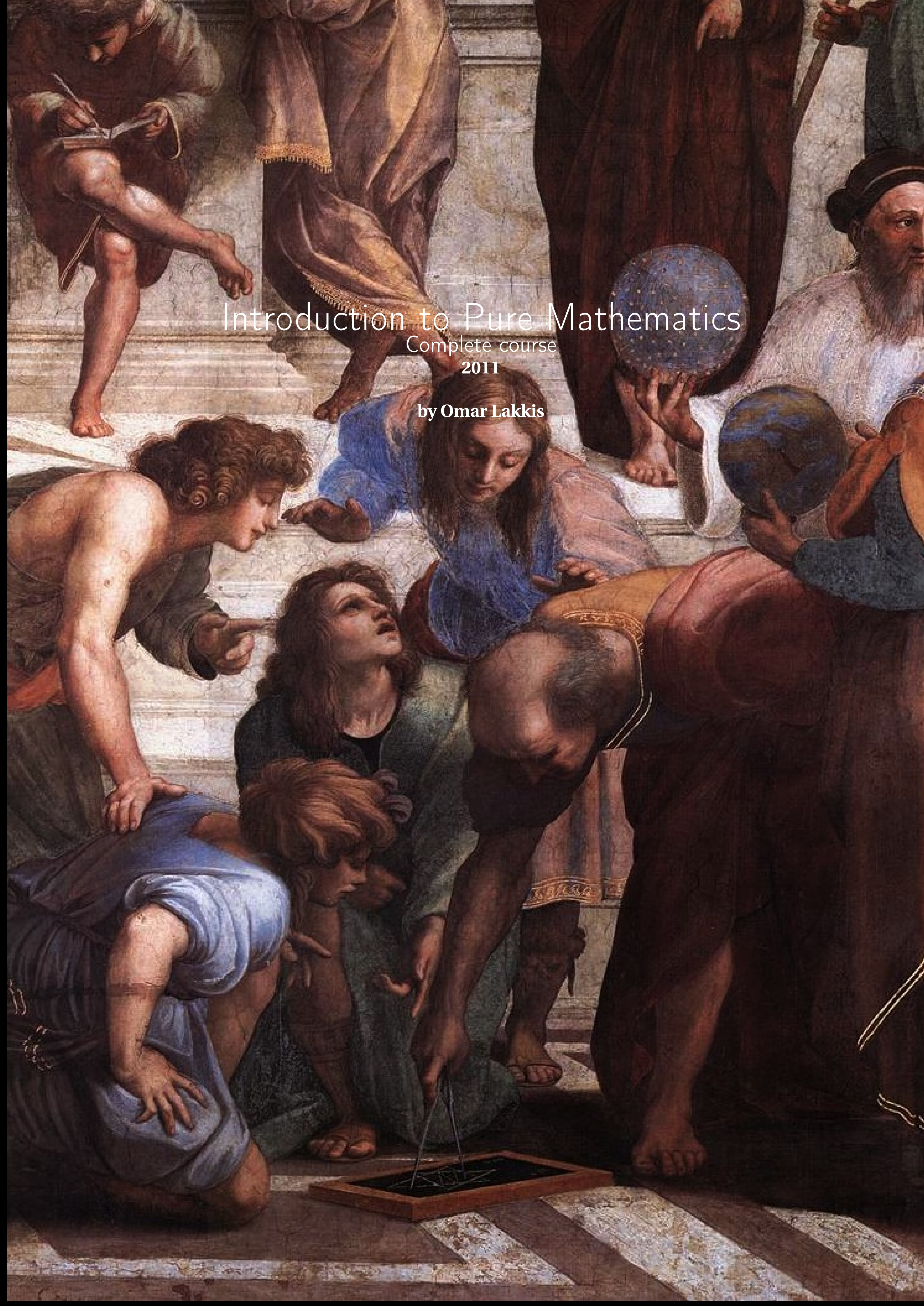
Introduction to Pure Mathematics

Complete course

2011

by Omar Lakkis



Copyright and copyleft notice

About the cover picture. "Euclid (or Archimedes) and his students", a detail of Raffaello Sanzio's fresco "Scuola di Atene/School of Athens", 1511, in the Apostolic Palace, Vatican City, Rome. Public Domain.

Contents

Copyright and copyleft notice 1

About the cover picture 1

What's this? 17

Acknowledgments 17

Course synopsys 17

Chapter 0. Preliminaries 1

0.1. About this course 1

0.2. Letters 2

Chapter 1. Numbers 5

1.1. Numbers and operations 5

1.1.1. Phony numbers vs. proper numbers 5

1.1.2. Sum of numbers as an operation 6

1.1.3. Definition of binary operation 6

1.1.4. Example (other binary operation on $\mathbb{Z}$) 7

1.1.5. Definition of unary operation/operator 7

1.1.6. Example (operators on $\mathbb{Z}$) 7

1.1.7. Example (composition of operations) 7

1.2. Algebraic properties of operations 8

1.2.1. Addition (also known as sum) on $\mathbb{Z}$ 8

1.2.2. Remark (this is not just plain silly) 8

1.2.3. Definition of group 9

1.2.4. Multiplication (also known as product) on $\mathbb{Z}$ 9

1.2.5. Exercise (groups) 10

1.2.6. Exercise (basic consequences of algebraic laws in $\mathbb{Z}$) 10

1.2.7. Exercise (multiplication of rational numbers) 10

1.2.8. Definition of ring 10

1.2.9. Example (rings) 11

1.2.10. Real numbers 12

1.2.11. Complex numbers 12

1.3. Polynomials 13

1.3.1. Definition of power 13

1.3.2. Definition of monomial 13

1.3.3. Example (monomials) 14

1.3.4. Definition of polynomial 14

1.3.5. Operations on polynomials 14

1.4. Sets and sequences 14

1.4.1. Finite sets and infinite sets 15

1.4.2. Example (finite sets) 15
1.4.3. Example (infinite sets) 15
1.4.4. Exercise (describing (infinite) sets) 15
1.4.5. Remark (terminology) 16
1.4.6. Remark (a mathie's shopping lists) 16
1.4.7. Operations on sets 16
1.5. Sequences 17
1.5.1. Definition of sequences (informal) 17
1.5.2. Example (sequence) 17
1.5.3. Example (decimal expansion) 17
1.5.4. Properties of sequences 18
1.5.5. Symbolics: brackets matter 18
1.5.6. Example (order matters for sequences) 18
1.5.7. Terminology: is a sequence finite or infinite? 19
1.5.8. Constructing infinite sequences 19
1.5.9. Example (sequences through rules) 19
1.5.10. Exercise (building sequences via rules) 20
1.5.11. Recursion 20
1.5.12. Example (Fibonacci numbers) 20
1.5.13. Exercise (Fibonacci numbers) 20
1.5.14. Example (summation as step-dependent recursion) 21
1.5.15. Definition of summation 21
1.5.16. Problem (products of n terms) 21
1.6. The Induction Principle 21
1.6.1. Theorem (Principle of Mathematical Induction—PMI) 22
1.6.2. Induction: terminology and remarks 23
1.6.3. Example (Induction Principle in action) 23
1.7. Comparison and ordering of numbers 24
1.7.1. Ordering of (real) numbers $(\mathbb{R},+,\times,\leq)$ 24
1.7.2. Definition of weak (or loose) ordering, inequalities 25
1.7.3. Example (inequalities) 25
1.7.4. Transitivity 25
1.7.5. Anti-symmetry and weak anti-symmetry 25
1.7.6. Definition of positive and negative numbers 25
1.7.7. Notation for sets of positive numbers 26
1.7.8. Basic relations between ordering and algebra 26
1.7.9. Exercise (product, ordering and simplification) 26
1.7.10. Exercise (product and ordering) 27
1.7.11. Remark 27
1.7.12. Problem 27
Exercises and problems on numbers 28

Chapter 2. Basic Number Theory 33
2.1. Preliminaries: divisibility and prime numbers 33
2.1.1. Definition of segments of integers 33
2.1.2. Definition of divisor, factor, multiple, divides 33
2.1.3. Example 33

2.1.4. Exercise 34
2.1.5. Definition of prime 34
2.1.6. Definition of composite 34
2.1.7. Remarks 34
2.1.8. Prime notation 34
2.1.9. Proposition (divisors are smaller) 34
2.1.10. Exercise 35
2.1.11. Theorem (a sufficient condition for primeness) 35
2.2. A classical proof by contradiction that $\sqrt{2} \notin \mathbb{Q}$ 36
2.2.1. Definition of square root 36
2.2.2. Exercise 36
2.2.3. Properties of the square root 36
2.2.4. Theorem ($\sqrt{2}$ is irrational) 37
2.2.5. Exercise 37
2.3. The well-ordering principle (WOP) 37
2.3.1. Definition of Empty set 38
2.3.2. Definition of least element (minimum) of a set in $\mathbb{R}$ 38
2.3.3. Remark (minimum $\neq$ lower bound) 38
2.3.4. Remark (uniqueness of minimum) 38
2.3.5. Theorem (Well Ordering Principle) 38
2.3.6. Remark (What about $\mathbb{N}_0$?) 39
2.4. Euclidean division 39
2.4.1. Theorem (Euclidean division) 40
2.4.2. Definition of quotient, remainder and modulo 41
2.5. The Euclidean algorithm 41
2.5.1. Definition of largest or greatest element, or maximum, of a set 41
2.5.2. Exercise (uniqueness of maximum) 41
2.5.3. Existence of maximum in $\mathbb{N}$, $\mathbb{N}_0$ and $\mathbb{Z}$ 41
2.5.4. Remark (maximum's uniqueness) 41
2.5.5. Definition of above-bounded set, upper bound 41
2.5.6. Remark (about the upper bound) 42
2.5.7. Proposition (an above-bounded set of integers has a maximum) 42
2.5.8. Example (a bounded set without a maximum) 42
2.5.9. Definition of highest common factor or greatest common divisor 42
2.5.10. Exercise (generalised highest common factor) 43
2.5.11. Example (how to calculate the highest common factor (without wasting too much space)) 43
2.5.12. The Euclidean Algorithm to find the HCF 43
2.5.13. Termination of the Euclidean algorithm 45
2.5.14. Remark 46
2.5.15. Remark (For hackers only) 46
2.5.16. Remark 47
2.5.17. Theorem (Euclidean Algorithm returns the hcf) 47
2.5.18. Lemma (hcf simplification via mod) 47
2.5.19. Example (Lemma 2.5.18 is crucial) 47
2.6. Bézout's identity 48
2.6.1. Theorem (Bézout's identity) 48

2.6.2. Remark (linear combinations and coefficients) 48
2.6.3. Example (working out s,t in (2.6.1)) 49
2.6.4. Problem (extended Euclidean algorithm) 49
2.6.5. Problem (divisibility and Bézout imply HCF) 50
2.7. Coprimality 50
2.7.1. Definition of coprime pair 50
2.7.2. Example (coprime pair) 50
2.7.3. Remark 50
2.7.4. Theorem (coprime factorisation) 50
2.7.5. Example 51
2.7.6. Atomic property of primes 51
2.7.7. Exercise (atomic property of primes with arbitrary factors) 51
2.7.9. Example 51
2.7.10. Exercise 51
2.8. The fundamental theorem of arithmetic 51
2.8.1. Definition of long products 52
2.8.2. Remark 52
2.8.3. Fundamental theorem of arithmetic 52
2.8.4. Remark (alternative statement of the fundamental theorem of arithmetic) 53
2.8.5. Problem (there are infinitely many primes) 53
2.8.6. Remark (YAASOFTA) 53
2.9. Extras 54
Exercises and problems on basic number theory 55

Chapter 3. Logic 59
3.1. The rise of the set 59
3.2. Russell's Paradox: an upsetting fact 59
3.3. Dirty propositions 60
3.3.1. Example (statements) 60
3.3.2. Example (propositions and non-propositions) 61
3.4. A cleaner approach with some rules 61
3.4.1. Remark 62
3.4.2. Logical operations 62
3.4.3. A warning about "⇒" and its misuses 63
3.4.4. Necessary and sufficient conditions 64
3.4.5. Quantifiers 64
3.4.6. Truth tables 64
3.4.7. Remark (redundant rules) 64
Exercises and problems on logic 65

Chapter 4. Sets 67
4.1. The joy of sets 67
4.2. Extension and equality 68
4.2.1. Definition of belong, element, member 68
4.2.2. Example (supermarket set-theory) 68
4.2.3. Example (extension by enumeration) 68

4.2.4. Example (by description) 69
4.2.5. Definition of $\in$ 69
4.2.6. Remark (a club that is fussy about membership) 69
4.2.7. Definition of inclusion, subsets 70
4.2.8. Definition of inequality, proper subsets 70
4.2.9. Exercise 70
4.2.10. How to prove equality of two sets? 70
4.2.11. Definition of formula 71
4.2.12. Example (formula) 71
4.2.13. Definition of multivariate formula 72
4.3. Specification and intersection 72
4.3.1. Remark 72
4.3.2. Example (from “Real Life”) 73
4.3.3. Definition of intersection of two sets 73
4.3.4. Exercise (some basic properties of intersection) 73
4.3.5. Intersection of collections 74
4.3.6. Exercise 74
4.3.7. Example (intersection) 74
4.3.8. Example (multiples and intersection) 74
4.3.9. Example (more multiples and intersection) 75
4.3.10. Example (intersection of nested intervals) 75
4.3.11. Example 76
4.4. Difference and complementation 77
4.4.1. Definition of difference of two sets 77
4.4.2. Remark (complementation) 77
4.4.3. Example 77
4.5. Do sets exist at all? 77
4.5.1. The empty set 78
4.5.2. Proposition 78
4.5.3. Definition of empty set 78
4.5.4. Definition of disjoint sets 78
4.5.5. Definition of vacuous truth 78
4.5.6. Exercise 78
4.6. Union 78
4.6.1. Definition of union of two sets 78
4.6.2. Example (union) 79
4.6.3. Exercise (basic properties of set theoretic operations) 79
4.6.4. Exercise (inclusion, intersection and union) 79
4.6.5. Exercise (De Morgan’s laws) 80
4.7. Sets of Sets 80
4.7.1. Pairing of sets 80
4.7.2. Ordered pairs 80
4.7.3. Definition of ordered pair 81
4.7.4. Definition of ordered tuple or finite sequence 82
4.7.5. Remark 82
4.7.6. Exercise (n-tuples equality) 82
4.7.7. Exercise 82

4.7.8. Exercise (singletons) 82
4.7.9. Exercise (finite sets of sets) 82
4.7.10. Example (v-NUMBERS) 83
4.7.11. Exercise 84
4.7.12. The power set 84
4.7.13. Example (power of ℕ) 85
4.8. Cartesian products, graphs, relations 85
4.8.1. Definition of cartesian product of two sets 85
4.8.2. Definition of cartesian square 85
4.8.3. Remark (existence of the cartesian product) 85
4.8.4. Remark (graphical representation of $A \times B$) 85
4.8.5. Definition of (directed) graph 86
4.8.6. Example (use of graphs) 86
4.8.7. Definition of long cartesian product, cartesian power 87
Exercises and problems on sets 88

Chapter 5. Counting finite sets 91
5.1. An intuitive notion of counting 91
5.1.1. Translating counting to maths: cardinality 91
5.1.2. The intuition behind counting, cardinality, finite and infinite sets 92
5.1.3. Example (cardinality of even natural numbers smaller than 101) 93
5.1.4. Exercise (cardinality of the power set) 93
5.2. Functions and maps 93
5.2.1. The intuitive idea of a function 93
5.2.2. Example (algebraic functions) 94
5.2.3. Example 95
5.2.4. Definition of identity 96
5.3. Distinguished functions 96
5.3.1. Definition of surjections/surjective functions/onto maps 96
5.3.2. Example (surjections and Venn Diagrams) 96
5.3.3. Example (surjection with infinite domain) 97
5.3.4. Example (surjection with inifinte domain and codomain) 97
5.3.5. Example (surjective and non-surjective polynomial maps on ℝ) 97
5.3.6. Definition of polynomials and polynomials maps 97
5.3.7. Example (a warning about polynomial maps on ℤ) 98
5.3.8. Definition of injections/injective functions/one-to-one maps 99
5.3.9. Remark (where is the quantifier in injectivity?) 99
5.3.10. Remark (injectivity via the contrapositive) 99
5.3.11. Example (injections and Venn diagrams) 99
5.3.12. Example (some simple injections on numbers) 99
5.3.13. Example (polynomial maps on ℝ) 100
5.3.14. Example (an injective set-to-set mapping) 101
5.3.15. Definition of bijective function/bijections/one-to-one correspondence 101
5.3.16. Example (bijections and Venn diagrams) 101
5.3.17. Example (identity) 102
5.3.18. Example (bijective and non-bijective polynomial maps) 102
5.3.19. Example (a bijective set-to-set mapping) 103

5.3.20. Theorem (finite cardinality comparison) 103
5.3.21. Remark (well-posedness of Definition 5.1.2) 103
5.3.22. Theorem (inverse function) 103
5.4. The Binomial Theorem 104
5.4.1. Definition of binomial coefficients/combinations 104
5.4.2. Proposition 104
5.4.3. Exercise (counting subsets) 105
5.4.4. Problem (Binomial Theorem) 105
5.4.5. Problem (Power set and binomial distribution) 106
5.5. The Inclusion-Exclusion Principle 106
5.5.1. Counting the union of two sets 106
5.5.2. Theorem (inclusion-exclusion (baby case)) 106
5.5.3. Example 106
5.5.4. Theorem (inclusion-exclusion for 3 sets) 107
5.5.5. Remark (why “inclusion-exclusion”) 107
5.5.6. Cardinality of a mutually disjoint union 107
5.5.7. The general case: Inclusion-Exclusion Principle 108
5.6. Permutation count 109
5.6.1. Students, chairs and music 109
5.6.2. Cherry-pick-and-exhaust-students solution 109
5.6.3. Choose the seats and permute 110
5.7. Functions, cartesian products and power 110
5.7.1. Problem (internal map counting) 110
5.7.2. Problem (general map counting) 111
5.7.3. Theorem (cartesian product cardinality and power) 111
Exercises and problems on counting 112

Chapter 6. Functions and natural numbers 117
6.1. Formal definition of a function 117
6.1.1. What does a function “do”? 117
6.1.2. A set-theoretical approach to functions 118
6.1.3. Definition of function 119
6.1.4. Remark (functions are particular graphs) 119
6.1.5. Terminology and notation 119
6.1.6. Remark (abuse of notation) 120
6.1.7. Definition of Image, range, counterimage 120
6.1.8. Remark (counterimage set of a singleton) 120
6.2. Distinguished functions revisited 120
6.2.1. Surjections 120
6.2.2. Remark (equivalence relations: a sneak preview) 121
6.2.3. Proposition (map induced equivalences) 121
6.2.4. Injections 121
6.2.5. Bijections 121
6.3. Composition and inversion 121
6.3.1. Definition of equality of functions 122
6.3.2. Definition of composition of two functions 122
6.3.3. Example (from “Real Life”) 122

6.3.4. Example (from Real Analysis) 123
6.3.5. Remark (sufficient condition for composability) 123
6.3.6. Remark (composition is not commutative!) 123
6.3.7. Theorem (composition is associative) 124
6.3.8. Theorem (invariance of distinguished functions under composition) 124
6.3.9. Definition of identity function 124
6.3.10. Proposition (identity is neutral for composition) 125
6.3.11. Definition of inverse function 125
6.3.12. Definition of partial inverse 125
6.3.13. Theorem (inverse function) 125
6.3.14. Remark (inverting injections) 126
6.3.15. Composition as an algebraic operation 126
6.3.16. Definition of composition power 126
6.4. Natural Numbers 126
6.4.1. What are natural numbers? 126
6.4.2. The origin of Peano's Axiom(s) 128
6.4.3. Definition of Peano's Axioms 128
6.4.4. Remark (successor map) 129
6.4.5. Remark (all sets are regular) 129
6.4.6. Definition of inductive set 130
6.4.7. Remark 130
6.4.8. Theorem (definition of $\mathbb{N}_0$) 130
6.4.9. Theorem 130
6.4.10. Definition of set of natural numbers, unit 130
6.4.11. Remark (successor and predecessor) 130
6.4.12. Remark (Principle of Mathematical Induction revisited) 131
6.4.13. Theorem (recursion) 131
6.5. Arithmetic of natural numbers 131
6.5.1. Definition of addition on $\mathbb{N}_0$ 132
6.5.2. Definition of ordering in $\mathbb{N}_0$ 132
6.5.3. Definition of multiplication in $\mathbb{N}$ 132
6.6. Finite sets 132
6.6.1. Definition of finite, infinite 132
6.6.2. Remark (rigorous and practical finiteness are the same) 133
6.6.3. Theorem (Pigeonhole Principle (version 1)) 133
6.6.4. Corollary (finite invertibility criterion) 133
6.6.5. Lemma (finite set injection into finite sets) 133
6.7. Cardinality of finite sets 134
6.7.1. Lemma (finite set injection into numbers) 134
6.7.2. Theorem (finite counting) 134
6.7.3. Remark (uniqueness of cardinality) 135
6.7.4. Theorem (finite set cardinality) 135
6.7.5. Definition of finite cardinality 135
6.7.6. Proposition 135
6.7.7. Natural numbers as model finite sets 135
6.7.8. Problem 135
6.7.9. Theorem (Pigeonhole Principle (version 2)) 135

6.8. Permutations 136
6.8.1. Definition of (finite) permutation 136
6.8.2. Definition of symmetric group 136
6.8.3. Notation and examples 136
6.8.4. Proposition (the symmetric group is a... group) 137
6.8.5. Proposition (symmetric group cardinality) 137
6.8.6. Definition of transposition, cycle 137
6.8.7. Example (cycles) 138
6.8.8. Lemma (cycle decomposition of permutations) 138
6.8.9. Example 138
6.8.10. Lemma (transposition decomposition of cycles) 139
6.8.11. Theorem (decomposition of permutation into transpositions) 139
Exercises and problems on functions 140

Chapter 7. Relations and algebraic extensions 143
7.1. Relations 143
7.1.1. Definition of relation 143
7.1.2. Example ("divisor of" as a relation) 143
7.1.3. Example (coprime-with and cofactored-with) 144
7.1.4. Example (geometric relations) 144
7.1.5. Example (logical relations) 144
7.1.6. Example (square root) 144
7.1.7. Example (ordering in $\mathbb{R}$) 144
7.2. Special relations 144
7.2.1. Definition of reflexive property 145
7.2.2. Example (reflexive relations) 145
7.2.3. Remark (geometric interpretation) 145
7.2.4. Definition of symmetric property 145
7.2.5. Example (symmetric relations) 145
7.2.6. Definition of transitive property 145
7.2.7. Example 145
7.2.8. Example (inclusion relation) 145
7.2.9. Example (strict inclusion relation) 146
7.2.10. Example (equipotency) 146
7.2.11. Definition of antisymmetric property 146
7.2.12. Example (antisymmetric relations) 146
7.2.13. Definition of weak antisymmetry 147
7.2.14. Remark 147
7.2.15. Exercise (antisymmetry and weak antisymmetry) 147
7.2.16. Remark 147
7.2.17. Example (weakly antisymmetric relations) 147
7.3. Ordering relations 147
7.3.1. Definition of ordering 147
7.3.2. Example 148
7.3.3. Definition of weak and strict orderings 148
7.3.4. Proposition (a nonstrict ordering is not a strict ordering) 148
7.3.5. Example 148

7.3.6. Proposition (strict orderings ⇄ partial orderings) 148
7.3.7. Definition of linear ordering 148
7.3.8. Definition of quasi-ordering 149
7.3.9. Example 149
7.3.10. Exercise (quasi-oredering quotienting) 149
7.3.11. Exercise (injection is quasi-ordering) 149
7.4. Equivalence relations 149
7.4.1. Definition of equivalence relations 149
7.4.2. Example (modular congruences in $\mathbb{Z}$) 150
7.4.3. Example (parallel lines in plane) 150
7.4.4. Example (equipollent segments in plane) 150
7.4.5. Definition of equivalence class, residue class, congruence class 151
7.4.6. Example (congruence modulo 5) 151
7.4.7. Proposition (properties of equivalence classes) 151
7.4.8. Definition of partition of a set 152
7.4.9. Example (partitions) 152
7.4.10. Remark (partition language) 152
7.4.11. Example (citizenship as an equivalence) 152
7.4.12. Example (congruence modulo $p \in \mathbb{N}$ in $\mathbb{Z}$) 153
7.4.13. Definition of representative, transversal, quotient set 154
7.4.14. Example (representative, transversal, quotient set in politics) 154
7.4.15. Theorem (natural map and transversal-quotient isomorphism) 154
7.4.16. Example (quotient of $\equiv_n$ in $\mathbb{Z}$) 155
7.4.17. Proposition (partition-induced equivalences) 155
7.4.18. Example (sign equality is an equivalence on $\mathbb{R}$) 155
7.4.19. Example (same argument as an equivalence on $\mathbb{C}$) 155
7.4.20. Definition of compatible operations and relations 156
7.4.21. Example (algebraic operations and congruence) 156
7.4.22. Example (a multiplicative but not additive equivalence on $\mathbb{R}$) 157
7.4.23. Proposition (quotient (also known as induced) operation) 157
7.4.24. Example (plus and times tables in modular arithmetic) 157
7.4.25. Exercise (plus and times tables in modular arithmetic) 157
7.4.26. Proposition (quotient relation) 158
7.4.27. Example (sign equality as equivalence) 158
7.4.28. Exercise (complex multiplication and geometry) 158
Exercises and problems on relations 159

Chapter 8. Real Numbers 163
8.1. Why real numbers? 163
8.2. Ordered fields 164
8.2.1. Definition of (algebraic) field 164
8.2.2. Definition of ordered field 164
8.2.3. Proposition (ordering's compatibility with product rules) 165
8.2.4. Proposition (product and ordering properties) 165
8.2.5. Definition of nonstrictly ordered field 165
8.2.6. Definition of positive and negative 166
8.2.7. Proposition (sign invariance under reciprocation) 166

8.2.8. Definition of absolute value, modulus 166
8.2.9. Proposition (properties of the absolute value) 166
8.2.10. Definition of interval 167
8.2.11. Proposition (an ordered field is infinite) 167
8.2.12. Characteristic of a field 168
8.3. Bounded sets, extremums and best bounds 168
8.3.1. Definition of upper and lower bound in an ordered field 168
8.3.2. Example (upper bound) 169
8.3.3. Exercise (upper bound's transitivity) 169
8.3.4. Definition of extremum 169
8.3.5. Exercise (maximum is an upper bound) 169
8.3.6. Definition of least upper bound and greatest lower bound 169
8.3.7. Exercise (basic facts about supremum and infimum) 169
8.3.8. Proposition (uniqueness of the least upper bound) 170
8.3.9. Definition of archimedean field 170
8.3.10. Lemma (vanishing constant) 170
8.4. Rational numbers 170
8.4.1. Definition of $\mathbb{Q}$, the set of rational numbers 170
8.4.2. Corollary (rationals are embedded in all ordered fields) 171
8.4.3. Proposition (rationals are archimedean) 171
8.4.4. Example (extremums and best bounds in $\mathbb{Q}$) 171
8.4.5. Least upper bound (supermum) vs. greatest element (maximum) 171
8.4.6. Exercise 172
8.5. Sequences in archimedean fields 172
8.5.1. Definition of $\mathbb{F}$-valued sequence and sequence in $\mathbb{F}$ 172
8.5.2. Definition of convergence to 0 and vanishing sequences 172
8.5.3. Definition of convergence, convergent sequences and limit 172
8.5.4. Proposition (uniqueness of limit (in archimedean fields)) 173
8.5.5. Definition of bounded above, bounded below and bounded sequence 173
8.5.6. Proposition (bounded means absolutely bounded) 173
8.5.7. Definition of divergence to infinity and infinite limits 173
8.5.8. Proposition (comparison of divergent sequences) 174
8.5.9. Lemma (Cauchy's necessary criterion) 174
8.5.10. Definition of Cauchy sequence 174
8.5.11. Theorem (convergent sequences are bounded) 174
8.5.12. Theorem (algebra of limits) 175
8.5.13. Lemma (squeezing) 175
8.5.14. Remark 175
8.6. Real numbers 176
8.6.1. Incompleteness of $\mathbb{Q}$ 176
8.6.2. Axiomatic definition of real numbers 176
Completeness Axiom 176
8.6.3. Rational numbers are real 176
8.6.4. Some reals are irrational 177
8.7. Power and root 177
8.7.1. Definition of (arithmetic) power function 177
8.7.2. Remark 177

8.7.3. Proposition (square root of positive reals) 177
8.7.4. Problem (square root of positive reals) 178
8.7.5. Proposition (root of an integer power) 178
8.7.6. Definition of n-th root function 178
8.7.7. Integer power and root laws 178
8.7.8. Definition 178
8.7.9. Proposition (power laws) 179
8.8. Montone sequences 179
8.8.1. Definition of monontone, increasing, and decreasing sequences 179
8.8.2. Definition of best bounds of a sequence 179
8.8.3. Theorem (monotone sequence) 179
8.8.4. Remark (completeness is essential for Monotone Sequence Theorem) 180
8.8.5. Example (Napier–Bernoulli–Euler constant) 180
8.9. Rational and constructive approaches to reals 181
8.9.1. Dedekind cuts 181
8.9.2. Theorem (Dedekind cuts) 181
8.9.3. Cauchy sequences of rational numbers 182
Exercises and problems on real numbers 183

Chapter 9. Infinity 187
9.1. Infinite sets 187
9.1.1. Lemma (characterisation of infinity via $\mathbb{N}_0$) 187
9.1.2. Remark 188
9.1.3. Proposition ($\mathbb{N}_0$ and $\mathbb{N}$ are infinite) 188
9.2. Countable and uncountable sets 188
9.2.1. Definition of countable 188
9.2.2. Remark (finite sets are countable) 188
9.2.3. Proposition 189
9.2.7. Theorem (Cantor's diagonal) 189
9.3. The Axiom Schema of Replacement 189
9.4. The Axiom of Choice 190
Exercises and problems on infinity 191

Chapter 10. Groups, rings and fields 193
10.1. Groups 193
10.1.1. Definition 193
10.2. Rings 193
10.3. Polynomials 193

Chapter 11. Analysis of sequences 261
11.1. Metric spaces 261
11.1.1. Definition of metric space 261
11.1.2. Example (fields are metric spaces) 261
11.1.3. Example 261
11.1.4. Definition of metric space balls 262
11.1.5. Definition of bounded set 262
11.2. Sequences in a metric space 262
11.2.1. Definition of *convergence* in a metric space 262

11.2.2. Remark 262
11.2.3. Definition of limit of a sequence 262
11.2.4. Proposition (uniqueness of limits) 262
11.2.5. Definition of Cauchy sequence 262
11.2.6. Definition of bounded sequence 263
11.3. Cauchy sequences 263
11.3.1. Lemma (boundedness of Cauchy sequences) 263
11.3.3. Theorem (completeness of $\mathbb{R}$) 263
11.4. Contractive sequences 263
11.4.1. Definition of contractive sequence 263
11.4.3. Theorem (contraction principle) 263
11.5. Real-valued sequences 263
11.5.1. Definition of lower limit (also known as liminf) and upper limit (also known as limsup) 264
11.5.2. Exercise 264
11.5.3. Lemma (inf and sup squeezing) 264
11.5.4. Exercise (squeezing) 264
11.5.5. Theorem (Stolz–Cesàro) 264
Exercises and problems on sequences 266

Chapter 12. Sums and series 271
12.1. From finite to infinite sums 271
12.1.1. Infinite sums as limits of finite sums 271

Chapter 13. Limits of functions and continuity 273
13.1. The limit of a function 273
13.1.1. Definition of sequential limit of a function 273
13.1.2. Definition of topological limit of a function 273
13.1.3. Lemma (sequential and topological limit are the equivalent) 273
13.1.4. Lemma (separation from zero) 273
13.1.6. Theorem (intermediate value) 274
Exercises and problems on limits of functions and continuity 275

Chapter 14. Complex numbers 277
14.1. Algebraic approach 277
14.1.1. Axiomatic definition of complex numbers 277
14.2. Geometric approach 277
14.2.1. Cartesian plane and planar vectors 277
14.2.2. Linear maps 278
14.2.3. Definition 279
14.2.4. Proposition 279
14.2.5. Definition of canonical basis 279
14.2.6. Definition of 2×2 matrix 279
14.2.7. Proposition 280
14.2.8. Example (rotations) 280
14.2.9. Definition of planar rotation matrices 281
14.2.10. Proposition 281
14.2.11. Example (rotodilations) 281

14.2.12. Invertibility of linear maps 281
14.2.13. Conformal and anticonformal maps 281
14.2.14. Proposition (linear conformal maps are rotodilations) 282
14.2.15. Proposition (field property) 282

Chapter 15. Polynomials and power series 283
15.1. The fundamental theorem of algebra 283
15.1.1. Definition of complex polynomials 283

Bibliography 285

Index 287

What's this?

I mainly compiled this booklet to help the students taking the course "Introduction to Pure Mathematics" at the University of Sussex in 2011. It contains all the material covered in class as well as the exercises and their solutions, and a little bit more than that.

Acknowledgments

Parts of this document were initially typeset by Raquel Barreira and Christos Papaioannou, who were postgraduates assisting me in the course and to whom I am quite grateful. Special thanks go also to Vanessa Styles, James Hirschfeld, Charlie Elliott and Tom Armour as well for their contributions. Of course, the sole responsible for mistakes, typos and omissions remains the author, who will be glad to receive any kind of comments on this text.

Course synopsys

This course (or booklet, if you're not a student at Sussex) covers the very basic elements of university level so-called "pure" Mathematics. Although dividing Mathematics into pure and applied is a futile activity, my aim here is to introduce mostly British-educated freshers straight from A-levels (where most maths taught nowadays is, unfortunately, so-called "applied") to concepts such as proofs, sets, formal logic and basic combinatorics. This is done, by following the modern British approach whereby number theory and elementary algebra are the driving motivations to develop the examples, without which proofs, sets, logic and combinatorics would become tedious abstract tasks.

Other recommended texts that are useful are given in the references at the end of the notes.

Johnson, 1998: A nicely read book, full of nice exercises and interesting problems and upon which I based my original course, from where these lecture notes were born.

Smith, 1998: This book covers similar material to the one covered in class. It gives a somewhat different perspective and has interesting exercises.

Halmos, 1974: This is a more "advanced" text, which is an excellent *second reading.* It is slim (though concentrated) and has all what a working mathematician really needs from set theory, which is covered in this course.

Graham, Knuth and Patashnik, 1994: An excellent book, way too big for our course, but full of resources for this course, as well for many other courses. Though intended primarily for informatics students, this book covers a large part of combinatorics, basic probabilty and other topics that are very useful to mathematicians. As the title hints, the approach is hands-on, whithout too much mathematical “sophistication”. This a good complement to Halmos, 1974, an excellent but somewhat “dry” classic.

Although the mathematical contents of these notes reflect an agreed curriculum by the Department of Mathematics at the University of Sussex, the personal opinions that appear in the text are to be attributed solely to me. These are not required for the final exam.

December 2011, Omar Lakkis.

CHAPTER 1

Numbers

> God made the integers; all else is the work of man.
> — Leopold Kronecker

It should not come as a surprise, if a first course in pure mathematics starts by discussing numbers. Along with geometric objects (lines, polygons, circles), numbers constitute the fundamental objects of mathematics. Numbers are used in all aspects of life, but their usage is not our (primary) source of interest in them. As mathematicians, we are interested in understanding how numbers "work". One of the main objectives of this course is to ultimately understand what numbers really are, what kind of relations among them and how they function.

In this chapter we review some basic facts about numbers, some of which may be known to you and some not.

1.1. Numbers and operations

1.1.1. Phony numbers vs. proper numbers. The word "number" has a very loose meaning in practise and it is usually used in two very different contexts. We have to make sure we understand what we mean intuitively by number.

We often use the word "number" in expressions like *phone numbers*, or *lottery ticket number*, or even *bus number*. But we also use it in some other contexts, like *the number of people in this room*, or *the number of molecules in a container*, or *the number of days until the end of term*, the *arrival number (position) of an athlete at a race*, etc. At first sight there seems to be no difference between the use of number in the first case and in the second case, but it is worth having a closer inspection. In the second case the concept of number is richer than in the first one. Indeed, in the second case, the number of people in a room can be manipulated whereas in the first case manipulation is not very useful. For example, we can *add* some to the number of people in this room when the late ones sneak in after the lecture has started, or we can *subtract* from it (if someone receives an unignorable mobile phone call and needs to jump out of the room to talk freely), or we can *multiply* it by 2 if we wanted to know how many eyes are staring at the lecturer, or divide it by 3 if we wanted say to form groups of three people for a computing project, etc. The same for the number of molecules, it makes physical sense for it to be added to, subtracted from, multiplied by or divided by another number of molecules. The position of the athlete can be also used in many ways, for example, one can calculate an athlete's performance over a period of time by averaging her position in all the races she has participated in, and two different athletes can be *ordered* judging on their respective position numbers.

Now try to think (or talk to someone) about ordering people by their telephone numbers, or adding lottery ticket numbers, or multiplying house or bus numbers together.

In the best scenario, you can be sure the person you talk to will be switching pavements next time they cross you on the street.
The moral is that the numbers that interest mathematicians are those of the second kind. That is numbers which can be used to operate with and to compare. This is the only kind of numbers we shall be dealing with.

1.1.2. Sum of numbers as an operation. The *sum* (also known as addition), usually denoted by "+", is a *binary operation* (binary because it requires 2 arguments or operands to be fed in). The sum of two numbers, say a and b, denoted by $a+b$ and has the following properties: it is *commutative*, in that we get the same result if we do $b+a$, and *associative*, that is,

$$a+(b+c)=(a+b)+c \tag{1.1.1}$$

for any three numbers a, b and c.
The sum operation is defined for any pair of numbers and it can be thought as some kind of machine, let's call it sum, that takes any 2 numbers as input and returns 1 number as output. We could be writing the result of summing x and y as $\operatorname{sum}(x,y)$, but this is not very customary, so we will do it only in this section to play a bit around. A convenient and meaningful way of writing this is

$$\begin{array}{rccc}\operatorname{sum}: & \mathbb{Z}\times\mathbb{Z} & \to & \mathbb{Z} \\ & (x,y) & \mapsto & x+y\end{array} . \tag{1.1.2}$$

Let us stop and deconstruct this notation for once:

- ⋆ the symbol "sum" here stands for the *operation* that we are describing,
- ⋆ $\mathbb{Z}$ denotes the *set of all integers* (this is a standard symbol used by almost all mathematicians), in symbols we may write

$$\mathbb{Z} := \{0, 1, -1, 2, -2, 3, -3, \ldots\}, \tag{1.1.3}$$

- ⋆ the *input set* of the operation sum is $\mathbb{Z}\times\mathbb{Z}$, i.e., two copies of $\mathbb{Z}$, one for the left (or first) argument and one for the right (or second) argument, (note that $\mathbb{Z}$ is different than $\mathbb{Z}\times\mathbb{Z}$ and that using second one is needed for the operation is binary)
- ⋆ the *output set* of the operation sum is $\mathbb{Z}$
- ⋆ the *effect* of the operation on a generic *pair* of numbers (x,y) is defined as being $x+y$, this could have been written as $\operatorname{sum}(x,y) := x+y$.

Addition (or sum) is an example of the following definition.

1.1.3. Definition of binary operation. A *binary operation B* on a set S is a operation that has a *set of inputs* (also known as *arguments/(independent) variables*) consisting of two copies of S, a *set of outputs* (also known as *images/values/dependent variable*) consisting of (a single copy of) S, and a given *rule* or *law*. This definition is a bit loose, because the terms "law", "rule", "set" were not introduced. In this chapter, and the next, we will be satisfied with this level of rigour, reserving the right to come back and make all this more rigorous later.

1.1.4. Example (other binary operation on $\mathbb{Z}$).

(a) The product (also known as multiplication) of two numbers is a binary operation on $\mathbb{Z}$.

(b) The difference (also known as subtraction) of two numbers is a binary operation on $\mathbb{Z}$.

(c) The division of two numbers is *not* a binary operation on $\mathbb{Z}$, because not every pair of inputs provides an output, e.g., $(3,4)$ or $(9,0)$ do not have an image because $3/4$ is not an integer whereas $9/0$ is not defined at all.

1.1.5. Definition of unary operation/operator. A simpler concept than binary operation is that of unary operation which take a single argument rather than two of them.

A *unary operation*, or *operator*, on a set S is a operation that takes inputs in the set S and returns outputs in S. Unary operations are also referred to as *operators* on S. An operator, say A, on S is usually denoted as

$$\begin{array}{rccc} A: & S & \rightarrow & S \\ & x & \mapsto & A(x) \end{array}, \tag{1.1.4}$$

and many times, when no confusion might occur, we drop the brackets by writing Ax for $A(x)$.

1.1.6. Example (operators on $\mathbb{Z}$).

- ⋆ An operator (one argument) is the negation, usually denoted by "$-$", and here denoted also by $\operatorname{neg}(\cdot)$. Thus $\operatorname{neg}(2) = -2$ or $\operatorname{neg}(-12) = 12$.
- ⋆ Another example of operator on $\mathbb{Z}$ is given by the law

$$\begin{array}{rccc} \tau: & \mathbb{Z} & \rightarrow & \mathbb{Z} \\ & n & \mapsto & 3n \end{array}, \tag{1.1.5}$$

  also written as $\tau n := 3n$, $n \in \mathbb{Z}$.
- ⋆ A further example of operator on $\mathbb{Z}$ is

$$\begin{array}{rccc} \sigma: & \mathbb{Z} & \rightarrow & \mathbb{Z} \\ & n & \mapsto & n^2 \end{array}. \tag{1.1.6}$$

- ⋆ A law that is not an operator on $\mathbb{Z}$ is the square root. Indeed, the square root of -4 is not a number in $\mathbb{Z}$, because (as we shall show below) all squares of integers must be non-negative. Since an operator must provide an output *for all inputs* square root cannot be an operator on $\mathbb{Z}$.
- ⋆ Similarly γ defined by $\gamma n := n/2$, $n \in \mathbb{Z}$, is not an operator on $\mathbb{Z}$, because $1/2$ is not in $\mathbb{Z}$.

  Note that γ becomes an operator if the set $\mathbb{Z}$ is substituted by $\mathbb{Q}$, where $\mathbb{Q}$ is the set of all numbers that can be written as fractions of integers, also called *rational numbers.*

1.1.7. Example (composition of operations). Operations (binary and unary) can be combined or *composed* to give rise to new operations. For example, subtraction is a composition of the sum and the negation as follows

$$a - b := \operatorname{sum}(a, \operatorname{neg} b). \tag{1.1.7}$$

1.2. Algebraic properties of operations

Algebra is the science that studies, among other things, the manipulation of numbers and their properties. Algebraists look at a the set $\mathbb{Z}$ not merely as a bucket with a whole lot of numbers in it, but as a *system* in which they include the operations $+$ (addition) and $\times$ (multiplication). This is conveniently denoted by $(\mathbb{Z}, +, \times)$. We list now the most basic properties of the operations in $(\mathbb{Z}, +, \times)$.

1.2.1. Addition (also known as sum) on $\mathbb{Z}$.

(i) Addition is *associative* in $\mathbb{Z}$, i.e.,

$$\forall\, n, m, p \in \mathbb{Z} : n + (m + p) = (n + m) + p. \tag{1.2.1}$$

(ii) A *neutral element e* in $\mathbb{Z}$, is defined, as satisfying $n + e = e + n = n$ for all $n \in \mathbb{Z}$. You probably know that 0 satisfies such a property. Thus, we say that addition has a neutral element in $\mathbb{Z}$, usually called zero and denoted 0. In symbols, this sentence can be written as

$$\underbrace{\exists\, 0 \in \mathbb{Z} :}_{\text{there exists an element 0 in } \mathbb{Z} \text{ such that}} \qquad \forall\, n \in \mathbb{Z} : 0 + n = n + 0 = n. \tag{1.2.2}$$

(iii) Each number a in $\mathbb{Z}$ has an *opposite* (also known as *additive inverse*) number, denoted by $(-a)$, in $\mathbb{Z}$, which satisfies $a + (-a) = (-a) + a = 0$. In symbols, we write this as

$$\underbrace{\forall\, n \in \mathbb{Z} :}_{\text{for all } n \text{ in } \mathbb{Z}} \qquad \underbrace{\exists\, (-n) \in \mathbb{Z} :}_{\text{there exists an element } (-n) \text{ in } \mathbb{Z} \text{ such that}} \qquad n + (-n) = (-n) + n = 0. \tag{1.2.3}$$

(It is customary to denote "$(-a)$" simply as "$-a$" and to write "$n - m$" instead of the more convoluted "$n + (-m)$".)

(iv) Addition is *commutative* in $\mathbb{Z}$, i.e., for any n, m numbers in $\mathbb{Z}$, we have $n + m = m + n$. This sentence can be written in symbols as

$$\underbrace{\forall\, n, m \in \mathbb{Z} :}_{\text{for all } n, m \text{ elements of set } \mathbb{Z} \text{ it holds}} \qquad n + m = m + n. \tag{1.2.4}$$

1.2.2. Remark (this is not just plain silly).

1. The properties of the addition listed above are "obvious" to us, because we use the sum everyday and we take them for granted.[1] However, many operations do not satisfy these properties. For example, subtraction, which is a binary operation on $\mathbb{Z}$, is neither commutative nor associative therein, as it is easily seen by producing some example.[∗]

[∗]: Check!

2. Note that is it is important to state the set on which an operation satisfies a certain property. For example, if we denote by $\mathbb{N}$ the subset of $\mathbb{Z}$ that consists of all numbers (and only those) that are strictly positive, then $+$ does not satisfy the neutral element property (ii) nor the opposite element property (iii).

3. Property (iii), implicitly defines a unary operator *opposition* (also known as additive inversion, or algebraic negation) on $\mathbb{Z}$.

[1] The word "obvious" should be banned from mathematical texts. If some fact is obvious why write it? The urge to say something makes it not obvious. We shall refrain from using the word "obvious".

4. Note how mathematicians try to confuse each other (and sometime themselves) by denoting two objects with the same symbol. For example "−" is used as a sign to denote the negation (unary) operation defined in (iii) as well as the subtraction (binary) operation. Logically these are two different operations. This economy in recycling the same symbol for two different objects is called "overloading" in informatics. It is quite useful, but it can be dangerous, so it must be handled with care and you should always make sure you understand the meaning of a symbol when you use it, just as you do with words.[2]

5. As if overloading was not enough, mathematicians like to call the same object with different names. For example, $-a$ is called by some *opposite* of a, by others *negative* a, by others yet *negated* a or *additive inverse* of a. The reason for such a proliferation of terminology is that different people in different disciplines discover the same object. While the object remains the same across disciplines, the terminology changes. We have to live with this sorry state of affairs.

1.2.3. Definition of group. Suppose S is a set (it could be $\mathbb{Z}$ but not necessarily so) and $*$ some binary operation on S (it could be $+$ but not necessarily so), we say that the algebraic system $(S,*)$ forms a *group* if and only if all the following hold true:

(i) $*$ is *associative* on S,

(ii) $*$ has a *neutral element* in S,

(iii) each $x \in S$, has an $*$-*inverse* in S.[3]

If $*$ happens to be also *commutative* on S, then we say that $(S,*)$ is an *Abelian group*. Groups are important structures that are found in most branches of Mathematics (and Physics). Group Theory constitutes one of the main areas of Modern Algebra. An example of (Abelian) group is $(\mathbb{Z},+)$. Note that $(\mathbb{N},+)$ is *not a group*.[∗]

[∗]: Check!

1.2.4. Multiplication (also known as product) on $\mathbb{Z}$. A second fundamental operation on $\mathbb{Z}$ is the *product* (also known as *multiplication*) ×. Note that although the operation is referred to using the symbol ×, we usually write "ab" instead of "$a \times b$". The arguments of × (i.e., a and b in the previous sentence) are called *factors*. When both factors are written in digits, we do use the symbol × (many in Britain use the alternative symbol ·). Example: to write "twelve times thirty four" then "12×34" or "$12 \cdot 34$", or "(12)(34)", are all OK, but "12 34" or "12.34" would be too confusing.

The product of numbers × satisfies the following properties.

(i) × is associative on $\mathbb{Z}$, in symbols

$$\forall\, a,b,c \in \mathbb{Z} : a(bc) = (ab)c. \tag{1.2.5}$$

(ii) × has a neutral element in $\mathbb{Z}$, called *one* (also known as *unit* or *identity*), and written 1. In symbols

$$\exists\, 1 \in \mathbb{Z} : a1 = 1a = a. \tag{1.2.6}$$

(iii) × is commutative on $\mathbb{Z}$, in symbols

$$\forall\, a,b \in \mathbb{Z} : ab = ba. \tag{1.2.7}$$

[2] Fortunately, most mathematicians seem to tackle these ambiguities successfully.

[3] The symbol $*$ is usually called asterisk.

