_1) = \mathbb{Z}$ . Therefore  $H_1^\Delta(W) = \mathbb{Z}/\mathbb{Z} = 0$ . Finally,  $\partial_0$  is the zero mapping, and thus the kernel is  $\ker(\partial_0) = \mathbb{Z}^3$ . Therefore  $H_0^\Delta(W) = \mathbb{Z}^3/\mathbb{Z}^2 = \mathbb{Z}$

**Example 40.1.22** Let  $W = \{[a, b], [b, c], [a, c]\}$ , a triangle without its interior. Then we have that  $C_2 = 0$  and  $C_1 = \mathbb{Z}\{[a, b], [b, c], [a, c]\} = \mathbb{Z}^3$ .  $\partial_2$  is the zero mapping and thus  $\text{Im}(\partial_2) = 0$ . From the previous example we saw that  $\text{Im}(\partial_1) = \mathbb{Z}^2$ , and thus  $\ker(\partial_1) = \mathbb{Z}$ . Thus  $H_1^\Delta(W) = \mathbb{Z}/0 = \mathbb{Z}$ . Finally,  $H_0^\Delta(W) = \mathbb{Z}^3/\mathbb{Z}^2 = \mathbb{Z}$ .

We see that removing the interior of the triangle changed what the homology groups are. As will be seen later, the homology groups are way of determining what the “holes,” of the space are.

**Example 40.1.23** Let  $W$  be the union of a triangle  $[a, b, c]$  and a line segment  $[c, d]$ . Then  $C_3 = 0$ , and  $C_n = 0$  for all  $n > 3$ . The chain is then  $0 \rightarrow C_2 \rightarrow C_1 \rightarrow C_0$ .

**Theorem 40.1.14.** *If  $W$  is contractible and connected, then  $H_0^\Delta(W) = \mathbb{Z}$  and, for all  $n > 0$ ,  $H_n^\Delta(W) = 0$ .*

### Singular Homology

Singular homology is formed in the same way that simplicial homology is, with the requirement that the chains be formed by  $\Delta_n(X)$  being relaxed. Now, only the continuity of the  $\sigma$  are required. Therefore, when  $X$  is a space that can be “triangulated,” we have that the simplicial and singular homologies are the same. Suppose  $X$  and  $Y$  are orientable manifolds  $f: X \rightarrow Y$  continuous. What’s degree? Take the  $n$ th top homology,  $h_n(Z) - h_n(Y, Z)$ . This is the map  $f^*$  which is induced by  $f$ . Each one of these is a copy of  $\mathbb{Z}$ . So each one of these is a homomorphism. So it has to send the number  $1 \in \mathbb{N}$  somewhere. So the image of  $1$  is the degree of the map  $f$ . For example if  $f$  is constant, the degree is  $0$ . If  $f$  is the identity, then the degree is  $1$ . If you take a loop that wraps around  $S^1$  twice then the degree is  $2$ . This looks like the winding number, but is not quite that.

**Part XXV**

**Knot Theory**



## CHAPTER 41

---

### Knot Theory

---

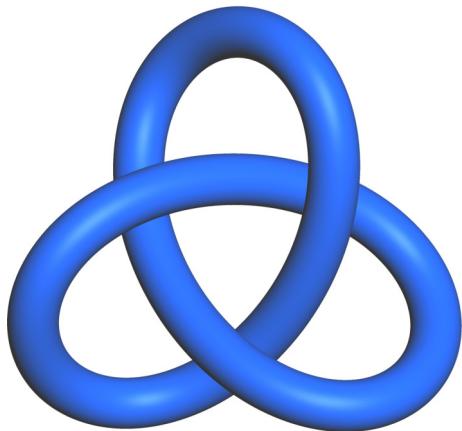


Fig. 41.1: A Trefoil Knot

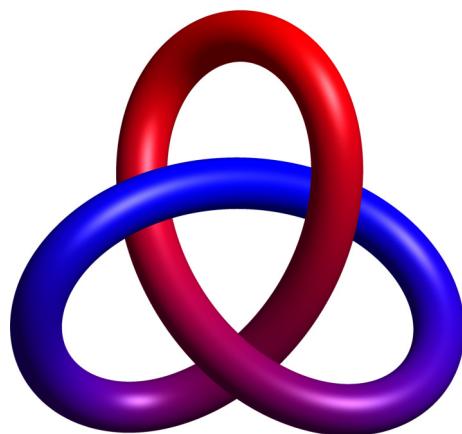


Fig. 41.2: A Colorful Trefoil Knot

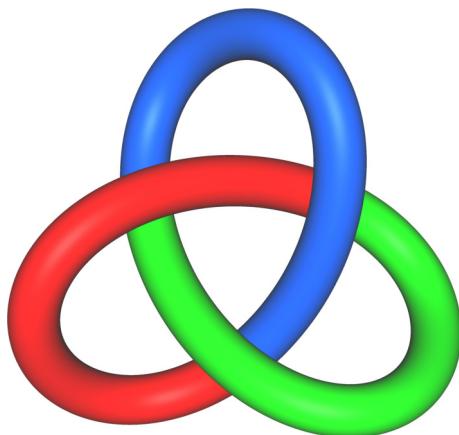


Fig. 41.3: Tricoloring of the Trefoil

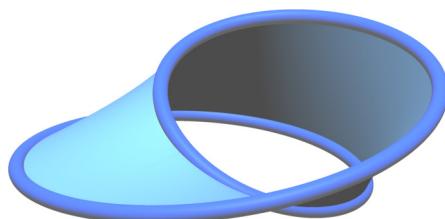


Fig. 41.4: Möbius Strip

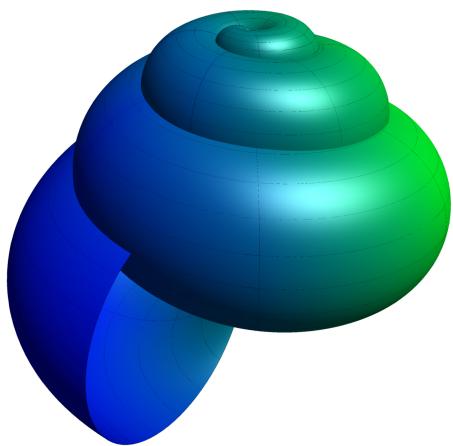


Fig. 41.5: A Sea Shell

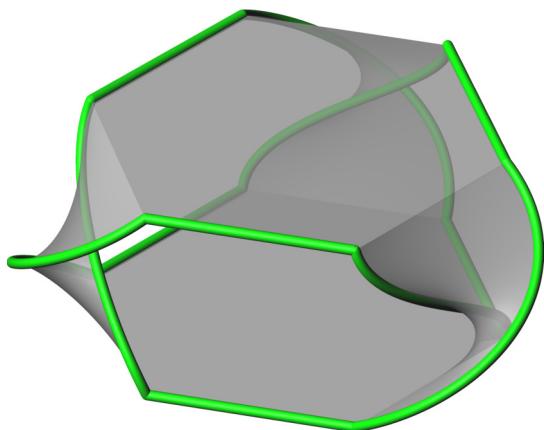


Fig. 41.6: Seifert Surface for a Trefoil Knot

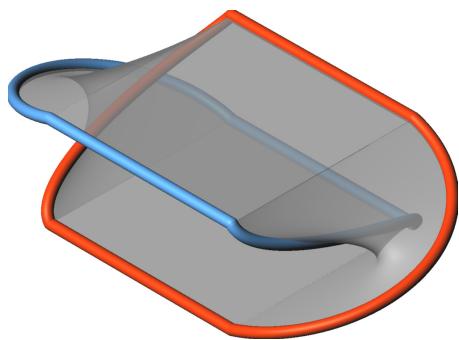


Fig. 41.7: Seifert Surface for a Hopf Link

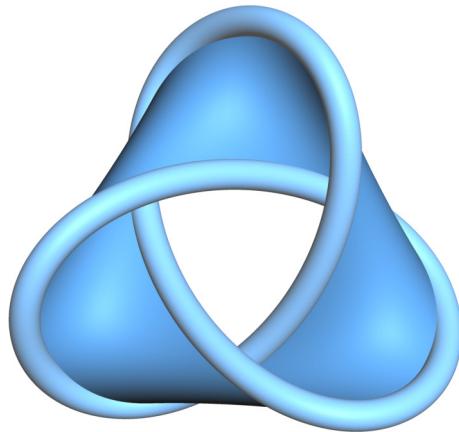


Fig. 41.8: Non-Oriented Surface With Trefoil Boundary

Book Seven

Physics



**Part XXVI**

# **Classical Mechanics**



## 41.1 Notes

### 41.1.1 Old Notes on Lagrangians

The Lagrangian is defined as  $\mathcal{L} = T - V$ , where  $T$  is the kinetic energy and  $V$  is the potential energy. Hamilton's Principle, which we take upon as an axiom of nature, states that the action of a system is an extremum. The action is defined as:

$$S = \int_{t_1}^{t_2} \mathcal{L}(x, \dot{x}, t) dt$$

This means that  $\mathcal{L}$  satisfies the Euler-Lagrange equation:

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = 0$$

This is the equation of motion in a one-dimensional Cartesian system. For general coordinates, we use:

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = 0.$$

Let's suppose we have a nice "Physics I," style Lagrangian. By that I mean that  $T = T(t)$ , and  $V \neq V(t)$ . That is, the potential energy is a function of position, and not of time. Then the equation of motion is:

$$\frac{\partial}{\partial q} (T - V) - \frac{d}{dt} \left( \frac{\partial}{\partial \dot{q}} (T - V) \right) = 0 \Rightarrow \frac{\partial V}{\partial q} + \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}} \right) = 0$$

Since  $\frac{\partial T}{\partial q} = 0$  and  $\frac{\partial V}{\partial \dot{q}} = 0$ . In a "Physics I" problem,  $T = \frac{1}{2}m\dot{q}^2$ . This is the kinetic energy. So we have:

$$m\ddot{q} = -\frac{\partial V}{\partial q}$$

Recall that Newton's Second Law says that:

$$m\ddot{x} = -\frac{\partial V}{\partial x}$$

This final result then looks very much like Newton's Second law. We thus define the following:

**Definition 41.1.1** The Generalized Momentum of a system is defined as  $\frac{\partial \mathcal{L}}{\partial \dot{q}}$ .

**Definition 41.1.2** The Generalized Force of a system is defined as  $\frac{\partial \mathcal{L}}{\partial q}$ .

When  $\mathcal{L}$  is the nice "Physics I," Lagrangian that we're familiar with, we see that Newton's Second Law appears. That is, the time derivative of momentum

is equal to minus the gradient of the potential energy. For any Lagrangian, if we use generalized momentum and generalized force, then the mathematics becomes very similar to Newton's Second Law. We obtain the "Generalized," Newton's Second Law:

$$\text{Generalized Force} = \frac{d}{dt}(\text{Generalized Momentum})$$

Let's consider the example  $q = \theta$ . That is, our generalized coordinate is the angle made between the particle and the  $x$  axis. The generalized momentum is just angular momentum, and the generalized force is angular force (Also known as torque). Let's consider the following system:

$$\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\ddot{\theta}$$

This is particle going around in a circle. Then we have from the equation of motion that:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}}\right) = mr^2\ddot{\theta} \Rightarrow \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\dot{\theta} + c \Rightarrow \mathcal{L} = \frac{1}{2}mr^2\dot{\theta} + g(\theta) + c\dot{\theta}$$

Remember from calculus that since we are taking partial derivatives, when we perform integration we get a function of integration, and not a just constant of integration. This function of integration is the  $\tilde{g}(\theta)$  we have in the equation above. Let's further add the requirement that  $\ddot{\theta} = 0$ . That is, there is no angular acceleration. This represent a particle going around in a circle at constant angular velocity. With this information, we can determine  $g(\theta)$ . We have that

$$\frac{\partial \mathcal{L}}{\partial \theta} = 0 \Rightarrow g'(\theta) = 0$$

which means  $g(\theta)$  is a constant. The Lagrangian of this problem is  $\mathcal{L} = \frac{1}{2}mr^2\dot{\theta}^2 + c_1\dot{\theta} + c_2$ . Now what of the generalized momentum? We have that:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}}\right) = \frac{\partial \mathcal{L}}{\partial \theta} = 0 \Rightarrow \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = \text{constant}$$

But our generalized momentum is simply the angular momentum. That is, we have shown that angular momentum is conserved.

**Part XXVII**

# **Diffraction Theory**



# CHAPTER 42

---

## Diffraction Theory

---

### 42.1 Maxwell's Equations

At the heart of electromagnetism are Maxwell's equations. They are:

$$\text{curl}(\mathbf{E}) = -\frac{\partial \mathbf{B}}{\partial t} \quad (42.1.1a) \quad \text{curl}(\mathbf{B}) = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad (42.1.1c)$$

$$\text{div}(\mathbf{E}) = \frac{\rho}{\epsilon_0} \quad (42.1.1b) \quad \text{div}(\mathbf{B}) = 0 \quad (42.1.1d)$$

With this, we will derive the Fresnel-Huygens principle.

### 42.2 Fresnel-Fraunhofer Theory

### 42.3 Fresnel's Approximation

### 42.4 Fresnel Inversion

#### 42.4.1 Diffraction Through a Square Well

We wish to solve for  $\hat{T}(\rho_0)$ . We have that:

$$\hat{T}(\rho_0) = \frac{1-i}{2F} \int_{-\infty}^{\infty} T(\rho) \exp\left(i\frac{\pi}{2}\left(\frac{\rho-\rho_0}{F}\right)^2\right) d\rho \quad (42.4.1)$$

For the square well of height  $M$  starting at  $a$  and ending at  $b$ :

$$T(\rho) = \begin{cases} 0, & \rho < a \\ M, & a \leq \rho \leq b \\ 0, & b < \rho \end{cases} \quad (42.4.2)$$

Let  $H_M(\rho_0, F; a, b)$  denote the solution. So, we have:

$$H_M(\rho_0, F; a, b) = \frac{1-i}{2F} \int_a^b M \exp\left(\frac{i\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho \quad (42.4.3a)$$

$$= \frac{M}{F} \frac{1-i}{2} \int_a^b \exp\left(\frac{i\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho \quad (42.4.3b)$$

Now from Euler's formula,  $\exp(i\theta) = \cos(\theta) + i \sin(\theta)$ . Using this we obtain:

$$\exp\left(\frac{i\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) = \cos\left(\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) + i \sin\left(\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) \quad (42.4.4)$$

Recall the Fresnel Cosine Integral and Sine Integrals  $(C(x), S(x))$  are defined as:

$$S(x) = \int_0^x \sin(t^2) dt \quad (42.4.5a) \qquad C(x) = \int_0^x \cos(t^2) dt \quad (42.4.5b)$$

Letting  $u = \sqrt{\frac{\pi}{2}}x$ , we obtain:

$$\int_a^b \sin\left(\frac{\pi}{2}x^2\right) dx = \int_0^b \sin\left(\frac{\pi}{2}x^2\right) dx - \int_0^a \sin\left(\frac{\pi}{2}x^2\right) dx \quad (42.4.6a)$$

$$= \sqrt{\frac{2}{\pi}} \left[ \int_0^{\sqrt{\frac{\pi}{2}}b} \sin(u^2) du - \int_0^{\sqrt{\frac{\pi}{2}}a} \sin(u^2) du \right] \quad (42.4.6b)$$

$$= \sqrt{\frac{2}{\pi}} \left[ S\left(\sqrt{\frac{\pi}{2}}b\right) - S\left(\sqrt{\frac{\pi}{2}}a\right) \right] \quad (42.4.6c)$$

$$\int_a^b \cos\left(\frac{\pi}{2}x^2\right) dx = \int_0^b \cos\left(\frac{\pi}{2}x^2\right) dx - \int_0^a \cos\left(\frac{\pi}{2}x^2\right) dx \quad (42.4.6d)$$

$$= C(b) - C(a) \quad (42.4.6e)$$

We can now compute  $H_M(\rho_0, F; a, b)$ . Let  $x = \frac{\rho - \rho_0}{F}$ . Then  $dx = \frac{d\rho}{F}$ . We have:

$$\begin{aligned} H_M(\rho_0, F; a, b) &= \frac{M}{F} \frac{1-i}{2} \int_a^b \exp\left(\frac{i\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho \\ &= \frac{M}{F} \frac{1-i}{2} \left[ \int_a^b \cos\left(\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho + i \int_a^b \sin\left(\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho \right] \\ &= M \frac{1-i}{2} \left[ \int_{\frac{a-\rho_0}{F}}^{\frac{b-\rho_0}{F}} \cos\left(\frac{\pi}{2}x^2\right) dx + i \int_{\frac{a-\rho_0}{F}}^{\frac{b-\rho_0}{F}} \sin\left(\frac{\pi}{2}x^2\right) dx \right] \\ &= M \frac{1-i}{2} \left[ \left( C\left(\frac{b-\rho_0}{F}\right) - C\left(\frac{a-\rho_0}{F}\right) \right) + i \left( S\left(\frac{b-\rho_0}{F}\right) - S\left(\frac{a-\rho_0}{F}\right) \right) \right] \end{aligned}$$

#### 42.4.2 Diffraction Through an Inverted Square Well

This time we have:

$$T(\rho) = \begin{cases} M, & \rho < a \\ 0, & a \leq \rho \leq b \\ M, & b < \rho \end{cases}$$

So  $T(\rho) = M - T_{Sq}(\rho)$ , where  $T_{Sq}(\rho)$  is the impulse from the standard square well (See 42.4.1). So we wish to solve:

$$\hat{T}(\rho_0) = \frac{1-i}{2F} \int_{-\infty}^{\infty} (M - T_{Sq}(\rho)) \exp\left(i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho$$

Simplifying, we have:

$$\begin{aligned} \hat{T}(\rho_0) &= \frac{1-i}{2F} \int_{-\infty}^{\infty} (M - T(\rho)) \exp\left(i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho \\ &= \frac{M}{F} \frac{1-i}{2} \left( \int_{-\infty}^{\infty} \exp\left(i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho - \int_a^b \exp\left(i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho \right) \end{aligned}$$

But from the previous derivation:

$$\int_a^b \exp\left(i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho = F \left( C\left(\frac{b - \rho_0}{F}\right) - C\left(\frac{a - \rho_0}{F}\right) \right) + i \left( S\left(\frac{b - \rho_0}{F}\right) - S\left(\frac{a - \rho_0}{F}\right) \right)$$

And:

$$\int_{-\infty}^{\infty} \exp\left(i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho = \lim_{a \rightarrow \infty} \int_{-a}^a \exp\left(i\frac{\pi}{2}\left(\frac{\rho - \rho_0}{F}\right)^2\right) d\rho = F(1+i)$$

Piecing this together, and using the notation from before, we have:

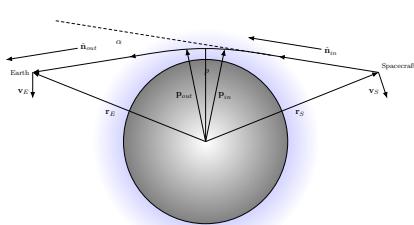
$$\hat{T}(\rho_0) = M - H_M(\rho_0, F; a, b)$$



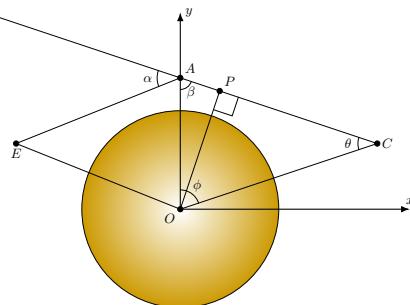
# CHAPTER 43

## Geometry

### 43.1 Titan Geometry



43.1.1: Geometry of an Occultation of Titan



43.1.2: Geometry of the Bending Angle

Fig. 43.1: Various Geometries for Titan

The following definitions are used:

1.  $O$  is the center of Titan.
2.  $E$  is the Earth.
3.  $C$  is the Cassini spacecraft.
4.  $\mathbf{r}_E = \overrightarrow{OE}$
5.  $\mathbf{v}_E = \dot{\mathbf{r}}_E$
6.  $\mathbf{r}_S = \overrightarrow{OC}$

7.  $\mathbf{v}_S = \dot{\mathbf{r}}_S$
8.  $\mathbf{p}_{in}$  is the projection of  $O$  onto  $\overline{AC}$
9.  $\mathbf{p}_{out}$  is the projection of  $O$  onto  $\overline{EA}$
10.  $\alpha$  is the bending angle ( $\pi - \angle EAC$ )
11.  $\hat{\mathbf{n}}_{in}$  is the direction of the emission.
12.  $\hat{\mathbf{n}}_{out}$  is the direction of the reception.
13. The ray plane lies in the plane  $OEC$
14.  $\phi = \angle AOC$
15.  $\theta = \angle ACO$
16.  $\beta = \angle OAC$
17.  $A$  is the intersection of the lines starting at  $C$  and  $E$ , parallel to  $\hat{\mathbf{n}}_{in}$  and  $\hat{\mathbf{n}}_{out}$ , respectively.

Where  $\dot{\mathbf{r}}$  denotes the time derivative of  $\mathbf{r}$ . The following assumptions are made:

1.  $\angle OAE = \angle OAC$
2.  $A$  lies in the plane  $OEC$

**Theorem 43.1.1.** *The ray plane is perpendicular to  $\hat{\mathbf{z}} = \frac{\mathbf{r}_S \times \mathbf{r}_E}{\|\mathbf{r}_S \times \mathbf{r}_E\|}$*

*Proof.* As the ray plane is the plane  $OEC$ ,  $\mathbf{r}_S$  and  $\mathbf{r}_E$  lie parallel to this plane. Moreover, during an occultation,  $\mathbf{r}_S$  and  $\mathbf{r}_E$  are not parallel and therefore  $OEC$  is uniquely determined by  $\mathbf{r}_E$ ,  $\mathbf{r}_S$ , and the point  $O$ . But  $\hat{\mathbf{z}} = \frac{\mathbf{r}_S \times \mathbf{r}_E}{\|\mathbf{r}_S \times \mathbf{r}_E\|}$  is perpendicular to both  $\mathbf{r}_E$  and  $\mathbf{r}_S$ . Therefore  $\hat{\mathbf{z}}$  is perpendicular to the ray plane.  $\square$

**Theorem 43.1.2.**  $\mathbf{r}_E \cdot \mathbf{p}_{out} = \|\mathbf{p}_{out}\|^2$

*Proof.*  $\mathbf{p}_{out}$  is the projection of the  $O$  onto  $\overline{EA}$ . But  $\overline{EA}$  lies parallel to  $\hat{\mathbf{n}}_{out}$ , and therefore  $\mathbf{p}_{out}$  and  $\hat{\mathbf{n}}_{out}$  are orthogonal, and thus  $\mathbf{p}_{out} \cdot \hat{\mathbf{n}}_{out} = 0$ . Moreover,  $\mathbf{r}_E = \mathbf{p}_{out} + (\mathbf{r}_E \cdot \hat{\mathbf{n}}_{out})\hat{\mathbf{n}}_{out}$ . But then:

$$\begin{aligned} \mathbf{p}_{out} \cdot \mathbf{r}_E &= \mathbf{p}_{out} \cdot (\mathbf{p}_{out} + (\mathbf{r}_E \cdot \hat{\mathbf{n}}_{out})\hat{\mathbf{n}}_{out}) \\ &\Rightarrow \mathbf{p}_{out} \cdot \mathbf{r}_E = \mathbf{p}_{out} \cdot \mathbf{p}_{out} + (\mathbf{r}_E \cdot \hat{\mathbf{n}}_{out})\mathbf{p}_{out} \cdot \hat{\mathbf{n}}_{out} \\ &\Rightarrow \mathbf{p}_{out} \cdot \mathbf{r}_E = \mathbf{p}_{out} \cdot \mathbf{p}_{out} \end{aligned}$$

Therefore  $\mathbf{p}_{out} \cdot \mathbf{r}_E = \|\mathbf{p}_{out}\|^2$   $\square$

**Theorem 43.1.3.**  $\alpha = \cos^{-1}(\hat{\mathbf{n}}_{in} \cdot \hat{\mathbf{n}}_{out})$

*Proof.* By definition,  $\alpha = \pi - \angle EAC$ . But  $\hat{\mathbf{n}}_{out}$  lies parallel to  $\overrightarrow{AE}$ , and  $-\hat{\mathbf{n}}_{in}$  lies parallel to  $\overrightarrow{AC}$ . Therefore:

$$-\hat{\mathbf{n}}_{out} \cdot \hat{\mathbf{n}}_{in} = \hat{\mathbf{n}}_{out} \cdot (-\hat{\mathbf{n}}_{in}) = \|\hat{\mathbf{n}}_{out}\| \|\hat{\mathbf{n}}_{in}\| \cos(\angle EAC)$$

But  $\hat{\mathbf{n}}_{in}$  and  $\hat{\mathbf{n}}_{out}$  are unit vectors, and therefore  $\|\hat{\mathbf{n}}_{out}\| = \|-\hat{\mathbf{n}}_{in}\| = 1$ . Therefore:

$$\angle EAC = \cos^{-1}(-\hat{\mathbf{n}}_{out} \cdot \hat{\mathbf{n}}_{in})$$

But  $\alpha = \pi - \angle EAC$ , and  $\cos^{-1}(-x) = \pi - \cos^{-1}(x)$ . Therefore:

$$\alpha = \pi - \angle EAC = \pi - (\pi - \cos^{-1}(\hat{\mathbf{n}}_{out} \cdot \hat{\mathbf{n}}_{in})) = \cos^{-1}(\hat{\mathbf{n}}_{out} \cdot \hat{\mathbf{n}}_{in})$$

□

**Theorem 43.1.4.**  $\theta = \cos^{-1}\left(\frac{(-\mathbf{r}_S) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_S\|}\right)$

*Proof.* For  $\theta = \angle OCA$ . But  $\hat{\mathbf{n}}_{in}$  is parallel with  $\overrightarrow{CA}$ , and  $(-\mathbf{r}_S)$  is parallel with  $\overrightarrow{CO}$ . Therefore:

$$\begin{aligned} (-\mathbf{r}_S) \cdot \hat{\mathbf{n}}_{in} &= \|(-\mathbf{r}_S)\| \|\hat{\mathbf{n}}_{in}\| \cos(\theta) \\ \Rightarrow \theta &= \cos^{-1}\left(\frac{(-\mathbf{r}_S) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_S\|}\right) \end{aligned}$$

□

**Theorem 43.1.5.**

$$\beta = \pi - \frac{1}{2} \cos^{-1}\left(\frac{\mathbf{r}_s \cdot \mathbf{r}_E}{\|\mathbf{r}_s\| \|\mathbf{r}_E\|}\right) - \frac{1}{2} \cos^{-1}\left(\frac{\mathbf{r}_E \cdot \hat{\mathbf{n}}_{out}}{\|\mathbf{r}_E\|}\right) - \frac{1}{2} \cos^{-1}\left(\frac{(-\mathbf{r}_s) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_s\|}\right)$$

*Proof.* The sum of the angles in  $OEAC$  is  $2\pi$ . But  $\angle OAE = \angle OAC = \phi$ , and therefore:

$$\begin{aligned} 2\beta &= \angle EAC \\ \Rightarrow 2\pi &= 2\beta + \angle AEO + \angle EOC + \angle OCA \\ \Rightarrow \beta &= \pi - \frac{\angle AEO}{2} - \frac{\angle EOC}{2} - \frac{\angle OCA}{2} \end{aligned}$$

But:

$$\begin{aligned} (-\hat{\mathbf{n}}_{out}) \cdot (-\hat{\mathbf{r}}_E) &= \|\mathbf{r}_E\| \cos(\angle AEO) \\ \Rightarrow \angle AEO &= \cos^{-1}\left(\frac{\hat{\mathbf{n}}_{out} \cdot \mathbf{r}_E}{\|\mathbf{r}_E\|}\right) \end{aligned}$$

Also:

$$\begin{aligned}\mathbf{r}_E \cdot \mathbf{r}_S &= \|\mathbf{r}_E\| \|\mathbf{r}_S\| \cos(\angle EOC) \\ \Rightarrow \angle EOC &= \cos^{-1} \left( \frac{\mathbf{r}_E \cdot \mathbf{r}_S}{\|\mathbf{r}_E\| \|\mathbf{r}_S\|} \right)\end{aligned}$$

But  $\angle OCA = \theta = \cos^{-1} \left( \frac{(-\mathbf{r}_s) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_s\|} \right)$ . Therefore:

$$\beta = \pi - \frac{1}{2} \cos^{-1} \left( \frac{\mathbf{r}_s \cdot \mathbf{r}_E}{\|\mathbf{r}_s\| \|\mathbf{r}_E\|} \right) - \frac{1}{2} \cos^{-1} \left( \frac{\mathbf{r}_E \cdot \hat{\mathbf{n}}_{out}}{\|\mathbf{r}_E\|} \right) - \frac{1}{2} \cos^{-1} \left( \frac{(-\mathbf{r}_s) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_s\|} \right)$$

□

**Theorem 43.1.6.**  $\alpha = \pi - 2\beta$

*Proof.*  $\alpha$  and  $\angle EAC$  are supplementary to the ray  $\overrightarrow{CA}$ , and therefore  $\alpha + \angle EAC = \pi$ . But  $\angle EAC = \angle EAC + \angle OAC = 2\beta$ . Therefore  $\alpha + 2\beta = \pi$ . Thus,  $\alpha = \pi - 2\beta$ . □

**Theorem 43.1.7.**  $\theta = \frac{\pi}{2} + \frac{\alpha}{2} - \phi$

*Proof.* As the angles of a triangle sum to  $\pi$ ,  $\theta + \beta + \phi = \pi$ . But  $\alpha = \pi - 2\beta \Rightarrow \beta = \frac{\pi}{2} - \frac{\alpha}{2}$ . So we have:

$$\begin{aligned}\theta + \phi + \beta &= \pi \\ \Rightarrow \theta + \phi + \frac{\pi}{2} - \frac{\alpha}{2} &= \pi \\ \Rightarrow \theta &= \frac{\pi}{2} + \frac{\alpha}{2} - \phi\end{aligned}$$

□

**Theorem 43.1.8.**

$$\phi = \frac{1}{2} \cos^{-1} \left( \frac{\mathbf{r}_s \cdot \mathbf{r}_E}{\|\mathbf{r}_s\| \|\mathbf{r}_E\|} \right) + \frac{1}{2} \cos^{-1} \left( \frac{\mathbf{r}_E \cdot \hat{\mathbf{n}}_{out}}{\|\mathbf{r}_E\|} \right) - \frac{1}{2} \cos^{-1} \left( \frac{(-\mathbf{r}_s) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_s\|} \right)$$

*Proof.* For:

$$\pi = \beta + \theta + \phi$$

$$\theta = \cos^{-1} \left( \frac{(-\mathbf{r}_s) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_s\|} \right)$$

$$\beta = \pi - \frac{1}{2} \cos^{-1} \left( \frac{\mathbf{r}_s \cdot \mathbf{r}_E}{\|\mathbf{r}_s\| \|\mathbf{r}_E\|} \right) - \frac{1}{2} \cos^{-1} \left( \frac{\mathbf{r}_E \cdot \hat{\mathbf{n}}_{out}}{\|\mathbf{r}_E\|} \right) - \frac{1}{2} \cos^{-1} \left( \frac{(-\mathbf{r}_s) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_s\|} \right)$$

$$\Rightarrow \phi = \frac{1}{2} \cos^{-1} \left( \frac{\mathbf{r}_s \cdot \mathbf{r}_E}{\|\mathbf{r}_s\| \|\mathbf{r}_E\|} \right) + \frac{1}{2} \cos^{-1} \left( \frac{\mathbf{r}_E \cdot \hat{\mathbf{n}}_{out}}{\|\mathbf{r}_E\|} \right) - \frac{1}{2} \cos^{-1} \left( \frac{(-\mathbf{r}_s) \cdot \hat{\mathbf{n}}_{in}}{\|\mathbf{r}_s\|} \right)$$

□

**Theorem 43.1.9.**  $p = \|\mathbf{p}_{in}\| = \|\mathbf{r}_S\| \cos(\phi - \frac{\alpha}{2})$

*Proof.* As  $P$  is the orthogonal projection of  $O$  onto  $\overline{CA}$ ,  $\angle OPC = \frac{\pi}{2}$ . But then:

$$|\overline{OP}| = |\overline{OC}| \sin(\angle OCP)$$

But  $|\overline{OP}| = \|\mathbf{p}_{in}\|$ ,  $|\overline{OC}| = \|\mathbf{r}_S\|$ , and  $\angle OCP = \theta$ . Therefore:

$$\|\mathbf{p}_{in}\| = \|\mathbf{r}_S\| \sin(\theta)$$

But  $\theta = \frac{\pi}{2} + \frac{\alpha}{2} - \phi$ , and  $\sin(\frac{\pi}{2} + x) = \cos(x)$ . Therefore:

$$\|\mathbf{p}_{in}\| = \|\mathbf{r}_S\| \cos\left(\frac{\alpha}{2} - \phi\right)$$

□

**Theorem 43.1.10.**  $\|\mathbf{p}_{in}\| = |\overline{OA}| \sin(\beta)$

*Proof.* For  $\overline{OP}$  is perpendicular to  $\overline{CA}$ , and therefore  $\Delta OPA$  is a right-angled triangle, and  $\overline{OA}$  is the hypotenuse. Moreoever  $\angle PAO = \beta$ . But then:

$$\begin{aligned} |\overline{OP}| &= |\overline{OA}| \sin(\angle PAO) \\ \Rightarrow |\overline{OP}| &= |\overline{OA}| \sin(\beta) \end{aligned}$$

But  $\|\mathbf{p}_{in}\| = |\overline{OP}|$ , and thus  $\|\mathbf{p}_{in}\| = |\overline{OA}| \sin(\beta)$

□

**Theorem 43.1.11.**  $\|\mathbf{p}_{in}\| = \|\mathbf{p}_{out}\|$

*Proof.* For  $\angle OAE = \angle OAC = \beta$ , and thus:

$$\|\mathbf{p}\|_{out} = |\overline{OA}| \sin(\angle OAE) = |\overline{OA}| \sin(\beta) = \|\mathbf{p}\|_{in}$$

□

## 43.2 Ring Geometry

**Theorem 43.2.1.** If  $\hat{\mathbf{u}}$  and  $\hat{\mathbf{z}}$  are unit vectors and  $\hat{\mathbf{u}} \times \hat{\mathbf{z}} \neq \mathbf{0}$ , then:

$$\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\} = \left\{ \left( \frac{\hat{\mathbf{u}} \times \hat{\mathbf{z}}}{\|\hat{\mathbf{u}} \times \hat{\mathbf{z}}\|} \right) \times \hat{\mathbf{z}}, \frac{\hat{\mathbf{u}} \times \hat{\mathbf{z}}}{\|\hat{\mathbf{u}} \times \hat{\mathbf{z}}\|}, \hat{\mathbf{z}} \right\} \quad (43.2.1)$$

is an orthonormal basis of  $\mathbb{R}^3$ .

*Proof.* Since  $\hat{\mathbf{u}} \times \hat{\mathbf{z}}$  is a non-zero vector,  $\|\mathbf{u} \times \mathbf{z}\| \neq 0$ . Thus, let  $\hat{\mathbf{y}} = \frac{\hat{\mathbf{u}} \times \hat{\mathbf{z}}}{\|\hat{\mathbf{u}} \times \hat{\mathbf{z}}\|}$  and let  $\hat{\mathbf{x}} = \hat{\mathbf{y}} \times \hat{\mathbf{z}}$ . Then  $\hat{\mathbf{y}} \cdot \hat{\mathbf{z}} = 0$ ,  $\hat{\mathbf{y}} \cdot \hat{\mathbf{x}} = 0$ , and  $\hat{\mathbf{x}} \cdot \hat{\mathbf{z}} = 0$ . Both  $\hat{\mathbf{z}}$  and  $\hat{\mathbf{y}}$  are unit vectors by definition, and  $\hat{\mathbf{x}}$  is the cross product of two orthogonal unit vectors, and is therefore itself a unit vector. But then  $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$  is a set of 3 mutually orthogonal unit vectors. By the Vector Space Dimension Theorem,  $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$  is an orthonormal basis of  $\mathbb{R}^3$ .

□

We define our Saturnian Coordinate System to be the Cartesian Coordinate System where  $\mathbf{u}$  is the vector from Earth to the Spacecraft,  $\hat{\mathbf{z}}$  is Saturn's Pole vector, and let  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{y}}$  be as defined in Eqn. 43.2.1. The origin is taken to be Saturn's Center. The ring plane of Saturn is the plane perpendicular to  $\hat{\mathbf{z}}$  which contains the origin.

**Theorem 43.2.2.** *Saturn's ring plane is the  $xy$  plane.*

*Proof.* This is a restatement of the fact that  $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$  is an orthonormal system (Thm. 43.2.1) and from the definition of Saturn's ring plane.  $\square$

**Theorem 43.2.3.** *The Earth-Spacecraft line,  $\mathbf{u}$ , lies parallel to the  $xz$  plane.*

*Proof.* It suffices to show that  $\hat{\mathbf{u}}$  is orthogonal to  $\hat{\mathbf{y}}$ . But:

$$\hat{\mathbf{u}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{u}} \cdot \frac{\hat{\mathbf{u}} \times \hat{\mathbf{z}}}{\|\hat{\mathbf{u}} \times \hat{\mathbf{z}}\|} \quad (43.2.2)$$

And for any two vectors  $\mathbf{a}$  and  $\mathbf{b}$ ,  $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$  and therefore  $\hat{\mathbf{u}}$  is orthogonal to  $\hat{\mathbf{y}}$ . Thus  $\hat{\mathbf{u}}$  is parallel to the  $xz$  plane.  $\square$

**Theorem 43.2.4.** *In the Saturn Reference frame, Earth lies on the  $xz$  plane if and only if the line from Earth to Cassini lies in it.*

*Proof.* If  $\hat{\mathbf{u}}$  lies in the  $xz$  plane, then Earth must also lie in it from the definition of  $\mathbf{u}$ . And from Thm. 43.2.4  $\hat{\mathbf{u}}$  lies parallel to the  $xz$  plane. Thus, if Earth lies in the  $xz$  plane, so must the line from Earth to Cassini.  $\square$

**Theorem 43.2.5.** *If  $\hat{\mathbf{z}}$  and  $\hat{\mathbf{u}}$  are defined by:*

$$\hat{\mathbf{z}} = z_1 \hat{\mathbf{x}}_E + z_2 \hat{\mathbf{y}}_E + z_3 \hat{\mathbf{z}}_E \quad (43.2.3a)$$

$$\hat{\mathbf{u}} = u_{E_x} \hat{\mathbf{x}}_E + u_{E_y} \hat{\mathbf{y}}_E + u_{E_z} \hat{\mathbf{z}}_E \quad (43.2.3b)$$

then:

$$\hat{\mathbf{y}} = y_1 \hat{\mathbf{x}}_E + y_2 \hat{\mathbf{y}}_E + y_3 \hat{\mathbf{z}}_E \quad (43.2.3c)$$

Where:

$$y_1 = \frac{z_2 u_{E_z} - z_3 u_{E_y}}{\sqrt{(z_2 u_{E_z} - z_3 u_{E_y})^2 + (z_3 u_{E_x} - z_1 u_{E_z})^2 + (z_1 u_{E_y} - z_2 u_{E_x})^2}} \quad (43.2.3d)$$

$$y_2 = \frac{z_3 u_{E_x} - z_1 u_{E_z}}{\sqrt{(z_2 u_{E_z} - z_3 u_{E_y})^2 + (z_3 u_{E_x} - z_1 u_{E_z})^2 + (z_1 u_{E_y} - z_2 u_{E_x})^2}} \quad (43.2.3e)$$

$$y_3 = \frac{z_1 u_{E_y} - z_2 u_{E_x}}{\sqrt{(z_2 u_{E_z} - z_3 u_{E_y})^2 + (z_3 u_{E_x} - z_1 u_{E_z})^2 + (z_1 u_{E_y} - z_2 u_{E_x})^2}} \quad (43.2.3f)$$

*Proof.* From the definition given in Thm. 43.2.1,  $\hat{\mathbf{y}}$  is defined as  $\frac{\hat{\mathbf{z}} \times \mathbf{u}_0}{\|\hat{\mathbf{z}} \times \mathbf{u}_0\|}$ . This equation is the cross-product divided by the norm.  $\square$

**Theorem 43.2.6.** If  $\hat{\mathbf{z}}$  and  $\mathbf{u}_0$  are as defined in Eqn. 43.2.3a and 43.2.3b, then:

$$\hat{\mathbf{x}} = x_1 \hat{\mathbf{x}}_E + x_2 \hat{\mathbf{y}}_E + x_3 \hat{\mathbf{z}}_E \quad (43.2.4a)$$

Where:

$$x_1 = \frac{z_3(z_3 u_{E_x} - z_1 u_{E_z}) - z_2(z_1 u_{E_y} - z_2 u_{E_x})}{\sqrt{(z_2 u_{E_z} - z_3 u_{E_y})^2 + (z_3 u_{E_x} - z_1 u_{E_z})^2 + (z_1 u_{E_y} - z_2 u_{E_x})^2}} \quad (43.2.4b)$$

$$x_2 = \frac{z_1(z_1 u_{E_y} - z_2 u_{E_x}) - z_3(z_2 u_{E_z} - z_3 u_{E_y})}{\sqrt{(z_2 u_{E_z} - z_3 u_{E_y})^2 + (z_3 u_{E_x} - z_1 u_{E_z})^2 + (z_1 u_{E_y} - z_2 u_{E_x})^2}} \quad (43.2.4c)$$

$$x_3 = \frac{z_2(z_2 u_{E_z} - z_3 u_{E_y}) - z_1(z_3 u_{E_x} - z_1 u_{E_z})}{\sqrt{(z_2 u_{E_z} - z_3 u_{E_y})^2 + (z_3 u_{E_x} - z_1 u_{E_z})^2 + (z_1 u_{E_y} - z_2 u_{E_x})^2}} \quad (43.2.4d)$$

*Proof.*  $\hat{\mathbf{x}}$  is defined as  $\hat{\mathbf{y}} \times \hat{\mathbf{z}}$ . This equation is merely that product.  $\square$

Thus if we have  $\hat{\mathbf{u}}$  and  $\hat{\mathbf{z}}$  in an Earth based system, we can easily compute the geometry in our Saturnian coordinate system. At the very least, a computer can easily compute this.

**Theorem 43.2.7.** If  $(S_x, S_y, S_z)$  is location of the center of Saturn with respect to the center of the Earth and  $(x_E, y_E, z_E)$  is a point in  $\mathbb{R}^3$  with respect to the center of the Earth, then the change of coordinates to the Saturn-based system is:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} \begin{pmatrix} x_E - S_x \\ y_E - S_y \\ z_E - S_z \end{pmatrix}$$

*Proof.* The point  $(x_E - S_x, y_E - S_y, z_E - S_z)$  translates the point  $(x_E, y_E, z_E)$  to the center of Saturn. The rotation matrix then aligns the Earth-based coordinates to the Saturn-based coordinates.  $\square$

### 43.3 Derivations of the Fresnel Kernel

Let  $\hat{\mathbf{u}}$  be the unit vector pointing from Earth to the spacecraft. Let  $\hat{\mathbf{z}}$  be the pole direction Saturn. To make the arguments easier, we assume the line from Earth to Saturn and the line from Earth to Voyager are parallel. That is, we assume that Saturn is infinitely far away. Define the following:

$$B = \sin^{-1}(\hat{\mathbf{z}} \cdot \hat{\mathbf{u}}) \quad (43.3.1)$$

$$\hat{\mathbf{y}} = \frac{\hat{\mathbf{u}} \times \hat{\mathbf{z}}}{\|\hat{\mathbf{u}} \times \hat{\mathbf{z}}\|} \quad (43.3.2a) \qquad \hat{\mathbf{x}} = \hat{\mathbf{y}} \times \hat{\mathbf{z}} \quad (43.3.2b)$$

We take the origin as Saturn's center. From Thm. 43.2.3,  $\hat{\mathbf{u}}$  lies parallel to the  $xz$  plane, and thus there are numbers  $a_1, a_2$  such that:

$$\hat{\mathbf{u}} = a_1 \hat{\mathbf{x}} + a_2 \hat{\mathbf{z}} \quad (43.3.3)$$

We can compute for  $a_1$  and  $a_2$  by using the definition of  $B$  in Eqn. 43.3.1.

$$\hat{\mathbf{u}} = \cos(B) \hat{\mathbf{x}} + \sin(B) \hat{\mathbf{z}} \quad (43.3.4)$$

Let  $\rho_0$  be the vector pointing from Saturn to the ring intercept point, and let  $\rho$  be a vector in the ring plane. Let  $\phi_0$  and  $\phi$  be the angles made with  $\rho_0$  and  $\rho$  to the  $x$  axis, respectively. Then:

$$\rho_0 = \rho_0 (\cos(\phi_0) \hat{\mathbf{x}} + \sin(\phi_0) \hat{\mathbf{y}}) \quad (43.3.5a)$$

$$\rho = \rho (\cos(\phi) \hat{\mathbf{x}} + \sin(\phi) \hat{\mathbf{y}}) \quad (43.3.5b)$$

Let  $\mathbf{R}_c$  be the vector pointing from Saturn to Voyager. Let  $D$  be the distance from the ring intercept point to Voyager. We thus have the following:

$$\mathbf{R}_c = \rho_0 + D \hat{\mathbf{u}} \quad (43.3.6a)$$

$$= (\rho_0 \cos(\phi_0) + D \cos(B)) \hat{\mathbf{x}} + \rho_0 \sin(\phi_0) \hat{\mathbf{y}} + D \sin(B) \hat{\mathbf{z}} \quad (43.3.6b)$$

We wish to compute  $\hat{\mathbf{u}} \cdot \rho + \|\mathbf{R}_c - \rho\|$ . We have:

$$\hat{\mathbf{u}} \cdot \rho = (\cos(B) \hat{\mathbf{x}} + \sin(B) \hat{\mathbf{z}}) \cdot (\rho (\cos(\phi) \hat{\mathbf{x}} + \sin(\phi) \hat{\mathbf{y}})) \quad (43.3.7a)$$

$$= \rho \cos(B) \cos(\phi) \quad (43.3.7b)$$

And also:

$$\|\mathbf{R}_c - \rho\| = \sqrt{(\mathbf{R}_c - \rho) \cdot (\mathbf{R}_c - \rho)} \quad (43.3.8a)$$

$$= \sqrt{\|\mathbf{R}_c\|^2 + \|\rho\|^2 - 2\mathbf{R}_c \cdot \rho} \quad (43.3.8b)$$

But, since  $\hat{\mathbf{u}}$  is a unit vector and  $\rho_0 \cdot \rho_0 = \rho_0^2$ , we have:

$$\|\mathbf{R}_c\|^2 = \rho_0^2 + D^2 + 2\rho_0 D \rho_0 \cdot \hat{\mathbf{u}} \quad (43.3.9a)$$

$$= \rho_0^2 + D^2 + 2D\rho_0 \cos(\phi_0) \cos(B) \quad (43.3.9b)$$

Furthering the computation we have:

$$\mathbf{R}_c \cdot \rho = \rho \cos(\phi) (\rho_0 \cos(\phi_0) + D \cos(B)) + \rho \rho_0 \sin(\phi) \sin(\phi_0) \quad (43.3.10a)$$

$$= \rho \rho_0 (\cos(\phi) \cos(\phi_0) + \sin(\phi) \sin(\phi_0)) + \rho D \cos(\phi) \cos(B) \quad (43.3.10b)$$

$$= \rho \rho_0 \cos(\phi - \phi_0) + \rho D \cos(\phi) \cos(B) \quad (43.3.10c)$$

So we have:

$$\begin{aligned} \|\mathbf{R}_c - \rho\|^2 &= \rho^2 + \rho_0^2 + D^2 - 2\rho \rho_0 \cos(\phi - \phi_0) \\ &\quad + 2D \cos(B) (\rho_0 \cos(\phi_0) - \rho \cos(\phi)) \end{aligned} \quad (43.3.11)$$

Now the definition of  $\hat{T}$  is:

$$\hat{T} = \frac{E_c}{E_0} e^{-ik\hat{\mathbf{u}} \cdot \mathbf{R}_c} \quad (43.3.12)$$

So we define  $\psi$  as:

$$\psi = k(\|\mathbf{R}_c - \rho\|^2 + \hat{\mathbf{u}} \cdot \boldsymbol{\rho} - \hat{\mathbf{u}} \cdot \mathbf{R}_c) \quad (43.3.13)$$

Trudging along, we have:

$$\hat{\mathbf{u}} \cdot \mathbf{R}_c = \rho_0 \cos(\phi_0) \cos(B) + D \cos^2(B) + D \sin^2(B) \quad (43.3.14a)$$

$$= \rho_0 \cos(\phi_0) \cos(B) + D \quad (43.3.14b)$$

Let's define the following:

$$\xi = \frac{\cos(B)(\rho \cos(\phi) - \rho_0 \cos(\phi_0))}{D} \quad (43.3.15)$$

$$\eta = \frac{\rho_0^2 + \rho^2 - 2\rho\rho_0 \cos(\phi - \phi_0)}{D^2} \quad (43.3.16)$$

Please note that  $\xi$  differs in sign from the definition found in MTR86. This is done intentionally in order for problem to lend itself more naturally to the use of Legendre polynomials. Using this, we finally obtain:

$$\psi = kD[\sqrt{1 + \eta - 2\xi} - (1 - \xi)] \quad (43.3.17)$$

An important configuration to consider is when  $\phi = \phi_0$ . Evaluating  $\psi$ , we obtain:

$$\begin{aligned} \psi_{\phi=\phi_0} &= kD \left[ \cos(\phi_0) \cos(B) \left( \frac{\rho - \rho_0}{D} \right) - 1 \right. \\ &\quad \left. + \sqrt{1 + \left( \frac{\rho - \rho_0}{D} \right)^2 - 2 \cos(\phi_0) \cos(B) \left( \frac{\rho - \rho_0}{D} \right)} \right] \end{aligned} \quad (43.3.18)$$

Define the following:

$$x = \frac{\rho - \rho_0}{D} \quad (43.3.19a) \quad \alpha = \cos(\phi_0) \cos(B) \quad (43.3.19b)$$

Then we may rewrite  $\psi_{\phi=\phi_0}$  as:

$$\psi_{\phi=\phi_0} = kD \left[ \alpha x - 1 + \sqrt{1 + x^2 - 2x} \right] \quad (43.3.20)$$

This is where the Legendre polynomials come into play. Letting  $P_n(\alpha)$  denote the  $n^{th}$  Legendre polynomial, the generating function is:

$$\sum_{n=0}^{\infty} P_n(\alpha)x^n = \frac{1}{\sqrt{1 + x^2 - 2\alpha x}} \quad (43.3.21)$$

We don't quite have this, but we can produce a differential equation that will lead us to this. First note the following:

$$\frac{d}{dx} \left( \sqrt{1 + x^2 - 2\alpha x} \right) = \frac{x - \alpha}{\sqrt{1 + x^2 - 2\alpha x}} \quad (43.3.22)$$

But from Eqn. 43.3.21, we have:

$$\frac{d}{dx} \left( \sqrt{1 + x^2 - 2\alpha x} \right) = (x - \alpha) \sum_{n=0}^{\infty} P_n(\alpha) x^n \quad (43.3.23)$$

Integrating across, we then obtain:

$$\sqrt{1 + x^2 - 2\alpha x} = 1 + \sum_{k=0}^{\infty} P_k(\alpha) \frac{x^{k+2}}{k+2} - \alpha \sum_{k=0}^{\infty} P_k(\alpha) \frac{x^{k+1}}{k+1} \quad (43.3.24)$$

And so finally:

$$\sqrt{1 + x^2 - 2\alpha x} + \alpha x - 1 = \sum_{k=0}^{\infty} \left( P_k(\alpha) - \alpha P_{k+1}(\alpha) \right) \frac{x^{k+2}}{k+2} \quad (43.3.25)$$

We can use this to evaluate and approximate  $\psi$ . Define the following sequence:

$$b_k = \frac{P_k(\alpha) - \alpha P_{k+1}(\alpha)}{k+2} \quad (43.3.26)$$

Then  $\psi_{\phi=\phi_0}$  may be expressed as follows:

$$\psi_{\phi=\phi_0} = kD \sum_{k=0}^{\infty} b_k x^{k+2} \quad (43.3.27)$$

# CHAPTER 44

---

## Occultation Observations

---

### 44.1 Diffraction Theory for Occultations

#### 44.1.1 Reduction to a Single Integral

We have modelled our problem from the Fresnel-Huygens principle. Given a plane wave of wavelength  $\lambda$  travelling in the  $\hat{\mathbf{u}}_i$  direction incident on a thin plane screen, with plane angle  $B$ , the complex transmittance at the point  $\mathbf{p}_0 = (\rho_0, \phi_0)$  measured by an observer at the point  $\mathbf{R}_c$  can be modelled by the following equation:

$$\hat{T}(\mathbf{p}_0) = \frac{\mu_0}{i\lambda} \int_0^{2\pi} \int_0^\infty T(\rho) \exp(ik\hat{\mathbf{u}}_i \cdot (\mathbf{p} - \mathbf{R}_c)) \frac{\exp(ik\|\mathbf{R}_c - \mathbf{p}\|)}{\|\mathbf{R}_c - \mathbf{p}\|} \rho d\rho d\phi \quad (44.1.1)$$

Where  $\mathbf{p} = (\rho, \phi)$  and  $k$  is the wavenumber  $k = \frac{2\pi}{\lambda}$ , and  $\mu_0$  is defined as:

$$\mu_0 = \sin(|B|) \quad (44.1.2)$$

If  $\mathbf{R}_c$  does not lie in the plane of the screen, and if the transmittance  $T(\mathbf{p})$  is bounded, then this integral converges. If we let  $\mathbf{D} = \mathbf{R}_c - \mathbf{p}_0$ , then we can collect all of the exponential terms together as:

$$\psi(\rho_0, \phi_0, \rho, \phi_s) = kD \left[ \sqrt{1 - 2\xi + \eta} - 1 + \eta \right] \quad (44.1.3)$$

Where  $\eta$  and  $\xi$  are defined by:

$$\xi = \cos(B) \left( \frac{\rho \cos(\phi) - \rho_0 \cos(\phi_0)}{D} \right) \quad (44.1.4a)$$

$$\eta = \frac{\rho^2 + \rho_0^2 - 2\rho\rho_0 \cos(\phi - \phi_0)}{D^2} \quad (44.1.4b)$$

Please note that  $\eta$  here is defined as the negative of the  $\eta$  defined in Marouf et. al 1986. This convention is adopted to make it clear later that Legendre polynomials can be applied to the problem. The integral becomes:

$$\hat{T}(\rho_0) = \frac{\sin(B)}{i\lambda} \int_0^{2\pi} \int_0^\infty \rho T(\rho) \frac{\exp(i\psi(\rho_0, \phi_0; \rho, \phi))}{\|\mathbf{R}_c - \rho\|} d\rho d\phi \quad (44.1.5)$$

We impose that  $T$  is non-negative (There's no such thing as 'negative' transmittance). From this we may use a complex version of Fubini's theorem to obtain:

$$\hat{T}(\rho_0) = \frac{\sin(B)}{i\lambda} \int_0^\infty \rho T(\rho) \int_0^{2\pi} \frac{\exp(i\psi(\rho_0, \phi_0; \rho, \phi))}{\|\mathbf{R}_c - \rho\|} d\phi d\rho \quad (44.1.6)$$

In general, this is as far as one can simplify the problem. For ring systems, which we are studying, we can suppose that  $T(\rho) = T(\rho)$ . That is, the screen is radially symmetric. The task is, given the data  $\hat{T}(\rho)$ , can we determine  $T(\rho)$ ? Unfortunately, we don't have an entire planar set of data, but rather some curve. The need then arises to try to collapse this problem down to a single integral by some means of approximation. The Stationary Phase Approximation works well here. Recall that if we have an equation like:

$$I(k) = \int_{\Omega} f(x) \exp(ikg(x)) dx \quad (44.1.7)$$

Where  $g$  is a smooth function with a minimum  $x_0$ , then for large  $k$  we can approximate  $I$  as:

$$I(k) \approx \exp(ikg(x_0)) \sqrt{\frac{2\pi i}{k|g''(c)|}} \quad (44.1.8)$$

### 44.1.2 The Inversion Approximation

The main equation we wish to study is:

$$\hat{T}(\rho_0) = \frac{1-i}{2F} \int_{-\infty}^{\infty} T(\rho) \exp(i\psi(\rho_0, \phi_0, \rho, \phi_s)) d\rho \quad (44.1.9)$$

We wish to solve for  $T(\rho)$ . In general, this is not possible and indeed uniqueness is not always guaranteed. For suppose  $\hat{T}$  is the zero function, and  $\psi$  is a constant. Then any function  $T$  whose integral on the real line is zero will be a solution, and there are infinitely many such functions. Let us suppose that  $\psi$  has the form:

$$\psi = \sum_{n=0}^{\infty} a_n (\rho - \rho_0)^n \quad (44.1.10)$$

It is still not the case that we may solve this. What we wish to do is use the Convolution theorem, which states the following:

**Theorem 44.1.1.** If  $f, g \in L^1(\mathbb{R})$ , then:

$$f * g = \mathcal{F}^{-1}(\mathcal{F}_\xi(f) \cdot \mathcal{F}_\xi(g)) \quad (44.1.11)$$

The requirement that  $f, g \in L^1(\mathbb{R})$  is necessary. For suppose we have:

$$\int_{-\infty}^{\infty} \exp(-\rho^2) \exp(2\pi i(\rho - \rho_0)) d\rho = T * e^{i2\pi\rho} \quad (44.1.12)$$

Taking the Fourier Transform, we get:

$$\mathcal{F}_{\rho_0}(e^{-\rho^2}) \cdot \mathcal{F}_{\rho_0}(e^{2\pi i\rho}) = \sqrt{\pi} \exp(-\pi^2 \rho_0^2) \int_{-\infty}^{\infty} \exp(2\pi i(\rho - \rho_0)) d\rho \quad (44.1.13)$$

And the integral on the right does not converge. We may speak in terms of distributions, or generalized functions (Most commonly, the delta function), but this makes numerical application difficult. Going back to our original problem, the  $\psi$  we are concerned with comes from the Fresnel kernel. That is::

$$\psi(\rho_0, \phi_0, \rho, \phi_s) = kD \left[ \sqrt{1 - 2\xi + \eta} - 1 + \eta \right] \quad (44.1.14)$$

Where  $\eta$  and  $\xi$  are defined by:

$$\eta = \cos(B) \left( \frac{\rho \cos(\phi) - \rho_0 \cos(\phi_0)}{D} \right) \quad (44.1.15a)$$

$$\xi = \frac{\rho^2 + \rho_0^2 - 2\rho\rho_0 \cos(\phi - \phi_0)}{D^2} \quad (44.1.15b)$$

Please note that  $\eta$  here is defined as the negative of the  $\eta$  defined in Marouf et. al 1986. This convention is adopted to make it clear later that Legendre polynomials can be applied to the problem. Unfortunately, for all values of  $\psi$ , we have:

$$\int_{-\infty}^{\infty} |\exp(i\psi)| d\rho = \int_{-\infty}^{\infty} 1 d\rho = +\infty \quad (44.1.16)$$

So it is never have the case that  $\exp(i\psi) \in L^1(\mathbb{R})$ . However, there are many examples of  $\psi$  where the conclusion of the theorem still seems to hold. Let:

$$\psi = \frac{\pi}{2} \rho^2 \quad (44.1.17a) \quad T(\rho) = \exp\left(-\frac{\pi}{2} \rho^2\right) \quad (44.1.17b)$$

Then evaluate the convolution, we have:

$$T * e^{i\psi} = \int_{-\infty}^{\infty} e^{-\frac{\pi}{2}\rho^2} e^{i\frac{\pi}{2}(\rho_0 - \rho)^2} d\rho = \sqrt{1+i} e^{-\frac{1+i}{4}\rho_0^2} \quad (44.1.18)$$

And  $\mathcal{F}_\xi(\sqrt{1+i} e^{-\frac{1+i}{4}\rho_0^2}) = \sqrt{2}(1+i)e^{-(1+i)2\pi\xi^2}$ . Now,  $\mathcal{F}_\xi(e^{-\frac{\pi}{2}\rho^2}) \cdot \mathcal{F}_\xi(e^{i\frac{\pi}{2}\rho^2}) = \sqrt{2}e^{-2\pi\xi^2}(1+i)e^{-2\pi i\xi^2} = \sqrt{2}(1+i)e^{-(1+i)2\pi\xi^2}$ . So we see that, even though  $\psi \notin L^1(\mathbb{R})$ , we still that the result still holds here.

### 44.1.3 Window Functions

We start with the following definition for resolution.

**Definition 44.1.1** The resolution of a reconstruction is:

$$\Delta R = \frac{2F^2}{W_{eff}} \frac{\frac{b^2}{2}}{e^{-b} + b - 1} = \frac{F^2}{W_{eff}} \frac{b^2}{e^{-b} + b - 1} \quad (44.1.19)$$

Where  $W_{eff}$  and  $b$  have the following definitions:

$$W_{eff} = \frac{W}{N_{eq}} \quad (44.1.20a) \qquad b = \frac{\sigma^2 \omega^2}{2\dot{\rho}} W_{eff} \quad (44.1.20b)$$

Here  $N_{eq}$  is the normalized equivalent width,  $\sigma$  is the Allen Deviation,  $\omega$  is the angular frequency, and  $\dot{\rho}$  is the time-derivative of  $\rho$ , where  $\rho(t)$  is the ring radius of the ring intercept point. Let:

$$\alpha = \frac{\sigma^2 \omega^2}{2\dot{\rho}} \quad (44.1.21)$$

Then  $b = \alpha W_{eff}$ . So we have:

$$\Delta R = \frac{F^2}{W_{eff}} \frac{\alpha^2 W_{eff}^2}{\exp(-\alpha W_{eff}) + \alpha W_{eff} - 1} \quad (44.1.22a)$$

$$= \alpha F^2 \frac{\alpha W_{eff}}{\exp(-\alpha W_{eff}) + \alpha W_{eff} - 1} \quad (44.1.22b)$$

While it may seem as though we've done some redundant separation of variables, this last expression can be inverted in terms of the Lambert  $W$  function. Recall that if  $f(x) = x \exp(x)$ ,  $x \in \mathbb{R}^+$ , then there is an inverse function called the Lambert  $W$  function,  $L_W(x)$ , such that:

$$x = L_W(x) \exp(L_W(x)) \quad (44.1.23)$$

Returning to  $b$  once again, we've reduced  $\Delta R$  down to:

$$\Delta R = \alpha F^2 \frac{b}{e^{-b} + b - 1} \Rightarrow \frac{R}{\alpha F^2} = \frac{b}{e^{-b} + b - 1}$$

Letting  $y = \Delta R / \alpha F^2$ , we have:

$$y = \frac{b}{e^{-b} + b - 1} \quad (44.1.24)$$

This is invertable in terms of  $L_W$ :

$$b = L_W\left(\frac{y}{1-y} e^{\frac{y}{1-y}}\right) - \frac{y}{1-y}$$

## 44.2 Problems with Fresnel Inversion

### 44.2.1 Radial Shift from a Linear Phase Offset

**Theorem 44.2.1.** If  $\hat{T}_0(\rho_0) = \hat{T}(\rho_0)e^{i(a\rho+b)}$ , and if  $\hat{f}(\xi) = \mathcal{F}_\xi(\hat{T}(\rho_0))$ , then:

$$\mathcal{F}_\xi(\hat{T}_0(\rho_0)) = e^{ib}\hat{f}\left(\frac{a}{2\pi} + \xi\right) \quad (44.2.1)$$

*Proof.* Let  $\hat{T}(\rho_0)$  be the complex amplitude,  $\hat{T} = \|\hat{T}\|e^{i\theta}$ , where  $\theta = \theta(\rho_0)$  is the phase. Let  $\hat{f}(\xi)$  denote the Fourier Transform of  $\hat{T}(\rho_0)$  onto  $\xi$ . If there is a linear offset in the phase  $a\rho_0 + b$ , we have  $\hat{T}_0 = \hat{T}e^{i(a\rho+b)}$ . Taking the Fourier Transform of this, we have the following:

$$\begin{aligned} \mathcal{F}_\xi(\hat{T}_0) &= \int_{-\infty}^{\infty} \hat{T}(\rho_0)e^{i(a\rho+b)}e^{i2\pi\rho_0\xi}d\rho_0 \\ &= e^{ib}\int_{-\infty}^{\infty} \hat{T}(\rho_0)e^{i\rho_0(a+2\pi\xi)}d\rho_0 \end{aligned}$$

Letting  $\eta = \frac{a}{2\pi} + \xi$ , we have:

$$\begin{aligned} \mathcal{F}_\xi(\hat{T}_0) &= e^{ib}\int_{-\infty}^{\infty} \hat{T}(\rho_0)e^{i2\pi\rho_0\eta}d\rho_0 \\ &= e^{ib}\hat{f}(\eta) \\ &= e^{ib}\hat{f}\left(\frac{a}{2\pi} + \xi\right) \end{aligned}$$

Thus, we have that:

$$\mathcal{F}_\xi(\hat{T}_0) = e^{ib}\hat{f}\left(\frac{a}{2\pi} + \xi\right)$$

So we see that a linear offset in phase creates a horizontal shift in the Fourier transform.  $\square$

We want the effects on  $T(\rho)$ .

**Theorem 44.2.2.** If  $\hat{T}_0(\rho_0) = \hat{T}(\rho_0)e^{i(a\rho+b)}$ ,  $F \neq 0$ ,  $T(\rho) = \int_{-\infty}^{\infty} \hat{T}(\rho_0)e^{i\frac{\pi}{2}\left(\frac{\rho-\rho_0}{F}\right)^2}d\rho_0$ , and if  $T_0(\rho) = \int_{-\infty}^{\infty} \hat{T}_0(\rho_0)e^{i\frac{\pi}{2}\left(\frac{\rho-\rho_0}{F}\right)^2}d\rho_0$ , then  $\|T_0(\rho)\| = \|T(\rho - \frac{aF^2}{\pi})\|$

*Proof.*

$$\begin{aligned} T_0(\rho) &= \int_{-\infty}^{\infty} \hat{T}_0(\rho_0) e^{i\frac{\pi}{2}\left(\frac{\rho-\rho_0}{F}\right)^2} d\rho_0 \\ &= \int_{-\infty}^{\infty} \hat{T}(\rho_0) e^{i(a\rho_0+b)} e^{i\frac{\pi}{2}\left(\frac{\rho-\rho_0}{F}\right)^2} d\rho_0 \\ &= e^{ib} \int_{-\infty}^{\infty} \hat{T}(\rho_0) e^{i\frac{\pi}{2}\left[\left(\frac{\rho-\rho_0}{F}\right)^2 + \frac{2a}{\pi}\rho_0\right]} d\rho_0 \end{aligned}$$

Expanding the terms in the exponential and simplifying, we have:

$$\begin{aligned} \left(\frac{\rho-\rho_0}{F}\right)^2 + \frac{2a}{\pi}\rho_0 &= \frac{\rho_0^2 - 2\rho\rho_0 + \rho^2 + \frac{2aF^2}{\pi}\rho_0}{F^2} \\ &= \frac{\rho_0^2 - 2\rho_0\left(\rho - \frac{aF^2}{\pi}\right) + \rho^2}{F^2} \\ &= \frac{\left(\rho_0 - \left(\rho - \frac{aF^2}{\pi}\right)\right)^2 - (\rho - \frac{aF^2}{\pi})^2 + \rho^2}{F^2} \\ &= \frac{\left(\rho_0 - \left(\rho - \frac{aF^2}{\pi}\right)\right)^2 + \frac{2aF^2}{\pi}\rho - \frac{a^2F^4}{\pi^2}}{F^2} \\ &= \frac{\left(\rho_0 - \left(\rho - \frac{aF^2}{\pi}\right)\right)^2}{F^2} + \frac{2a}{\pi}\rho - \frac{a^2F^2}{\pi^2} \end{aligned}$$

The integral is over  $\rho_0$ , so we may write:

$$T_0(\rho) = e^{ib} e^{i\frac{\pi}{2}\left(\frac{2a}{\pi}\rho - \frac{a^2F^2}{\pi^2}\right)} \int_{-\infty}^{\infty} \hat{T}(\rho_0) e^{i\frac{\pi}{2}\left(\frac{\rho_0 - (\rho - \frac{aF^2}{\pi})}{F}\right)^2} d\rho_0$$

Let  $u = \rho - \frac{aF^2}{\pi}$ . Then we have:

$$\int_{-\infty}^{\infty} \hat{T}(\rho_0) e^{i\frac{\pi}{2}\left(\frac{\rho_0 - u}{F}\right)^2} d\rho_0 = T(u)$$

Therefore:

$$T_0(\rho) = e^{ib} e^{i\frac{\pi}{2}\left(\frac{2a}{\pi}\rho - \frac{a^2F^2}{\pi^2}\right)} T\left(\rho - \frac{aF^2}{\pi}\right)$$

Computing reconstructed power takes the norm  $\|T_0(\rho)\|$ , and  $\|e^{ib} e^{i\frac{\pi}{2}\left(\frac{2a}{\pi}\rho - \frac{a^2F^2}{\pi^2}\right)}\| = 1$ , for all values of  $a, F, \rho$  (This is from Euler's theorem). Thus:

$$\|T_0(\rho)\| = \|T\left(\rho - \frac{aF^2}{\pi}\right)\|$$

So a linear offset  $a\rho_0 + b$  in the phase creates a radial offset in the reconstructed power of  $-\frac{aF^2}{\pi}$ . □

### 44.2.2 Notes on the Fresnel Approximation

$$\begin{aligned}
\psi &= \left( 1 + \frac{\rho^2 + \rho_0^2 - 2\rho\rho_0 \cos(\phi - \phi_0)}{D^2} - 2 \cos(B) \left( \frac{\rho \cos(\phi) - \rho_0 \cos(\phi_0)}{D} \right) \right)^{1/2} \\
&\quad - (1 - \cos(B)) \left( \frac{\rho \cos(\phi) - \rho_0 \cos(\phi_0)}{D} \right) \\
&= \sqrt{1 + \eta - 2\xi} - (1 - \xi) \\
\Rightarrow \frac{\partial \psi}{\partial \phi} &= \frac{1}{2\sqrt{1 + \eta - 2\xi}} \left( \frac{\partial \eta}{\partial \phi} - 2 \frac{\partial \xi}{\partial \phi} \right) + \frac{\partial \xi}{\partial \phi} \\
\Rightarrow \frac{\partial^2 \psi}{\partial \phi^2} &= \frac{-1}{4(1 + \eta - 2\xi)^{3/2}} \left( \frac{\partial \eta}{\partial \phi} - 2 \frac{\partial \xi}{\partial \phi} \right)^2 + \frac{1}{2\sqrt{1 + \eta - 2\xi}} \left( \frac{\partial^2 \eta}{\partial \phi^2} - 2 \frac{\partial^2 \xi}{\partial \phi^2} \right) + \frac{\partial^2 \xi}{\partial \phi^2}
\end{aligned}$$

Now, from the definitions of  $\eta$  and  $\xi$ , we have:

$$\begin{aligned}
\eta_{\phi=\phi_0} &= \left( \frac{\rho - \rho_0}{D} \right)^2 & \xi_{\phi=\phi_0} &= \cos(B) \cos(\phi_0) \left( \frac{\rho - \rho_0}{D} \right) \\
\frac{\partial \eta}{\partial \phi}_{\phi=\phi_0} &= 0 & \frac{\partial \xi}{\partial \phi}_{\phi=\phi_0} &= -\cos(B) \frac{\rho \sin(\phi_0)}{D} \\
\frac{\partial^2 \eta}{\partial \phi^2}_{\phi=\phi_0} &= \frac{2\rho\rho_0}{D^2} & \frac{\partial^2 \xi}{\partial \phi^2}_{\phi=\phi_0} &= -\cos(B) \frac{\rho \cos(\phi_0)}{D}
\end{aligned}$$

Let  $\alpha = \cos(B) \cos(\phi_0)$  and  $x = (\frac{\rho - \rho_0}{D})^2$ . From this, we obtain:

$$\begin{aligned}
\psi_{\phi=\phi_0} &= \sqrt{1 + x^2 - 2\alpha x} + \alpha x - 1 \\
\frac{\partial \psi}{\partial \phi}_{\phi=\phi_0} &= \cos(B) \sin(\phi_0) \frac{\rho}{D} \left( \frac{1}{\sqrt{1 + x^2 - 2\alpha x}} - 1 \right) \\
\frac{\partial^2 \psi}{\partial \phi^2}_{\phi=\phi_0} &= \frac{-\rho^2 \cos^2(B) \sin^2(\phi_0)}{D^2(1 + x^2 - 2\alpha x)^{3/2}} + \frac{\rho\rho_0}{D^2(1 + x^2 - 2\alpha x)^{1/2}} + \frac{\alpha\rho}{D} \left( \frac{1}{\sqrt{1 + x^2 - 2\alpha x}} - 1 \right)
\end{aligned}$$

A nice property emerges here related to Legendre polynomials. The following is true:

$$\frac{1}{\sqrt{1 - 2\alpha x + x^2}} = \sum_{n=0}^{\infty} P_n(\alpha) x^n$$

Where  $P_n(\alpha)$  is the  $n^{th}$  Legendre polynomial. That is, this is the *generating function* for the Legendre polynomials. Using this we can easily compute the Taylor series expansions for  $\psi$  and  $\psi_\phi$  about  $x = 0$ .  $\psi_{\phi\phi}$  is a nastier type of monster. Using this, we obtain the following equations:

$$\begin{aligned}
\psi_{\phi=\phi_0} &= \frac{1 - \alpha^2}{2} x^2 + \frac{\alpha(1 - \alpha^2)}{3} x^3 + \frac{-5\alpha^4 + 6\alpha^2 - 1}{8} x^4 + \dots \\
\frac{\partial \psi}{\partial \phi}_{\phi=\phi_0} &= \cos(B) \sin(\phi_0) \frac{\rho}{D} \left( \alpha x + \frac{3\alpha^2 - 1}{2} x^2 + \frac{35\alpha^4 - 30\alpha^2 + 3}{8} x^4 + \dots \right)
\end{aligned}$$

We wish to find the point  $\phi$  for which  $\psi_\phi = 0$ . We can use Newton-Raphson with initial guess  $\phi = \phi_0$ . Note that, from analyticity:

$$\psi = \sum_{k=0}^{\infty} \psi^{(k)}(\phi = \phi_s) \frac{(\phi - \phi_s)^k}{k!} = \psi_s + \psi'_s(\phi - \phi_s) + \frac{1}{2}\psi''_s(\phi - \phi_s)^2 + \dots$$