A fundamental property that relates × to + in $\mathbb{Z}$ is the *distributive law*

$$\forall\, a,b,c \in \mathbb{Z} : a(b+c) = ab+ac. \tag{1.2.8}$$

1.2.5. Exercise (groups).
(a) Is $(\mathbb{Z},\times)$ a group? Explain your answer.

(b) Why is $(\mathbb{N},+)$ not a group?

(c) Let $\mathbb{N}_0 := \{0,1,2,3,\dots\}$ (i.e., the set of all non-negative integers). Is $(\mathbb{N}_0,+)$ a group?

(d) Which of the group properties of $(\mathbb{Z},+)$ remain valid for $(\mathbb{N},+)$ and which do not? Same question for $(\mathbb{N}_0,+)$.

1.2.6. Exercise (basic consequences of algebraic laws in $\mathbb{Z}$). *Each result in this exercise depends on the preceeding ones, but they can be solved indipendently.*
(a) Suppose $e,a \in \mathbb{Z}$ and $e+a=a$. Show, based on first principles, that $e=0$. Justify each single step with basic algebraic properties of addition.

(b) Show, based on first principles (i.e., the basic properties of the sum) that if $a+a=a$, $a\in\mathbb{Z}$, then it must be $a=0$.

(c) Show that 0 is an absorbing element for $\times$, i.e., that

$$\forall\, a \in \mathbb{Z} : 0a = 0, \tag{1.2.9}$$

using the basic properties of + and × in $\mathbb{Z}$ and the previous results.

Hint. Start by summing $0a$ to itself and work that out.

(d) Denote by -1 the opposite of 1. Drawing from first principles and previously shown results, prove that $(-1)a = -a$ for any $a \in \mathbb{Z}$.

Hint. Prove that $a+(-1)a=0$.

1.2.7. Exercise (multiplication of rational numbers). *Let $\mathbb{Q}$ be the set of all possible fractions of integers, that is all numbers of the form m/n with $m,n\in\mathbb{Z}$ and $n\neq 0$, with the convention that two fractions m/n and m'/n' represent the same number, i.e., $m/n = m'/n'$, if and only if $mn' = m'n$.*
(a) Recalling that multiplication of fractions is defined by

$$\frac{m}{n}\times\frac{k}{l} := \frac{mk}{nl}, \tag{1.2.10}$$

and using the properties of × in $\mathbb{Z}$ show that × is associative, commutative and that $1/1$ is its neutral element in $\mathbb{Q}$.

(b) Show that $m/m = 1/1$ in $\mathbb{Q}$, for any $m\in\mathbb{Z}$ and $m\neq 0$.

(c) Let $\mathbb{Q}^*$ be the set $\mathbb{Q}$ with the fraction $0/m$ taken away (note that $0/n = 0/m$ for any $n,m\in\mathbb{N}$). Explain why $(\mathbb{Q}^*,\times)$ is a group.

Hint. Show that for each number $q = m/n \in \mathbb{Q}^*$ it is possible to find a number $q' \in \mathbb{Q}^*$ such that $qq' = 1/1$ (i.e., each q has a multiplicative inverse q').

(d) Is $(\mathbb{Q},\times)$ a group?

1.2.8. Definition of ring. A ring is an algebraic structure on a set A with two binary operations, denoted + and $\star$ such that $(A,+)$ is an Abelian group and $\star$

(i) is associative,

(ii) it has a neutral element, called A's unit and denoted by 1_A, $1_\star$ or simply 1, and

(iii) it is distributive with respect to $+$, i.e.,

$$a \star (b + c) = a \star b + a \star c \text{ and } (a + b) \star c = a \star b + a \star c. \tag{1.2.11}$$

Note that in general we do not require $\star$ to be commutative (which is why we need two statements for the left and right distributivity in (1.2.11)) and many useful rings are not commutative. When $\star$ is commutative we say that $(A, +, \star)$ is a *commutative ring*. An element $a \in A$ has a $\star$-inverse if there exits $a' \in A$ such that

$$a' \star a = a \star a' = 1_\star. \tag{1.2.12}$$

The $\star$-inverse is denoted $a\star^{-1}$. When $\star$ is the a multiplication, or understood from context, it is omitted.

1.2.9. Example (rings). The following are rings:

$$(\mathbb{Z}, +, \times), (, +, \times), (\mathbb{R}, +, \times), (\mathbb{C}, +, \times). \tag{1.2.13}$$

These are not

$$(\mathbb{N}, +, \times), (\mathbb{N}_0, +, \times). \tag{1.2.14}$$

The set of 2×2 matrices with (say) integer, rational, real or even complex coefficients is a ring. Should you not know what a matrix is all you need to know is that these are tables (or arrays) of the form

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix}, \tag{1.2.15}$$

which we *add* as follows

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} + \begin{bmatrix} s & u \\ t & v \end{bmatrix} = \begin{bmatrix} a + s & c + u \\ b + t & d + v \end{bmatrix} \tag{1.2.16}$$

and *multiply* as follows

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} s & u \\ t & v \end{bmatrix} = \begin{bmatrix} as + ct & au + cv \\ bs + dt & bt + dv \end{bmatrix}. \tag{1.2.17}$$

Matrices are very useful devices in Geometry and Algebra; this is just a warm-up example. Their set is indicated by $X^{2\times2}$ where X can be any one of $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$ or any structure that has a sum and a product (both needed to define the matrix sum and product). A useful exercise is to convince yourself that if X is a ring then $X^{2\times2}$ is also a ring; and, very importantly, that $X^{2\times2}$ may not commutative even in cases where X is. For example, posing

$$\boldsymbol{A} := \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}, \boldsymbol{B} := \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \text{ we get } \boldsymbol{AB} = \begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix} \neq \begin{bmatrix} -2 & 1 \\ 3 & 1 \end{bmatrix} = \boldsymbol{BA}. \tag{1.2.18}$$

1.2.10. Real numbers. So far we have seen the sets of numbers $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$. Each of theses sets is a proper subset of the next; in symbols we write this fact as

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q}. \tag{1.2.19}$$

By *proper subset* we mean subset and different, which explains the use of " $\subsetneq$ " to emphasise the difference. There is one more set of numbers, which contains all these. From what we have learned so far we see that $+$ and $\times$ are associative, commutative and they satisfy the distributive law in $\mathbb{Q}$ (and thus in $\mathbb{Z}$ and $\mathbb{N}$). Furthermore, $(\mathbb{Z},+)$ and $(\mathbb{Q},+)$ are Abelian groups, and so is $(\mathbb{Q}^*,\times)$.

So $\mathbb{Q}$ seems like a great set, where all the numbers which we could think of are to be found. But this is not true. In fact, it can be shown (we shall do it later in this chapter, see also Johnson Theorem 1.2) that there are no number q in $\mathbb{Q}$ such that $q^2 = 2$. On the other hand, by Pythagoras Theorem, we know that such a number s should exist, in order to measure the hypotenuse of a square triangle with side measuring 1 (because $1^2 + 1^2 = 2 = q^2$. In other words $\sqrt{2}$ is not a number in $\mathbb{Q}$. This looks like a grim prospect, except it is possible to construct a set of number $\mathbb{R}$, of which $\mathbb{Q}$ is a proper subset, and which contains $\sqrt{2}$. The set $\mathbb{R}$ is called the set of real numbers and $+$ and $\times$ are defined therein. Furthermore $(\mathbb{R},+)$ is also a group and $(\mathbb{R}^*,\times)$ too (here $\mathbb{R}^*$ is the set of all reals numbers except 0) and $\times$ is distributive with respect to $+$ in $\mathbb{R}$. The set of real numbers is very important in *Analysis*, the branch of mathematics which studies quantities that become very, very, very small, in fact infinitely small.

1.2.11. Complex numbers. With $\mathbb{R}$ we own all the numbers that are needed to measure length of lines, area of surfaces, volume of solids, mass, energy, money in the bank, geometric ratios, etc. However, there is more to mathematics than "measuring" and "quantifying". For example, one may wish to solve equations of the type $x^2 = r$ for x where $r \in \mathbb{R}$. It turns out that there are solutions $x \in \mathbb{R}$ if $r \geq 0$, but no solution $x \in \mathbb{R}$ for $r < 0$. It is however, possible to build a proper superset of $\mathbb{R}$, denoted $\mathbb{C}$, and called the *set of complex numbers*, where we can find solutions x to the equation $x^2 = r$, for all $r \in \mathbb{R}$. In fact, it turns out to be possible to solve $x^2 = r$ in $\mathbb{C}$ for all $r \in \mathbb{C}$ and more generally, the celebrated Fundamental Theorem of Algebra, proves that any polynomial with coefficients in $\mathbb{C}$ has at least one root in $\mathbb{C}$. This fact to say that $\mathbb{C}$ is *algebraically closed.*

A quick and dirty way of defining complex numbers is to introduce an imaginary unit i such that $\mathrm{i}^2 = -1$. Since for any real number x we have $x^2 \geq 0$, this means that i is *not a real number.* We then consider all expressions of the type $a + b\,\mathrm{i} = a + \mathrm{i}\,b$ where $a, b \in \mathbb{R}$ with the usual algebraic rules (associativity, commutativity and distributivity) for sum and product. For example,

$$(a + b\,\mathrm{i})(c + d\,\mathrm{i}) = ac + ad\,\mathrm{i} + bc\,\mathrm{i} + bd\,\mathrm{i}^2 = (ac - bd) + (ad + bc)\mathrm{i}. \tag{1.2.20}$$

It follows that $\mathbb{C}$ is a field, and thanks to $\mathrm{i}^2 = -1$ we are able to solve any quadratic equation $x^2 = y$ with negative y. It turns out in fact that this is solvable for *any* $y \in \mathbb{C}$, say $y = a + b\,\mathrm{i}$, for some $a, b \in \mathbb{R}$. Indeed writing $x = \xi + \eta\,\mathrm{i}$, for some $\xi, \eta \in \mathbb{R}$ we get

$$a + b\,\mathrm{i} = y = x^2 = (\xi + \eta\,\mathrm{i})^2 = \xi^2 - \eta^2 + 2(\xi\eta)\mathrm{i}, \tag{1.2.21}$$

which is equivalent to

$$\xi^2 - \eta^2 = a \text{ and } \xi\eta = \frac{b}{2}. \tag{1.2.22}$$

Using the second equation to substitute $\eta = {}^{b}/_{2\xi}$ in the first we get

$$4\xi^4 - 4a\xi^2 - b^2 = 0, \tag{1.2.23}$$

which is a quadratic equation for $\zeta := \xi^2$, whence

$$\zeta = \frac{1}{4}\left(2a \pm \sqrt{4a^2 + 4b^2}\right) = \frac{1}{2}\left(a \pm \sqrt{a^2 + b^2}\right). \tag{1.2.24}$$

Since $a \leq \sqrt{a^2 + b^2}$ for any $a, b \in \mathbb{R}$, the equation has a real solution ξ with ξ^2 if and only if $\zeta \geq 0$ and thus

$$\zeta = \frac{1}{2}\left(a + \sqrt{a^2 + b^2}\right), \tag{1.2.25}$$

which implies the following solutions

$$\xi = \pm\sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}} \text{ and } \eta = \frac{\pm b}{\sqrt{2\left(q + \sqrt{a^2 + b^2}\right)}}. \tag{1.2.26}$$

1.3. Polynomials

1.3.1. Definition of power. If $(G, \star)$ is a group (usually thought as a multiplicative group with neutral element $1_\star$, or simply 1 when $\star$ is the usual multiplication or clear from the context), and $x \in G$, we define the n-th power of x $x\star^n$, or simply x^n when $\star$ is the usual multiplication or clear from the context, as

$$x\star^n := \begin{cases} 1_\star & \text{if } n = 0 \\ x \star \left(x\star^{n-1}\right) & \text{if } n \geq 1. \end{cases} \tag{1.3.1}$$

1.3.2. Definition of monomial. A *monomial* in one variable (or one indeterminate) with coefficients in A is an expression of the type

$$ax^n \tag{1.3.2}$$

for a fixed integer $n \geq 0$ and $a \in A$ in a ring; the x is a *variable* or *indeterminate*. Unless otherwise stated, the x is simply a symbol. If $a \neq 0$, the integer $n \geq 0$ is called the monomial's *degree*. When $n = 0, 1, 2, 3, 4, 5$ and $a \neq 0$, the monomial is called *constant*, *linear*, *quadratic*, *cubic*, *quartic*, *quintic*. For higher degrees, we just say a monomial of degree n. By convention 0 is considered a monomial of any degree. When we need to stress the coefficient set, we also employ the word *A-monomial*.
A *monomial* in two variables is an expression of the type

$$ax^iy^j, \tag{1.3.3}$$

whose *degree* (if $a \neq 0$) is defined as $i + j$. Similarly a *monomial* in k variables is an expression of the type

$$a{x_1}^{i_1} \cdots {x_k}^{i_k} \tag{1.3.4}$$

whose degree, when $a \neq 0$, is defined to be $i_1 + \cdots + i_k$. For notational convenience, we use vectors $\boldsymbol{x} = (x_1, \ldots, x_k)$, $\boldsymbol{i} = (i_1, \ldots, i_k)$ and the generalised power notation

$$\boldsymbol{x}^{\boldsymbol{i}} = {x_1}^{i_1} \ldots {x_k}^{i_k}; \tag{1.3.5}$$

thus the monomial in (1.3.4) is simply written as

$$a\boldsymbol{x}^{\boldsymbol{i}}. \tag{1.3.6}$$

Monomials in one (respectively one or more) variable(s) are called *univariate* (respectively *multivariate*). Unless otherwise indicated when we say monomial we mean a univariate one.

1.3.3. Example (monomials). The following are monomials

$$7x^{13}, \frac{2}{3}x, \sqrt{2}x^3, xy^2, -\pi xyz, (3-\mathrm{i}/2)x^2. \tag{1.3.7}$$

The first three and the last one are univariate. The degrees are $13, 1, 3, 3, 3, 2$ respectively. The minimal coefficient fields for each are $\mathbb{Z}, \mathbb{Q}, , , , \mathbb{C}$.[4]

1.3.4. Definition of polynomial. A *polynomial* in one variable (or indeterminate) with coefficients in a ring A is an expression that can be written as the sum of one ore more monomials:

$$p(x) := a_0 + a_1 x + \dots + a_n x^n, \tag{1.3.8}$$

where n. The highest power i for which $a_i \neq 0$ is called the polynomial's degree and is denoted as $\deg p(x)$.

Similarly a polynomial in k variables is a sum of one or more monomials in those variables

$$\sum_{\boldsymbol{i}} a_{\boldsymbol{i}} \boldsymbol{x}^{\boldsymbol{i}}, \tag{1.3.9}$$

where we use the multivariable and multi-index notation introduced in and the degree of the polynomial is the highest monomial degree appearing in the sum with a nonzero coefficient.

As for monomials a polynomial is called univariate or multivariate depending on the number of its variables (or indeterminates) and we say A-polynomial when we want to emphasise the set of coefficients. (When $A = \mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$, we actually say integer polynomial, rational polynomial, real polynomial or complex polynomial, respectively.)

The set of all A-polynomials in a variable x, is denoted by $A[x]$, or $A[x_1, \dots, x_k]$ when k variables are in use.

1.3.5. Operations on polynomials. All operations on A can be ported to operations on $A[x]$ (or $A[x_1, \dots, x_k]$ for some k). For example, polynomials can be summed: for $p(x) = \sum_{i=0}^n a_i x^i$ and $q(x) = \sum_{i=0}^n b_i x^i$

$$p(x) + q(x) = \sum_{i=0}^n a_i x^i \tag{1.3.10}$$

1.4. Sets and sequences

We have been using *sets* for a while now. It is time we say something specific about them. Closely related to sets, but different in many aspects, are *sequences.* The main obstacle to understanding both sets and sequences is to get into manipulating *infinite* versions such objects. The following discussion is just a brush-up and quite informal. We will have the opportunity to come back to these concepts with a deeper look later in this course.

[4] We say minimal because that's the smallest possible ring where we can pick the coefficients.

1.4.1. Finite sets and infinite sets. Sets come in two main flavours: finite sets and infinite sets.

Finite sets are those sets that can be enumerated completely by listing their elements one by one.

Infinite sets cannot be listed completely by spelling out their elements. Nevertheless, the elements of an infinite set can be characterised by some *property* which allows us to define very clearly what the set is.

While working with finite sets is quite straightforward, as it relies on the so-called "common sense", infinite sets can be tricky (and thus more interesting) to work with. One of the objectives of this course is to get you used to working with infinite sets.

1.4.2. Example (finite sets). $\{4,5,17\}$, {persons in this room}, $\{1\}$ (note that this is different than 1), {all atoms in the universe}.[5]

Notation (curly braces)**.** Note the use of the curly braces "{ }" to enclose the *elements* of a set. *Use only curly braces to list a set's elements.*

It is important to keep the distinction between a set, as one single entity, and the list of its elements which may be a plurality. For this reason, instead of listing all their elements, sets are commonly indicated by a letter.

Also, some sets are tedious to list completely (especially the inifinite ones) so we use an English sentence to describe its elements, instead of the list.

1.4.3. Example (infinite sets). Commonly encountered infinite sets are the set of *natural numbers*

$$\mathbb{N} = \left\{n : n \text{ is a positive integer}\right\} = \{1,2,3,\dots\}, \tag{1.4.1}$$

the set of *(relative) integers*

$$\mathbb{Z} = \left\{n : n \text{ is an integer}\right\} = \{\dots,-3,-2,-1,0,1,2,3,\dots\}, \tag{1.4.2}$$

the set of all points in the plane, the set of all points on a line, the set of all lines in the plane, the set of all rational numbers (also known as fractions),

$$\mathbb{Q} = \{m/n : m \in \mathbb{Z}, n \in \mathbb{N}\} \tag{1.4.3}$$

where $m/n = m'/n'$ if and only if $mn' = m'n$), time, the set of prime numbers, the set of odd numbers, the set of even numbers $= \{2k : k \in \mathbb{Z}\}$, the set of integers that are multiples of $17 = \{17k : k \in \mathbb{Z}\}$, etc.

Note that we still use the curly braces to "enclose" the elements of an infinite set. Although many sets like $\mathbb{N}$ and $\mathbb{Z}$ can be easily written using suggestive "dots", it is important, for the more involved sets like $\mathbb{Q}$ or the set of prime numbers, to learn how to describe a set by specifying a property that characterises its elements.

1.4.4. Exercise (describing (infinite) sets). *Describe concisely (i.e., in symbols and using the curly braces) the set of prime numbers (without using the word "prime"), the set of odd numbers (without using the word "odd") and the set of rational numbers that are not integers (without using the word "integer").*

[5] It is commonly agreed by physicists that the Universe—that is the one we live in—is finite.

1.4.5. Remark (terminology). In §1.2 we have introduced the concept of *group*. You probably know by now that words used in Maths are not exactly the same as words used in common life. Mathematicians tend to be more precise with their terminology. For example, we say "group" in Maths we mean the algebraic structure defined in 1.2.3, whereas a "set" may not have such a structure. Because a group is defined on an underlying set (with an operation on that set obeying certain rules), we view each group as a set, but not every set is a group (and worse, on the same set there may be more than one group structure!).

Therefore we have placed a ban on ourselves, in that we are not allowed to use the word "group" and phrases like "a group of numbers" in Maths, except when we are really talking about a group in the sense of 1.2.3. When no particular structure is intended, a mathematician would talk about a "set of numbers". It is a good idea to mind your (mathematical) language as to avoid misunderstandings.[6]

Note that mathematicians do use synonyms. For example, synonyms for "set" are *collection, class, locus,* etc.

1.4.6. Remark (a mathie's shopping lists). Sets are like shopping lists, the order in which you list its elements does not matter (as long as you bring everything home). There is a small difference with shopping lists though: repetition does not matter, you may list an item 33 times, it will be counted only once.

For example, we have

$$\{1,2,4\} = \{4,2,1\} = \{1,2,1,2,1,2,4\} = \{4,4,1,2\} \tag{1.4.4}$$

and so on.

1.4.7. Operations on sets. Just like numbers, sets accept operations on them. Note however that the operations defined on sets are *different* than those defined on numbers.

Given two sets, say A and B, their *intersection* (also known as *meeting, meet*) is the set denoted $A \cap B$ which consists of exactly those elements which are element of A and element of B simultaneously. For example, if $A = \{1,2,3\}$ and $B = \{2,4,6\}$ then $A \cap B = \{2\}$.

Given two sets A and B, their *union* (also known as *joining, join*) is the set denoted $A \cup B$ which consists of exactly those elements that are elements of A or elements of B. For example, with A and B as above, we have

$$A \cup B = \{1,2,3,4,6\}. \tag{1.4.5}$$

Similarly the *set difference* (also known as *taking, complement*) of A and B, is the set denoted $A \smallsetminus B$ which consists of exactly those elements that are in A, but not in B. With A and B as in the previous example, we have

$$A \smallsetminus B = \{1,3\} \text{ and } B \smallsetminus A = \{4,6\}. \tag{1.4.6}$$

Finally the *cartesian product* of two sets A and B, denoted $A \times B$, consists of all the possible *ordered pairs* (x, y) where x is an element of A and y an element of B. For example, using $A = \{1,2\}$ and $B = \{3,5\}$ we have that

$$A \times B = \{(1,3),(1,5),(2,3),(2,5)\}. \tag{1.4.7}$$

[6] It must be said that the word "group" is not a very lucky one as it reminds us of no "structure", but the word was introduced before sets were even defined and has since stuck.

1.5. Sequences

Sequences are similar to sets, except they have more structure than sets. Sequences are ordered whereas sets do not have a particular order.

1.5.1. Definition of sequences (informal). Intuitively a *sequence* with *terms* in a given set X (X could be $\mathbb{N}$, or $\mathbb{Z}$, or $\mathbb{Q}$ or whatever set you one can get hold of) is an enumeration of elements of X.
Each appearance of an element is called a *term* of the sequence and has an *index* telling which position that term takes in the sequence.
The set X is the set of *values* of the sequence.
In this course, we will denote sequence with fat letter like $\boldsymbol{s}$.

1.5.2. Example (sequence). For example, you can think of X being the characters in a play and the sequence $\boldsymbol{s}$ as being the names in the dialogue:

$$X = \left\{\text{Harry, Sally, Pedro}\right\} \tag{1.5.1}$$

and

$$\boldsymbol{s} = \left(\text{Sally, Harry, Sally, Harry, Sally, Pedro, Sally, Sally, Harry,}\ldots\right). \tag{1.5.2}$$

while Sally and Sally, appearing in the sequence, are understood to be the same character (same *element*) of the *set* X, they are different instances (different *terms*) in the *sequence*. Like a film director, you should appreciate how the order of appearance matters in a sequence whereas you wouldn't care too much on where and how the cast sits when outside the camera's field.

1.5.3. Example (decimal expansion). Consider the set of *digits*

$$X := \{n \in \mathbb{N}_0 : 0 \le n \le 9\}. \tag{1.5.3}$$

X is a finite set, as it has only ten distinct elements. The digits appearing on the "right-hand side" of (i.e., "after") the decimal dot "." in the decimal expansion of a number $x \in \mathbb{R}$ forms a sequence.
For example, if $x = 1/2$ then the sequence is

$$(5, 0, 0, 0, 0, \ldots) \tag{1.5.4}$$

(note the infinite repetition of 0). If $x = 3/8$ then the sequence is

$$(3, 7, 5, 0, 0, \ldots). \tag{1.5.5}$$

If $x = 3/2$ then we get

$$(5, 0, 0, 0, 0, \ldots) \tag{1.5.6}$$

again, because we are deliberately ignoring the "left-hand side" of the expansion.
To shorten writing, denote by $\boldsymbol{d}_x$ the sequence defined above for $x \in \mathbb{R}$. Then

$$\begin{aligned}
\boldsymbol{d}_{1/9} &= (1, 1, 1, 1, 1, \ldots) && \bigl(\text{all terms have value 1}\bigr), \\
\boldsymbol{d}_{10/7} &= (4, 2, 8, 5, 7, 1, 4, 2, 8, 5, 7, 1, 4, 2, \ldots) && \bigl(\text{the pattern repeats every 6 terms}\bigr), \\
\boldsymbol{d}_{\sqrt{2}} &= (4, 1, 4, 2, 1, 3, 5, 6, 2, 3, 7, 3, 1, \ldots) && \bigl(\text{the pattern } \textit{does not} \text{ repeat}\bigr). \\
\boldsymbol{d}_{\pi} &= (1, 4, 1, 5, 9, 2, 6, 5, 3, 5, 8, 9, 7, \ldots) && \bigl(\text{the pattern } \textit{does not} \text{ repeat}\bigr).
\end{aligned} \tag{1.5.7}$$

Note how in all these examples the *values of the sequence* are finite, but the number of terms in each sequence is *infinite*. Of particular importance in mathematics is the fact that the two last sequences have no repeating pattern.[7]

1.5.4. Properties of sequences. A sequence $\boldsymbol{s}$ in a set X satisfies the following properties:

1. All terms of $\boldsymbol{s}$ belong to (or take value in) the given set X.
2. Each term has an index, indicating its appearance order in the sequence. Terms of a sequence $\boldsymbol{s}$ are denoted by s_k, where k is the *index* in $\mathbb{N}$ (or $\mathbb{N}_0$ if you are a hacker[8])
3. There is a first (element) term, let us denote it by s_1 (s_0 for hackers).
4. Each term has at most one predecessor.
5. Each term has at most one successor, we denote the successor of s_k by s_{k+1}.
6. At most one term may be the last. If there is a last term we say that $\boldsymbol{s}$ is a *finite sequence*, otherwise (i.e., if there is no last element) then we say that $\boldsymbol{s}$ is an infinite sequence.

Notation (for (finite and infinite) sequences)**.** A finite sequence is usually written in one of the following ways

$$(s_1, s_2, \dots, s_N),\text{ or }(s_1, \dots, s_N),\text{ or }(s_n)_{n=1,\dots,N},\text{ or }(s)_n 1 \leq n \leq N. \tag{1.5.8}$$

Here the number N is called the *length* of the sequence $\boldsymbol{s}$.
An infinite sequence is usually written as follows

$$(s_1, s_2, \dots),\text{ or }(s_1, s_2, \dots, s_n, \dots),\text{ or }(s_n)_{n\in\mathbb{N}}. \tag{1.5.9}$$

Infinite sequences have no length associated to them.

1.5.5. Symbolics: brackets matter. Note that we use curly braces "{ }"to enclose elements of a set (whether we list them or we give a property characterising them), whereas sequences are listed by using round brackets (parentheses) "( )".
In this course, we denote sequences by bold face sans-serifed type, e.g., $\boldsymbol{a}$, $\boldsymbol{g}$, $\boldsymbol{x}$, and the corresponding terms with the corresponding letter in the italic type, e.g., a_n, g_k, x_i, respectively. Note that the index is not always n, but it may be some other symbol.

1.5.6. Example (order matters for sequences). For sequences we have

$$(5,5,5,5,5) \neq (5)\text{ and }(1,2,3) \neq (3,2,1), \tag{1.5.10}$$

but for sets we have

$$\{5,5,5,5,5,5\} = \{5\}\text{ and }\{1,2,3\} = \{3,2,1\}. \tag{1.5.11}$$

[7] The pedant will have noticed that the notation $\boldsymbol{d}_x$ is a bit dangerous, in that for some $x \in \mathbb{R}$ there are *two* different sequences of digits for the expansion, e.g., $\boldsymbol{d}_1 = (0,0,0,\dots)$ but also $\boldsymbol{d}_1 = (9,9,9,\dots)$. We leave the pedant to sorting this out, while we avoid dwelving into resolving such sophistication. . . yet.

[8] In this course, we use the word "hacker" for what it originally meant, i.e., *a person who likes to solve problems by writing computer code.* In a sensationalim ridden society, the word "hacker" has been hi-jacked by (mostly computer-illiterate) journalists to indicate a "cracker", is a person who wastes time trying to sneak in and unlawfully extract information from other's computer systems (a bit like journalists). It is a (journalist fuelled) myth that "good" crackers must be good hackers. It is a true fact, though, that many good mathematicians are good hackers, and vice-versa (maybe because they both need to think logically). Hackers are thus generally not dangerous persons, in fact they can be very useful (unlike crackers and journalists), but they are known to start counting from 0, rather than 1.

1.5.7. Terminology: is a sequence finite or infinite? Since infinite sequences are more interesting than finite sequences (at least for mathematicians), it is commonly understood by “sequence” to mean “infinite sequence”, unless otherwise stated. We will be following this convention and we shall only specify it only when a sequence is finite.

1.5.8. Constructing infinite sequences. We have seen that many infinite sets can be completely described by giving some property characterising their elements. Can we do something similar for (infinite) sequences?
One way to build sequences, is to note that $\mathbb{N}$ yields a natural sequence structure:

$$\boldsymbol{N} := (1,2,3,4,\dots). \tag{1.5.12}$$

We may think of the sequence $\boldsymbol{N}$ as the mother of all sequences. In fact a sequence $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$ can be seen as a *rule* that assigns to each element $n \in \mathbb{N}$ an element a_n in a given set. If we know the rule, we completely determine the sequence. In fact, we may think of the sequence itself *being* the rule. Note that, in spite them being very similar, we still like to think that the *set* $\mathbb{N}$ and the *sequence* $\boldsymbol{N}$ are two separate objects.

1.5.9. Example (sequences through rules). We show now how to “build” sequences based on the basic sequence $\boldsymbol{N}$.

- ⋆ Consider the rule $n \mapsto n^2$. This produces the sequence

  $$(1,4,9,16,25,\dots). \tag{1.5.13}$$

  Note that the rule itself is much more precise than the attempt to enumerate the sequence which inevitably fails with those dots. In fact the displayed expression does not rule out the possibility of having the 6-th term of the sequence being 3, whereas the rule clearly tells us this must be 36.
- ⋆ The rule $k \mapsto 4\sin(\pi k/6)^2$, produces the sequence

  $$(1,3,4,3,1,0,1,3,4,3,\dots) \tag{1.5.14}$$

  Again note how the rule is much more useful in determining any term of the sequence.
- ⋆ Consider rule $k \mapsto 1/k$, then the sequence produced is

  $$(1,1/2,1/3,1/4,\dots) \tag{1.5.15}$$

  This is a sequence in $\mathbb{Q}$. Sequences are not only made of whole numbers.
- ⋆ Consider the rule where h_k is the highest factor of $k+1$ different than $k+1$ itself. This yields sequence

  $$\boldsymbol{h} = (1,1,2,1,3,1,4,3,5,1,6,1,7,5,8,1,9,1,10,7,\dots). \tag{1.5.16}$$

  This example is quite telling: if you saw the first terms of the sequence written like that, it would not be very easy to guess the whole sequence. This shows the superiority of defining a sequence through a rule, rather than just listing some of its initial terms and hoping that the reader will come up with the rule.

1.5.10. Exercise (building sequences via rules). *This exercise needs a colleague.*

part 1. *Hiding your work from your colleague, and following Example 1.5.9, make up three "rules" that transform natural numbers into some other numbers. On sheet A, write down the first* 10 *terms of each of the sequences, and on sheet B write* 10 *more terms for each sequence.*

part 2. *Challenge your colleague to find your transformation rules by showing them only sheet A. You get a point if they fail to find the next* 10 *terms, otherwise your colleague gets the point.*

part 3. *Reverse the roles, whereby your colleague makes up three rules and challenges you to guess them.*

part 4. *Report your examples and the overall score.*

1.5.11. Recursion. We have seen a general way to define sequences, by producing some rule that binds elements of $\mathbb{N}$ to each term of the sequence. But there are other ways of producing sequences, such as *recursive sequences.*

To build a sequence, say $\boldsymbol{a}$, in a set X recursively, we first need a unary operator on X

$$\begin{array}{rccc} F: & X & \to & X \\ & x & \mapsto & F(x) \end{array}, \tag{1.5.17}$$

which we shall call the *generating function* and a *seed* $s \in X$. The *recursive sequence* with generating function F and seed s is the sequence $\boldsymbol{a} = (a_1, a_2, \ldots, a_n, \ldots)$ defined by

$$a_n := \begin{cases} s & \text{if } n = 1, \\ F(a_{n-1}) & \text{if } n \geq 2. \end{cases} \tag{1.5.18}$$

This construction can be generalised further by considering a binary operator on X,

$$\begin{array}{rccc} F: & X \times X & \to & X \\ & (x, y) & \mapsto & F(x, y) \end{array} \tag{1.5.19}$$

and *two seeds,* s and t. Then we may define a sequence $\boldsymbol{a}$ through the double recursion

$$a_n := \begin{cases} s & \text{if } n = 1, \\ t & \text{if } n = 2, \\ F(a_{n-2}, a_{n-1}) & \text{if } n \geq 3. \end{cases} \tag{1.5.20}$$

1.5.12. Example (Fibonacci numbers). A famous example of a double recursion is the construction of the sequence of *Fibonacci numbers,* $\boldsymbol{f}$ whereby

$$f_n := \begin{cases} 1, & \text{if } n = 1, 2, \\ f_{n-2} + f_{n-1}, & \text{if } n \geq 3. \end{cases} \tag{1.5.21}$$

1.5.13. Exercise (Fibonacci numbers).

(a) *Identify the generating function of the Fibonacci sequence.*

(b) *Write down the first eight terms of the Fibonacci sequence.*

Recursion can be also generalised in another direction, where the generating function can depend on the index. To make this clear let us focus on a particular case.

1.5.14. Example (summation as step-dependent recursion). Let $\boldsymbol{a}$ be a given sequence, and define

$$s_n := \begin{cases} a_1, & \text{if } n = 1, \\ s_{n-1} + a_n, & \text{if } n \geq 2. \end{cases} \tag{1.5.22}$$

The sequence $\boldsymbol{s}$ thus obtained is actually the sequence of summations, meaning

$$s_n = (\cdots((a_1 + a_2) + a_3) \cdots + a_n). \tag{1.5.23}$$

While this notation makes intuitive sense, the recursion definition is much more precise as it does away with all those dots. The generating function here changes with each $n \in \mathbb{N}$, namely

$$F_2(x) = x + a_2, \dots, F_n(x) = x + a_n, \dots, \tag{1.5.24}$$

so in fact we are dealing with a sequence of generating functions. Nevertheless, this is still a good way to define a sequence, it is called a *step-dependent recursion.* Generally, suppose $(F_n)_{n\in\mathbb{N}}$ is a sequence of unary operators on X, and let $s \in X$, then we can define the step-dependent recursive sequence through

$$s_n := \begin{cases} s, & \text{if } n = 0, \\ F_n(s_{n-1}), & \text{if } n \geq 1. \end{cases} \tag{1.5.25}$$

One of the most important examples of step-dependent recursive definitions is the summation of which we now give a definition which includes the notation $\sum$.

1.5.15. Definition of summation. Suppose $\boldsymbol{a}$ is a sequence of numbers (finite or infinite) we define

$$\sum_{k=1}^{n} a_k = \begin{cases} 0 & \text{for } n = 0, \\ \left(\sum_{k=1}^{n-1} a_k\right) + a_n & \text{for } n \geq 1 \end{cases} \quad \big(\text{and } n \leq \operatorname{length} \boldsymbol{a} \text{ if } \boldsymbol{a} \text{ is finite}\big). \tag{1.5.26}$$

1.5.16. Problem (products of n terms). *Given a sequence $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$, write down a step-dependent recursion that defines the sequence $\left(\prod_{k=1}^{n} a_k\right)$ whose n-th term is the product of the first n terms of $\boldsymbol{a}$.*
Hint. It is very much like $\sum$ but make sure you put the right seed.

1.6. The Induction Principle

Closely related to recursion is the so-called *Principle of Mathematical Induction (PMI)*, more simply referred to as *Induction Principle* or *Induction.* Induction a technique that is quite useful to make sure that certain statements about numbers are true. The process of making sure that something is true is called (mathematical) *verification* or *proof.*

Big Fat Note (a warning about Induction)**.** While Induction is a powerful tool to prove statements, many statements cannot be proved using Induction. So it is not a universal way of proving things. In fact, there is no universal technique to prove all statements (which, incidentally, is why mathematicians still have a place under the sun).

After this sad note, let us proceed with happier affairs and see how Induction can be used. Induction is the mathematicians domino effect, except we have infinitely many domino pieces. Let us call them d_n and put them into a sequence $\boldsymbol{d}$. To build a domino effect we need to make sure of two things:

(a) when we set up the dominoes, we make it in such a way that *if* a domino d_n falls, *then* the next one d_{n+1} will fall too, and this must be true *for each* pair of dominoes (d_n, d_{n+1})

(b) once the set-up is done, we must hit the first domino, d_1, in the sequence $\boldsymbol{d}$.

If we ensure both (a) and (b) happen, then we can be sure that all dominoes will fall. To phrase this in more mathematical terms, suppose we have a certain *formula* (also known as *predicate*) let's call it $P(\cdot)$ (or just P), so that applied to each $n \in \mathbb{N}$ it gives us a *proposition* $P(n)$, which could be true or false. It makes sense to ask whether $P(n)$ is true, for all $n \in \mathbb{N}$. To fix ideas, let us consider some examples of formulas:

$$P(n) \Leftrightarrow (n \text{ is a prime number}) \tag{1.6.1}$$

$$Q(n) \Leftrightarrow \left(n^2 + 2n + 1 \text{ is not prime}\right) \tag{1.6.2}$$

$$S(n) \Leftrightarrow \left(\sum_{k=1}^{n} k = n(n+1)/2\right) \tag{1.6.3}$$

Recall that a number $m \in \mathbb{N}$ is called prime if the only ways it can be written as a product are $m \times 1$ and $1 \times m$.

In the first case, we know that $P(n)$ cannot be true for all $n \in \mathbb{N}$. Indeed $P(4)$ is false because $4 = 2 \times 2$ which makes 4 not prime.

In the second case, we know, from basic algebra that $n^2 + 2n + 1 = (n+1)^2$ and for $n \geq 1$, being $n + 1 \geq 2$ it follows that $n^2 + 2n + 1$ is composite (i.e., not prime) and thus $Q(n)$ is true for all $n \in \mathbb{N}$.

In the third case, it is not so easy to tell. We could try some cases, just to see if we get something wrong

$$\begin{gathered}
\sum_{k=1}^{1} k = 1 \text{ and } \frac{1(1+1)}{2} = 1 \\
\sum_{k=1}^{2} k = 1 + 2 = 3 \text{ and } \frac{2(2+1)}{2} = 3 \\
\sum_{k=1}^{3} k = 1 + 2 + 3 = 6 \text{ and } \frac{3(3+1)}{2} = 6.
\end{gathered} \tag{1.6.4}$$

Nothing wrong so far: $S(n)$ is true for $n = 1, 2, 3$. It is futile, though, to continue along these lines, because we would have to check $S(n)$ for an infinite number of cases, and life is too short for that. This is where Induction kicks in: it allows us to prove infinitely many statements in one go.

1.6.1. Theorem (Principle of Mathematical Induction—PMI). *Let P be a formula that provides a statement $P(n)$ for each $n \in \mathbb{N}$, and suppose that the following hold:*

(a) if $P(n)$ were true, then $P(n+1)$ would also be true, for all $n \in \mathbb{N}$,

(b) $P(1)$ is true.

Then $P(n)$ is true for each $n \in \mathbb{N}$.

Proof A proof of the Principle of Mathematical Induction can be derived from the so-called "Axiom of Infinity" (also known as Peano's Axiom), to be studied later in this course. Such a proof is beyond the scope of this course though.[9] □

1.6.2. Induction: terminology and remarks. Condition (a) is called the *Inductive Step*. It is important to appreciate that the Inductive Step *does not say that* $P(n)$ *is actually true*; rather, it says that $P(n+1)$ *would* be true *if* $P(n)$ *happens to be true*. In symbols the Inductive Step can be written as

$$(P(n) \Rightarrow P(n+1)). \tag{1.6.5}$$

$P(n)$ is called the *inductive hypothesis*, and as the name says it, it is quite hypothetical at this stage.

Note that (1.6.5) is very different from saying

$$\forall\, n \in \mathbb{N} : P(n) \text{ is true,} \tag{1.6.6}$$

which is just an abbreviated version of the sentence

$$\text{for all } n \in \mathbb{N}, \text{ the proposition } P(n) \text{ holds true.} \tag{1.6.7}$$

Indeed, proposition (1.6.7)—or its equivalent in symbols (1.6.6)— constitutes the conclusion of Theorem 1.6.1 and tells us that $P(n)$ is true for all $n \in \mathbb{N}$, whereas proposition (1.6.5) does not say that $P(n)$ is true.

The logical difference between (a) and the actual conclusion of the Theorem is, in our physical domino analogy, the difference between *setting up* each pair of successive dominoes so that the next falls *if* the previous one, and the *actual fall of the dominoes*, respectively.

Condition (b) of Theorem 1.6.1 is called the *base case* of the induction process. Without it the conclusion cannot be guaranteed and it must be checked. It is usually not a very hard task. The main difficulty being in checking that (a) works.

In order to apply the PMI, we need to check that both the inductive step (a) (for *all* $n \in \mathbb{N}$) and the base case (b) are satisfied. Once this is done it follows that the statement $P(n)$ is true for all $n \in \mathbb{N}$.

1.6.3. Example (Induction Principle in action). Let us go back to the example given by (1.6.3). The proposition is true.

Problem. *Using the Induction Principle, prove that*

$$\forall\, n \in \mathbb{N} : \sum_{k=1}^{n} k = n(n+1)/2. \tag{1.6.8}$$

Solution. In this case the formula P is given by

$$P(m) \Leftrightarrow \left(\sum_{k=1}^{m} k = m(m+1)/2 \right). \tag{1.6.9}$$

[9] If you are interested in the details about the Axiom of Infinity, you should check Halmos, 1974. But, make sure you master this course before embarking on such more advanced reading.

To apply Induction we need to prove the inductive step and the base case. To organise the work let us define

$$L(n) = \sum_{k=1}^{n} k \text{ and } R(n) = n(n+1)/2, \tag{1.6.10}$$

so that $P(n)$ can be written as $L(n) = R(n)$.

Inductive step: To show this, fix $n \in \mathbb{N}$, and suppose that $P(n)$ is true, i.e., $L(n) = R(n)$: this is the *inductive hypothesis.*

We want to deduce that $P(n+1)$ is true, i.e., $L(n+1) = R(n+1)$

This can be done as follows

$$\begin{aligned} L(n+1) &= \sum_{k=1}^{n+1} k \\ &= L(n) + (n+1) \qquad \text{(by definition of } L) \\ &= R(n) + (n+1) \qquad \text{(by the inductive hypothesis).} \end{aligned} \tag{1.6.11}$$

On the other hand we have

$$\begin{aligned} R(n+1) &= \frac{1}{2}(n+1)((n+1)+1) \\ &= \frac{1}{2}(n+1)(n+2) \\ &= \frac{1}{2}(n+1)n + (n+1) \\ &= R(n) + (n+1) \end{aligned} \tag{1.6.12}$$

Therefore $L(n+1) = R(n+1)$, as desired.

Since the fixed n was chosen arbitrarily in $\mathbb{N}$, this means that we have shown the inductive step for all $n \in \mathbb{N}$, as required to satisfy (a) in Theorem 1.6.1.

base case: This is easily satisfied: $L(1) = \sum_{k=1}^{1} k = 1$ (by definition) and $R(1) = 1 \times (1+1)/2 = 1$, thus $L(1) = R(1)$, as required.

By Theorem 1.6.1, it follows that $L(n) = R(n)$ for all $n \in \mathbb{N}$.

1.7. Comparison and ordering of numbers

In the introductory discussion about numbers, we have seen that numbers can be compared, on top of being operated on. Comparison of numbers, means that we are able to order them, from the smallest to the biggest. While ordering can be done in a meaningful way on $\mathbb{R}$ (and thus on all its subsets, such as $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{N}_0$ and $\mathbb{N}$), in $\mathbb{C}$ it turns out that ordering is not so useful. Let us therefore have a quick glance at ordering in $\mathbb{R}$.

1.7.1. Ordering of (real) numbers $(\mathbb{R}, +, \times, \leq)$**.** Given two numbers, say x and y in $\mathbb{R}$, then one and only one of the following (mutually exclusive) situations will happen:

(a) $x < y$, i.e., x is *less (or smaller) than* y,

(b) $y < x$,

(c) $x = y$, i.e., x is *equal to* (or *equals*) y.

This means that any two real numbers can always be compared (and either one of the two wins, or they are the same number). This property is called *trichotomy* of the real numbers.
The notation "$x < y$" is often equivalently written as "$y > x$", and we may say that y is *more (or greater, or bigger) than x*.
For example, the following are all true: $2 < 3$, $0 < 1.2$, $-14 < 3.5$, $-\pi < \pi$, $4 > 3$, $7/2 > 3$, $-1/2 = -(-1/2)$, $7/2 = 3.5$, $3.5 > 3$, etc. Note that each relation excludes the other two: for example knowing that $7/2 > 3$ means that $7/2 \neq 3$ and $7/2 \nless 3$. (When some relation such as $x = y$ is *not true* we cross it by a "slash", $x \neq y$.)
Situations like $x < y$ or $x > y$ are referred to as *inequalities*, whereas $x = y$ is called an *equality* (or *identity* when it happens to many x and y's). For example, $2 = 4/2$ is usually called an equality, and $2a/4 = a/2$ is called an identity (because it is true for all a).

1.7.2. Definition of weak (or loose) ordering, inequalities. Sometime it is useful to indicate that only two of the situations (a)– (c) may occur. Given $x, y \in \mathbb{R}$, we say that x is *less* (than) *or equal* to y, if and only if $x < y$ or $x = y$. We say that x is *greater* (than) *or equal* to y, if and only if $x > y$ or $x = y$ and we write $x \geq y$ in this case. Notice that "$x < y$ or $x > y$" is the same as "$x \neq y$" because of trichotomy.

1.7.3. Example (inequalities). The following are true $3 \leq 3.5$, $0 \leq 1$, $0 \leq a^2$, $0 \geq -a^2$, etc.

1.7.4. Transitivity. A basic property of the ordering $<$ is the so-called *transitive property* (or transitivity) which says that if $x < y$ and $y < z$ then we have $x < z$, where $x, y, z \in \mathbb{R}$. This seemingly trivial rule is quite important.
It is not very hard to show that $\leq$ also satisfies the transitive property on $\mathbb{R}$.[∗] [∗]: Check!

1.7.5. Anti-symmetry and weak anti-symmetry. Another property orderings is the *anti-symmetric property* (or *anti-symmetry*). The strict ordering $<$ is anti-symmetric, which is to say

$$x < y \Rightarrow y \nless x, \tag{1.7.1}$$

for all $x, y \in \mathbb{R}$. This is a direct consequence of the mutual exclusivity of (a)–(c) in §1.7.1.[∗] [∗]: Check!
On the other hand, it may very well happen that $x \leq y$ and $y \leq x$ are satisfied. But, again due to the mutual exclusivity of (a)–(c), it follows that x must equal y if both inequalities are satisfied. In symbols, we have

$$\big(x \leq y \text{ and } y \leq x\big) \Rightarrow x = y. \tag{1.7.2}$$

This property of $\leq$ is quite handy and it is used often (especially in Analysis) to show that two numbers are in fact the same number.

1.7.6. Definition of positive and negative numbers. The number zero, 0, plays an important role. It sorts the real numbers, a number $x \in \mathbb{R}$ is called

- ⋆ *positive* whenever $x > 0$,
- ⋆ *negative* whenever $x < 0$.

Some authors use the term positive (negative) for $x \geq 0$ ($x \leq 0$ respectively). We shall say *non-negative* (*non-positive*) for $x \geq 0$ ($x \leq 0$ respectively).