Where all derivatives are taken with respect to  $\phi$ , and  $\psi_s$  denotes the derivatives evaluated at  $\phi = \phi_s$ . We may form a sequence of points and functions defined by the Newton-Raphson method as follows:

$$\begin{aligned}\phi_{n+1} &= \phi_n - \psi'_n / \psi''_n \\ \psi &\approx \psi_n + \psi'_n(\phi_{n+1} - \phi_n) + \frac{1}{2}\psi''_n(\phi_{n+1} - \phi_n)^2 \\ &= \psi_n + \psi'(-\psi'_n / \psi''_n) + \frac{1}{2}\psi''_n(-\psi'_n / \psi''_n)^2 \\ &= \psi_n - \frac{1}{2}\psi'^2_n / \psi''_n\end{aligned}$$

Now, let's use the equations we've found before to evaluate the first iteration of this approximation. First we only consider the leading terms of each expansion. The leading term of  $\partial^2\psi/\partial\phi^2$  occurs when we evaluate at  $x = 0$ . This occurs only when  $\rho = \rho_0$ . From this, we have the following:

$$\begin{aligned}\psi_{\phi=\phi_0} &\approx \frac{1}{2}(1 - \cos^2(B)\cos^2(\phi_0))\left(\frac{\rho - \rho_0}{D}\right)^2 \\ \frac{\partial\psi}{\partial\phi}_{\phi=\phi_0} &\approx \cos^2(B)\sin(\phi_0)\cos(\phi_0)\frac{\rho}{D}\left(\frac{\rho - \rho_0}{D}\right) \\ \frac{\partial^2\psi}{\partial\phi^2}_{\phi=\phi_0} &\approx (1 - \cos^2(B)\sin^2(\phi_0))\left(\frac{\rho_0}{D}\right)^2\end{aligned}$$

From this, the first approximation becomes:

$$\begin{aligned}\phi_1 &= \phi_0 - \frac{\cos^2(B)\sin(\phi_0)\cos(\phi_0)}{1 - \cos^2(B)\sin^2(\phi_0)} \frac{\rho}{\rho_0} \left(\frac{\rho - \rho_0}{\rho_0}\right) \\ \psi &\approx \frac{1}{2}(1 - \cos^2(B)\cos^2(\phi_0))\left(\frac{\rho - \rho_0}{D}\right)^2 - \frac{1}{2}\frac{\cos^4(B)\sin^2(\phi_0)\cos^2(\phi_0)}{1 - \cos^2(B)\sin^2(\phi_0)} \frac{\rho^2}{\rho_0^2} \left(\frac{\rho - \rho_0}{D}\right)^2 \\ &\approx \frac{\sin^2(B)}{1 - \cos^2(B)\sin^2(\phi_0)} \left(\frac{\rho - \rho_0}{D}\right)^2\end{aligned}$$

The second approximation comes from the fact that we have assumed that  $\rho^2/\rho_0^2 \approx 1$  in the algebra for this final expression. When we apply the weight of  $kD$ , where  $k$  is the wavenumber, we get the classic Fresnel approximation:

$$\psi_Q = \frac{\pi}{2} \left(\frac{\rho - \rho_0}{F}\right)^2 \quad F^2 = \frac{\lambda D}{2} \frac{1 - \cos^2(B)\sin^2(\phi_0)}{\sin^2(B)}$$

This gives us a quadratic approximation that is both very easy to compute and gives decent diffraction reconstructions for many occultation observations. In particular, this applies very well to the Rev007E Cassini occultation. During more pathological geometries, there is a need to use a better approximation and to use higher order terms. For ease of analysis, and to better take advantage of these Legendre polynomials, we will continue to use the approximation that  $\rho^2/\rho_0^2 \approx 1$ . Since  $\rho_0$  is usually greater than 65,000 (Radius of Saturn) and  $\rho - \rho_0$  is bounded by the window width, the worst case scenario one could reasonably expect is 4,000 kilometer windows, in which case the ratio is bounded by 1.06. For 10,000 kilometer windows the ratio is bounded by 1.14. The approximation can thus be generalized to:

$$\psi \approx \sum_{k=0}^{N-1} b_k x^{k+2} - \frac{1}{2} \frac{\cos^2(B) \sin^2(\phi_0)}{1 - \cos^2(B) \sin^2(\phi_0)} \left( \sum_{k=1}^N P_k(\alpha) x^k \right)^2$$

Where  $b_k$  is the  $k^{th}$  term of the expansion for  $\psi$ , and  $P_n$  is the  $n^{th}$  Legendre Polynomial. When  $N = 1$ , we obtain the Fresnel approximation we previously derived. It would be convenient if we could write the  $b_k$  in terms of the  $P_k(\alpha)$ . We can do this by solving the following simple differential equation.

$$\begin{aligned} \frac{d}{dx} \left( \sqrt{1 + x^2 - \alpha x} \right) &= \frac{x - \alpha}{\sqrt{1 + x^2 - \alpha x}} \\ &= (x - \alpha) \sum_{k=0}^{\infty} P_k(\alpha) x^k \\ &= \sum_{k=0}^{\infty} P_k(\alpha) x^{k+1} - \alpha \sum_{k=0}^{\infty} P_k(\alpha) x^k \\ \Rightarrow \sqrt{1 + x^2 - 2\alpha x} &= 1 + \sum_{k=0}^{\infty} P_k(\alpha) \frac{x^{k+2}}{k+2} - \alpha \sum_{k=0}^{\infty} P_k(\alpha) \frac{x^{k+1}}{k+1} \\ \Rightarrow \sqrt{1 + x^2 - 2\alpha x} + \alpha x - 1 &= \sum_{k=0}^{\infty} \left( P_k(\alpha) - \alpha P_{k+1}(\alpha) \right) \frac{x^{k+2}}{k+2} \end{aligned}$$

Where we used a shift of index to obtain the last equation. This gives us our formula:

$$b_k = \frac{P_k(\alpha) - \alpha P_{k+1}(\alpha)}{k+2}$$

We have the following first order approximation for  $\psi$ . Higher order approximations can be computed by evaluating  $\psi_n - \psi_n'^2/\psi_n''$  at better approximations obtained by the Newton-Raphson method for  $\psi$ .

$$\psi \approx \sum_{k=0}^{N-1} \frac{P_k(\alpha) - \alpha P_{k+1}(\alpha)}{k+2} x^{k+2} - \frac{1}{2} \frac{\cos^2(B) \sin^2(\phi_0)}{1 - \cos^2(B) \sin^2(\phi_0)} \left( \sum_{k=1}^N P_k(\alpha) x^k \right)^2$$

In the cases where  $\psi_{\phi\phi}$  can not be assumed constant, we need to keep track of the how it's Taylor expansion behaves beyond zeroth order. We have:

$$\begin{aligned}\frac{\partial^2 \psi}{\partial \phi^2}_{\phi=\phi_0} &= \frac{-\rho^2 \cos^2(B) \sin^2(\phi_0)}{D^2(1+x^2-2\alpha x)^{3/2}} + \frac{\rho \rho_0}{D^2(1+x^2-2\alpha x)^{1/2}} + \frac{\alpha \rho}{D} \left( \frac{1}{\sqrt{1+x^2-2\alpha x}} - 1 \right) \\ &\approx \frac{\rho^2}{D^2} \left( \frac{-\cos^2(B) \sin^2(\phi_0)}{(1+x^2-2\alpha x)^{3/2}} + \frac{1}{(1+x^2-2\alpha x)^{1/2}} + \frac{\alpha D}{\rho} \left( \frac{1}{\sqrt{1+x^2-2\alpha x}} - 1 \right) \right) \\ &\approx \frac{\rho^2}{D^2} \left( \sum_{n=0}^{\infty} P_n(\alpha) x^n \left( 1 - \frac{\cos^2(B) \sin^2(\phi_0)}{1+x^2-2\alpha x} \right) + \frac{\alpha D}{\rho} \left( \sum_{n=1}^{\infty} P_n(\alpha) x^n \right) \right)\end{aligned}$$

Where the approximation is because we have once again assumed that  $\rho/\rho_0 \approx 1$ . The ratio of  $\psi'^2/\psi''$  then becomes:

$$\frac{\psi'}{\psi''}_{\phi=\phi_0} \approx \frac{\cos^2(B) \sin^2(\phi_0) \left( \sum_{k=1}^N P_k(\alpha) x^n \right)^2}{\sum_{n=0}^{\infty} P_n(\alpha) x^n \left( 1 - \frac{\cos^2(B) \sin^2(\phi_0)}{1+x^2-2\alpha x} \right) + \frac{\alpha D}{\rho} \left( \sum_{n=1}^{\infty} P_n(\alpha) x^n \right)}$$

Using the fact that  $P_0(\alpha) = 1$ , we can simplify this to:

$$\frac{\psi'}{\psi''}_{\phi=\phi_0} \approx \frac{\cos^2(B) \sin^2(\phi_0) \left( \sum_{k=1}^N P_k(\alpha) x^n \right)^2}{\sum_{n=0}^{\infty} P_n(\alpha) x^n \left( 1 + \frac{\alpha D}{\rho} - \frac{\cos^2(B) \sin^2(\phi_0)}{1+x^2-2\alpha x} \right) - \frac{\alpha D}{\rho}}$$

Stopping the sum at  $n = 0$ , we retrieve our previous approximation. The only limitation to this approximation is the validity of  $\rho/\rho_0 \approx 1$ . In most cases, this is reasonable.

**Part XXVIII**

# **Electromagnetism**



# CHAPTER 45

---

## Electromagnetism I

---

### 45.1 Homework Sets

#### 45.1.1 Homework I

Wangsness Chapter 1 - Problems: 2, 3, 4, 5, 8, 9

**Problem 45.1.1** Given  $\mathbf{A} = 2\hat{\mathbf{x}} - 3\hat{\mathbf{y}} - 4\hat{\mathbf{z}}$  and  $\mathbf{B} = 6\hat{\mathbf{x}} + 5\hat{\mathbf{y}} + \hat{\mathbf{z}}$ , find the magnitudes and angles made with the  $x$ ,  $y$ , and  $z$  axes for  $\mathbf{A} + \mathbf{B}$  and  $\mathbf{A} - \mathbf{B}$ .

*Solution* First, we need to find  $\mathbf{A} + \mathbf{B}$  and  $\mathbf{A} - \mathbf{B}$ :

$$\mathbf{A} + \mathbf{B} = (2\hat{\mathbf{x}} - 3\hat{\mathbf{y}} - 4\hat{\mathbf{z}}) + (6\hat{\mathbf{x}} + 5\hat{\mathbf{y}} + \hat{\mathbf{z}}) \quad (45.1.1a)$$

$$= (2+6)\hat{\mathbf{x}} + (5-3)\hat{\mathbf{y}} + (1-4)\hat{\mathbf{z}} \quad (45.1.1b)$$

$$= 8\hat{\mathbf{x}} + 2\hat{\mathbf{y}} - 3\hat{\mathbf{z}} \quad (45.1.1c)$$

$$\mathbf{A} - \mathbf{B} = (2\hat{\mathbf{x}} - 3\hat{\mathbf{y}} - 4\hat{\mathbf{z}}) - (6\hat{\mathbf{x}} + 5\hat{\mathbf{y}} + \hat{\mathbf{z}}) \quad (45.1.2a)$$

$$= (2-6)\hat{\mathbf{x}} - (3+5)\hat{\mathbf{y}} - (4+1)\hat{\mathbf{z}} \quad (45.1.2b)$$

$$= -4\hat{\mathbf{x}} - 8\hat{\mathbf{y}} - 5\hat{\mathbf{z}} \quad (45.1.2c)$$

The magnitude of a vector  $\mathbf{A} = a_1\hat{\mathbf{x}}_1 + \cdots + a_N\hat{\mathbf{x}}_N$ , also called its *norm*, is:

$$\|\mathbf{A}\| = \sqrt{\sum_{i=1}^N a_i^2} \quad (45.1.3)$$

Using this, we have:

$$\|\mathbf{A} + \mathbf{B}\| = (8^2 + 2^2 + 3^2)^{1/2} \quad (45.1.4a)$$

$$= \sqrt{77} \quad (45.1.4b)$$

$$\|\mathbf{A} - \mathbf{B}\| = (4^2 + 8^2 + 5^2)^{1/2} \quad (45.1.4c)$$

$$= \sqrt{105} \quad (45.1.4d)$$

The *direction angle* between  $\mathbf{A}$  and the  $\xi$  axis is:

$$\alpha_\xi = \cos^{-1} \left( \frac{\mathbf{A} \cdot \hat{\boldsymbol{\xi}}}{\|\mathbf{A}\| \|\hat{\boldsymbol{\xi}}\|} \right) = \cos^{-1} \left( \frac{\mathbf{A} \cdot \hat{\boldsymbol{\xi}}}{\|\mathbf{A}\|} \right) \quad (45.1.5)$$

The direction angles of  $\mathbf{A} + \mathbf{B}$  and  $\mathbf{A} - \mathbf{B}$  for  $\hat{\mathbf{x}}$ ,  $\hat{\mathbf{y}}$ , and  $\hat{\mathbf{z}}$  are:

$$\alpha = \cos^{-1} \left( \frac{(\mathbf{A} + \mathbf{B}) \cdot \hat{\mathbf{x}}}{\|\mathbf{A} + \mathbf{B}\|} \right) = \cos^{-1} \left( \frac{8}{\sqrt{77}} \right) = 24.3^\circ \quad (45.1.6a)$$

$$\beta = \cos^{-1} \left( \frac{(\mathbf{A} + \mathbf{B}) \cdot \hat{\mathbf{y}}}{\|\mathbf{A} + \mathbf{B}\|} \right) = \cos^{-1} \left( \frac{2}{\sqrt{77}} \right) = 76.8^\circ \quad (45.1.6b)$$

$$\gamma = \cos^{-1} \left( \frac{(\mathbf{A} + \mathbf{B}) \cdot \hat{\mathbf{z}}}{\|\mathbf{A} + \mathbf{B}\|} \right) = \cos^{-1} \left( \frac{-3}{\sqrt{77}} \right) = 110^\circ \quad (45.1.6c)$$

For  $\mathbf{A} - \mathbf{B}$ :

$$\alpha = \cos^{-1} \left( \frac{(\mathbf{A} - \mathbf{B}) \cdot \hat{\mathbf{x}}}{\|\mathbf{A} - \mathbf{B}\|} \right) = \cos^{-1} \left( \frac{-4}{\sqrt{105}} \right) = 113^\circ \quad (45.1.7a)$$

$$\beta = \cos^{-1} \left( \frac{(\mathbf{A} - \mathbf{B}) \cdot \hat{\mathbf{y}}}{\|\mathbf{A} - \mathbf{B}\|} \right) = \cos^{-1} \left( \frac{-8}{\sqrt{105}} \right) = 141.3^\circ \quad (45.1.7b)$$

$$\gamma = \cos^{-1} \left( \frac{(\mathbf{A} - \mathbf{B}) \cdot \hat{\mathbf{z}}}{\|\mathbf{A} - \mathbf{B}\|} \right) = \cos^{-1} \left( \frac{-5}{\sqrt{105}} \right) = 119.2^\circ \quad (45.1.7c)$$

**Problem 45.1.2** Find the relative position vector  $\mathbf{R}$  of  $\mathbf{P} = (2, -2, 3)$  with respect to  $\mathbf{P}' = (-3, 1, 4)$ . What are the direction angles of  $\mathbf{R}$ ?

*Solution* The relative position vector of  $\mathbf{B}$  with respect to  $\mathbf{A}$  is:

$$\mathbf{R}_{\mathbf{A} \rightarrow \mathbf{B}} = \mathbf{B} - \mathbf{A} \quad (45.1.8)$$

Thus, we have:

$$\mathbf{R} = \mathbf{P} - \mathbf{P}' \quad (45.1.9a)$$

$$= (2\hat{\mathbf{x}} - 2\hat{\mathbf{y}} + 3\hat{\mathbf{z}}) - (-3\hat{\mathbf{x}} + \hat{\mathbf{y}} + 4\hat{\mathbf{z}}) \quad (45.1.9b)$$

$$= (2 + 3)\hat{\mathbf{x}} + (-2 - 1)\hat{\mathbf{y}} + (3 - 4)\hat{\mathbf{z}} \quad (45.1.9c)$$

$$= 5\hat{\mathbf{x}} - 3\hat{\mathbf{y}} - \hat{\mathbf{z}} \quad (45.1.9d)$$

The direction angles for  $\mathbf{R}$  are:

$$\alpha = \cos^{-1} \left( \frac{\mathbf{R} \cdot \hat{\mathbf{x}}}{\|\mathbf{R}\|} \right) = \cos^{-1} \left( \frac{5}{\sqrt{35}} \right) = 32.5^\circ \quad (45.1.10a)$$

$$\beta = \cos^{-1} \left( \frac{\mathbf{R} \cdot \hat{\mathbf{y}}}{\|\mathbf{R}\|} \right) = \cos^{-1} \left( \frac{-3}{\sqrt{35}} \right) = 120^\circ \quad (45.1.10b)$$

$$\gamma = \cos^{-1} \left( \frac{\mathbf{R} \cdot \hat{\mathbf{z}}}{\|\mathbf{R}\|} \right) = \cos^{-1} \left( \frac{-1}{\sqrt{35}} \right) = 99.7^\circ \quad (45.1.10c)$$

**Problem 45.1.3** Given  $\mathbf{A} = \hat{\mathbf{x}} + 2\hat{\mathbf{y}} + 3\hat{\mathbf{z}}$  and  $\mathbf{B} = 4\hat{\mathbf{x}} - 5\hat{\mathbf{y}} + 6\hat{\mathbf{z}}$ , find the angle between them. Find the component of  $\mathbf{A}$  in the direction of  $\mathbf{B}$ .

*Solution* The definition of the *angle* between two vectors  $\mathbf{A}$  and  $\mathbf{B}$  is:

$$\theta = \cos^{-1} \left( \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} \right) \quad (45.1.11)$$

We have that:

$$\mathbf{A} \cdot \mathbf{B} = 1 \cdot 4 - 2 \cdot 5 + 3 \cdot 6 = 12 \quad (45.1.12)$$

The norms of  $\mathbf{A}$  and  $\mathbf{B}$  are computed as follows:

$$\|\mathbf{A}\| = \sqrt{1^2 + 2^2 + 3^2} \quad (45.1.13a) \quad \|\mathbf{B}\| = \sqrt{4^2 + 5^2 + 6^2} \quad (45.1.13c)$$

$$= \sqrt{14} \quad (45.1.13b) \quad = \sqrt{77} \quad (45.1.13d)$$

Using this, we obtain:

$$\theta = \cos^{-1} \left( \frac{12}{\sqrt{14}\sqrt{77}} \right) = 68.6^\circ \quad (45.1.14)$$

The *component* of  $\mathbf{A}$  in the direction of  $\mathbf{B}$  is defined as:

$$\text{comp}_{\mathbf{B}}(\mathbf{A}) = \mathbf{A} \cdot \frac{\mathbf{B}}{\|\mathbf{B}\|} \quad (45.1.15)$$

Using this, we have:

$$\text{comp}_{\mathbf{B}}(\mathbf{A}) = \mathbf{A} \cdot \frac{\mathbf{B}}{\|\mathbf{B}\|} = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{B}\|} = \frac{12}{\sqrt{77}} \approx 1.37 \quad (45.1.16)$$

**Problem 45.1.4** Given the vectors  $\mathbf{A} = 2\hat{\mathbf{x}} + 3\hat{\mathbf{y}} - 4\hat{\mathbf{z}}$  and  $\mathbf{B} = -6\hat{\mathbf{x}} - 4\hat{\mathbf{y}} + \hat{\mathbf{z}}$ , find the component of  $\mathbf{A} \times \mathbf{B}$  along the direction of  $\mathbf{C} = \hat{\mathbf{x}} - \hat{\mathbf{y}} + \hat{\mathbf{z}}$ .

*Solution* The *cross product* of  $\mathbf{A}$  with  $\mathbf{B}$  is:

$$\mathbf{A} \times \mathbf{B} = (A_y B_z - A_z B_y) \hat{\mathbf{x}} + (A_z B_x - A_x B_z) \hat{\mathbf{y}} + (A_x B_y - A_y B_x) \hat{\mathbf{z}} \quad (45.1.17)$$

Note that  $\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}$ . A way to remember this formula is using matrices:

$$\mathbf{A} \times \mathbf{B} = \det \begin{pmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{pmatrix} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \quad (45.1.18)$$

We have:

$$\mathbf{A} \times \mathbf{B} = (2\hat{\mathbf{x}} + 3\hat{\mathbf{y}} - 4\hat{\mathbf{z}}) \times (-6\hat{\mathbf{x}} - 4\hat{\mathbf{y}} + \hat{\mathbf{z}}) \quad (45.1.19a)$$

$$= (3 - 16)\hat{\mathbf{x}} + (24 - 2)\hat{\mathbf{y}} + (-8 + 18)\hat{\mathbf{z}} \quad (45.1.19b)$$

$$= -13\hat{\mathbf{x}} + 22\hat{\mathbf{y}} + 10\hat{\mathbf{z}} \quad (45.1.19c)$$

The component along the direction of  $\mathbf{C}$  is:

$$\text{comp}_{\mathbf{C}}(\mathbf{A} \times \mathbf{B}) = (\mathbf{A} \times \mathbf{B}) \cdot \frac{\mathbf{C}}{\|\mathbf{C}\|} = -\frac{25}{\sqrt{3}} \quad (45.1.20)$$

**Problem 45.1.5** Given a family of hyperbolas in the  $xy$  plane  $u = xy$ , find  $\text{grad}(u)$ . If  $\mathbf{A} = 3\hat{\mathbf{x}} + 2\hat{\mathbf{y}} + 4\hat{\mathbf{z}}$ , find the component of  $\mathbf{A}$  in the direction of  $\text{grad}(u)$  at the point on the curve for which  $u = 3$  and  $x = 2$ .

*Solution*  $\text{grad}(u)$  is called the *gradient* of  $u$ . It is also common to write  $\nabla(u)$ . In Cartesian coordinates this can be written as:

$$\text{grad}(u) = \sum_{i=1}^N \frac{\partial u}{\partial x_i} \hat{\mathbf{x}}_i \quad (45.1.21)$$

Where  $\frac{\partial u}{\partial x_i}$  is the partial derivative of  $u$  with respect to the  $i^{th}$  coordinate. Thus:

$$\text{grad}(u) = \frac{\partial(xy)}{\partial x} \hat{\mathbf{x}} + \frac{\partial(xy)}{\partial y} \hat{\mathbf{y}} = y\hat{\mathbf{x}} + x\hat{\mathbf{y}} \quad (45.1.22)$$

When  $u = 3$  and  $x = 2$ , we have  $y = 3/2$ . We then obtain:

$$\text{grad}(u)_{(2,3/2)} = \frac{3}{2}\hat{\mathbf{x}} + 2\hat{\mathbf{y}} \quad (45.1.23a) \quad \|\text{grad}(u)_{(2,3/2)}\| = \frac{5}{2} \quad (45.1.23b)$$

So the component of  $\mathbf{A}$  along  $\text{grad}(u)$  when  $u = 3$  and  $x = 2$  is:

$$\text{comp}_{\text{grad}(u)}(\mathbf{A}) = \mathbf{A} \cdot \frac{\text{grad}(u)}{\|\text{grad}(u)\|} = (3\hat{\mathbf{x}} + 2\hat{\mathbf{y}} - 4\hat{\mathbf{z}}) \cdot \left( \frac{3\hat{\mathbf{x}} + 4\hat{\mathbf{y}}}{5} \right) = \frac{17}{5} \quad (45.1.24)$$

**Problem 45.1.6** An ellipsoid is define by the equation:

$$u = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \quad (45.1.25)$$

Find the unit vector normal to the surface of each point of an ellipsoid.

*Solution* The vector normal to a surface  $u$  is the gradient. The unit vector normal to a surface would then be:

$$\hat{\mathbf{n}} = \frac{\text{grad}(u)}{\|\text{grad}(u)\|} \quad (45.1.26)$$

We have:

$$\text{grad}(u) = \frac{\partial u}{\partial x} \hat{\mathbf{x}} + \frac{\partial u}{\partial y} \hat{\mathbf{y}} + \frac{\partial u}{\partial z} \hat{\mathbf{z}} \quad (45.1.27a)$$

$$= 2a^{-2}x \hat{\mathbf{x}} + 2b^{-2}y \hat{\mathbf{y}} + 2c^{-2}z \hat{\mathbf{z}} \quad (45.1.27b)$$

The norm of  $\text{grad}(u)$  is then:

$$\|\text{grad}(u)\| = 2\sqrt{a^{-4}x^2 + b^{-4}y^2 + c^{-4}z^2} \quad (45.1.28)$$

From the definition of  $\hat{\mathbf{n}}$ , we obtain:

$$\hat{\mathbf{n}} = \frac{\text{grad}(u)}{\|\text{grad}(u)\|} = \frac{a^{-2}x \hat{\mathbf{x}} + b^{-2}y \hat{\mathbf{y}} + c^{-2}z \hat{\mathbf{z}}}{\sqrt{a^{-4}x^2 + b^{-4}y^2 + c^{-4}z^2}} \quad (45.1.29)$$

## 45.1.2 Homework II

Wangness Chapter 1 - Problems: 11, 12, 13, 14, 15

**Problem 45.1.7** (Wangness 1-11) Calculate the path integral of the vector valued function  $\mathbf{A}$  defined by:

$$\mathbf{A}(x, y, z) = x^2 \hat{\mathbf{x}} + y^2 \hat{\mathbf{y}} + z^2 \hat{\mathbf{z}} \quad (45.1.30)$$

Integrating along the path shown in Fig. 45.1.1 by integrating over  $y$ .

*Solution* The *path integral* of  $\mathbf{A}$  along a path  $C$  is:

$$\int_C \mathbf{A} \cdot d\ell = \int_C \mathbf{A} \cdot (dx \hat{\mathbf{x}} + dy \hat{\mathbf{y}} + dz \hat{\mathbf{z}}) \quad (45.1.31)$$

We have  $\mathbf{A} = x^2 \hat{\mathbf{x}} + y^2 \hat{\mathbf{y}} + z^2 \hat{\mathbf{z}}$ . Using this, we obtain:

$$\int_C \mathbf{A} \cdot d\ell = \int_C (x^2 \hat{\mathbf{x}} + y^2 \hat{\mathbf{y}} + z^2 \hat{\mathbf{z}}) \cdot (dx \hat{\mathbf{x}} + dy \hat{\mathbf{y}} + dz \hat{\mathbf{z}}) \quad (45.1.32a)$$

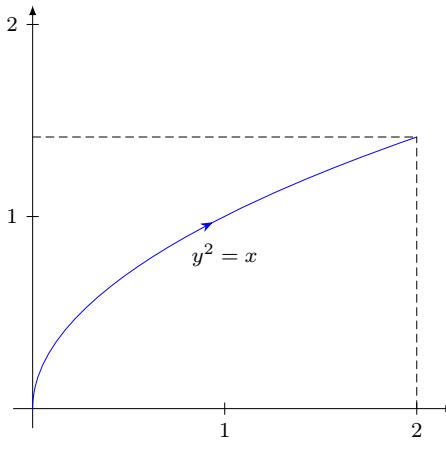
$$= \int_C (x^2 dx + y^2 dy) \quad (45.1.32b)$$

Along the path of integration, we have  $x = y^2$ , and therefore  $dx = 2y dy$ . Thus:

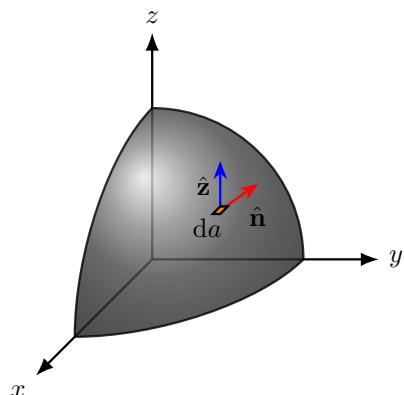
$$\int_C \mathbf{A} \cdot d\ell = \int_C (x^2 dx + y^2 dy) = \int_0^{\sqrt{2}} ((y^2)^2 (2y dy) + y^2 dy) \quad (45.1.33)$$

We can simplify this further to obtain:

$$\int_C \mathbf{A} \cdot d\ell = \int_0^{\sqrt{2}} (2y^5 + y^2) dy = \frac{1}{3} [y^6 + y^3]_0^{\sqrt{2}} = \frac{2}{3} (4 + \sqrt{2}) \quad (45.1.34)$$



45.1.1: Path of Integration for Wangsness 1-11.



45.1.2: Geometry for Wangsness 1-12.

Fig. 45.1: Figures for Problems 45.1.7 and 45.1.8, Respectively.

**Problem 45.1.8** (Wangsness 1-12) Find the surface integral of  $\mathbf{r}$  and the volume integral of  $\operatorname{div}(\mathbf{r})$  for a sphere of radius  $a_0$  centered at the origin.

*Solution* The *divergence* of a vector field  $\mathbf{A}$  in Cartesian coordinates is:

$$\operatorname{div}(\mathbf{A}) = \sum_{k=1}^n \frac{\partial A_k}{\partial x_k} \quad (45.1.35)$$

It is also common to write  $\nabla \cdot \mathbf{A}$  for the divergence. The *surface integral* of  $\mathbf{A}$  over a closed surface  $\partial\Sigma$  is defined as:

$$\iint_{\partial\Sigma} \mathbf{A} \cdot d\mathbf{a} = \iint_{\partial\Sigma} \mathbf{A} \cdot \hat{\mathbf{n}} d\mathbf{a} \quad (45.1.36)$$

Where  $\hat{\mathbf{n}}$  is the unit normal to the surface  $\partial\Sigma$ . For a sphere, we have:

$$\hat{\mathbf{n}} = \frac{\operatorname{grad}(u)}{\|\operatorname{grad}(u)\|} = \frac{2x\hat{\mathbf{x}} + 2y\hat{\mathbf{y}} + 2z\hat{\mathbf{z}}}{\sqrt{4x^2 + 4y^2 + 4z^2}} = \frac{x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}}{\sqrt{x^2 + y^2 + z^2}} \quad (45.1.37)$$

Thus, we have:

$$\iint_{\partial\Sigma} \mathbf{r} \cdot \hat{\mathbf{n}} d\mathbf{a} = \iint_{\partial\Sigma} (x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}) \cdot \left( \frac{x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}}{\sqrt{x^2 + y^2 + z^2}} \right) d\mathbf{a} \quad (45.1.38a)$$

$$= \iint_{\partial\Sigma} \sqrt{x^2 + y^2 + z^2} d\mathbf{a} \quad (45.1.38b)$$

But recall that  $x^2 + y^2 + z^2 = a_0^2$ , so we have:

$$\oint\!\oint_{\partial\Sigma} \mathbf{r} \cdot d\mathbf{a} = a_0 \oint\!\oint_{\partial\Sigma} da \quad (45.1.39)$$

And this last integral is simply the surface area of  $\partial\Sigma$ . And the surface area of the sphere is  $4\pi a_0^2$ . So, we obtain:

$$\oint\!\oint_{\partial\Sigma} \mathbf{r} \cdot d\mathbf{a} = 4\pi a_0^3 \quad (45.1.40)$$

Using spherical coordinates is much easier.

$$\oint\!\oint_{\partial\Sigma} \mathbf{r} \cdot d\mathbf{a} = \int_0^{2\pi} \int_0^\pi a_0 \hat{\mathbf{r}} \cdot \hat{\mathbf{r}} a_0^2 \sin(\theta) d\theta d\varphi \quad (45.1.41a)$$

$$= \int_0^{2\pi} \int_0^\pi a_0^3 \sin(\theta) d\theta d\varphi \quad (45.1.41b)$$

$$= 2\pi a_0^3 \int_0^\pi \sin(\theta) d\theta \quad (45.1.41c)$$

$$= 4\pi a_0^3 \quad (45.1.41d)$$

To compute the *volume integral* within  $\Sigma$ , we compute  $\text{div}(\mathbf{r})$  and integrate:

$$\text{div}(\mathbf{r}) = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3 \quad (45.1.42a)$$

$$\iiint_{\Sigma} \text{div}(\mathbf{r}) d\tau = \iiint_{\Sigma} 3 d\tau = 3 \iiint_{\Sigma} d\tau = 3 \frac{4}{3}\pi a_0^3 = 4\pi a_0^3 \quad (45.1.42b)$$

**Problem 45.1.9** (Wangness 1-13) Given the vector field  $\mathbf{A} = xy\hat{\mathbf{x}} + yz\hat{\mathbf{y}} + xz\hat{\mathbf{z}}$ , evaluate the flux of  $\mathbf{A}$  through a parallelepiped of sides  $a, b, c$  shown in Fig. 45.2.1. Compute the volume integral of  $\text{div}(\mathbf{A})$ .

*Solution* There are six sides we must integrate over. We have:

$$\oint\!\oint_{\partial\Sigma} \mathbf{A} \cdot d\mathbf{a} = \oint\!\oint_{\partial\Sigma} (xy dy dz + yz dx dz + xz dx dy) \quad (45.1.43)$$

The  $dy dz$  integral is performed over the front and back faces. We obtain:

$$\begin{aligned} \iint_{\text{Front}} xy dy dz - \iint_{\text{Back}} xy dy dz &= \int_0^c \int_0^b ay dy dz - \int_0^c \int_0^b 0y dy dz \quad (45.1.44a) \\ &= \frac{ab^2 c}{2} \end{aligned}$$

$$(45.1.44b)$$

The  $dx dz$  integral is performed over the left and right faces. We compute:

$$\iint_{\text{Right}} yz \, dx \, dz - \iint_{\text{Left}} yz \, dx \, dz = \int_0^c \int_0^a bz \, dx \, dz - \int_0^c \int_0^a 0z \, dx \, dz \quad (45.1.45a)$$

$$= \frac{abc^2}{2} \quad (45.1.45b)$$

The final  $dx dz$  integral is performed over the top and bottom faces. We have:

$$\iint_{\text{Top}} xz \, dx \, dy - \iint_{\text{Bottom}} xz \, dx \, dy = \int_0^b \int_0^a cx \, dx \, dy - \int_0^b \int_0^a 0x \, dx \, dy \quad (45.1.46a)$$

$$= \frac{a^2bc}{2} \quad (45.1.46b)$$

Summing the results, we obtain:

$$\iint_{\partial\Sigma} \mathbf{A} \cdot \mathbf{d}\mathbf{a} = \frac{abc}{2}(a + b + c) \quad (45.1.47)$$

To compute the volume integral, we first compute  $\operatorname{div}(\mathbf{A})$ . We have:

$$\operatorname{div}(\mathbf{A}) = \frac{\partial(xy)}{\partial x} + \frac{\partial(yz)}{\partial y} + \frac{\partial(xz)}{\partial z} = x + y + z \quad (45.1.48)$$

Thus:

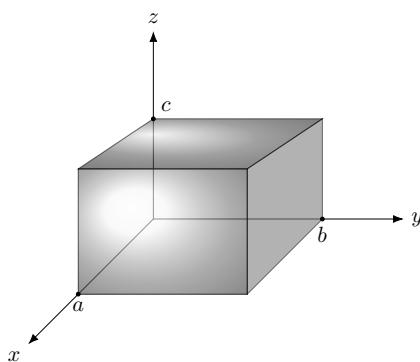
$$\iiint_{\Sigma} \operatorname{div}(\mathbf{A}) \, d\tau = \iiint_{\Sigma} (x + y + z) \, d\tau = \int_0^c \int_0^b \int_0^a (x + y + z) \, dx \, dy \, dz \quad (45.1.49)$$

Computing the integrals separately, we obtain:

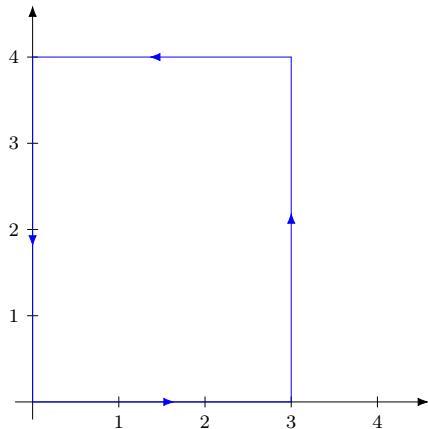
$$\begin{aligned} \iiint_{\Sigma} \operatorname{div}(\mathbf{A}) \, d\tau &= \int_0^c \int_0^b \int_0^a x \, dx \, dy \, dz + \int_0^c \int_0^b \int_0^a y \, dx \, dy \, dz \\ &\quad + \int_0^c \int_0^b \int_0^a z \, dx \, dy \, dz \end{aligned} \quad (45.1.50a)$$

$$= \frac{a^2bc}{2} + \frac{ab^2c}{2} + \frac{abc^2}{2} \quad (45.1.50b)$$

$$= \frac{abc}{2}(a + b + c) \quad (45.1.50c)$$



45.2.1: Wangsness 1-13



45.2.2: Wangsness 1-14

Fig. 45.2: Figures for problems 45.1.9 and 45.1.10, Respectively.

**Problem 45.1.10** (Wangsness 1-14) Given  $\mathbf{A} = -y\hat{\mathbf{x}} + x\hat{\mathbf{y}}$ , calculate the line integral over the closed path in the  $xy$  plane shown in Fig. 45.2.2. Compute the surface integral of  $\text{curl}(\mathbf{A})$ .

*Solution* Computing the line integral, we get:

$$\oint_{\partial S} \mathbf{A} \cdot d\ell = \oint_{\partial S} (-y\hat{\mathbf{x}} + x\hat{\mathbf{y}}) \cdot (dx\hat{\mathbf{x}} + dy\hat{\mathbf{y}}) \quad (45.1.51a)$$

$$= \underbrace{\int_0^3 (-y dx + x dy)}_{y=0, dy=0} + \underbrace{\int_0^4 (-y dx + x dy)}_{x=3, dx=0} \\ + \underbrace{\int_3^0 (-y dx + x dy)}_{y=4, dy=0} + \underbrace{\int_4^0 (-y dx + x dy)}_{x=0, dx=0} \quad (45.1.51b)$$

Piecing this together, we obtain:

$$\oint_{\partial S} \mathbf{A} \cdot d\ell = 0 + 12 + 12 + 0 = 24 \quad (45.1.52)$$

Next, we compute the area integral. We have  $\text{curl}(\mathbf{A}) = 2\hat{\mathbf{z}}$ . Thus:

$$\iint_S (\nabla \times \mathbf{A}) \cdot d\mathbf{a} = \iint_S (2\hat{\mathbf{z}}) \cdot (dy dz\hat{\mathbf{x}} + dx dz\hat{\mathbf{y}} + dx dy\hat{\mathbf{z}}) \quad (45.1.53a)$$

$$= \int_0^4 \int_0^3 2 dy dx \quad (45.1.53b)$$

$$= 24 \quad (45.1.53c)$$

**Problem 45.1.11** (Wangsness 1-15) Given the vector field  $\mathbf{A}$  defined by:

$$\mathbf{A} = x^2y\hat{\mathbf{x}} + xy^2\hat{\mathbf{y}} + a^3e^{-\beta y} \cos(\alpha x)\hat{\mathbf{z}} \quad (45.1.54)$$

compute the line integral along the path in Fig. 45.3. Compute the surface integral of  $\text{curl}(\mathbf{A})$  over the same region.

*Solution* Along the entire contour, we have  $z = 0$  and  $dz = 0$ . Thus, we have:

$$\oint_{\partial S} \mathbf{A} \cdot d\ell = \oint_{\partial S} (x^2y\hat{\mathbf{x}} + xy^2\hat{\mathbf{y}} + a^3e^{-\beta y} \cos(\alpha x)\hat{\mathbf{z}}) \cdot (dx\hat{\mathbf{x}} + dy\hat{\mathbf{y}}) \quad (45.1.55a)$$

$$= \oint_{\partial S} (x^2y \, dx + xy^2 \, dy) \quad (45.1.55b)$$

Along the first path we have  $x = 0$  and  $dx = 0$ . Along the second path, we have  $y = \sqrt{2k}$  and thus  $dy = 0$ . Along the third path we have  $y^2 = kx$ , and therefore  $dx = 2y \, dy/k$ . So:

$$\oint_{\partial S} \mathbf{A} \cdot d\ell = \int_{C_1} \mathbf{A} \cdot d\ell + \int_{C_2} \mathbf{A} \cdot d\ell + \int_{C_3} \mathbf{A} \cdot d\ell \quad (45.1.56a)$$

$$= \int_0^{\sqrt{2k}} 0y^2 \, dy + \int_0^2 x^2\sqrt{2k} \, dx + \int_{\sqrt{2k}}^0 \left( \frac{2y^6}{k^3} + \frac{y^4}{k} \right) \, dy \quad (45.1.56b)$$

$$= \frac{8}{3}\sqrt{2k} + \int_{\sqrt{2k}}^0 \left( \frac{2y^6}{k^3} + \frac{y^4}{k} \right) \, dy \quad (45.1.56c)$$

$$= \frac{8}{3}\sqrt{2k} - \frac{16}{7}\sqrt{2k} - \frac{4k\sqrt{2k}}{5} \quad (45.1.56d)$$

$$= \sqrt{2k} \left( \frac{8}{21} - \frac{4}{5}k \right) \quad (45.1.56e)$$

Next we compute the area Note that  $d\mathbf{a} = \hat{\mathbf{z}} \, dx \, dy$ . The  $\hat{\mathbf{z}}$  component for  $\text{curl}(\mathbf{A})$  is  $(y^2 - x^2)\hat{\mathbf{z}}$ . We have:

$$\iint_{\Sigma} (\nabla \times \mathbf{A}) \cdot d\mathbf{a} = \int_0^2 \int_{\sqrt{kx}}^{\sqrt{2k}} (x^2 - y^2) \, dy \, dx \quad (45.1.57a)$$

$$= \int_0^2 \left[ x^2y - \frac{y^3}{3} \right]_{\sqrt{kx}}^{\sqrt{2k}} \, dx \quad (45.1.57b)$$

$$= \int_0^2 \left[ x^2\sqrt{2k} - \frac{2k\sqrt{2k}}{3} - x^2\sqrt{kx} + \frac{kx\sqrt{kx}}{3} \right] \, dx \quad (45.1.57c)$$

Splitting the integral into four parts, we obtain:

$$\iint_{\Sigma} (\nabla \times \mathbf{A}) \cdot d\mathbf{a} = \sqrt{2k} \int_0^2 x^2 dx - \frac{2k}{3} \sqrt{2k} \int_0^2 dx - \sqrt{k} \int_0^2 x^{\frac{5}{2}} dx + \frac{k\sqrt{k}}{3} \int_0^2 x^{\frac{3}{2}} dx \quad (45.1.58)$$

$$= \frac{8}{3} \sqrt{2k} - \frac{4k}{3} \sqrt{2k} - \frac{16}{7} \sqrt{2k} + \frac{8k}{15} \sqrt{2k} = \frac{8}{21} \sqrt{2k} - \frac{4}{5} k \sqrt{2k} = \sqrt{2k} \left( \frac{8}{21} - \frac{4}{5} k \right) \quad (45.1.59)$$

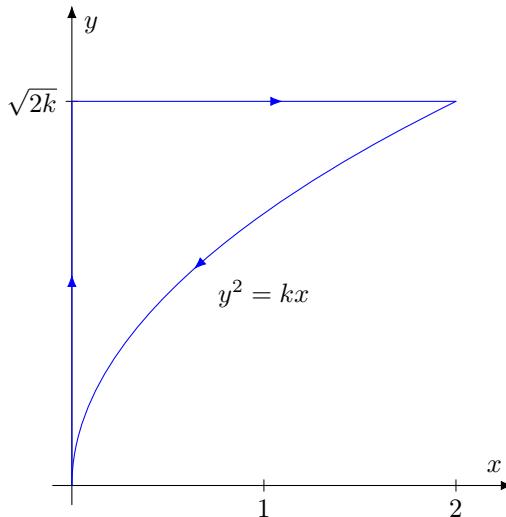


Fig. 45.3: Figure for problem 45.1.11

### 45.1.3 Homework III

Wangness Chapter 1 - Problems: 19, 20, 21, 22, 23, 24, 26

**Problem 45.1.12** (Wangness 1-19) Let  $\mathbf{A} = a\hat{\rho} + b\hat{\varphi} + c\hat{z}$ , where  $a, b, c$  are constants. Is  $\mathbf{A}$  a constant vector? Find  $\nabla \cdot \mathbf{A}$  and  $\nabla \times \mathbf{A}$ . Find the rectangular and spherical components of  $\mathbf{A}$ , expressing in terms of  $x, y, z$  and  $r, \theta, \varphi$ , respectively.

*Solution.* If  $a$  or  $b$  are non-zero, then  $\mathbf{A}$  is not a constant, for  $\hat{\rho}$  and  $\hat{\varphi}$  are non-constant functions of  $\varphi$ . To compute  $\nabla \cdot \mathbf{A}$ , we use  $\nabla$  in cylindrical coordinates and do:

$$\nabla \cdot \mathbf{A} = \frac{\partial A_\rho}{\partial \rho} + \frac{A_\rho}{\rho} + \frac{1}{\rho} \frac{\partial A_\phi}{\partial \varphi} + \frac{\partial A_z}{\partial z} = \frac{\partial a}{\partial \rho} + \frac{a}{\rho} + \frac{1}{\rho} \frac{\partial b}{\partial \varphi} + \frac{\partial c}{\partial z} = \frac{a}{\rho}$$

For  $\nabla \times \mathbf{A}$ , we again use cylindrical coordinates and do:

$$\begin{aligned}\nabla \times \mathbf{A} &= \begin{vmatrix} \frac{1}{\rho} \hat{\mathbf{p}} & \hat{\boldsymbol{\varphi}} & \frac{1}{\rho} \hat{\mathbf{z}} \\ \frac{\partial}{\partial \rho} & \frac{\partial}{\partial \varphi} & \frac{\partial}{\partial z} \\ A_\rho & \rho A_\varphi & A_z \end{vmatrix} = \frac{\hat{\mathbf{p}}}{\rho} \left( \frac{\partial A_z}{\partial \varphi} - \frac{\partial(\rho A_\varphi)}{\partial z} \right) + \hat{\boldsymbol{\varphi}} \left( \frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial \rho} \right) + \hat{\mathbf{z}} \left( \frac{1}{\rho} \frac{\partial(\rho A_\varphi)}{\partial \rho} - \frac{1}{\rho} \frac{\partial A_\rho}{\partial \varphi} \right) \\ &= \hat{\mathbf{p}} \left( \frac{1}{\rho} \frac{\partial A_z}{\partial \varphi} - \frac{\partial A_\varphi}{\partial z} \right) + \hat{\boldsymbol{\varphi}} \left( \frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial \rho} \right) + \hat{\mathbf{z}} \left( \frac{1}{\rho} \frac{\partial(\rho A_\varphi)}{\partial \rho} - \frac{1}{\rho} \frac{\partial A_\rho}{\partial \varphi} \right) \\ &= \hat{\mathbf{p}}(0 - 0) + \hat{\boldsymbol{\varphi}}(0 - 0) + \hat{\mathbf{z}} \left( \frac{A_\phi}{\rho} + 0 - 0 \right) = \frac{b}{\rho} \hat{\mathbf{z}}\end{aligned}$$

Rectangular coordinates of  $\mathbf{A}$ :

$$\begin{aligned}\mathbf{A} &= a(\cos(\phi)\hat{\mathbf{x}} + \sin(\phi)\hat{\mathbf{y}}) + b(-\sin(\phi)\hat{\mathbf{x}} + \cos(\phi)\hat{\mathbf{y}}) + c\hat{\mathbf{z}} \\ &= a \left( \frac{x\hat{\mathbf{x}} + y\hat{\mathbf{y}}}{\sqrt{x^2 + y^2}} \right) + b \left( \frac{-y\hat{\mathbf{x}} + x\hat{\mathbf{y}}}{\sqrt{x^2 + y^2}} \right) + c\hat{\mathbf{z}} = \frac{ax - by}{\sqrt{x^2 + y^2}} \hat{\mathbf{x}} + \frac{ay + bx}{\sqrt{x^2 + y^2}} \hat{\mathbf{y}} + c\hat{\mathbf{z}}\end{aligned}$$

For spherical coordinates:

$$\begin{aligned}\hat{\mathbf{p}} &= \sin(\theta)\hat{\mathbf{r}} + \cos(\theta)\hat{\boldsymbol{\theta}} & \mathbf{A} &= a(\sin(\theta)\hat{\mathbf{r}} + \cos(\theta)\hat{\boldsymbol{\theta}}) + b\hat{\boldsymbol{\varphi}} + c(\cos(\theta)\hat{\mathbf{r}} - \sin(\theta)\hat{\boldsymbol{\theta}}) \\ \hat{\mathbf{z}} &= \cos(\theta)\hat{\mathbf{r}} - \cos(\theta)\hat{\boldsymbol{\theta}} & &= (a \sin(\theta) + c \cos(\theta))\hat{\mathbf{r}} + b\hat{\boldsymbol{\varphi}} + (a \cos(\theta) - c \sin(\theta))\hat{\boldsymbol{\theta}}\end{aligned}$$

□

**Problem 45.1.13** (Wangness 1-20) Let  $\mathbf{A} = a\hat{\mathbf{r}} + b\hat{\boldsymbol{\theta}} + c\hat{\boldsymbol{\varphi}}$ . Is  $\mathbf{A}$  a constant vector? Find  $\nabla \cdot \mathbf{A}$  and  $\nabla \times \mathbf{A}$ . Find the rectangular and cylindrical components of  $\mathbf{A}$ , expressing in terms of  $x, y, z$  and  $\rho, \varphi, z$ , respectively.

*Solution.* If  $a$  or  $b$  or  $c$  are non-zero, then  $\mathbf{A}$  is not a constant vector, for  $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}$ , and  $\hat{\boldsymbol{\varphi}}$  are non-constant functions of  $r, \theta, \varphi$ . To compute  $\nabla \cdot \mathbf{A}$ , we use spherical coordinates and do:

$$\nabla \cdot \mathbf{A} = \frac{1}{r^2} \frac{\partial(r^2 A_r)}{\partial r} + \frac{1}{r \sin(\theta)} \frac{\partial(\sin(\theta) A_\theta)}{\partial \theta} + \frac{1}{r \sin(\theta)} \frac{\partial A_\varphi}{\partial \varphi} = \frac{2a}{r} + \frac{b \cos(\theta)}{r \sin(\theta)}$$

For  $\nabla \times \mathbf{A}$ :

$$\begin{aligned}\nabla \times \mathbf{A} &= \begin{vmatrix} \frac{1}{r^2 \sin(\theta)} \hat{\mathbf{r}} & \frac{1}{r \sin(\theta)} \hat{\boldsymbol{\theta}} & \frac{1}{r} \hat{\boldsymbol{\varphi}} \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \varphi} \\ A_r & r A_\theta & r \sin(\theta) A_\varphi \end{vmatrix} \\ &= \frac{\hat{\mathbf{r}}}{r \sin(\theta)} \left( \frac{\partial(\sin(\theta) A_\varphi)}{\partial \theta} - \frac{\partial A_\theta}{\partial \varphi} \right) + \frac{\hat{\boldsymbol{\theta}}}{r} \left( \frac{1}{\sin(\theta)} \frac{\partial A_r}{\partial \varphi} - \frac{\partial(r A_\varphi)}{\partial r} \right) + \frac{\hat{\boldsymbol{\varphi}}}{r} \left( \frac{\partial(r A_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) \\ &= \frac{\cos(\theta)}{r \sin(\theta)} \hat{\mathbf{r}} - \frac{c}{r} \hat{\boldsymbol{\theta}} + \frac{b}{r} \hat{\boldsymbol{\varphi}}\end{aligned}$$

In rectangular coordinates, we have:

$$\begin{aligned}
 \mathbf{A} &= a(\sin(\theta) \cos(\varphi) \hat{\mathbf{x}} + \sin(\theta) \sin(\varphi) \hat{\mathbf{y}} + \cos(\theta) \hat{\mathbf{z}}) + \\
 &\quad b(\cos(\theta) \cos(\varphi) \hat{\mathbf{x}} + \cos(\theta) \sin(\varphi) \hat{\mathbf{y}} - \sin(\theta) \hat{\mathbf{z}}) + \\
 &\quad c(-\sin(\varphi) \hat{\mathbf{x}} + \cos(\varphi) \hat{\mathbf{y}}) \\
 &= \frac{a}{\sqrt{x^2 + y^2 + z^2}} \left( x \hat{\mathbf{x}} + y \hat{\mathbf{y}} + z \hat{\mathbf{z}} \right) + \\
 &\quad \frac{b}{\sqrt{x^2 + y^2 + z^2}} \left( \frac{xz}{\sqrt{x^2 + y^2}} \hat{\mathbf{x}} + \frac{yz}{\sqrt{x^2 + y^2}} \hat{\mathbf{y}} - \sqrt{x^2 + y^2} \hat{\mathbf{z}} \right) + \\
 &\quad - \frac{y}{\sqrt{x^2 + y^2}} \hat{\mathbf{x}} + \frac{x}{\sqrt{x^2 + y^2}} \hat{\mathbf{y}} \\
 &= \left( \frac{ax}{\sqrt{x^2 + y^2 + z^2}} + \frac{bxz}{\sqrt{x^2 + y^2} \sqrt{x^2 + y^2 + z^2}} - \frac{y}{\sqrt{x^2 + y^2}} \right) \hat{\mathbf{x}} + \\
 &\quad \left( \frac{ay}{\sqrt{x^2 + y^2 + z^2}} + \frac{byz}{\sqrt{x^2 + y^2} \sqrt{x^2 + y^2 + z^2}} + \frac{x}{\sqrt{x^2 + y^2}} \right) \hat{\mathbf{y}} + \\
 &\quad \left( \frac{az}{\sqrt{x^2 + y^2 + z^2}} - \frac{b\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2 + z^2}} \right) \hat{\mathbf{z}}
 \end{aligned}$$

We can then use this to convert to cylindrical, recalling that  $r^2 = \rho^2 + z^2$ :

$$\mathbf{A} = \left( \frac{a\rho + bz}{\sqrt{\rho^2 + z^2}} \right) \hat{\mathbf{p}} + c\hat{\Phi} + \left( \frac{az - b\rho}{\sqrt{\rho^2 + z^2}} \right) \hat{\mathbf{z}}$$

□

**Problem 45.1.14** (Wangsness 1-21) Find  $\nabla \cdot \mathbf{r}$  for the position vector  $\mathbf{r}$  expressed in rectangular, cylindrical, and spherical coordinates.

*Solution.* In rectangular coordinates we have  $\mathbf{r} = x \hat{\mathbf{x}} + y \hat{\mathbf{y}} + z \hat{\mathbf{z}}$ . So:

$$\nabla \cdot \mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 1 + 1 + 1 = 3$$

In cylindrical coordinates,  $\mathbf{r} = \rho \hat{\mathbf{p}} + z \hat{\mathbf{z}}$ , So:

$$\nabla \cdot \mathbf{r} = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho^2) + \frac{1}{\rho} \frac{\partial}{\partial \phi} (0) + \frac{\partial z}{\partial z} = 2 + 0 + 1 = 3$$

In spherical coordinates we have:

$$\nabla \cdot \mathbf{r} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 r) = 3$$

□

**Problem 45.1.15** (Wangsness 1-22) Let  $\mathbf{A} = a\rho\hat{\rho} + b\hat{\phi} + cz\hat{z}$ , where  $a$ ,  $b$ , and  $c$  are constants. Find  $\oint \mathbf{A} \cdot d\mathbf{a}$  over the surface of a right circular cylinder of length  $L$  and radius  $\rho_0$  whose axis is along the positive  $z$  axis and the origin is the center of the lower circular face (See Fig. 45.4.1). Find  $\iiint \nabla \cdot \mathbf{A} d\tau$  over the volume of the cylinder.

*Solution.* We have that:

$$\oint \mathbf{A} \cdot d\mathbf{a} = \int_{Top} \mathbf{A} \cdot d\mathbf{a} + \int_{Cylinder} \mathbf{A} \cdot d\mathbf{a} + \int_{Bottom} \mathbf{A} \cdot d\mathbf{a}$$

On the cylindrical surface,  $d\mathbf{a} = \rho_0 d\phi dz$ , so the integral is:

$$\int_0^L \int_0^{2\pi} (a\rho_0\hat{\rho} + b\hat{\phi} + cz\hat{z}) \cdot \hat{\rho}\rho_0 d\phi dz = \int_0^L \int_0^{2\pi} a\rho_0^2 d\phi dz = a\rho_0^2(2\pi)L$$

For the top and bottom,  $d\mathbf{a} = \pm \rho d\rho d\phi \hat{z}$ , respectively. On the bottom surface  $z = 0$  and thus the integral is zero. On the top we get:

$$\int_0^{2\pi} \int_0^{\rho_0} cL\rho d\rho d\phi = \pi c L \rho_0^2$$

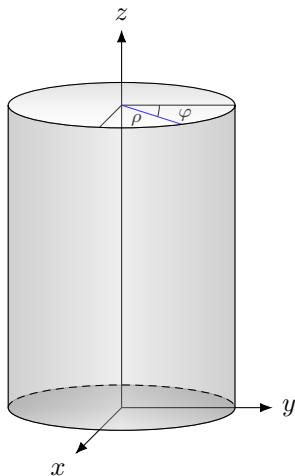
Thus:

$$\oint \mathbf{A} \cdot d\mathbf{a} = \pi L \rho_0^2 (2a + c)$$

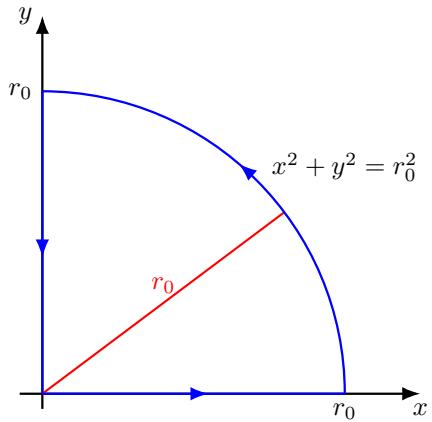
Computing the divergence, we get  $\nabla \cdot \mathbf{A} = 2a + c$ . Therefore:

$$\iiint_C \nabla \cdot \mathbf{A} d\tau = (2a + c)V = \pi L \rho_0^2 (2a + c)$$

□



45.4.1: Wangsness 1-22



45.4.2: Wangsness 1-23

Fig. 45.4: Figures for problems 45.1.15 and 45.1.16, respectively.

**Problem 45.1.16** (Wangsness 1-23) Let  $\mathbf{A} = 4\hat{\mathbf{r}} + 3\hat{\theta} - 2\hat{\varphi}$ . Find the line integral around the closed path shown in Fig. 45.4.2. Find the surface integral of  $\nabla \times \mathbf{A}$  over the enclosed area.

*Solution.* We have:

$$\oint_{\partial S} \mathbf{A} \cdot d\ell = \sum_i \int_{\partial S_i} \mathbf{A} \cdot d\ell$$

Along the first path  $\varphi = 0$ ,  $\theta = \frac{\pi}{2}$ , and  $d\ell = dr\hat{\mathbf{r}}$ . The integral is then  $4r_0$ .

$$\int_{\partial S_1} \mathbf{A} \cdot d\ell = \int_0^{r_0} (4\hat{\mathbf{r}} + 3\hat{\theta} - 2\hat{\varphi}) \cdot (\hat{\mathbf{r}} dr) = 4 \int_0^{r_0} dr = 4r_0$$

Along the second path  $r = r_0$ ,  $\theta = \frac{\pi}{2}$ , and  $d\ell = r_0 d\varphi \hat{\varphi}$ .

$$\int_{\partial S_2} \mathbf{A} \cdot d\ell = \int_0^{\frac{\pi}{2}} (4\hat{\mathbf{r}} + 3\hat{\theta} - 2\hat{\varphi}) \cdot (r_0 d\varphi \hat{\varphi}) = -2 \int_0^{\frac{\pi}{2}} r_0 d\varphi = -\pi r_0$$

Along the final path,  $\varphi = \frac{\pi}{2}$ ,  $\theta = \frac{\pi}{2}$ , and  $d\ell = dr\hat{\mathbf{r}}$ .

$$\int_{\partial S_3} \mathbf{A} \cdot d\ell = \int_{r_0}^0 (4\hat{\mathbf{r}} + 3\hat{\theta} - 2\hat{\varphi}) \cdot (\hat{\mathbf{r}} dr) = 4 \int_{r_0}^0 dr = -4r_0$$

Therefore:

$$\oint_{\partial S} \mathbf{A} \cdot d\ell = \sum_i \int_{\partial S_i} \mathbf{A} \cdot d\ell = 4r_0 - \pi r_0 - 4r_0 = -\pi r_0$$

The curl is:  $\nabla \times \mathbf{A} = \frac{-2 \cot(\theta)}{r} \hat{\mathbf{r}} + \frac{2}{r} \hat{\theta} + \frac{3}{r} \hat{\phi}$ . For the plane,  $d\mathbf{a} = \hat{\mathbf{z}} r dr d\varphi$ ,  $\theta = \frac{\pi}{2}$ . So:

$$\iint \nabla \times \mathbf{A} \cdot d\mathbf{a} = \int_0^{\frac{\pi}{2}} \int_0^{r_0} \left( \frac{-2 \cot(\theta)}{r} \hat{\mathbf{r}} + \frac{2}{r} \hat{\theta} + \frac{3}{r} \hat{\phi} \right) \cdot \hat{\mathbf{z}} r dr d\varphi$$

$$= \int_0^{\frac{\pi}{2}} \int_0^{r_0} \left( -2 \cot(\theta) \cos(\theta) - 2 \sin(\theta) \right) dr d\varphi = -2 \int_0^{\frac{\pi}{2}} \int_0^{r_0} dr d\varphi = -$$

□

**Problem 45.1.17** (Wangsness 1-24) Verify that  $\nabla \times (u\mathbf{A}) = \nabla(u) \times \mathbf{A} + u(\nabla \times \mathbf{A})$

*Solution.* Let  $\mathbf{A} = \langle A_x, A_y, A_z \rangle$  and  $u = u(x, y, z)$ . Using the product rule, we get:

$$\begin{aligned} \nabla \times (u\mathbf{A}) &= \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ uA_x & uA_y & uA_z \end{vmatrix} \\ &= \hat{\mathbf{x}} \left( \frac{\partial u}{\partial y} A_z + u \frac{\partial A_z}{\partial y} - \frac{\partial u}{\partial z} A_y - u \frac{\partial A_y}{\partial z} \right) + \\ &\quad \hat{\mathbf{y}} \left( \frac{\partial u}{\partial z} A_x + u \frac{\partial A_x}{\partial z} - \frac{\partial u}{\partial x} A_z - u \frac{\partial A_z}{\partial x} \right) + \\ &\quad \hat{\mathbf{z}} \left( \frac{\partial u}{\partial x} A_y + u \frac{\partial A_y}{\partial x} - \frac{\partial u}{\partial y} A_x - u \frac{\partial A_x}{\partial y} \right) \end{aligned}$$

But:

$$\nabla(u) \times \mathbf{A} = \hat{\mathbf{x}} \left( \frac{\partial u}{\partial y} A_z - \frac{\partial u}{\partial z} A_y \right) + \hat{\mathbf{y}} \left( \frac{\partial u}{\partial z} A_x - \frac{\partial u}{\partial x} A_z \right) + \hat{\mathbf{z}} \left( \frac{\partial u}{\partial x} A_y - \frac{\partial u}{\partial y} A_x \right)$$

and

$$u(\nabla \times \mathbf{A}) = \hat{\mathbf{x}} \left( u \frac{\partial A_z}{\partial y} - u \frac{\partial A_y}{\partial z} \right) + \hat{\mathbf{y}} \left( u \frac{\partial A_x}{\partial z} - u \frac{\partial A_z}{\partial x} \right) + \hat{\mathbf{z}} \left( u \frac{\partial A_y}{\partial x} - u \frac{\partial A_x}{\partial y} \right)$$

Summing these, we have  $\nabla \times (u\mathbf{A}) = \nabla(u) \times \mathbf{A} + u\nabla \times \mathbf{A}$  □

**Problem 45.1.18** (Wangsness 1-26) Verify that  $\oint_S u d\mathbf{a} = \int_V \nabla(u) d\tau$  and  $\oint_S \mathbf{A} \times d\mathbf{a} = - \int_V \nabla \times \mathbf{A} d\tau$

*Solution.* Let  $\mathbf{C}$  be an arbitrary constant vector. Then:

$$\mathbf{C} \cdot \left( \iint_S u d\mathbf{a} - \iiint_V \nabla(u) d\tau \right) = \iint_S u \left( \mathbf{C} \cdot d\mathbf{a} \right) - \iiint_V \left( \mathbf{C} \cdot \nabla(u) \right) d\tau$$

But, as  $\mathbf{C}$  is constant,  $\nabla \cdot \mathbf{C} = 0$ , and thus we have:

$$\begin{aligned}\iiint_V \mathbf{C} \cdot \nabla(u) d\tau &= \iiint_V \nabla \cdot (u\mathbf{C}) d\tau - \iiint_V u \nabla \cdot \mathbf{C} d\tau \\ &= \iiint_V \nabla(u\mathbf{C}) d\tau\end{aligned}$$

But from the divergence theorem:

$$\iiint_V \nabla(u\mathbf{C}) d\tau = \oint_S u\mathbf{C} \cdot d\mathbf{a}$$

Thus,  $\mathbf{C} \cdot (\oint_S u\mathbf{d}\mathbf{a} - \int_V \nabla(u) d\tau) = 0$ . As  $\mathbf{C}$  is any arbitrary vector,  $\oint_S u\mathbf{d}\mathbf{a} - \int_V \nabla(u) d\tau = 0$  and thus  $\oint_S u\mathbf{d}\mathbf{a} = \int_V \nabla(u) d\tau$ . It is a simple exercise in vector geometry to show that if  $\mathbf{A} \cdot \mathbf{C} = 0$  for all vectors  $\mathbf{C}$ , then  $\mathbf{A} = \mathbf{0}$ . In an analogous manner:

$$\mathbf{C} \cdot \left( \oint_S \mathbf{A} \times d\mathbf{a} + \iiint_V \nabla \times \mathbf{A} d\tau \right) = 0 \Rightarrow \oint_S \mathbf{A} \times d\mathbf{a} = - \iiint_V \nabla \times \mathbf{A} d\tau$$

□

#### 45.1.4 Homework IV

Wangsness Chapter 2: Problems 3, 7, 8

Wangsness Chapter 3: Problems 9, 10

**Problem 45.1.19** (Wangsness 2-3) Consider 8 equal point charges  $q$  located on the corners of a cube of length  $a$ , as in Fig. 45.5.1. Find the total force on the charge at the origin.

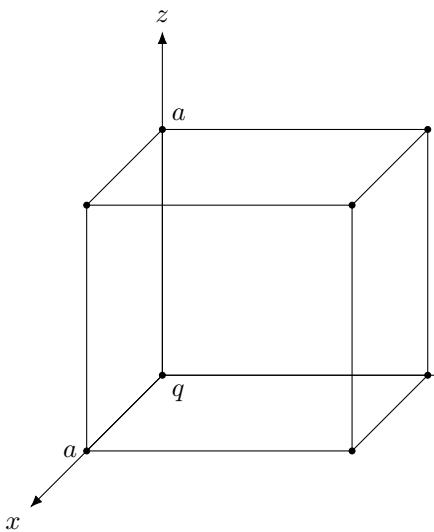
*Solution.* The force on a charged particle  $q$  at a point  $\mathbf{R}$  exerted by a distribution of charges  $q_i$  at points  $\mathbf{R}_i$  is defined by:

$$\mathbf{F}_q = \sum_{i=1}^N \frac{qq_i \hat{\mathbf{r}}_i}{4\pi\epsilon_0 \|\mathbf{R} - \mathbf{R}_i\|}$$

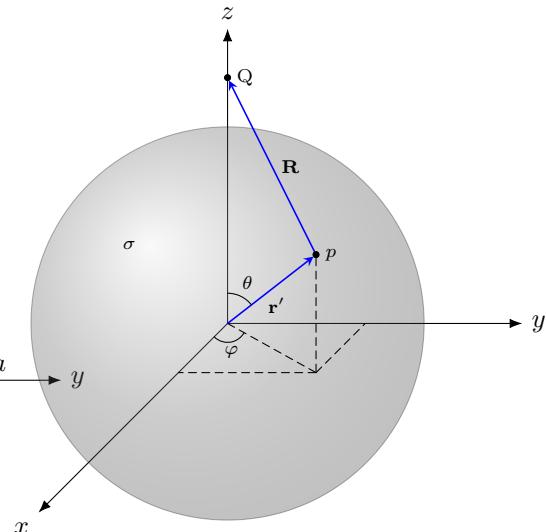
Where  $\hat{\mathbf{r}}_i$  is the unit vector pointing from  $\mathbf{R}$  to  $\mathbf{R}_i$ . The charge is located at the origin, so we have:

$$\mathbf{F}_q = \sum_{i=1}^N \frac{qq'_i}{4\pi\epsilon_0 R_i^2} \hat{\mathbf{r}}_i = -\frac{q^2}{4\pi\epsilon_0 a^2} \left(1 + \frac{2}{2^{3/2}} + \frac{1}{3^{3/2}}\right) (\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}) \approx -1.9 \frac{q^2}{4\pi\epsilon_0 a^2} (\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})$$

□



45.5.1: Drawing for Wangsness 2-3



45.5.2: Drawing for Wangsness 2-8

Fig. 45.5: Figures for problems 45.1.19 and 45.1.21

**Problem 45.1.20** (Wangsness 2-7) Given a line charge of length  $L$  with constant charge density lying along the positive  $z$  axis with its ends located at  $z_0$  and  $z_0 + L$ , find the total force exerted on this by a uniform spherical charge distribution with center at the origin and radius  $a < z_0$ .

*Solution.* If the charge is distributed over a length  $L$  along the  $z$ -axis with charge per unit length  $\lambda$ , the force over the length  $L$  is given by:

$$\mathbf{F}_{Lz} = \frac{\rho a^3}{3\epsilon_0} \hat{\mathbf{z}} \int_{z_0}^{z_0+L} \frac{\lambda dz}{z^2} = \frac{\rho \lambda a^3}{3\epsilon_0} \left( \frac{L}{z_0(z_0+L)} \right) \hat{\mathbf{z}}$$

□

**Problem 45.1.21** (Wangsness 2-8) Consider the sphere in Fig. 45.5.2 of radius  $a$  with a constant surface charge density  $\sigma$ . What is the total charge  $Q'$  on the sphere? Find the force produced by this charge distribution on a point  $q$  on the  $z$  axis for both  $z > a$  and  $z < a$ .

*Solution.* We have that:

$$Q' = \iint_S \sigma da = \sigma \iint_S da = \sigma 4\pi a^2$$

The relative position vector  $\mathbf{R}$  of the point  $Q$  with respect to a point  $\mathbf{r}'$  on the sphere is  $z\hat{\mathbf{z}} - a\hat{\mathbf{r}}'$ . So:

$$\mathbf{F}_q = \frac{q}{4\pi\epsilon_0} \iint_S \frac{\sigma da' \mathbf{R}}{R^3} = \frac{q\sigma}{4\pi\epsilon_0} \int_0^{2\pi} \int_0^\pi \frac{(z\hat{\mathbf{z}} - a\hat{\mathbf{r}}') a^2 \sin(\theta') d\theta' d\varphi'}{(z^2 + a^2 - 2az \cos(\theta'))^{3/2}}$$

Writing  $\hat{\mathbf{r}}' = \sin(\theta') \cos(\varphi') \hat{\mathbf{x}} + \sin(\theta') \sin(\varphi') \hat{\mathbf{y}} + \cos(\theta') \hat{\mathbf{z}}$  leads us to conclude the  $x$  and  $y$  component vanish as  $\int_0^{2\pi} \cos(\varphi') d\varphi' = \int_0^{2\pi} \sin(\varphi') d\varphi' = 0$ . Thus, we have:

$$\mathbf{F}_q = \frac{q\sigma\hat{\mathbf{z}}}{4\pi\epsilon_0} \int_0^{2\pi} \int_0^\pi \frac{(z - a \cos(\theta)) a^2 \sin(\theta')}{(z^2 + a^2 - 2az \cos(\theta'))^{3/2}} d\theta' d\varphi'$$

Let  $u = \cos(\theta')$ . Then  $du = -\sin(\theta')d\theta'$ . We obtain:

$$\begin{aligned} \mathbf{F}_q &= a^2 \frac{q\sigma\hat{\mathbf{z}}}{2\epsilon_0} \int_{-1}^1 \frac{u}{(z^2 + a^2 - 2azu)^{3/2}} du &= a^2 \frac{q\sigma\hat{\mathbf{z}}}{2\epsilon_0} \frac{\partial}{\partial z} \left[ \frac{1}{za} \sqrt{a^2 + z^2 - 2azu} \right]_{-1} \\ &= a^2 \frac{q\sigma\hat{\mathbf{z}}}{2\epsilon_0} \int_{-1}^1 \frac{\partial}{\partial z} \left( \frac{1}{\sqrt{a^2 + z^2 - 2azu}} \right) du &= a^2 \frac{q\sigma\hat{\mathbf{z}}}{2\epsilon_0} \frac{\partial}{\partial z} \left( \frac{|z - a| - |z + a|}{az} \right) \\ &= a^2 \frac{q\sigma\hat{\mathbf{z}}}{2\epsilon_0} \frac{\partial}{\partial z} \int_{-1}^1 \frac{1}{\sqrt{a^2 + z^2 - 2azu}} du \end{aligned}$$

Now, if  $z > a$ , then  $|z - a| - |z + a| = (z - a) - (z + a) = -2a$ , and thus:

$$\mathbf{F}_q = a^2 \frac{q\sigma\hat{\mathbf{z}}}{2\epsilon_0} \frac{\partial}{\partial z} \left( \frac{-2a}{az} \right) = a^2 \frac{q\sigma}{\epsilon_0 z^2} \hat{\mathbf{z}} = \frac{qQ}{4\pi\epsilon_0 z^2} \hat{\mathbf{z}} \quad (z > a)$$

If  $z < a$ , then  $|z - a| - |z + a| = 2z$ , and thus:

$$\mathbf{F} = \mathbf{F}_q = a^2 \frac{q\sigma\hat{\mathbf{z}}}{2\epsilon_0} \frac{\partial}{\partial z} \left( \frac{2z}{az} \right) = \mathbf{0} \quad (z < a)$$

□

**Problem 45.1.22** (Wangsness 3-9) Given two infinite plane sheets with equal and opposite constant surface charge densities  $\sigma$  that are parallel and a distance  $\pm a$  to the  $xy$  plane, find  $\mathbf{E}$  everywhere.