1.7.7. Notation for sets of positive numbers. If $\mathbb{E}$ is one of the sets $\mathbb{R}$, $\mathbb{Q}$ or $\mathbb{Z}$ we denote by $\mathbb{E}^+$ the set of all positive numbers in $\mathbb{E}$, in symbols

$$\mathbb{E}^+ := \{x \in \mathbb{E} : \ x > 0\} \tag{1.7.3}$$

Similarly we define

$$\mathbb{E}^- := \{x \in \mathbb{E} : \ x < 0\}, \tag{1.7.4}$$

$$\mathbb{E}_0^+ := \{x \in \mathbb{E} : \ x \geq 0\}, \tag{1.7.5}$$

$$\mathbb{E}_0^- := \{x \in \mathbb{E} : \ x \leq 0\}, \tag{1.7.6}$$

$$\mathbb{E}^* := \{x \in \mathbb{E} : \ x \neq 0\}. \tag{1.7.7}$$

1.7.8. Basic relations between ordering and algebra. Ordering of real numbers interacts with the algebraic operations + and × therein. The rules of interaction are called *compatibility* and can be written as follows.

cancellation and compatibility of ordering with addition (or sum): The following properties are equivalent

$$x < y, \tag{1.7.8}$$

$$\exists\, z \in \mathbb{R} : x + z < y + z, \tag{1.7.9}$$

$$\forall\, z \in \mathbb{R} : x + z < y + z. \tag{1.7.10}$$

The case "$x+z < y+z \Rightarrow x < y$" is usually called *cancellation property*. The case "$x < y \Rightarrow x + z < y + z$" usually called *invariance of inequalities under addition* or *compatiblity of ordering with addition* (or *sum*).

cancellation and invariance under multiplication: Suppose $z \in \mathbb{R}^+$ where $\mathbb{R}^+ := \{x \in \mathbb{R} : \ x > 0\}$, then the following facts are equivalent

$$x < y, \tag{1.7.11}$$

$$\exists\, z \in \mathbb{R}_+ : xz < yz, \tag{1.7.12}$$

$$\forall\, z \in \mathbb{R}_+ : xz < yz. \tag{1.7.13}$$

Again we talk about *cancellation* when $xz < yz \Rightarrow x < y$ and *invariance* (or *compatibility*) when $x < y \Rightarrow xz < yz$. Note that z *must be strictly positive* for the last two conditions to be true.

positivity's characterisation: The following are equivalent

$$x > 0, \tag{1.7.14}$$

$$\exists\, y, z \in \mathbb{R} : y < z \text{ and } xy < xz, \tag{1.7.15}$$

$$y < z \Rightarrow xy < xz. \tag{1.7.16}$$

1.7.9. Exercise (product, ordering and simplification). *Refering to statements in §1.7.8, can you replace "$<$" by "$\leq$"? Rewrite the above with $\leq$ instead of $<$ and check whether the statements are true.*
Hint. Be careful with the 0 *and multiplication.*

1.7.10. Exercise (product and ordering).

(a) *Deduce from the basic relations between ordering and algebra, that for $x<0$ and $y<z$, one has $xy>xz$.*

Hint. Argue by contraposition: suppose $y<z$ and $xy\leq xz$ and show that it must then be $x\geq 0$.

(b) *Based on the cancellation laws the algebraic properties and the fact $1\neq 0$ for all prove that $1>0$.*

Hint. Use a contradiction argument.

1.7.11. Remark. The relations between ordering and algebra will work also, as stated, if $\mathbb{R}$ is replaced by one its subsets, such as $\mathbb{Q}, \mathbb{Z}$, $\mathbb{N}_0$ and $\mathbb{N}$.
The last two subsets though, have an interesting property that $\mathbb{R}$, $\mathbb{Q}$ and $\mathbb{Z}$ do not have:

There exists a smallest element in $\mathbb{N}$ and $\mathbb{N}_0$.

We will come back to this important property of positive (or non-negative) integers later on in this chapter.

1.7.12. Problem. *In all this problem take $a,b,c,d\in\mathbb{R}$. Prove the following results using only first principles in the Lecture Notes and results derived so far.*

(a) *If $a>0$ then $-a<0$.*

(b) *If $a>0$ and $b\geq 0$ then $a+b>0$.*

Hint. Use transitivity as explained in §1.7.4.

(c) *$a<b$ if and only if there exists $r\in\mathbb{R}_+$ such that $a+r=b$.*

(d) *If $a<b$ and $c<d$ then $a+c<b+d$.*

Solution. (a) Using invariance of order under addition of $-a$, from $a>0$ we get $0=a-a>0-a=-a$. Thus $-a<0$.

(b) Using invariance we have that $a+b>0+b=b$, thus $a+b\geq b$ but we also have $b\geq 0$. Hence, thanks to transitivity of $\leq$, we obtain $a+b\geq 0$. To finish we need to exclude the equality: by contradiction suppose $a+b=0$, then, subtracting a from both sides, we get $b=-a$. Recalling that $a>0$ and using (a) we obtain $b<0$, which is contrast with $b\geq 0$. So $a+b\neq 0$ and it follows that $a+b>0$ as claimed.

(c) Suppose first that $a<b$, then by invariance under addition, we have $a-a<b-a$, which means that $b-a>0$ and $b-a=:r$ is what we want.

Conversely, suppose $a+r=b$ for $r>0$ then $b-a=r>0$ and by invariance we obtain $b=b-a+a>0+a=a$, thus $b>a$.

(d) Let $a<b$ and $c<d$ from (c) we have that $a+r=b$ and $c+s=d$. By invariance it follows that $a+r+c+s=b+d$, and thus $(a+c)+t=b+d$ with $t>0$. Thus, thanks to the equivalence in (c), we get that $a+c<b+d$ as required.

An alternative proof uses transitivity directly: since $a<b$, then by adding c and using invariance of $<$ under $+c$, we get

$$a+c<b+c. \tag{1.7.17}$$

Similarly, starting from $c<d$ and adding b we get

$$b+c<b+d. \tag{1.7.18}$$

Using transitivity we obtain $a+c<b+d$.

Exercises and problems on numbers

Exercise 1.X.1 (groups). (a) Is $(\mathbb{Z}, \times)$ a group? Explain your answer.

(b) Why is $(\mathbb{N}, +)$ not a group?

(c) Let $\mathbb{N}_0 := \{0, 1, 2, 3, \ldots\}$ (i.e., the set of all non-negative integers). Is $(\mathbb{N}_0, +)$ a group?

(d) Which of the group properties of $(\mathbb{Z}, +)$ remain valid for $(\mathbb{N}, +)$ and which do not? Same question for $(\mathbb{N}_0, +)$.

Exercise 1.X.2 (multiplication of rational numbers). Let $\mathbb{Q}$ be the set of all possible fractions of integers, that is all numbers of the form m/n with $m, n \in \mathbb{Z}$ and $n \neq 0$, with the convention that two fractions m/n and m'/n' represent the same number, i.e., $m/n = m'/n'$, if and only if $mn' = m'n$.

(a) Recalling that multiplication of fractions is defined by

$$\frac{m}{n} \times \frac{k}{l} := \frac{mk}{nl}, \tag{1.X.2.1}$$

and using the properties of × in $\mathbb{Z}$ show that × is associative, commutative and that $1/1$ is its neutral element in $\mathbb{Q}$.

(b) Show that $m/m = 1/1$ in $\mathbb{Q}$, for any $m \in \mathbb{Z}$ and $m \neq 0$.

(c) Let $\mathbb{Q}^*$ be the set $\mathbb{Q}$ with the fraction $0/m$ taken away (note that $0/n = 0/m$ for any $n, m \in \mathbb{N}$). Explain why $(\mathbb{Q}^*, \times)$ is a group.

Hint. Show that for each number $q = m/n \in \mathbb{Q}^*$ it is possible to find a number $q' \in \mathbb{Q}^*$ such that $qq' = 1/1$ (i.e., each q has a *multiplicative inverse* q').

(d) Is $(\mathbb{Q}, \times)$ a group?

Exercise 1.X.3 (describing (infinite) sets). Describe concisely (i.e., in symbols and using the curly braces) the set of prime numbers (without using the word "prime"), the set of odd numbers (without using the word "odd") and the set of rational numbers that are not integers (without using the word "integer").

Exercise 1.X.4 (sets and sequences of numbers). (a) For each of the following statements, say whether it is true or false. Explain briefly your answer.

$$\{1, 2, 3, 4\} = \{4, 3, 2, 1\}, \tag{1.X.4.1}$$

$$\{1, 2, 3\} \neq \{1, 2, 3, 2\}, \tag{1.X.4.2}$$

$$(1, 2, 3, 4) = (4, 3, 2, 1), \tag{1.X.4.3}$$

$$(1, 2, 3) \neq (1, 2, 3, 2). \tag{1.X.4.4}$$

(b) Give four different infinite sets contained in $\mathbb{N}$. (Remember that while an infinite set cannot be listed, it can be described by a "rule" that characterises its elements, e.g., the set of even numbers is the set of numbers of the form $2n$ where $n \in \mathbb{N}$, briefly written as $\{2n : n \in \mathbb{N}\}$.) Give a good reason why the sets you are stating are infinite and why they are different.

(c) Give one example of each of

(i) a sequence that is infinite and has infinitely many values,
(ii) a sequence that is infinite but has finitely many values.

(d) Can you put all the (relative) integers in a sequence? What about all the rational numbers (fractions)? You are allowed to have repeated values, as long as you cover all the numbers. Explain your answers; "pub-quiz" answers, such as "yes" or "no", have zero value unless backed by explanations.

Exercise 1.X.5 (arithmetic progressions)**.** Calculate (or simplify) using pen, pencil and paper only, the following sums

$$1+2+3+\cdots+498+499+500, \tag{1.X.5.1}$$

$$5+6+7+\cdots+53+54+55, \tag{1.X.5.2}$$

$$(n+1)+\cdots+(n+m-1)+(n+m), \text{ for } m,n\in\mathbb{N} \tag{1.X.5.3}$$

$$20+22+24+\cdots+106+108+110, \tag{1.X.5.4}$$

Exercise 1.X.6 (the sum of successive squares)**.** (a) Using induction, prove that

$$\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1), \ \forall\, n\in\mathbb{N}. \tag{1.X.6.1}$$

(b) Using the previous result show that for any integer $n\in\mathbb{N}$, (at least) one of the three numbers n, $n+1$, $2n+1$ must be divisible by 3.

Exercise 1.X.7 (the sum of successive cubes)**.** Johnson:98:book:Elements
(a) Prove that for every natural number n,

$$\sum_{k=1}^{n} k^3 = \frac{1}{4}n^2(n+1)^2. \tag{1.X.7.1}$$

(b) Can you think of any natural reason for the fact that this is equal to $\left(\sum_{k=1}^{n} k\right)^2$?

Exercise 1.X.8 (factorial and sum)**.** (a) Let $n\in\mathbb{N}_0$ be a nonnegative integer number, give the recursive definition of its *factorial* $n!$.
(b) Prove that for each $n\in\mathbb{N}$

$$\sum_{k=1}^{n} k(k!) = (n+1)!-1. \tag{1.X.8.1}$$

Problem 1.X.9 (basic consequences of algebraic laws in $\mathbb{Z}$)**.** Each result in this exercise depends on the preceeding ones, but they can be solved indipendently.
(a) Suppose $e,a\in\mathbb{Z}$ and $e+a=a$. Show, based on first principles, that $e=0$. Justify each single step with basic algebraic properties of addition.
(b) Show, based on first principles (i.e., the basic properties of the sum) that if $a+a=a$, $a\in\mathbb{Z}$, then it must be $a=0$.
(c) Show that 0 is an *absorbing element* for ×, i.e., that

$$\forall\, a\in\mathbb{Z}: 0a=0, \tag{1.X.9.1}$$

using the basic properties of + and × in $\mathbb{Z}$ and the previous results.

Hint. Start by summing $0a$ to itself and work that out.

(d) Denote by -1 the opposite of 1. Drawing from first principles and previously shown results, prove that $(-1)a=-a$ for any $a\in\mathbb{Z}$.

Hint. Prove that $a+(-1)a=0$.

Problem 1.X.10 (Induction for inequalities). (a) Guess, by trial and error, the smallest $n_0 \in \mathbb{N}$ for which the statement

$$2^n > n^3, \text{ for all } n \geq n_0, \tag{1.X.10.1}$$

is true.

(b) Prove that this statement is true. (Use induction.)

(c) Could you do the second part of this exercise by trial and error? Why?

Problem 1.X.11 (sum rules). Let $(a_k)_{k=1,\dots,n}$ and $(b_k)_{k=1,\dots,n}$ be two finite sequences of numbers and c a number.

(a) Complete the boxes in the following recursive definition of $\sum_{k=1}^{n} a_k$:

$$\sum_{k=1}^{n} a_k := \begin{cases} a_1, & \text{if } \boxed{\quad\texttt{[01]}}, \\ \boxed{\qquad\qquad\texttt{[02]}}, & \text{if } n \geq 2. \end{cases} \tag{1.X.11.1}$$

(b) Show, by using the basic properties of sum and product and induction on n, *only one* of the following well-known formulas (also known as "sum-rules") are true (do not worry too much about their names) for each $n \in \mathbb{N}$:

$$c\left(\sum_{k=1}^{n} a_k\right) = \sum_{k=1}^{n} (c\,a_k) \qquad \text{(distributive law)}, \tag{1.X.11.2}$$

$$\sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} (a_k + b_k) \qquad \text{(additivity)}, \tag{1.X.11.3}$$

$$\sum_{k=1}^{n} a_k = \sum_{i=1}^{n} a_i \qquad \text{(index rename)}, \tag{1.X.11.4}$$

$$\sum_{k=2}^{n} a_k = \sum_{k=1}^{n-1} a_{k+1} \qquad \text{(index shift)}, \tag{1.X.11.5}$$

$$\sum_{k=1}^{m} a_k + \sum_{k=m+1}^{n} a_k = \sum_{k=1}^{n} a_k, \ \forall\, m : 1 \leq m \leq n-1 \qquad \text{(associative law)}. \tag{1.X.11.6}$$

The last one requires a double induction on m and n.

Hint. Although in theory the base case has to be proven only for $n = 1$, you will find it useful to try $n = 2, 3, 4$ before proceeding to the general inductive step.

(c) Suppose now that $(b_k)_{k=1,\dots,n}$ is a *permutation* of $(a_k)_{k=1,\dots,n}$, i.e., for each k between 1 and n, there exists one $j = j(k)$ between 1 and n, such that $a_k = b_{j(k)}$. (In other words, one of the sequences can be obtained from the other by reordering it.) Show, by induction on n and using (1.X.11.6), the *(generalised) commutative law*

$$\sum_{k=1}^{n} a_k = \sum_{j=1}^{n} b_j. \tag{1.X.11.7}$$

Hint. Distinguish two possible cases: (a) $a_n = b_n$ and (b) $a_n \neq b_n$.

(d) Deduce from (1.X.11.7) and previous rules that

$$\sum_{k=1}^{n} a_k = \begin{cases} \sum_{k=1}^{n} a_{n-k+1}, \text{ for } n \in \mathbb{N} & (\text{index reversal}), \\ \sum_{k=1}^{n/2} (a_{2k-1} + a_{2k}), \text{ for } n \in \mathbb{N} \text{ and } n \text{ even} & (\text{index doubling}). \end{cases} \tag{1.X.11.8}$$

Problem 1.X.12 (sum-product rules)**.** Let $(a_k)_{k\in\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}}$ be two infinite sequences.

(a) Prove, by double induction, the *general distributive law,*

$$\left(\sum_{k=1}^{m} a_k\right)\left(\sum_{k=1}^{n} b_k\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_i b_j\right), \text{ for all } m, n \in \mathbb{N}. \tag{1.X.12.1}$$

(b) Prove, by induction on n, that

$$\sum_{i=1}^{n}\left(\sum_{j=i}^{n} a_i b_j\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{j} b_j a_i\right), \tag{1.X.12.2}$$

$$\left(\sum_{k=1}^{n} a_k\right)^2 = \sum_{k=1}^{n} a_k^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_i a_j. \tag{1.X.12.3}$$

Problem 1.X.13 (Cassini's identity)**.** Suppose $(a_n)_{n\in\mathbb{N}_0}$ is the Fibonacci sequence, defined recursively by

$$a_0 := a_1 := 1, \text{ and } a_{n+1} := a_n + a_{n-1}, \text{ for } n \geq 1. \tag{1.X.13.1}$$

show by induction the following identity

$$a_{n+1}a_{n-1} - {a_n}^2 = (-1)^{n+1} \tag{$P(n)$}$$

for all $n \geq 1$.

Hint. To prove the inductive step $P(n) \Rightarrow P(n+1)$, start from the Inductive Hypothesis $P(n)$ written as follows

$$(-1)^{n+1} = a_{n+1}a_{n-1} - {a_n}^2, \tag{1.X.13.2}$$

and substitute a_{n+1} (using a_{n+2} and a_n) and a_{n-1} (using a_{n+1} and a_n).

Exercise 1.X.14 (product, ordering and simplification)**.** Recall the following properties of operations and ordering among real numbers: The following properties are equivalent

$$x < y, \tag{1.X.14.1}$$

$$\exists\, z \in \mathbb{R} : x + z < y + z, \tag{1.X.14.2}$$

$$\forall\, z \in \mathbb{R} : x + z < y + z. \tag{1.X.14.3}$$

Suppose $z \in \mathbb{R}^+$ where $\mathbb{R}^+ := \{x \in \mathbb{R} : x > 0\}$, then the following facts are equivalent

$$x < y, \tag{1.X.14.4}$$

$$\exists\, z \in \mathbb{R}_+ : xz < yz, \tag{1.X.14.5}$$

$$\forall\, z \in \mathbb{R}_+ : xz < yz. \tag{1.X.14.6}$$

The following are equivalent

$$x > 0, \tag{1.X.14.7}$$

$$\exists\, y, z \in \mathbb{R} : y < z \text{ and } xy < xz, \tag{1.X.14.8}$$

$$y < z \Rightarrow xy < xz. \tag{1.X.14.9}$$

To what extent can you replace "$<$" by "$\le$"? Rewrite the above with $\le$ instead of $<$ and check whether the statements are true.
Hint. Be careful with the 0 and multiplication.

Exercise 1.X.15 (product and ordering). (a) Deduce from the basic relations between ordering and algebra, that for $x < 0$ and $y < z$, one has $xy > xz$.

Hint. Argue by contraposition: suppose $y < z$ and $xy \le xz$ and show that it must then be $x \ge 0$.

(b) Based on the cancellation laws the algebraic properties and the fact $1 \neq 0$ for all prove that $1 > 0$.

Hint. Use a *contradiction argument.*

Exercise 1.X.16 (uniqueness of neutral and inverses). Let $(G, \star)$ be a group with a neutral element e. Show the following.

(a) G has a *unique* neutral element e with respect ot the operation $\star$.

Hint. Start by assuming there is another one, say f and showing that $f = e$.

(b) The inverse x' of each element $x \in G$, is unique.

Hint. Start by assuming there is another inverse, say x'' and showing that $x' = x''$.

CHAPTER 2

Basic Number Theory

> God may not play dice with the universe, but something strange is going on with the prime numbers.
> –Pál Erdős (attributed to him by Carl Pomerance)

In this chapter we shall be playing mainly with integers. The branch of mathematics that studies integers and their behaviour is called *Number Theory*. Number Theory is a very fine area (and many times a hard one) of mathematical research with applications as surprising as computer-data security. However, unlike most of advanced mathematics, many of Number Theory's results can be stated in very simple terms which the average person with a minimum of mathematical knowledge, say a high school student, can comprehend (which does not mean that one does not need sophisticated maths as one progresses in its study).

Here we shall look at very elementary results, which nevertheless, constitute a good ground for learning the basics of mathematical thinking and the techniques for proving a result.

At the end of this chapter we will be able to prove rigorously that $\sqrt{2} \notin \mathbb{Q}$ and that there are infinitely many prime numbers.

2.1. Preliminaries: divisibility and prime numbers

We begin with some preliminary definitions and basic facts.

2.1.1. Definition of segments of integers. Given $m, n \in \mathbb{Z}$ we define the *segment of integers* starting at m and ending at n, or starting at m and of length $m - n + 1$, to be the set

$$\{k \in \mathbb{Z} : \ m \leq k \text{ and } k \leq n\} =: [m \dots n], \tag{2.1.1}$$

where the latter defined new notation. The *length* of the segment $[m \dots n]$ is $m - n + 1$. Of particular use are the segments of positive numbers of the form $[1 \dots n]$ and $[0 \dots n]$.

2.1.2. Definition of divisor, factor, multiple, divides. Suppose $m \in \mathbb{Z}$ and $n \in \mathbb{Z}$. We say that m *divides* n, or that m is a *divisor (or factor)* of n, or that n is a *multiple* of n, and we write "$m \downharpoonleft n$" if and only if

$$\exists\, k \in \mathbb{Z} : n = km. \tag{2.1.1}$$

We say that m does not divide n if there is no k such that $n = km$; this case is denoted by $m \not\downharpoonleft n$.

2.1.3. Example. The following are true: 2 divides 16, 18 and -36; $3 \downharpoonleft 9$; $9 \not\downharpoonleft 3$, $13 \downharpoonleft -52$, $-15 \not\downharpoonleft -32$. Note that ± 1 divides any number $n \in \mathbb{Z}$, trivially, because $n = (\pm 1)(\pm n)$. Also any number $n \in \mathbb{Z}$ divides itself and $-n$. 0 on the other hand does not divide any $n \neq 0$. (Does 0 divide 0?)

2.1.4. Exercise.

(a) Suppose a number $n \in \mathbb{N}_0$*, show that* n *must have at least one divisor in* $\mathbb{N}_0$*.*

(b) Show that 1 *is the only element in* $\mathbb{N}_0$ *that has exactly 1 divisor in* $\mathbb{N}_0$*.*

(c) How many divisors does 0 *have in* $\mathbb{N}_0$*?*

2.1.5. Definition of prime. A natural number $p \in \mathbb{N}$ is called *prime* if it has exactly 2 divisors in $\mathbb{N}$.

[∗]: Check! An equivalent statement is[∗]

$$\big(p \geq 2\big) \text{ and } \forall\, d \in \mathbb{N} : \big(d \rfloor p \Rightarrow d = 1 \text{ or } d = p\big). \tag{2.1.1}$$

[∗]: Check! This can be rewritten in a logical statement as follows[∗]

$$\forall\, d \in \mathbb{N} : \big(\big(d \rfloor p \text{ and } d \neq 1\big) \Rightarrow d = p\big). \tag{2.1.2}$$

(It is a good idea to go through as many possible equivalent versions of a definition by stating equivalent propositions.)

2.1.6. Definition of composite. A natural number $n \in \mathbb{N}$ is called *composite* if it has 3 or more divisors.

2.1.7. Remarks.

(a) Since 1 has only one divisor (1 itself) in $\mathbb{N}$, then 1 *is not a prime number.*

(b) The sequence of consecutive prime numbers starts with 2, 3, 5, 7, 11, 13, 17, etc. There is no easy rule to tell whether a number is a prime or not.

(c) Prime numbers are still the object of fascinating mathematical research. A substantial part of Number Theory is devoted to understand prime numbers and related constructs. These topics play a central part in application such as *Cryptology*, the science that studies secret codes and the techniques to crack them and provides us with things, such as secure monetary transactions over the internet and hole-in-the-walls.

(d) An important basic result in number theory, due to Euclid, is that *prime numbers are infinite.* You are going to show that very soon in one of your homework exercises.

2.1.8. Prime notation. We denote the set of all prime numbers by

$$\mathbb{N}' := \{2, 3, 5, 7, 11, \ldots\}. \tag{2.1.1}$$

We also use the following notation for any $n \in \mathbb{N}$

$$[1 \ldots n]' := [1 \ldots n] \cap \mathbb{N}'. \tag{2.1.2}$$

2.1.9. Proposition (divisors are smaller). *Let* $m, n \in \mathbb{N}$ *such that* $m \rfloor n$*; then* $m \leq n$*. If* $m, n \in \mathbb{Z}$ *are such that* $m \rfloor n \neq 0$*, then* $|m| \leq |n|$*.*

Proof We use *contradiction* to prove this result. Since this is our first proof by contradiction, let us explain what we mean by that.

Suppose you want to prove that

$$A \Rightarrow B. \tag{2.1.1}$$

This is equivalent to the statement that either A is false (and B whatever) or that A is true and B is true. In other words, the only situation that is excluded by (2.1.1) is the one where

$$A \text{ is true and } B \text{ is false,} \tag{2.1.2}$$

which may be also written as

$$A \text{ and not} B. \tag{2.1.3}$$

A proof by contradiction is one that excludes the situation and it works as follows, suppose the statement in (2.1.3) is true and conclude something which is absurd.
In this case, the statement we would like to prove is, given $m, n \in \mathbb{N}$,

$$m \mid n \Rightarrow m \leq n. \tag{2.1.4}$$

So let us suppose $m \mid n$ (i.e., A is true) and $m > n$ (i.e., B is false); now we want to get an absurdity out of our assumption. Then we have that $mk = n$ for some $k \in \mathbb{N}$ (because $m \mid n$) and, on the other hand, because $m > n$, we have $mk > nk$ from the basic rules of multiplication and order in $\mathbb{N}$. Also, since $k \in \mathbb{N}$ we have $nk > n$. Thus $mk > nk$, but we know that $nk \geq n$ (because $k \geq 1$), and thus $mk > n$. But this is a contradiction with the fact that $mk = n$. This proves the result for $m, n > 0$.
If $m, n \in \mathbb{Z}$ and $m \mid n \neq 0$, then we have that $|m| \mid |n|$, so from the just proven result we have that $|m| \leq |n|$. □

2.1.10. Exercise. *Give a direct proof of Proposition 2.1.9, that is a proof that does not use contradition.*

2.1.11. Theorem (a sufficient condition for primeness). *Let $n \in \mathbb{N}$ and $n \geq 2$ and supppose that*

$$\big(d \in \mathbb{N} \text{ and } 1 < d^2 \leq n\big) \Rightarrow d \nmid n, \tag{2.1.1}$$

then n is prime.
Proof We now use a *proof by contraposition* (or *contrapositive proof*) which means that we exploit the fact that, given two statements A and B the statement

$$A \Rightarrow B \tag{2.1.2}$$

is equivalent to the statement

$$\text{not} B \Rightarrow \text{not} A. \tag{2.1.3}$$

The statement in (2.1.3) is called the *contrapositive* of the statement in (2.1.2).[1]

$$A :\Leftrightarrow \big(d \in \mathbb{N} \text{ and } 1 < d^2 < n\big) \Rightarrow d \nmid n \tag{2.1.4}$$

$$B :\Leftrightarrow n \text{ is prime} \tag{2.1.5}$$

and thus

$$\text{not} A \Leftrightarrow \exists\, d \in \mathbb{N} : 1 < d^2 < n \text{ and } d \mid n \tag{2.1.6}$$

$$\text{not} B \Leftrightarrow n \text{ is not prime.} \tag{2.1.7}$$

Suppose that (not B) holds true (i.e., B is false), our proof now consists in showing that (not A) holds true.
B false means that n is not prime, which means, by Definition of prime in §2.1.5, that there exist $m, k \in \mathbb{N}$ such that $m, k \geq 2$ and $n = mk$. Two cases are possible:
Case 1. If $m^2 \leq n$, let us take $d := m$: this makes (not A) true as required.[∗] [∗]: Check!

[1] Contraposition and contradiction, not only do they sound similar, but they look similar to the untrained eye, because both arguments start by supposing that B is false. Note however that in contradiction *we also assume that A is true* from the outset (and we try to obtain a logical conflict), whereas

Case 2. If $m^2 > n$, then $k^2m^2 > k^2n$, which implies that $n^2 > k^2n$. Using $n > 0$, we can simplify this inequality to $n > k^2$. Now taking $d := k$ makes the proposition (not A) true in this case too.

We have shown that from (not B) we get (not A). But this is the contrapositive (and thus equivalent to show) that $A \Rightarrow B$, which is what we wanted to show. □

2.2. A classical proof by contradiction that $\sqrt{2} \notin \mathbb{Q}$

2.2.1. Definition of square root. Given a real number $a \in \mathbb{R}$, such that $a \geq 0$, we say that $s \in \mathbb{R}$ is the *square root* of a if and only if

$$s \geq 0 \text{ and } s^2 = a. \tag{2.2.1}$$

For example 2 is the square root of 4, while -2 is not because, although $(-2)^2 = 4$, the statement $-2 \geq 0$ is false.[2]

2.2.2. Exercise. *Prove that, according to our definition of square root, there can be no more than one square root of a number $a \in \mathbb{R}$. We say that the square root of a non-negative real number is unique (if one exists).*
Hint. Consider $s, t \in \mathbb{R}$ such that, $s, t \geq 0$ and $s^2 = a = t^2$. You want to show that $s = t$. Start from $0 = a - a = t^2 - s^2$ and do not forget that $t, s \geq 0$ and the basic rules of algebra and ordering.

2.2.3. Properties of the square root. An important theorem of Analysis proves that *a square root of $a \in \mathbb{R}_{0+}$ always exists.* It can be shown (this is a bit easier) that for each a number in $\mathbb{R}_{0+}$ there is only one number s in $\mathbb{R}_{0+}$ such that $s^2 = a$ (if $a = 0$, only $s = 0$ satisfies $s^2 = 0$, and if $a \neq 0$ if there was another $r \in \mathbb{R}_{0+}$ such that $r^2 = a$, then $r^2 - s^2 = a - a = 0$, thus $(r - s)(r + s) = 0$, which, using cancellation implies $0 = r - s$, i.e., $r = s$). The previous two facts, i.e., uniqueness and existence of the square root of a positive real number, make the *square root* a well-defined operator (or function)

$$\begin{array}{rccl} \sqrt{\ }: & \mathbb{R}_{0+} & \to & \mathbb{R}_{0+} \\ & x & \mapsto & y \text{ such that } y^2 = x \end{array} \tag{2.2.1}$$

is well defined (i.e., for each input there is exactly one output). In particular for any $n \in \mathbb{N}$, its square root is well defined in $\mathbb{R}_{0+}$. An interesting question is whether $\sqrt{\ }$ is well defined as an operator from $\mathbb{N}$ into $\mathbb{N}$, and the answer is no, because $\sqrt{1} = 1$ and $\sqrt{4} = 2$ and since

$$\forall\, x, y \in \mathbb{R}_{0+} : x^2 \leq y^2 \Leftrightarrow x < y \tag{2.2.2}$$

[∗]: Check! (a fact that can be proved from the basic rules of multiplication and order in $\mathbb{R}$[∗]) then it must be $1 < \sqrt{2} < 4$. Hence $\sqrt{2} \notin \mathbb{N}$. So the next question is to ask whether $\sqrt{2} \in \mathbb{Q}$. It turns out that the answer to this question is also no. This fact, is telling: it means that the set of real numbers $\mathbb{R}$ is *really* (no pun intended) bigger than $\mathbb{Q}$. Let us prove it

in contraposition *we (try to) conclude that A is false.* A proof by contradition that does not make an active use of the fact that A is true is merely a goofy proof by contraposition. Goofy is OK, as long as it is true.

[2] Some other authors may *define* a square root without the condition that $s \geq 0$, which is perfectly legitimate as long as they stick to their definition throughout their work. In this course *square root* means the *non-negative square root.*

2.2.4. Theorem ($\sqrt{2}$ is irrational). *There is no number $q \in \mathbb{Q}$ such that*

$$q^2 = 2. \tag{2.2.1}$$

Proof By contradiction, suppose that there exists $q \in \mathbb{Q}$ such that $q^2 = 2$. Of course, $q \neq 0$, so it must be $q < 0$ or $q > 0$. Then there exist $n, d \in \mathbb{N}$ such that $q = n/d$ (if $q > 0$) or $-q = n/d$ (if $q < 0$) and

$$\boxed{n \text{ and } d \text{ have no common factors except for 1.}} \tag{2.2.2}$$

In either case, we have that

$$\frac{n^2}{d^2} = q^2 = 2. \tag{2.2.3}$$

Hence $n^2 = 2d^2$, so that $2 \rfloor n^2$. But

$$\boxed{2 \rfloor k^2 \Rightarrow 2 \rfloor k, \forall k \in \mathbb{Z}.} \tag{2.2.4}$$

Therefore conclude that $2 \rfloor n$, i.e., $n = 2m$ for some $m \in \mathbb{N}$. It follows that

$$2d^2 = (2m)^2 = 2 \times (2m^2) \tag{2.2.5}$$

and, by cancellation, that

$$d^2 = 2m^2. \tag{2.2.6}$$

But this means that $2 \rfloor d^2$ and using (2.2.4) again it follows that $2 \rfloor d$. It follows thus that d and n have a common factor 2 which is different than 1; but this is a contradiction with (2.2.2). □

Note that in the above proof we have used two facts in boxes, which are not obvious. To make this proof completely rigorous we ought to prove them. The boxed fact (2.2.4) is not too hard to prove, given a couple of hints.

2.2.5. Exercise.

(a) *Fill in the dots in the following statement:*
"A number n is odd if and only if there exists a number $k \in \dots$ such that $n = \dots$"

(b) *Show that if n is odd then n^2 is odd.*

(c) *Deduce that if $2 \rfloor n^2$ then $2 \rfloor n$.*

(d) *Can you state any important theorem that uses the last result?*

On the other hand, proving that *each rational number can be written as a fraction of two numbers which have no common factors except* 1 requires a technical tool about natural numbers known as the well-ordering principle. We present this topic in the next section and leave the proof of this fact as an exercise, to come back to once the well-ordering principle has been absorbed.

2.3. The well-ordering principle (WOP)

We mentioned a while ago, that an important property of $\mathbb{N}$ is that it has a smallest element. At first sight, this seems to be quite silly, but we can observe that there are (ordered) sets which do not have a smallest element. To be concise, let us focus on sets of numbers: $\mathbb{Z}$ does not have a smallest element, and neither does $\mathbb{Q}$ nor does $\mathbb{R}$. What is even more particular about $\mathbb{N}$ (or $\mathbb{N}_0$) is that anyone of its subsets, call it S, must have a smallest element, provided there is some element in S to start with. This property of $\mathbb{N}$ is known as the *Well-ordering Principle (WOP)* thereof. Some like to say

that $\mathbb{N}$ is well-ordered by $<$ (which is the relation that allows us to talk about "smallest"). In everydays language this means that each chunk of $\mathbb{N}$ has a beginning. Since this property of $\mathbb{N}$ comes handy many times in (a mathematician's) life, let us check it in more detail and derive it by using the induction principle.
We begin with some definitions.

2.3.1. Definition of Empty set. A set is called empty, and denoted $\varnothing$. If it contains no elements. The empty set is quite handy. It's like an empty shopping list, which is different than no shopping list at all.
For example, an empty shopping list tells you something: it tells you there are no items to be bought, you can go directly to the pub. On the other hand, if you lost the shopping list your partner/friend/mum gave you (i.e., you have no list) then you may still decide to go to the pub, but that guilt feeling and the prospect of a heated discussion will persist. An empty set (like an empty shopping list) will avoid you unnecessary arguments.

2.3.2. Definition of least element (minimum) of a set in $\mathbb{R}$. Let S be some fixed subset of $\mathbb{R}$. We say that S has a least element if and only if there exists m such that

$$m \in S \tag{2.3.1}$$

$$n \in S \Rightarrow n \geq m. \tag{2.3.2}$$

Such an element m is called a *minimum* of S and is denoted by $\min S$.

2.3.3. Remark (minimum $\neq$ lower bound). It is quite important to remember that the minimum of S must belong to S by definition. For example, taking the set S to be $\mathbb{Q}^+$ does not have a minimum: even though 0 satisfies (2.3.2), it does not satisfy (2.3.1). In this case 0 is called *lower bound*, but cannot be considered a minimum according to Definition 2.3.2. A minimum must be a lower bound, but a lower bound may not be a minimum

2.3.4. Remark (uniqueness of minimum). A subset of $\mathbb{R}$ S may fail to have a minimum. But if it has one then there is exactly that one minimum. Mathematicians will say that "the minimum is unique" (understating the "when it exists"). To see this, suppose m' is another minimum of S, we can easily see that $m' = m$, which means that m and m' are one and the same element of S. Indeed, we have $m \leq m'$ because of property (2.3.2) holds for m and property (2.3.1) holds for m'. Exchanging roles we have $m \geq m'$, and thus $m = m'$.

2.3.5. Theorem (Well Ordering Principle). *Suppose $S \subseteq \mathbb{N}$ (i.e., $(x \in S \Rightarrow x \in \mathbb{N})$) and $S \neq \varnothing$ (i.e., S has at least one element). Then S has a least element.*
In other words, if we introduce

$$\mathscr{N}^* := \{S \subseteq \mathbb{N} : S \neq \varnothing\}, \tag{2.3.1}$$

that is $\mathscr{N}^*$ is the set of all non-empty subsets of $\mathbb{N}$ the well ordering principle says that the rule

$$\begin{array}{rccl} \min : & \mathscr{N}^* & \rightarrow & \mathbb{N} \\ & S & \mapsto & \min S := \textit{smallest element of } S \end{array} \tag{2.3.2}$$

is a well-defined operation, which has exactly one output for each input.

Proof Let us argue by contraposition: suppose $S \subseteq \mathbb{N}$, such that no $m \in S$ satisfies both (2.3.1) and (2.3.2), we show that it must be $S = \emptyset$.
Because $S \subseteq \mathbb{N}$ then all we need to show is that

$$\forall\, n \in \mathbb{N} : n \notin S. \tag{2.3.3}$$

Let us do it by induction on $\mathbb{N}$.

Inductive step: (We use cumulative induction: i.e., if all dominoes up to m have fallen then m falls.) Suppose $1, 2, \dots, m-1 \notin S$. We want to show that $m \notin S$.

Let $n \in S$. Then $n \neq 1, 2, \dots, m-1$ and $n \in \mathbb{N}$. Hence, $n \geq m$. Thus m satisfies (2.3.2), so it cannot satisfy (2.3.1) by the assumption on S. Therefore $m \notin S$, as required.

Base case: We have to show that $1 \notin S$.

Note that 1 satisfies is the least element of $\mathbb{N}$ so it is smaller than any element in S, i.e., 1 satisfies (2.3.2). Thus, by the assumptions on S we cannot have $1 \in S$. I.e., $1 \notin S$.

□

2.3.6. Remark (What about $\mathbb{N}_0$?) It is not hard to see that $\mathbb{N}_0$ satisfies also the well ordering principle.
Let $S \subseteq \mathbb{N}_0$ and $S \neq \emptyset$. We have two cases:

Case 1. $0 \notin S$: in this case $S \subseteq \mathbb{N}$ and since $\mathbb{N}$ satisfies WOP then S must have a minimum.

Case 2. $0 \in S$: in this case 0 is the minium of S. Indeed, (2.3.1) is satisfied by assumption and for $n \in S$ we have $n \geq 0$ (either $n = 0$, or $n \in \mathbb{N}$ and $n > 0$).

2.4. Euclidean division

We've all been taught, quite early in our lives that in order to divide a (whole) number n by another (whole) number d we should count the biggest number q of times that d can fit in completely in n. We are then left with a rest $n - qd$ which is surely less than d. Basically, it is drilled in people's minds that *there is an operation*

$$\begin{array}{rccc} \operatorname{div}: & \mathbb{Z} \times \mathbb{N} & \to & \mathbb{Z} \times \mathbb{Z} \\ & (n, d) & \mapsto & (q, r) \end{array} \tag{2.4.1}$$

such that

$$n = qd + r \tag{2.4.2}$$

and

$$r = 0, \dots, n-1. \tag{2.4.3}$$

Intuitively, all this makes sense; but assuming we want to rely less on intuition and more on a logical process we may ask some questions:

(a) Is div well defined as an operation? I.e., are we certain that given any pair (n, d) that (a) a pair (q, r) satisfying the two properties exist, (b) that there is only one such pair. This is not plain silliness, because suppose you want to program a computer that does the job, one must be sure that these two requirements are satisfied.

(b) Provided we can answer positively the first question, is there a way to construct (q, r) out of the given pair (n, d)? This question seems less silly than the first one, but it turns out that they are two faces of the same coin.

It is WOP that gives a rigorous positive answer to question (a) and question (b) by the same token.

2.4.1. Theorem (Euclidean division). *For each given $n \in \mathbb{Z}$ and $d \in \mathbb{N}$ there exists $q \in \mathbb{Z}$ and $r \in \mathbb{Z}$ such that*

$$n = qd + r \tag{2.4.1}$$

and

$$0 \le r \le d - 1. \tag{2.4.2}$$

Furthermore when $n \neq 0$ such pair (q, r) is unique, i.e., if a (possibly different) pair $(q', r') \in \mathbb{Z} \times \mathbb{Z}$ satisfies (2.4.1) and (2.4.2) in lieu of (q, r) then $q = q'$ and $r = r'$ (i.e., the pairs cannot be different).

Proof Let $n \in \mathbb{Z}$ and $d \in \mathbb{N}$. We want to find $q, r \in \mathbb{Z}$ such that

$$0 \le r \le d - 1 \tag{2.4.3}$$

and

$$n = qd + r. \tag{2.4.4}$$

Let us try to find r first; this can be achieved by following the usual intuition for dividing whole numbers: keep on subtracting multiples of the divisor d from the dividand n and look at the remainder until this is less than d, but still positive.[3] Since we want r such that $n = qd + r$, and this is equivalent to $r = n - qd$, let us look at all the numbers $s \in \mathbb{N}_0$ of the form

$$s = n - kd, \text{ for some } k \in \mathbb{N}_0, \tag{2.4.5}$$

then "pick" the smallest such number and call it r. To make sure it is possible to "pick" the smallest s, we will use WOP. Namely, consider the set

$$S := \{s \in \mathbb{N}_0 : \ s = n - kd \text{ for some } k \in \mathbb{N}_0\}. \tag{2.4.6}$$

S is not empty because $n \in S$ (with $k = 0$), so by the WOP, $S \subseteq \mathbb{N}_0$ must have a least element. Let us call it r.

By definition of S and $r \in S$, there exists $q \in \mathbb{N}_0$ such that $r = n - qd$. Hence,

$$n = qd + r \tag{2.4.7}$$

which is (2.4.4).

We still have to show (2.4.3). By definition, $r \ge 0$ so all we need is to prove that $r < d$. Suppose, by contradiction, that $r \ge d$. It follows that $n - qd \ge d$, which implies that

$$0 \le n - qd - d = n - (q+1)d. \tag{2.4.8}$$

Thus, $r' = n - (q+1)d$ is of the form (2.4.5) and thus $r' \in S$. But we have also $r' = n - (q+1)d < n - qd = r$ which is in contrast with the fact that $r = \min S$.

If $n \in \mathbb{Z} \setminus \mathbb{N}_0$, then the same argument will work, but with some extra care. Define

$$S := \{s \in \mathbb{N}_0 : \ s = n - kd \text{ for some } k \in \mathbb{Z}\} \tag{2.4.9}$$

and note that for $k = n$ one has $n - kd = n(1 - d)$ and since $n < 0$ and $1 - d \le 0$ we obtain that $n - kd \ge 0$ and thus $n - kd \in S$. The argument then works the same as before. □

[3] The intuition works well when n and d are both positive; but the proof turns out to be correct for when one of them is negative.

2.4.2. Definition of quotient, remainder and modulo. Given a pair $(n,d) \in \mathbb{Z} \setminus \{0\} \times \mathbb{N}$ with $n \neq 0$. The pair (q,r), uniquely defined by Theorem 2.4.1, is called the *result of the (Euclidean) division of n by d*, q is the (whole) *quotient* of $\operatorname{div}_d n$ and r is the *rest*, or the *remainder* of $\operatorname{div}_d n$.
The remainder r of $\operatorname{div}_d n$ goes also by the name of n *modulo* d, and we will denote it by $r = \operatorname{mod}_d n$ or $r = n \,(\operatorname{mod} d)$.

2.5. The Euclidean algorithm

2.5.1. Definition of largest or greatest element, or maximum, of a set. An analogue of minimum is the so-called maximum of a set. Given a set $S \subseteq \mathbb{R}$, we say that S has a largest (greatest, highest) element or simply a maximum if there exist $m \in \mathbb{R}$ such that

$$m \in S \tag{2.5.1}$$

$$x \in S \Rightarrow x \leq m. \tag{2.5.2}$$

Such a number, if it exists, is unique in which case it is called *the maximum* (also known as *greatest element* or *largest element*) of S and denoted by $\max S$.

2.5.2. Exercise (uniqueness of maximum). *Let S be a subset of $\mathbb{R}$. Check that (provided it exists)* $\max S$ *is unique.*
This justifies the notation $m = \max S$, whenever m is the maximum of S.

2.5.3. Existence of maximum in $\mathbb{N}$, $\mathbb{N}_0$ and $\mathbb{Z}$. Unlike for the minimum, it is not true that the maximum of a non-empty subset $S \subseteq \mathbb{N}$ must exists. For example the set of even numbers in $\mathbb{N}$ does not have a maximum. However *if the set $S \in \mathbb{N}$, in addition to not being empty, has an upper bound, i.e., there is a number $r_0 \in \mathbb{N}$ such that $x \leq r_0$ for all $x \in S$, then* $\max S$ *exists.*
To see this, one can apply the WOP in $\mathbb{N}_0$ to the set

$$R := \{r_0 - x : x \in S\}. \tag{2.5.1}$$

(Note that $r \in R$ if and only if $r_0 - r \in S$.) Since S is not empty it follows that R is not empty either. Thus $\min R$ exists, by WOP in $\mathbb{N}_0$. It follows that $M = r_0 - \min R$ satisfies the properties of $\max S$. Indeed, since $\min R \in R$, it follows as noted after (2.5.1) that $M = r_0 - \min R$ belongs to S. Furthermore if $x \in S$ then $r_0 - x \in R$ and $r_0 - x \geq \min R$. Thus $\max S = r_0 - \min R \geq r_0 - (r_0 - x) = x$.

2.5.4. Remark (maximum's uniqueness). Although a maximum may not exist for a given set S, there can be no more than one maximum for S. As an exercise, show this claim. This justifies our use of the definite article *the*, when talking about $\max S$.[∗]

[∗]: Check!

2.5.5. Definition of above-bounded set, upper bound. A set $S \subseteq \mathbb{R}$ is called *bounded above* if there exists a number $K \in \mathbb{R}$ such that

$$x \in S \Rightarrow x \leq K. \tag{2.5.1}$$

Any number K that satisfies (2.5.1) is called an *upper bound* of S.

2.5.6. Remark (about the upper bound). Note that there may be more than one upper bound on a set S. Indeed, if there is one upper bound, there are infinitely many others.[∗] [∗]: Check!
Note also that if K is an upper bound of S, then K need not be an element of S.

2.5.7. Proposition (an above-bounded set of integers has a maximum). *Let S be a given a subset of$\mathbb{N}_0$ ($S \subseteq \mathbb{N}_0$). If S is bounded above then S has a maximum element.*
Proof We can use WOP. Consider the set R given by

$$R = \{n \in \mathbb{N}_0 : \; n \geq m, \; \forall \, m \in S\}. \tag{2.5.1}$$

Surely, because S is finite, there is an element in R, so that R is not empty. Thanks to WOP, we conclude that there is a smaller element in R. Define $M := \min R$. We will show that $M := \max S$. First, because $M \in R$, by (2.3.1), then $M \geq m$ for all $m \in S$ and thus M satisfies (2.5.2).
We still have to show (2.5.1), i.e., that $M \in S$. Suppose, by contradiction that $M \notin S$, and consider $M' = M - 1$.
We claim that $M' \in R$ (contradicting thus the fact that M is the smallest element of R). To show $M' \in R$ all we need to do is show that $m \leq M'$ for all $m \in S$. So suppose $m \in S$, then $m \leq M$, but equality is not allowed because we supposed that $M \notin S$, so $m < M$. Thus $m \leq M - 1 = M'$, which confirms our claim because m is arbitrary in S.
□

2.5.8. Example (a bounded set without a maximum). Note that $S \subseteq \mathbb{N}_0$ is crucial in Proposition 2.5.7. That is, if S is not a subset of $\mathbb{N}_0$ then Proposition 2.5.7 no longer applies. As an example take the set

$$S := \left\{ \frac{n-1}{n} : \; n \in \mathbb{N} \right\}. \tag{2.5.1}$$

Then S is bounded above (ensure you believe this by finding an upper bound), yet it has no maximum. Indeed, if $x \in S$ (one of the two properties for x to be a maximum of S) then x is equal to $(n-1)/n$ for some $n \in \mathbb{N}$. But taking the number $y := n/(n+1)$ we see that $y \in S$ and $y > x$, which means that x does not satisfie the other property for being a maximum of S.

2.5.9. Definition of highest common factor or greatest common divisor. Given two integers $m, n \in \mathbb{Z}$, with at least one of them not null (say $n \neq 0$), their *highest common factor* (also known as *greatest common divisor*) is defined as:

$$\max \underbrace{\{d \in \mathbb{N} : \; d \rfloor n \text{ and } d \rfloor m\}}_{=: D(m,n)} \tag{2.5.1}$$

Note that since the set $D(m,n)$ is bounded above by $|n|$ (since from 2.1.9 a divisor of $n \neq 0$ cannot exceed $|n|$), $\max D(m,n)$ must exist by Proposition 2.5.7. A notation for the highest common factor of two numbers m, n is

$$\operatorname{hcf}\{m, n\}. \tag{2.5.2}$$

Since the order in which m and n are taken does not influence the outcome, we may safely use curly braces for the arguments of hcf. In fact, the definition of highest common factor can be made for general subsets of $\mathbb{Z}$.

2.5.10. Exercise (generalised highest common factor).

(a) *Can you generalise the definition of highest common factor to that of three integers l, m, n, of which one is not zero? How about any set of integers, A? Make sure you generalise the fact that one of the elements of A must be non-zero.*

(b) *What happens if $A = \{n\}$ for a given $n \in \mathbb{Z} \smallsetminus \{0\}$, i.e., A is a set with exactly one element (also known as a singleton)?*

(c) *Suppose A_1, A_2 are two subsets of $\mathbb{Z}$ such that $A_i \smallsetminus \{0\} \neq$, for both $i = 1, 2$, show that*

$$\operatorname{hcf} A_1 \cup A_2 = \operatorname{hcf}\{\operatorname{hcf} A_1, \operatorname{hcf} A_2\}. \tag{2.5.1}$$

(d) *Let $A \subseteq \mathbb{Z}$ and define*

$$B := \{|n| : \ n \in A\}. \tag{2.5.2}$$

Show that

$$\operatorname{hcf} A = \operatorname{hcf} B. \tag{2.5.3}$$

This reduces the study of highest common factor of subsets of $\mathbb{Z}$ to only the subsets of $\mathbb{N}$.