*Solution.* For an infinite sheet,  $\mathbf{E} = \frac{\sigma}{2\epsilon_0}$ . Using the principle of superposition, we get:

$$\mathbf{E} = \begin{cases} 0, & |z| > a \\ -\frac{\sigma}{\epsilon_0} \hat{\mathbf{z}}, & |z| \leq a \end{cases}$$

□

**Problem 45.1.23** (Wangsness 3-10) A circular arc of radius  $a$  with arc angle  $2\alpha$  that lies in the  $xy$  plane and has a constant linear charge density  $\lambda$  and center of curvature at the origin. Find  $\mathbf{E}$  at an arbitrary point on the  $z$  axis. Show that when the arc becomes a complete circle you obtain  $\mathbf{E} = \frac{\lambda az\hat{\mathbf{z}}}{2\epsilon_0(a^2+z^2)^{3/2}}$

*Solution.* We have that:

$$\mathbf{E} = \frac{\lambda}{4\pi\epsilon_0} \int \frac{dl\hat{\mathbf{r}}}{R^2} = \frac{\lambda}{4\pi\epsilon_0} \int \frac{a d\varphi' \hat{\mathbf{r}}}{a^2 + z^2}$$

And

$$\hat{\mathbf{r}} = \frac{\mathbf{R}}{R} = \frac{-\rho' \hat{\rho} + z \hat{\mathbf{z}}}{\sqrt{a^2 + z^2}} = \frac{-a \cos(\phi') \hat{x} - a \sin(\phi') \hat{y} + z \hat{\mathbf{z}}}{\sqrt{a^2 + z^2}}$$

So, we obtain the following:

$$\mathbf{E} = \frac{\lambda}{4\pi\epsilon_0} \int_{-\alpha}^{\alpha} \frac{-a \cos(\phi') \hat{x} - a \sin(\phi') \hat{y} + z \hat{\mathbf{z}}}{(a^2 + z^2)^{3/2}} a d\varphi' = \frac{\lambda a [-a \sin(\alpha) \hat{x} + z \alpha \hat{\mathbf{z}}]}{2\pi\epsilon_0 (a^2 + z^2)^{3/2}}$$

If  $\alpha = \pi$ , we get:

$$\mathbf{E} = \frac{\lambda az}{2\epsilon_0 (a^2 + z^2)^{3/2}} \hat{\mathbf{z}}$$

□

### 45.1.5 Homework V

Wangsness Chapter 4 - Problems: 3, 5, 6, 7, 11, 12

**Problem 45.1.24** (Wangsness 4-3) An infinitely long line is surrounded by an infinitely long cylinder of radius  $\rho_0$  whose axis coincides with a line of charge (See Fig. 45.6). The surface of the cylinder carries a charge of constant surface density  $\sigma$ . Find  $\mathbf{E}$  everywhere. What particular value of  $\sigma$  will make  $\mathbf{E} = \mathbf{0}$  for all points outside of the charged cylinder?

*Solution.* For  $\rho < \rho_0$  choose a Gaussian cylinder concentric with the line. From Gauss' Law we have:

$$\oint \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{in}}{\epsilon_0} = E(2\pi\rho\ell) \Rightarrow \mathbf{E} = \frac{\lambda}{2\pi\epsilon_0\rho} \hat{\rho}$$

Where  $\lambda$  is the linear charge density. For  $\rho > \rho_0$  choose a similar Gaussian cylinder. We get:

$$\oint \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{in}}{\epsilon_0} \Rightarrow \mathbf{E} = \frac{\lambda\ell + \sigma 2\pi\rho_0\ell}{2\pi\epsilon_0\rho\ell} \hat{\rho} = \frac{\lambda + 2\pi\rho_0\sigma}{2\pi\epsilon_0\rho} \hat{\rho}$$

To make  $\hat{\mathbf{E}} = \mathbf{0}$  for  $\rho > \rho_0$  we need  $\sigma = -\frac{\lambda}{2\pi\rho_0}$ .

□

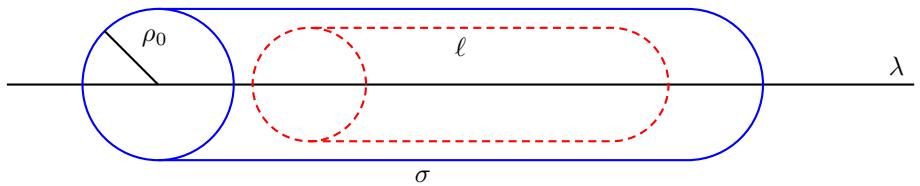


Fig. 45.6: Drawing for problem 45.1.24

**Problem 45.1.25** (Wangsness 4-5) A sphere of radius  $a$  has a charge density that varies with distance  $r$  from the center according to  $\rho = Ar^{1/2}$ , where  $A$  is a constant. Find  $\mathbf{E}$  everywhere.

*Solution.* For  $r > a$ , choose a Gaussian sphere concentric with the origin one. We have:

$$\oint \oint \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{in}}{\epsilon_0} \Rightarrow E(4\pi r^2) = \iiint Ar'^{1/2}r'^2 \sin(\theta') dr' d\theta' d\varphi' = 4\pi \frac{2}{7} \frac{a^{7/2}}{\epsilon_0} A \Rightarrow \mathbf{E} = \frac{2Aa^{7/2}}{7\epsilon_0} \hat{\mathbf{r}}$$

For  $r < a$ :

$$E(4\pi r^2) = 4\pi A \frac{2}{7} r^{7/2} \Rightarrow \mathbf{E} = \frac{2Ar^{3/2}}{7\epsilon_0} \hat{\mathbf{r}}$$

□

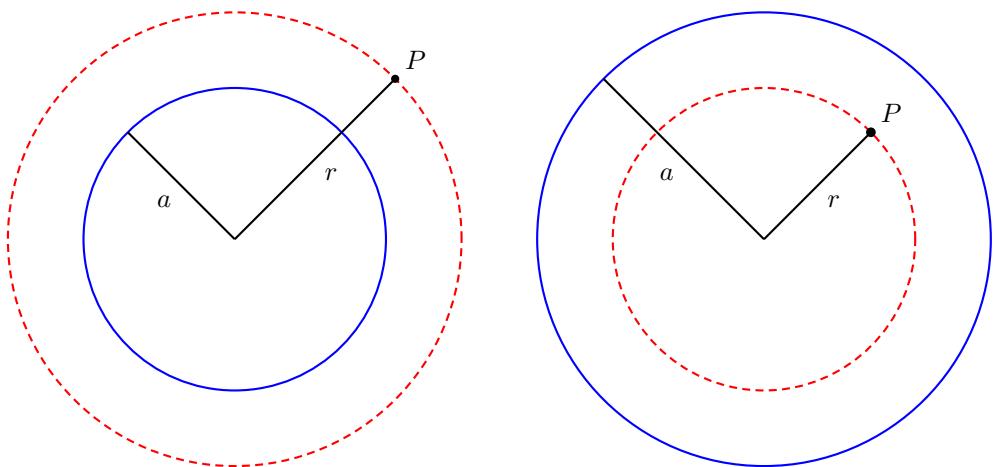


Fig. 45.7: Drawing for problem 45.1.25. Dashed red represents the Gaussian surface, and blue represents the charged sphere.

**Problem 45.1.26** (Wangsness 4-6) Two concentric spheres have radii  $a$  and  $b$  with  $b > a$ . The region between them ( $a \leq r \leq b$ ) is filled with a charge of constant density. The charge density is zero everywhere else. Find  $\mathbf{E}$  everywhere and express it in terms of the total charge  $Q$ . What happens as  $a \rightarrow 0$ ?

*Solution.* From the symmetry of the problem we have, for  $r < a$ , that:

$$\oint\oint \mathbf{E} \cdot d\mathbf{a} = 0 \Rightarrow \mathbf{E} = \mathbf{0}$$

For  $a \leq r \leq b$ :

$$\oint\oint \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{in}}{\epsilon_0} = \frac{\rho_{ch}(\frac{4}{3}\pi)(r^3 - a^3)}{\epsilon_0}$$

Where  $\rho_{ch}$  is the charge density  $\rho_{ch} = \frac{Q}{\frac{4}{3}\pi(b^3 - a^3)}$ . Therefore:

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0 r^2} \left( \frac{r^3 - a^3}{b^3 - a^3} \right) \hat{\mathbf{r}}$$

For  $r \geq b$ , this is similar to the uniform sphere problem:

$$\oint\oint \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{in}}{\epsilon_0} \Rightarrow \mathbf{E} = \frac{Q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}$$

In the limit as  $a \rightarrow 0$  we have:

$$\mathbf{E} = \begin{cases} \frac{Qr}{4\pi\epsilon_0 b^3}, & r \leq b \\ \frac{Q}{4\pi\epsilon_0 r^2}, & r > b \end{cases}$$

This is expected, for in the limit as  $a \rightarrow 0$  we obtain a uniformly charged sphere of radius  $b$ .  $\square$

**Problem 45.1.27** (Wangsness 4-7) Given an infinite cylinder whose circular cross section area has radius  $a$ , and with constant surface charge density  $\sigma_{ch}$ , find  $\mathbf{E}$  everywhere.

*Solution.* Choose as a Gaussian surface a cylinder of length  $L$  and radius  $\rho$  that is concentric with the infinite cylinder. For  $\rho < a$  we have:

$$\oint\oint \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{in}}{\epsilon_0} = \oint\oint \mathbf{E} \cdot d\mathbf{a} = \iint_{\text{Top}} \mathbf{E} \cdot d\mathbf{a} + \iint_{\text{Bottom}} \mathbf{E} \cdot d\mathbf{a} + \iint_{\text{Cylinder}} \mathbf{E} \cdot d\mathbf{a}$$

From symmetry, we have that  $\mathbf{E}$  and  $d\mathbf{a}$  are orthogonal along the top and bottom faces of the cylinder, leaving only the cylindrical body left to integrate over. We have:

$$E(2\pi\rho L) = \frac{\rho_{ch}\pi a^2 L}{\epsilon_0}$$

where  $\rho_{ch}$  is the charge density. Combining this together, we obtain:

$$\mathbf{E} = \begin{cases} \frac{\rho_{ch} a^2}{2\epsilon_0 \rho} \hat{\mathbf{p}}, & \rho > a \\ \mathbf{0}, & \rho < a \end{cases}$$

We see that the electric field goes like  $\rho^{-1}$ , which is consistent with the result obtained from 4-11.  $\square$

**Problem 45.1.28** (Wangsness 4-11) If  $\mathbf{E}$  is an electric field defined by  $\mathbf{E} = E_0(\rho/a)^3 \hat{\mathbf{p}}$  for  $0 < \rho < a$  and  $\mathbf{E} = \mathbf{0}$  otherwise. Find the charge density  $\rho_{ch}$  everywhere.

*Solution.* From the differential form of Gauss' Law, we have:

$$\nabla \cdot \mathbf{E} = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho E_\rho) + \frac{1}{\rho} \frac{\partial E_\phi}{\partial \phi} + \frac{\partial E_z}{\partial z} = \frac{4E_0 \rho^2}{a^3} = \frac{\rho_{ch}}{\epsilon_0}$$

$$\Rightarrow \rho_{ch} = \begin{cases} \frac{4\epsilon_0 E_0 \rho^2}{a^3}, & \rho < a \\ 0, & \rho > a \end{cases}$$

$\square$

**Problem 45.1.29** (Wangsness 4-12) If  $\mathbf{E}$  is an electric field defined by  $\mathbf{E} = (\cos(\theta)\hat{\mathbf{r}} + \sin(\theta)\hat{\mathbf{\phi}})2A/r^3$ , find the charge density  $\rho_{ch}$  everywhere.

*Solution.* We have from Gauss' Law:

$$\nabla \cdot \mathbf{E} = \frac{\rho_{ch}}{\epsilon_0} \Rightarrow \rho_{ch} = \epsilon_0 \nabla \cdot \mathbf{E}$$

But  $\nabla \cdot \mathbf{E} = 0$ , and thus  $\rho_{ch} = 0$ .  $\square$

## 45.1.6 Homework VI

**Problem 45.1.30** (Wangsness 5-1) Can  $\mathbf{E} = (yz - 2x)\hat{\mathbf{x}} + xz\hat{\mathbf{y}} + xy\hat{\mathbf{z}}$  be a possible electrostatic field? If so, find a possible potential function  $\phi$ .

*Solution.* We have that:

$$\nabla \times \mathbf{E} = (x - x)\hat{\mathbf{x}} - (y - y)\hat{\mathbf{y}} + (z - z)\hat{\mathbf{z}} = \mathbf{0}$$

Therefore  $\mathbf{E}$  is a possible electrostatic field. Indeed, writing  $\mathbf{E} = -\nabla(\phi)$ , we get:

$$\begin{aligned} -\frac{\partial \phi}{\partial x} &= yz - 2x \Rightarrow \phi = -xyz - x^2 + g(y, z) \\ -\frac{\partial \phi}{\partial y} &= xz \Rightarrow \frac{\partial g}{\partial y} = 0 \Rightarrow g(y, z) = g(z) \\ -\frac{\partial \phi}{\partial z} &= xy \Rightarrow \frac{\partial g}{\partial z} = 0 \Rightarrow g = \text{constant} \end{aligned}$$

The reason  $g$  is a function of  $y$  and  $z$  is because we are taking partial derivatives, and thus the “constants” of integration can be functions of the other variables. Thus we must verify what  $g(y, z)$  is. Taking partial derivatives with respect to  $y$  and  $z$  revealed to us that  $g$  is indeed just a constant. So, we have:

$$\phi(x, y, z) = -xyz + x^2 + C$$

where  $C$  is some constant. We may choose  $C$  as we desire, so let  $C = 0$  to make things easy. The path integral from the origin to a point  $(x, y, z)$  is independent of path and may be computed by using the fundamental theorem of gradients. This theorem says that, if  $\phi$  is differentiable (All of its partial derivatives exists), then:

$$\int_{P_1}^{P_2} \nabla(\phi) \cdot d\ell = \phi(P_2) - \phi(P_1)$$

That is, the path integral of the gradient of  $\phi$  is independent of the path. It only depends on the endpoints. This is *multivariate* form of the Fundamental Theorem of Calculus. Using this, we have:

$$\int_C \mathbf{E} \cdot d\ell = - \int_{(0,0,0)}^{(x,y,z)} \nabla(\phi) \cdot d\ell = \phi(0, 0, 0) - \phi(x, y, z) = xyz - x^2$$

□

**Problem 45.1.31** (Wangsness 5-3) Give two point charges  $q$  and  $-q$  on the  $z$  axis at  $z = a$  and  $z = -a$ , respectively, Find  $\phi$  everywhere. Show that the  $xy$  plane is an equipotential surface.

*Solution.* The potential is defined as:

$$\phi = \sum_k \frac{q_k}{4\pi\epsilon R_k}$$

Using this, we obtain:

$$\phi = \frac{1}{4\pi\epsilon_0} \left( \frac{q}{R_+} - \frac{q}{R_-} \right) = \frac{q}{4\pi\epsilon_0} \left( \frac{1}{\sqrt{x^2 + y^2 + (z-a)^2}} - \frac{1}{\sqrt{x^2 + y^2 + (z+a)^2}} \right)$$

Evaluating at  $z = 0$ , we have:

$$\phi_{z=0} = \frac{q}{4\pi\epsilon_0} \left( \frac{1}{\sqrt{x^2 + y^2 + a^2}} - \frac{1}{\sqrt{x^2 + y^2 + a^2}} \right) = 0$$

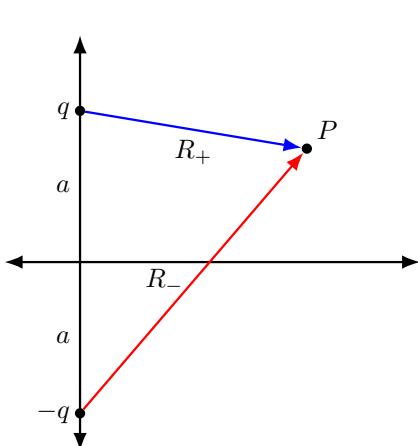
Thus, the entire  $xy$  plane is an equipotential surface with  $\phi = 0$ . □

**Problem 45.1.32** (Wangsness 5-4) Consider the charge distribution show in Fig. 45.8.2. Find  $\phi$  at the center of the square. Why can't you compute  $\mathbf{E}$  at this point from your result?

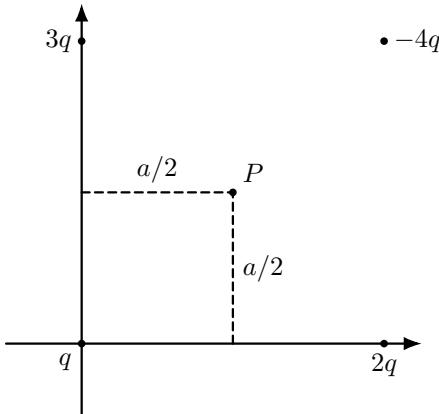
*Solution.*

$$\phi = \frac{1}{4\pi\epsilon_0} \sum \frac{q_i}{R_i} = \frac{1}{4\pi\epsilon_0} \left( \frac{q}{\sqrt{a^2/4}} + \frac{2q}{\sqrt{a^2/2}} - \frac{4q}{\sqrt{a^2/2}} + \frac{3q}{\sqrt{a^2/2}} \right) = \frac{q}{\sqrt{2}\pi\epsilon_0 a}$$

To know  $\mathbf{E}$  from  $\phi$ ,  $\phi$  must be known in some region about the point, not just at the point. To be precise in mathematical terms, we must know  $\phi$  in some open set about the point in order to compute  $\nabla(\phi)$ . This is analogous to functions in calculus. Suppose  $f$  is a function and  $a$  is a real number, and suppose we know the value of  $f(a)$ . Can we determine what  $f'(a)$  is? The answer is no, there is not enough information. If we know what  $f(x)$  is in some interval  $(a - \epsilon, a + \epsilon)$ , then we can compute  $f'(a)$ .  $\square$



45.8.1: Drawing for Wangsness 5-3



45.8.2: Drawing for Wangsness 5-4

Fig. 45.8: Drawings for problems 45.1.31 and 45.1.32

**Problem 45.1.33** (Wangsness 5-10) Given a sphere with radius  $a$  which has a charge density that varies by  $\rho_{ch}(r) = Ar^{1/2}$  for  $r < a$ , where  $A$  is a constant, find  $\phi$  at all points inside and outside of the sphere by using the path integral definition.

*Solution.* The potential difference between two points  $P_1$  and  $P_2$  is defined as:

$$\Delta\phi = - \int_{P_1}^{P_2} \mathbf{E} \cdot d\ell$$

We have that:

$$\mathbf{E} = \begin{cases} \frac{2Aa^{7/2}}{7\epsilon_0 r^2} \hat{\mathbf{r}}, & r > a \\ \frac{2Ar^{3/2}}{7\epsilon_0} \hat{\mathbf{r}}, & r < a \end{cases}$$

Now,  $\mathbf{d}\ell = -d\mathbf{r}(-\hat{\mathbf{r}}) = d\mathbf{r}$ . Let  $\phi$  be 0 at the origin. If  $r_0 < a$ , we have:

$$\phi = - \int_0^{r_0} \mathbf{E} \cdot d\ell = - \frac{2A}{7\epsilon_0} \int_0^{r_0} r^{3/2} dr = - \frac{4Ar_0^{5/2}}{35\epsilon_0}$$

If  $r_0 > a$ , then we have:

$$\phi = - \int_0^a \mathbf{E} \cdot d\ell - \int_a^{r_0} \mathbf{E} \cdot d\ell = - \frac{4Aa^{5/2}}{35\epsilon_0} - \frac{2Aa^{7/2}}{7\epsilon_0} \int_a^{r_0} \frac{1}{r^2} dr = - \frac{4Aa^{5/2}}{35\epsilon_0} - \frac{2Aa^{7/2}}{7\epsilon_0} \left( \frac{1}{r} - \frac{1}{a} \right)$$

□

**Problem 45.1.34** (Wangsness 5-11) Given two concentric spheres with radii  $a$  and  $b$ , with  $a < b$ , such that the region between them is filled with a charge of constant density  $\rho_{ch}$ , find  $\phi$  at all points and express the answer in terms of  $\rho_{ch}$ .

We have solved for the  $\mathbf{E}$  field in a previous problem, and have that:

$$\mathbf{E} = \begin{cases} \mathbf{0}, & r < a \\ \frac{Q}{4\pi\epsilon_0} \left( \frac{r^3 - a^3}{b^3 - a^3} \right) \hat{\mathbf{r}}, & a \leq r \leq b \\ \frac{Q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}, & r > b \end{cases}$$

We thus split the integral into three regions and compute:

$$\Delta\phi = \int \mathbf{E} \cdot d\ell = \int_0^a \mathbf{E} \cdot d\mathbf{r} + \int_a^b \mathbf{E} \cdot d\mathbf{r} + \int_b^\infty \mathbf{E} \cdot d\mathbf{r} \Rightarrow \phi(\mathbf{r}) = \begin{cases} \frac{\rho_{ch}(b^3 - a^3)}{3\epsilon_0 r}, & r > b \\ \frac{\rho_{ch}}{3\epsilon_0} \left( \frac{3}{2}b^2 - \frac{r^2}{2} - \frac{a^3}{r} \right), & a \leq r \leq b \\ \frac{\rho_{ch}}{2\epsilon_0} (b^2 - a^2), & r < a \end{cases}$$

**Problem 45.1.35** (Wangsness 5-14) Given a sphere of radius  $a$  with constant surface charge density  $\rho_{ch}$ , but no volume charge density, find  $\phi$  everywhere.

*Solution.*  $\phi$  may be defined as follows:

$$\phi = \frac{1}{4\pi\epsilon_0} \iint_{\Sigma} \frac{\sigma_{ch} da'}{R} \quad (45.1.60)$$

Here,  $R = |\mathbf{r} - \mathbf{r}'| = \sqrt{a^2 + r'^2 - 2ar \cos(\theta')}$ , and  $da' = a^2 \sin(\theta') d\theta' d\phi'$ . Thus:

$$\phi = \frac{\sigma_c}{4\pi\epsilon_0} \int_0^{2\pi} \int_0^\pi \frac{a^2 \sin(\theta') d\theta' d\phi'}{\sqrt{a^2 + r^2 - 2ar \cos(\theta')}} \quad (45.1.61a)$$

$$= \frac{\sigma_c a^2}{2\epsilon_0} \int_0^\pi \frac{\sin(\theta') d\theta'}{\sqrt{a^2 - r^2 - 2ar \cos(\theta')}} \quad (45.1.61b)$$

$$= \begin{cases} \frac{a^2 \sigma}{\epsilon_0 r}, & r > a \\ \frac{a \sigma}{\epsilon_0}, & r < a \end{cases} \quad (45.1.61c)$$

□

### 45.1.7 Homework VII

**Problem 45.1.36** (Wangness 6-6) Suppose there is a capacitor  $C_1$  that is charged to a potential difference  $\Delta\phi$  between its plates, and another capacitor  $C_2$  is uncharged. If one of the plates of  $C_1$  is connected to  $C_2$  by a conductor of negligible capacitance, and if the remaining plates are similarly connected, for the resultant equilibrium state, find the charge on each capacitor and the potential difference  $\Delta\phi'$  between their respective plates.

*Solution.* Before the connection we have  $Q = C_1\Delta\phi$ . After the connection,  $Q_1 = C_1\Delta\phi'$ ,  $Q_2 = C_2\Delta\phi'$ , and from the conservation of charge we have  $Q_1 + Q_2 = Q$ . Putting this together, we obtain:

$$Q_1 + Q_2 = (C_1 + C_2)\Delta\phi' = Q = C_1\Delta\phi$$

Thus:

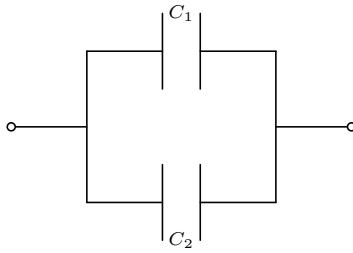
$$\Delta\phi' = \frac{C_1}{C_1 + C_2}\Delta\phi \quad Q_1 = \frac{C_1^2}{C_1 + C_2}\Delta\phi \quad Q_2 = \frac{C_1 C_2}{C_1 + C_2}\Delta\phi$$

□

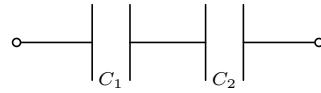
**Problem 45.1.37** (Wangness 6-7) Suppose the plates of two capacitors  $C_1$  and  $C_2$  are connected by conductors of negligible capacitance in parallel as shown in Fig. 45.9.1. If a potential difference is applied across the terminals of  $\Delta\phi$ , show that this combination is equivalent to a single capacitor of capacitance  $C_p = C_1 + C_2$ . Similarly, do this for series (Fig. 45.9.2) and show that  $1/C_s = 1/C_1 + 1/C_2$ .

In parallel we have  $Q_1 = C_1\Delta\phi$  and  $Q_2 = C_2\Delta\phi$ , where  $Q_1$  and  $Q_2$  are the charges on the plates  $C_1$  and  $C_2$ , respectively. The total charge is  $Q_1 + Q_2$ . Thus we have:

$$Q = Q_1 + Q_2 = C_1\Delta\phi + C_2\Delta\phi = (C_1 + C_2)\Delta\phi = C_p\Delta\phi \Rightarrow C_p = C_1 + C_2$$



45.9.1: Parallel Circuit.



45.9.2: Series Circuit.

Fig. 45.9: Circuits for problem 45.1.37

For series we have  $\Delta\phi = \Delta\phi_1 + \Delta\phi_2$ , where  $\Delta\phi_1$  and  $\Delta\phi_2$  are the potential differences across  $C_1$  and  $C_2$ , respectively. If a charge  $Q$  is on the left plate of  $C_1$ , then there is a charge  $-Q$  on the right plate, and therefore there is a charge  $Q$  on the left plate of  $C_2$  as well. Thus,  $Q = Q_1 = Q_2$ . So:

$$\Delta\phi = \frac{Q_1}{C_1} + \frac{Q_2}{C_2} = \frac{Q}{C_1} + \frac{Q}{C_2} = Q\left(\frac{1}{C_1} + \frac{1}{C_2}\right) = \frac{Q}{C_s} \Rightarrow \frac{1}{C_s} = \frac{1}{C_1} + \frac{1}{C_2}$$

**Problem 45.1.38** (Wangsness 6-9) Suppose that the potential difference between the plates of a spherical capacitor are kept constant at  $\Delta\phi$ . Show that the electric field at the surface of the inner sphere will be a minimum when  $a = b/2$ . Find this minimum value of  $E$ .

*Solution.* The charge on a capacitor is  $Q = C\Delta\phi$ . The electric field for a spherical capacitor at  $r = a$  is:

$$E = \frac{Q}{4\pi\epsilon_0 a^2}$$

For a spherical capacitor we have:

$$C = \frac{4\pi\epsilon_0 ab}{b-a} \Rightarrow E = \frac{b\Delta\phi}{a(b-a)}$$

To minimize this, we solve the following:

$$\frac{\partial E}{\partial a} = 0 \Rightarrow \frac{b\Delta\phi}{a^2(b-a)^2}(b-2a) = 0 \Rightarrow a = \frac{b}{2}$$

To see that this is a minimum, we look at the second partial derivative:

$$\frac{\partial^2 E}{\partial a^2} \Big|_{a=\frac{b}{2}} = \frac{32\Delta\phi}{b} > 0$$

Therefore  $a = b/2$  is a minimum. Evaluating  $E$  at this point, we have  $E = 4\Delta\phi/b$  □

**Problem 45.1.39** (Wangness 6-10) Given a capacitor made from two infinitely long conductors with coaxial cylindrical surfaces, like the ones shown in Fig. 45.10, show that the capacitance of a length  $L$  is given by  $C = 2\pi\epsilon_0 L / \ln(b/a)$ .

*Solution.* We have that  $E = \frac{\lambda}{2\pi\epsilon_0\rho}$ , where  $\lambda$  is the linear charge density,  $\lambda = \frac{Q}{L}$ . Thus,

$$\Delta\phi = - \int_b^a \frac{\lambda}{2\pi\epsilon_0\rho} d\rho = \frac{\lambda}{2\pi\epsilon_0} \ln\left(\frac{b}{a}\right)$$

Therefore:

$$C = \frac{Q}{\Delta\phi} = \frac{2\pi\epsilon_0 L}{\ln\left(\frac{b}{a}\right)}$$

□

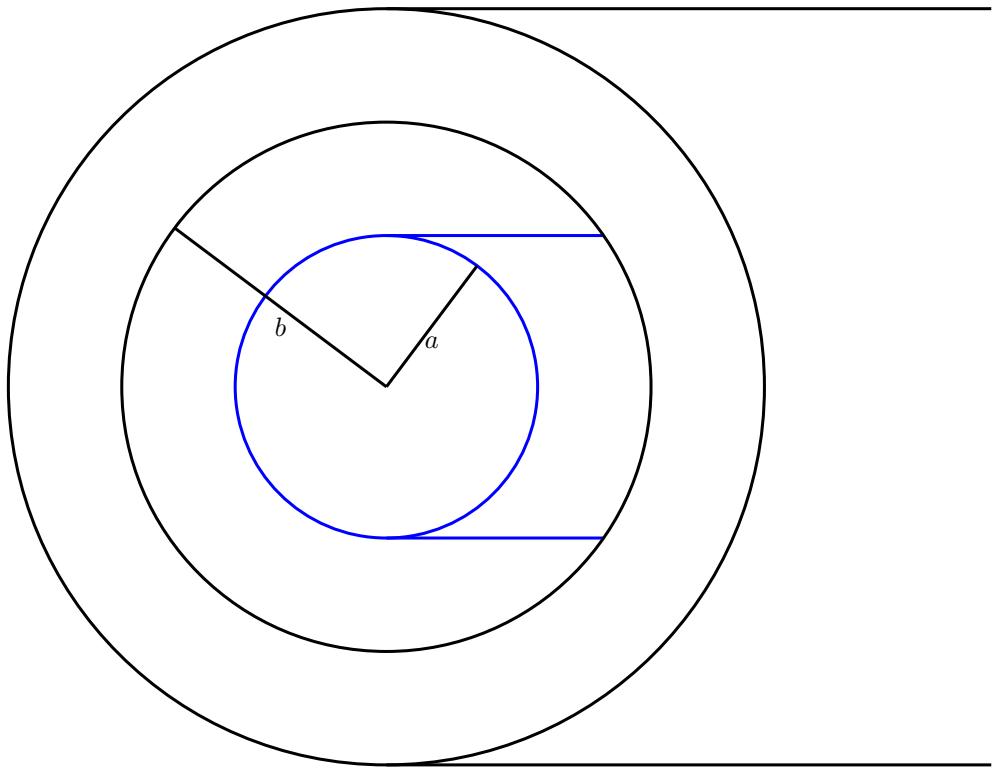


Fig. 45.10: Infinite Cylinders for problem 45.1.39

### 45.1.8 Homework VIII

#### Wangsness 7-2

$U_e = \sum_{\text{All Pairs}} \frac{q_i q_j}{r \pi \epsilon_0 R_{ij}}$ , where  $R_{ij} = |\mathbf{r}_i - \mathbf{r}_j| = \sqrt{r_i^2 + r_j^2 - 2\mathbf{r}_i \cdot \mathbf{r}_j}$ . Computing the sum, we get  $U_e = \frac{q^2}{4\pi\epsilon_0 a} \left( 12 + \frac{12}{\sqrt{2}} + \frac{4}{\sqrt{3}} \right)$ .

#### Wangsness 7-4

$\rho_c = Ar^n$ , where  $A$  is a constant and  $n \geq 0$ . Thus,  $U_e = \frac{1}{2} \int \rho_c(\mathbf{r}) \phi(\mathbf{r}) d\tau$ . We have that  $E = \frac{Aa^{n+3}}{\epsilon_0(n+3)r^2}$  for  $r \geq a$ , and  $E = \frac{Ar^{n+1}}{\epsilon_0(n+3)}$  for  $r \leq a$  from Gauss' Law. Now, all of the charges are located within  $r \leq a$ , and so we must compute  $\phi$  in this region. Letting  $\phi(r) \rightarrow 0$  as  $r \rightarrow \infty$ , we may compute  $\phi$  as  $\phi(r) = - \int_\infty^r \mathbf{E} \cdot d\mathbf{l} = - \int_\infty^a \mathbf{E} \cdot d\mathbf{l} - \int_a^r \mathbf{E} \cdot d\mathbf{l} = \frac{Aa^{n+3}}{\epsilon_0(n+3)a} + \frac{Aa^{n+2}}{\epsilon_0(n+2)(n+3)} - \frac{Ar^{n+2}}{\epsilon_0(n+3)(n+2)}$ . We can now compute the potential energy.  $U_e = \frac{1}{2} \int \rho_c(\mathbf{r}) \phi(\mathbf{r}) d\tau = \int_0^{2\pi} \int_0^\pi \int_0^a Ar^n \left( \frac{Aa^{n+3}}{\epsilon_0(n+3)a} + \frac{Aa^{n+2}}{\epsilon_0(n+2)(n+3)} - \frac{Ar^{n+2}}{\epsilon_0(n+3)(n+2)} \right) r^2 \sin(\theta) dr d\theta d\phi = \frac{2\pi A^2}{\epsilon_0(n+3)} \left[ \frac{a^{2n+5}}{n+3} + \frac{a^{2n+5}}{(n+2)(n+3)} - \frac{a^{2n+5}}{(n+2)(2n+5)} \right]$ . Taking the limit as  $n \rightarrow 0$ , we get  $\frac{3}{5} \left[ \frac{Q^2}{4\pi\epsilon_0 a} \right]$ , as expected for a constant spherical charge density.

#### Wangsness 7-6

$U_e = \frac{1}{2} \int_S \sigma_c(\mathbf{r}) \phi(\mathbf{r}) da$ . In the region between the cylinders we have that  $E = \frac{\lambda}{2\pi\epsilon_0 \rho}$ , and thus  $\phi = \frac{\lambda 2\pi\epsilon_0}{\ln} \left( \frac{\rho_0}{\rho} \right)$ , where  $\rho_0$  is the zero of  $\phi$ . We can now compute  $U_e$  and we get  $U_e = \frac{\lambda L}{4\pi\epsilon_0} \ln \left( \frac{b}{a} \right)$ . As  $U_e = \frac{1}{2} \frac{Q^2}{C}$ , we get  $C = \frac{2\pi\epsilon_0 L}{\ln \left( \frac{b}{a} \right)}$

#### Wangsness 7-9

The electric field is  $\mathbf{E}_i = \frac{Qr}{4\pi\epsilon_0 a^3} \hat{\mathbf{r}}$  inside the distribution, and  $\mathbf{E}_o = \frac{Q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}$ . The energy density inside is thus  $\mu_{e_i} = \frac{1}{2} \epsilon_0 E_i^2 = \frac{Q^2 r^2}{32\pi^2 \epsilon_0 a^6}$  and outside is  $\mu_{e_o} = \frac{1}{2} \epsilon_0 E_o^2 = \frac{Q^2}{32\pi^2 \epsilon_0 r^4}$ . The total energy is  $\int_{\text{Inside}} \mu_{e_i} d\tau + \int_{\text{Outside}} \mu_{e_o} d\tau$ . Computing this integral, we get  $U_e = \frac{3}{5} \left( \frac{Q^2}{4\pi\epsilon_0} \right)$ , in agreement with before.

## Wangsness 7-10

$U_e = \frac{\epsilon_0}{2} \int_{\text{All Space}} E^2 d\tau$ . The  $\mathbf{E}$ -Field in between the spheres is  $\frac{Q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}$ . The energy associated to this region is thus  $\int_0^{2\pi} \int_0^\pi \int_a^b \frac{\epsilon}{2} \frac{Q^2}{16\pi^2 \epsilon_0^2 r^4} r^2 \sin(\theta) dr d\theta d\phi = \frac{Q^2}{8\pi\epsilon_0} \left( \frac{b-a}{ab} \right)$ . We have that  $C = \frac{1}{2} \frac{Q^2}{U_e}$ , and thus  $C = \frac{4\pi\epsilon_0 ab}{b-a}$ .

## Wangsness 7-17

$\mathbf{E} = \frac{\lambda}{2\pi\epsilon_0\rho} \hat{\mathbf{r}}$ ,  $\Delta\phi = \frac{\lambda}{2\pi\epsilon_0 \ln\left(\frac{b}{a}\right)}$ , and therefore  $\lambda = \frac{2\pi\epsilon_0 \Delta\phi}{\ln\left(\frac{b}{a}\right)}$  and  $E = \frac{\Delta\phi}{\rho \ln\left(\frac{b}{a}\right)}$ . So,  $f_e = \mu_e = \frac{1}{2}\epsilon_0 E^2 \Big|_{\rho=a} = \frac{1}{2}\epsilon_0 \left[ \frac{\Delta\phi}{a \ln\left(\frac{b}{a}\right)} \right]$ . Thus,  $\mathbf{F}_{Tot} = \int f_e \mathbf{da} = f_e \int \mathbf{da} = 0$ .