Solution.

(a) If A is a subset of $\mathbb{Z}$ such that $A \smallsetminus \{0\}$ is not empty, then we may define

$$\operatorname{hcf} A := \max\{d \in \mathbb{N} : \ d \rfloor n \text{ for each } n \in A\}. \tag{2.5.4}$$

Denote by $D(A)$ the set on the

2.5.11. Example (how to calculate the highest common factor (without wasting too much space)). Find the highest common factor between 492 and 294.

$$\begin{aligned} \operatorname{mod}_{294} 492 &= 198 \\ \operatorname{mod}_{198} 294 &= 96 \\ \operatorname{mod}_{198} 96 &= 6 \\ \operatorname{mod}_{6} 96 &= 0. \end{aligned} \tag{2.5.1}$$

Since 6 is the last remainder that is not zero in this sequence of Euclidean divisions, we declare it to the be the hcf. This method is known as the Euclidean Algorithm. An important mathematical question about this algorithm is "Why does it work"?[4] By "work" we mean that if we replaced 492 and 294 by two general numbers, say m and n,

(1) Does the algorithm ever "hit" a 0 remainder?

(2) Does it yield the desired result? (I.e., can we insure that the remainder just before the 0 the hcf?)

2.5.12. The Euclidean Algorithm to find the HCF. Taking inspiration from Example 2.5.11, we want to write down an algorithm to calculate the hcf. This algorithm is known as *Euclid's algorithm* or the *Euclidean algorithm.*

Given $(m, n) \in \mathbb{N} \times \mathbb{N}_0$, we may define a recursive sequence of pairs r_k, for $k \geq -1$, as follows

$$r_{-1} := m \text{ and } r_0 := n \tag{2.5.1}$$

[4] If you were already asking this question, then you are already thinking mathematically.

and for $k \geq 1$

$$(q_{k+1}, r_{k+1}) := \begin{cases} \operatorname{div}_{r_k} r_{k-1}, & \text{if } r_k \neq 0, \\ (q_k, r_k), & \text{if } r_k = 0. \end{cases} \tag{2.5.2}$$

Although the sequence of quotients q_k, $k = 1, 2, \ldots$, thus produced are not needed to actually compute $\operatorname{hcf}\{m, n\}$, they are useful for many other things, such as the theoretical analysis of the Euclidean algorithm and their role in Bézout's identity. In practice, the k-th step of this recursive definition can be also written using the "mod" notation as follows

$$r_{k+1} := \begin{cases} \operatorname{mod}_{r_k} r_{k-1}, & \text{if } r_k \neq 0, \\ 0, & \text{if } r_k = 0. \end{cases} \tag{2.5.3}$$

In practice, the algorithm is stopped as soon as $r_k = 0$ for some $k \in \mathbb{N}_0$. A computer implementation (for those in the know) can be written as follows:

Algorithm (Euclidean)**. Require:** $m \in \mathbb{Z}$, $n \in \mathbb{Z} \setminus \{0\}$
Ensure: $\operatorname{hcf}\{m, n\}$

```
procedure EUCLID($m, n$)                      ▷ computes the hcf of $m$ and $n$
    $r_{-1} \leftarrow m$, $r_0 \leftarrow n$, $k \leftarrow 1$
    while $r_k \neq 0$ do                     ▷ answer is reached when $r_k = 0$
        $r_k \leftarrow \operatorname{mod}_{r_{k-2}} r_{k-1}$
        $k \leftarrow k + 1$
    end while
    return $r_{k-1}$                          ▷ hcf is the last non-zero remainder
end procedure
```

Let us have a closer look at what the algorithm is doing. Suppose that for some given $N \in \mathbb{N}_0$, we have that the remainder r_{N+1} is 0, while all the previous ones, $r_{-1}, \ldots, r_N$, are all non-zero, (we will soon show that such an N must exist), then we have

$$r_{N-1} = q_{N+1} r_N + r_{N+1} = q_{N+1} r_N \tag{2.5.4}$$

$$r_{N-2} = q_N r_{N-1} + r_N \tag{2.5.5}$$

$$\vdots$$

$$r_1 = q_3 r_2 + r_3 \tag{2.5.6}$$

$$r_0 = q_2 r_1 + r_2 \tag{2.5.7}$$

$$r_{-1} = q_1 r_0 + r_1. \tag{2.5.8}$$

The first equality says basically that $r_{N-1} = s_1 r_N$, where s_1 is just another notation for q_{N+1}. Using this first equality to substitute r_{N-1} in (2.5.5), we see that:

$$r_{N-2} = q_N s_1 r_N + r_N = \underbrace{\big(q_N s_1 + 1\big)}_{=: s_2} r_N. \tag{2.5.9}$$

It then follows that

$$r_{N-3} = \underbrace{\big(q_{N-1} s_2 + s_1\big)}_{=: s_3} r_N \tag{2.5.10}$$

Working our way back recursively (in an informal yet intuitive manner) down to r_{-1} we obtain

$$n = r_0 = s_N r_N \text{ and } m = r_{-1} = s_{N+1} r_N \tag{2.5.11}$$

for appropriate choice of s_{N-1} and s_N in $\mathbb{N}$. We infer thus the following

Claim *Let $r_{-1}, r_0, \dots, r_N$, be the sequence of all non-zero numbers given by (2.5.1) and (2.5.2), we have*

$$\forall\, j = 0, \dots, N+1 : \exists\, s_j \in \mathbb{N} : r_{N-j} = s_j r_N. \tag{2.5.12}$$

This claim can be (rigorously) proved by induction on j and is left as an exercise.[$*$] [$*$]: Check! Thus r_N divides both m and n, so the result of Algorithm 2.5.12 is indeed a common factor of the input m, n. We still have to prove that

(i) the process does terminate (i.e., $r_{N+1} = 0$ will actually happen for some N), and

(ii) r_N is not only a common factor of m and n but the *largest* one.

2.5.13. Termination of the Euclidean algorithm. To understand why i must be true, it is sufficient to observe that $r_{-1} > r_0$, and thus $r_0 > r_1$ (unless $r_0 = 0$), and thus $r_1 > r_2$ (unless $r_1 = 0$), etc. This means that the sequence of postive rests r_k is a strictly decreasing sequence of natural numbers and must "bottom up" and cannot be infinite and there must be an N such that $r_N > 0$ but $r_{N+1} = 0$. We do this, a bit more rigorously, next.

Proposition (termination of the Euclidean algorithm)**.** *Given $m, n \in \mathbb{N}$. Consider the sequence $(r_k)_{k\geq -1}$ as defined by the recursion (2.5.1)– (2.5.2), then there exists an index $N \in \mathbb{N}_0$ such that $r_N > 0$ and $r_{N+k} = 0$ for all $k \geq 1$. (That is, a bit more sloppily but also more suggestively, for some $N \in \mathbb{N}_0$, the sequence $r_{-1}, r_0, r_1, \dots, r_k, \dots$ satisfies*

$$r_{-1} > r_0 > \dots > r_{N-1} > r_N > r_{N+1} = r_{N+2} = \dots = 0.) \tag{2.5.1}$$

Proof The main point here is to observe that for each $k \in \mathbb{N}_0$ we either have

$$r_{k-1} \neq 0, \text{ and} \tag{2.5.2}$$

$$r_{k-1} > r_k \tag{2.5.3}$$

or

$$r_{k-1} = 0 \text{ and } r_k = 0. \tag{2.5.4}$$

This is a consequence of (2.5.2), which defines r_k for $k \geq 1$. If $r_{k-1} = 0$ then $r_k = 0$ by definition. If $r_{k-1} > 0$, then $r_k = \operatorname{mod}_{r_{k-1}} r_{k-2}$, which, in view of the Euclidean Division Theorem 2.4.1, implies that $0 \leq r_k < r_{k-1}$.

By an induction argument, we deduce that for all $i, j \in \mathbb{N}_0$ we have the following monotonicity property

$$i < j \Rightarrow r_j < r_i \text{ or } r_j = r_i = 0, \tag{2.5.5}$$

which is equivalent, in the contrapositive, to say that

$$\left(r_i > 0 \text{ and } r_i \leq r_j\right) \Rightarrow i \geq j. \tag{2.5.6}$$

(The base case $j = 1$ has been proved, the inductive step $j \to j+1$ is left to the reader.) Note that inequality (2.5.3) does not exclude $r_k = 0$. In fact, to finish the proof it is enough to show that it does happen for some k which we will call $N+1$, so that $r_N > 0$ and $r_{N+1} = 0$.

To "find" such an N, consider the set $S = \{r_k : k \geq 0 \text{ and } r_k > 0\}$. By definition, $S \subseteq \mathbb{N}$ and it is not empty (because $r_0 \in S$, since $r_0 = n \neq 0$), so by WOP we have that $\min S$ exists, which is the same as saying that there is there exists $N \in \mathbb{N}_0$ such that

$$r_N = \min S. \tag{2.5.7}$$

It follows, from (2.5.5) that for if for some $k \in \mathbb{N}$ we have $r_k > 0$, then $r_k \in S$, so $r_k \geq r_N$, thus $k \leq N$. Hence if $k \geq N+1$, then $r_k < r_N$, which implies that $r_k \notin S$, and therefore $r_k = 0$. □

2.5.14. Remark. The above proposition and its proof tell us that the sequence output by Algorithm 2.5.12 is strictly decreasing and that the Algorithm in fact terminates (because the while loop is broken as soon as r_{k+1} becomes 0.

2.5.15. Remark (For hackers only). In practice, if we are interested only in hcf$\{m,n\}$ we do not need to store all the sequence r_k, all we need is to keep track of two code variables, for example, `rold`, `r`, `rnew` replacing r_{k-1}, r_k, r_{k-1} in the loop. The q_k's are not really needed, if all we want is the hcf.
The following are 2 Octave files, which will probably run on Matlab® too.

Printout of file `IPM/Code/divEuclid.m`

```
%% -*- mode: octave -*-
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [q,r] = divEuclid(m,d)
%% function [q,r] = divEuclid(m,d)
%%
%% returns quotient q and rest r of Euclidean division of integers m by d
%% for positive m and strictly positive d
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
  r = m;
  q = 0;
  while(r>=d)
    r = r-d;
    q = q+1;
  end
```

You should test this file, for instance, with the line

    divEuclid(32,5)

Printout of file `IPM/Code/hcfEuclid.m`

```
%% -*- mode: octave; -*-
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function rold = hcfEuclid(m,n)
  %% function rold = hcfEuclid(m,n)
  %%
  %% returns the highest common factor (greatest common divisor) of m
  %$ and n
  rold = m;
  rnew = n;
  sold = 1;
  snew = 0;
  told = 0;
  tnew = 1;
  while(rnew != 0)
    r = rnew;
    s = snew;
    t = tnew;
    [q,rnew] = divEuclid(rold,r);
    s = s*q-sold;
    t = t*q-told;
    rold = r;
  end
  if(s*m+t*n==r)
    print("gcd's all righty!\n");
end
```

A good test line, for example, would be

```
hcfEuclid(492,294)
```

2.5.16. Remark. Thanks to Proposition 2.5.13, we now know that the Euclidean Algorithm 2.5.12 terminates.
We also have noted that the output of this algorithm is a common divisor of the input. We still have to show that it is the *highest* common divisor.

2.5.17. Theorem (Euclidean Algorithm returns the hcf). *Let $m,n \in \mathbb{N}$ and let $N \in \mathbb{N}_0$ be the index such that $r_N = 0$ and $r_{N-1} > 0$ returned by Algorithm 2.5.12 then $r_N =$* $\operatorname{hcf}\{m,n\}$. The proof of this Theorem relies on the following basic Lemma.

2.5.18. Lemma (hcf simplification via mod). *Let $a \in \mathbb{N}$ and $b \in \mathbb{N}_0$, then*

$$\operatorname{hcf}\{a,b\} = \begin{cases} a, & \text{if } b = 0, \\ \operatorname{hcf}\{b, \operatorname{mod}_b a\}, & \text{if } b > 0. \end{cases} \tag{2.5.1}$$

Proof Let $a \in \mathbb{N}$ and $b \in \mathbb{N}_0$. If $b = 0$, then all natural numbers divide b. This includes a. So a is a common factor of a and $b = 0$. Since $a > 0$, a has no factor bigger than it. Hence $a = \operatorname{hcf}\{a,0\}$.
If $b \neq 0$, then $b > 0$. Define $c := \operatorname{mod}_b a$, so that $a = qb + c$ for some $q \in \mathbb{N}_0$. Let $M := \{k \in \mathbb{N} : k \rfloor a \text{ and } k \rfloor b\}$ and $L := \{k \in \mathbb{N} : k \rfloor c \text{ and } k \rfloor b\}$, so that $\operatorname{hcf}\{a,b\} = \max M$ and $\operatorname{hcf}\{b,c\} = \max L$. To show that the two maximums coincide, it is sufficient to check that the two sets M and L are equal.[5]
Suppose $k \in M$, then $k \rfloor a$ and $k \rfloor b$, i.e., $a = ik$ and $b = jk$ for some $i, j \in \mathbb{N}$, hence $c = a - qb = ik - qjk = (i - qj)k$ and $k \rfloor c$. Since $k \rfloor b$ also, then $k \in L$. Conversely, if $k \in L$, then $k \rfloor c$ and $k \rfloor b$, i.e., $c = lk$ and $b = jk$ for soem $l, j \in \mathbb{N}$. Hence $a = qb + c = (qj + l)k$, which means that $k \rfloor a$ and $k \in M$.
So the two sets M and L coincide, and so must their greatest elements. Hence $\operatorname{hcf}\{a,b\} = \operatorname{hcf}\{b,c\}$, as desired. □

2.5.19. Example (Lemma 2.5.18 is crucial). To appreciate Lemma 2.5.18, suppose we want to calculate the hcf of 492 and 294, then the Lemma says that

$$\operatorname{hcf}\{492,294\} = \operatorname{hcf}\{294, \operatorname{mod}_{294} 492\} = \operatorname{hcf}\{294,198\}. \tag{2.5.1}$$

Applying the Lemma again we have

$$\operatorname{hcf}\{294,198\} = \operatorname{hcf}\{198, \operatorname{mod}_{198} 294\} = \operatorname{hcf}\{198,96\}. \tag{2.5.2}$$

And again

$$\operatorname{hcf}\{198,96\} = \operatorname{hcf}\{96, \operatorname{mod}_{96} 198\} = \operatorname{hcf}\{96,6\}. \tag{2.5.3}$$

And again

$$\operatorname{hcf}\{96,6\} = \operatorname{hcf}\{6, \operatorname{mod}_{6} 96\} = \operatorname{hcf}\{6,0\}. \tag{2.5.4}$$

[5] We say that $A = B$ if the sets A and B have exactly the same elements, i.e., $x \in A \Leftrightarrow x \in B$. Recalling that $x \in A \Rightarrow x \in B$ is $A \subseteq B$ by definition, it is sufficient to check that $M \subseteq L$ and $L \subseteq M$.

And, one final time (this time using the $b = 0$ case)

$$\mathrm{hcf}\{6,0\} = 6. \tag{2.5.5}$$

Joining all equalities (2.5.1) through (2.5.5), it follows that

$$\mathrm{hcf}\{492,294\} = 6. \tag{2.5.6}$$

So in fact, this Lemma explains why the Euclidean Algorithm yields the highest common factor.
Proof of Theorem 2.5.17 Noting that $r_k = \mathrm{mod}_{r_{k-1}} r_{k-2}$ for $1 \leq k \leq N$, and using Lemma 2.5.18, it follows that

$$\mathrm{hcf}\{r_{-1}, r_0\} = \mathrm{hcf}\{r_0, r_1\} = \ldots = \mathrm{hcf}\{r_{N-1}, r_N\}. \tag{2.5.7}$$

But $r_N = 0$ (and $r_{N-1} > 0$), hence the last term equals r_{N-1}. Thus

$$\mathrm{hcf}\{m,n\} = r_{N-1}. \tag{2.5.8}$$

□

2.6. Bézout's identity

Theorem 2.5.17 has a very useful consequence, which can be stated as follows.

2.6.1. Theorem (Bézout's identity). *For each pair $(m,n) \in \mathbb{N}^2$ there exists a pair $(s,t) \in \mathbb{Z}^2$ such that*

$$\mathrm{hcf}\{m,n\} = sm + tn. \tag{2.6.1}$$

Equivalently, $\mathrm{hcf}\{m,n\}$ *is a $\mathbb{Z}$-linear combination of m, n.*

2.6.2. Remark (linear combinations and coefficients).
(a) Since $\mathrm{hcf}\{m,n\}$ is positive, then one of s, t must be negative and the other positive; therefore *we have to take $\mathbb{Z}$ as the set of possible coefficients.*

(b) Let us bust some jargon before we proceed. The expression $\mathbb{Z}$-*linear combination* comes from Linear Algebra, where linear combinations are used very often. Do not worry too much about where it comes from now. All we need to know is that *linear combination of m and n* is an English expression for a mathematical expression like "$sm + tn$".

The only operations allowed in a linear combination are + and multiplication with the coefficients s and t. For example: $2n + 3m$ is a linear combination of n and m. Also $4m - 3n$ and $am + bn + cm$ are linear combinations of m and n.

The s and t appearing in the linear combination $sm + tn$ are called the *coefficients* (or *scalars*) of the linear combination. To stress the fact that the coefficients s, t in (2.6.1) must be in $\mathbb{Z}$ we (use "$\mathbb{Z}$-" as a prefix and) say that $sm + tn$ is a $\mathbb{Z}$-*linear combination* of m and n.

2.6.3. Example (working out s, t in (2.6.1)). To understand Theorem 2.6.1, let us go through an example. From Example 2.5.11 we know that $\text{hcf}\{492, 294\} = 6$. The Theorem tells us that there are numbers $s, t \in \mathbb{Z}$ such that $6 = 492s + 294t$. Unfortunately, this Theorem says nothing about *how* to find s and t.
To find s and t, we could exploit the divisions that we made in order to find 6 in 2.5.11. Namely, we know that

$$492 = 1 \times 294 + 198, 294 \quad = 1 \times 198 + 96, 198 = 2 \times 96 + 6, \tag{2.6.1}$$

Working via a "back-substitution" we get

$$\begin{aligned} 6 &= -2 \times 96 + 198 \\ &= -2(-198 + 294) + 198 \\ &= 3 \times 198 - 2 \times 294 \\ &= 3(492 - 294) - 2 \times 294 \\ &= 3 \times 492 - 5 \times 294. \end{aligned} \tag{2.6.2}$$

(It is a good idea to check that this is indeed true.)
Proof of Theorem 2.6.1 From Theorem 2.5.17, we know that $r_{N-1} = \text{hcf}\{m, n\}$. We will show that r_{N-1} is a $\mathbb{Z}$-linear combination of m and n. In fact, we show that for all $j \in [-1 \dots N-1]$ we have that $r_j = s_j m + t_j n$ for some s_j and t_j in $\mathbb{Z}$, which can be also written as

$$\forall\, j \in \mathbb{N}_0 : \exists\, s_j, t_j \in \mathbb{Z} : r_j = s_j m + t_j n. \tag{2.6.3}$$

For the base case of (2.6.3), i.e., for $j = -1, 0$, it is enough to take $s_{-1} = 1$, $t_{-1} = 0$ and $s_0 = 0$, $t_0 = 1$.
To prove the inductive step, fix $j < N - 1$, it follows from the the Euclidean Algorithm (2.5.2) that

$$r_{j-1} = q_j r_j + r_{j+1} \qquad (\text{since } j < N-1) \tag{2.6.4}$$

and thus a simple algebraic manipulation gives

$$\begin{aligned} r_{j+1} &= r_{j-1} - q_j r_j \\ \big(\text{inductive hypothesis}\big) \quad &= (s_{j-1} m + t_{j-1} n) - q_j (s_j m + t_j n) \\ \big(\text{algebraic manipulation}\big) \quad &= (s_{j-1} - q_j s_j) m + (t_{j-1} - q_j t_j) n. \end{aligned} \tag{2.6.5}$$

Defining

$$s_{j+1} := s_{j-1} - q_j s_j \text{ and } t_{j+1} := t_{j-1} - q_j t_j \tag{2.6.6}$$

we obtain that

$$r_{j+1} = s_{j+1} m + t_{j+1} n. \tag{2.6.7}$$

□

2.6.4. Problem (extended Euclidean algorithm).

(a) *Based on the Euclidean algorithm and Bézout's identity, write a pseudocode for (or explain in plain English) an algorithm that, for each given $m, n \in \mathbb{N}$, computes their highest common factor,* $\text{hcf}\{m, n\}$*, and two integers $s, t \in \mathbb{Z}$ such that*

$$\text{hcf}\{m, n\} = sm + tn. \tag{2.6.1}$$

(b) Show that this algorithm is certifying, i.e., that you can easily use its output to check that the result is correct.

Hint. You will find it useful to recall that $\operatorname{hcf}\{m,n\}$ *is the only number that both divides both, and is a* $\mathbb{Z}$*-linear combination, of* m *and* n*.*

(c) Implement the algorithm with a computer language of your choice.

2.6.5. Problem (divisibility and Bézout imply HCF). *Let* $m,n\in\mathbb{Z}$ *and* $m\neq 0$*, show that if* $g\in\mathbb{N}$ *safisfies*

$$g \rfloor m, g \rfloor n, \text{ and} \tag{2.6.1}$$

$$g = sm + tn \text{ for some } s,t\in\mathbb{Z}, \tag{2.6.2}$$

then

$$g = \operatorname{hcf}\{m,n\}. \tag{2.6.3}$$

2.7. Coprimality

We close this chapter, by discussing some important properties of prime and coprime numbers.

2.7.1. Definition of coprime pair. We say that a pair of integers m and n is *coprime*, if and only if $\operatorname{hcf}\{m,n\}=1$. Sometimes we say that *m and n are coprime.* More generally, we say that a set of integers A is *coprime* if and only if $\operatorname{hcf} A = 1$. A set A is *pairwise coprime* if each pair $\{m,n\}\subseteq A$, with $m\neq n$, is coprime. Some authors employ *relatively prime* or *mutually prime* for coprime and the notation

$$m \perp n. \tag{2.7.1}$$

2.7.2. Example (coprime pair). The following pairs are coprime:

$$\{15,8\},\{19,24\},\{14,15\},\{10,99\}. \tag{2.7.1}$$

The set of all prime numbers is coprime.

2.7.3. Remark. If p is a prime number then for any $k\in\mathbb{N}$, $k<p$ we have that (k,p) is coprime. Indeed, $p \not\rfloor k$ and $k \not\rfloor p$, unless $k=1$. Therefore if $d\geq 2$ is a divisor of k and of p, then $d\leq k<p$ and thus $d=1$.

2.7.4. Theorem (coprime factorisation). *If* n *and* a *are coprime and* $n \rfloor ab$ *then* $n \rfloor b$*.*

Proof For n and a are coprime, we have $\operatorname{hcf}\{n,a\}=1$. Thus, by Theorem 2.6.1, $1 = sn+ta$, for some suitable $s,t\in\mathbb{Z}$.

Then multiplying by b both sides we get that $b = snb + tab$. But $n \rfloor ab$, so $nk = ab$ for some $k\in\mathbb{N}$, hence

$$b = sbn + tkn = (sb+tk)n, \tag{2.7.1}$$

which means that $n \rfloor b$, as we wanted to show. □

2.7.5. Example. As an application of 2.7.4, consider 26 and 35. It is not hard to check that

$$35 = 1\times 26 + 9,\ 26 = 2\times 9 + 8,\ 9 = 1\times 8 + 1\,(,\ 8 = 8\times 1) \tag{2.7.1}$$

so $\operatorname{hcf}\{35, 26\} = 1$. So 26 and 35 are coprime. Now, as an example consider

$$26 \mid 1820 = 35\times 52 \tag{2.7.2}$$

then Theorem 2.7.4 says that 26 must divide 52 (which, of course, could be inferred directly without using the Theorem).

2.7.6. Atomic property of primes. A very useful consequence of 2.7.4 is the so-called *atomic property of prime numbers,* which says that a prime divisor of a product must divide at least one of the two factors. Let us state and prove this property.

Theorem (atomic property of primes)**.** *If n is a prime number and $n \mid ab$ then $n \mid a$ or $n \mid b$.*

Proof Suppose n is prime and $n \mid ab$. If $n \mid a$ then there is nothing to prove, so suppose also $n \nmid a$. We want to show that $n \mid b$. For this note that since $n \nmid a$, then $\operatorname{hcf}\{a, n\} < n$. But since the only other divisor of n is 1, then $\operatorname{hcf}\{a, n\} = 1$, i.e., (a, n) is a coprime pair. Therefore, by Theorem 2.7.4, we conclude that $n \mid b$. □

2.7.7. Exercise (atomic property of primes with arbitrary factors). *Prove the following statement. Suppose n is a prime number and $n \mid a_1 \cdots a_k$, where $a_1, \dots, a_k \in \mathbb{N}$, then $n \mid a_j$ for some $j = 1, \dots, k$.*

Problem 2.7.8. The atomic property of primes states that a necessary condition for $n \in \mathbb{N}$, $n \geq 2$, to be prime, is that for all $a, b \in \mathbb{Z}$

$$n \mid ab \Rightarrow n \mid a \text{ or } n \mid b. \tag{2.7.1}$$

Show that this is also a sufficient condition, and thus a characterisation of prime numbers.

2.7.9. Example. Suppose 19 divides $4 \times k$, for some $k \in \mathbb{N}$. What does this tell us about k? Well, since 19 is prime and $19 \nmid 4$, we can conclude that 19 divides k.

2.7.10. Exercise. *Check, from the definition of prime number, that if p and q are prime numbers and $p \mid q$ then $p = q$.*

2.8. The fundamental theorem of arithmetic

We close this chapter with a well-known corollary of 2.7.6: the so-called *fundamental theorem of arithmetic,* also known as the *unique factorisation theorem,* which says that each natural number can be written as a product of primes in a unique way (up to a permutation of the factors, of course).

2.8.1. Definition of long products. Suppose $\boldsymbol{a} = (a_i)_{i\in X}$ (X could be $\mathbb{N}$ or $[1\dots N]$), is a sequence of real (or complex) numbers,[6] the product of the first n numbers terms is defined with the following recursion

$$\prod_{i=1}^{n} a_i := \prod_{1\leq i\leq n} a_i := \begin{cases} 1 & \text{if } n = 0, \\ \left(\prod_{i=1}^{n-1} a_i\right) a_n, & \text{if } n \geq 1. \end{cases} \tag{2.8.1}$$

Though weird at first sight, there is nothing wrong with the bottom case ($n = 0$), which is known as the *empty product*. It makes sense to define the empty product to be 1 as the recursion makes perfect sense with this choice.

2.8.2. Remark. Special cases of the long products are:

$a_i = i$ we get the *factorial*, with the shorter traditional notation $n!$ for $\prod_{i=1}^{n} i$ (note that definition implies $0! = 1$);

$a_i = a$ for all i we get the usual power denoted a^n for $\prod_{i=1}^{n} a$ when $n \geq 1$, whereas for $n = 0$ and any a in $\mathbb{R}$ (or any unitary ring) we have $a^0 = 1$ (or the unit of said ring), it is worth noting that when $a = 0$, it follows that $0^0 = 1$ by this definition.[7]

2.8.3. Fundamental theorem of arithmetic.

Recalling the notation for the set of prime numbers as $\mathbb{N}'$, given $n \in \mathbb{N}$ then there exists a unique choice of $m \in \mathbb{N}_0$, $p_1, \dots, p_m \in \mathbb{N}$ such that:

(i) p_i is prime for each $i = 1, \dots, l$, (equivalently $p_i \in \mathbb{N}'$ for each $i = 1, \dots, l$),

(ii) $p_i \leq p_{i+1}$ for each $i = 1, \dots, l-1$,
and

(iii) $n = \prod_{i=1}^{l} p_i$.

Proof The existence of a factorisation of a given $n \in \mathbb{N}$ into a product of primes can be proved by (cumulative) induction. If $n = 1$, then n equals the empty product and this proves the base case. If $n \geq 2$, then either n is prime (ok) or n is composite, meaning that $n = ab$ for some $a, b = 2, \dots, n-1$. By the inductive hypothesis each of a, b can be written as a product of primes, and thus n is a product of primes.
To conclude, we have to prove uniqueness. Suppose that for a given n we have two factorisations:

$$n = \prod_{j=1}^{l} q_j = \prod_{i=1}^{m} p_i \tag{2.8.1}$$

with both sequences $\left(q_j\right)_{j=1,\dots,l}$ and $\left(p_i\right)_{i=1,\dots,m}$, satisfying (i) and (ii). In order to show uniqueness, all we have to show is

$$l = m, \text{ and } q_i = p_i, \ \forall\, i = 1, \dots, m. \tag{2.8.2}$$

We proceed by induction on n.

[6] In fact this definition works with anything that can be multiplied, including square matrices and elements of a group or multiplicative algebra.

[7] One argument for $0^0 = 1$ being a "good" definition is that the binomial theorem is valid when 0 is one of the summands:

$$(a+0)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} 0^i = a^n + 0 = a^n$$

.

The base case $n = 1$ is trivial. The only way 1 can be written as a product of primes is the empty product.[8] If you don't like empty products, you may start with the base case $n = 2$ and forget about empty products. Since 2 is prime, there is only one way to write it as a product of primes: trivially, $2 = \prod_{i=1}^{1} p_i$ where $p_1 = 2$.
To perform the inductive step, let $n \geq 3$ and suppose the result is true for all $N < n$, we want to show that it holds also for n. Since

$$p_1 \cdots p_m = q_1 \cdots q_l, \tag{2.8.3}$$

it follows that $p_m \downharpoonleft q_1 \cdots q_l$ and, by Exercise 2.7.7, we must have $p_m \downharpoonleft q_k$, for some $k = 1, \ldots, l$. Since q_k is prime, it follows that $p_m = q_k$, but $q_k \leq q_l$ and thus $p_m \leq q_l$. The same argument with roles exchanged, leads to $q_l \leq p_m$. It follows that $p_m = q_l$. Simplifying (2.8.3), we obtain that

$$N := p_1 \cdots p_{m-1} = q_1 \cdots q_{l-1} \text{ and } N < n. \tag{2.8.4}$$

By the inductive hypothesis, we know that the factorisation of N in a product of non-decreasing primes is unique. Thus

$$l - 1 = m - 1 \text{ which implies } l = m. \tag{2.8.5}$$

Furthermore $p_i = q_i$ for all $i = 1, \ldots, m-1$; since we also proved $p_m = q_m$, the proof is complete. □

2.8.4. Remark (alternative statement of the fundamental theorem of arithmetic). Each natural number n can be written as a product of powers of all different primes, i.e., there exists $I \in \mathbb{N}_0$, $p_1, \ldots, p_I$ all primes with $p_i < p_i + 1$ for all $i = 1, \ldots, I - 1$ and exponents $a_1, \ldots, a_I \in \mathbb{N}$ such that

$$n = p_1^{a_1} \cdots p_I^{a_I} = \prod_{i=1}^{I} p_i^{a_i}. \tag{2.8.1}$$

The choice of such I, $p_1, \ldots, p_I$ and $a_1, \ldots, a_I$ is unique.

2.8.5. Problem (there are infinitely many primes). *Euclid's theorem on the infinity of primes says that among the positive integers there are infinitely many primes. Using the fundamental theorem of arithmetic, which states that every integer $n \geq 2$ is a product of prime numbers, show by contradiction.*

2.8.6. Remark (YAASOFTA). Let $p_1, p_2, \ldots$ be the sequence of consecutive primes, $p_1 = 2$, $p_2 = 3$, etc. For each natural number n there exists a unique sequence of exponents $a_1, a_2, \ldots$ in $\mathbb{N}_0$, such that $a_{N+k} = 0$ for all $k \geq 1$ and some $N \geq 0$,[9]

$$n = \prod_{i=1}^{N} p_i^{a_i}. \tag{2.8.1}$$

[8] Empty products may cause distress for someone reading about them the first time, but there is nothing wrong about *defining* $\prod_{i=1}^{0} := \prod_{i \in \emptyset} p_i := 1$ in the recursive definition of $\prod$.

[9] Some people call such sequences *ultimately vanishing*.

2.9. Extras

Number theory is one of the most exciting fields of mathematics. It also has one of the hardest problems to be solved, including the Goldbach conjecture, the Riemann hypothesis, the twin-prime conjecture and it is connected to an important conjecture in theoretical computer science affectionately known as $P \neq NP$. The literature is also immense. A nice account on this (skipping technical details) was provided by Tao, 2009.

Exercises and problems on basic number theory

Exercise 2.X.1 (divisibility). (a) Fill in the "…" in the following definition of divisibility:
"Given $n \in \mathbb{Z}$ and $m \in \mathbb{N}$, we say that m *divides* n (or m is a *divisor* of n), and we write $m \rfloor n$, if and only if there exists a … number q such that … "

Hint. Review the definitions in the lecture notes.

(b) Based on the above definition show that $8 \rfloor 16$. (You just need to find the appropriate q in the definition of divisor.)

(c) Show that if $m, n \in \mathbb{N}$ and $m \rfloor n$, then $m \leq n$.

Hint. Use the basic relations between the ordering of numbers, "$\leq$", and the algebraic operations $+$ and $\times$.

(d) Show that if $m, n \in \mathbb{N}$ and $m > n$, then m does not divide n.

(e) Show that 8 does not divide 17.

Hint. Suppose by contradiction that $8 \rfloor 17$, then recall that $8 \rfloor 16$, deduce that $8 \rfloor 1$ and use the above result to conclude.

Exercise 2.X.2 (an even square is a square of even). (a) Fill in the dots in the following statement:
"A number n is *odd* if and only if there exists a number $k \in \ldots$ such that $n = \ldots$"

(b) Show that *if* n is odd *then* n^2 is odd.

(c) Deduce that *if* $2 \rfloor n^2$ *then* $2 \rfloor n$.

(d) Can you state any important theorem that uses the last result?

Exercise 2.X.3 (minimum and maximum). Let $S \subseteq \mathbb{R}$.
(a) Write down a precise definition of $\min S$ and $\max S$.

(b) Prove that the there is at most one number $x \in \mathbb{R}$ such that $x = \min S$.

(c) Find a set $T \subseteq \mathbb{Q}_0^+$ for which $\min T$ does not exist. Here $\mathbb{Q}_0^+$ denotes $\{x \in \mathbb{Q} : x \geq 0\}$.

Exercise 2.X.4 (well-ordering). (a) Does $\mathbb{Z}$ satisfy the Well Ordering Principle? Explain.

(b) Does $\mathbb{Q}$ satisfy the Well Ordering Principle? Explain.

(c) What about $\mathbb{Q}_0^+ := \{r \in \mathbb{Q} : r \geq 0\}$? Explain.

Problem 2.X.5 (cubic root of five is irrational). (a) Deduce from the Atomic Property of Primes that "if $5 \rfloor n^3$ then $5 \rfloor n$", for any $n \in \mathbb{N}$.

(b) Use the above result to show that the cubic root of 5 is an irrational number. (Make sure you know the definition of the cubic root of 5 first!)

Problem 2.X.6 (Euclidean division). Using the Well Ordering Principle, prove that

$$\forall\, n \in \mathbb{Z}, d \in \mathbb{N} : \exists\, q, r \in \mathbb{Z} : n = qd + r \text{ and } 0 \leq r < d. \tag{2.X.6.1}$$

Hint. Consider the set

$$S := \{s \in \mathbb{N}_0 : \exists m \in \mathbb{Z} : s = n - dm\}. \tag{2.X.6.2}$$

Show that is a non-empty subset of $\mathbb{N}_0$ and apply the well-ordering principle of $\mathbb{N}_0$. Relate the minimum of S to (2.X.6.1) and conclude.
You may peak at Johnson (1998, Theorem 1.14).

Problem 2.X.7 (reducing a fraction to its lowest terms). Let $r \in \mathbb{Q}$.
(a) Prove that r has a least possible denominator n, i.e., $r = m/n$ for some $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ and that for no $n' < n$ is it possible to write $r = m'/n'$. We say that such n is *denominator of* r *in lowest terms* and the fraction m/n is the *lowest terms* representation of r as a fraction.

Hint. Introduce the set

$$S := \{l \in \mathbb{N} : \exists\, k \in \mathbb{Z} : r = k/l\}. \tag{2.X.7.1}$$

Show that S is non-empty and apply the well ordering property of $\mathbb{N}$.

(b) Show that if m/n is the lowest term fraction of $r \neq 0$ then m and n are co-prime.

(c) Show that there is a unique pair $(m, n) \in \mathbb{Z} \times \mathbb{N}$ for which $r = m/n$ in lowest terms.

Problem 2.X.8 (lowest common multiple). Given $a, b \in \mathbb{N}$.
(a) Using WOP, prove that they have a least common multiple, which is denoted as $\operatorname{lcm}\{a, b\}$

(b) Prove that

$$ab = \operatorname{hcf}\{a, b\} \operatorname{lcm}\{a, b\}. \tag{2.X.8.1}$$

Problem 2.X.9 (HCF as a $\mathbb{Z}$-linear combination). Find the highest common factor h of 366 and 305. Find $s, t \in \mathbb{Z}$ such that

$$h = 366s + 305t. \tag{2.X.9.1}$$

Problem 2.X.10 (HCF is a multiple of all common factors). Let $m \in \mathbb{Z}$, $n \in \mathbb{N}$, and $h := \operatorname{hcf}\{m, n\}$. Show that if $d \rfloor m$ and $d \rfloor n$ then $d \rfloor h$.

Problem 2.X.11 (extended Euclidean algorithm). (a) Based on the Euclidean algorithm and Bézout's identity, write a pseudocode for (or explain in plain English) an algorithm that, for each given $m, n \in \mathbb{N}$, computes their highest common factor, $\operatorname{hcf}\{m, n\}$, and two integers $s, t \in \mathbb{Z}$ such that

$$\operatorname{hcf}\{m, n\} = sm + tn. \tag{2.X.11.1}$$

(b) Show that this algorithm is *certifying*, i.e., that you can easily use its output to check that the result is correct.

Hint. You will find it useful to recall that $\operatorname{hcf}\{m, n\}$ is the only number that both divides both, and is a $\mathbb{Z}$-linear combination, of m and n.

(c) Implement the algorithm with a computer language of your choice.

Problem 2.X.12 (linear Diophantine equations). *Diophantine equations* are ones that have coefficients and unknowns in $\mathbb{Z}$. Let $a, b, c \in \mathbb{Z}$.
(a) Find a characterisation of a, c such that the following equation is solvable for the single-unknown $x \in \mathbb{Z}$

$$ax = c. \tag{2.X.12.1}$$

(b) What can you say about uniqueness of solutions x of (2.X.12.1).

(c) Find a characterisation of a, b, c such that the following equation has a solution $x, y \in \mathbb{Z}$

$$ax + by = c. \tag{2.X.12.2}$$

(d) What can you say about uniqueness of solutions x, y of (2.X.12.2).

Problem 2.X.13 (Chinese remainder theorem with two equations). Let n_1, n_2 be coprime, and a_1, a_2, we are interested in finding $x \in \mathbb{Z}$ such that

$$0 \le x < n_1 n_2, \tag{2.X.13.1}$$

$$x = a_1 \,(\operatorname{mod} n_1), \tag{2.X.13.2}$$

$$x = a_2 \,(\operatorname{mod} n_2). \tag{2.X.13.3}$$

This is known as the *Chinese remainder* problem. Can you observe analogies between this and Euclidean division?
Hint. Find two integers m_1 and m_2 such that $m_1 n_1 + m_2 n_2 = 1$ (why is this possible?). Then show that $x := a_1 m_2 n_2 + a_2 m_1 n_1$ works.

Problem 2.X.14 (factors and powers). Let $m, l, k \in \mathbb{Z}$, $k \ge 1$, show that $l \rfloor m$ if and only if $l^k \rfloor m^k$. Deduce, for any two integers m and n, and $k \in \mathbb{N}$ that

$$\operatorname{hcf}\{m^k, n^k\} = (\operatorname{hcf}\{m, n\})^k. \tag{2.X.14.1}$$

Hint. Start by proving that $l \rfloor m$ if and only if $l^k \rfloor m^k$.

Problem 2.X.15 (coprime invariance under power). Let $m, n \in \mathbb{Z}$, $k \in \mathbb{N}_0$ such that m and n are coprime, prove that m^k and n^k are coprime.

Problem 2.X.16 (infinity of primes — Euclid). Euclid's theorem on the infinity of primes says that *among the positive integers there are infinitely many primes.* Using the fundamental theorem of arithmetic, which states that *every integer $n \ge 2$ is a product of prime numbers*, show by contradiction.

Problem 2.X.17 (lower bound on prime gaps example). Prove that none of the numbers $2310 + j$ for $j = 2, 3, \ldots, 12$ is prime.
Hint. Avoid brute force[10] and start by noting that $2310 = 2 \times 3 \times 5 \times 7 \times 11$.

Problem 2.X.18 (lower bound on prime gaps). Prove that for any number k there exists a sequence of k consecutive integers, none of which is prime. A *prime gap* is the difference between two consecutive primes and represents how many integers you have to count before reaching the next prime. This problem shows that there are prime gaps that are arbitrarily long.

Problem 2.X.19 (divisibility and Bézout imply HCF). Let $m, n \in \mathbb{Z}$ and $m \neq 0$, show that if $g \in \mathbb{N}$ safisfies

$$g \rfloor m, g \rfloor n, \text{ and} \tag{2.X.19.1}$$

$$g = sm + tn \text{ for some } s, t \in \mathbb{Z}, \tag{2.X.19.2}$$

then

$$g = \operatorname{hcf}\{m, n\}. \tag{2.X.19.3}$$

Exercise 2.X.20 (uniqueness of maximum). Let S be a subset of $\mathbb{R}$. Check that (provided it exists) $\max S$ is unique.
This justifies the notation $m = \max S$, whenever m is *the* maximum of S.

[10] With a computer you could check each number, but there are more elegant approaches.

CHAPTER 3

Logic

> The truth comes as conqueror only because we have lost the art of receiving it as guest.
> — Rabindranath Tagore *The Fourfold Way of India*

3.1. The rise of the set

In the late 19th century it became clear to mathematicians that natural numbers were not the most fundamental objects of mathematics that one could imagine, but that they could be derived from more "elementary" objects which were named "sets", which correspond, give or take some precise definition, to our intuitive notion of sets. Thanks to sets, mathematics, after 5000 years of succesful achievements, was finally laid on solid logical foundations and everybody could sleep tight at night (except Kurt Gödel, a wierd Austrian, but that's a different story) not having to worry that all this work was wasted nonsense. Most of the work on sets was developped by Georg Cantor, a German mathematician, by developing the *Mengenlehre*, which is roughly translated as "Set Theory". This theory, which provoked one of the largest controversies in mathematics,[1] eventually proved quite successful and was refined, sometime too much, through out the 20th century opening the door for very deep philosophical questions. We adopt a practical view-point, advocated by Halmos, 1974 in his "Naive Set Theory" where we view Set Theory as a tool, rather than an object of study for its own sake. We will thus avoid, as much as possible, logico-philosophical discussion, but without compromising the reach of this important mathematical theory.

3.2. Russell's Paradox: an upsetting fact

Intuitively, as we have already seen, a set is a "bag of stuff", where the "stuff" is composed of elements. This analogy is at the same time useful and misleading. In fact, it can even lead to important logical pitfalls.[2]

To make an example of such pitfalls, let us follow the footsteps of Bertrand Russell, a Britton from Cambridge, who on a sunny aftenoon of last century, after eating a particularly buttery scone at the Orchard, dreamt of a set, $\mathscr{S}$, containing of all of the other sets including itself (some kind of a god-set). Then Russell tried to figure out what a certain subset, call it $\mathscr{N}$, of $\mathscr{S}$ might look like. The subset $\mathscr{N}$ that Russell had

[1] Leading 19th century mathematicians, such as L. Kronecker and H. Poincaré, disliked so much Cantor's set theory that they personally insulted him.

[2] As a rule of thumb, finite sets are "tame" and infinite sets are "wild". The sets (or non-sets for that matter) that can provide "interesting" examples are usually objects which are liable to have infinite element. With time, you will learn to deal with infinite sets and you will see that they are not so harmful. In fact they are quite fun.

in mind consisted of precisely those sets X (which are elements of $\mathscr{S}$) that are not elements of themselves, in symbols $\mathscr{N} = \{X \in \mathscr{S} : X \notin X\}$. [3]

When the quiz-time arrived, Russell asked himself, is $\mathscr{N} \in \mathscr{N}$ or is $\mathscr{N} \notin \mathscr{N}$? Of course, either one of these two facts must be true and the other one false (this is a basic rule of Logic as we all agree).

If ($\mathscr{N} \in \mathscr{N}$) is true, given that $\mathscr{N} \in \mathscr{S}$—because $\mathscr{S}$ contains all sets—then $\mathscr{N}$ is one of those X that satisfy $X \notin X$. Thus $\mathscr{N} \notin \mathscr{N}$ which means that ($\mathscr{N} \in \mathscr{N}$) is false. This is a contradiction, and our starting point should be refuted.

Therefore ($\mathscr{N} \notin \mathscr{N}$) must true. But if this were the case, then $\mathscr{N}$ is one of those X that satisfy $X \notin X$ and hence $\mathscr{N} \in \mathscr{N}$, i.e., ($\mathscr{N} \in \mathscr{N}$) is true. So again we are in a contradiction.

What shall we do? We cannot have neither $\mathscr{N} \in \mathscr{N}$ nor $\mathscr{N} \notin \mathscr{N}$. One way out of this quagmire[4] is to forbid the construction of both $\mathscr{S}$ and $\mathscr{N}$ by declaring that

$$\text{\textit{The set of all sets does not exist.}} \tag{3.2.1}$$

So, now that we know what a set cannot be, we ask

$$\text{\textit{Given an "object" how can we decide whether it is a set or not?}} \tag{3.2.2}$$

We will answer this question at the end of this lecture. In fact, we shall see that we cannot define a set, except through certain rules that allow us to identify one when we see it walking down the street by the very way it walks.

But before we embark on the Set Theoretical Boat, we need to "review" a bit of Logic.

3.3. Dirty propositions

The basic construct in logic is what we call *proposition.* A proposition is a particular case of what we call *statement* and a statement is a sentence formed according to the syntactic rules of a language which can be given a complete meaning.

3.3.1. Example (statements). The following are some English statements:

(1) India is the most populated nation in Asia.
(2) It is not raining now.
(3) It will rain tomorrow.
(4) Some bananas are yellow.
(5) All bananas are green.
(6) Tony Blair was Britain's best prime minister ever.
(7) There are no dragons.
(8) My cat has the blues.
(9) The white horse has eaten the green grass.
(10) To be or not to be.

[3] You make ask yourself "how can a set (which is a bag) belong to itself to start with?". Beside this not being the main issue here, if you think that there are no set can belong to itself, then you simply think that $\mathscr{N} = \mathscr{S}$. Of course, you soon realise that this is not true because $\mathscr{S} \in \mathscr{S}$ (by definition of $\mathscr{S}$) and that there's more to $\mathscr{S}$ than merely $\mathscr{N}$. This is a digestible scone.

[4] Another way out would be to accept the fact that a proposition may be neither true nor false. This is an un-orthodox school of thought which deserves more attention than mathematicians usually give it, but not in such an orthodox course as ours. Most mathematicians, with notable exceptions, follow our way.

Intuitively, we want a *proposition* to be a statement that can be assigned a definite truth value; i.e., a proposition can be decided to be true or false.

3.3.2. Example (propositions and non-propositions). Some propositions among the statements above are 1 (false), 4 (true), 5 (false), 2 (have a look outside and tell), 10 (true). Indeed, we can decide whether each one of these statements is true or false. Some statements which are *not* propositions are 6, 3, 7 and 8. Indeed, none of these statements can be unequivocally decided to be true or false, as the response to them is a matter of personal opinion (including the cat's).

Note that in real life it is not always easy to decide whether a statement is a proposition or not. Think about the following statement: "My aunt has 33043 hairs on her head." Is this a proposition?

A problem we face with the above sentence is that we need to define exactly what a head is: should we count what grows on our aunt's neck? And what a hair really means. Are the hairs that have fallen considered hairs? And, assuming they are, should they be still on the skull, or a hair accidentally on auntie's tongue is considered also a hair? Are eyebrows "hair"? What about a wig? And you may not even have an aunt to start with!

We see here that deciding whether an "everyday's life" statement is a proposition or not is not an easy business. Mathematical modelling (using Maths to describe the real world) is an interesting topic, but in this course, we are pure minded, and we will not be facing the dirty propositions of everyday's life. We will deal with purely mathematical object, as those figments of our imagination called sets, for which propositions become quite easy to handle.

3.4. A cleaner approach with some rules

In mathematics, all propositions can be built from more elementary propositions according to some rules. All mathematical propositions can be built by using sets only.[5] These rules can be summarised (with some redundancy) as follows:

Rule 1. For any two sets x and y are sets the statement $\left(x \in y\right)$ is a proposition (i.e., either true or false). This proposition is also pronounced "*x is an element of* the set y", or simply "*x belongs to y*".

Rule 2. If P, Q are propositions, then

$$P \text{ and } Q \tag{3.4.1}$$

is a proposition.

Rule 3. If P, Q are propositions, then

$$P \text{ or } Q \tag{3.4.2}$$

is a proposition.

Rule 4. If P is a proposition, then

$$\text{not}\, P \tag{3.4.3}$$

is a proposition.

[5] Mathematicians forget very soon that this is a fact and seldom reduce their sentences to the basic rules; but they (almost) all agree that this is possible.

Rule 5. If P, Q are propositions, then

$$P \Rightarrow Q \tag{3.4.4}$$

is a proposition. The proposition in (3.4.4) is usually read as "*if* P is true *then* Q is true" or simply "P implies Q". A strange way to pronounce this proposition is "Q is true only if P is true" and a yet stranger way to pronounce it is "P is false if Q is false".

Rule 6. If P, Q are propositions, then

$$P \Leftrightarrow Q \tag{3.4.5}$$

is a proposition. The proposition (3.4.5) is usually pronounced as "P is true *if and only if* Q is true" or simply "P is equivalent to Q".

Rule 7. If $P(\cdot)$ is a proposition that depends on a variable x and A is a given set, then

$$\forall\, x \in A \,:\, P(x) \tag{3.4.6}$$

is a proposition. The proposition in (3.4.6) is pronounced "*for any* (or *for all*) element(s) x of the set A *we have that* $P(x)$ is true".

Rule 8. Suppose that $P(\cdot)$ is a proposition that depends on x and A a given set. Then

$$\exists\, x \in A \,:\, P(x) \tag{3.4.7}$$

is a proposition. The proposition in (3.4.7) is pronounced "*there exists* (at least) one element x of the set A *such that* $P(x)$ is true".

3.4.1. Remark. We will explain next in some detail what all these rules are about. Notice that rule 1 will be explained in details in Chapter 4.

Rules 2, 3, 4, 5, 6 do not involve sets explicitly and we study them next. Each rule defines a logical operator.

Finally, rules 7 and 8 are called "quantification" rules and will be also discussed next.

3.4.2. Logical operations. Rules 2–6 are useless unless we define what the involved operations mean. Of course, we want these definitions coincide with the linguistic (English) meaning of *and, or, not, if/then,* and *if and only if,* so our definitions must match our intuition.

To define a logical operation, it is enough to examine its truth value for all the possible combinations of truth values of the operands.

We start with the operation " and " which is defined by

$$(P \text{ and } Q) \begin{cases} \text{is false,} & \text{when } P \text{ is false and } Q \text{ is false,} \\ \text{is false,} & \text{when } P \text{ is false and } Q \text{ is true,} \\ \text{is false,} & \text{when } P \text{ is true and } Q \text{ is false,} \\ \text{is true,} & \text{when } P \text{ is true and } Q \text{ is true}. \end{cases} \tag{3.4.1}$$

The operation " or " is similarly defined by

$$(P \text{ or } Q) \begin{cases} \text{is false,} & \text{when } P \text{ is false and } Q \text{ is false,} \\ \text{is true,} & \text{when } P \text{ is false and } Q \text{ is true,} \\ \text{is true,} & \text{when } P \text{ is true and } Q \text{ is false,} \\ \text{is true,} & \text{when } P \text{ is true and } Q \text{ is true}. \end{cases} \tag{3.4.2}$$

The operation “not”, known as logical *negation*, is defined by

$$(\text{not}\, P) \begin{cases} \text{is true}, & \text{when } P \text{ is false}, \\ \text{is false}, & \text{when } P \text{ is true}. \end{cases} \tag{3.4.3}$$

The operation “ ⇒ ”, known as *implication*, is defined by

$$(P \Rightarrow Q) \begin{cases} \text{is true}, & \text{when } P \text{ is false and } Q \text{ is false}, \\ \text{is true}, & \text{when } P \text{ is false and } Q \text{ is true}, \\ \text{is false}, & \text{when } P \text{ is true and } Q \text{ is false}, \\ \text{is true}, & \text{when } P \text{ is true and } Q \text{ is true}. \end{cases} \tag{3.4.4}$$

Finally, the operation “ ⇔ ”, known as logical *equivalence*, is defined by

$$(P \Leftrightarrow Q) \begin{cases} \text{is true}, & \text{when } P \text{ is false and } Q \text{ is false}, \\ \text{is false}, & \text{when } P \text{ is false and } Q \text{ is true}, \\ \text{is false}, & \text{when } P \text{ is true and } Q \text{ is false}, \\ \text{is true}, & \text{when } P \text{ is true and } Q \text{ is true}. \end{cases} \tag{3.4.5}$$

3.4.3. A warning about “⇒” and its misuses. The logical operation “ ⇒ ” may cause some confusion when comparing it with the linguistic usage of “if–then”. The reason behind this confusion is that we tend to forget that P is hypothetical, so it may be false even if the proposition $(P \Rightarrow Q)$ is true. Also, the implication is used often to perform logical inference (described next) in which one knows $(P \Rightarrow Q)$ is true and tries to use this to conclude something about Q or P based on some knowledge of the other proposition.

A typical situation is when we would like to prove Q is true and we know of a theorem that says $(P \Rightarrow Q)$. It is enough to show that P is true; indeed, if $(P \Rightarrow Q)$ is true and P is true then, this rules out the three first cases in (3.4.4), and by the fourth case it must be that Q is true. This is called *logical inference* and is the most popular use of the implication.

As an example, suppose $P \Leftrightarrow$“it is raining” and $Q \Leftrightarrow$“there are clouds”, then we know $(P \Rightarrow Q)$ is true. Now suppose you get out of the house with closed eyes and you feel that drizzle on you skin (i.e., P is true), then you immediately know that Q is true.

Another typical use of the implication is the *backward (or negative) inference*. Suppose that you know, say by some theorem, that $(P \Rightarrow Q)$ is true and you manage to show that Q is false, then the last three cases in (3.4.4) are ruled out and the only possibility left for P is to be false.

Sticking to our example, suppose that before stepping out of the house you see and absolute clear sky (i.e., Q is false) then you know automatically that you won’t get any rain (i.e.,. P is false).

Note that if we know that P is false, then knowing that $(P \Rightarrow Q)$ is true is inconclusive as to whether Q is true or not. The **typical mistake** (made by many daily) is “$(P \Rightarrow Q)$ is true and P is false hence Q is false”. Indeed, coming back to our example, P is false means that it is not raining. But, as you know too well, this is not enough for us to infer that there are no clouds.

3.4.4. Necessary and sufficient conditions. In mathematical (and logical) parlance when $(P \Rightarrow Q)$ is true, we say that Q is a *necessary condition* for P, and that P is a *sufficient condition* for Q.

In our meteorological example, "to be raining" is a sufficient, but not necessary, condition for it "to be cloudy" and viceversa, "being cloudy" is necessary, but not sufficient, for it "to be raining".

When we have $(P \iff Q)$ we say that Q, respectively P, is a *necessary and sufficient condition* for P, or Q. Sometime the terms *characterisation* or *criterion* are used to indicate a necessary and sufficient condition.

3.4.5. Quantifiers. The quantifier $\forall$ in (3.4.6) is called the *universal quantifier.* Its meaning in English is "for all" (or "for any"). The proposition in (3.4.6) is defined by

$$(\forall\, x \in A : P(x)) \begin{cases} \text{is true} & \text{when } P(x) \text{ is true for } \textit{each } x \text{ of the set } A, \\ \text{is false} & \text{when } P(x) \text{ is false for } \textit{at least one } x \text{ of the set } A. \end{cases} \tag{3.4.1}$$

As an example consider the proposition $P(x) :\iff$ "x is positive". If we take $A = \mathbb{N}$, then $(\forall\, x \in A : P(x))$ is true. If however we take $A = \mathbb{N}_0$, then $(\forall\, x \in A : P(x))$ is false.

The other quantifier $\exists$ in (3.4.7) is called the *specification* or existential quantifier. The English translation of $\exists$ is "there is at least one" (or "there exists one"). Therefore, it makes sense to define the proposition in (3.4.7) by

$$(\exists\, x \in A : P(x)) \begin{cases} \text{is true} & \text{when } P(x) \text{ is true} \text{for } \textit{at least one } x \text{ of the set } A, \\ \text{is false} & \text{when } P(x) \text{ is false} \text{for } \textit{each } x \text{ of the set } A. \end{cases} \tag{3.4.2}$$

3.4.6. Truth tables. The definitions of " and , or , not, $\Rightarrow$, $\iff$" can be summarised in truth tables. These are just handy devices (and nothing more than that!) to make the writing shorter and reading faster. Also $\forall$ and $\exists$ can be summarised in truth tables (which may be infinite tables, but nevertheless tables, if the set A in Rule 7 and 8 is infinite).

For example the truth table for $\Rightarrow$ can be written as follows

P	Q	P⇒ Q
is false	is false	is true
is false	is true	is true
is true	is false	is false
is true	is true	is true

(3.4.1)

When people get bored with writing " is false" and " is true" (a phenomenon that happens quite fast usually), they replace them with "0" and "1", respectively, without attaching too much numerical meaning to these symbols. That's it.

As an exercise, write down the truth tables for all the logical operations introduced so far.

3.4.7. Remark (redundant rules). Some of the rules 2–6 can be derived from other rules. For example, rule 5 is in fact a consequence of rule 3 and rule 4 because

$$(P \Rightarrow Q) \text{ is equivalent to } (Q \text{ or not} P). \tag{3.4.1}$$

In fact we just need rules 1, 4 and 7: you may enjoy an afternoon by showing that all the other rules are in fact consequences of these three.

Exercises and problems on logic

Exercise 3.X.1 (implication). (a) Let P and Q be two propositions. Show that

$$(P \Rightarrow Q) \text{ is logically equivalent to } (\text{not}\, Q \Rightarrow \text{not}\, P) \tag{3.X.1.1}$$

$$(P \Leftrightarrow Q) \text{ is logically equivalent to } ((Q \Rightarrow P) \text{ and } (P \Rightarrow Q)) \tag{3.X.1.2}$$

Hint. You just have to show that the right-hand side and left-hand side have the same truth-value for all possible combinations of truth-value for P and Q. You may use truth tables.

(b) Suppose P, Q are propositions and you want to prove a Theorem that says $P \Rightarrow Q$. Based on the definition of this operation, given in class, explain why it is enough to follow the procedure: "assume P is true and show that Q is true". In other words, what about the other 3 cases of truth-value of P and Q?

Exercise 3.X.2 (De Morgan Laws). Let P and Q be two propositions. Show the following are true:

$$\text{not}(\text{not}\, P) \Leftrightarrow P, \tag{3.X.2.1}$$

$$\text{not}(P \text{ and } Q) \Leftrightarrow \text{not}\, P \text{ or } \text{not}\, Q, \tag{3.X.2.2}$$

$$\text{not}(P \text{ or } Q) \Leftrightarrow \text{not}\, P \text{ and } \text{not}\, Q. \tag{3.X.2.3}$$

The last two equivalences are called De Morgan's laws.
Hint. We are using the symbol " $\Leftrightarrow$ " as a shortcut to "is logically equivalent to". You just need to show that the left-hand side of each equivalence has the same truth-value of the right-hand side for all the possible combinations of truth-value of P and Q. For the first equivalence distinuish two cases (P is false, P is true) and for the last two, distinguish four cases (P is false and Q is false, etc.). You may use truth tables if you like them.

Exercise 3.X.3 (ifisornot). This exercise's aim is to write "$\Rightarrow$" (if-then) by using " or " and "not".

(a) Let P and Q be two propositions, show that

$$(P \Rightarrow Q) \Leftrightarrow (Q \text{ or } \text{not}\, P). \tag{3.X.3.1}$$

(b) The affectionate name of equivalence (3.X.3.1) is *Ifisornot.* Can you explain the origin of this word?

(c) If a logician sends you a postcard which says: "I'll be delighted or you won't come to visit me", what should be your polite response?

Problem 3.X.4 (logic and chocolate). A sweet maker produces two types of chocolate: dark and milk, with the rule that *the dark chocolate must be wrapped in red or blue paper.* A workshop sample tray is composed of three wrapped chocolates: one yellow, one red and one blue.

(i) What is the minimum number of chocolates one must unwrap, in order to ensure the rule holds true for this sample? Explain.

(ii) Two persons, of which one is a dark-only chocolate lover and one is a milk-only lover, are presented with the tray. Both persons know the rule and are asked to choose exactly one of the wrapped chocolates. Which one is guaranteed to satisfy her taste? Explain.

Exercise 3.X.5 (transitivity of implication). (a) Suppose $(P \Rightarrow Q)$ is true and $(Q \Rightarrow R)$ is true show that $(P \Rightarrow R)$ is true.

(b) What can you then say about

$$((P \Rightarrow Q) \text{ and } (Q \Rightarrow R)) \Rightarrow (P \Rightarrow R). \tag{3.X.5.1}$$

CHAPTER 4

Sets

No one shall expel us from the Paradise that Cantor has created.
– David Hilbert (on Georg Cantor's work on set theory)

Il semble que la perfection soit atteinte non quand il n'y a plus rien à ajouter, mais quand il n'y a plus rien à retrancher.
It seems that perfection is attained not when there is nothing left to add, but when there is nothing left to take away.
– Antoine de Saint-Exupèry *Terre des Hommes.*

4.1. The joy of sets

Just like we did with (clean) propositions, we shall build Set Theory by enumerating some fundamental rules that allow to build sets from known sets. These fundamental rules are known as *axioms.*

The axioms of Set Theory have been tested against time (and many minds) and they constitute a minimal list[1] that is required and sufficient to build all the mathematical knowledge in a rigorous way, i.e., by following the rules of logic and no more than that. Believe it or not, some egregious mathematicians and logicians, such as Ernst Zermelo and Abraham Fraenkel, building on the work of Set Theory pioneers, such as, Giuseppe Peano (who identified the fundamental properties of natural numbers and arithmetic) and Georg Cantor (who produced the first version of axiomatic set theory), spent part of their lives during the early 20th century proving that these axioms do not contradict one another (consistency) and that all the known mathematical theories can be derived from them (completeness).

Then came Kurt Gödel, a young man who showed that any axiomatic system, say $\mathscr{F}$, that can generate natural numbers (i.e., that has some basic arithmetic) and thus propositions (by a smart correspondence between numbers and logical propositions), will also have a proposition $G_{\mathscr{F}}$ (known as $\mathscr{F}$'s Gödel's sentence) which cannot be proved true nor false by $\mathscr{F}$ *alone.* This implies (for us) two things: (1) our previous "definition" of proposition (as a sentence that is either true or false) is actually wrong and we have to review it, (2) if an axiomatic system is powerful enough to generate propositions, then it is too weak to also prove all such propositions.

We will not do all this here (phew!) and we shall live by parasitically enjoying the results of these great people's hard work with confidence. We will just list the axioms and show *how to use them.*

[1] Strictly speaking, we should say almost minimal list. Some of the axioms we present here become redundant once all the axioms have been preseneted. Redundancy is not a problem for us, consistency (i.e., that a proposition cannot be true and false at the same time), or the lack thereof, would be. There are many ways of tweaking the axioms of Set Theory in a way to avoid inconsistencies, but they all eventually lead to similar results as far as "traditional mathematics" goes.

There are many approaches to axiomatic set theory, and we shall follow the most standard one (delegating the discussion of whether this is good or bad choice to people less busy than we are). According to Zermelo and Fraenkel's theory, there are 8 axioms in all. An extra axiom was added later, that is the (in)famous *Axiom of Choice*. These axioms, 9 in total, constitute the logical basis of (most) contemporary mathematics known by the name of *Zermelo–Fraenkel–Choice* (*ZFC*).
Let's begin.

4.2. Extension and equality

4.2.1. Definition of belong, element, member. The atomic predicate (or elementary proposition) in set theory is "x *belongs* to y", also pronounced "x is an *element* (also known as *member*) of y" it is also denoted $x \in y$.

Axiom 1 (extension)**.** *Two sets are equal are if and only if they have exactly the same elements.*

4.2.2. Example (supermarket set-theory). Suppose you have two shopping carts. The first shopping cart contains: a box of tea, a toothbrush, a tomato and a courgette. The second shopping cart contains: a courgette, a tomato, a box of tea and a toothbrush. For the (apprentice) set-theorist, who sees the two carts as two sets through his newly acquired set-theoretical shades, these two carts are equal.
The annoying (expert) set-theorist will tell you that the tomato in cart 1 is different than the tomato in cart 2, so the carts aren't really the same set. Of course the expert has got a point here, but we shall ignore this annoyance for the moment and say that the tomato in cart 1 is the same tomato as in cart 2 (and the same for the courgette, toothbrush, etc.)
As with all "everyday's life" examples, there are limitations. So let us turn to something more mathematical. It is time to nail down an important notation which comes implictly with the Axiom of Extension. Sets can be enumerated by listing their elements inside *curly braces*.
There are two ways to do this:

(a) We list all the elements, and this procedure is possible only with finite sets (because our life is finite and our descendants may refuse to engage in such an interesting procedure as listing the elements of an infinte set).
(b) We give a property that describes all the elements of the set. By a "property that describes" we mean a "decisive test" that will tell you whether something belongs to the set or not.

Note that notational orthodoxy wants us to enclose the elements of a set in *curly braces* in order to indicate the set, regardless of whether we use method (a) or method (b).

4.2.3. Example (extension by enumeration). The enumeration procedure and the Axiom of Extension imply that the sets

$$A := \{1,2,3\}, B := \{3,1,2\}, C := \{3,3,2,1,2,3,3,3\}, \tag{4.2.1}$$

are, despite the appearances, equal.

4.2.4. Example (by description). The descriptive approach is more time-saving (and more powerful). For example the sets

$$H := \{\text{humans}\}, \tag{4.2.1}$$

$$G := \{\text{women and men and children}\}, \tag{4.2.2}$$

$$F := \left\{\text{earthling creatures that sometime speak, think, cry or silently listen to music}\right\} \tag{4.2.3}$$

are the same sets (without going into deep philosophical discussions, unrelated to set-theory).
Propositions built using Rule 1 and Extension Axiom 1 are

- ⋆ apple $\in H$ (false),
- ⋆ Lily Allen $\in H$ (true),
- ⋆ $H = G$ (true),
- ⋆ $F = G$ (true).

4.2.5. Definition of $\in$. We have been using the symbol "$\in$" to indicate *membership* and that is how we will continue using it. So whenever we see $x \in X$, we mean (and we say) that x (and x could be "anything") is an *element* of X. And viceversa.
If x is not a member, or element, of a set A we write $x \notin A$.

4.2.6. Remark (a club that is fussy about membership). Notice that there is subtlety in the use of "$\in$" which may appear as an excess of fuss, but which is quite useful to respect, especially for the beginners. For example, suppose P is the set of all point in the plane. Suppose x and y are two different points in the plane, then we may write $x \in P$ and $y \in P$. Now consider the set $\left\{x, y\right\}$, which contins exactly x and y and ask "Is $(\left\{x, y\right\} \in P)$ true?
The answer, surprisingly to many, is "No." and we write $\left\{x, y\right\} \notin P$. Indeed, we declared P to be the set whose elements are exactly all the points in the plane, but we did not say that its elements are the "sets of two points in the plane".
In fact, we will be so fussy as to require $\{x\} \notin P$. This stems from the fact that x and $\{x\}$ are two different objects. The first is merely "the point x" whereas the second is "the set that contains x as its only element". Thus we have $x \in P$ but $\{x\} \notin P$. This may seem a pathologically obsessive excess of zeal, but the record show this distinction has saved (or the lack of it has doomed) many mathematical arguments from (to) failure. We shall therefore make this distinction.
Suppose, to make a further example, that ℓ is the straight line in the plane passing through the points x and y, and let z be a third point such that x, y and z are the vertices of a triangle. Can you tell which one of the follwing propositions is true and

which one is false:

$$x \in \ell, \tag{4.2.1}$$

$$y \notin \ell, \tag{4.2.2}$$

$$\{y\} \notin \ell, \tag{4.2.3}$$

$$z \in \ell, \tag{4.2.4}$$

$$z \notin \ell, \tag{4.2.5}$$

$$\ell \in P. \tag{4.2.6}$$

4.2.7. Definition of inclusion, subsets. Let A and B be two sets. We say that A is a *subset* of B, or A is a *part* of B, or A is *included* in B, or B is a *superset* of A, of B *includes A*if and only if

$$(x \in A \Rightarrow x \in B) \text{ is true}. \tag{4.2.1}$$

An alternative (equivalent) definition is

$$\forall\, x \in B : x \in A. \tag{4.2.2}$$

When A is a subset of B, we write "$A \subseteq B$" or "$A \subseteq B$". We prefer the second notation, because it dispells any doubt about whether A is allowed or not to be equal to B. We will therefore use $\subseteq$ and eschew $\subseteq$, with a warning in that we allow ourselves slip over this rule inadvertently (especially when we are in a hurry).

4.2.8. Definition of inequality, proper subsets. Let A and B be two sets. Whenever $(A = B)$ is false we say that A and B are *different* and we write

$$A \neq B. \tag{4.2.1}$$

If $A \subseteq B$ and $A \neq B$ then we write $A \subsetneq B$

4.2.9. Exercise. *Make up some examples that illustrate situations in which $\neq$ and $\subsetneq$ are used.*
Write a proposition that characterises "$A \neq B$" by using only the logical operations and symbols introduced in 3.4.
Write a proposition that characterises "$A \subsetneq B$".

4.2.10. How to prove equality of two sets? By Axiom 1, to show that two sets, say A and B, are equal it is enough to show two things:

$$(A \subseteq B) \text{ and } (B \subseteq A). \tag{4.2.1}$$

Examples will make this often used procedure clear.

1. Let us nail down this technique by making a silly example first: Consider $E := \{3,1,2\}$ and $F := \{3,3,2,1,2,3,3,3\}$. We claim that $E = F$.

   Indeed, we have that $E \subseteq F$: just go through each element of E and chekc that it belongs to F:

$$3 \in F \text{ is true}, 1 \in F \text{ is true}, 2 \in F \text{ is true}. \tag{4.2.2}$$

Conversely, going through each element of F (which is a bit more tedious than for E) we see that (omitting the " is true"):

$$3 \in E, 3 \in E, 2 \in E, 1 \in E, \dots, 3 \in E, \tag{4.2.3}$$

and therefore $F \subseteq E$.

2. Let us now make an example which still somewhat silly but a bit more interesting: consider the set X of natural numbers that are *even and divisible only by* 1 *and themselves* and the set Y of integers that are *primes and less than or equal to 2*. We claim that $X = Y$.

Indeed, let us first show that $X \subseteq Y$. Pick a generic element $n \in X$, since $n \in \mathbb{N}$ and n is divisible only by 1 and itself, then either $n = 1$ or n is prime. But n is even, so it must be also divisible by 2. Therefore $n = 2$. It follows that n is a prime which is less than or equal to 2 and can be admitted in Y. Since n is a generic element of X we have $X \subseteq Y$. (Note that incidentally we proved, what many knew all the way from the start, that $X = \{2\}$. This makes all this talk sound a bit too long for such a simple fact... but we're not after economic discourse. Not yet, at least.)

Conversely, suppose now n is a generic element of Y, we want to show that $n \in X$. Since n must be prime, then it is clearly divisible only by 1 and itself and $n \geq 2$ (by definition of prime). But we know also that $n \leq 2$ (because of $n \in Y$), so we must have $n = 2$, which is a member of X. Thus $n = 2$. (Again we have incidentally shown that $Y = \{2\}$, which may have simplfied our discourse...)

3. Let us finish with in example that requires some work (and involves infinite sets). Consider the set P made of integers that are squares of some other integer and R the set of integers that are squares of rational numbers. While you may find it easy to list the first few elements of P by using some recipe (and a calculator), you may be a bit puzzled about R at first sight. In fact we shall prove that $R = P$ and get the fog out of our way.

Let us first prove $P \subseteq R$ which is not so hard. Let $n \in P$. Then by definition of P there exists a $k \in \mathbb{Z}$ such that $n = k^2$. Since $k \in \mathbb{Q}$ then it follows that $n \in R$.

The converse inclusion $R \subseteq P$ requires a bit more work. Let $n \in R$, then by definition of R, there exist $r \in \mathbb{Q}$ such that $n = r^2$. Let $l \in \mathbb{Z}$ and $k \in \mathbb{N}$, be coprime and such that $r = l/k$. We want to show that $r \in \mathbb{Z}$, i.e., $k = 1$. Note first that l^2 and k^2 are coprime (if you haven't proved this yet, do it now), hence by Bézout's identity we have, for some $s, t \in \mathbb{Z}$, that

$$1 = sl^2 + tk^2 = snk^2 + tk^2 = (snk + tk)k, \tag{4.2.4}$$

which means that $k \rfloor 1$, and thus $k = 1$.

4.2.11. Definition of formula. Let A be a set. A *univariate formula* $S(x)$ *with domain (or defined on)* A is a statement involving set-theoretical and logical constructs (according to the rules of Logic stated in 3.4) such that for each instance of the *variable* x as an element of the set A, then the statement becomes a proposition.

4.2.12. Example (formula). Consider the statment that depends on the variable n

$$n \text{ is an even number.} \tag{4.2.1}$$

This is a formula defined on the set of natural numbers. It is also defined on the set of integers (or even reals for that matter). But it is not defined on the set of points in

the 3-dimensional space, because the concep of even number simply does not make sense for such objects as points in the space.[2]
It is customary to give a short name to formulas. For instance, call the formula in (4.2.1) $S(n)$. Then we may say "the proposition $S(2)$ is true", "$S(3)$ is false", etc.

4.2.13. Definition of multivariate formula. A formula may depend on more than one variable, if this is the case we call it a *multivariate formula*. For example the statement

$$F(n,m) :\Longleftrightarrow n \text{ and } m \text{ are parallel} \tag{4.2.1}$$

is a formula defined on the set of pairs of straight lines in the plane. It is alsow a formula on the set of straigh lines (and even planes) in 3-dimensional space. Note that it is not defined on the set of natural numbers, $\mathbb{N}$, simply because we do not have the concept of parallelism for numbers.
A multivariate formula, with one (or more) of its variables specified, known as *parameters*, becomes a formula in all the remaining variables which we call *free variables*. In symbols, suppose $F(n_1,\dots,n_k,\dots,n_{k+l})$ is a formula for some integers $k,l \geq 1$. and suppose $n_{k+1},\dots,n_{k+l}$ have each a specified value. Then we have defined a (sub) formula as

$$G(n_1,\dots,n_k) := F(n_1,\dots,n_k,\dots,n_{k+l}). \tag{4.2.2}$$

One says that G is F with *free variables* $n_1,\dots,n_k$ and *parameters* $n_{k+1},\dots,n_{k+l}$.

4.3. Specification and intersection

Axiom 2 (specification)**.** *Given a set A and a formula P defined on A , then there exists a set B formed of all elements x of A that satisfy $P(x)$, i.e., all elements x that make $P(x)$ a true proposition.*

4.3.1. Remark.

(i) The set B appearing in Axiom 2 is denoted by

$$B := \{x \in A : P(x)\} \tag{4.3.1}$$

(ii) An immediate consequence is that $B \subseteq A$.

(iii) Axiom 2 is also known as the *Axiom of Comprehension* by some and the *Axiom of Separation* by others.

(iv) Strictly speaking the Axiom of Specification is an axiom of axioms, in that, for each formula $P(x)$ with free variable in x, we have can build a set Y from a set X on which P is a formula. This is why some authors refer to it as an *axiom schema*, and it is often found as the *Axiom Schema of Specification* (or *Comprehension*, or *Separation*).

(v) Note that we are building B by specifying it from an "ambient set" A. The reasons for doing this are two: the first reason is that the formula P needs a domain to be well defined and the second reason is that by doing so we avoid pitfalls such as Russell's Paradox.

[2] If we were to take the plane you could get away with this being a formula. This is slightly off the topic though.

4.3.2. Example (from "Real Life"). Suppose we want to define formally the set of married people.
First we need an ambient set, so consider $H :=$ {all humans} and let

$$P(x) :\Leftrightarrow \text{"}x \text{ is married".} \tag{4.3.1}$$

The set of elements of H that satisfy P (i.e., the set of humans h such that h is married) is defined by specifying P on H

$$M := \{h \in H : P(h)\}. \tag{4.3.2}$$

Similarly we could have built the set of wives (married women) W. This can be done in more than one way. One way to do it is

$$F := \{x \in H : x \text{ is female}\} \tag{4.3.3}$$

$$W := \{x \in F : P(x)\}. \tag{4.3.4}$$

But we could have also built a new formula

$$Q(x) :\Leftrightarrow (x \text{ is a female}) \text{ and } P(x) \tag{4.3.5}$$

$$W : \{x \in H : Q(x)\}. \tag{4.3.6}$$

Alternative ways to build W is to write

$$W := \{x \in F : x \in M\} \tag{4.3.7}$$

$$W := \{x \in H : (x \in F) \text{ and } (x \in M)\}. \tag{4.3.8}$$

The last two constructions use, without naming it, the concept of intersection, which we shall define next.

4.3.3. Definition of intersection of two sets. Let X and Y be two sets, the *intersection* (also known as the *meet*) of X with Y is defined as the set

$$\{x \in X : x \in Y\}. \tag{4.3.1}$$

The intersection of X with Y is briefly denoted by

$$X \cap Y. \tag{4.3.2}$$

The fussy will note that for any given (sets) x and Y, "$x \in Y$" is a proposition thanks to Logic Rule 1. Therefore, in the terminology of Axiom of Specification (Axiom 2), we are using the formula $F(x, Y) :\Leftrightarrow (x \in Y)$, with Y as a parameter and x as a free variable. In view of the Axiom of Specification, the intersection of two sets is thus a set.

4.3.4. Exercise (some basic properties of intersection). *Show that if A, B and C are given sets then*

$$A \cap B = B \cap A \qquad (\textit{commutativity of intersection}) \tag{4.3.1}$$

$$(A \cap B) \cap C = A \cap (B \cap C) \qquad (\textit{associativity of intersection}) \tag{4.3.2}$$

$$A \cap B \subseteq A \qquad (\textit{minimality of intersection}) \tag{4.3.3}$$

$$A \subseteq B \Leftrightarrow A \cap B = A \qquad (\textit{monotonicity of intersection}) \tag{4.3.4}$$

$$A \cap \varnothing = \varnothing \qquad (\textit{absorbing element for intersection}). \tag{4.3.5}$$

4.3.5. Intersection of collections. Note that associativity allows to "drop the brackets" because the result is the same for any choice of "grouping". In fact, associativity allows to define the intersection for any finite set of sets. For example let $\mathscr{X} = \{X_1, X_2, X_3, X_4, X_5\}$ be a collection (or family) consisting of 5 sets, then $\mathscr{X}$'s intersection is defined as

$$\bigcap_{i=1}^{5} X_i := \{x \in X_1 : \forall\, i \in [2 \dots 5] : x \in X_i\}. \tag{4.3.1}$$

The intersection operation can be extended to any *collection* or *family*. Collection (or family) is just another word for set; the difference in terminology lies in the use we make of such sets: when we want to emphasise that we are dealing with a set of sets and manipulate the whole, we refer to it as a collection (or family). Let $\mathscr{C}$ be a nonempty collection (of sets) then we define its intersection to be the set

$$\bigcap_{X \in \mathscr{C}} X := \{x \in X_0 : \forall\, X \in \mathscr{C} : x \in X\}, \tag{4.3.2}$$

where $X_0 \in \mathscr{C}$.

This notation is quite confusing, because it is too short, and most mathematicians prefer to make it longer by putting an index as follows: let

$$\mathscr{C} = \{X_i : X_i \text{ is a set for each } i \in \mathscr{I}\}, \tag{4.3.3}$$

where $\mathscr{I}$ is a given set of indexes. The intersection of $\mathscr{C}$ can be then defined as

$$\bigcap_{i \in \mathscr{I}} X_i := \left\{x \in X_{i_0} : \forall\, i \in \mathscr{I} : x \in X_i\right\} \tag{4.3.4}$$

where $i_0 \in \mathscr{I}$.

4.3.6. Exercise. *Show that the intersection of a collection $\mathscr{C}$, as defined by* (4.3.3) *is the same if we exchanged the set X_0 with some other set, say $X_1 \in \mathscr{C}$. We refer to this fact by saying that the definition of intersection* (4.3.3) *is independent of, or invariant with respect to, the choice of the "intersection base" set X_0.*

4.3.7. Example (intersection). Consider the sets

$$E := \{1,2,3,4\}, \quad F := \{1,3,5,0\}, \quad G := \{0,1,4,9\}. \tag{4.3.1}$$

Then we have

$$E \cap F = \{1,3\}, \quad E \cap G = \{1,4\} \quad F \cap G = \{0,1\}, E \cap F \cap G = \{1\}. \tag{4.3.2}$$

4.3.8. Example (multiples and intersection). Consider the following sets

$$M_2 := \{n \in \mathbb{Z} : (\exists\, k \in \mathbb{Z} : n = 2k)\} \qquad \big(\text{sometime written as } \{2k : k \in \mathbb{Z}\}\big), \tag{4.3.1}$$

$$M_3 := \{n \in \mathbb{Z} : (\exists\, k \in \mathbb{Z} : n = 3k)\} = \{3k : k \in \mathbb{Z}\}. \tag{4.3.2}$$

In words M_2 is the set of integers which are multiples of 2 (or simply even numbers) and M_3 is the set of those integers multiples of 3. An "intuitive" (but less rigorous) way to write these sets is

$$M_2 = \{\dots,-6,-4,-2,0,2,4,6,\dots\} \tag{4.3.3}$$

$$M_3 = \{\dots,-6,-3,0,3,6,\dots\}. \tag{4.3.4}$$

What is the intersection of M_2 and M_3?

From the intuitive writing we may "guess" that

$$M_2 \cap M_3 = \{\dots, -6, 0, 6, \dots\} \tag{4.3.5}$$

and that must be the set of integers that are multiples of 6. To prove this rigorously, it is sufficient to show that

$$M_2 \cap M_3 \subseteq M_6 \text{ and } M_6 \subseteq M_2 \cap M_3 \tag{4.3.6}$$

where

$$M_6 = \{n \in \mathbb{Z} : \ n \text{ is a multiple of } 6\}. \tag{4.3.7}$$

Check that this is possible.

4.3.9. Example (more multiples and intersection). Let

$$M_k := \{n \in \mathbb{Z} : \ n \text{ is a multiple of } k\}. \tag{4.3.1}$$

Let us describe concisely the set $S := M_2 \cap M_3 \cap M_4 = \bigcap_{k=2}^{4} M_k$.
The set S consists of exactly those those numbers n that are multiples of 2, 3 and 4, at the same time. Since all multiples of 4 are also multiples of 2, then S is the set of exactly those numbers n that are multiples of 4 and 3. We claim now that this n must be a multiple of 12. Indeed being a multiple of 3 means that $3 \rfloor n$, and being a multiple of 4 it means that $n = 4k$ for some $k \in \mathbb{Z}$. By a Theorem from elementary number theory which says that if $a, b \in \mathbb{Z}$ $n \in \mathbb{N}$, $\operatorname{hcf}\{a, n\} = 1$ and $n \rfloor ab$ then $n \rfloor b$, since $3 \rfloor 4k$, but $\operatorname{hcf}\{3, 4\} = 1$ we must have $3 \rfloor k$. Thus $k = 3l$ for some $l \in \mathbb{Z}$ and thus $n = 4 \times 3l = 12l$, which means that n is a multiple of 12, as claimed. So $S \subseteq M_{12}$. Conversely, $M_{12} \subseteq S$; indeed, picking $n \in M_{12}$, then $n = 12k = 4 \times (3k) = 3 \times (4k)$ which means that $n \in M_4$ (and thus in M_2) and $n \in M_3$.
So the answer is

$$M_2 \cap M_3 \cap M_4 = M_{12}. \tag{4.3.2}$$

4.3.10. Example (intersection of nested intervals). Consider the collection $\mathscr{C}$ of the following subsets of the real line

$$S_n = \left\{ x \in \mathbb{R} : \ \frac{-1}{n} < x < \frac{1}{n} \right\}, \text{ for } n \in \mathbb{N}. \tag{4.3.1}$$

(These are open intervals with endpoints $\pm 1/n$, which some strange people like 2nd year maths students denote by $(-1/n, 1/n)$.) The set of indexes of the collection $\mathscr{C}$ is $\mathbb{N}$. We ask, what is the intersection, denoted by $\bigcap_{n \in \mathbb{N}} S_n$, of this collection. Note that this collection does not have a "smallest element", in the sense that any element of $\mathscr{C}$, say S_n, strictly includes another element, e.g., S_{n+1} (make sure you understand this), so the intersection is not that easy to figure out, but it is quite plausible that

$$\bigcap_{n \in \mathbb{N}} S_n = \{0\}. \tag{4.3.2}$$

To see this equality you must prove that the set on the left equals the set on the right, by showing the inclusion on both sides. Since $0 \in S_n$ for each $n \in \mathbb{N}$, it follows that $\{0\} \subseteq \bigcap_{n \in \mathbb{N}} S_n$ and that was easy. On the other hand, suppose $y \bigcap_{n \in \mathbb{N}} S_n$, then we have

$$\forall n \in \mathbb{N} : -\frac{1}{n} < y < \frac{1}{n}. \tag{4.3.3}$$

So the only possibility left for y is $y = 0$. (Do you beleive this?[3]) Thus $\bigcap_{n\in\mathbb{N}} S_n = \{0\}$.

4.3.11. Example. This example demands some concentration. Consider the collection

$$\mathscr{D} = \{D_i : D_i \text{ is the set of all factors of } i, \text{ for } i \in \mathbb{N} \text{ and } i \geq 2\}. \tag{4.3.1}$$

For example, the following sets are elements of the collection $\mathscr{D}$

$$D_9 = \{1,3,9\}, \qquad D_{15} = \{1,3,5,15\}, \tag{4.3.2}$$

$$D_{18} = \{1,2,3,6,9,18\}, \qquad D_{19} = \{1,19\}. \tag{4.3.3}$$

Note that if i is prime then $D_i = \{1, i\}$. Suppose now that $i, j \in \mathbb{N}$ and $i \neq j$, what can you say about $D_i \cap D_j$? Try an example, and convince yourself that this must be the set of common factors of D_i and D_j, i.e., the set of numbers that divide both i and j. What can you say about $\max(D_i \cap D_j)$? Try an example. What happens if one of the i and j is prime? What if i, j are co-prime?

Now that we understand the intersection of two elements of $\mathscr{D}$, let us try three elements. Suppose $i_1, i_2, i_3 \in \mathbb{N}$, what can we say about $D_{i_1} \cap D_{i_2} \cap D_{i_3}$? Again, try some examples with different choices of i_1, i_2 and i_3 (suggestions: $i_1 = 9$, $i_2 = 15$, $i_3 = 18$, and $i_1 = 3$, $i_2 = 6$, $i_3 = 14$).

Let us try now to take the intersection of some infinite subcollection of $\mathscr{D}$, for example let us start by asking what kind of set is

$$I := \bigcap_{D\in\mathscr{D}} D = \bigcap_{i\geq 2} D_i? \tag{4.3.4}$$

The answer is $\{1\}$. The discussion above provides a heuristic explanation of this: the intersection contains those numbers which are common factors to all numbers bigger than or equal to 2. Since there is only one such factor, 1, then the answer is apparent. To nail down this rigorously, it is enough to note that $D_2 \cap D_3 = \{1\}$ and since $I \subseteq D_2 \cap D_3$ (why?) we get $I \subseteq \{1\}$. Also, since 1 is a factor of any natural number, then we have $I = \{1\}$.

Now let us try to figure out the following interesection

$$\bigcap_{i\geq k} D_i, \text{ for some fixed } k \in \mathbb{N}, k \geq 2. \tag{4.3.5}$$

Try some examples first with $k = 3, 4, 5$. Do you see a pattern? The answer is always $\{1\}$. What's funny about this example, is that none of the elements of $\mathscr{D}$ is equal to $\{1\}$.

Indeed, consider the following, a bit crazy but quite legitimate intersection

$$\bigcap_{i\in M_4} D_i, \tag{4.3.6}$$

where M_4 is the set of multiples of 4. What do you get?

[3] If you do believe this, you are intuitively using the so-called *Archimedean property of real numbers* which you shall encounter shortly in your mathematical career; if you don't believe this, you are justified in your skepticism, because it is not entirely "obvious" and the Archimedean property, once encountered and understood, will help dispell any doubts about the truth of this fact.

4.4. Difference and complementation

A set operation that is related to intersection is *set differrence*, which you should be very careful not to confuse with the difference of two numbers.

4.4.1. Definition of difference of two sets. Given two sets A, B, the between the set A and B is (also spronounced "*A take B*") is the set, denoted by $A \smallsetminus B$ defined as

$$\{x \in A : x \notin B\}. \tag{4.4.1}$$

4.4.2. Remark (complementation). In the definition in 4.4.1, we do not require that $B \subseteq A$. However, when B happens to be a subset of A then the terminology and notation changes. In this case, we call $A \smallsetminus B$ the of B with respect to A.
If A is understood from context, it is called the *universe* in which B lies, and we denote $A \smallsetminus B$ by B^{c}, or $\bar{B}$, or B'.
Note that despite the set A being called a *universe*, but there is really nothing universal about it.[4]

4.4.3. Example.
(a) Suppose that the universe is $\mathbb{Z}$. Then $\mathbb{N}^{\mathrm{c}} = \mathbb{Z} \smallsetminus \mathbb{N} = \{x \in \mathbb{Z} : x \notin \mathbb{N}\}$, which is the set of non-positive integers, $\{\dots, -3, -2, -1, 0\}$. So we have

$$-10 \in \mathbb{N}^{\mathrm{c}},\ 10 \notin \mathbb{N}^{\mathrm{c}} \text{ and } \frac{3}{2} \notin \mathbb{N}^{\mathrm{c}}. \tag{4.4.1}$$

(b) Suppose now we switch our universe to $\mathbb{R}$. Then the set $\mathbb{N}^{\mathrm{c}}$ will be different, indeed now we have

$$\frac{3}{2} \in \mathbb{N}^{\mathrm{c}}. \tag{4.4.2}$$

4.5. Do sets exist at all?

We have been living quite dangerously so far because with all these Axioms and Rules, we're still not sure there is a set at all to start with. Indeed, both Axioms 1 and 2 start from *having* one or two sets. So, to dispell all kinds of fears about building the theory of naught, we shall postulate the following.

Axiom 3 (existence)**.** *There exists a set.*

To be truly orthodox to the ZFC theory, this just a subaxiom, or a consequence of a slightly more consistent axiom which postulates the existence of an *infinite set*. But since we do not know what an infinite set is, so we will stick to this "Baby-"Axiom for now and return to it when we know what an infinite set really is.

[4] A more appropriate name for A would be the complemetation's *context*, but this is a one-man's campaign which we shall not insist upon.

4.5.1. The empty set. An *empty set* is defined a set that has no elements. An empty set must exist, in view of Axiom 2 and Axiom 3: indeed, the latter insures we have a set, say S that exists, and using specification we can build the set

$$E := \{x \in S : x \notin S\}. \tag{4.5.1}$$

An element of E, if such a thing existed, would both belong and not belong to S which is not possible, because a proposition would be true and false at the same time (and this, as you well know by now violates the laws of Logic). Therefore E has no elements. Also, two empty sets, say E and F, must be equal. Indeed, it is a two-line exercise in Logic to show that if $x \in E$ then $x \in F$. This is a so-called *vacuously true* implication because the premise is always false (which makes the implication always true). So $E \subseteq F$, and by exchanging roles, $F \subseteq E$. By Axiom 1 it follows that $E = F$. Summarising this discussion, we have proved the following result.

4.5.2. Proposition. *There exists a unique empty set.*

4.5.3. Definition of empty set. *The empty set*, defined above, is denoted by the stylised Nordic letter $\varnothing$ (and not by the Greek letters ϕ or Φ). Another way of writing the empty set is to use the curly brace notation: $\varnothing = \{\}$.

4.5.4. Definition of disjoint sets. Two sets, say A and B, are called *disjoint* if and only if $A \cap B = \varnothing$.

4.5.5. Definition of vacuous truth. Suppose P is a formula on some set X. A *vacuous truth* is a proposition of the following form

$$\forall\, x \in \varnothing : P(x). \tag{4.5.1}$$

4.5.6. Exercise. *Suppose P is a formula on some set X. Prove that a vacuous truth of the form $\forall\, x \in \varnothing : P(x)$ is a well-defined proposition which is always true.*

4.6. Union

You are probably familiar with the notion of or *join* of two sets, but let us remind ourselves of what this means.

4.6.1. Definition of union of two sets. Given two sets, say A and B, their union, denoted by $A \cup B$ is the set

$$\{x : x \in A \text{ or } x \in B\}. \tag{4.6.1}$$

There is a slight problem though with this definition in that what we are building is not guaranteed to be a set. In fact we cannot use specification unless we knew that A and B are contained in a bigger set, call it $\mathbb{U}$ (for Universe), and then we may use specification by defining

$$A \cup B = \{x \in \mathbb{U} : x \in A \text{ or } x \in B\}. \tag{4.6.2}$$

So we could require a small Axiom that postulates the existence of a Universal Set $\mathbb{U}$. But, with Russell's Paradox 3.2 looming at large, it may be not a very safe road to pursue[5] and we shall postulate a weaker fact: simply we will assume that the union of two sets is a set. In fact, we shall assume that the union of any collection of sets is a set.

Axiom 4 (union)**.** *Given a collection of sets, say $\mathscr{C} = \{X_i : i \in \mathscr{I}\}$, then there exists a set U such that*

$$x \in U \iff \exists\, i \in \mathscr{I} : x \in X_i. \tag{4.6.3}$$

The set U in 4 is unique and is denoted by $\bigcup_{i\in\mathscr{I}} X_i$, or simply by $\bigcup_{X\in\mathscr{C}} X$, and if $\mathscr{C}$ consists of only two elements, call then X and Y, then the union is denoted by $X \cup Y$.

4.6.2. Example (union).

(a) Let $A = \{1,2\}$, $B = \{2,3\}$, $C = \{4\}$. Then

$$\begin{gathered} A\cup B = \{1,2,3\}, \quad A\cap B = \{2\}, \quad A\cap C = \varnothing, \\ A\cup C = \{1,2,4\}, \quad (A\cup B)\cup C = \{1,2,3,4\} \text{ and } A\cup(B\cup C) = \{1,2,3,4\}. \end{gathered} \tag{4.6.1}$$

Notice that $(A\cup B)\cup C$ is equal to $A\cup(B\cup C)$: we shall see very soon that this is a result of the operation $\cup$ enjoying the associative property.

4.6.3. Exercise (basic properties of set theoretic operations). *Given three sets A, B and C prove only four of the following:*

(a) $A\cap B = B\cap A$ (Commutativity of the intersection).
(b) $A\cup B = B\cup A$ (Commutativity of the union).
(c) $(A\cap B)\cap C = A\cap(B\cap C)$ (Associativity of the intersection).
(d) $(A\cup B)\cup C = A\cup(B\cup C)$ (Associativity of the union).
(e) $A\cap(B\cup C) = (A\cap B)\cup(A\cap C)$ (Distributivity of the intersection over the union).
(f) $A\cup(B\cap C) = (A\cup B)\cap(A\cup C)$ (Distributivity of the union over the intersection).
(g) $A\cup\varnothing = A$ (Neutrality of the empty set for union).
(h) $A\cap\varnothing = \varnothing$ (Absorbing property of $\varnothing$ for the intersection).

4.6.4. Exercise (inclusion, intersection and union).

(a) Let A, B be two sets prove the following inclusions

$$\varnothing \subseteq A \tag{4.6.1}$$

$$A\cap B \subseteq A \tag{4.6.2}$$

$$A \subseteq A\cup B \tag{4.6.3}$$

(b) Let $\mathscr{C}$ be a collection of sets, say

$$\mathscr{C} = \{X_i : i \in \mathscr{I}\} \tag{4.6.4}$$

for a set of indexes $\mathscr{I}$. Show that for any $j \in \mathscr{I}$, the following two inclusions hold:

$$\bigcap_{i\in\mathscr{I}} X_i \subseteq X_j \subseteq \bigcup_{i\in\mathscr{I}} X_i. \tag{4.6.5}$$

(c) Make up the most interesting example, with $\mathscr{I}$ possibly infinite, that you can think of that illustrates the situation in (b).

[5] In fact, there are some Set Theories which use the concept of Universe without incurring into Russell's Paradox, but this is not a road that we are interested in pursuing here.