### 45.1.9 Homework IX

#### Wangsness 8-5

The monopole moment is  $Q = \sum q_i = -3q - 2q - q + q + 2q + 3q + 4q + 5q = 9q$ . The dipole moment is  $\mathbf{p} = \sum q_i \hat{\mathbf{r}}_i = (-3q)\mathbf{0} + (-2q)a\hat{\mathbf{x}} + (-q)(a\hat{\mathbf{x}} + a\hat{\mathbf{y}}) + qa\hat{\mathbf{y}} + 2q(a\hat{\mathbf{y}} + a\hat{\mathbf{z}}) + 3q(a\hat{\mathbf{x}} + a\hat{\mathbf{y}} + a\hat{\mathbf{z}}) + 4q(a\hat{\mathbf{x}} + a\hat{\mathbf{z}}) + 5qa\hat{\mathbf{z}} = 4qa\hat{\mathbf{x}} + 5qa\hat{\mathbf{y}} + 14aq\hat{\mathbf{z}}$

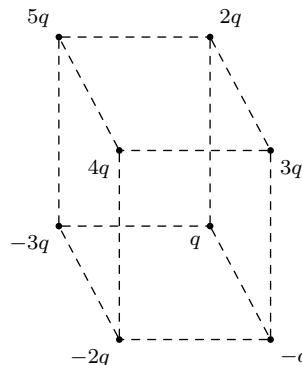


Fig. 45.11: Drawing for Wangsness 8-5

It is possible to find an origin about which the dipole moment will vanish. Consider an arbitrary charge distribution with the center of charge designated as c.c. The position vector of this is  $\mathbf{r}_{c.c.} = \frac{\int \mathbf{r}' \rho d\tau'}{\int \rho d\tau'} = \frac{\sum \mathbf{r}_i q_i}{\sum q_i}$ . Shifting the origin by  $\mathbf{r}_{c.c.}$  the dipole moment becomes zero. The original dipole moment is  $\mathbf{p}_o = \int \mathbf{r}' d\tau' = \sum \mathbf{r}_i q_i$ . The new dipole moment will be  $\mathbf{p}_N = \int (\mathbf{r}' - \mathbf{r}_{c.c.}) d\tau' = \int \mathbf{r}' d\tau' - \mathbf{r}_{c.c.} \int \rho d\tau' = \sum \mathbf{r}_i q_i - \mathbf{r}_{c.c.} \sum q_i = \sum \mathbf{r}_i q_i - \frac{\sum \mathbf{r}_i q_i}{\sum q_i} \sum q_i = \mathbf{0}$ .

For the problem at hand, this equates to  $\mathbf{r}_{c.c} = \langle \frac{4}{9}a, \frac{5}{9}a, \frac{14}{9}a \rangle$  (In Cartesian Coordinates).

### Problem 45.1.40 (Wangsness 8-8)

*Solution.*

$$Q = \int_{S'} \sigma da' = \int_0^{2\pi} \sin(\theta) \cos(\theta) d\theta = \int_0^{2\pi} \int_0^\pi \sigma_0 \cos(\theta) a^2 \sin(\theta) d\theta d\phi = 2\pi\sigma_0 = 0$$

$$\mathbf{p} = \int_{S'} \sigma \mathbf{r}' da' = \sigma a^3 \int_0^{2\pi} \int_0^\pi \cos(\theta) (\sin(\theta) \cos(\phi) \hat{\mathbf{x}} + \sin(\theta) \sin(\phi) \hat{\mathbf{y}} + \cos(\theta) \hat{\mathbf{z}}) \sin(\theta) d\theta d\phi$$

$$\phi \approx \frac{1}{4\pi\epsilon_0} \frac{\mathbf{p} \cdot \hat{\mathbf{r}}}{r^2} = \frac{\sigma_0 a^3}{3\epsilon_0^2 r^2} \cos(\theta)$$

□

### Wangsness 9-1

Surface of Separation between regions 1 and 2 is a plane  $f = 2x + y + z = 1$ .  $\mathbf{E}_1 = 4\hat{\mathbf{x}} + \hat{\mathbf{y}} - 3\hat{\mathbf{z}}$  is given. Find the normal and tangential component of  $\mathbf{E}_1$ : The unit vector is the normal to the plane which is  $\hat{\mathbf{n}} = \frac{\nabla(f)}{|\nabla(f)|} = \frac{2\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}}{\sqrt{6}}$ . The normal component of  $\mathbf{E}_1$  is  $\mathbf{E}_1 \cdot \hat{\mathbf{n}} = \sqrt{6}$ . Thus,  $\mathbf{E}_{1n} = (\mathbf{E}_1 \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} = 2\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}$ . The tangential component is  $\mathbf{E}_1 - \mathbf{E}_{1n} = 2\hat{\mathbf{x}} - 4\hat{\mathbf{z}}$ .

### Wangsness 9-3

We are given the density  $\sigma = \sigma_0 \cos(\theta) = \frac{\sigma_0 z}{a}$  and  $\mathbf{E}_1 = \alpha\hat{\mathbf{x}} + \beta\hat{\mathbf{y}} + \gamma\hat{\mathbf{z}}$ . The boundary conditions are  $E_{2t} = E_{1t}$  and  $E_{2n} - E_{1n} = \frac{\sigma}{\epsilon_0}$ . We can write  $\mathbf{E}_1 = E_{1t}\hat{\mu} + E_{1n}\hat{\mathbf{r}}$  where  $\hat{\mathbf{r}}$  is the normal to the spherical surface and  $\hat{\mu}$  is the tangent to the sphere. On the outside,  $\mathbf{E}_2 = E_{2t}\hat{\mu} + E_{2n}\hat{\mathbf{r}}$ . Now, using the boundary conditions the  $E$ -field on the outside is  $\mathbf{E}_2 = E_{1t}\hat{\mu} + \left(\frac{\sigma}{\epsilon_0} + E_{1n}\right)\hat{\mathbf{r}} = E_{1t}\hat{\mu} + E_{1n}\hat{\mathbf{r}} + \frac{\sigma}{\epsilon_0}\hat{\mathbf{r}} = \alpha\hat{\mathbf{x}} + \beta\hat{\mathbf{y}} + \gamma\hat{\mathbf{z}} + \frac{\sigma z}{\epsilon_0 a} \left( \frac{x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}}{a} \right)$ . So,  $\mathbf{E}_2 = \left(\alpha + \frac{\sigma_0 zx}{\epsilon_0 a^2}\right)\hat{\mathbf{x}} + \left(\beta + \frac{\sigma_0 zy}{\epsilon_0 a^2}\right)\hat{\mathbf{y}} + \left(\gamma + \frac{\sigma_0 z^2}{\epsilon_0 a^2}\right)\hat{\mathbf{z}}$

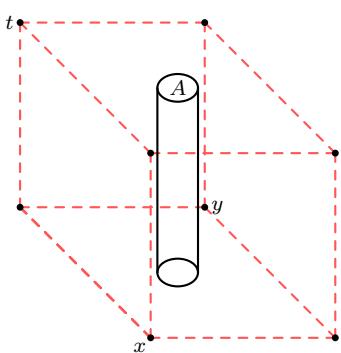
### 45.1.10 Homework X

#### Wangsness 10-3

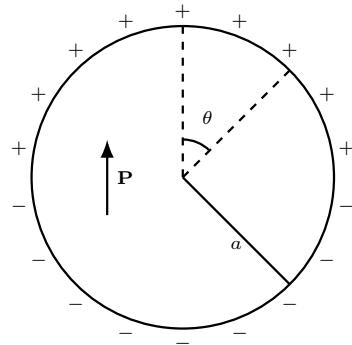
$\mathbf{P} = P(1 + \alpha z)\hat{\mathbf{z}}$ , where  $P$  and  $\alpha$  are constants. Volume charge density is  $\rho_b = -\nabla \cdot \mathbf{P} = -\alpha P$ . The surface charge densities are  $\mathbf{P} \cdot \hat{\mathbf{n}}$ , so  $\sigma_{top} = P(1 + \alpha t)$  and  $\sigma_{bottom} = -P$ . On the left and right sides the charge density is zero as the normals to the sides are at right angles with  $\mathbf{P}$ . So,  $Q = \int_V \rho d\tau' + \int_{top} \sigma_{top} da' + \int_{bottom} \sigma_{bottom} da' = \int_V -\alpha P d\tau' + \int_{top} P(1 + \alpha t) da' + \int_{bottom} -P da' = \alpha P A t + PA - \alpha P A t - PA = 0$

**Wangsness 10-6**

We are given that  $\mathbf{P} = P_0 \hat{\mathbf{k}}$ . Now  $\rho_b = -\nabla \cdot \mathbf{P}$ , and as  $\mathbf{P}$  is uniform,  $-\nabla \cdot \mathbf{P} = 0$ . Thus  $\rho_b = 0$ .  $\sigma_b = \mathbf{P} \cdot \hat{\mathbf{n}} = P_0 \hat{\mathbf{k}} \cdot \hat{\mathbf{n}} = P_0 \cos(\theta)$ . The positive charge is thus located in the region  $\theta < \frac{\pi}{2}$ . So  $Q_b^+ = \int_0^{\pi/2} \int_0^{2\pi} P_0 \cos(\theta) a^2 \sin(\theta) d\theta d\phi = \pi a^2 P_0$ .



45.12.1: Drawing for Wangsness 10-3



45.12.2: Drawing for Wangsness 10-6

**Wangsness 10-17**

Choose a spherical Gaussian surface outside the sphere concentric with the given sphere.  $\oint \mathbf{D} \cdot d\mathbf{a} = Q_f$ , so  $D_o(4\pi r^2) = q$ , and thus  $\mathbf{D}_o = \frac{q}{4\pi r^2} \hat{\mathbf{r}}$ . From  $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$ , we have that  $\mathbf{D}_o - \epsilon_0 \mathbf{E}_o = \mathbf{P}_o$ .  $\mathbf{E}_o = \frac{q}{4\pi \epsilon_0 r^2} \hat{\mathbf{r}}$ , and thus  $\mathbf{P}_o = 0$ . There is no dielectric outside of the sphere. Choosing a Gaussian surface inside of the sphere, we get  $\oint \mathbf{D} \cdot d\mathbf{a} = Q_f$ , for  $D(4\pi r^2) = q$ , and thus  $\mathbf{D}_i = \frac{q}{4\pi r^2} \hat{\mathbf{r}}$ .  $\mathbf{E}_i = \frac{\mathbf{D}_i}{\epsilon} = \frac{\mathbf{D}_i}{\kappa_e \epsilon_0} = \frac{q}{4\pi \kappa_e \epsilon_0 r^2} \hat{\mathbf{r}}$ . So  $\mathbf{P}_i = \mathbf{D}_i - \epsilon_0 \mathbf{E}_i = (1 - \frac{1}{\kappa_e}) \frac{q}{4\pi r^2} \hat{\mathbf{r}}$ . Finally,  $Q_b^{surface} = \int_S \sigma_b da' = \iint \mathbf{P} \cdot \hat{\mathbf{n}} da' = \int_0^\pi \int_0^{2\pi} \frac{\kappa_e - 1}{\kappa_e} \frac{q}{4\pi} \sin(\theta) d\theta d\phi = \frac{\kappa_e - 1}{\kappa_e} q$ .

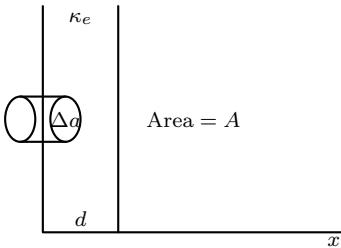
**Wangsness 10-18**

$\oint \mathbf{D} \cdot d\mathbf{a} = q$ .  $\mathbf{D} = \frac{q}{4\pi r^2} \hat{\mathbf{r}}$  for all  $r$  inside the cavity or in the dielectric.  $\rho_b = 0$  since  $\rho_f = 0$  in the dielectric. In the dielectric  $\mathbf{E} = \frac{\mathbf{D}}{\epsilon} = \frac{\mathbf{D}}{\kappa_e \epsilon_0}$ , so  $\mathbf{E} = \frac{q}{4\pi \kappa_e \epsilon_0 r^2} \hat{\mathbf{r}}$ .  $\mathbf{P} = \mathbf{D} - \epsilon_0 \mathbf{E} = \frac{\kappa_e - 1}{\kappa_e} \frac{q}{4\pi r^2} \hat{\mathbf{r}}$  at the surface of the cavity  $r = a$ .  $\sigma_b = \mathbf{P} \cdot \hat{\mathbf{n}} = \frac{\kappa_e - 1}{\kappa_e} \frac{q}{4\pi a^2} \hat{\mathbf{r}} \cdot (-\hat{\mathbf{r}}) = -\frac{\kappa_e - 1}{\kappa_e} \frac{q}{4\pi a^2}$ .  $Q_b^{cavity} = \int \sigma da = -\frac{\kappa_e - 1}{\kappa_e} q$ .

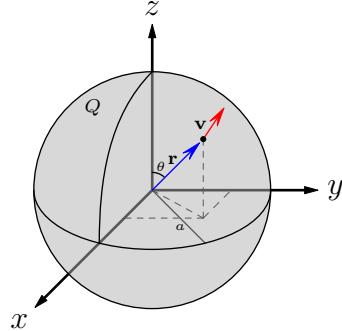
**Wangsness 10-25**

$\kappa_e(x) = \alpha + \beta x$  (The dielectric constant varies linearly with  $x$ .  $\alpha$  and  $\beta$  are constants). Find  $\mathbf{D}$  between the plates.  $\int_{GaussianSurface} \mathbf{D} \cdot d\mathbf{a} = Q_f^{enc}$  ( $D$  is

uniform between plates).  $D\Delta a = Q_f^{enc}$ , and thus  $D = \frac{Q_f^{enc}}{\Delta a} = \sigma = \frac{Q}{A}$ , where  $Q$  is the total charge of the plate and  $A$  is the area of the plate.  $E = \frac{D}{\epsilon} = \frac{Q}{\kappa_e_0 A} = \frac{Q}{\epsilon_0 A(\alpha + \beta x)}$ . At  $x = 0$ ,  $\kappa_e = \kappa_{e1}$ , so  $\alpha + \beta(0) = \kappa_{e1}$ , and thus  $\alpha = \kappa_{e1}$ . At  $x = d$ ,  $\kappa_e = \kappa_{e2}$ , and so  $\beta = \frac{\kappa_{e2} - \kappa_{e1}}{d}$ . The potential difference between the plates is  $\Delta\phi = -\int_{-}^{+} \mathbf{E} \cdot d\ell = \int_{-}^{+} Edx = \frac{Q}{\epsilon_0 A} \int_0^d \frac{dx}{\alpha + \beta x} = \frac{Q}{\epsilon_0 A \beta} \ln(\alpha + \beta x) \Big|_0^d = \frac{Q}{\epsilon_0 A \beta} \ln\left(\frac{\alpha + \beta d}{\alpha}\right) = \frac{Q}{\epsilon_0 A \beta} \ln\left(\frac{\kappa_{e2}}{\kappa_{e1}}\right) = \frac{Q}{C}$ . Hence  $C = \frac{\epsilon_0 A \beta}{\ln\left(\frac{\kappa_{e2}}{\kappa_{e1}}\right)} = \frac{(\kappa_{e2} - \kappa_{e1})\epsilon_0 A}{d \ln\left(\frac{\kappa_{e2}}{\kappa_{e1}}\right)}$



45.13.1: Drawing for Wangsness 10-25



45.13.2: Drawing for Wangsness 12-3

## Wangsness 10-27

$\kappa = \kappa_{e1}$  for  $a \leq \rho < \rho_0$ ,  $\kappa = \kappa_{e2}$  for  $\rho_0 \leq \rho \leq b$ . First get  $D$  by assuming a charge per unit length  $\lambda$  on the inner cylinder and  $-\lambda$  on the outer.  $\int \mathbf{D} \cdot d\mathbf{a} = Q_f^{enc} = D(2\pi\rho L) = \lambda L$ . So  $\mathbf{D} = \frac{\lambda}{2\pi\rho} \hat{\rho}$ .  $\Delta\phi = -\int_{-}^{+} \mathbf{E} \cdot d\ell = \int_a^b \frac{\lambda}{2\pi\rho\epsilon} d\rho = \int_a^{\rho_0} \frac{\lambda}{2\pi\epsilon_0\kappa_{e1}\rho} d\rho + \int_{\rho_0}^b \frac{\lambda}{2\pi\epsilon_0\kappa_{e2}\rho} d\rho = \frac{\lambda}{2\pi\epsilon_0} \left[ \frac{1}{\kappa_{e1}} \ln\left(\frac{\rho_0}{a}\right) + \frac{1}{\kappa_{e2}} \ln\left(\frac{b}{\rho_0}\right) \right]$ . From  $\Delta\phi = \frac{Q}{C}$ , and  $Q = \lambda L$ , we get  $C = \frac{2\pi\epsilon_0 L}{\frac{1}{\kappa_{e1}} \ln\left(\frac{\rho_0}{a}\right) + \frac{1}{\kappa_{e2}} \ln\left(\frac{b}{\rho_0}\right)}$

## 45.1.11 Homework XI

### Wangsness 12-3

$\mathbf{J} = \rho \mathbf{v}$ .  $\rho = \frac{Q}{\frac{4}{3}\pi a^3} = \frac{3q}{4\pi a^3}$ .  $\mathbf{u} = \omega \times \mathbf{r} = \omega \hat{\mathbf{z}} \times r \hat{\mathbf{r}} = \omega r \sin(\theta) \hat{\phi}$ . So, we have that  $\mathbf{J} = \frac{3Q}{4\pi a^3} \omega r \sin(\theta) \hat{\phi}$ .  $d\mathbf{a} = r dr d\theta \hat{\phi}$ , so  $I = \int \mathbf{J} \cdot d\mathbf{a} = \frac{3Q\omega}{4\pi a^3} \int_0^\pi \int_0^a r^2 \sin(\theta) dr d\theta = \frac{Q\omega}{2\pi}$

### Wangsness 13-4

We will calculate the force exerted by  $C'$  on  $C$ .  $\mathbf{F}_{C' \rightarrow C} = \frac{\mu_0}{4\pi} \oint_C \oint_{C'} \frac{I d\ell \times (I' d\ell' \times \hat{\mathbf{r}})}{R^2}$ . We use the  $BAC - CAB$  rule:  $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ . We can rewrite

the previous integral as  $\mathbf{F}_{C' \rightarrow C} = -\frac{\mu_0 II'}{4\pi} \oint_C \oint_{C'} [\mathbf{d}\ell' \cdot (\mathbf{d}\ell \times \frac{\hat{\mathbf{r}}}{R^2}) - \frac{\hat{\mathbf{r}}}{R^2} \mathbf{d}\ell \cdot \mathbf{d}\ell']$ . Recall that  $\nabla(\frac{1}{R}) = \frac{\hat{\mathbf{r}}}{R^2}$ . Using this, we have  $\mathbf{F}_{C' \rightarrow C} = -\frac{\mu_0 II'}{4\pi} \oint_C \oint_{C'} \mathbf{d}\ell' [\mathbf{d}\ell \cdot \nabla(\frac{1}{R}) - \frac{\hat{\mathbf{r}}}{R^2} \mathbf{d}\ell' \cdot \mathbf{d}\ell]$ . From the fundamental theorem of gradients,  $\oint \nabla(f) \cdot \mathbf{d}\ell = 0$  for any function  $f$ . Thus  $\oint \nabla(\frac{1}{R}) \cdot \mathbf{d}\ell = 0$ . From this we have  $\mathbf{F}_{C \rightarrow C'} = -\frac{\mu_0 II'}{4\pi} \oint_C \oint_{C'} \frac{\hat{\mathbf{r}}}{R^2} \mathbf{d}\ell' \cdot \mathbf{d}\ell$ . We now compute this integral along all four paths of the problem.  $\mathbf{d}\ell' = dz' \hat{\mathbf{z}}$  for all paths. Along path  $I$ ,  $\mathbf{r} = \hat{\mathbf{x}}d + \hat{\mathbf{z}}z$ ,  $\mathbf{d}\ell = \hat{\mathbf{z}}dz$ . Along path  $III$ ,  $\mathbf{r} = \hat{\mathbf{x}}(a+d) + \hat{\mathbf{z}}z$ ,  $\mathbf{d}\ell = \hat{\mathbf{z}}dz$ . Along paths  $II$  and  $IV$ ,  $\mathbf{d}\ell \cdot \mathbf{d}\ell' = 0$ . Piecing this together,  $\mathbf{F}_{C \rightarrow C'} = -\frac{\mu_0 II'}{4\pi} \int_0^b \int_{-\infty}^{\infty} \frac{\hat{\mathbf{x}}d + \hat{\mathbf{z}}(z-z')}{(d^2 + (z-z')^2)^{3/2}} dz' dz - \frac{\mu_0 II'}{4\pi} \int_b^0 \int_{-\infty}^{\infty} \frac{\hat{\mathbf{x}}(d+a) + \hat{\mathbf{z}}(z-z')}{((d+a)^2 + (z-z')^2)^{3/2}} dz' dz$ . Making the substitution  $t = z' - z$ , we get an integral of the form  $\int_{-\infty}^{\infty} \frac{t+z'}{(A+t^2)^{3/2}} dt$ . This is an odd function that is integrated over symmetric bounds, and thus the integral is zero. The only part left is the  $\hat{\mathbf{x}}$  contribution. Evaluating this integral, we get  $\mathbf{F}_{C \rightarrow C'} = -\frac{\mu_0 II' ab}{2\pi d(a+d)} \hat{\mathbf{x}}$ .

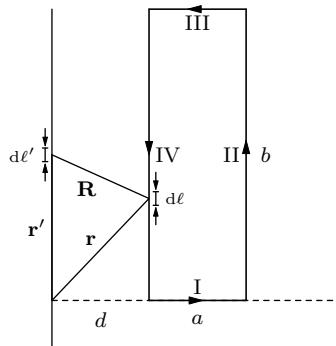


Fig. 45.14: Drawing for Wangsness 13-4

### Wangsness 14-7

$\mathbf{R} = \mathbf{r} - \mathbf{r}'$ , where  $\mathbf{r} = z\hat{\mathbf{z}}$  and  $\mathbf{r}' = a \cos(\phi')\hat{\mathbf{x}} + a \sin(\phi')\hat{\mathbf{y}}$ . We have  $d\ell' = ad\phi' \hat{\boldsymbol{\phi}}$ . Putting this together, we have  $\mathbf{R} = z\hat{\mathbf{z}} - a(\cos(\phi')\hat{\mathbf{x}} + \sin(\phi')\hat{\mathbf{y}})$ . So:

$$\begin{aligned} \mathbf{B} &= \frac{\mu_0 I'}{4\pi} \int \frac{d\ell \times \mathbf{R}}{R^3} = \frac{\mu_0 I'}{4\pi} \int_{-\alpha}^{\alpha} \frac{ad\phi' \hat{\boldsymbol{\phi}} \times (z\hat{\mathbf{z}} - a \cos(\phi')\hat{\mathbf{x}} - a \sin(\phi')\hat{\mathbf{y}})}{(z^2 + a^2)^{3/2}} \\ &= \frac{\mu_0 I' a}{4\pi(z^2 + a^2)^{3/2}} \int_{-\alpha}^{\alpha} (-\sin(\phi')\hat{\mathbf{x}} + \cos(\phi')\hat{\mathbf{y}}) \times (-a \cos(\phi')\hat{\mathbf{x}} - a \sin(\phi')\hat{\mathbf{y}} + z\hat{\mathbf{z}}) d\phi' \\ &= \frac{\mu_0 I' a}{4\pi(z^2 + a^2)^{3/2}} \int_{-\alpha}^{\alpha} (z \cos(\phi')\hat{\mathbf{x}} + z \sin(\phi')\hat{\mathbf{y}} + a\hat{\mathbf{z}}) d\phi' \end{aligned}$$

Sine is an odd function, and the limit is over a symmetric interval, and thus the  $\hat{\mathbf{y}}$  component is zero. So we have:

$$\mathbf{B} = \frac{\mu_0 I' a}{2\pi(z^2 + a^2)^{3/2}} (z \sin(\alpha) \hat{\mathbf{x}} + a \alpha \hat{\mathbf{z}})$$

### Wangsness 14-15

The force on  $q$  is given by  $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ . We first get  $\mathbf{B}$  at  $q$ .  $\mathbf{B} = \frac{\mu_0}{4\pi} \int \frac{I' d\ell' \times \hat{\mathbf{f}}}{R^2}$ . For this problem,  $\mathbf{R} = -\rho' \hat{\boldsymbol{\rho}}$ . We need only compute the integral along paths  $I$  and  $III$ , for along  $II$  and  $IV$  we have that  $d\ell$  and  $\mathbf{R}$  are parallel. So, we have  $\mathbf{B} = \frac{\mu_0}{4\pi} \int_0^\pi \frac{I'(-ad\phi' \hat{\boldsymbol{\phi}}) \times (-a\hat{\boldsymbol{\rho}})}{a^3} + \frac{\mu_0}{4\pi} \int_0^\pi \frac{I'(bd\phi' \hat{\boldsymbol{\phi}}) \times (-b\hat{\boldsymbol{\rho}})}{b^3} = \frac{\mu_0 I'}{4} \frac{b-a}{ab} \hat{\mathbf{z}}$ . The force is  $\mathbf{F} = qv\hat{\mathbf{y}} \times \frac{\mu_0 I}{4} \frac{b-a}{ab} \hat{\mathbf{z}} = \frac{qv\mu_0 I'}{4} \frac{b-a}{ab} \hat{\mathbf{x}}$

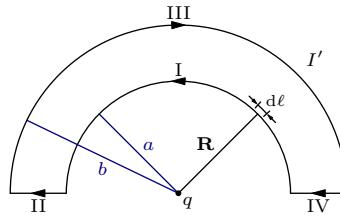


Fig. 45.15: Drawing for Wangsness 14-15

### 45.1.12 Homework XII

#### Wangsness 15-7

For  $\rho \leq a$ , path (1) has  $\oint \mathbf{B} \cdot d\ell = \mu_0 I_{enc}$ , where  $I_{enc} = I \frac{\rho^2}{a^2}$ . So  $\mathbf{B} = \frac{\mu_0 I \rho}{2\pi a^2} \hat{\boldsymbol{\phi}}$ . For  $a \leq \rho \leq b$ ,  $I_{enc} = I$ . So  $\mathbf{B} = \frac{\mu_0 I}{2\pi \rho} \hat{\boldsymbol{\phi}}$ . For  $b \leq \rho \leq c$ ,  $I_{enc} = I = I \frac{\rho^2 - b^2}{c^2 - b^2} = I \frac{c^2 - \rho^2}{c^2 - b^2}$ . So  $\mathbf{B} = \frac{\mu_0 I}{2\pi \rho} \frac{c^2 - \rho^2}{c^2 - b^2} \hat{\boldsymbol{\phi}}$ . Finally, or  $\rho \geq c$ ,  $I_{enc} = 0$ , so  $\mathbf{B} = 0$ .

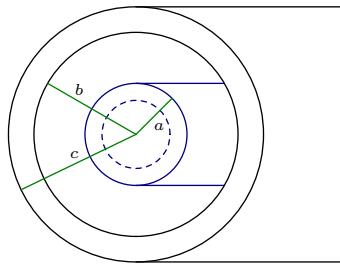


Fig. 45.16: Drawing for Wangsness 15-7

### Wangsness 15-8

$$\mathbf{B} = \begin{cases} 0, & \rho < a \\ \frac{\mu_0 I}{2\pi\rho} \frac{\rho^2 - a^2}{b^2 - a^2} \hat{\Phi}, & a < \rho < b \\ \frac{\mu_0 I}{2\pi\rho} \hat{\Phi}, & \rho > b \end{cases}$$

By definition,  $\mu_0 \mathbf{J} = \nabla \times \mathbf{B}$ . So:

$$\mathbf{J} = \begin{cases} 0, & \rho < a \\ \frac{I}{\pi(b^2 - a^2)}, & a < \rho < b \\ 0, & \rho > b \end{cases}$$

The current  $I$  is distributed uniformly over the volume between two coaxial cylinders of inner radius  $a$  and outer radius  $b$  in the direction of the cylinder axis.

### Wangsness 16-10

The field point is on the  $z$ -axis.  $\mathbf{r} = z\hat{\mathbf{z}}$ ,  $\mathbf{r}' = a \cos(\phi')\hat{\mathbf{x}} + a \sin(\phi')\hat{\mathbf{y}}$ .  $d\ell' = d\mathbf{r}' = ad\phi'\hat{\Phi} = ad\phi'(-\sin(\phi')\hat{\mathbf{x}} + \cos(\phi')\hat{\mathbf{y}})$ .  $\mathbf{A} = \frac{\mu_0 I'}{4\pi} \int \frac{d\ell'}{R} = \frac{\mu_0 I'}{4\pi} \frac{a}{\sqrt{a^2 + z^2}} \int_{-\alpha}^{\alpha} (-\sin(\phi')\hat{\mathbf{x}} + \cos(\phi')\hat{\mathbf{y}}) d\phi' = \frac{\mu_0 I' a}{2\pi \sqrt{a^2 + z^2}} \sin(\alpha)\hat{\mathbf{y}}$ . To find  $\mathbf{B}$  from  $\mathbf{A}$ , we need to evaluate  $\nabla \times \mathbf{A}$ . We don't know about  $\mathbf{A}$  for a general point, and thus we can't evaluate the  $x$  and  $y$  derivatives.

### 45.1.13 Homework XIII

#### Wangsness 17-3

The  $\mathbf{B}$  field associated with  $I$  is  $\mathbf{B} = \frac{\mu_0 I}{2\pi\rho} \hat{\Phi}$ . In the plane of the paper,  $\phi$  is into the paper.  $\Phi = \int \mathbf{B} \cdot d\mathbf{a} = \int \frac{\mu_0 I}{2\pi\rho} \cdot bd\rho \hat{\Phi} = \frac{\mu_0 Ib}{2\pi} \int_d^{d+a} \frac{d\rho}{\rho} = \frac{\mu_0 Ib}{2\pi} \ln\left(\frac{d+a}{d}\right) = \frac{\mu_0 b I_0 e^{-\lambda t}}{2\pi} \ln\left(\frac{d+a}{d}\right) = \frac{\mu_0 I_0 \lambda b}{2\pi} \ln\left(\frac{d+a}{d}\right) e^{-\lambda t}$ . The induced current is clockwise around the loop to produce a field which goes into the paper to counteract the decreasing  $\mathbf{B}$  due to  $I_0$ .

#### Wangsness 17-4

The  $\mathbf{B}$ -field at distance  $\rho$  from the wire at points in the plane of the paper is  $\mathbf{B} = \frac{\mu_0 I}{2\pi\rho} \hat{\mathbf{y}}$ . The flux of  $\mathbf{B}$  through the loop is  $\Phi = \int \mathbf{B} \cdot d\mathbf{a} = \iint \frac{\mu_0 I}{2\pi\rho} \rho d\theta d\rho$ . We have  $\rho = b + r \cos(\theta)$ . So  $\Phi = \frac{\mu_0 I}{2\pi} \int_0^a \int_0^{2\pi} \frac{rd\theta dr}{b+r \cos(\theta)} = \frac{\mu_0 I}{2\pi} \int_0^a \frac{2r}{\sqrt{b^2 - r^2}} \tan^{-1} \left[ \frac{\sqrt{b^2 - r^2} \tan(\theta/2)}{b+r} \right]_0^{2\pi} \Rightarrow \tan^{-1} \left[ \frac{\sqrt{b^2 - r^2}}{b+r} \tan(\pi) \right] - \tan^{-1} \left[ \frac{\sqrt{b^2 - r^2}}{b+r} \tan(0) \right]$ . So  $\Phi = \mu_0 I [b - \sqrt{b^2 - a^2}]$ . The loop moves with constant speed  $v$  along the  $x$ -axis away from the current  $I$ ,  $v = \frac{db}{dt}$ . So  $\xi = -\frac{d\Phi}{dt} = -\mu_0 I \frac{d}{dt} [b - \sqrt{b^2 - a^2}]$ .

$\sqrt{b^2 - a^2}] = -\mu_0 I [v - \frac{bv}{\sqrt{b^2 - a^2}}] = \mu_0 N I v [\frac{b}{\sqrt{b^2 - a^2}} - 1]$ . The current will be clockwise trying to increase the flux which is decreasing due to motion away from the wire.

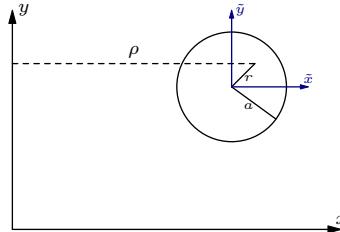


Fig. 45.17: Drawing for Wangness 17-4

### Wangness 17-19

$\Phi_{12} = IM_{12}$ . The flux due to 1 through 2 is  $\Phi_{12} = \int \mathbf{B}_1 \cdot d\mathbf{a}_2$ .  $\mathbf{B}_1 = \frac{\mu_0 I}{2\pi} \left( \frac{1}{\rho+d} - \frac{1}{\rho+d+D} \right)$ . So we have that  $\Phi_{12} = \int_0^a \frac{\mu_0 I}{2\pi} \left( \frac{1}{\rho+d} - \frac{1}{\rho+d+D} \right) bd\rho = \frac{\mu_0 Ib}{2\pi} \left[ \ln\left(\frac{a+d}{a+d+D}\right) - \ln\left(\frac{d}{d+D}\right) \right]$ . Thus, we have  $M = \frac{\mu_0 b}{2\pi} \ln\left(\frac{a+d}{d}\right)$

### Wangness 17-20

The field inside the toroid is  $\mathbf{B} = \frac{\mu_0 NI}{2\pi\rho} \hat{\phi}$ . The flux through a single turn is  $\Phi^1 = \frac{\mu_0 NI}{2\pi} \int_0^a \int_0^{2\pi} \frac{r}{b+r\cos(\theta)} d\theta dr$ . We've done this integral before, and we get  $\Phi^1 = \mu_0 NI [b - \sqrt{b^2 - r^2}]$ .  $\Phi = N\Phi^1$ .  $L = \frac{\Phi}{I} = \frac{\mu_0 N^2 I}{I} [b - \sqrt{b^2 - r^2}] = \mu_0 N^2 [b - \sqrt{b^2 - r^2}]$

## 45.2 Exams

### 45.2.1 Exam I

#### Question I

Give the vector field  $\mathbf{A} = c\hat{\theta}$ , where  $c$  is a constant, find  $\nabla \times \mathbf{A}$ . Is this a conservative vector field? Explain.

#### Question II

Verify the Divergence Theorem for  $\mathbf{A}$  given in problem 1 in spherical coordinates for a hemisphere of radius  $a_0$  resting on the  $xy$ -plane with the center

of the flat base of the hemisphere at the origin and the symmetry axis of the hemisphere along the positive  $z$ -axis.

### Question III

A semicircular charged line of radius  $a$  carries uniform linear charge density  $\lambda$ . It has the equation  $x^2 + y^2 = a^2$ ,  $x \geq 0$ ,  $z = 0$ . That is, the half circle resting on the  $xy$ -plane of radius  $a$ . Find the electric field at a point  $P$  on the  $z$ -axis a distance  $z$  from the origin.

## 45.2.2 Exam II

### Question I

A conducting sphere of radius  $a$ , centered at the origin carries charge  $Q_1$ . This sphere is surrounded by a hollow concentric conducting spherical shell of inner radius  $b$  and outer radius  $c$  with  $a < b < c$ . The outer hollow conducting shell carries a total charge  $Q_2$ .

1. What is the electric field everywhere?
2. What is the potential everywhere, assuming  $\lim_{r \rightarrow \infty} \phi(r) = 0$ ?
3. How much charge is on the inner and outer surfaces of the conducting shell at  $r = b$  and  $r = c$ ?

### Question II

The outer conductor of problem I is now grounded.

1. What is the electric field everywhere?
2. What is the potential everywhere?
3. How much charge is on the surfaces at  $r = b$  and  $r = c$ ?
4. What is the capacitance of the system of conductors?
5. Calculate the electrostatic potential energy of the configuration assuming the energy resides in the charges.
6. Calculate the electrostatic potential energy of the configuration assuming the energy resides in the electric field.

**Question III**

A semicircular arc of radius  $a$  in the  $y - z$  plane with center on the  $y - axis$  at the origin and the top of the arc on the positive  $y - axis$  carries linear charge density  $\lambda = \lambda_0 \cos(\theta')$ , where  $\lambda_0$  is constant and  $\theta'$  is measured with respect to the positive  $z - axis$ .

1. What is the electric monopole moment of this distribution?
2. What is the electric dipole moment of this distribution?
3. What is the electric potential at a distance  $r$  from the origin for this distribution where  $r > a$ , accurate to order  $\frac{1}{r^2}$ ?

**45.2.3 Exam III****Question I**

The electric field in a spherical region of space of radius  $a$  is given by  $\mathbf{E} = E_0 \frac{r^2}{a^2} \hat{\mathbf{r}}$  for  $r < a$ , where  $E_0$  is a constant. This region is surrounded concentrically by a grounded conducting spherical shell of inner radius  $b$  and outer radius  $c$  with  $a < b < c$ . There is no charge in the region  $a < r < b$ .

1. What is the electric charge density in the region  $r < a$ ?
2. What is the electric field in the region  $a < r < b$ ?
3. How much charge is on the surfaces at  $r = b$  and  $r = c$ ?
4. What is the electric field for  $r > c$ ?
5. What is the electric potential  $\phi$  at  $r = 0$  assuming that ground potential is  $\phi = 0$ .

**Question II**

A capacitor  $C_1$  is charged to a potential difference  $\Delta\phi$  between its plates. A second capacitor  $C_2$  is uncharged. One plate of  $C_2$  is now connected to a plate of  $C_1$  by a conductor of negligible capacitance, the remaining plates are similarly connected.

1. For the resultant equilibrium state, find the charge on each capacitor and the potential difference  $\Delta\phi$  between their respective plates.
2. Compare the energy stored in capacitor  $C_1$  before connecting it to  $C_2$ , to the energy of the combination after connecting them. Are these energies the same? If not, which is larger and where did any additional energy come from, or where did any lost energy go?