4.6.5. Exercise (De Morgan's laws). *Suppose $\mathbb{U}$ is a set and $A, B \subseteq \mathbb{U}$. Prove the following*

$$(A \cup B)^{\mathrm{c}} = A^{\mathrm{c}} \cap B^{\mathrm{c}} \tag{4.6.1}$$

$$(A \cap B)^{\mathrm{c}} = A^{\mathrm{c}} \cup B^{\mathrm{c}}, \tag{4.6.2}$$

where the universe for the complementation is $\mathbb{U}$.
Hint. Use the de Morgan laws for logic binary operations " and " and " or ".

4.7. Sets of Sets

There is nothing wrong with a set being an element of another set. In fact, we have been using such things with the undercover name of "collection". But collections are, intuitively at least, nothing but sets whose elements happen to be sets. What we still lack, in our list of Axioms, are mechanisms that would permit to build such sets from preexisting ones, or rules that would "allow" certain intuitive collections to be sets. We dedicate this section to some Axioms which do precisely this. We start with the simplest.

4.7.1. Pairing of sets.

Axiom 5 (pairing). *Given two sets, say U and V, then there exists a set whose elements are exactly U and V.*

Big Fat Note. The set of which in Axiom 5 can be written, using the curly brace notation, as $\{U, V\}$. You must be well aware that this set is *not the union $U \cup V$* (nor the intersection $U \cap V$ for that matter). Indeed, we have

$$U, V \in \{U, V\}, \text{ yet } U, V \notin U \cup V \text{ in general.} \tag{4.7.1}$$

Also

$$U, V \subseteq U \cup V, \text{ but } U, V \nsubseteq \{U, V\} \text{ in general.} \tag{4.7.2}$$

4.7.2. Ordered pairs. An interesting application of the Axiom of Pairing (Axiom 5), is the possibility to define rigorously ordered pairs. This is also a good exercsie in reverse engineering if you like such things.
You are probably familiar with the notion of ordered pairs from some other course. Usually an ordered pair is denoted using the round brackets (parentheses) to distinguish it from an unordered pair. For example, in analysis or geometry you may have encountered ordered pairs representing points on the Cartesian plane(also known as the xy-plane).
While sets are not "order-sensitive", ordered pairs are (which explains their name). For example, while the sets represented by $\{1,2\}$ and $\{2,1\}$ are *equal* (by the Extension Axiom 1), the ordered pairs $(1,2)$ and $(2,1)$ represent *different* points in the plane. Also the set $\{1,1\}$, is not really a pair, as it can be collapsed to $\{1\}$ (by Extension), whereas $(1,1)$ is different than (1), whatever the last thing means. The following constitutes a possible definition for ordered pairs.

4.7.3. Definition of ordered pair. Given $a \in S$ and $b \in T$, where S and T are two non-empty sets, the *ordered pair* (a,b) is a construct which satisfies the following defining property:

$$(a,b)=(c,d) \Leftrightarrow a=c \text{ and } b=d, \tag{4.7.1}$$

for all possible c and d. This definition would make sense if we could show that these "contructs" are possible. Given our pretense that we can build all mathematics from set theory, we must now show that this construct is in fact a result of set manipulations.

So suppose we are given two elements, $a \in S$ and $b \in T$, how can we build a set that represents (a,b)? In other words we want to build a set, say P, which depends only on a, b, and given c, d the corresponding set, call it Q, must satisfy (4.7.1). Recall that we can use only Axioms and consequent results proved so far.

If you give it a thought all you need to "identify" an ordered pair (a,b), is the set of its elements $\{a,b\}$ (or $\{b,a\}$ if you prefer, order doesn't—yet—matter) and to know which element comes first, that is a. Therefore the set $\{a,\{a,b\}\}$ should suffice. Except, for reasons apparent later on, this set will not have as easy properties as the seemingly more complicated set $\boldsymbol{p} := \{\{a\},\{a,b\}\}$. Similarly we build $\boldsymbol{q} := \{\{c\},\{c,d\}\}$.

To make all this work useful, we need to answer positively these two questions:

Q1. Is it possible to build such a set $\boldsymbol{p}$ using the known axioms?

Q2. Do $\boldsymbol{p}$ and $\boldsymbol{q}$ satisfy the defining property of ordered pairs (4.7.1)? I.e., is true that

$$\underbrace{\boldsymbol{p}=\boldsymbol{q}}_{L} \Leftrightarrow \underbrace{a=c \text{ and } b=d}_{R}. \tag{4.7.2}$$

The answer, of course, is yes to both questions. For Q1 let us "deconstruct" the set $\boldsymbol{p}$, it consists of two elements, $\{a\}$ and $\{a,b\}$, which are themselves sets. So if these sets are possible, $\boldsymbol{p}$ is possible in view of the Axiom of Pairing (Axiom 5). Now $\{a\}$ is possible which can be built by using the Specification Axiom 2 $\{a\} := \{x \in S : x = a\}$. Finally, to see that is $\{a,b\}$ possible, use the Union Axiom 4 and Specification Axiom 2 to get $\{a,b\} := \{x \in S \cup T : x = a \text{ or } x = b\}$.

To answer positively Q2, we must show (4.7.2) for any c and d. Since we want to prove the equivalence $L \Leftrightarrow R$, we have to do prove the two implications: $L \Leftarrow R$ and $L \Rightarrow R$.

$(L \Leftarrow R)$ This is not hard to show. Suppose R is true, then $a = c$ and $b = d$, then $\{a\}=\{c\}$ and $\{a,b\}=\{c,d\}$ and therefore $\boldsymbol{p} \subseteq \boldsymbol{q}$ and $\boldsymbol{q} \subseteq \boldsymbol{p}$ which implies $\boldsymbol{p}=\boldsymbol{q}$, i.e., L is true.

$(L \Rightarrow R)$ This implication takes a bit more work. Suppose $\boldsymbol{p}=\boldsymbol{q}$, it means that

$$\{\{a\},\{a,b\}\}=\{\{c\},\{c,d\}\}. \tag{4.7.3}$$

It follows that $\{a\}=\{c\}$ or $\{a\}=\{c,d\}$, so let us analyse what happens in each case.

Case (a) $\{a\}=\{c,d\}$ then we obtain $c=a=d$ and thus $a=c$ and $\{c,d\}=\{a\}=\{c\}$, so the set on the right hand side of (4.7.3). It follows that $\{a,b\}=\{c\}$, and thus $b=a=c=d$, from which $b=d$. (Of course, in this case arises in a very particular situation where the ordered pairs have the same entry.)

Case (b) $\{a\}=\{c\}$, so it follows $a=c$ and we still have to prove $b=d$. Again let us distinguish two (sub-) cases:

Subcase (i) $\{a,b\}=\{c\}$. In this case we will be in a similar situation as in a (but with the roles inverted) which will lead to $a=b=c=d$, thus $b=d$ as required.

Subcase (ii) $\{a, b\} = \{c, d\}$. Now it follows that $b = d$ and we are done, or that $b = c$, but then $a = c = b$ and it follows that $d = b$, so we are also done.

4.7.4. Definition of ordered tuple or finite sequence. Now that we know how to build an ordered pair, we may also define a triple: given a, b and c, three elements of three corresponding sets, then we may define the ordered *triple* $(a, b, c) := ((a, b), c)$. Indeed, let (a, b) play the role of the first coordinate and c play the role of the second coordinate in the construction of an ordered pair.
Similarly can define also an ordered *quadruple* $(a, b, c, d) := ((a, b, c), d)$, and a *quintuple*, *sextuple*, *septuple*, etc.
In fact for any $n \in \mathbb{N}$, $n \geq 3$, using recursion, given n elements, say $a_1, \dots, a_n$, then the n-*tuple* $(a_1, \dots, a_n)$ is defined as the ordered pair $((a_1, \dots, a_{n-1}), a_n)$. To make this work even for one element, we can define the *onetuple* (a) to be simply $\{a\}$ and take the 0-*tuple* to be the empty set. Another name, that you might already know, for an n-tuple is *finite sequence* of length n.

4.7.5. Remark. Note that officially we do not know what numbers are, and neither do we know what induction or recursion are, but we allow ourselves to cheat a bit in order not to postpone useful definitions. We will come back to the definition of n-tuples and finite sequences once we establish the theory of natural numbers.

4.7.6. Exercise (n-tuples equality). *Prove, by induction on n, that if $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \dots, b_n)$ are two n-tuples for some $n \in \mathbb{N}$ and $\boldsymbol{a} = \boldsymbol{b}$ then $a_i = b_i$, for all $i \in [1 \dots n]$.*

4.7.7. Exercise. *Write the quadruple (a, a, b, b) by using only curly braces. What about (a, a, a, a)?*

4.7.8. Exercise (singletons). *Recall that a singleton is a set which consists of only one element. Another funny (and surprisingly useful) application of the Pairing Axiom is the following fact: Given a set A, there exists a singleton set whose element is A, shortly denote by $\{A\}$. Prove that this is possible.*

4.7.9. Exercise (finite sets of sets). *This exercise generalises the concept of pairing to the gathering of an arbitrary, but finite*[6]*, number of sets.*
Show that given a finite number of sets $A_1, \dots, A_n$ it is possible to build a set which contains precisely all the A_i, $i = 1, \dots, n$, as its elements. I.e., show that $\{A_1, \dots, A_n\}$ is a set.
Hint. Use the result in 4.X.12 to form $\{A_i\}$, for each $i \in [1 \dots n]$, and then take their union.

[6] For the experts: to extend this to an infinite collection of sets we would need to start from an infinite collection of sets, which is tautological in the sense that we end up with the same thing we started from.

4.7.10. Example (v-NUMBERS). of the non-negative integers] We have already seen, earlier in §4.5.1 this Chapter, that the empty set exists. Using this fact, and the possibility of building sets by pairing, we now present an interesting construction, allegedly introduced by John von Neumann which leads (almost)[7] to the rigerous concept of natural number.

Let us denote the empty set (only in this example!) $\varnothing$ by . Since is a set, then, by the Axiom of Pairing we know that the singleton {} is also a set. Let us denote it with

$$\mathbb{1} := \{\}. \tag{4.7.1}$$

Note the following properties

$$\mathbb{1} = \cup\{\}, \quad \mathbb{1} \neq \text{ and } \in \mathbb{1}. \tag{4.7.2}$$

(Of course we also have $\subseteq \mathbb{1}$ but that's not very exciting news since $= \varnothing$.) Further, using the Axiom of Pairing we may define the set

$$:= \mathbb{1} \cup \{\mathbb{1}\}. \tag{4.7.3}$$

Note that reverting to the original notation we see that

$$\mathbb{1} = \{\varnothing, \{\varnothing\}\} \text{ and } = \{, \mathbb{1}\} = \{\varnothing, \{\varnothing\}\}, \tag{4.7.4}$$

which causes pain to our eyes, so we better stick the "new" notation. Some slightly more interesting facts about are

$$\mathbb{1} \in \text{ and } \mathbb{1} \subsetneq . \tag{4.7.5}$$

(The diligent reader, you, will appreciate the difference between the last two statements.) Before we stop let us define

$$:= \cup\{\} = \{, \mathbb{1}, \} \text{ and } := \cup\{\} = \{, \mathbb{1}, , \}. \tag{4.7.6}$$

It should be clear at this point how go on forever, and define , , . More generally, we may define *successor* of any set X as follows:

$$\operatorname{suc} X := X \cup \{X\}, \tag{4.7.7}$$

for each given set X (using the Axioms of Pairing and Union). And then start from and recursively generate the sequence $\mathbb{1} := \operatorname{suc}$, $:= \operatorname{suc}\mathbb{1}$, $:= \operatorname{suc}$, etc. For some reason, we may like to call this sequence of sets as *von Neumann's Union of Member and Bracket Enclosed Recursive Sequence* and abbreviate it to *v-NUMBERS.*

In fact, if we repeated "infinitely" many times the suc recursion starting from , we could obtain a set $\mathscr{N}$ which satisfies the following properties:

$$\varnothing = \in \mathscr{N}, \tag{4.7.8}$$

$$\ltimes \in \mathscr{N} \Rightarrow \operatorname{suc} \ltimes \in \mathscr{N}. \tag{4.7.9}$$

Two nontrivial issues about this construction are

(i) nothing guarantees that $\operatorname{suc} \ltimes \neq \ltimes$ for all $\ltimes$ in $\mathscr{N}$,
(ii) we do not know whether a set such as $\mathscr{N}$ is a set.

[7] To make this totally rigorous, within set theory, we need an axiom, which allows the construction of infinite sets by induction. This will be done later in the course.

Both issues turn out to need axioms to be answered.
Issue (i) is settled with the Axiom of Regularity (or Axiom of Good Foundation, or, grotesquelier, Axiom of Well-Foundedness). You should resist the temptation of proving the result by induction: this is in fact possible, but the PMI needs the existence of natural numbers, which is exactly what we are constructing, which means that such a argument would be circular (assuming what it proves).
Issue (ii) is setteled by the Axiom of Infinity (also known as Peano's Axioms), which we study in §6.4, postulates that von Neumann's construction yields a set in the framework of Set Theory. The set $\mathscr{N}$ thus obtained behaves exactly like good old $\mathbb{N}_0$ (which we know only informally so far). So this will finally settle the rigorous definition of the set of natural numbers promised earlier in the course, by showing that

$$\in \mathbb{1} \in \ \in \ \in \cdots \text{ and } \subsetneq \mathbb{1} \subsetneq \ \subsetneq \ \in \cdots \tag{4.7.10}$$

Once this is settled, we can safely exchange the notation with the "usual one"

$$\begin{gathered} 0 \leftarrow \ , 1 \leftarrow \mathbb{1}, 2 \leftarrow \ , 3 \leftarrow \\ n+1 \leftarrow \operatorname{suc} n. \end{gathered} \tag{4.7.11}$$

The laws of usual arithmetic and order of natural numbers can then be inferred.

4.7.11. Exercise. *Denote by* $, \mathbb{1}, , \ldots$ *the sets appearing in v-NUMBERS. Show that we have*

$$\mathbb{1} \subsetneq \ \textit{and}\ \mathbb{1} \in \ . \tag{4.7.1}$$

4.7.12. The power set. Suppose $a \neq b$, what are all the possible subsets of $\{a, b\}$ and what is their number?
The answer is

$$\varnothing, \{a\}, \{b\}, \{a, b\}. \tag{4.7.1}$$

Hence $\{a, b\}$ has 4 subsets in total.
The same question with the set $\{a, b, c\}$ leads to

$$\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}. \tag{4.7.2}$$

Thus $\{a, b, c\}$ has 8 subsets in total.
Based on 4.7.9, the sets in (4.7.1) and (4.7.2), can be gathered into the sets of sets

$$\{\varnothing, \{a\}, \{b\}, \{a, b\}\}, \text{ and} \tag{4.7.3}$$

$$\{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}, \tag{4.7.4}$$

respectively. These sets are known as the *power sets* of the sets $\{a, b\}$ and $\{a, b, c\}$, respectively. The reason for the name power will be apparent later on, for now it suffices to note that a set of 2 elements has a power set of 2^2 elements and a set of 3 elements has a power set of 2^3 elements.
This example should have convinced you that given a set with finitely many elements it is possible to build its power set, based on Axioms 1 through 5. However, in Mathematics we often encounter sets which may not be finite and of which we would still like to build the corresponding power set. So, to be sure, this prompts us to ask for yet another Axiom.

Axiom 6 (power)**.** *Given a set A, there exists a set, denoted $\wp(A)$, whose elements are precisely all the subsets of A.*

The set $\wp(A)$ is called the *power set* (or the *set of parts*, or the *set of subsets*) of A.

4.7.13. Example (power of $\mathbb{N}$). Consider the set . Then all the subsets of $\mathbb{N}$ are elements of this set. For example,

$$\varnothing \in \wp(\mathbb{N}), \{1,2\} \in \wp(\mathbb{N}), [3\dots 13] \in \wp(\mathbb{N}), \{2n : n \in \mathbb{N}\} \in \wp(\mathbb{N}). \tag{4.7.1}$$

Notice that since $\mathbb{N}$ is an infinite set, $\wp(\mathbb{N})$ must be an infinite set (do you see why?) and it must also contain elements which themselves are infinite sets.

4.8. Cartesian products, graphs, relations

4.8.1. Definition of cartesian product of two sets. Given two sets A and B, their *cartesian product*, $A \times B$, is defined as the set of all possible ordered pairs with first component in A and second component in B. In symbols, we may write this as

$$A \times B := \{(a,b) : a \in A \text{ and } b \in B\}. \tag{4.8.1}$$

4.8.2. Definition of cartesian square. Given a set A, the *cartesian square* of A is $A \times A$.

4.8.3. Remark (existence of the cartesian product). The cartesian product is a well defined set. Indeed, thanks to the Axiom of pairing we know that $(a,b) := \{\{a\},\{a,b\}\}$ is well defined for each $a \in A$ and each $b \in B$. Furthermore, since $\{a\}$ and $\{a,b\} \in \wp(A \cup B)$ we have that

$$(a,b) = \{\{a\},\{a,b\}\} \subseteq \wp(A \cup B), \tag{4.8.1}$$

which is equivalent to saying

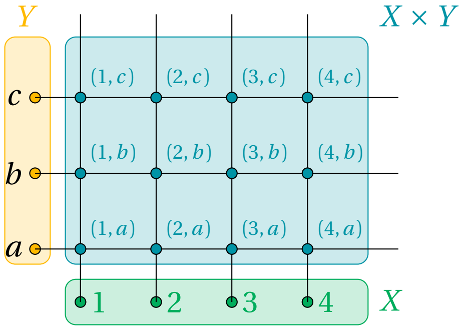
$$(a,b) \in \wp\big(\wp(A \cup B)\big). \tag{4.8.2}$$

In other words, the cartesian product must therefore be sought as a subset of $\wp\big(\wp(A \cup B)\big)$. Now from the Axiom of Union and the Axiom of Power we know that $\wp\big(\wp(A \cup B)\big)$ is a set, so by the Axiom of Specification we get that the following defines a set

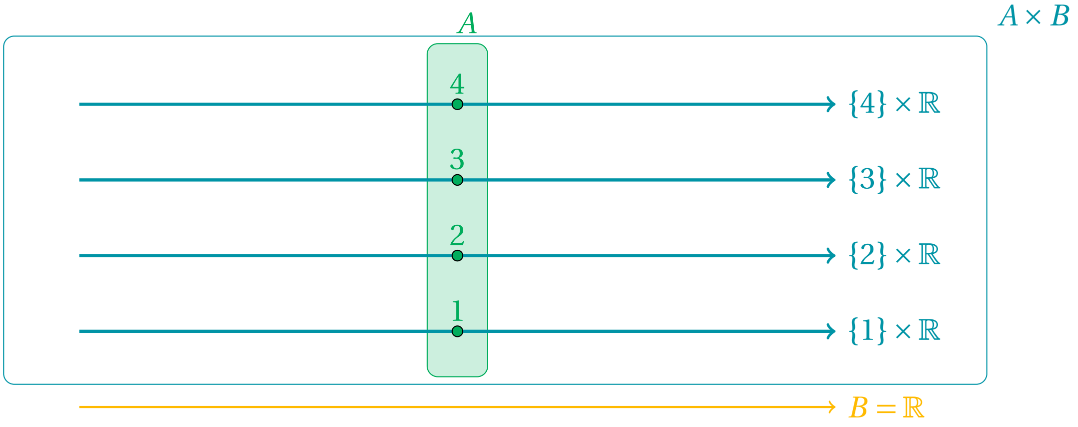
$$\Big\{P \in \wp\big(\wp(A \cup B)\big) : P = (a,b) \text{ with } a \in A, b \in B\Big\} =: A \times B. \tag{4.8.3}$$

4.8.4. Remark (graphical representation of $A \times B$). A handy way of representing the set $A \times B$ is to think of A and B be represented as two intersecting lines (or sequences of dots if discrete).

(a) For example, if $X = \{1,2,3,4\}$ and $Y = \{a,b,c\}$, then we may "draw" the cartesian product as follows



(b) Another example, with $A = \{a, b, c, d\}$ (all elements distinct) and $B = \mathbb{R}$ (the set of all real numbers), for which $A \times B$ can be sketched as follows:



(c) A typical example of cartesian product, is the familiar cartesian plane $\mathbb{R} \times \mathbb{R}$ from Geometry and Calculus.

4.8.5. Definition of (directed) graph. Given two sets A, B, a (directed) *graph* from A into B is a subset of $A \times B$. Graphs are very handy to represent relationships between elements of A and elements of B.[8]

4.8.6. Example (use of graphs). Let A be the set of students in a maths class and let B be a set of book titles. Consider the "rule"

$$x \in A \text{ has read } y \in B. \tag{4.8.1}$$

This is easily represented as a graph. To make a concise example, let the class be quite small, say

$$A = \{\text{Al}, \text{Ben}, \text{Cat}, \text{Dan}, \text{Ed}\} \tag{4.8.2}$$

and the books

$$B = \{\text{Aleph}, \text{Blue}, \text{Catch 22}, \text{Dune}\}. \tag{4.8.3}$$

Then we may summarise who read what in a *table* known as the *adjacency matrix* or *adjacency array* of the graph

	Al	Ben	Cat	Dan	Ed
Aleph	0	1	0	0	0
Blue	0	0	0	0	0
Catch 22	1	1	1	1	1
Dune	1	1	1	0	0

(4.8.4)

[8] In many mathematical fields, the word "graph" has slightly more specific meanings. For example, in mathematical analysis it is used to indicate *set-valued functions*, whereas in discrete mathematics and computer science it is used for *undirected graphs*. Here we use the word in its most general meaning of *undirected graph*, reserving the right to redefine it when specialised purposes occur.

or represent this situation as a graph

$$\begin{aligned} R := \{&(\text{Al,Catch 22}), (\text{Al,Dune}),\\ &(\text{Ben,Aleph}), (\text{Ben,Catch 22}), (\text{Ben,Dune}),\\ &(\text{Cat,Catch 22}), (\text{Cat,Dune}),\\ &(\text{Dan,Catch 22}), (\text{Ed,Catch 22})\} \end{aligned} \tag{4.8.5}$$

4.8.7. Definition of long cartesian product, cartesian power. For some integer n, let A_i be a given set for each $i = 1, \dots, n$ then for any integer $k \leq n$, we may define the *long cartesian product* of the first k sets recursively by

$$\prod_{i=1}^{k} A_i := \begin{cases} \{\varnothing\} \text{ for } k = 0, \\ \left(\prod_{i=1}^{k-1} A_i\right) \times A_k. \end{cases} \tag{4.8.1}$$

It follows that $\prod_{i=1}^{1} A_i$ is in a one-to-one correspondence with A_1 and can be thought of as A_1. In fact, for any $k = 1, \dots, n$ we have

$$\prod_{i=1}^{k} A_i \leftrightarrows \{(a_1, \dots, a_k) : \ a_i \in A_i \text{ for each } i = 1, \dots, k\}. \tag{4.8.2}$$

If for some set A, we have $A_i = A$ for all $i = 1, \dots, n$, then $\prod_{i=1}^{k} A_i$ is denoted A^k and called the k-th cartesian power of A, with the particular names for the case $k = 2$, *cartesian square*, and $k = 3$, *cartesian cube*.

Exercises and problems on sets

Exercise 4.X.1 (playing with sets). Let $A := \{1,2,3\}$, $B := \{1,2\}$, $C := \{1,3\}$, $D := \{2,3\}$, $E := \{1\}$, $F := \{2\}$, $G := \{3\}$, $H := \varnothing$.
Simplify the following expressions; in each case the answer should be one of the sets $A, \dots, H$.

(a) $A \cap B$

(b) $A \cup C$

(c) $A \cap (B \cap C)$

(d) $(C \cup A) \cap B$

(e) $A \smallsetminus B$

(f) $C \smallsetminus A$

(g) $(D \smallsetminus F) \cup (F \smallsetminus D)$

(h) $G \smallsetminus D$

(i) $A \cup ((B \smallsetminus C) \smallsetminus F)$

(j) $H \cup H$

(k) $A \cap A$

(l) $((B \cup C) \cap C) \cup H$

Exercise 4.X.2 (examples of sets). In this question, each part contains a description of three sets. In every case two of the sets are the same, and one is different. Find the set which is different. When we say two sets are the same, we mean that they are equal (in the the sense of the Extension Axiom).

(a) $A := \varnothing$,
$B := \{\}$,
$C := \{\varnothing\}$.

(b) $A := \{x \in \mathbb{Z} :\ 0 < x < 1\}$,
$B := \big\{y \in \mathbb{Z} :\ 0 < y < 1\big\}$,
$C := \{z :\ z \subseteq \{\varnothing\}\}$.

(c) $A := \{x \in \mathbb{N} :\ 1 \leq x < 8\}$,
$B := \big\{r \in \mathbb{Z} :\ 1 \leq r^2 < 64\big\}$,
$C := \big\{\zeta \in \mathbb{Z} :\ 1 \leq \zeta^3 < 512\big\}$.

(d) $A := \{\lambda \in \mathbb{Z} :\ \lambda \geq 0\}$,
$B := \mathbb{N}$,
$C := \{\nu :\ \nu \in \mathbb{N}\}$.

(e) $A := \varnothing \cup \varnothing$,
$B := \{\varnothing, \varnothing\}$,
$C := \varnothing \smallsetminus \varnothing$.

(f) Recalling that eis Napier's basis of the natural exponential $2.71 < \mathrm{e} < 2.72$, and π is the circumference of a circle of diameter 1, $3.14 < \pi < 3.15$. $A := \big\{\rho \in \mathbb{N} : \mathrm{e} < \rho < \pi\big\}$,
$B := \{\sigma \in \mathbb{N} : -\pi < \sigma < -\mathrm{e}\}$,
$C := \big\{\gamma \in \mathbb{N} : \pi < \gamma < \mathrm{e}\big\}$.

(g) $A := \big\{\beta \subseteq \mathbb{N} : \beta^{\mathrm{c}} \text{ is finite}\big\}$,
$B := \big\{\mu \subseteq \mathbb{N} : \mu \text{ is infinite}\big\}$,
$C := \{\nu \subseteq \mathbb{N} :\ \nu^{\mathrm{cc}} \text{ is infinite}\}$. Here a set X's complement, X^{c}, is taken with respect to $\mathbb{N}$.

(h) $A := \mathbb{N} \cup \{\tau : -\tau \in \mathbb{N}\}$,
$B := \mathbb{Z} \smallsetminus \{0\}$,
$C := \{\alpha : \exists A \subseteq \mathbb{N} : \alpha \in A\}$.

(i) This is not a question, it is a defintion: In the rest of this exercise we consider the notation

$$[r, s) := \{x \in \mathbb{R} :\ r \leq x < s\},$$

which defines the *closed-bottom-open-top interval* of real numbers with endpoints $r, s \in \mathbb{R}$.

(j) $A := [0,1) \cap [-1,0)$,
$B := [1,0) \cap [0,-1)$,
$C := \{0\}$.

(k) With the same notation as above:
$A := \bigcap_{\lambda > 0} [0, \lambda)$,
$B := \bigcap_{\lambda \geq 0} [0, \lambda)$,
$C := \bigcap_{\lambda < 0} [0, \lambda)$

(l) $A := \mathbb{Q} \cap \mathbb{Z}$,
$B := \mathbb{Q} \cap \mathbb{N}$,
$C := \mathbb{R} \cap \mathbb{N}$.

(m) $A := (\varnothing \cup \{\varnothing\}) \cup \{0\}$,
$B := \{\varnothing, \{\varnothing\}, 0\}$,
$C := \{\varnothing, 0\}$.

Exercise 4.X.3 (from logic to set theory). (a) Consider a set $\mathbb{U}$ and let A and B be two distinct, subsets of $\mathbb{U}$ with non-empty intersection and such that neither $A \subseteq B$ nor $B \subseteq A$. Draw a Venn Diagram describing this situation.

(b) Indicate by coloring Venn Diagrams the following sets

$$\begin{gathered} I := \{x \in \mathbb{U} : x \in A \text{ and } x \in B\}, \quad U := \{x \in \mathbb{U} : x \in A \text{ or } x \in B\} \\ S := \{x \in \mathbb{U} : x \in A \Rightarrow x \in B\} \quad N := \{x \in \mathbb{U} : x \in B \Rightarrow x \in A\} \\ E := \{x \in \mathbb{U} : x \in A \Leftrightarrow x \in B\} \quad X := \{x \in \mathbb{U} : \text{not}\, x \in A \Leftrightarrow x \in B\} \end{gathered} \tag{4.X.3.1}$$

(c) Using only $\cup$, $\cap$, $A, B, A^{\mathrm{c}}, B^{\mathrm{c}}$ (where A^{c} is the complement of A in $\mathbb{U}$, defined by $\mathbb{U} \smallsetminus A$), express the sets I, U, S, N, E, X defined in (4.X.3.1).

Exercise 4.X.4 (a non-commutative set operation). Exactly one of the four set operations $\cup$, $\cap$, $\smallsetminus$, $\triangle$ is not commutative. Say which one and give a counterexample to the commuatative law for it (i.e., an example where the commutative law fails).

Exercise 4.X.5. (a) Let $E := \{3n : n \in \mathbb{N}_0\}$, $F := \{4n : n \in \mathbb{Z}\}$ and $G := \{k + 6 : k \in E\}$. For each of the following, indicate "true" or "false":

(i) $0 \in E \cap F$,

(ii) $\varnothing = E \smallsetminus G$,

(iii) $\varnothing = G \smallsetminus E$,

(iv) $24 \in G \cup F$,

(v) $3 \in G$.

(b) With the E and F from (a), give a concise expression of $E \cap F$.

(c) Let $n \in \mathbb{N}$, given n sets A_i, for $i = [1 \dots n]$, define their union $\bigcup_{i=1}^{n} A_i$.

(d) Exhibit two subsets of $\mathbb{Z}$, call them X and Y, such that $X \cup Y$ is infinite, but $X \cap Y$ is finite.

Exercise 4.X.6. (a) Let $E := \{2n : n \in \mathbb{Z}\}$ and $F := \{7n : n \in \mathbb{Z}\}$. For each of the following statements, say whether it is true or false:

(i) $-2 \in E$,

(ii) $-2 \in F$,

(iii) $14 \in F$,

(iv) $-14 \in E$,

(v) $17 \in F$.

Give a concise expression for the intersection $E \cap F$.

(b) Given two sets A and B write a definition for their cartesian product.

(c) Give an example of two sets $X, Y \subseteq \mathbb{Z}$ such that $X \cap Y$ and $X \smallsetminus Y$ are both infinite and $Y \smallsetminus X$ is finite and non-empty.

Exercise 4.X.7 (De Morgan laws for sets). Suppose $\mathbb{U}$ is a set and $A, B \subseteq \mathbb{U}$. Prove the following

$$(A \cup B)^{\mathrm{c}} = A^{\mathrm{c}} \cap B^{\mathrm{c}} \tag{4.X.7.1}$$

$$(A \cap B)^{\mathrm{c}} = A^{\mathrm{c}} \cup B^{\mathrm{c}}, \tag{4.X.7.2}$$

where the universe for the complementation is $\mathbb{U}$.
Hint. Use the de Morgan laws for logic binary operations " and " and " or ".

Problem 4.X.8 (set operations and their basic properties). Given three sets A, B and C prove *only four* of the following:

(a) $A \cap B = B \cap A$ (Commutativity of the intersection).
(b) $A \cup B = B \cup A$ (Commutativity of the union).
(c) $(A \cap B) \cap C = A \cap (B \cap C)$ (Associativity of the intersection).
(d) $(A \cup B) \cup C = A \cup (B \cup C)$ (Associativity of the union).
(e) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Distributivity of the intersection over the union).
(f) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (Distributivity of the union over the intersection).
(g) $A \cup \varnothing = A$ (Neutrality of the empty set for union).
(h) $A \cap \varnothing = \varnothing$ (Absorbing property of $\varnothing$ for the intersection).

Problem 4.X.9 (set inclusion and operations). (a) Let A, B be two sets prove the following inclusions

$$\varnothing \subseteq A \tag{4.X.9.1}$$

$$A \cap B \subseteq A \tag{4.X.9.2}$$

$$A \subseteq A \cup B \tag{4.X.9.3}$$

(b) Let $\mathscr{C}$ be a collection of sets, say

$$\mathscr{C} = \{X_i : i \in \mathscr{I}\} \tag{4.X.9.4}$$

for a set of indexes $\mathscr{I}$. Show that for any $j \in \mathscr{I}$, the following two inclusions hold:

$$\bigcap_{i \in \mathscr{I}} X_i \subseteq X_j \subseteq \bigcup_{i \in \mathscr{I}} X_i. \tag{4.X.9.5}$$

(c) Make up the most interesting example, with $\mathscr{I}$ possibly infinite, that you can think of that illustrates the situation in (b).

Exercise 4.X.10 (some basic properties of intersection). Show that if A, B and C are given sets then

$$A \cap B = B \cap A \qquad \left(\text{commutativity of intersection}\right) \tag{4.X.10.1}$$

$$(A \cap B) \cap C = A \cap (B \cap C) \qquad \left(\text{associativity of intersection}\right) \tag{4.X.10.2}$$

$$A \cap B \subseteq A \qquad \left(\text{minimality of intersection}\right) \tag{4.X.10.3}$$

$$A \subseteq B \iff A \cap B = A \qquad \left(\text{monotonicity of intersection}\right) \tag{4.X.10.4}$$

$$A \cap \varnothing = \varnothing \qquad \left(\text{absorbing element for intersection}\right). \tag{4.X.10.5}$$

Exercise 4.X.11. Show that the intersection of a collection $\mathscr{C}$, as defined by (4.3.3) is the same if we exchanged the set X_0 with some other set, say $X_1 \in \mathscr{C}$. We refer to this fact by saying that the definition of intersection (4.3.3) is *independent of*, or *invariant with respect to, the choice of the "intersection base" set* X_0.

Exercise 4.X.12 (singletons). Recall that a *singleton* is a set which consists of only one element. Another funny (and surprisingly useful) application of the Pairing Axiom is the following fact: *Given a set A, there exists a singleton set whose element is A, shortly denote by* $\{A\}$. Prove that this is possible.

Exercise 4.X.13 (equality of tuples). Prove, by induction on n, that if $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \dots, b_n)$ are two n-tuples for some $n \in \mathbb{N}$ and $\boldsymbol{a} = \boldsymbol{b}$ then $a_i = b_i$, for all $i \in [1 \dots n]$.

CHAPTER 5

Counting finite sets

> It was an allegory—speaking at the highest level—of how scandalously, how outrageously a meaning can take up residence in a system without becoming a term in it. Did you notice how, whenever I tried to pin you down you slipped away? I noticed.
>
> — J.M. Coetzee *Life & Times of Michael K*

5.1. An intuitive notion of counting

We take now a break from all those axioms and, instead, play an ancient game practised (at least) since shepherds (and sheep) appeared on earth. That is *counting*.[1] [2] Before we begin with anything rigorous, let us pay some thought on the *process of counting*. Suppose you have a herd of sheep and you want to count them, then you must make sure that

(C1) you don't count anyone more than once,
(C2) and you don't miss anyone out.

So, to be completely certain, you grab some mud (the shepherd's equivalent of spray paint) and you go through the sheep, and mark on each sheep a number starting from 1, then 2, 3, etc. Of course, you use each number exactly once, if you encounter a sheep that is already marked you ignore it and you stop when there are no more sheep left to mark. If you are a real-life shepherd, your herd is finite and sooner or later you will stop counting (an advantage of being a poor shepherd is that you stop counting quite soon and you have more time to do more entertaining things). Assuming you don't fall asleep during the process, you will end up with some number $n \geq 1$. You declare then this number to be the number of your sheep.

5.1.1. Translating counting to maths: cardinality. From a mathematical view-point, all you have a done is *putting your herd of sheep in a one-to-one correspondence with the set* $[1 \dots n]$. I.e., you have built a "rule" (or a "table") that associates to each natural number between 1 and n exactly one sheep (element) taken from the herd (set to be counted). There are no infinite herds in real-life, but in mathematics there are infinite sets, so counting this way will work only for *finite sets*. Once we formalise the concept of function we will be able to give a rigorous definition of finite and infinite sets, but for the moment let us rely on the following intuitive characterisation: a set S is called

[1] Anthropologists and linguists have shown that isolated tribes that live from hunting and gathering only have a very limited amount of numbers (up to 10), whereas populations that hold cattle, and other animals, have a much more developed sense of numbers and counting. Traces of counting can be found in early vocabulary which is common to many languages.

[2] However, a stunning discovery was published in 2006, by Daniel Everett, who discovered that the Pirahã people in the Amazon have no words to describe *number* nor *time*.

finite if there exists a number $n \in \mathbb{N}_0$ such that $[1 \ldots n]$ can be put in a one-to-one correspondence with S.[3] In other words, there exists a rule

$$q : [1 \ldots n] \leftrightarrows S, \tag{5.1.1}$$

such that

$$q(k) = q(j) \Rightarrow k = j, \tag{5.1.2}$$

and

$$\forall\, x \in S : \exists\, k \in [1 \ldots n] : q(k) = x. \tag{5.1.3}$$

The proposition in (5.1.2) ensures condition (C1), whereas (5.1.3) ensures condition (C2).
In contrast, a non-empty set that cannot be put in a one-to-one correspondence with a segment of the natural numbers $[1 \ldots n]$, for any $n \in \mathbb{N}$ is called *infinite*. For example, $\mathbb{N}$ itself cannot be put in a one-to-one correspondence with $[1 \ldots n]$, for any $n \in \mathbb{N}$ and it is hence infinite.
In §

5.1.2. The intuition behind counting, cardinality, finite and infinite sets. If A is a finite set, its *cardinality* as the number of its elements. This number is usually denoted by $\operatorname{card} A$, or $\#A$, or $|A|$. More precisely, the cardinality of a finite set A is the number n for which A and $[1 \ldots n]$ are in a one-to-one correspondence.
In simple words, when A is finite, the *cardinality* of A is just a highly-sounding name for the "numbers of elements in A" and in order to *count* them we need to have an unambiguous way to associate each number from 1 to n with a unique element of A. When this process terminates, we call the set A *finite*, otherwise it is *infinite*.
When A is infinite, then its *cardinality* is not a natural number and is declared to be *infinity* also denoted ∞. Thus $\#A = \infty$, simply means that the set is not finite.[4]

Big Fat Note. Our definition of cardinality 5.1.2 is not (yet) completely rigorous, because we rely only on the intuitive notions of a *number* and a *one-to-one correspondence*. In fact, we still do not "know" what finite and infinite really mean from a rigorous view-point. This issue will be settled when we introduce the notion of function and natural numbers. Meanwhile, we live with intuition rather than rigour.
Note also, once and for all, that

∞ is not a natural number, nor a real number nor a complex number!

This said, there are mathematical theories of infinite numbers, e.g., *cardinal numbers* and *ordinals*, but we shall not deal with this theory until Chapter 9. Suffices it to say here that natural numbers constitute a tiny subset of the very, very large set of cardinal numbers and the set of ordinals.

[3] Note that the segment $[1 \ldots 0]$ is the empty set. Therefore, according to our definition, the empty set is a finite set.

[4] In fact, in proper theories of inifinite numbers, one sees that there are (infinitely!) many infinite cardinal numbers, e.g., $\aleph_0$, $\mathfrak{c}$, $\aleph_1$

5.1.3. Example (cardinality of even natural numbers smaller than 101**).** Let us use the definition of cardinality to "count" the elements of some sets. Let E be the set of all even natural numbers, i.e., $E = \{2, 4, 6, \dots\}$, and let $F = \{n \in E : 1 \leq n \leq 100\}$. We ask what is the cardinality of (or how many element are in) F. One way to do this is to line up all the numbers of F and associate a natural number starting from 1 as follows:

$$\begin{array}{cccccccccc} & 2 & 4 & 6 & \cdots & i & \cdots & 98 & 100 \\ R & \downarrow & \downarrow & \downarrow & \cdots & \downarrow & \cdots & \downarrow & \downarrow \\ & 1 & 2 & 3 & \cdots & i/2 & \cdots & 49 & 50 \end{array} \tag{5.1.1}$$

So the cardinality must be 50. To be completely sure, we must check that the "rule" R which transforms each $i \in F$ into $i/2$ is a "one-to-one correspondence" and that it transforms F exactly into $[1 \dots 50]$(this fact may be apparent, but it is worth checking it for the sake of exercise). We write $R(i)$ for the transformed element.

For each $i \in F$ since $2 \leq i \leq 100$ it must be $1 \leq i/2 \leq 100/2 = 50$, thus $R(i) \in F$. This means that the set F is transformed into a subset of $[1 \dots 50]$, so R is a correspondence from F into $[1 \dots 50]$.

We still have to check that R is a one-to-one correspondence. So have to check two things:

(a) R covers all of $[1 \dots 50]$. Indeed, let $j \in [1 \dots 50]$, define $i = 2j$. Then $i \in F$ (because j is even and $1 \leq 2 \times 1 \leq 2j \leq 2 \times 50 = 100$).

(b) R cannot transform two different elements into the same one. Indeed, suppose $R(i) = R(k)$, for some $i, k \in F$ then $2i = 2k$ which implies, $i = k$.

5.1.4. Exercise (cardinality of the power set). *We want to show that for any finite set A we have*

$$\#\wp(A) = 2^{\#A}, \tag{5.1.1}$$

where # denotes the cardinality of a set.

(a) First make some examples with $\#A = 0, 1, 2, 3$ to make sure you understand the purpose.

(b) Now produce a "cherry picking" argument that shows the result in an intuitive way.

(c) Finally, write down a rigorous proof by induction on #A.

5.2. Functions and maps

We aim at making a bit more rigorous the concept of one-to-one correspondence. For this we need to introduce *functions* (also known as *maps* or *mappings*) which consitute the object of this section.

5.2.1. The intuitive idea of a function. Given two sets, call them A and B, we have seen that it is very useful to consider *unary operators*, or *functions*, or *maps*, or *transformations*, or *correspondences* that take an *input* (or *independent variable*, or *operand*, or *argument*) in A and returns an *output* (or *image*, or *value*, or *outcome*, or *realisation*) in B.

For a “rule” ϕ to be a function, for each given argument, say $a \in A$, the image, say $b \in B$ has to *exist* and be *unique*. In symbols we may write this as

$$\underbrace{\big(\forall\, a \in A : \exists\, b \in B : b = \phi(a)\big)}_{\text{existence of image}}, \text{ and } \underbrace{\big(c = \phi(a) \text{ and } b = \phi(a) \Rightarrow\ c = b\big)}_{\text{uniqueness of image}}. \qquad (5.2.1)$$

If this is the case we write, somewhat more shortly,

$$\begin{array}{rccl} \phi: & A & \to & B \\ & a & \mapsto & \phi(a) = \boxed{\qquad\qquad\qquad \texttt{[01]}} \end{array} \qquad (5.2.2)$$

where the box (i) is filled by some kind of formula or sentence. This is also writen as

$$\text{let } \phi(a) = \boxed{\qquad\qquad\qquad \texttt{[02]}} \in B, \text{ for each } a \in A. \qquad (5.2.3)$$

When the formula defining the function is undertstood (or unknown, or generic) we often abbreviate this notation to $\phi : A \to B$.
The set A on which the function ϕ *acts* is called the *domain* of ϕ. The set B where the function ϕ takes its values is called the *codomain* (or *range*) of ϕ.

5.2.2. Example (algebraic functions). Many functions of numbers (integers, rational, real or complex) can be constructed by using arithmetic and algebraic operations such as +, −, ×, / and powers.

(a) Let

$$\psi(a) := 3a + 4, \text{ for } a \in \mathbb{N}. \qquad (5.2.1)$$

This defines a function $\psi : \mathbb{N} \to \mathbb{N}$. Indeed, for each $a \in \mathbb{N}$ we have that $3a + 4 \in \mathbb{N}$ and this outcome is unique (by construction).

(b) The rule

$$\begin{array}{rccl} g: & \mathbb{Z} & \to & \mathbb{Q} \\ & x & \mapsto & g(x) := x/3 \end{array} \qquad (5.2.2)$$

is a function. Indeed, for each $x \in \mathbb{N}$, the fraction $x/3$ defines a unique rational number.

(c) The expression

$$h(\xi) = \frac{1}{\xi}, \text{ for } \xi \in \mathbb{R}^+ \qquad (5.2.3)$$

defines a function with values in $\mathbb{R}$. (Note how we carefully avoid having 0 in the domain of h.)

(d) Consider the rule

$$\begin{array}{rccl} f: & \mathbb{Z} & \to & \mathbb{N}_0 \\ & n & \mapsto & f(n) := n^2 \end{array} . \qquad (5.2.4)$$

Here f is a function. Indeed, for each $n \in \mathbb{Z}$, n^2 is well defined, it is unique and it belongs to $\mathbb{N}_0$.
All the examples of functions seen so far involve some kind of algebraic construction to define them, but sometimes functions can be defined in a very abstract way. You should get used to the fact that functions need not be algebraic.

Big Fat Note (domain and codomain)**.** It is very, very, very, very[5] important, when talking about a function, to keep track of its domain and codomain. We emphasise

[5] , very

this fact because in many instances, in the rush of our mathematical thoughts, we are "tempted" to forget about A and B by focusing only on the rule that defines the function (which is usually the most intriguing part). There is nothing wrong with this forgetfulness, *as long as* when you are done with the maths brainstorming, you go back to what you have written and make sure that you are really handled a function. A typically abused "function" is the square root, as we shall see next.

5.2.3. Example. Consider the function s defined by

$$y = s(x) \Leftrightarrow y^2 = x. \tag{5.2.1}$$

Is this a function?

In a basic analysis course, one of the first interesting results is to show that for each $x \in \mathbb{R}_{0+} := \{x \in \mathbb{R} : \ x \geq 0\}$ there exists $y \in \mathbb{R}$ such that $y^2 = x$. (And in some advanced algebra course you may even prove that for each $x \in \mathbb{R}$ there exists $y \in \mathbb{C}$ such that $y^2 = x$.) These look like very good grounds to say that the s as defined by (5.2.1) is a function.

We have to be careful with the domain of s. We cannot take it to be $\mathbb{R}$, because if $x < 0$ there is no number $y \in \mathbb{R}$ such that $y^2 = x$. Of course, if you know $\mathbb{C}$ then you can go and look for a y therein. So you could define either

$$\begin{array}{rccl} s: & \mathbb{R} & \to & \mathbb{C} \\ & x & \mapsto & y : y^2 = x \end{array} \tag{5.2.2}$$

or

$$\begin{array}{rccl} s: & \mathbb{R}_{0+} & \to & \mathbb{R} \\ & x & \mapsto & y : y^2 = x \end{array} . \tag{5.2.3}$$

So we have managed to define the square root as a function. Or did we?

In fact, we did not. We have successfully solved (assuming we know Analysis and Advanced Algebra courses) the problem of the *existence* of a square root y to each x in the domain of s, but we forgot about uniqueness. *And this is very dangerous*, because if we do not specify something more about y we may end up saying something really troublesome. Indeed, supposing s is a function, we "know" that

$$a = b \Rightarrow s(a) = s(b). \tag{5.2.4}$$

Taking $a = 4 = b$ we get $s(a) = 2$, but also $s(b) = -2$, and thus

$$2 = -2, \tag{5.2.5}$$

which of course is absurd. We got into trouble because we were careless about uniqueness of the values of the rule s. As it stands the "rule" s is *not a function* because it may be multiply-valued.

In fact, to make the rule s appearing in (5.2.3) (forget $\mathbb{C}$ for now) a function we have to either exclude $y < 0$ from the definition or (equivalently) chop away some of its codomain. For examplw, we may define

$$\begin{array}{rccl} s: & \mathbb{R}_{0+} & \to & \mathbb{R} \\ & x & \mapsto & y : y \geq 0 \text{ and } y^2 = x \end{array} \quad \text{or} \quad \begin{array}{rccl} s: & \mathbb{R}_{0+} & \to & \mathbb{R}_{0+} \\ & x & \mapsto & y : y^2 = x \end{array} . \tag{5.2.6}$$

Now s is a well-defined function and we can really sleep tight at night, without the nightmarish prospect of risking $2 = -2$.

That is the reason why in the definition of square root in the begining of the course, we have stressed the fact that it must be non-negative.

5.2.4. Definition of identity. Given any set D we define the *identity* on D as being the function that returns the input as an output. (Sounds boring, but the identity turns out to be quite useful, in fact as useful as 0 is for + and 1 for ×, as we shall see.) The identity on a set D is denoted by id_D, meaning

$$\begin{array}{rccl} \mathrm{id}_D : & D & \to & D \\ & x & \mapsto & \mathrm{id}_D(x) := x \end{array} \quad . \tag{5.2.1}$$

When the set in question is really clear from the context we write id instead of id_D; this is sloppy though and we shall refrain from such sloppiness.[6]

5.3. Distinguished functions

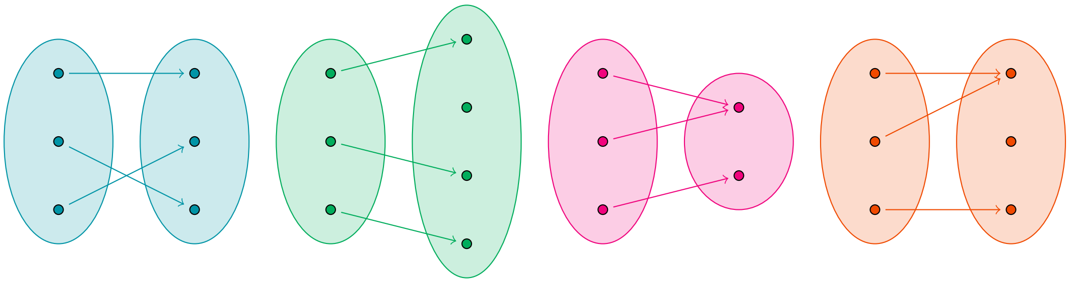
Some functions have nicer behaviour than other functions. Although a function associates to each argument one and only one (i.e., exactly one) value, it may fail to do the "inverse". I.e., for the function $f : X \to Y$, and element y in the codomain Y may not be the image of any element x in the domain X. Likewise, it is possible for an element $y \in Y$ to be the image of two different elements $x_1, x_2 \in X$. Distinguished functions are functions that exclude one (or both) of these possibilities. One-to-one correspondences (also known as bijections) are the "nicest" possible functions, in the sense that they allow finite sets to be "counted". We study distinguished functions in this section.

5.3.1. Definition of surjections/surjective functions/onto maps. We say that the function $f : X \to Y$ is *surjective*, or that it is a *surjection*, or that it *maps* X *onto* Y if and only if each element of the codomain, $y \in Y$, is the image of some element in the domain $x \in X$.
In symbols the condition for being a surjection can be written as

$$\forall\, y \in Y : \exists\, x \in X : f(x) = y. \tag{5.3.1}$$

Intuitively, this means that all elements of Y are "targetted" by some element of X, which explains the term *onto*.

5.3.2. Example (surjections and Venn Diagrams). The following are Venn Diagrams depicting surjective and nonsurjective functions between finite sets



surjective; not surjective; surjective; not surjective;

[6] Just as there is no set of all sets, there is no *universal identity function* which returns the input for *any* input. We will not explain this in this course, but from Russels Paradox we are warned about sloppiness towards the concept of sets.

Big Fat Note. Though Venn Diagrams are useful in transmitting the most important feature regarding surjections, do not get attached too much to them, because they are hopeless tools regarding situations with infinite set (unless you really know what you are talking about, which means that you should not be reading this stuff).

5.3.3. Example (surjection with infinite domain). The function $R : \mathbb{Z} \to [0 \dots 4]$ defined by

$$R(n) = r \text{ where } (q, r) = \operatorname{div}(n, 5), \tag{5.3.1}$$

is surjective. Note that the same definition with codomain $\mathbb{N}_0$ instead of $[0 \dots 4]$ would not be surjective. It is therefore very important to specifiy the codomain when talking about a function's properties.

5.3.4. Example (surjection with inifinte domain and codomain). The function

$$\begin{array}{rccl} \chi : & \mathbb{N} & \to & \mathbb{N} \\ & n & \mapsto & \text{number of divisors of } n \end{array} \tag{5.3.1}$$

is surjective. Indeed, χ is a function, the number of divisors of a number $n \in \mathbb{N}$ is clearly unique and it is a positive integer. For example, $\chi(1) = 1$, $\chi(2) = 2$, $\chi(3) = 2$, $\chi(4) = 3$, $\chi(5) = 2$, $\chi(6) = 4$, etc.
Also, given any $k \in \mathbb{N}$, then the divisors of 2^{k-1} are exactly k. Indeed, if $d \rfloor 2^{k-1}$, then by the Fundamental Theorem of Arithmetic, it must be of the form 2^j with $j \in [0 \dots k-]$ which means that all the possible divisors are

$$1, 2, 2^2 = 4, \dots, 2^{k-2}, 2^{k-1}. \tag{5.3.2}$$

Note that the same argument works with 2 replaced by 3, or any prime number, for that matter.

5.3.5. Example (surjective and non-surjective polynomial maps on $\mathbb{R}$). Consider the functions k and l given by

$$k(x) = x^3 \text{ and } l(x) = x^2 + 1, \text{ for } x \in \mathbb{R}, \tag{5.3.1}$$

with codomain $\mathbb{R}$. Then the function k is surjective, while l is not. Indeed, it can be shown (Analysis G5085) that for all $\beta \in \mathbb{R}$ there exists $\alpha \in \mathbb{R}$ such that $\alpha^3 = \beta$, so k is surjective.
As for l, note that for $x \in \mathbb{R}$, $x^2 \geq 0$ and thus $l(x) \geq 1$. So for $y = 0$ there are no $x \in \mathbb{R}$ such that $l(x) = y$.

5.3.6. Definition of polynomials and polynomials maps. You may be wondering what a *polynomial map* is, and you will be excused if you do not know. In fact, it would be surprising if you did, because no high-school teacher dares telling you what a *polynomial* is and they usually mislead you into thinking that a polynomial and polynomial map are the same thing: *they are not!*

So a *polynomial*[7] is a (finite) sequence of algebraic operations involving only $+$, $\times$ and their inverses. For example, $1+3X^2$ can be seen as the sequence of operations:

$$X \mapsto X^2 \mapsto 3X^2 \mapsto 3X^2+1. \tag{5.3.1}$$

What is intersting about a polynomial is that it does not care about *what X really is*, as long as X can be added and multiplied. For this reason X is called the *indeterminate* (which you should not be confusing with "unknown"!).

Polynomial maps arise from a polynomial when we want to settle for a certain set where to pick the X. For example, we could decide to form a polynomial map out of the polynom (oops!) $1+3X^2$ over $\mathbb{R}$. Since sum and product are possible on $\mathbb{R}$ we obtain the polynomial map

$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & 1+3x^2 \end{array} . \tag{5.3.2}$$

But X itself could have been anything that can be summed, squared and multiplied. For example, we could have chosen X to be a 2×2 matrix in $\mathbb{R}^{2\times 2}$ (ask your Geometry tutor about matrixes please!) and that would have provided us with *another polynomial map*

$$\begin{array}{ccc} \mathbb{R}^{2\times 2} & \to & \mathbb{R}^{2\times 2} \\ \boldsymbol{A} & \mapsto & \mathbf{I}+3\boldsymbol{A}\boldsymbol{A} \end{array} . \tag{5.3.3}$$

Also

$$\begin{array}{ccc} \mathbb{Z} & \to & \mathbb{Z} \\ n & \mapsto & 1+3n^2 \end{array} , \tag{5.3.4}$$

is another polynomial map.

The bottom line is:

Polynom(ial): is a rule given by a finite number of algebraic (sum, multiply) operations.

Polynomial map: is a function (rule on a domain with values in a codomain) whose rule is given by a polynom(ial).

5.3.7. Example (a warning about polynomial maps on $\mathbb{Z}$). The function

$$\begin{array}{cccc} \tilde{k}: & \mathbb{Z} & \to & \mathbb{Z} \\ & n & \mapsto & \tilde{k}(n) := n^3 \end{array} \tag{5.3.1}$$

is *not surjective*. Indeed, the element $2 \in \mathbb{Z}$ has no "counterimage" $n \in \mathbb{Z}$, i.e., for any $n \in \mathbb{Z}$, $n^3 = 2$. Indeed, the cubic root of 2, if it exists, cannot be a rational number, and it is thus not an integer. Note how $\tilde{k}$ and k from the previous example have the same "rule", yet one is surjective and the other one is not. It is quite important to know what the domain of the funcion is.[8]

[7] Polynomial is another treacherous misnomer of Mathematics, it should have been called "polynom". In fact, "poly" means "many" in Greek and the verb *nomeon* means "to reckon", so polynom means "many calculations". "Polynomial" is an adjective, which has been morphed into a noun with time by shortening the expression "polynomial calculation", which is already redundant. Unfortunately, this has the side effect of making people think that polynomial is a shortening of "polynomial map", which is wrong of course. Maybe we should be using "manireckonig" instead?

[8] You may be familiar with the so-called "graph test" (Intermediate Value Theorem) to see if a continuous function $f : \mathbb{R} \to \mathbb{R}$ is surjective. You have to be careful though that this test does not apply when the domain that is not $\mathbb{R}$ or an interal thereof.

5.3.8. Definition of injections/injective functions/one-to-one maps. We say that the function $f : X \longrightarrow Y$ is *injective*, or that it is a *injection*, or that it maps X *one-to-one into* Y if each element of the codomain $y \in Y$ can be the image of no more than one element in the domain $x \in X$. This is equivalent to say that *any* two different elements, say $x_1 \neq x_2$ in X, map to two different elements $f(x_1) \neq f(x_2) \in Y$. In symbols, we may write this as

$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2. \tag{5.3.1}$$

5.3.9. Remark (where is the quantifier in injectivity?) Some care is needed in this statement, because we are understating the fact that $x_1, x_2 \in X$ here. In practice, if you tend to “forget” about the domain of the function this may lead to trouble, so, to be emphatic, let us write this statement in the somewhat redundant form:

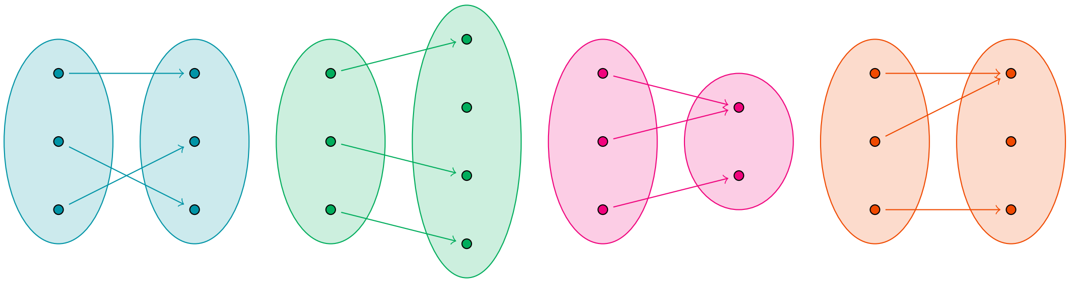
$$\forall\, x_1, x_2 \in X : f(x_1) = f(x_2) \Rightarrow x_1 = x_2. \tag{5.3.1}$$

5.3.10. Remark (injectivity via the contrapositive). At this point of the course, you should have been manipulating contraposition in logic enough times as to recognise that proposition (5.3.1) is equivalent to say

$$\forall\, x_1, x_2 \in X : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2). \tag{5.3.1}$$

(Here the quantifier is *really* needed, because $(x_1 \neq x_2)$, unlike $\big(f(x_1) \neq f(x_2)\big)$, does not understate the fact that $x_1, x_2 \in X$.) This contrapositive definition of injectivity is very often used in practice, to show that a map is injective.

5.3.11. Example (injections and Venn diagrams). The following are Venn diagrams of injective and noninjective functions.



injective; injective; not injective; not injective;

As usual, remember that Venn diagrams and infinite sets do not mix too well. It is better to resort to more rigorous methods for identifying injective maps from or to an infinite set.

5.3.12. Example (some simple injections on numbers). Consider the set of even natural numbers (positive integers) E, then the map

$$\begin{array}{rccc} d: & E & \longrightarrow & \mathbb{N} \\ & n & \longmapsto & n/2 \end{array} \tag{5.3.1}$$

is an injective map. But also the map

$$\begin{array}{rccc} e: & E & \longrightarrow & \mathbb{N} \\ & n & \longmapsto & 3n \end{array} \tag{5.3.2}$$

is injective. What is funny about e is that it is not surjective, whereas d is also surjective.
Can you find an injective map $f:\mathbb{N}\to E$?

5.3.13. Example (polynomial maps on $\mathbb{R}$). We know the set of real numbers only intuitivel so far (Analysis makes $\mathbb{R}$ a rigorous topic), but we assume that the reader is familiar with the fact that for each $n\in\mathbb{N}$ and each non-negative real number $s\in\mathbb{R}_{0+}$ there exists a number $r\in\mathbb{R}_{0+}$, such that

$$r^n = s. \tag{5.3.1}$$

Note that this fact is not true if we look for r in $\mathbb{Q}_0^+$, even for $s\in\mathbb{Q}_0^+$, instead of $\mathbb{R}_{0+}$ (think of $s=2$ and $n=2$). It is therefore *essential* to work in $\mathbb{R}_{0+}$.
Consider the following polynomial maps:

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^2 \\ g: & \mathbb{R}^+ & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^2 \\ h: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & t & \longmapsto & t^3 \\ k: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & y & \longmapsto & y^3-y \end{array} \tag{5.3.2}$$

We ask which are injective and which are not.
The function f is not injective. Indeed, we have

$$f(1)=1^2=1=(-1)^2=f(-1), \tag{5.3.3}$$

but (of course) $1\neq -1$ and the injectivity condition (5.3.1) is thus violated.
The function g is injective. Note that, in spite of being defined by the same polynomial as f, the domains of f and g are different. Suppose that $u,v\in\mathbb{R}^+$ are such that

$$g(u)=g(v), \tag{5.3.4}$$

then

$$u^2=v^2, \tag{5.3.5}$$

which (by the basic rules of algebra, cf. 1.2.10) implies that

$$u=v \text{ or } u=-v. \tag{5.3.6}$$

But since $u,v>0$, the second equality is impossible and it follows that $u=v$. We have thus proved that the injectivity condition (5.3.1) holds true for g.
The function h is injective. Indeed, suppose that $t^3=s^3$, we want to show that $t=s$. For any $t,s\in\mathbb{R}$, we have that

$$t^3=s^3 \Rightarrow 0=t^3-s^3=(t-s)(t^2+ts+s^2). \tag{5.3.7}$$

The either $t = s$, and we are done, or $t^2 + ts + s^2 = 0$. We now show that the second equality cannot be true, unless $s = t = 0$, in which case we are done too. Indeed, for any $t, s \in \mathbb{R}$, we have

$$0 \le (t+s)^2 = t^2 + 2st + s^2 \tag{5.3.8}$$

and, by manipulating this a bit, we get

$$-st \le \frac{t^2+s^2}{2}. \tag{5.3.9}$$

But $(t^2 + s^2)/2 > 0$ when one (or both) of s, t is non-zero. Thus, adding this to the right-hand side of (5.3.9) we obtain

$$-st < t^2 + s^2, \text{ i.e., } t^2 + st + s^2 > 0. \tag{5.3.10}$$

The function k is not injective. Indeed for $x = 1$ and $y = -1$ we have

$$x \ne y \text{ and } k(x) = 1^3 - 1 = 0 = -1 + 1 = (-1)^3 + 1 = k(y). \tag{5.3.11}$$

5.3.14. Example (an injective set-to-set mapping). Let S be a non-empty set. Let $s_0 \in S$ and denote by $S^* := S \setminus \{s_0\}$. Consider now the mapping

$$\begin{array}{rccl} \psi : & \wp(S^*) & \to & \wp(S) \\ & X & \mapsto & \psi(X) := X \cup \{s_0\} \end{array} . \tag{5.3.1}$$

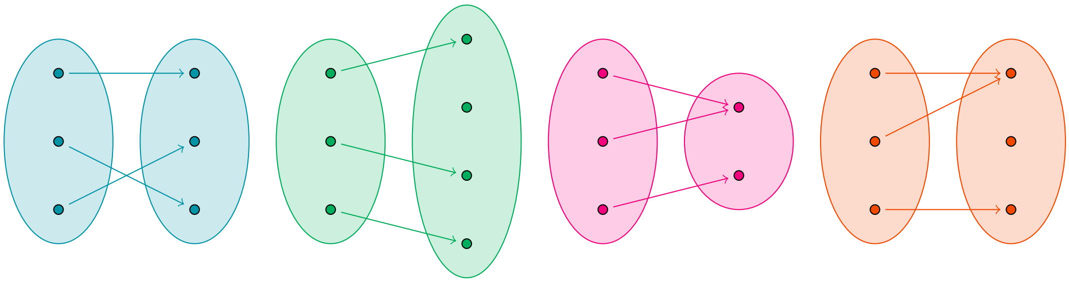
We will see that ψ is injective.
Indeed, let $X, Y \in \wp(S^*)$, i.e., $X, Y \subseteq S^*$, and suppose that $\psi(X) = \psi(Y)$, we want to show that $X = Y$. Since X and Y are sets, it is enough to show that $X \subseteq Y$ and, viceversa, that $Y \subseteq X$. The situation being perferctly symmetric, it is enough to show one inclusion, as the other inclusion is totally analogous. Noting that

$$\begin{aligned} x \in X &\Rightarrow x \in \psi(X) \qquad (\text{by definition } X \subseteq X \cup \{s_0\}\psi(X)) \\ &\Rightarrow x \in \psi(Y) \qquad (\text{from } \psi(X) = \psi(Y)) \\ &\Rightarrow x \in Y \text{ or } x = s_0 \qquad (\text{by definition of } \psi(Y)) \\ &\Rightarrow x \in Y \qquad (\text{since } x = s_0 \text{ is ruled out by } s_0 \notin X \ni x), \end{aligned} \tag{5.3.2}$$

if follows that $X \subseteq Y$, as desired.

5.3.15. Definition of bijective function/bijections/one-to-one correspondence. A function is called *bijective* (also known as bijection or one-to-one correspondence) if and only if it is simultaneously injective and surjective.

5.3.16. Example (bijections and Venn diagrams). The following picture has bijective and non-bijective functions.



surjective, injective, bijective;

not surjective, injective, not bijective;

surjective, not injective, not bijective;

not surjective, not injective, not bijective;

5.3.17. Example (identity). Given a set S, the identity function on S,

$$\begin{array}{rccc} \mathrm{id}: & S & \longrightarrow & S \\ & x & \longmapsto & x \end{array}, \tag{5.3.1}$$

is a bijection.
The proof of this is trivial.
A seemigly silly, but important property is the fact that

$$\#[1\ldots n] = n,\ \forall\, n \in \mathbb{N}_0. \tag{5.3.2}$$

Indeed, recalling Definition 5.1.2, and the fact that $\mathrm{id}: [1\ldots n] \to [1\ldots n]$ is a bijection, the conclusion is clear.

5.3.18. Example (bijective and non-bijective polynomial maps). The function

$$\begin{array}{rccc} q: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^3 \end{array}, \tag{5.3.1}$$

[∗]: Check! is bijective.[∗]
The restriction of q to $\mathbb{Z}$,

$$\begin{array}{rccc} k: & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & x & \longmapsto & x^3 \end{array}, \tag{5.3.2}$$

is *not bijective*. Though injective, the map k is not surjective: e.g., $2 \in \mathbb{Z}$ has no counterimage $x \in \mathbb{Z}$ such that $k(x) = 2$.
The function

$$\begin{array}{rccc} s: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^2 \end{array}, \tag{5.3.3}$$

[∗]: Check! is not bijective.[∗]
The restriction of s to $\mathbb{R}_{0+}$,

$$\begin{array}{rccc} r: & \mathbb{R}_{0+} & \longrightarrow & \mathbb{R}_{0+} \\ & x & \longmapsto & x^2 \end{array} \tag{5.3.4}$$

[∗]: Check! is bijective (provided we restrict the codomain of s too!).[∗]
The function

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^3 + x \end{array}, \tag{5.3.5}$$

is bijective. This can be checked using techniques from Analysis (or Foundation of Analytical Skills) by looking at the sign of the derivative of f and is left as an exercise. To prove that f is bijective using purely algebraic techniques is not that easy and this is not required for the Introduction to Pure Mathematics exam.

5.3.19. Example (a bijective set-to-set mapping). Let S be a set, consider the *complementation* map

$$\begin{array}{rccl} \phi: & \wp(S) & \to & \wp(S) \\ & X & \mapsto & \phi(X) := S \setminus X \end{array} . \tag{5.3.1}$$

The funciton ϕ is bijective.
Let us first prove that ϕ is injective: suppose $X, Y \subseteq S$, then

$$\begin{aligned} \phi(X) = \phi(Y) &\Rightarrow S \setminus X = S \setminus Y \\ &\Rightarrow X = S \setminus (S \setminus X) = S \setminus (S \setminus Y) = Y. \end{aligned} \tag{5.3.2}$$

So (5.3.1) is satisfied.
To conclude we prove that ϕ is surjective. Let $Y \in \wp(S)$, we want to find $X \in \wp(S)$ such that $\phi(X) = Y$. This is easily achieved by taking $X := S \setminus Y$. We conclude this section with the chief purpose of injective, surjective and bijective maps, which consists in counting.

5.3.20. Theorem (finite cardinality comparison). *Suppose A and B are finite sets and let $\phi : A \to B$ be a mapping between them.*

(a) If ϕ is surjective then $\#A \geq \#B$.
(b) If ϕ is injective then $\#A \leq \#B$.
(c) If ϕ is bijective then $\#A = \#B$.

Proof The proof will be given in Chapter 6. □

5.3.21. Remark (well-posedness of Definition 5.1.2). An important consequence of the Finite Cardinality Comparison Theorem is that Definition 5.1.2 makes sense. Namely, if $m, n \in \mathbb{N}_0$ are such that there exist two bijections

$$\phi : [1 \dots m] \to S \text{ and } \psi : [1 \dots n] \to S, \tag{5.3.1}$$

then $m = n$. There is no risk, thus, that a finite set S ends up having two different cardinalities.
To prove this fact, we need to introduce the *composition* of two functions $f : A \to B$ and $g : B \to C$ as the function $h : A \to C$ such that $h(x) = g(f(x))$ for all $x \in A$ with the notation $g \circ f$. to show that the composition $\psi \circ \phi$ is a bijection and then use Theorem 5.3.20. This will be done in Chapter 6.

5.3.22. Theorem (inverse function). *A function $f : A \to B$ is bijective if and only if there exists a function $g : B \to A$ such that*

$$f \circ g = \mathrm{id}_B \text{ and } g \circ f = \mathrm{id}_A. \tag{5.3.1}$$

For each $f : A \to B$ if a map g satisfying (5.3.1) exists then g is unique, is called the inverse function of f and is denoted f^{-1}.
Proof The proof of this important result will be given in 6. □

5.4. The Binomial Theorem

We now put the theory developped so far to good use to answer a fundamental question about finite sets and prove a very important result known as the *Binomial Theorem* in this section. Let S be finite and let $n := \#S \in \mathbb{N}$. Suppose $k \in [0 \dots n]$, we ask ourselves how many subsets of exactly k elements are there in S. If $k = 0$, the answer is easy: there is only one set with no elements, the empty set (which we know to be unique), so the number of subsets with 0 elements in S is 1. Also for $k = n$ the answer is quite easy, there is only one subset of S which has n elements: S itself. Thus the number of subsets with n elements in S is also 1. Now what if $0 < k < n$, which may happen as soon as $n \geq 2$.

5.4.1. Definition of binomial coefficients/combinations. Given $n \in \mathbb{N}_0$ and $k = 0, \dots, n$ we define the *binomial coefficient n choose k* (also known as *the number of combinations of k elements out of n*), denoted by $\binom{n}{k}$, as the number of all possible subsets of $[1 \dots n] := \{1, \dots, n\}$ consisting of exactly k elements. In symbols this means

$$\binom{n}{k} := \#\left\{X \in \wp([1 \dots n]) : \#X = k\right\}. \tag{5.4.1}$$

The binomial coefficient $\binom{n}{k}$ is pronounced shortly as *n choose k*. Sometime $\binom{n}{k}$ is denoted by $C(n,k)$, $C_{n,k}$ or C_k^n, where the letter C stands for "combinations".

5.4.2. Proposition. *Let S be a finite set of n (distinct) elements, then for each $k \in \mathbb{N}_0$, $0 \leq k \leq n$, we have*

$$\#\{X \subseteq S : \#X = k\} = \binom{n}{k} \tag{5.4.1}$$

Proof To prove the result it is enough to establish a one-to-one correspondence between the set on the left-hand side of (5.4.1), which we denote by $\mathscr{K}$, and the set on the right-hand side of (5.4.1), which we denote by $\mathscr{C}$. From §5.3.21 we know that there exists a bijection $\phi : [1 \dots n] \to S$, this allows us to define the following mapping

$$\begin{array}{rccl} \psi : & \mathscr{C} & \to & \mathscr{K} \\ & X & \mapsto & \psi(X) := \left\{\phi(x) : x \in X\right\} \end{array} . \tag{5.4.2}$$

We claim that the map ψ is bijective.
ψ is surjective. Indeed, for $Y \in \mathscr{K}$ (i.e., $Y \subseteq S$ and $\#Y = k$) consider the set

$$X = \left\{\phi^{-1}(y) : y \in Y\right\}. \tag{5.4.3}$$

Then $\psi(X) = Y$, because

$$\begin{aligned} u \in \psi(X) &\Leftrightarrow u = \phi(x) \text{ for some } x \in X \\ &\Leftrightarrow u = \phi(\phi^{-1}(y)) \text{ for some } y \in Y \\ &\Leftrightarrow u \in Y. \end{aligned} \tag{5.4.4}$$

ψ is injective. Suppose $\psi(X) = \psi(X')$ for $X, X' \in \mathscr{C}$, then

$$\begin{aligned} x \in X &\Rightarrow \phi(x) \in \psi(X) = \psi(X') \\ &\Rightarrow x \in X', \end{aligned} \tag{5.4.5}$$

which means $X \subseteq X'$. Similarly $X' \subseteq X$ and thus $X = X'$. □

5.4.3. Exercise (counting subsets). *Given $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$, and consider the following definition of the binomial coefficient*

$$\binom{n}{k} := \#\left\{K \in \wp(N) : \#K = k\right\} \tag{5.4.1}$$

where $N = [1 \dots n]$. In words, $\binom{n}{k}$ is the number of all possible subsets of N which have exactly k elements.

(a) Make examples with $n = 2, 3, 4$ to understand of how $\binom{n}{k}$ behaves with respect to a variable k from 0 to n.

(b) Show that for all $n \in \mathbb{N}_0$,

$$\binom{n}{0} = 1, \binom{n}{n} = 1, \text{ and } \forall\, k \in [1 \dots n-1] : \binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \tag{5.4.2}$$

Hint. Take a generic subset K of k elements in $[1 \dots n]$, and analyse the two mutually exclusive cases: (a) $n \in K$, (b) $n \notin K$. Show in each case that you can recondut yourself to picking a subset K' from $[1 \dots n-1]$.

(c) Based on the identity (5.X.11.2)*, build the so-called Pascal triangle up to 4 rows. Compare your results with those obtained in (a).*

(d) Using identity (5.X.11.2) show that

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}. \tag{5.4.3}$$

5.4.4. Problem (Binomial Theorem). *The goal of this exercise is for you to prove Newton's Binomial Formula:*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k, \tag{5.4.1}$$

for each $n \in \mathbb{N}_0$. Make sure you understand Exercises 5.1.4 and 5.X.11 before you proceed.

(a) Write down what the formula says in the particular cases $n \in [0 \dots 4]$.

(b) To get a feel for the general case, concentrate first on $n = 4$ and expand $(a+b)^n$ as to obtain

$$(a+b)^4 = (a\,a\,a\,a + b\,a\,a\,a + a\,b\,a\,a + b\,b\,a\,a + a\,a\,b\,a + b\,a\,b\,a + \dots + a\,b\,b\,b + b\,b\,b\,b). \tag{5.4.2}$$

Using a binomial tree (or a table with 4 columns and $0/1$ entries, or what you know about $\wp([1 \dots 4])$) count how many of the summands appearing in the right-hand side of (5.X.12.2) *will have 0 times the factor a, how many will have 1 time the factor a, how many will have 2 time the factor a, etc. Relate these numbers to the numbers $\binom{4}{k}$ from Exercise 5.1.4.*

(c) Repeating the intuitive argument, with n instead 4, explain why you expect the Binomial Formula (5.X.12.1) *to be true.*

(d) Now it's time to do the grown-ups part of the exercise. Use induction on n (and the recursive formula in Exercise 5.X.11) to show (5.X.12.1).

Hint. You may want to try first the inductive step for the particular case where you go from 3 to 4, to get "inspiration" for the general case from $n-1$ to n for any $n \in \mathbb{N}$.

5.4.5. Problem (Power set and binomial distribution).
(a) Using the results from Exercises 5.1.4 and the definition of $\binom{n}{k}$ *(see 5.X.11, e.g.), explain why*

$$2^n = \sum_{k=0}^{n} \binom{n}{k}, \tag{5.4.1}$$

must be true.

(b) Using (5.X.12.1) *show, algebraically, that (5.4.1) must be true.*

Hint. You may want to recall the elementary, yet quite useful fact, that $1+1=2$.

(c) Let $n \in \mathbb{N}$, *find the sum*

$$\sum_{\substack{k\in\mathbb{N}\\ 0\leq k\leq n/2}} \binom{n}{2k}. \tag{5.4.2}$$

5.5. The Inclusion-Exclusion Principle

We address now a different problem in counting which is so intuitive that we have used it “subconsciously” in some of the previous sections.

5.5.1. Counting the union of two sets. Suppose we are given two finite sets A and B, say with $\#A = n$ and $\#B = m$, what can we say about the cardinality of $A \cup B$?
An inspection shows that

$$\min\{\#A, \#B\} \leq \#(A \cup B) \leq \#A + \#B. \tag{5.5.1}$$

The lower bound is attained whenever one set is a subset of the other, i.e., $A \subseteq B$ or $B \subseteq A$, whereas the upper bound is attained when the two subsets are disjoint. In fact, we may strengthen this statement as follows.

5.5.2. Theorem (inclusion-exclusion (baby case)). *Given two finite sets* A, B *we have*

$$\#(A \cup B) = \#A + \#B - \#(A \cap B). \tag{5.5.1}$$

Proof We leave the proof as an exercise (e.g., by induction on $n = \#A$). □

5.5.3. Example.

Problem. *Suppose that in a class* C *each student either smokes or drinks, or both. Suppose the class* C *has* 37 *students, of which* 27 *drink and* 17 *smoke. How many students both drink and smoke?*

Solution. Let S be the set of smokers and D the set of drinkers, we want to know how much is $\#(D \cap S)$.
We know that $C = D \cup S$, because each student either drinks or smokes. Thus

$$37 = \#C = \#(D \cup S) = \#D + \#S - \#(D \cap S) = 27 + 17 - \#(D \cap S) \tag{5.5.1}$$

and the answer is

$$\#(D \cap S) = 27 + 17 - 37 = 7. \tag{5.5.2}$$

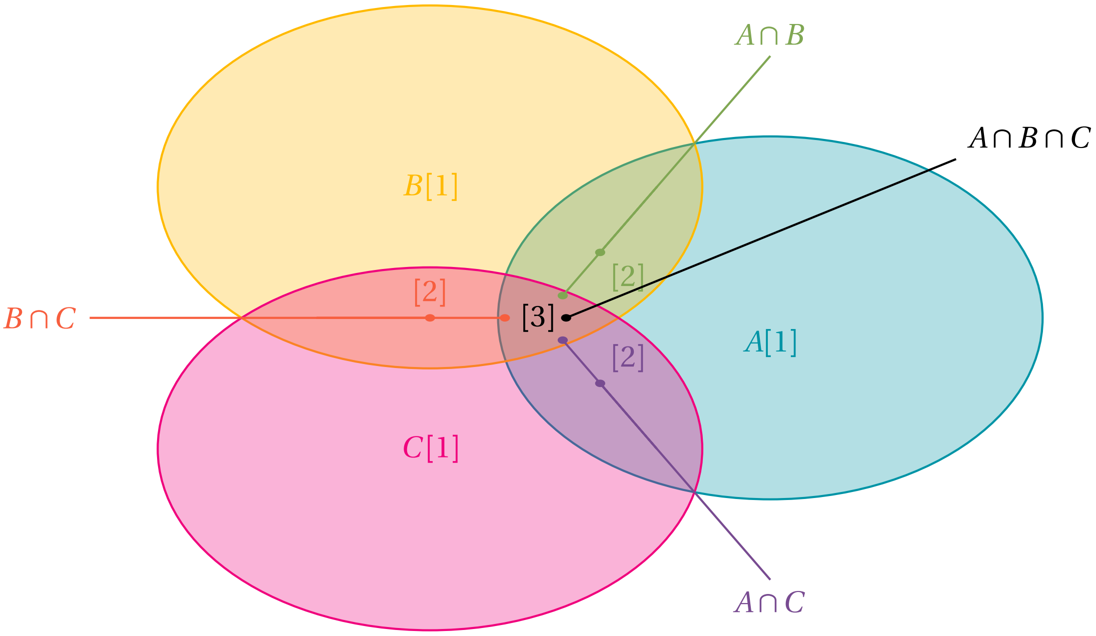
Theorem 5.5.2 can be generalised to the case of three sets.

5.5.4. Theorem (inclusion-exclusion for 3 sets). *Given three finite sets A, B and C we have*

$$\#(A\cup B\cup C)=\#A+\#B+\#C \\ -(\#(A\cap B)+\#(A\cap C)+\#(B\cap C))+\#(A\cap B\cap C). \quad (5.5.1)$$

Proof The proof of this theorem is also left as an exercise. □

5.5.5. Remark (why "inclusion-exclusion"). Informally, relation (5.5.1) can be understood as *including* in the count each element in A, B and C, once, but then having to *exclude* the elements which were counted twice, i.e., those who are in $A\cap B$, $A\cap C$ or $B\cap C$, but then we have to *include* again those who were excluded twice, i.e., those that are in $A\cap B\cap C$. This inclusion-exclusion process explains the name of the Theorem.
It is useful for this situation to be "drawn" using Venn diagrams as follows.



The numbers in square brackets inside each delimited area denote the number of inclusion/exclusions performed on each element of that area, in order to count the elements in the union of the whole lot.

5.5.6. Cardinality of a mutually disjoint union. Before we proceed to the general case of the Inclusion-Exclusion Principle, we treat the simple situation where we have a finite *mutually disjoint* collection $\{A_1,\cdots,A_n\}$ where the set A_i is finite for each $i\in[1\dots n]$. Mutually disjoint means that

$$i\neq j \Rightarrow A_i\cap A_j=\varnothing. \quad (5.5.1)$$

Theorem. *Suppose* $\{A_1,\cdots,A_n\}$ *is a mutually disjoint collection of sets, then*

$$\#\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\#A_i. \quad (5.5.2)$$

Proof By induction on n. The result is true for $n = 2$ by Theorem 5.5.2. For $n \geq 2$ suppose the result is true for $n-1$, then

$$\begin{aligned}\#\left(\bigcup_{i=1}^{n} A_i\right) &= \#\bigcup_{i=1}^{n-1} A_i + \#A_n \qquad \text{(2 disjoint sets)}\\ &= \sum_{i=1}^{n-1} \#A_i + \#A_n \qquad \text{(inductive hypothesis)} \qquad (5.5.3)\\ &= \sum_{i=1}^{n} \#A_i.\end{aligned}$$

□

5.5.7. The general case: Inclusion-Exclusion Principle. *This material is not required for exam.*

Suppose now that we are give a finite collection of finite sets, i.e.,

$$\{A_i : i \in I\} \tag{5.5.1}$$

where each A_i is finite and I is a finite set of indexes.

Then the cardinality of the union

$$A_1 \cup \cdots \cup A_n = \bigcup_{i \in I} A_i \tag{5.5.2}$$

is given by

$$\begin{aligned}\sum_{i=1}^{n} \#A_i - \sum_{1 \leq i < j \leq n} \#\left(A_i \cap A_j\right) + \sum_{1 \leq i_1 < i_2 < i_3 \leq n} \#\left(A_{i_1} \cap A_{i_2} \cap A_{i_3}\right)\\ + \cdots + (-1)^k \sum_{1 \leq i_1 < \cdots < i_k \leq n} \#\left(\bigcap_{j=1}^{k} A_{i_j}\right) \qquad (5.5.3)\\ + \cdots + (-1)^n \#\left(\bigcap_{i=1}^{n} A_i\right).\end{aligned}$$

This can be summarised as

$$\#\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \#\left(\bigcap_{j=1}^{k} A_{i_j}\right). \tag{5.5.4}$$

This can be compressed even further as

$$\sum_{J \subseteq I} (-1)^{\#J} \#\left(\bigcap_{i \in J} A_i\right) = 0, \tag{5.5.5}$$

with the convention that

$$\bigcap_{i \in \varnothing} A_i := \bigcup_{i \in I} A_i. \tag{5.5.6}$$

Proof not required for exam We proceed by induction on $n = \#I$. The base case, for $n = 2$, is given by Theorem 5.5.2. Let $n \geq 3$ and suppose that the formula is true when $\#I = n-1$, we want to show that it is true for $\#I = n$. For simplicity let us take $I = [1\dots n]$. Then, by the associative property of summations,

$$\sum_{J\subseteq I}(-1)^{\#J}\#\left(\bigcap_{i\in J}A_i\right) = \sum_{\substack{J\subseteq I\\ J\not\ni n}}(-1)^{\#J}\#\left(\bigcap_{i\in J}A_i\right) + \sum_{\substack{J\subseteq I\\ J\ni n}}(-1)^{\#J}\#\left(\bigcap_{i\in J}A_i\right). \tag{5.5.7}$$

Denote by $I' = I \setminus \{n\}$, then the first term on the right-hand side simplifies to

$$\sum_{\substack{J\subseteq I\\ J\not\ni n}}(-1)^{\#J}\#\left(\bigcap_{i\in J}A_i\right) = \sum_{J\subseteq I'}(-1)^{\#J}\#\left(\bigcap_{i\in J}A_i\right) = 0 \tag{5.5.8}$$

thanks to the inductive hypothesis, which may be applied to the collection $\{A_1,\dots,A_{n-1}\}$ with the index set I' since $\#I' = n-1$.
To treat the second term on the right-hand side of (5.5.7), introduce the sets

$$J' := J \setminus \{n\},\ \forall\ J \subseteq I, A'_i := A_i \cap A_n,\ \forall\ i \in I'. \tag{5.5.9}$$

Then we have

$$\sum_{\substack{J\subseteq I\\ J\ni n}}(-1)^{\#J}\#\left(\bigcap_{i\in J}A_i\right) = -\sum_{J'\in I'}(-1)^{\#J'}\#\left(\bigcap_{i\in J'}A'_i\right) = 0, \tag{5.5.10}$$

which is again true by the inductive hypothesis with the index set I', but on the collection $\left\{A'_1,\cdots,A'_{n-1}\right\}$ this time. The conclusion follows by summing up. □

5.6. Permutation count

5.6.1. Students, chairs and music. It is often useful to think of a finite set as being "ordered". For example, while a bunch of students constitute your Pure class, you may be interested in understanding how the students are sitting in the class. And although most people tend to sit in the same place when they return to a lecture hall, some people change or switch places.
A basic problem in counting, is to understand in how many different ways you can sit a class of n students in and N-seat lecture theatre.
Of course, the problem's solution is easy if $N < n$, this is the *musical chair problem* and there is clearly no way of seating n students on N chairs (we are working under the assumption that there can be no more than one student per chair). So if $N < n$ the number of possible arrangements is 0.
Let us turn to the situation of $N \geq n$ and try to solve it via a cherry-picking argument. In fact we give two different solutions.
As we have seen, when we count a set, we give it an (arbitrary) initial order, for example a class is usually given in alphabetical order, say $(s_1, s_2, \dots, s_n)$. (Note the use of round brackets, to indicate that we care about order, not just the set of students $\{s_1, s_2, \dots, s_n\}$.)

5.6.2. Cherry-pick-and-exhaust-students solution. Let us now go through all the students s_i, $i = 1,\dots,n$ and choose a seat for each one:

- ⋆ $i = 1$, s_1 can be seated on any one of the N available seats,
- ⋆ $i = 2$, s_2 can be seated on any one of the remaining $N-1$ seats,
- ⋆ …
- ⋆ i, s_i can be seated on any one of the remaining $N+1-i$ seats,

- ⋆ …
- ⋆ $i = n$, s_n can be seated on any one of the remaining $N+1-n$ seats.

Therefore the number of all possible seating configurations is

$$N(N-1)\cdots(N+1-n). \tag{5.6.1}$$

Note that this number can be also written as

$$N(N-1)\cdots(N+1-n) = \frac{N!}{(N-n)!} = \binom{N}{n} n! \tag{5.6.2}$$

which can be interpreted as the following alternative answer to the counting question: there are $\binom{N}{n}$ ways of chi

5.6.3. Choose the seats and permute. Another way of solving the problem is to

- ⋆ first choose the n seats, out of the N available ones, that will be occupied,
- ⋆ then find all the possible "permutations" of the n students for each choice of seats.

To do the first step, we find all possible subsets of n elements among N possible ones, and this is just the binomial coefficient $\binom{N}{n}$.
The second step can be performed by general reasoning, as we did in §5.6.2, except now we have n instead of N possible seats to choose from. This yields

$$n(n-1)\cdots 2\times 1 = n! \tag{5.6.1}$$

possible permutations.
Therefore the number of possible seating configurations of n people in N-seat theatre is

$$\binom{N}{n} n!, \tag{5.6.2}$$

which, unsurprisingly, equals the first solution.

5.7. Functions, cartesian products and power

We close this section by solving some problems in function counting.

5.7.1. Problem (internal map counting). *Consider a finite set X with exactly n elements, say*

$$X = \{x_1, x_2, \dots, x_n\}. \tag{5.7.1}$$

How many functions are there from X into itself (internal maps)?
Note that this is different than counting the permutations of X because we do not require the internal maps to be bijective.

Solution. Let $\phi : X \to X$ be the generic map we want to choose. There are n different ways to choose $\phi(x_1)$, n different ways to choose $\phi(x_2)$, etc. Hence the total number of different maps is

$$\underbrace{n\,n\cdots n}_{n \text{ factors}} = n^n. \tag{5.7.2}$$

5.7.2. Problem (general map counting). *Consider two finite sets X and Y. To keep it simple, let's place ourselves in the particular case where $X = [1 \dots n]$ and $Y = [1 \dots m]$, with $m, n \in \mathbb{N}_0$.*
How many different functions are there from X into Y?
Note that Problem 5.7.1 is a particular case of this problem, with $X = Y$.

Solution. Let $\phi : X \longrightarrow Y$ be the generic map we want to choose. There are k different ways to choose $\phi(1)$, k different ways to choose $\phi(2)$, etc. up to $\phi(n)$. Hence the total number of different maps is

$$\underbrace{k\,k \cdots k}_{n \text{ factors}} = k^n. \tag{5.7.1}$$

To make a more rigorous argument, note that counting the functions from X into Y, amounts to counting the n-tuples

$$\left(y_1, \dots, y_n\right) \text{ where } y_i \in Y, \text{ for } i = 1, \dots, n. \tag{5.7.2}$$

The set of these n-tuples is nothing but the cartesian product

$$\underbrace{Y \times Y \times \cdots \times Y}_{n \text{ factors}} =: Y^n. \tag{5.7.3}$$

This reduces the problem to the following one.

5.7.3. Theorem (cartesian product cardinality and power). *Let $X_1, \dots, X_m$ be a finite sequence of finite sets. Then*

$$\#X_1 \times X_2 \times \dots \times X_m = \#X_1 \#X_2 \dots \#X_m. \tag{5.7.1}$$

Proof By induction on m. For $m = 1$ (base case) the formula is trivial.[9]
To prove the inductive step suppose

$$\#X_1 \times X_2 \times \dots \times X_{m-1} = \#X_1 \#X_2 \dots \#X_{m-1}. \tag{5.7.2}$$

Denote by $X'_m := X_1 \times X_2 \dots \times X_{m-1}$ and note that

$$X_1 \times X_2 \times \cdots \times X_m := X'_m \times X_m. \tag{5.7.3}$$

We already know (cf.5.X.1) that for two sets A and B we have

$$\#A \times B = \#A \#B. \tag{5.7.4}$$

Using this fact, and the inductive hypothesis (5.7.2) we obtain

$$\#(X_1 \times \cdots \times X_m) = \#X'_m \#X_m = (\#X_1 \cdots \#X_{m-1}) \#X_m, \tag{5.7.5}$$

as claimed. □

[9] The purist may want to start at $m = 0$, but this is a bit confusing and requires a proper definition of the empty cartesian product as being $\{\varnothing\}$ which has cardinality 1.

Exercises and problems on counting

Exercise 5.X.1 (cardinality and products). Consider the sets $A = \{a_1, a_2, \dots, a_n\}$ (with $a_i \neq a_j$) and $B = \{b_1, \dots, b_m\}$.

(a) Draw an graph representing $A \times B$, for $n = 5$ and $m = 3$, and identify the element (a_4, b_2).

(b) How many elements does $A \times B$ have in your example?

(c) How would you complete the following formula

$$\#(A \times B) = \boxed{\texttt{[01]}} \tag{5.X.1.1}$$

in the general case where $n, m \in \mathbb{N}$.

(d) Prove the formula.

Exercise 5.X.2 (quadratic functions). Consider the quadratic function $q : \mathbb{Z} \to \mathbb{Z}$ given by

$$q(x) = ax^2 + bx + c, \tag{5.X.2.1}$$

where $a, b, c \in \mathbb{Z}$ are fixed numbers.

(a) What can you say about $q^{-1}\{n\}$ for $n \in \mathbb{Z}$? Explain.

(b) Can you choose a such that q is injective? Explain.

(c) Can choose a and b such that q is bijective? Explain.

Exercise 5.X.3 (distinguished functions). For each of the following four functions answer both of the following:

(a) Is the function injective? Explain.
(b) Is the function surjective? Explain.

You must explain each of your answers: to prove it, give an algebraic argument, or to disprove it, give a counterexample.

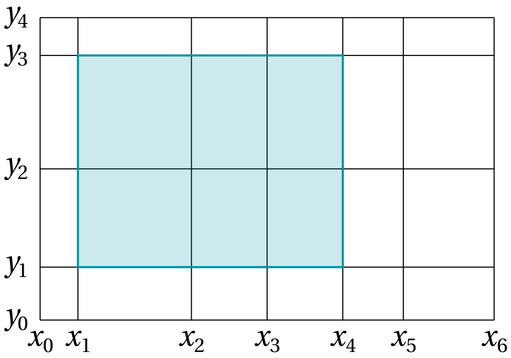
(i) $f : \mathbb{Z} \to \mathbb{Z}$ with $f(x) := x^2$.
(ii) $g : \mathbb{N} \to \mathbb{N}$ with $g(x) := x^2$.
(iii) $h : \mathbb{Z} \to \mathbb{Z}$ with $h(x) := x + 3$.
(iv) $k : \mathbb{Z} \to \mathbb{Z}$ with $k(x) := 2x + 3$.

Exercise 5.X.4 (counting lines through 6 points). In the plane, consider 6 points, no 3 of which are aligned.

(a) Count the number of all possible lines that you can obtain by joining each pair of points.
(b) How many intersection points, besides the original 6 points, are there *at most*?

Exercise 5.X.5 (counting subrectangles)**.**

Consider a large rectangle. Cut its base into m parts and its height into n parts. Then partition the rectangle into $m \times n$ *elementary rectangles*, as in the picture. How many rectangles can you count in the figure thus obtained? A nonelementary rectangle is shaded in the picture.



Hint. Establish a one-to-one correspondence between the set of all possible rectangles and the set

$$\left\{(x_i, x_j), (y_k, y_l) \colon 0 \leq i < j \leq m \text{ and } 0 \leq k < l \leq n\right\}. \tag{5.X.5.1}$$

Try an example first, say with $m = 3$ or 4 and $n = 2$ or 3 (start counting "by hand" and once you understand this do the general problem).

Exercise 5.X.6 (counting digitals)**.** (a) Among all digital expansions of exactly 7 digits, how many are there with only "1" and "2" appearing in them (e.g., 1111211 and 1212121)? Explain your answer.

(b) How many are there with exactly 4 figures "1" and 3 figures "2" (e.g., 1121122)? Explain your answer.

Exercise 5.X.7. Consider the digital expansions of $\mathbb{N}_0$ numbers with 4 or less digits, i.e., from 0 to 9999, with the convention that we write the "trailing zeros" for numbers less that 1000, e.g., 0013 for 13 and 0103 for 103.

(a) How many are there with all different digits (e.g., 1409)?
(b) How many are there with decreasing digits (e.g., 7541)?
(c) How many are there with only even digits (e.g., 2008)?
(d) How many are there whose digits are respectively (i.e., in this order) even, even, odd, odd (e.g., 2019)?
(e) How many are there with 2 even digits and 2 odd digits (in any order)?

Exercise 5.X.8 (symmetric group is not abelian)**.** Recalling that a group is called *Abelian* if its operation is commutative.
Show, by providing a counterexample, that the symmetric group $\operatorname{Sym}(X)$ is not Abelian in general.
Hint. Find an example where $f, g \in \operatorname{Sym}(A)$ and $a \in A$ such that

$$f(g(a)) \neq g(f(a)). \tag{5.X.8.1}$$

Problem 5.X.9 (power sums—constructively revisited)**.** The first time induction is introduced, students are asked to show the formulas summing a given power of the first d integers

$$\sum_{n=1}^{N} n = \frac{1}{2}N^2 + \frac{1}{2}N, \quad \sum_{n=1}^{N} n^2 = \frac{1}{3}N^3 + \frac{1}{6}N^2 + \frac{1}{2}N, \text{ and so on for successive powers.} \tag{5.X.9.1}$$

While that is a good exercise to learn induction, it does not say how those formulas can be derived constructively and imposes them on the students. In fact, it is unsettling that the solution is already known and the whole exercise appears a bit futile (in a sense it is). The purpose of this problem is to redress the impression that induction is useless and in fact provide a *constructive recursive expression* for the coefficients appearing in the power sums. Start by defining

$$s^{d+1}(N) := \sum_{n=0}^{N} n^d \text{ for } N \in \mathbb{N}_0 \text{ and } d \in \mathbb{N}_0. \tag{5.X.9.2}$$

(a) Show that

$$s^1(N) = N + 1 \tag{5.X.9.3}$$

(hence a polynomial of degree 1 in N) and let $s^1(X) := X + 1$, where X is used as the polynomials *indeterminate* (also known as *placeholder*).