**Question III**

Find the electric dipole moment of an hourglass configuration of charge consisting of two identical right circular cones of radius  $a$  and height  $a$  with symmetry axes aligned apex to apex along the  $z$ -axis with the apexes touching at the origin. The top cone has charge density  $\sigma_0$  on its surface, and the bottom cone has charge density  $-\sigma_0$  on its surface.

**45.2.4 Practice Final Exam****Problem I**

The electric field in a region of space is given in spherical coordinates as  $\mathbf{E} = cr\hat{\mathbf{r}}$ , where  $c$  is constant.

1. Find the charge density at a point  $(r, \theta, \phi)$
2. Find the total charge inside a sphere of radius  $a$  centered at the origin.

**Problem II**

A battery is used to charge an ideal parallel plate capacitor to a potential difference  $\Delta\phi = V_0$ . The battery is then disconnected. The separation between the plates is now increasing from  $d$  to  $\alpha d$ , where  $\alpha > 1$ . The area of the plates is  $A$ .

1. What is the ratio of the new energy to the original energy?
2. Is the energy increases or decreased?
3. Where does this energy come from or go to?
4. Compute the change in energy  $\Delta U_e$  expressing your answer in terms of the given quantities  $V_0, d, A, \alpha$  and fundamental constants.

**Problem III**

A dielectric sphere of radius  $a$  and permittivity  $\epsilon$  contains a free charge density  $\rho_f = cr$ , where  $c$  is a constant. The sphere is centered at the origin. Find the electric potential at the center of the sphere assuming that the potential is zero at an infinite distance from the center.

**Problem IV**

A thick slab extending from  $z = -a$  to  $z = a$  carries a uniform volume current density  $\mathbf{J} = J_0\hat{\mathbf{x}}$ . The slab is infinite in the  $xy$ -plane. Find the magnetic field  $B$  as a function of  $z$  inside and outside the slab. Plot  $B$  as a function of  $z$  for  $-b < z < b$  where  $b > a$ .

**Problem V**

An ideal long solenoid of radius  $a$ , carrying  $n$  turns per unit length, is looped by a wire with resistance  $R$ .

- If the current in the solenoid is increasing at a constant rate  $\frac{dI}{dt} = k$ , what current flows in the loop, and which way (Left or right) does it pass through the resistor?
- If the current  $I$  in the solenoid is constant but the solenoid is pulled out of the loop and reinserted in the opposite direction, what total charge passes through the resistor?

**45.2.5 Final Exam****Problem I**

- Write down Maxwell's Equations in differential form.
- Convert them to integral form and show derivations.
- Name each equation.

*Solution.*

- Gauss' Law:  $\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \Rightarrow \frac{Q_{encl}}{\epsilon_0} = \iiint_V \frac{\rho}{\epsilon_0} d\tau = \iiint_V (\nabla \cdot \mathbf{E}) d\tau = \oint_{\partial V} \mathbf{E} \cdot d\mathbf{a}$
- Faraday's Law:  $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \Rightarrow -\frac{d\Phi_B}{dt} = \iint_S -\frac{\partial \mathbf{B}}{\partial t} da = \iint_S (\nabla \times \mathbf{E}) da = \oint_{\partial S} \mathbf{E} \cdot d\ell$
- Gauss' Law of Magnetism:  $\nabla \cdot \mathbf{B} = 0 \Rightarrow 0 = \iiint_V (\nabla \cdot \mathbf{B}) d\tau = \oint_{\partial V} \mathbf{B} \cdot d\mathbf{a}$
- Ampere's Law:  $\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \Rightarrow \mu_0 I_{encl} + \mu_0 \epsilon_0 \frac{d\Phi_E}{dt} = \iint_S (\mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}) da = \iint_S (\nabla \times \mathbf{B}) da = \oint_{\partial S} \mathbf{B} \cdot d\ell$

□

**Problem II**

A conduction sphere of radius  $a$  carries a charge  $Q_1$ . It is surrounded by a conducting spherical shell of inner radius  $b$  and outer radius  $c$  with  $a < b < c$ . The charge on the conducting shell is  $Q_2$ . The region between the conductors  $a < r < b$  is filled with linear isotropic dielectric of permittivity  $\epsilon$ . Find the following in the regions  $r < a$ .  $a < r < b$ .  $b < r < c$ .  $c < r$ :

- The electric displacement  $\mathbf{D}$

2. The electric field  $\mathbf{E}$
3. The polarization vector  $\mathbf{P}$
4. Find the free charge on the conductors at  $r = a, b, c$ .
5. The bound volume charge in the dielectric.
6. The bound surface charge density at the inner and outer surfaces of the dielectric.
7. The electric potential at the origin assuming the potential is zero as  $r$  goes to infinity.

*Solution.*

$$1. \mathbf{D} = \begin{cases} \mathbf{0} & r < a \\ \frac{Q_1}{4\pi r^2} & a < r < b \\ \mathbf{0} & b < r < c \\ \frac{Q_1+Q_2}{4\pi r^2} & c < r \end{cases}$$

$$3. \mathbf{P} = \begin{cases} \mathbf{0} & r < a \\ \frac{Q_1}{4\pi r^2}(1 - \frac{\epsilon_0}{\epsilon}) & a < r < b \\ \mathbf{0} & b < r < c \\ \mathbf{0} & c < r \end{cases}$$

$$2. \mathbf{E} = \begin{cases} \mathbf{0} & r < a \\ \frac{Q_1}{4\pi\epsilon_0 r^2} & a < r < b \\ \mathbf{0} & b < r < c \\ \frac{Q_1+Q_2}{4\pi\epsilon_0 r^2} & c < r \end{cases}$$

$$4. Q = \begin{cases} Q_1 & r = a \\ -Q_1 & r = b \\ Q_1 + Q_2 & r = c \end{cases}$$

$$5. \rho_b = \nabla \cdot \mathbf{P}, \rho_b = 0.$$

$$6. \sigma_{b,a} = -\frac{Q_1}{4\pi a^2} \left(1 - \frac{\epsilon_0}{\epsilon}\right), \sigma_{b,b} = \frac{Q_1}{4\pi b^2} \left(1 - \frac{\epsilon_0}{\epsilon}\right).$$

$$7. \phi = -\int_0^\infty \mathbf{E} \cdot d\ell = \int_c^\infty \mathbf{E} \cdot d\ell + \int_b^c \mathbf{E} \cdot d\ell + \int_a^b \mathbf{E} \cdot d\ell + \int_0^a \mathbf{E} \cdot d\ell = \frac{Q_1+Q_2}{4\pi\epsilon_0 c} + \frac{Q_1}{4\pi\epsilon_0 a} - \frac{Q_1}{4\pi\epsilon_0 b}.$$

□

### Problem III

A sphere of radius  $a$  carries charge density  $\rho = \rho_0(r/a)$ , where  $\rho_0$  is a constant. Find the work done to assemble the charge distribution.

*Solution.* We find  $\mathbf{E}$  inside and outside using Gauss' law.

$$\iint_S \mathbf{E} \cdot d\mathbf{a} = \frac{Q_{encl}}{\epsilon_0} = \int_0^r \int_0^\pi \int_0^{2\pi} \rho_0 \frac{r}{a} r^2 \sin(\theta) d\varphi d\theta dr = \frac{4\pi\rho_0}{a\epsilon_0} \frac{r^4}{4} = E(4\pi r^2).$$

So  $\mathbf{E} = \frac{\rho_0 r^2}{4a\epsilon_0} \hat{\mathbf{r}}$ . Outside we have  $\oint_S \mathbf{E} \cdot d\mathbf{a} = \int_0^{2\pi} \int_0^\pi \int_0^a \rho \frac{r}{a} r^2 \sin(\theta) dr d\theta d\varphi$ , so  $\mathbf{E} = \frac{\rho_0 a^3}{4r^2 \epsilon_0} \hat{\mathbf{r}}$ . The work is

$$\begin{aligned} \frac{\epsilon_0}{2} \int_{All\ Space} E^2 d\tau &= \frac{\epsilon_0}{2} \int_0^{2\pi} \int_0^\pi \int_0^\infty E^2 r^2 \sin(\theta) dr d\theta d\varphi \\ &= \int_0^{2\pi} \int_0^\pi \int_0^a E^2 r^2 \sin(\theta) dr d\theta d\varphi + \int_0^{2\pi} \int_0^\pi \int_a^\infty E^2 r^2 \sin(\theta) dr d\theta d\varphi \\ &= \frac{\pi \rho_0^2 a^5}{7\epsilon_0} \end{aligned}$$

So  $\mathbf{E} = \frac{\rho_0 r^2}{4a\epsilon_0} \hat{\mathbf{r}}$

□

### Problem IV

1. Could the vector field  $\mathbf{F} = ax\hat{\mathbf{x}} + by\hat{\mathbf{y}} + cz\hat{\mathbf{z}}$  be a possible magnetic field, where  $a + b + c \neq 0$ ? Explain why or why not.
2. An electric field is given by  $\mathbf{E} = ax\hat{\mathbf{y}}$ , where  $a$  is a constant. Is this a conservative field? Explain why or why not.
3. Find the possible magnetic field  $\mathbf{B}$  associated to  $\mathbf{E}$ .

*Solution.*

1. No, for  $\nabla \cdot \mathbf{F} = a + b + c \neq 0$ , and therefore  $\mathbf{F}$  cannot be a magnetic field.
2. No, for  $\nabla \times \mathbf{E} = a\hat{\mathbf{z}} \neq 0$ , and thus  $\mathbf{E}$  is not a conservative field.
3.  $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = a\hat{\mathbf{z}}$ , so  $\mathbf{B} = -at\hat{\mathbf{z}} + \mathbf{B}_0$ , where  $\mathbf{B}_0$  is some constant vector. Here  $\mathbf{B}$  is increasing with time in the  $-z$  direction.

□

### Problem V

Two infinitely long coaxial cylindrical infinitesimally thin conducting shells concentric with the  $z$ -axis carry oppositely directed currents of equal magnitude in the  $+$  and  $-z$ -directions. The radius of the inner shell is  $a$  and that of the outer shell is  $b$ . What is the self-inductance of a length  $\ell$  of this system?

*Solution.* The flux carried by the inner shell cuts through the area of a rectangle of length  $\ell$  and width  $b - a$ . So  $\Phi = \int \mathbf{B} \cdot d\mathbf{a} = \int_a^b \frac{\mu_0 I}{2\pi\rho} \ell d\rho = \frac{\mu_0 I \ell}{2\pi} \ln\left(\frac{b}{a}\right)$ . So,  $L = \frac{\Phi}{I} = \frac{\mu_0 \ell}{2\pi} \ln\left(\frac{b}{a}\right)$ .

□



# CHAPTER 46

---

## Electromagnetism II

---

### 46.1 Homework Sets

#### 46.1.1 Homework I

Wangness Chapter 8 - Problems: 2, 4, 7, 8

**Problem 46.1.1** Given a single point charge  $q$  located at the point  $(a, b, c)$ , Find  $Q$ ,  $\mathbf{p}$ , and all components of  $Q_{jk}$  for this system. Which coordinates are changed there is a charge  $-q$  at the origin?

*Solution* The total charge is:

$$Q = \sum_{n=1}^N q_n \quad (46.1.1)$$

The only charge in this system is  $q$ , so we have  $Q = q$ . Next,  $\mathbf{p}$  is defined as:

$$\mathbf{p} = \sum_{n=1}^N q_n \mathbf{r}_n \quad (46.1.2)$$

Using this, we have:

$$\mathbf{p} = q(a\hat{\mathbf{x}} + b\hat{\mathbf{y}} + c\hat{\mathbf{z}}) \quad (46.1.3)$$

Finally, the quadrupole moments are defined as:

$$Q_{jk} = \sum_{i=1}^N q_i (3x_i y_i - r_i^2 \delta_{jk}) \quad (46.1.4)$$

Using this, we obtain the following table:

$Q_{ij}$	x	y	z
x	$q(3a^2 - (a^2 + b^2 + c^2))$	$3qab$	$3qac$
y	$3qab$	$q(3b^2 - (a^2 + b^2 + c^2))$	$3qbc$
z	$3qac$	$3qbc$	$q(3c^2 - (a^2 + b^2 + c^2))$

Table 46.1: Quadrupole Moments for Problem 46.1.1.

Since both the dipole and quadrupole moments are weighted by the coordinates of the charges, adding a charge to the original does not affect these. However, the total charge will now be zero.

**Problem 46.1.2** Compute  $\phi$  using the multi-pole expansion for a charge  $2q$  at  $(0, \ell, 0)$  and charges  $-q$  located at  $(-a, 0, 0)$  and  $(a, 0, 0)$ .

*Solution* The multi-pole expansion of  $\phi$  is:

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \left( \frac{Q}{r} + \frac{\mathbf{P} \cdot \hat{\mathbf{r}}}{r^2} + \frac{1}{2r^3} \sum_{j,k} \ell_j \ell_k Q_{kj} + \dots \right) \quad (46.1.5)$$

For this system, we have the following:

$$Q = 0 \quad (46.1.6a)$$

$$\mathbf{P} = 2q\ell\hat{\mathbf{y}} \quad (46.1.6b)$$

The dipole portion of the potential is then:

$$\phi_{\mathbf{P}} = \frac{2q\ell y}{r^3} \quad (46.1.7)$$

Computing the quadrupole moments, we obtain the table below:

$Q_{ij}$	x	y	z
x	$-2q(\ell^2 + 2a^2)$	0	0
y	0	$2q(2\ell^2 + a^2)$	0
z	0	0	$2q(-\ell^2 + a^2)$

Table 46.2: Quadrupole Moment for Problem 46.1.

The Quadrupole portion of  $\phi$  is then:

$$\phi_Q = \frac{q}{r^5} \left[ -x^2(\ell^2 + 2a^2) + y^2(2\ell^2 + a^2) + z^2(-\ell^2 + a^2) \right] \quad (46.1.8)$$

Recalling that  $r^2 = x^2 + y^2 + z^2$ , and using everything up to the quadrupole moment, we obtain the following approximation for  $\phi$ :

$$\phi(r) \approx \frac{q}{2\pi\epsilon_0 r^3} \left( \ell y + \frac{1}{2r^2} a [\ell^2(3y^2 - r^2) - a^2(3x^2 - r^2)] \right) \quad (46.1.9)$$

**Problem 46.1.3** Given a line of charge of length  $L$  with constant charge density  $\lambda$  that lies in the first quadrant of the  $xy$  plane with one end at the origin and making an angle  $\alpha$  with the  $x$  axis, find  $Q$ ,  $\mathbf{p}$ , and all components of  $Q_{jk}$ .

*Solution* The total charge  $Q$  is:

$$Q = \int_C \lambda d\ell \quad (46.1.10)$$

Since the path  $C$  is a line, and the density  $\lambda$  is a constant, we can integrate this to obtain:

$$Q = \int_0^L \lambda dr' = \lambda L \quad (46.1.11)$$

$$\mathbf{p} = \int_C \lambda \mathbf{r} d\ell \quad (46.1.12)$$

For this distribution of charge, we have:

$$\mathbf{p} = \lambda \int_0^L r' \hat{\mathbf{r}}' dr' \quad (46.1.13a)$$

$$= \lambda \int_0^L r' (\cos(\alpha) \hat{\mathbf{x}} + \sin(\alpha) \hat{\mathbf{y}}) dr' \quad (46.1.13b)$$

$$= \frac{\lambda L^2}{2} (\cos(\alpha) \hat{\mathbf{x}} + \sin(\alpha) \hat{\mathbf{y}}) \quad (46.1.13c)$$

The quadrupole components can be obtained by:

$$Q_{jk} = \int_C \lambda(\mathbf{r}) (3x_j x_k - r^2 \delta_{jk}) d\ell \quad (46.1.14)$$

Using this, we obtain the following table:

$Q_{jk}$	$x$	$y$	$z$
$x$	$\frac{\lambda L^3}{3} (3 \cos^2(\alpha) - 1)$	$\lambda L^3 \cos(\alpha) \sin(\alpha)$	0
$y$	$\lambda L^3 \cos(\alpha) \sin(\alpha)$	$\frac{\lambda L^3}{3} (3 \sin^2(\alpha) - 1)$	0
$z$	0	0	$-\frac{\lambda L^3}{3}$

Table 46.3: Caption

Writing  $\mathbf{r}$  using Cartesian unit vectors, we may then approximate  $\phi$  as:

$$\begin{aligned} \phi(r) &= \frac{\lambda L}{4\pi\epsilon_0 r} \left[ 1 + \frac{L}{2r} (\cos(\alpha) \cos(\theta) + \sin(\alpha) \sin(\theta)) \right. \\ &\quad \left. + \frac{L^2}{6r^4} \left( 3(x \cos(\alpha) + y \sin(\alpha))^2 - r^2 \right) \right] \end{aligned} \quad (46.1.15)$$

**Problem 46.1.4** Given a sphere of radius  $a$  with a surface charge density  $\sigma = \sigma_0 \cos(\theta)$ , where  $\sigma_0$  is a constant, and such that the center of the sphere lies at the origin, calculate  $Q$ ,  $\mathbf{p}$ , and all of the coordinates of  $Q_{jk}$ .

*Solution* The total charge distributed on a surface  $S$  is:

$$Q = \iint_S \sigma \, da \quad (46.1.16)$$

Using this, we have:

$$Q = \int_0^{2\pi} \int_0^\pi \sigma_0 \cos(\theta) a^2 \sin(\theta) \, d\theta \, d\phi = 2\pi\sigma_0 \int_0^\pi \sin(\theta) \cos(\theta) \, d\theta = 0 \quad (46.1.17)$$

The dipole moment can be obtained by:

$$\mathbf{p} = \iint_S \sigma \mathbf{r} \cdot da \quad (46.1.18)$$

From this, we have:

$$\mathbf{p} = \int_0^{2\pi} \int_0^\pi \sigma_0 \cos(\theta) \mathbf{r} a^2 \sin(\theta) \, d\theta \, d\phi \quad (46.1.19a)$$

$$p_x = \sigma_0 a^3 \int_0^{2\pi} \int_0^\pi \cos(\theta) \sin^2(\theta) \cos(\phi) \, d\theta \, d\phi = 0 \quad (46.1.19b)$$

$$p_y = \sigma_0 a^3 \int_0^{2\pi} \int_0^\pi \cos(\theta) \sin^2(\theta) \sin(\phi) \, d\theta \, d\phi = 0 \quad (46.1.19c)$$

$$p_z = \sigma_0 a^3 \int_0^{2\pi} \int_0^\pi \cos^2(\theta) \sin(\theta) \, d\theta \, d\phi = \frac{4\pi\sigma_0 a^3}{3} \quad (46.1.19d)$$

Thus the dipole moment is:

$$\mathbf{p} = \frac{4\pi\sigma_0 a^3}{3} \hat{\mathbf{z}} \quad (46.1.20)$$

Finally, the quadrupole moments can be obtained from:

$$Q_{jk} = \int_S \sigma (3x_j x_k - r^2 \delta_{jk}) \, da \quad (46.1.21)$$

The charge distribution has axial symmetry.

$$Q_{zz} = \int_0^{2\pi} \int_0^\pi \sigma \cos(\theta) (3z^2 - r^2) a^2 \sin(\theta) \, d\theta \, d\phi \quad (46.1.22a)$$

$$= 2\pi\sigma_0 a^4 \left[ 3 \int_0^\pi \cos^3(\theta) \sin(\theta) \, d\theta - \int_0^\pi \cos(\theta) \sin(\theta) \, d\theta \right] \quad (46.1.22b)$$

$$= 0 \quad (46.1.22c)$$

We can approximate  $\phi$  as:

$$\phi = \frac{1}{2\pi\epsilon_0} \frac{\mathbf{p} \cdot \hat{\mathbf{r}}}{r^2} = \frac{\sigma_0 a^3}{3\epsilon_0 r^2} \cos(\theta) \quad (46.1.23)$$

**Problem 46.1.5** The nucleus of an atom can be approximated as a uniform distribution of positive charge  $ne$ , where  $n$  is the number of protons in the nucleus and  $e = 1.6 \times 10^{-19} C$ . Certain nuclei like  $^{208}\text{Pb}$  are spherical in shape, and others like  $^{184}\text{W}$  are ellipsoidal. Consider  $^{184}\text{W}$  to be an ellipsoidal nucleus with uniform charge density  $\rho$  and total charge  $74e$ . The equation of an ellipsoid is:

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{b^2} = 1 \quad (46.1.24)$$

Where  $z$  is the symmetry axis, and  $a$  and  $b$  are the semi-major and semi-minor axes, respectively. Let  $a = 6.85 \times 10^{-15} \text{ m}$  and  $b = 5.570 \times 10^{-15} \text{ m}$ . Calculate the quadrupole moment of this nucleus in units of  $e$ .

*Solution* We are given that  $\rho_c$  is a constant and  $Q = 74e$  is the total charge. From axial symmetry, the quadrupole moment is:

$$Q^a = \iiint_V \rho_c (3z^2 - r^2) d\tau \quad (46.1.25)$$

We can rewrite this in cylindrical coordinates, noting that  $r^2 = \rho^2 + z^2$ , to obtain:

$$Q^a = \int_{-a}^a \int_0^{2\pi} \int_0^{b\sqrt{1-z^2/a^2}} \rho_c (2z^2 - \rho^2) \rho d\rho d\phi dz \quad (46.1.26a)$$

$$= 2\pi\rho_c \int_{-a}^a \int_0^{b\sqrt{1-z^2/a^2}} (2z^2\rho - \rho^3) d\rho dz \quad (46.1.26b)$$

$$= 2\pi\rho_c \int_{-a}^a \left[ z^2\rho^2 - \frac{\rho^4}{4} \right]_0^{b\sqrt{1-z^2/a^2}} dz \quad (46.1.26c)$$

$$= 2\pi\rho_c \int_{-a}^a \left[ z^2b^2 \left( 1 - \frac{z^2}{a^2} \right) - \frac{b^4}{4} \left( 1 - \frac{z^2}{a^2} \right)^2 \right] dz \quad (46.1.26d)$$

$$= \pi\rho_c b^2 \int_0^a \left[ 4z^2 - 4\frac{z^4}{a^2} - b^2 + \frac{2b^2 z^2}{a^2} - \frac{b^2 z^4}{a^4} \right] dz \quad (46.1.26e)$$

$$= \pi\rho_c b^2 \left[ \frac{4}{3}a^3 - \frac{4}{5}a^3 - ab^2 + \frac{2ab^2}{3} - \frac{ab^2}{5} \right] \quad (46.1.26f)$$

$$= \frac{8\pi\rho_c ab^2}{15} (a^2 - b^2) \quad (46.1.26g)$$

But we know that  $\rho_c = Q/V$ , where  $Q = 74e$  and  $V$  is the volume of an ellipsoid:

$$V = \frac{4\pi}{3}ab^2 \quad (46.1.27)$$

And thus we have:

$$Q^a = \frac{2}{5}(74e)(a^2 - b^2) \quad (46.1.28)$$

Using the numerical values for  $e$ ,  $a$ , and  $b$ , we obtain  $Q^a = 4.27 \times 10^{-28}e$

### 46.1.2 Homework II

Wangness Chapter 11 - Problems 3, 4, 8, 9

**Problem 46.1.6** Given a point charge  $q$  in the distribution shown in Fig. 46.1, where  $q$  lies in the  $xy$  plane near two intersecting planes which intersect at a right angle, and such that the  $z$  axis is the line of intersection, find and justify image charges that will give the potential  $\phi$  for all points in the first quadrant  $x, y \geq 0$ . Calculate  $E$  and  $\sigma_f$ .

*Solution* The configuration of the images makes both planes equipotentials with  $\phi = 0$ . So, if  $P_0$  is a point on the  $x$  axis, we have:

$$\phi_{P_0} = \frac{1}{4\pi\epsilon_0} \left[ \frac{q}{[(x-a)^2 + b^2]^{1/2}} - \frac{q}{[(x-a)^2 + b^2]^{1/2}} - \frac{q}{[(x+a)^2 + b^2]^{1/2}} + \frac{q}{[(x+a)^2 + b^2]^{1/2}} \right] \quad (46.1.29)$$

And this evaluates to zero on the entire  $x$  axis. If  $P_0$  lies in the  $xz$  plane, but not on the  $x$  axis, then there would be an additional  $z^2$  term under the square root in each denominator. So the sum would still evaluate to zero. Similarly for a point in the  $yz$  plane. Thus  $\phi$  evaluates to zero on the  $xz$  and the  $yz$  planes. The potential at a point  $P = (x, y, z)$  in the first octant is:

$$\phi_{P_0} = \frac{1}{4\pi\epsilon_0} \left[ \frac{q}{[(x-a)^2 + (y-b)^2 + z^2]^{1/2}} - \frac{q}{[(x-a)^2 + (y+b)^2 + z^2]^{1/2}} + \frac{q}{[(x+a)^2 + (y+b)^2 + z^2]^{1/2}} - \frac{q}{[(x+a)^2 + (y-b)^2 + z^2]^{1/2}} \right] \quad (46.1.30)$$

To compute the electric field, we take the gradient and negate it:

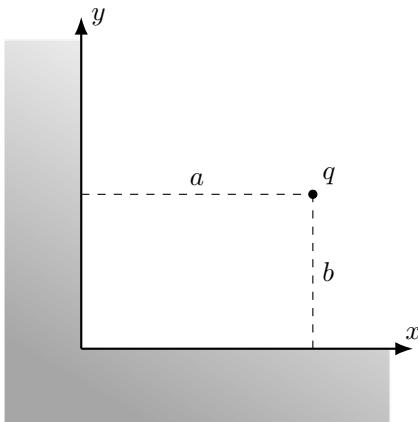
$$\mathbf{E} = -\nabla(\phi) \quad (46.1.31)$$

Computing the components, we have:

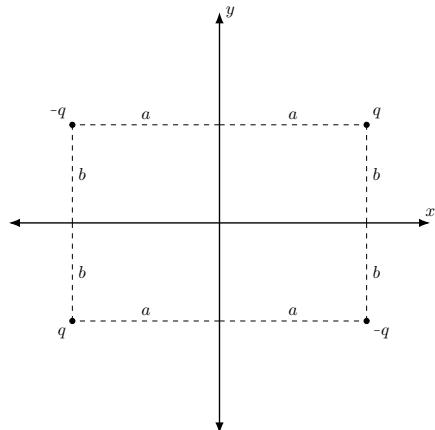
$$\begin{aligned} E_y = \frac{-1}{4\pi\epsilon_0} & \left[ \frac{q(y-b)}{[(x-a)^2 + (y-b)^2 + z^2]^{3/2}} - \frac{q(y+b)}{[(x-a)^2 + (y+b)^2 + z^2]^{3/2}} + \right. \\ & \left. \frac{q(y+b)}{[(x+a)^2 + (y+b)^2 + z^2]^{3/2}} - \frac{q(y-b)}{[(x+a)^2 + (y-b)^2 + z^2]^{3/2}} \right] \end{aligned} \quad (46.1.32)$$

From this, we see that in the  $yz$  plane that  $\mathbf{E} = \mathbf{0}$ . In the  $xz$  plane we have:

$$\sigma_f = \frac{qb}{2\pi} \left[ \frac{1}{[(x+a)^2 + b^2 + z^2]^{3/2}} - \frac{1}{[(x-a)^2 + b^2 + z^2]^{3/2}} \right] \quad (46.1.33)$$



46.1.1: Configuration of the Problem.



46.1.2: Location of Image Charges.

Fig. 46.1: Figures for Problem 46.1.6.

**Problem 46.1.7** Suppose that the angle between two conducting planes is  $60^\circ$  and that a charge  $q$  lies on the angle bisector of the two planes. Find the image charges that will compute  $\phi$  in the region containing  $q$ . What is the direction of the force on  $q$ ?

*Solution* Consider the distribution shown in Fig. 46.2. From the geometry, the potential on the lines through the origin at 0 and  $60$  degrees is thus zero, so we can use this to compute  $\phi$  for the two planes. Using this, we see that the net force will be towards the origin. In general, if we have two plates that are an angle  $\theta = \pi/n$ , where  $n$  is a positive integer, we will need  $2n - 1$  image charges.

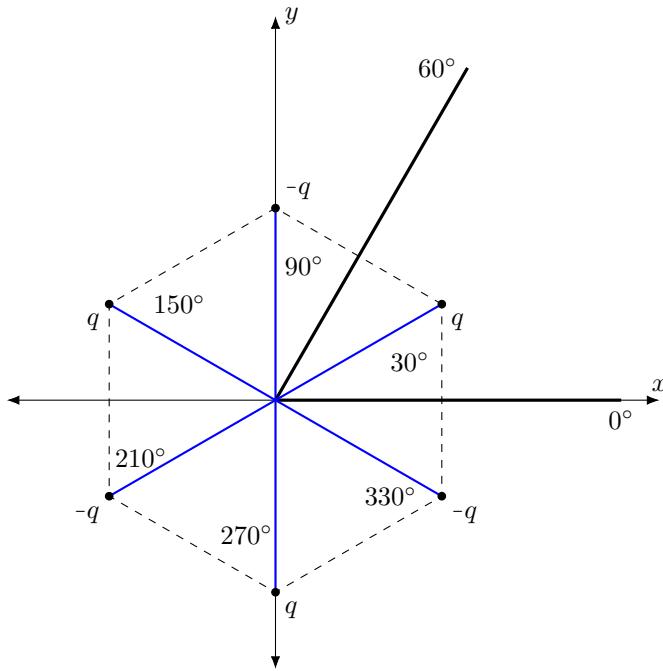


Fig. 46.2: Solution to Problem 46.1.7.

**Problem 46.1.8** Consider the geometry of Fig. 46.3. Find  $\phi$  at all points  $P$  outside of the sphere given that  $q$  lies a distance  $d$  from the origin and  $q' = -(a/d)q$  lies a distance  $a^2/d$ , and  $q'' = -q'$ . Find  $\mathbf{E}$  and the force on  $q$ , and calculate  $\sigma_f$ .

*Solution* The potential can be obtained from the usual rule of superposition:

$$\phi_P = \frac{1}{4\pi\epsilon_0} \left[ \frac{q}{R} - \frac{a}{d} \frac{q}{R'} + \frac{a}{d} \frac{q}{r} \right] \quad (46.1.34)$$

$$= \frac{q}{4\pi\epsilon_0} \left[ \frac{1}{[r^2 + d^2 - 2rd \cos(\theta)]^{1/2}} + \frac{ad^{-1}}{[r^2 + \frac{a^4}{d^2} - 2\frac{ra^2}{d} \cos(\theta)]^{1/2}} + \frac{ad^{-1}}{r} \right] \quad (46.1.35)$$

Evaluating at  $r = a$ , we have:

$$\phi = \frac{q}{4\pi\epsilon_0 d} \quad (46.1.36)$$

As expected. To find  $\mathbf{E}$  we take the gradient. Thus:

$$\mathbf{E} = -\nabla(\phi) = -\left[ \frac{\partial\phi}{\partial r} \hat{\mathbf{r}} + \frac{1}{r} \frac{\partial\phi}{\partial\theta} \hat{\theta} \right] \quad (46.1.37)$$

The components are:

$$E_r = \frac{q}{4\pi\epsilon_0} \left[ \frac{r - d \cos(\theta)}{R^3} - \frac{ad^{-1}[r - a^2d^{-1}\cos(\theta)]}{R'^3} \right] \quad (46.1.38)$$

$$E_\theta = \frac{qd\sin(\theta)}{4\pi\epsilon_0} \left[ \frac{1}{R^3} - \frac{a^3d^{-3}}{R'^3} \right] \quad (46.1.39)$$

The charge density is then:

$$\sigma_f = \epsilon_0 E_r \Big|_{r=a} = \frac{q}{4\pi} \left[ \frac{a^2 - d^2}{a[a^2 + d^2 - 2ad\cos\theta]^{3/2}} + \frac{1}{ad} \right] \quad (46.1.40)$$

The total charge on the sphere is just the integral over the surface of the sphere:

$$Q_f = \iint_S \sigma_f \, da = \int_0^{2\pi} \int_0^\pi \sigma_f a^2 \sin(\theta) \, d\theta \, d\phi = 0 \quad (46.1.41)$$

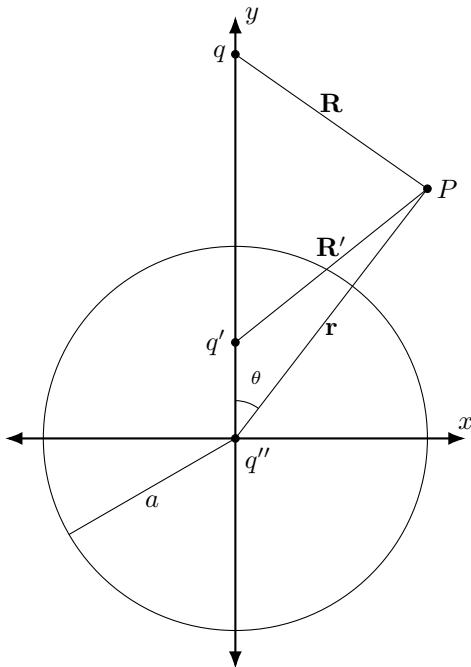


Fig. 46.3: Figure for Problem 46.1.8.

**Problem 46.1.9** Given a sphere of radius  $a$  concentric with the origin that is insulated and contains a total charge  $Q$  on it, and given a point charge  $q$  on the  $z$  axis a distance  $d$  from the origin, find  $\phi$  on the sphere and the force on  $q$ .

*Solution* By placing a point charge  $q' = Q$  at  $a^2 d^{-1}$ , the surface of the sphere will be equipotential with  $\phi = 0$ . We then place  $q''$  at the origin to bring  $\phi$  to the proper value. We have:

$$\phi_C = \frac{q}{4\pi\epsilon_0 d} + \frac{Q}{4\pi\epsilon_0 a} = \frac{q''}{4\pi\epsilon_0 a} \quad (46.1.42)$$

From this we obtain:

$$q'' = \frac{a}{d} q + Q \quad (46.1.43)$$

The force on  $q$  is then:

$$\mathbf{F} = \frac{-ad^{-1}q^2}{4\pi\epsilon_0(d - a^2d^{-1})^2}\hat{\mathbf{z}} + \frac{ad^{-1}q^2 + qQ}{4\pi\epsilon_0 d^2}\hat{\mathbf{z}} \quad (46.1.44)$$

### 46.1.3 Homework III

Wangness Chapter 11 - Problems 15, 23, 24, Bonus

**Problem 46.1.10** Given two semi-infinite conducting planes parallel to the  $yz$  plane, as shown in Fig. 46.4, find the surface charge density on the face of  $x = 0$ .

*Solution* The  $x$  component of  $\mathbf{E}$  is:

$$E_x = -\phi_0 \frac{4}{L} \sum_{n=1}^{\infty} \cos\left(\frac{(2n-1)\pi x}{L}\right) \exp\left(-\frac{(2n-1)\pi y}{L}\right) \quad (46.1.45)$$

From the definition of  $\sigma_f$ , we have:

$$\sigma_f = -\epsilon_0 \phi_0 \frac{4}{L} \sum_{n=1}^{\infty} \cos\left(\frac{(2n-1)\pi x}{L}\right) \exp\left(-\frac{(2n-1)\pi y}{L}\right) \quad (46.1.46)$$

Evaluating at zero, we get:

$$\sigma_f = -\epsilon_0 \phi_0 \frac{4}{L} \sum_{n=1}^{\infty} \exp\left(-\frac{(2n-1)\pi y}{L}\right) \quad (46.1.47a)$$

$$= -\epsilon_0 \phi_0 \frac{4}{L} \exp\left(\frac{\pi y}{L}\right) \sum_{n=1}^{\infty} \exp\left(-\frac{2\pi y}{L}\right)^n \quad (46.1.47b)$$

And this is just a geometric series, the argument of which is bounded by 1 for all  $y > 0$ . The formula for a geometric series is:

$$\sum_{n=1}^{\infty} x^n = \frac{1}{1-x} \quad (46.1.48)$$

Using this, we get:

$$\sigma_f = -\epsilon_0 \phi_0 \frac{4}{L} \exp\left(\frac{\pi y}{L}\right) \left( \frac{1}{1 - \exp(2\pi y/L)} \right) \quad (46.1.49)$$

Simplifying the exponential, and using the definition of the hyperbolic sine function  $\sinh$ , we get:

$$\sigma_f = \frac{-2\epsilon_0 \phi_0}{L \sinh(\pi y/L)} \quad (46.1.50)$$

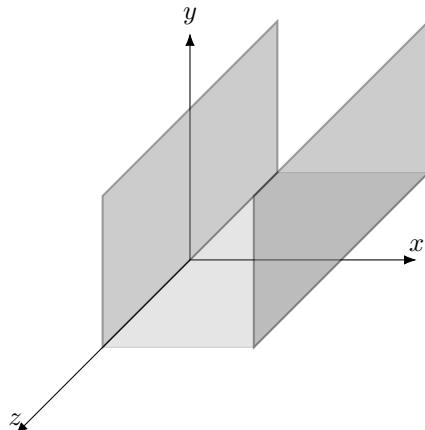


Fig. 46.4: Diagram for Problem 46.1.10.



---

## Acronyms

---

**AU** Astronomical Unit *Glossary:* [Astronomical Unit](#)

**BSR** Bistatic Radar *Glossary:* [Bistatic Radar](#)

**CFT** Continuous Fourier Transform *Glossary:* [Continuous Fourier Transform](#)

**DC** Direct Current *Glossary:* [Direct Current](#)

**DSN** Deep Space Network *Glossary:* [Deep Space Network](#)

**ERT** Earth Received Time *Glossary:* [Earth Received Time](#)

**ESA** European Space Agency *Glossary:* [European Space Agency](#)

**ET** Ephemeris Time *Glossary:* [Ephemeris Time](#)

**FFT** Fast Fourier Transform *Glossary:* [Fast Fourier Transform](#)

**FIR** Finite Impulse Response *Glossary:* [Finite Impulse Response](#)

**FOIL** First Outside Inner Last *Glossary:* [FOIL](#)

**GPS** Global Positioning System *Glossary:* [Global Positioning System](#)

**GR** General Relativity *Glossary:* [General Relativity](#)

**GW** Gravitation Wave *Glossary:* [Gravitational Wave](#)

**GWE** Gravitation Wave Experiment *Glossary:* [Gravitational Wave Experiment](#)

**HGA** High Gain Antenna *Glossary:* [High Gain Antenna](#)

**Hz** Hertz *Glossary:* [Hertz](#)

**IF** Intermediate Frequency *Glossary:* [Intermediate Frequency](#)

**JPL** Jet Propulsion Lab *Glossary:* [Jet Propulsion Laboratory](#)

**KAT** K<sub>a</sub>-band Translator *Glossary:* [K<sub>a</sub>-band transponder](#)

**LCP** Left-hand Circularly Polarized *Glossary:*

**NAIF** Navigation Ancillary Information Facility *Glossary:* [NAIF](#)

**NASA** National Aeronautic and Space Administration *Glossary:* [NASA](#)

**OD** Orbit Determination *Glossary:* [Orbit Determination](#)

**Opemode** Operations Mode *Glossary:* [Operations Mode](#)

**PDS** Planetary Data System *Glossary:* [Planetary Data System](#)

**PEMDAS** Parenthesis Exponents Multiplication Division Addition Subtraction *Glossary:* [PEMDAS](#)

**PLL** Phase-Lock Loop *Glossary:* [Phase-Lock Loop](#)

**PSA** Planetary Science Archive *Glossary:* [Planetary Science Archive](#)

**RCP** Right-Hand Circularly Polarized *Glossary:* [Right-Hand Circulary Polarized](#)

**Rev** Revolution *Glossary:* [Rev](#)

**RF** Radio Frequency *Glossary:* [Right-Hand Circulary Polarized](#)

**RMS** Root Mean Square *Glossary:* [Root Mean Square](#)

**RS** rs *Glossary:* [Radio Science](#)

**RSR** Radio Science Receiver *Glossary:* [Radio Science Receiver](#)

**RSS** Radio Science Subsystem *Glossary:* [Radio Science Subsystem](#)

**RTG** Radioisotope Thermonuclear Generator *Glossary:* [Radioisotope Thermonuclear Generator](#)

**SCE** Solar Conjunction Experiment *Glossary:* [Solar Conjunction Experiment](#)

**SNR** Signal-to-Noise Ratio *Glossary:* [Signal-to-Noise Ratio](#)

**SPICE** Spacecraft, Planet, Instrument, C-Matrix, Events *Glossary:* [SPICE](#)

**TDB** Barycentric Dynamical Time *Glossary:* [Barycentric Dynamical Time](#)

**TLM** Telemetry *Glossary:* [Telemetry](#)

**USO** Ultra-Stable Oscillator *Glossary:* [Ultra-Stable Oscillator](#)

**UTC** Universal Time Coordinated *Glossary:* [Universal Time Coordinated](#)

**ZFC** Zermelo-Fraenkel Set Theory with Choice [37](#), [46](#)

---

## Notation

---

Complex Numbers, $\mathbb{C}$	41	Less Than, $<$	44
Conjunction, $\wedge$	27, 29, 82	Less Than or Equal, $\leq$	40, 44
Containment, $\in$	13, 27, 39, 42, 81, 82, 91	Natural Numbers, $\mathbb{N}$	36, 38, 45
Defined by, $\equiv$	66	Positive Even Integers, $\mathbb{N}_e$	43, 52, 55
Defined by, $\Leftrightarrow$	27	Positive Odd Integers, $\mathbb{N}_o$	43, 52, 55
Defined by, $\neg$	27	Power Set, $\mathcal{P}(A)$	60
Defined by, $\Rightarrow$	27	Ordered Pair, $(a, b)$	46
Disjunction, $\vee$	27, 29	Proper Subset, $\subsetneq$	44
Empty Set, $\emptyset$	38	Rational Numbers, $\mathbb{Q}$	41, 45
Equality, $=$	27, 39, 43	Real Numbers, $\mathbb{R}$	40, 41, 45, 167
Greater Than, $>$	40	Subset, $\subseteq$	27, 40, 42, 44, 82
Integers, $\mathbb{Z}$	36, 38, 45	Union of Two Sets, $A \cup B$	51
Integers Modulo $n$ , $\mathbb{Z}_n$	36, 38	Union over a Collection, $\bigcup \mathcal{O}$	50,
Intersection of Two Sets, $A \cap B$	54		51
Intersection over a Collection, $\bigcap \mathcal{O}$	56, 57		



---

## Glossary

---

**Abelian Group** A group  $(G, *)$  where  $*$  is a commutative. *See:* [group](#) & [commutative operation](#), 194, 225, 226, 239.

**Associative Operation** A binary operation  $*$  where the following holds:

$$a * (b * c) = (a * b) * c$$

*See:* [binary operation](#), 133, 167.

**Axiom** A proposition that is affirmed to be true without proof or justification. *See:* [proposition](#) & [proof](#), 20.

**Bijective Function** A function that is both injective and surjective. *See:* [function](#), [injective function](#) & [surjective function](#), 124, 125, 211, 247.

**Binary Operation** A function  $* : A \times A \rightarrow A$ . That is, it is a function from the Cartesian product of  $A$  with itself that maps into  $A$ . For a binary operation  $*$  on  $A$ , and for elements  $a, b \in A$ , we denote the image  $*(a, b)$  by writing  $a * b$ . *See:* [Cartesian product](#), [function](#) & [set](#), 131, 133, 134, 135, 136, 138, 139, 141, 142, 143, 144, 167, 175.

**Cartesian Product** A set produced by two sets  $A$  and  $B$ , defined as:

$$A \times B = \{ (a, b) : a \in A \text{ and } b \in B \}$$

Where  $(a, b)$  denotes the ordered pair of  $a$  with respect to  $b$ . *See:* [ordered pair](#) & [set](#), 64, 100.

**Commutative Monoid** A monoid  $(G, *)$  where  $*$  is commutative. *See: monoid & commutative operation, 171.*

**Commutative Operation** A binary operation  $*$  where the following holds:

$$a * b = b * a$$

*See: binary operation & set, 143, 144, 168, 194.*

**Complex Conjugate** A complex number formed by reflecting a given complex number  $z$  across the  $x$  axis. That is, if  $z = x + iy$ , the complex conjugate is given by:

$$\bar{z} = x - iy$$

*See: complex number, 661.*

**Complex Number** An element of the Euclidean plane (Also called the Cartesian plane, or just written as  $\mathbb{R}^2$ ) with the following arithmetic:

$$\begin{aligned}(a, b) + (c, d) &= (a + c, b + d) \\ (a, b) \cdot (c, d) &= (ac - bd, ad + bc)\end{aligned}$$

We often write a complex number  $z = (x, y)$  as:

$$z = x + iy$$

where  $i$  is called the *imaginary unit*. [659.](#)

**Connective** A connective is a symbol used to take one or two variables or formulas and produce a new one. There are five basic connectives:

$a \wedge b$  True if and only if  $a$  is true and  $b$  is true

$a \vee b$  True if and only if  $a$  or  $b$  or both are true

$\neg a$  True if and only if  $a$  is false

$a \Leftrightarrow b$   $a$  and  $b$  are equivalent

$a \Rightarrow b$   $a$  implies  $b$

<b>Distinct Sets</b>	Sets $A$ and $B$ such that $A \cap B = \emptyset$ . See: <a href="#">intersection of two sets</a> & <a href="#">set</a> , 55, 58.
<b>Distributive Operation</b>	A distributive operation over a binary operation $+$ on a set $A$ is a binary operation $*$ on $A$ such that, for all $a, b, c \in A$ , the following is true:
	$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$
	<i>See: <a href="#">binary operation</a> &amp; <a href="#">set</a>, 144, 225, 226.</i>
<b>Domain (Relation)</b>	The domain of a relation $R$ on a set $A$ is the set: $\text{dom}(R) = \{a \in A \mid \exists b \in A \text{ such that } aRb\}$ <i>See: <a href="#">relation</a> &amp; <a href="#">set</a>, 102.</i>
<b>Empty Set</b>	The set that contains no elements. It is denoted $\emptyset$ , and for all $x$ it is true that $x \notin \emptyset$ . See: <a href="#">set</a> , 38, 80.
<b>Equal Sets</b>	Sets $A$ and $B$ where $A \subseteq B$ and $B \subseteq A$ . See: <a href="#">subset</a> , 43.
<b>Fiber</b>	The pre-image of a single point $y \in B$ under a function $f : A \rightarrow B$ . See: <a href="#">function</a> & <a href="#">pre-image</a> , 76.
<b>Field</b>	A field is a commutative division ring. That is, a set $\mathbf{F}$ with two operations $+$ and $\cdot$ , such that $(\mathbf{F}, +)$ is an Abelian group, $(\mathbf{F}, \cdot)$ is a commutative monoid, where $+$ distributes over $\cdot$ , and such that for all $a \in \mathbf{F}$ such that $a$ is not a unital element with respect to $\cdot$ (that is, $a$ is non-zero) it is true that $a$ is invertible with respect to $\cdot$ ( $a$ has a multiplicative inverse $a^{-1}$ ). See: <a href="#">Abelian group</a> , <a href="#">commutative monoid</a> , <a href="#">distributive operation</a> , <a href="#">invertible element</a> & <a href="#">unital element</a> , 167.

**Field (Relation)** The field of a relation  $R$  on a set  $A$  is the set:

$$\text{field}(R) = \text{dom}(R) \cup \text{ran}(R)$$

Where  $\text{dom}(R)$  is the domain of  $R$  and  $\text{ran}(R)$  is the range of  $R$ . See: [relation](#), [set](#), [domain & range](#), [102](#).

**Function** A subset  $f$  of the Cartesian product of two sets  $A$  and  $B$ , denoted  $f : A \rightarrow B$  such that, for all  $x \in A$  there is a unique  $y \in Y$  such that  $(x, y) \in f$ . See: [Cartesian product](#), [subset & set](#), [68](#), [73](#), [74](#), [75](#), [76](#), [122](#), [228](#), [239](#), [246](#).

**Group** A set  $G$  and a binary operation  $*$  on  $G$  such that  $*$  is associative, contains a unital element, and such that every element of  $G$  is invertible. See: [binary operation](#), [unital element & invertible element](#), [167](#), [192](#), [193](#), [194](#), [202](#), [211](#), [215](#).

**Group Homomorphism** A function  $\varphi : G \rightarrow G'$  from a group  $(G, *)$  to a group  $(G', \circ)$  such that for all  $a, b \in G$  it is true that:

$$\varphi(a * b) = \varphi(a) \circ \varphi(b)$$

See: [group](#) & [function](#), [211](#).

**Group Isomorphism** A bijective group homomorphism from a group  $(G, *)$  to a group  $(G', \circ)$ . See: [group](#), [group homomorphism](#) & [bijective function](#), [211](#).

**Image (Of a Point)** Given a function  $f : A \rightarrow B$  and a point  $x \in A$ , the image of  $x$  under  $f$  is the unique element  $y \in B$  such that  $f(x) = y$ . See: [function](#), [73](#).

**Image (Of a Set)** A set obtained from a function  $f : A \rightarrow B$  and a subset  $S \subseteq A$ , defined as:

$$f[S] = \{ f(x) \in B : x \in S \}$$

That is, the set points in  $B$  that  $S$  maps into by the function  $f$ . See: [function](#), [74](#).

---

<b>Implication</b>	The statement <i>if P, then Q</i> , where P and Q are propositions. <i>See: proposition, 19.</i>
<b>Injective Functions</b>	A function $f : A \rightarrow B$ between two sets A and B such that for all $x, y \in A$ such that $x \neq y$ , it is true that $f(x) \neq f(y)$ . <i>See: function, 123.</i>
<b>Intersection of Two Sets</b>	The intersection of two sets A and B is the set $A \cap B$ defined by:
	$A \cap B = \{ x \mid x \in A \text{ and } x \in B \}$ 54.
<b>Intersection over a Set</b>	The intersection over a collection of sets $\mathcal{O}$ is the set:
	$\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} = \{ x \mid \text{For all } \mathcal{U} \in \mathcal{O}, x \in \mathcal{U} \}$ 57.
<b>Inverse Element</b>	An inverse element of an element $x$ in a set $X$ with respect to a binary operation $*$ on $X$ is an element $x^{-1}$ such that $x * x^{-1}$ is a unital element of $*$ . <i>See: binary operation, set &amp; unital element.</i>
<b>Invertible Element</b>	An invertible element in a set $A$ with respect to a binary operation $*$ on $A$ is an element $a \in A$ such that there exists an element $b \in A$ such that $a * b$ and $b * a$ are unital elements. <i>See: binary operation, set &amp; unital element, 167, 172.</i>

**Module**

A module on a ring  $(R, +, \cdot)$  is a set  $M$  with a binary operation  $+$  and a function  $\star : R \times M \rightarrow M$  such that  $(M, +)$  is an Abelian group and such that, for all  $r_1, r_2 \in R$  and for all  $m_1, m_2 \in M$ , the following are true:

$$\begin{aligned}r_1 \star (m_1 + m_2) &= (r_1 \star m_1) + (r_1 \star m_2) \\(r_1 + r_2) \star m_1 &= (r_1 \star m_1) + (r_2 \star m_1) \\r_1 \star (r_2 \star m_1) &= (r_1 \cdot r_2) \star m_1 \\1 \star m_1 &= m_1\end{aligned}$$

Where  $1$  is the unital element of  $R$ . We denote a module by  $(M, +, \star)$ . See: [module](#), [ring](#), [binary operation](#) & [Abelian group](#), 239, 246, 247.

**Module Homomorphism**

A function from a module  $(M_1, +_1, \star_1)$  to a module  $(M_2, +_2, \star_2)$  over a ring  $(R, +, \cdot)$  such that for all  $x, y \in M_1$ , and for all  $r \in R$ , the following are true:

$$\begin{aligned}f(x +_1 y) &= f(x) +_2 f(y) \\f(r \star_1 x) &= r \star_2 f(x)\end{aligned}$$

See: [function](#), [module](#) & [ring](#), 246, 247.

**Module Isomorphism**

A module homomorphism that is also bijective. See: [module homomorphism](#) & [bijective function](#), 247.

**Modulus**

The size, or absolute value, of a complex number. This is the Euclidean distance in the complex plane from the point  $z = x + iy$  to the origin. By Euclidean distance, we mean that it is the distance that satisfies the Pythagorean formula:

$$|z| = \sqrt{x^2 + y^2}$$

If  $\bar{z}$  denotes the complex conjugate, we can compute the modulus by  $|z| = \sqrt{z\bar{z}}$ . See: [complex number](#) & [complex conjugate](#), 663.

<b>Monoids</b>	A set $G$ with a binary operation $*$ such that $*$ is associative and contains a unital element. That is, a set $G$ that is a semigroup and contains a unital element. <i>See:</i> <a href="#">binary operation</a> , <a href="#">unital element</a> , <a href="#">associative operation</a> & <a href="#">semigroup</a> , 167, 170, 172, 226, 240.
<b>Non-Empty Set</b>	A set $A$ that contains an element. That is, there exists an $x$ such that $x \in A$ . <i>See:</i> <a href="#">set</a> , 38, 202.
<b>Ordered Pair</b>	A set produced by two elements $x$ and $y$ defined by:
	$(x, y) = \{ \{x\}, \{x, y\} \}$
	This definition is due to Kazimierz Kuratowski in 1921. An alternative definition from Norbert Wiener, put forward in 1914, defines ordered pairs as:
	$(x, y)_W = \left\{ \{ \{x\}, \emptyset \}, \{ \{y\} \} \right\}$
	Note that, in both definitions, for distinct elements $x$ and $y$ , $(x, y) \neq (y, x)$ . 48, 167, 168, 175.
<b>Permutation</b>	A bijection from from a set $A$ to itself. <i>See:</i> <a href="#">bijective function</a> & <a href="#">set</a> , 125.
<b>Power Set</b>	The set of all subsets of a give set. That is:
	$\mathcal{P}(X) = \{ A \mid A \subseteq X \}$
	<i>See:</i> <a href="#">subset</a> , 61, 144, 459.
<b>Predicate</b>	A statement or sentence that takes in a set of variable and returns either true or false. It may thus be thought of as a <i>Boolean-Valued Function</i> . <i>See:</i> <a href="#">set</a> , 17, 18, 27.
<b>Pre-Image</b>	A set obtained from a function $f : A \rightarrow B$ and a subset $S \subseteq B$ , defined as:
	$f^{-1}(S) = \{ x \in X : f(x) \in S \}$
	That is, the set of all points in $X$ that map into $S$ . <i>See:</i> <a href="#">function</a> , 75.

<b>Proof</b>	A valid and rigorous argument which affirms a proposition. <i>See:</i> <a href="#">proposition</a> , 20.
<b>Proper Subsets</b>	A subset $A$ of a set $B$ such that $A \neq B$ <i>See:</i> <a href="#">subset</a> & <a href="#">equal set</a> , 44.
<b>Proposition</b>	An evaluation of a predicate on a particular set of variables which is either affirmed or denied as true or false. <i>See:</i> <a href="#">predicate</a> & <a href="#">set</a> , 18, 19, 20.
<b>Quantifier</b>	One of the two symbols $\forall$ and $\exists$ used to denote <i>for all</i> or <i>there exists</i> , respectively. 33.
<b>Range (Relation)</b>	The range of a relation $R$ on a set $A$ is the set: $\text{ran}(R) = \{b \in A \mid \exists a \in A \text{ such that } aRb\}$ <i>See:</i> <a href="#">relation</a> & <a href="#">set</a> , 102.
<b>Relation</b>	A relation on a set $A$ is a subset of $A \times A$ . We denote elements $(x, y) \in R$ by writing $xRy$ . <i>See:</i> <a href="#">Cartesian product</a> , <a href="#">subset</a> & <a href="#">set</a> , 100, 102.
<b>Ring</b>	A set $R$ with two binary operations $+$ and $\cdot$ such that $(R, +)$ is an Abelian group, $(R, \cdot)$ is a monoid, and such that $\cdot$ distributes over $+$ . The unital element of $(R, +)$ is often denoted 0. <i>See:</i> <a href="#">Abelian group</a> , <a href="#">binary operation</a> , <a href="#">monoid</a> & <a href="#">distributive operation</a> , 225, 226, 228, 239, 240, 246, 247.
<b>Ring Endomorphism</b>	A ring homomorphism from a ring $(R, +, \cdot)$ to itself. That is, a function $f : R \rightarrow R$ such that $f$ is a ring homomorphism. <i>See:</i> <a href="#">function</a> , <a href="#">ring</a> & <a href="#">ring homomorphism</a> , 228.

**Ring Homomorphism** A function  $f : R_1 \rightarrow R_2$  from a ring  $(R_1, +, \cdot)$  to a ring  $(R_2, +', *)$  such that, for all  $x, y \in R_1$ , the following are true:

$$\begin{aligned} f(x + y) &= f(x) +' f(y) \\ f(x \cdot y) &= f(x) * f(y) \\ f(1_{R_1}) &= 1_{R_2} \end{aligned}$$

Where  $1_{R_1}$  is the unital element of  $R_1$  and  $1_{R_2}$  is the unital element of  $R_2$ . See: [function](#) & [ring](#), 228.

**Rng** A set  $R$  with two binary operations  $+$  and  $\cdot$  such that  $(R, +)$  is an Abelian group,  $(R, \cdot)$  is a semigroup, and such that  $\cdot$  distributes over  $+$ . The unital element of  $(R, +)$  is often denoted 0. The unital element of  $(R, \cdot)$  is often denoted 1. See: [Abelian group](#), [binary operation](#), [semigroup](#) & [distributive operation](#), 225, 226.

**Semigroup** A set  $G$  with a binary operation  $*$  such that  $*$  is associative. See: [binary operation](#) & [associative operation](#), 167, 170, 225.

**Sets** A collection of objects, called elements, none of which is the set itself. We denote that  $x$  is an element of a set  $A$  by writing  $x \in A$ . We write that  $x$  is not an element by writing  $x \notin A$ . The requirement that sets do not contain themselves can be rephrased by stating that for any set  $A$ , it is true that  $A \notin A$ . 13, 17, 18, 38, 40, 43, 44, 48, 55, 61, 64, 68, 74, 76, 80, 100, 102, 122, 125, 131, 133, 134, 135, 136, 138, 139, 141, 142, 143, 144, 167, 175, 458, 461, 464.

**Subgroup** A subset  $(H, *)$  of a group  $(G, *)$  such that, for all  $x, y \in H$  it is true that  $x * y \in H$ ,  $x^{-1} \in H$ , and that  $e \in H$  where  $e$  is the unital element of  $(G, *)$ . See: [group](#), [unital element](#) & [inverse element](#), 247.

**Submodule** A submodule of a module  $(M, +, \star)$  over a ring  $(R, +, \cdot)$  is a subgroup  $(N, +)$  of  $(M, +)$  such that, for all  $n \in N$  and for all  $r \in R$  it is true that  $r \star n \in N$ . See: [subgroup](#) & [module](#), 247.

**Subsets** A subset of a set  $B$  is a set  $A$  such that, for all  $x \in A$ , it is true that  $x \in B$ . We denote this by  $A \subseteq B$ . [40](#), [68](#), [74](#), [75](#), [100](#), [202](#), [458](#), [464](#).

**Surjective Function** A function  $f : A \rightarrow B$  between two sets  $A$  and  $B$  such that  $[(A)] = B$ . That is, every point in  $B$  gets mapped onto. See: [function](#), [122](#).

**Topological Space** An ordered pair  $(X, \tau)$  where  $X$  is a set and  $\tau$  is a topology on  $X$ . See: [ordered pair](#), [set](#) & [topology](#), [461](#), [464](#).

**Topology** A subset  $\tau \subseteq \mathcal{P}(X)$  of the power set of a set  $X$  with the following three properties:

$$\begin{aligned} \emptyset \in \tau, \quad X \in \tau \\ A, B \in \tau \implies A \cap B \in \tau \\ \mathcal{O} \subseteq \tau \implies \bigcup \mathcal{O} \in \tau \end{aligned}$$

See: [set](#), [power set](#), [subset](#), [union over a set](#) & [intersection of two sets](#), [458](#), [461](#).

**Union of Two Sets** The union of two sets  $A$  and  $B$  is the set  $A \cup B$  defined by:

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

[51](#).

**Union over a Set** The union over a collection of sets  $\mathcal{O}$  is the set:

$$\bigcup_{\mathcal{U} \in \mathcal{O}} \mathcal{U} = \{x \mid \text{There exists } \mathcal{U} \in \mathcal{O} \text{ with } x \in \mathcal{U}\}$$

[50](#).

**Unital Elements**

A unital element in a set  $A$  with respect to a binary operation  $*$  on  $A$  is an element  $e \in A$  such that, for all  $a \in A$  it is true that:

$$a * e = e * a = a$$

*See: [set](#) & [binary operation](#), 136, 138, 141, 144, 167, 170.*



---

## Bibliography

---

- [Cam20] *Cambridge Dictionary*. Cambridge University Press, 2020. Truth, noun.
- [Can95] Georg Cantor. Beiträge zur begründung der transfiniten mengenlehre (1). *Mathematische Annalen*, 46:481–512, 1895.  
Cantor’s original definition of a set is found on the first page. I was unable to find an English translation.
- [Eul60] Leonhard Euler. De seriebus divergentibus. *Novi Commentarii academiae scientiarum Petropolitanae*, pages 205–237, 1760.  
Euler’s sum of  $1 + 2 + 3 + 4 + \dots$  is mentioned, as are many other sums such as Grandi’s and  $1 + 1 + 1 + \dots$  using varies methods. The original is in Latin and can be found via Dartmouth College’s Euler archive (also linked via the Mathematical Association of America), though German and partial English translations exist.
- [Gal38] Galileo Galilei. *Discorsi e dimostrazioni matematiche intorno a due nuove scienze*. 1638.  
One of Galileo’s groundbreaking text in the natural sciences. It contains Galileo’s paradox on infinity.
- [Gri03] Nicholas Griffin. *The Cambridge Companion to Bertrand Russell*. 2003.  
A biography of the life and works of Bertrand Russell. A citation crediting Berry’s paradox to G. G. Berry can be found on page 63.
- [Hau27] Felix Hausdorff. *Mengenlehre*. Verlag Walter de Gruyter, 1927.

Translated into English under the name *Set Theory* in 1957 by J. R. Aumann, published by Chelsea Publishing Co. This contains Hausdorff's definition of set and a treatise on the general theory.

- [Haw02] Stephen Hawking. Gödel and the end of physics. See <http://www.hawking.org.uk/godel-and-the-end-of-physics.html>, 2002.

Transcript of a talk given by Stephen Hawking containing reflections of the consequences of Gödel's incompleteness theorems on the quest for a theory of everything.

- [HY61] John G. Hocking and Gail S. Young. *Topology*. Dover, New York, 1961.

The terminology is definitely dated and may confuse the modern reader, but it dives into a plethora of topics including homology.

- [Jac76] N. Jacobson. *Lectures in Abstract Algebra I: Basic Concepts*. Springer, New York, 1976.

Uses outdated and perhaps confusing function notation, writing  $xf$  or  $x^f$  instead of  $f(x)$ . Other than that, it dives into a good deal of material on groups and rings in a reasonably self-contained manner. It is now made obsolete by Jacobson's later effort *Basic Algebra I*.

- [Kin11] *The Holy Bible: King James Version*. 1611.

Early accounts of the liar's paradox occur in Titus 1-12 and Psalm 116-11.

- [Mer20] *Merriam-Webster Dictionary*. Merriam-Webster, 2020. Truth.

- [Mun00] James Munkres. *Topology*. Prentice-Hall, Upper Saddle River, NJ, 2000.

One of the best point-set topology books known to me, providing a plethora of examples and counterexamples and sets the reader up perfectly for a course in algebraic topology or differential geometry. The second half discusses planar topology and homotopy.

- [OED20] *Oxford English Dictionary*. Oxford University Press, 2020. true, adj., n., adv., and int.

- [Pin10] C. Pinter. *A Book of Abstract Algebra*. Dover, New York, 2010.

A very gentle introduction to groups, rings, and fields. Definitely a readable treatise on the material that should be accessible to any undergraduate.

- [Ram57] Srinivasa Ramanujan. *Notebooks Volume I*. 1957.

The method of summation for the divergent series  $1 + 2 + 3 + \dots$  can be found in Chapter VIII, page 3. It omits mention of Grandi's series as an intermediate step.

- [Tar36] Alfred Tarski. Der wahrheitsbegriff in den formalisierten sprachen. *Studia Philosophica*, 1:261–405, 1936.

This contains Tarski's Undefinability Theorem. An English translation has been provided by J. H. Woodger in 1983, titled *The Concept of Truth in Formalized Languages*.

- [vW56] Wolfgang Sartorius von Waltershausen. *Gauss: zum Gedächtnis*. S. Hirzel, 1856.

This biography of Gauss, dated a year after his death, contains the famous story of his summing an arithmetic series at a young age in a matter of moments. To the best of my knowledge it is only available in German. The story is told on pages 12 and 13.



---

# Index

---

- Abelian Group, 172  
Absolute Geometry, 21  
Antisymmetry  
  of Containment, 83  
Aristotle, 18  
Associative Law  
  of Intersections, 91  
  of Unions, 87  
Axiom  
  Euclid's Fifth, 21  
  of Choice, 24, 77, 78  
  of Contraposition, 22  
  of Extensionality, 39, 40, 42,  
    44, 47  
  of Infinity, 44, 46  
  of Modus Ponens, 18, 22  
  of Modus Tollens, 22  
  of Pairing, 47, 51  
  of Regularity, 38, 59, 83, 168  
  of Separation, 46  
  of the Empty Set, 38  
  of the Power Set, 60, 63  
  of Union, 50, 51  
  of Unrestricted  
    Comprehension, 37, 45  
Schema of Replacement, 75  
Schema of Specification, 18,  
  39, 45, 46, 52, 63  
Axiomatic Method, 13  
Berry's Paradox, 9  
Berry, G. G., 9  
Binary Operation, 80, 131, 167  
  Associative, 68, 133, 167  
  Commutative, 131, 144, 168  
  Distributive, 144  
Binomial Coefficient, 62  
Bisection Method, 4  
Bolyai, János, 5  
Bolzano, Bernard, 12, 15  
Boolean Algebra, 3, 80  
Boolean Group, 188  
Boolos, George, 10  
Bounded Lattice, 145  
Bug-Eyed Line, 901  
C Programming Language, 12  
Cancellation Law, 186  
  Left, 187  
  Right, 187  
Cantor, Georg, 12  
Cartesian Plane, 65  
Cartesian Product, 49, 60, 64, 68,  
  78, 131  
Cayley Table, 173  
Cayley's Theorem, 182, 201, 217  
Cayley, Arthur, 173  
Chaotic Topology, 459, 460  
Clopen Set, 465  
Commutative

- Semigroup, 169  
Commutative Diagram, 126  
Commutative Law  
    of Intersections, 88  
    of Unions, 84  
Complement  
    of a Set, 95  
Conclusion, 19  
Conjunction, 28  
Connective (Logic), 28  
Containment  $\in$ , 14  
Continuous Function, 457, 459,  
    466  
Contrapositive, 24  
Converse, 23  
Cross Product, 167  
Curry's Paradox, 11
- De Morgan's Laws, 88  
De Morgan, Augustus, 12  
Descartes, René, 65  
Dirichlet, Peter Gustav Lejeune,  
    10  
Disconnected Space, 465  
Discrete  
    Metric, 459  
    Topology, 459  
Discrete Topology, 459  
Duck Typing, 12
- Einstein, Albert, 9  
Element, 14  
    Notation, 14  
Elements, The (Euclid), 13, 21, 80  
Elliptic Geometry, 22  
Empty Set, 14, 38, 80  
Epimorphism  
    Group, 212  
Equality, 40  
Equivalence Relation, 83  
Euclid of Alexandria, 65  
Euclidean Geometry, 21  
Euclidean Plane, 65
- Euclidean Space, 461  
Euler, Leonhard, 5
- False, 8  
Fiber, 76  
Field, 167, 240  
Formula, 28, 69  
Fraenkel, Abraham, 38  
Function, 39, 68, 71, 115, 131  
    Bijective, 125  
    Choice Function, 78  
    Composition, 126  
    Continuous, 457, 459  
    Injective, 119  
    Permutation, 125  
    Set of All, 73  
    Surjective, 75, 122
- Functor  
    Forgetful, 168
- Gödel, Kurt, 9  
Galilei, Galileo, 15  
Galileo's Paradox, 15  
Gauss, Carl Friedrich, 4
- Geometry  
    Absolute, 21  
    Elliptic, 22  
    Euclidean, 21  
    Hyperbolic, 22
- Grandi's series, 5  
Grandi, Luigi Guido, 5  
Grelling, Kurt, 10  
Grelling-Nelson Paradox, 10
- Group, 167, 172  
    Abelian, 172, 194  
    Boolean, 201  
    Center of, 206  
    Direct Product of, 202  
    Free Group, 209  
    Generated Subgroup, 207  
    Presentation, 208  
    Subgroup, 203  
        Sylow, 39

- Groupoid, 167
- Haskell Curry, 11
- Hausdorff, Felix, 13
- Hilbert, David, 12
- Homological Algebra, 127
- Homomorphism
- Group, 212
- Hypothesis, 19
- Idempotent, 134
- Idempotent Law
- of Intersections, 90
  - of Unions, 86
- Identity Law
- of Intersections, 89
  - of Unions, 86
- Image
- of a Point, 73
  - of a Set, 74
- Implication, 19
- Definition, 19
- Index Set, 50
- Indiscrete Topology, 459
- Infinite Regress, 13
- Integer Lattice, 66
- Integers, 36, 46
- Intersection
- of Two Sets, 55
- Interval
- Closed, 73
  - Open, 49
- Invertible Element, 142, 167
- Left Invertible, 141
  - Right Invertible, 138
  - Weakly Left Invertible, 139
  - Weakly Right Invertible, 136
- Isomorphism
- Group, 212
- Jacobi Identity, 167
- Khayyám, Omar, 21
- Kleene-Rosser Paradox, 11
- Kuratowski, Kazimierz, 47, 48, 78
- Löb's Paradox, 11
- Lambert, Johann Heinrich, 22
- Latin Square, 175
- Latin Square Property, 175, 188
- of a group, 188
- Law of the Excluded Middle, 24, 37
- Liar's Paradox, 8
- Line with Two Origins, 901
- Linear Algebra, 150
- Logic
- Propositional, 3
  - Sentential, 3
- Logical Fallacy, 23
- Affirming the Consequent, 23
  - Denying the Antecedent, 25
  - Inverse Fallacy, 23
  - Masked-Man Fallacy, 26
  - of the Converse, 23
  - of the Undistributed Middle, 25
- Lycaeous, Proclus, 21
- Matrix, 150
- Metric, 461
- Metric Space, 459, 461
- Model, 21
- Module, 239
- Module Homomorphism, 246
- Module Isomorphism, 247
- Modus Ponens, 18
- Monoid, 167
- Monomorphism
- Group, 213
- Mutually Disjoint Collection, 58
- Naive Set Theory, 37
- Natural Numbers, 36, 46
- Even, 43, 46, 52
  - Odd, 43, 46, 52
- Nelson, Leonard, 10

- Open Set, 458
- Ordered  $n$  Tuple, 68
- Ordered Pair, 46, 48, 49, 78
- Ordered Triple, 68
- Pairwise Disjoint, 58
- Paradox
  - Berry's, 9
  - Curry's, 11
  - Epimenides', 8
  - Galileo's, 15
  - Grelling-Nelson, 10
  - Kleene-Rosser, 11
  - Löb's, 11
  - Liar's, 8
  - Russell's, 12
- Partial Function, 167
- Partial Order, 82
- Partially Ordered Set, 49
- Permutation, 125, 210
- Plato, 13
- Platonic Realism, 13
- Playfair, John, 22
- Power Set, 61, 459
- Pre-Image, 76
- Predicate, 18, 27
  - Definition, 18
- Principle of Duality, 87
- Principle of Explosion, 11
- Product
  - of Sets, 78
- Proof
  - by Contradiction, 37
- Proposition, 18, 37, 45, 52
  - Definition, 18
- Ptolemy, Cladius, 21
- Pythagoras' Theorem, 462
- Python, 12
- Quantifier, 33
- Quasigroup, 175
- Ramanujan, Srinivasa, 6
- Rational Numbers, 41, 45
- Real Numbers, 46
- Reflexivity
  - of Equality, 82
  - of Inclusion, 82
- Relation
  - Antisymmetric, 82, 83
  - Reflexive, 82
  - Symmetric, 82
  - Transitive, 81
- Ring, 150
- Ring Endomorphism, 228
- Ring Homomorphism, 228
- Root Finding
  - Bisection, 4
- Russell's Paradox, 37, 45, 59, 82
- Russell, Bertrand, 11, 36, 48
- Semigroup, 167, 168
  - Commutative, 169
  - Empty Semigroup, 169
- Set, 13, 167
  - Definition, 13
  - Difference, 94
  - Disjoint Sets, 43, 55, 56
  - Element of, 14, 42
  - Empty, 38
  - Equal Sets, 40, 43, 44
  - Equality, 43
  - Finite, 36, 47, 65
  - Non-Empty, 38
  - of All Sets, 37, 45, 60
  - Power Set, 60, 61
  - Subset, 40–42
    - Proper, 44
  - Union, 51
- Set Difference, 95
- Set-Builder Notation, 45, 52
- Sierpinski Topology, 460
- Socrates, 18
  - Syllogism of, 18
- Square Root, 188
- Stolz, Otto, 15
- Stone Space, 3

- Submodule, 247  
Symmetric Difference, 197  
Symmetric Group, 216  
Symmetry  
    of Equality, 82
- Tarski's Undefinability Theorem,  
    7
- Tarski, Alfred, 7
- Theorem  
    Binomial Theorem, 62  
    Cantor's Power Set Theorem,  
        61  
    Cayley's, 182  
    Desargues's, 77  
    Diaconescu's, 79  
    Five Lemma, 127  
    Intermediate Value, 3  
    Russell's Paradox, 37  
    Tarski's Undefinability  
        Theorem, 7
- Theory of Everything, 9
- Topological Space, 457  
    Disconnected, 465  
    Stone Space, 3
- Topology, 457  
    Chaotic, 459  
    Discrete, 459  
    Indiscrete, 459
- Induced by Metric, 459, 462  
Trivial, 459
- Transitivity  
    of Inclusion, 81
- Trivial Topology, 459
- Truth, 8
- Truth Table, 19, 29
- Type Theory, 11, 48
- Union  
    of Two Sets, 52  
    over a Collection, 50, 51
- Unital Element, 136, 167  
    Left Unital, 134  
    Right Unital, 135
- Variable, 18
- Vector Space, 240
- Venn Diagram, 53, 56
- Vertical Line Test, 68, 71
- von Neumann, John, 14, 81
- Well Ordering, 9
- Whitehead, Alfred North, 11
- Wiener, Norbert, 48
- Zariski Topology, 466
- Zermelo, Ernst, 38
- Zermelo-Fraenkel Set Theory, 12,  
    14, 37, 38