(b) Using the binomial formula, for $d \geq 1$, obtain an expansion in the form

$$(n+1)^d - n^d = \binom{d}{1} n^{d-1} + \binom{d}{2} n^{d-2} + \cdots \tag{5.X.9.4}$$

with no powers of d (only strictly lower ones) appearing on the right-hand side.

(c) Summing up both sides of the expression above for $n = 0, \dots, N$, telescoping on the left-hand side and using the definition of s^d on the right-hand side, find

$$s^d(N) = \frac{1}{d}\left((N+1)^d - \sum_{n=0}^{d-1} \binom{d}{n} s^n(N) \right). \tag{5.X.9.5}$$

(d) Deduce by induction on d that for each $d \in \mathbb{N}_0$ $s^d(X)$ is a polynomial of degree d of the form

$$s^d(X) = \sum_{k=0}^{d} s^d_k X^k \left(= \sum_{k \geq 0} s^d_k X^k \text{ with } s^d_k = 0 \text{ if } k > d \right) \tag{5.X.9.6}$$

and find a recursive expression for the coefficients s^d_k.

(e) Complete the table

$$\begin{matrix} s^1_0 & s^2_0 & s^3_0 & s^4_0 \\ s^1_1 & s^2_1 & s^3_1 & s^4_1 \\ s^1_2 & s^2_2 & s^3_2 & s^4_2 \\ s^1_3 & s^2_3 & s^3_3 & s^4_3 \\ s^1_4 & s^2_4 & s^3_4 & s^4_4 \end{matrix} \tag{5.X.9.7}$$

and check that the coefficients you obtained match those from otherwise "known" formulas.

Problem 5.X.10 (binomial coefficients)**.** Show that

$$\sum_{k=0}^{20} \binom{50}{k} \binom{50}{20-k} = \binom{100}{20}. \tag{5.X.10.1}$$

Hint. Avoid algebra. Use the (proper) combinatorial definition of binomial coefficients.

Problem 5.X.11. Given $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$, and consider the following definition of the binomial coefficient

$$\binom{n}{k} := \#\left\{K \in \wp(N) \colon \#K = k\right\} \tag{5.X.11.1}$$

where $N = [1 \dots n]$. In words, $\binom{n}{k}$ is the number of all possible subsets of N which have exactly k elements.

(a) Make examples with $n = 2, 3, 4$ to understand of how $\binom{n}{k}$ behaves with respect to a variable k from 0 to n.

(b) Show that for all $n \in \mathbb{N}_0$,

$$\binom{n}{0} = 1, \binom{n}{n} = 1, \text{ and } \forall\, k \in [1 \dots n-1] \colon \binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \tag{5.X.11.2}$$

Hint. Take a generic subset K of k elements in $[1 \dots n]$, and analyse the two mutually exclusive cases: (a) $n \in K$, (b) $n \notin K$. Show in each case that you can recondut yourself to picking a subset K' from $[1 \dots n-1]$.

(c) Based on the identity (5.X.11.2), build the so-called *Pascal triangle* up to 4 rows. Compare your results with those obtained in (a).

(d) Using identity (5.X.11.2) show that

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}. \tag{5.X.11.3}$$

Problem 5.X.12 (binomial theorem)**.** The goal of this exercise is for you to prove Newton's Binomial Formula:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k, \tag{5.X.12.1}$$

for each $n \in \mathbb{N}_0$. Make sure you understand Exercises 5.1.4 and 5.X.11 before you proceed.

(a) Write down what the formula says in the particular cases $n \in [0 \dots 4]$.

(b) To get a feel for the general case, concentrate first on $n = 4$ and expand $(a+b)^n$ as to obtain

$$(a+b)^4 = (aaaa + baaa + abaa + bbaa + aaba + baba + \dots + abbb + bbbb). \tag{5.X.12.2}$$

Using a binomial tree (or a table with 4 columns and 0/1 entries, or what you know about $\wp([1 \dots 4])$) count how many of the summands appearing in the right-hand side of (5.X.12.2) will have 0 times the factor a, how many will have 1 time the factor a, how many will have 2 time the factor a, etc. Relate these numbers to the numbers $\binom{4}{k}$ from Exercise 5.1.4.

(c) Repeating the intuitive argument, with n instead 4, explain why you expect the Binomial Formula (5.X.12.1) to be true.

(d) Now it's time to do the grown-ups part of the exercise. Use induction on n (and the recursive formula in Exercise 5.X.11) to show (5.X.12.1).

Hint. You may want to try first the inductive step for the particular case where you go from 3 to 4, to get "inspiration" for the general case from $n-1$ to n for any $n \in \mathbb{N}$.

CHAPTER 6

Functions and natural numbers

You never know what is enough until you know what is too much.
— William Blake

We have seen functions (and maps) throughout the course, especially in Chapter 5 where we have learned how to use bijections to count finite sets. Functions, which are also known as maps, mappings, rules, correspondences, operators, applications, are ubiquitous objects in mathematics. While it is possible to imagine the development of mathematics without sets (a hard task, which some still make possible), no one mathematician can do without functions.

In this Chapter we develop rigorously, building on the Set Theory so far discussed, the main tools for manipulating functions. What we discuss in this chapter is fundamental for other courses such as Geometry, Linear Algebra, Foundations of Analytical Skills, and Analysis.

6.1. Formal definition of a function

6.1.1. What does a function "do"? As we have learned intuitively a vague definition of *function* is that it is a rule that takes elements, called *variable* or *argument*, in a set, called the *domain*, and returns elements, called the *value* or the *image*, in another set, called the *codomain*.

Given a "rule", for it to be a function $f : A \to B$, we want f to:

(i) to return *at least* one element $y \in B$, for each $x \in A$, in symbols:

$$\forall\, x \in A : \exists\, y \in B : y = f(x); \tag{6.1.1}$$

(ii) to return *at most* one such element, in symbols:

$$y = f(x) \text{ and } z = f(x) \Rightarrow\ y = z. \tag{6.1.2}$$

Of course, could say, in one sentence that for each element $x \in A$, there is exactly one $y \in B$ such that $y = f(x)$. This is sometime written as

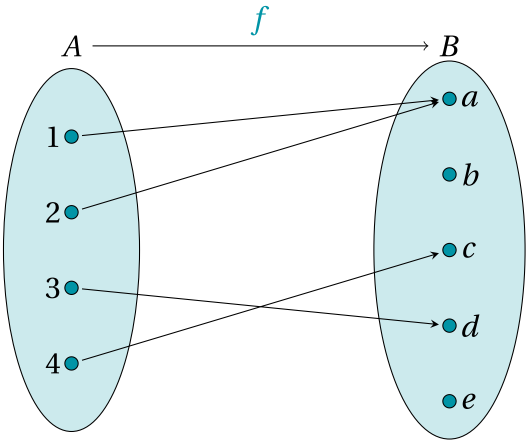
$$\forall\, x \in A : \exists!y \in B : y = f(x), \tag{6.1.3}$$

but something may be lost on the beginner, if we spell it in such a concentrated manner. So we rather think of the "there exists exactly one" as being two different statements that need to be checked separately.

In summary, a function f for each one input, it returns exactly one output. All this inputing and returning gives us a very neat picture of "action" for each function. On the other hand, sets are such "static" objects, sitting in the universe, hardly "moving".

6.1.2. A set-theoretical approach to functions. In order to build the concept of function from set theory, we need to make a somewhat artificial effort, in thinking of a function as being a static object, rather than some kind of machine which churns inputs into outputs.

Let us see how to obtain a proper definition of a function using sets, by working on a concrete example. Consider the following is a representation, by Venn diagrams of a very simple function.



An alternative representation of this function can be made using a table (or a hash) as follows

input	x	1	2	3	4
output	$f(x)$	a	a	d	c

(6.1.1)

Another way of representing the function f is to use the "rule notation"

$$f(1)=a,\ f(2)=a,\ f(3)=d,\ f(4)=c. \tag{6.1.2}$$

Also the "maps-to arrow" notation is sometimes used

$$f: 1 \mapsto a, 2 \mapsto a, 3 \mapsto d, 4 \mapsto c. \tag{6.1.3}$$

Some strange people (e.g., computer scientists) sometime use the "object.method notation"

$$1 \cdot f = a, 2 \cdot f = a, 3 \cdot f = d, 4 \cdot f = c; \tag{6.1.4}$$

we will not use this notation further in this course.[1]

[1] There are many other notations for functions with many inconstistencies, but these never constitute a big problem, examples include:

- subindexing the function with the variable, as in sequences $a_1, a_2, a_3, \dots$, which is a function from natural numbers (or a subset thereof) into some set;
- superindexing the function with the variable, as in the exponential function e^x;
- postfixing the variable without brackets as in linear function $\boldsymbol{Ax}$;
- superindexing the function, as in the power function or real (or complex) variable x^3, or the complement of a set variable X^{c}, or the transpose of a matrix variable $\boldsymbol{X}^{\mathsf{T}}$, or the derivative of a variable, say u, which is a function u' (a function of functions);
- by using typographical "embelishments", as in the complex conjugate of a complex variable $\bar{z}$, or the Fourier transform of a variable which is itself a function $\hat{u}$, or the Netwon-styled derivative $\dot{u}$;
- by using special type of brackets, as in the absolute value of a real (or complex) variable $|x|$ or the ceil or floor functions $\lceil x \rceil$ or $\lfloor x \rfloor$, respectively, of a real variable x;

and many more notations which we have surely missed.

Using set theoretical notation the function f can be represented by the following set of ordered pairs

$$S := \{(1,a),(2,a),(3,d),(4,c)\}. \tag{6.1.5}$$

In fact we can *think* of *the function f being this set*:

$$f := S. \tag{6.1.6}$$

This is the key to the proper definition of a function $f : A \to B$.

6.1.3. Definition of function. Given two sets A and B, a *function f* from A into B is a subset of the cartesian product $A \times B$, i.e., $f \subseteq A \times B$, which satisfies the following two properties:

(i) (existence of output) each input has at least one output

$$\forall\, x \in A : \exists\, y \in B : (x,y) \in f. \tag{6.1.1}$$

(ii) (uniqueness of output) each input has at most one output

$$(x,y) \in f \text{ and } (x,z) \in f \;\Rightarrow\; y = z. \tag{6.1.2}$$

When f is a function from A into B we write it as

$$f : A \to B, \tag{6.1.3}$$

and if $(x,y) \in f$ we use the more usual notation

$$f(x) = y \text{ or, occasionally, } f : x \mapsto y. \tag{6.1.4}$$

Thanks to the uniqueness of the output the "=" in this notation is not dangerous.[2]

6.1.4. Remark (functions are particular graphs). Recalling the definition of a graph (see Chapter 7 §4.8.5), we note that a function is just a particular type of graph from the set A into the set B.

6.1.5. Terminology and notation. The terminology and the notation regarding functions is so diverse and widespread that we cannot contain all of it in this Chapter. To make things worse, many mathematicians, including very good ones, are quite sloppy with their notation. We will be as precise as possible with our notation, because a good notation makes it easier to learn concepts. You are strongly advised to be as precise as possible too.

Assume in the rest of this paragraph that $f : A \to B$ is a given function from A into B. The set A is called *domain (of definition)* of f, the set B is the *codomain* (also known as *range*) of f. The generic element of A is called the *argument* (or *independent variable*) of f and for each $x \in A$, the corresponding element $f(x) \in B$ is called the *image* of x through f (in B).

For each given function, since its domain and codomain are unique we define

$$\operatorname{Dom} f := A \text{ and } \operatorname{Cod} f := B \tag{6.1.1}$$

Given a subset $X \subseteq A$, we define its *image set*

$$f(X) := \left\{ y \in B : \; \exists\, x \in X : f(x) = y \right\} = \left\{ f(x) : \; x \in X \right\}. \tag{6.1.2}$$

By the Axiom of Specification $f(X)$ is a subset of B.

[2] Try to figure out why the notation $y = f(x)$ would be "dangerous" if $f(x)$ had more than one output value, say $z = f(x)$ with $z \neq y$.

6.1.6. Remark (abuse of notation). Note the abuse of notation, although f is a function on A, when we "apply it" to a subset $X \subseteq A$ it, we are in fact definiting a new function

$$\begin{array}{rccl} f\triangleright: & \wp(A) & \to & \wp(B) \\ & X & \mapsto & f \triangleright A := \{ y \in B : \ y = f(x) \text{ for some } x \in X \} = \{ f(x) : \ x \in X \} \ . \end{array} \tag{6.1.1}$$

Conceptually f and $f\triangleright$ are two quite different function, but intuitively we like to identify them (by thinking of f as $f\triangleright$ acting on the singletons of A). Strictly speaking f has the elements of A as variables while $f\triangleright$ has the subsets of A as elements. But the function $f\triangleright$ is usually denoted by f and we shall commit this abuse of notation without further apology.

6.1.7. Definition of Image, range, counterimage. The set $f(A)$ is called the *image* of the function f. We denote

$$\operatorname{Img} f := f(A). \tag{6.1.1}$$

Some authors refer to $f(A)$ (somewhat confusingly) as the *range* of f. Some other authors use the word *range* for f's codomain, which may be different than (but always containing) $f(A)$.
Given a set $Y \subseteq B$, we define its *counterimage set* as

$$f \triangleleft Y := f^{-1} Y := \{ x \in A : \ f(x) \in Y \}. \tag{6.1.2}$$

That is $f^{-1}(Y)$ is the set of all elements in A whose image is in Y.

6.1.8. Remark (counterimage set of a singleton). Suppose $Y \in B$ is a singleton, i.e., $Y = \{y\}$ for some fixed $y \in B$. Then the counterimage of $\{y\}$, is denoted by $f^{-1}(y)$ (instead of $f^{-1}(\{y\})$). This is a double abuse of notation! Note that the *set* $f^{-1}(y)$ is *not necessarily a singleton.* In fact, it may have more than one element, or even be empty. (See the figure in the example above as an exercise and write down the counterimage of each element.) Again we will be often committing this abuse as well.

6.2. Distinguished functions revisited

We review in this section the definitions of surjective, injective and bijective map (or function).

6.2.1. Surjections. Recalling Definition 5.3.1 in Chapter 5, a function $f : A \to B$ is *surjective* (or *onto* or a *surjection*) if and only if

$$\forall\, y \in B : \exists\, x \in A : f(x) = y. \tag{6.2.1}$$

With the new terminology introduced in 6.1.5 we can express this property by saying that f is surjective if and only if

$$\forall\, y \in B : f^{-1}(y) \neq \varnothing. \tag{6.2.2}$$

Alternatively, having f surjective is equivalent to the image of A coinciding with the whole of B:

$$f(A) = B. \tag{6.2.3}$$

6.2.2. Remark (equivalence relations: a sneak preview). In Chapter 7 we have seen that equivalence relations and surjections are closely related. Namely, we have shown that for each equivalence relation $\sim$ on a set S, there is a natural map ϕ, which is surjective, such that $S : S/\sim \rightarrow$ which takes each element $x \in S$ into its equivalence class $[x]$. We now show that the converse is true too.

6.2.3. Proposition (map induced equivalences). *Suppose $f : A \rightarrow B$ is a surjective map. Then the relation $\sim$ defined by*

$$x \sim y :\Leftrightarrow f(x) = f(y), \text{ for } x, y \in A, \tag{6.2.1}$$

is an equivalence relation on A.
Furthermore the quotient set of A is given by

$$A/\sim = \left\{ f^{-1}(y) : y \in B \right\}. \tag{6.2.2}$$

In particular the collection of sets on the right-hand side of the above relation is a partition of A.
Proof The proof is not hard at all and is left as an exercise. □

6.2.4. Injections. Recall that by Definition 5.3.8 a function $f : A \rightarrow B$ is *injective* (or *one-to-one* or an *injection*)if and only if

$$f(x) = f(y) \Rightarrow x = y. \tag{6.2.1}$$

In our new notation we see that f is injective if and only if

$$\forall\, y \in B : \# f^{-1}(y) \leq 1. \tag{6.2.2}$$

Indeed, to say that $\# X \leq 1$ for a set X, is just a shortcut for saying that X has at most one element (X may still be empty though, so it is not exactly one).

6.2.5. Bijections. As we already know, a function $f : A \rightarrow B$ is called *bijective*, or a *one-to-one correspondence*, or a *bijection*, by definition if and only if it is both injective and surjective.
In view of the above, we see that f is a bijection if and only if

$$\forall\, y \in B : \# f^{-1}(y) = 1, \tag{6.2.1}$$

which means that each single element of B has one and only one element in its counterimage set.

6.3. Composition and inversion

As with set, or numbers, we may operate on functions too. The single most important operation on functions is their composition. You may be familiar, from other courses, with the concept of composition of functions. Roughly speaking, the composition of a function f with a function g is the function that is obtained by applying first f, say to some element x, and then g to the outcome of the previous, application $f(x)$. As usual, we want to be a bit careful about the domains and codomains. It pays off to concentrate a bit on these preliminary details.

6.3.1. Definition of equality of functions. Two functions f and g are said to be *equal* (sometime this is stressed as *identically equal*) if and only if they are equal as sets (see Definition 6.1.3 if you are confused), which boils down to:

$$\operatorname{Dom} f = \operatorname{Dom} g, \operatorname{Cod} f = \operatorname{Cod} g, \text{ and } f(x) = g(x), \forall\, x \in \operatorname{Dom} f. \tag{6.3.1}$$

Note that if two functions coincide on some, but not all of their arguments, they are not equal, according to this definition.

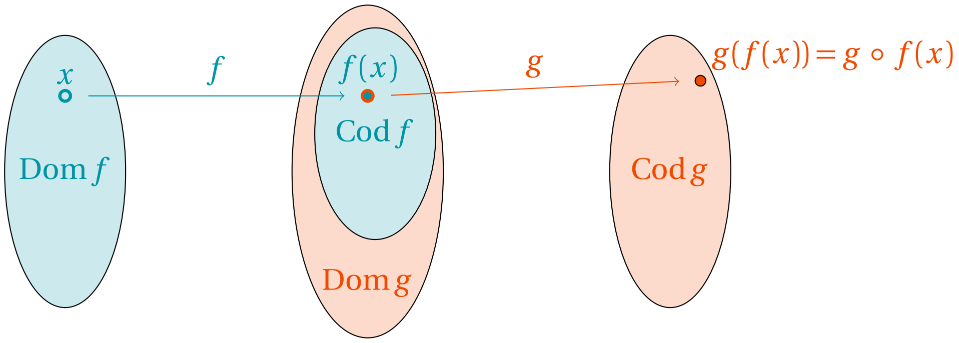
6.3.2. Definition of composition of two functions. Suppose f and g are two functions. We say that f is with g if and only if

$$\operatorname{Img} f \subseteq \operatorname{Dom} g. \tag{6.3.1}$$

If f and g are composable then, the (also known as or function) of f by (or then) g is the function

$$\begin{aligned} h: \quad \operatorname{Dom} f &\to \operatorname{Cod} g \\ x &\mapsto g(f(x)) \end{aligned} \quad . \tag{6.3.2}$$

The function h is usually denoted by $g \circ f$, though in some cases the notations $g f$ or $g(f(\cdot))$ are also used.
We summarise the idea of composition with a picture



6.3.3. Example (from "Real Life"). Consider the sets

$$A := \{\text{people}\} \text{ and } B := \{\text{males}\}. \tag{6.3.1}$$

and the functions

$$\begin{aligned} f: \quad A &\to B \\ p &\mapsto f(p) \text{ is the father of } p \\ g: \quad B &\to A \\ x &\mapsto g(x) \text{ is the firstborn child of } x \end{aligned} \tag{6.3.2}$$

Then the composition of f with g is defined by

$$g \circ f(p) = \begin{cases} \text{oldest sibling of } p & \text{if } p \text{ is not eldest,} \\ p & \text{if } p \text{ is eldest.} \end{cases} \tag{6.3.3}$$

for all $p \in A$.

6.3.4. Example (from Real Analysis). Consider

$$\begin{array}{rccc} \exp: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \mathrm{e}^x \end{array} \quad\text{and}\quad \begin{array}{rccc} s: & \mathbb{R}_{0+} & \longrightarrow & \mathbb{R}_{0+} \\ & x & \longmapsto & \sqrt{x} \end{array} . \tag{6.3.1}$$

Then exp is composable with s because a well known property of exp is that

$$\mathrm{e}^x > 0, \forall\, x \in \mathbb{R}, \tag{6.3.2}$$

which means that

$$\operatorname{Img} \exp \subseteq \mathbb{R}^+ \subseteq \mathbb{R}_{0+} = \operatorname{Dom} s. \tag{6.3.3}$$

We have

$$s \circ \exp(x) = \sqrt{\mathrm{e}^x} = \mathrm{e}^{x/2}, \forall\, x \in \mathbb{R}. \tag{6.3.4}$$

Note, that we are in a particular situatation where s is composable with exp too. Indeed, we have

$$\operatorname{Img} s = \mathbb{R}_{0+} \subseteq \mathbb{R} = \operatorname{Dom} \exp. \tag{6.3.5}$$

Furthermore

$$\exp \circ\, s(x) = \mathrm{e}^{\sqrt{x}}, \forall\, x \in \mathbb{R}_{0+}. \tag{6.3.6}$$

6.3.5. Remark (sufficient condition for composability). Since we have $\operatorname{Img} f \subseteq \operatorname{Cod} f$, a sufficient condition for f to be composable with g is

$$\operatorname{Cod} f \subseteq \operatorname{Dom} g. \tag{6.3.1}$$

6.3.6. Remark (composition is not commutative!) You may have noticed it, but just in case you did not, it is very important to realise that if f is composable with g, then g is not necessarily composable with f. Even when two functions are mutually composable, their compositions may be different. As an example, let

$$f : \mathbb{Z} \ni x \mapsto x^2 \in \mathbb{Z} \text{ and } g : \mathbb{Z} \ni x \mapsto 2x \in \mathbb{Z}. \tag{6.3.1}$$

In this case f and g are clearly mutually composable, since all domains and codomains are equal. However $f \circ g \neq g \circ f$. Indeed, we have

$$g \circ f(x) = g(f(x)) = 2x^2, \text{ and } f \circ g(x) = f(g(x)) = (2x)^2 = 4x^2, \tag{6.3.2}$$

for all $x \in \mathbb{Z}$. So if we take $x = 1$ we see that

$$g \circ f(1) = 2 \neq 4 = f \circ g(1), \tag{6.3.3}$$

hence $f \neq g$.

6.3.7. Theorem (composition is associative). *Let f, g, h be three function such that f is composable with g and g is composable with h. Then f is composable with $h \circ g$ and $g \circ f$ is composable with h. Furthermore*

$$h \circ (g \circ f) = (h \circ g) \circ f. \tag{6.3.1}$$

Proof From the assumption we have

$$\operatorname{Img} f \subseteq \operatorname{Dom} g \text{ and } \operatorname{Img} g \subseteq \operatorname{Dom} h. \tag{6.3.2}$$

Also, by the definition of composition we have

$$\operatorname{Dom}\big(h \circ g\big) = \operatorname{Dom} g. \tag{6.3.3}$$

It follows then that

$$\operatorname{Img} f \subseteq \operatorname{Dom}\big(h \circ g\big) \tag{6.3.4}$$

that is, f composable with $h \circ g$.

$$\operatorname{Cod}\big(g \circ f\big) = \operatorname{Cod} g \tag{6.3.5}$$

$$\operatorname{Img}\big(g \circ f\big) \subseteq \operatorname{Dom} h \tag{6.3.6}$$

and $g \circ f$ composable with h.
Introduce now

$$p := h \circ \big(g \circ f\big) \text{ and } q := \big(h \circ g\big) \circ f. \tag{6.3.7}$$

To show that $p = q$ it is enough to show that

$$x \in A \Rightarrow p(x) = q(x). \tag{6.3.8}$$

For each fixed $x \in A$ we have

$$\begin{aligned} p(x) &= h \circ (g \circ f)(x) = h(g \circ f(x)) = h(g(\underbrace{f(x)}_{=: y}) \\ &= h \circ g(y) = (h \circ g)(y) = h \circ g(f(x)) \\ &= (h \circ g) \circ f(x) = q(x), \end{aligned} \tag{6.3.9}$$

as desired. □

6.3.8. Theorem (invariance of distinguished functions under composition). *Suppose that $\phi : Q \to R$ is composable with $\psi : S \to T$ (i.e., $R \subseteq S$).*

(i) If ϕ and ψ are injective then $\psi \circ \phi$ is injective. This property is known as invariance of injectivity under composition.

(ii) If ϕ and ψ are surjective and $R = S$ then $\psi \circ \phi$ is surjective. This property is known as invariance of surjectivity under composition.

(iii) If ϕ and ψ are bijective and $R = S$ then $\psi \circ \phi$ is bijective. This property is known as invariance of bijectivity under composition.

Proof See Problem 6.X.4. □

6.3.9. Definition of identity function. Let us remind ourselves of the *identity* (function) on a set A. This is simply the function defined by

$$\begin{aligned} \mathrm{id}_A : \quad A &\to A \\ x &\mapsto x \end{aligned} . \tag{6.3.1}$$

This seemingly trivial map turns out to be quite useful.

6.3.10. Proposition (identity is neutral for composition). *If $f : A \to B$ is a function, then we have*

$$f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f. \tag{6.3.1}$$

Proof The proof (guess what?) is left as an exercise. □

6.3.11. Definition of inverse function. Suppose $f : A \to B$, we say that g is an *inverse* of f if and only if

$$g \circ f = \mathrm{id}_A \text{ and } g \circ f = \mathrm{id}_B. \tag{6.3.1}$$

If a function f has an inverse, we say that f is *invertible*. The inverse is usually denoted by f^{-1}.

Remark (notation!)**.** Note that the inverse function of f, if any, is denoted in the same way as the counterimage set function induced by f. Which meaning to give to f^{-1} depends then on the context. Note that the inverse function makes sense only when f is invertible, whereas the coutnerimage set function is always defined.
As if this was not confusing enough, the notation f^{-1} is also dangerous when inversion is used for number-valued functions. A typical example, that puzzles some students, rightly so, is $\cos^{-1} x$ (which means the invese of cos applied to x) and $\cos^2 x$ (which means the square of $\cos x$). As a hint, do not use the notation x^{-1} to indicate the division by x, unless strictly necessary and clearly understood from the context. Sticking to the notation $1/x$ guarantees more clarity.

6.3.12. Definition of partial inverse. Let $f : A \to B$ and $g : B \to A$.

⋆ We say that g is a *post-inverse* (also known as *right-inverse*) of f if and only if

$$g \circ f = \mathrm{id}_A. \tag{6.3.1}$$

⋆ We say that g is a *pre-inverse* (also known as *left-inverse*) of g if and only if

$$f \circ g = \mathrm{id}_B. \tag{6.3.2}$$

Comparing these definitions with Definition 6.3.11, g is an inverse of f if and only if g is both the post-inverse and pre-inverse of f.

6.3.13. Theorem (inverse function). *A function $f : A \to B$ is bijective if and only if it is invertible.*
Proof Suppose f is bijective, we will "construct" a function $g : B \to A$ which turns out to be its inverse. For each $y \in B$, since f is surjective, there exists at least one $x \in A$ such that $f(x) = y$. Because f is injective, there is only one such x. We pose then, by definition that

$$g(y) := x, \text{ where } x \in A \text{ is the unique element such that } f(x) = y. \tag{6.3.1}$$

Then $g : B \to A$ is clearly a function, it is composable with f and f is composable with g and, for each $y \in B$ and $x = g(y)$ we have

$$f(g(y)) = f(x) = y \text{ and } g(f(x)) = g(y) = x \qquad \text{(by definition of } g). \tag{6.3.2}$$

Thus f has an inverse, as required.
Conversely, suppose f has an inverse, f^{-1}, we want to show that f is bijective. Surjectivity first, let $y \in B$, consider $x := f^{-1}(y)$, then we have $f(x) = f(f^{-1}(y)) = y$, so y has x in its counterimage. For injectivity, suppose we have $f(x) = f(x')$, then applying f^{-1} to this element of B we obtain $x = f^{-1}(f(x)) = f^{-1}(f(x')) = x'$. □

6.3.14. Remark (inverting injections). An injection is close to being invertible. Indeed, it is a useful exercise to show that a map $\phi : A \to B$ is injective if and only if it has a post-inverse (also known as left-inverse). In fact, if one looks at the function $\hat{\phi} : A \to \operatorname{Img}\phi$, such that $\hat{\phi}(x) := \phi(x)$, then $\hat{\phi}$ is seen to be bijective, and thus invertible with inverse $\hat{\phi}^{-1} : \operatorname{Img}\phi \to A$. This construction is so useful, that we will use it and denote $\hat{\phi}^{-1}$ simply by ϕ^{-1}.

6.3.15. Composition as an algebraic operation. Consider the set of functions on S, $S^S := \{\phi : \phi : S \to S\}$ for a given set S; composition defines the following binary operation

$$\begin{array}{rccc} \circ: & S^S \times S^S & \to & S^S \\ & (\phi, \psi) & \mapsto & \phi \circ \psi \,. \end{array} \tag{6.3.1}$$

Theorem 6.3.7 tells us that this operation is associative. Definition 6.3.9 provides the neutral element, (also known as identity) for this operation

$$\mathrm{id}_S \circ \phi = \phi \circ \mathrm{id}_S = \phi. \tag{6.3.2}$$

Therefore S^S fulfills 2 of the 3 properties required to be a group (this is called by some a *monoid*) but not the third which requires every element to have an inverse. However if we restrict $\circ$ to $\operatorname{Bij}(S,S) = \operatorname{Sym}(S)$, the set of all bijective maps on S, then we obtain a group. This is known as the (general) symmetric group. In the case of finite S, the symmetric group is also denoted by S_n where $n := \#S$. (Note the different S for the set and S for the symmetric group S, not to be confused.)

6.3.16. Definition of composition power. Given a set S and a function ϕ thereon, $\phi : S \to S$, define, for each $n \in \mathbb{N}_0$, the n-th *composition power* of ϕ as follows

$$\phi^{\circ n} = \begin{cases} \mathrm{id}_S & \text{for } n = 0 \\ \phi \circ \phi^{\circ n-1} & \text{for } n \geq 1. \end{cases} \tag{6.3.1}$$

Thanks to the Recursion Theorem 6.4.13, the power function is a well-defined function. When there is no risk of confusion, the composition power is denoted by ϕ^n instead of $\phi^{\circ n}$.

6.4. Natural Numbers

In this section we will (finally) introduce rigorously the set of Natural Numbers, based solely on the axioms of Set Theory so far encounterd (plus a new one which make infinite sets exist).

We start with a review of what is it we want to obtain from our construction.

6.4.1. What are natural numbers? Natural numbers are characterised by *two structures.*

Algebraic structure: An *algebraic structure* is defined as a set S together with some operations on S.

Definition ((algebraic) operation)**.** An *operation* on a set S is a function

$$\rho : S \times S \to T \tag{6.4.1}$$

where T may or may not be S. If $T = S$ we say that the operation rho is *internal.*

In this Chapter we need only internal operations and we will assume all operations to be such without further notice.

The *sum* (also known as *addition*) is an operation on $\mathbb{N}$:

$$\begin{array}{rccl} \mathrm{sum}: & \mathbb{N}\times\mathbb{N} & \to & \mathbb{N} \\ & (x,y) & \mapsto & \mathrm{sum}(n,m)=n+m \end{array} . \tag{6.4.2}$$

(The notation $\mathrm{sum}(n,m)$ is a just wierd way of writing $n+m$ that "reminds" us that the sum is an operation according to our definition.)

As we know from Chapter 1, this operation comes with rules (laws) such as commutativity, associativity, etc, which need to be recovered in our construction of natural numbers.

Similarly the *product* (also known as *multiplication*) is an operation on $\mathbb{N}$ which enjoys also pleasant properties such as commutativity, associativity, etc. Also the product and the sum has common properties, such as distributivity.

In fact the sum and the product are the two basic operations in $\mathbb{N}$. (As you should realise by now the difference and division are not operations on $\mathbb{N}$. Can you say why?)

Ordering: Natural numbers come with the natural ordering $\leq$. We know that $\leq$ is transitive, reflexive, weakly antisymmetric and linear. This structure is essential to natural numbers (and then it trickles down to integers, rational and real—but not the complex—numbers).

Recalling sequences from Chapter 1, we note that the ordering of natural numbers is more than just a linear ordering. It has properties that even $\mathbb{Q}$ or $\mathbb{Z}$ do not enjoy. These properties are fundamental to "recognise" $\mathbb{N}$ from other ordered algebraic sets so let us list them briefly:

- $\star$ $(\mathbb{N},\leq)$ is *well-ordered* in that each non-empty subset of $\natural$ has a least element (also known as minimum). In particular, $\mathbb{N}$ itself has a "bottom" or first element, denoted by 1.[3]
- $\star$ $(\mathbb{N},\leq)$ is *discrete* in that each element n has a "next one", called *successor,n'*, such that $n<n'$ ($n\leq n'$ and $n\neq n'$) and there is no other number $m\in\mathbb{N}$ such that $n<m<n'$. (Of course, $n'=n+1$ but we are playing dumb in order not to get $+$ in the discussion, yet.)

Interaction: Algebra and ordering interact in $\mathbb{N}$ leading to a single "structure". These were discussed in Chapter 1 and you should review them to appreciate fully the rest of this section. The main properties are cancellation and monotonicity:

$$\forall\, a,b,c\in\mathbb{N}:(a\leq b \iff a+c\leq b+c) \text{ and } (a<b \iff ac\leq bc). \tag{6.4.3}$$

Any set X that has two operations that satisfy the basic properties of sum and product, that is well-ordered and whose order is linear, well-ordered and discrete has all the virtues of $\mathbb{N}$. Therefore such a set is identified with $\mathbb{N}$, i.e., $X=\mathbb{N}$. Our objective now is to build such a set, and from then on consider that set to *be* $\mathbb{N}$.

We have two ways of proceeding:

(a) either, we assume that such as set exists by outlining its basic properties,

[3] Many authors prefer to have 0 as the "first" element, and they build $\mathbb{N}_0$ instead of $\mathbb{N}$. There is nothing wrong about that approach, except the quite unnatural fact that the "first number" (or element number one if you will) is 0.

(b) or, we assume something apparently weaker (the Axiom of Infinity), and then, using the other axioms of Set Theory, we derive the properties postulated in (a).

It is more instructive to choose alternative (a) first and then by following in the footsteps of the Italian mathematician Luigi Peano. Once we know what we're looking for, we can tackle (b).

6.4.2. The origin of Peano's Axiom(s). Peano observed that in order to get the structure $(\mathbb{N},+,\times,\leq)$ with all its properties, you can start with a very economical set of rules (or axioms). It is a good idea to think that we are "walking" on $\mathbb{N}$.

Rule 1. You have to start somewhere. It is natural to start with 1 (a bit less natural, but equally legal, to start from 0).
Rule 2. You can always step forward by 1. That is, each element has a successor.

Observed facts:

Fact 1. You never cross one element more than once after starting.
Fact 2. If you repeat Rule 2 "endlessly", you get all of $\mathbb{N}$.

This is all nice and true, except it is a little bit vague. We want to phrase this in terms of sets and functions.

6.4.3. Definition of Peano's Axioms. A set-function (S,σ), where $\sigma : S \to S$ is said to satisfy *Peano's Axioms* for *natural numbers with* 0 and if and only if the following is true

PA1. the empty set is an element of S, $\varnothing \subseteq S$, in this context $\varnothing$ is called *zero* and denoted 0;[4]

PA2. $\sigma : S \to S$ is a well defined map on S, i.e.,

$$n = m \Rightarrow \sigma(n) = \sigma(m), \tag{6.4.1}$$

whereby, for each $n \in S$, $\sigma(n)$ is *the successor of* n (we say also that n is *a predecessor* of $\sigma(n)$, σ itself is called the *successor map*),

PA3. σ is injective, i.e., two numbers have the same successor only if they coincide (or each number has at most one predecessor), in symbols

$$\sigma(n) = \sigma(m) \Rightarrow n = m; \tag{6.4.2}$$

PA4. $0 \notin \operatorname{Img}\sigma$, i.e., the zero is not the successor of any element of S (or the zero has no predecessor);

PA5. if $X \subseteq S$ such that

$$0 \in X, \text{ and} \tag{6.4.3}$$

$$n \in X \Rightarrow \sigma(n) \in X \tag{6.4.4}$$

then

$$X = S. \tag{6.4.5}$$

[4] We are requiring $\varnothing$ to be an *element*, but not requiring $\varnothing$ to be a subset of S, which is always true.

6.4.4. Remark (successor map). In PA2 of §6.4.3, we have been a bit mysterious about the nature of the successor map σ, but in §4.7.10 we see a "concrete" example of such a map via the following *union-of-member-and-bracket-enclosed rule* (*UMBER*):

$$\sigma(n) := \operatorname{suc} n := n \cup \{n\}. \tag{6.4.1}$$

If n is a set, then by axioms 1–6 of Set Theory, $\operatorname{suc} n$ is a well-defined set. So v-NUMBERS make Peano's definition of an inductive set constructive, rather than just an abstract definition: for example, we see immediately that

$$0 := \varnothing \text{ is a set}, 1 := \operatorname{suc} 0 \text{ is a set}, 2 := \operatorname{suc} 1 \text{ is a set}. \tag{6.4.2}$$

We could continue this way "forever", or *ad infinitum* if you like Latin, but we have no axiom telling us that we can continue this way. In fact, two important issues that the axioms we have seen so far, from extension (Axiom 1) through to power (Axiom 6), cannot guarantee are:

(i) the construction of successor yields something new, i.e., for example that $n \neq \operatorname{suc} n$.
(ii) the existence of a set that contains all successor of any of its elements and 0.

For this we need new axioms. The answer to i will be given by the Axiom of Foundation (also known as Axiom of Regularity).

Axiom 7 (foundation)**.** *A nonempty set X must have a member x for which $x \cap X = \varnothing$. In logical notation, for any set X we have*

$$X \neq \varnothing \Rightarrow \exists\, x \in X : x \cap X = \varnothing. \tag{6.4.3}$$

6.4.5. Remark (all sets are regular). The Axiom of Foundation (Axiom 7) is also known as *Axiom of Regularity*. It ensures that any set X is *regular*, i.e., that X is not a member of itself. In logical notation this means that

$$\forall\, X \text{ set } (X \in X) \text{ is false}. \tag{6.4.1}$$

To see this, we argue by contraposition and show that $X \in X$ negates the Axiom of Regularity. Suppose that $X \in X$ for some set X, then this would imply that the successor of X, $X \cup \{X\}$ is the singleton $\{X\}$. This would make the following proposition true:

$$\forall\, x \in X : x = X, \tag{6.4.2}$$

whence, since $X \in X$

$$\forall\, x \in X : x \cap X = X. \tag{6.4.3}$$

But $X \neq \varnothing$ (because it is a singleton), therefore

$$\forall\, x \in X : x \cap X \neq \varnothing, \tag{6.4.4}$$

which negates the Axiom of Foundation.[5]

[5] For experts: the Axiom of Foundation is not really needed here, as it is possible to work by using PMI, after $\mathbb{N}_0$ and its arithmetic have been established. See Halmos (1974, end of §12)

6.4.6. Definition of inductive set. A set is X called *inductive* if and only if it statisfies

$$\varnothing \in X, \tag{6.4.1}$$

$$x \in X \Rightarrow \operatorname{suc} x \in X, \tag{6.4.2}$$

where suc is Von Neumann's successor (UMBER) map defined in (6.4.1).

Axiom 3 (infinity)**.** *There exists an inductive set.*

6.4.7. Remark. Note that Axiom of Existence (Axiom 3 on p.77) is now a subcase of the Axiom of Infinity (Axiom 3 on p.130) which is why the two axioms have the same number.

6.4.8. Theorem (definition of $\mathbb{N}_0$). *There exists a unique minimal inductive set, named $\mathbb{N}_0$. I.e., for any other inductive set X, we have $\mathbb{N}_0 \subseteq X$.*
Proof By the Axiom of Infinity there exists an inductive set, call it I. Consider the collection of all inductive subsets of I defined, by Axiom of Power and Axiom of Specification, as follows

$$\mathscr{I} := \left\{ J \in \wp(I) : \ J \text{ is inductive} \right\}. \tag{6.4.1}$$

After noting that $I \in \mathscr{I}$, we know that $\mathscr{I} \neq \varnothing$ and may consider its intersection

$$\mathbb{N}_0 := \bigcap_{J \in \mathscr{I}} J. \tag{6.4.2}$$

Now, we need to check that $\mathbb{N}_0$ is inductive. Firstly, $0 \in \mathbb{N}_0$, because 0 is member of all inductive sets, in particular to those J in $\mathscr{I}$. Second, we need to check that $x \in \mathbb{N}_0 \Rightarrow \operatorname{suc} x \in \mathbb{N}_0$. So assume $x \in \mathbb{N}_0$, this means that $x \in J$ for each $J \in \mathscr{I}$, and since each such J is inductive, we have $\operatorname{suc} x \in J$ for all $J \in \mathscr{I}$, whence $\operatorname{suc} x \in \mathbb{N}_0$, as required.
Finally, we need to show the uniqueness of minimal inductive sets like $\mathbb{N}_0$. Suppose K is another one, consider the set $J := K \cap I$ (where I is the set we have fixed at the beginning of the proof). Then $J \in \mathscr{I}$ (where $\mathscr{I}$ is defined by (6.4.1). By definition of $\mathbb{N}_0$ we have $\mathbb{N}_0 \subseteq J$, but $J \subseteq K$, hence (by transitivity of $\subseteq$) we get $\mathbb{N}_0 \subseteq K$. □

6.4.9. Theorem. *The set* $(\mathbb{N}_0, \operatorname{suc})$ *satisfies Peano's Axioms.*
Proof This is an intersting exercise which is left to your discretion, with an encouragment to read Halmos (1974, §12). □

6.4.10. Definition of set of natural numbers, unit. The set $\mathbb{N}_0$ called the *set of natural numbers with zero.* For many people this *is* the *set of natural numbers* $\mathbb{N}$, but we like to exclude 0: $\mathbb{N} := \mathbb{N}_0 \smallsetminus \{0\}$. The elements of $\mathbb{N}_0$ are called the *nonnegative integers*, and those of $\mathbb{N}$ the *natural numbers* or *positive integers*; we say *strictly positive integers* to exclude 0 when there is risk of ambiguity. The successor of 0 is called the *unit* and denoted by 1. From now on, we denote the successor of a number n with the usual arithmetic notation $n+1$, instead of $\operatorname{suc} n$.

6.4.11. Remark (successor and predecessor). From Peano's Axiom, we can immediately conclude the two following facts:

(a) Each number $n \in \mathbb{N}$ has a *successor* denoted by $\sigma(n)$ in Peano's Axioms and by $\operatorname{suc} n$ when using v-NUMBERS. This will be denoted by $n+1$ from now on.

(b) Each number $n \in \mathbb{N}$ that is *different than* 0 has a *predecessor,* denoted by $n-1$ such that

$$\operatorname{suc}(n-1) = n. \tag{6.4.1}$$

6.4.12. Remark (Principle of Mathematical Induction revisited). Note that PA5 is a reformulation of the Principle of Mathematical Induction 1.6.1. This means that all the results based upen the PMI, encountered earlier, e.g., ,the well-ordering principle (Theorem 2.3.5) can be now accepted as rigorous results.

6.4.13. Theorem (recursion). *Let X be a set and suppose a function $f_n : X \to X$ is given for each $n \in \mathbb{N}_0$. Let $a_0 \in X$, then there exists a function $a_\cdot : \mathbb{N}_0 \to X$ such that*

$$a_{\operatorname{suc} n} = f_n(a_n). \tag{6.4.1}$$

Proof This proof is not required for the exam and may be safely skipped at first reading; it is here for the truly curious. Before reading it, you need to understand the motivation for the theorem, while the process of construction is well-defined, it is not obvious that what is obtained is a sequence. That is why we need the Theorem, so we have to essentially show that there is a (unique) sequence that satisfies the recursive property.

We want to show that the sequence $(a_n)_{n\in\mathbb{N}_0}$ exists. Recall that a sequence is nothing but a function from $\mathbb{N}_0$ into X and that such a function is simply a graph from domain $\mathbb{N}_0$ into codomain X, i.e., a subset of the cartesian product $\mathbb{N}_0 \times X$, with the special property that for each element of the domain there corresponds exactly one element in the domain. We will build the function (sequence) $a_\cdot : \mathbb{N}_0 \to X$ as a subset of $\mathbb{N}_0 \times X$ and then verify that it sastisfies the function property. Consider the collection $\mathscr{B}$ of graphs b in $\mathbb{N}_0 \times X$ such that

$$(0, a_0) \in b \text{ and } (n, x) \in b \Rightarrow (n, f_n(x)) \in b. \tag{6.4.2}$$

Let a be the intersection of the collection $\mathscr{B}$, i.e.,

$$a := \bigcap_{b\in\mathscr{B}} b. \tag{6.4.3}$$

We argue by induction on n.

For the base case, we have $(0, a_0) \in a$ and $(0, a_0)$ is the only element in a of the form $(0, x)$ with $x \in X$. Indeed, suppose there was some $b_0 \neq a_0$ for which $(0, b_0) \in a$ with $b_0 \in X$, then taking $a' := a \smallsetminus (0, b_0)$ we obtain a graph a' in $\mathscr{B}$ such that $a' \subsetneq a$, which contradicts the fact that a is the intersection of $\mathscr{B}$.

To prove the inductive step, fix $n \in \mathbb{N}_0$ and suppose the inductive hypothesis that $(n, a_n) \in a$ and that

$$(n, x) \in a \Rightarrow x = a_n. \tag{6.4.4}$$

Then, by definition of $\mathscr{B}$ we know that any element in $b \in \mathscr{B}$ satisfies $(\operatorname{suc} n, f_n(a_n)) \in b$. Therefore $(\operatorname{suc} n, f_n(a_n)) \in a$ because a is the intersection of the collection $\mathscr{B}$. Furthermore, to prove uniqueness suppose $(\operatorname{suc} n, x) \in a$ with $x \neq a_{\operatorname{suc} n}$ then we have $x \neq f_n(a_n)$ and $a'' := a \smallsetminus \{(\operatorname{suc} n, x)\}$ belongs to $\mathscr{B}$ and $a'' \subsetneq a$, which is a contradiction with the fact that a is the intersection of $\mathscr{B}$. □

6.5. Arithmetic of natural numbers

Now that we have a rigorous notion of $\mathbb{N}_0$, we want to introduce algebra and ordering.

6.5.1. Definition of addition on $\mathbb{N}_0$. We start with *adding 1* to a natural number $n \in \mathbb{N}_0$ as being *by definition*

$$n + 1 := \operatorname{suc} n, \tag{6.5.1}$$

where suc is v-NUMBER's successor map defined by (4.7.7), and given a generic $m \in \mathbb{N}$ we define the *sum* (or *addition*) , recursively as follows:

$$n + \operatorname{suc} m := \operatorname{suc}(n + m). \tag{6.5.2}$$

This is operation is well-defined, using recursion given by Theorem 6.4.13.
It is now a useful, though straightfoward, exercise to prove (using induction) the two most familiar properties of the sum in $\mathbb{N}$: *associativity* and *commutativity.*

6.5.2. Definition of ordering in $\mathbb{N}_0$. Given two natural numbers n and m, we would like to derive a rule that can tell us whether $n \leq m$ or $m \leq n$. Again, recursion on one of the terms, say m, makes it not so hard.
If $m = 0$ then, *by definition* we put $0 \leq n$, for any n.
If $m \neq 0$ then we define

$$\begin{cases} n \leq m & \text{if } n \leq m - 1, \\ m \leq n & \text{if } n > m - 1, \end{cases} \tag{6.5.1}$$

where $m - 1$ is the predecessor of m defined in 6.4.11, and $a < b$ (or $b > a$, which is the same) is defined as $a \leq b$ and $a \neq b$.
After reading about orderings in Ch. 7.1, it becomes an exercise to show that $\leq$ is an ordering relation in $\mathbb{N}$.

6.5.3. Definition of multiplication in $\mathbb{N}$. Recursively we define

$$n \times m := \begin{cases} n \text{ if } m = 1 \\ n \times (m - 1) + m \text{ if } m > 1. \end{cases} \tag{6.5.1}$$

Following tradition, when "variables" are used, instead of digits, the sign × is usually omitted and thus $n \times m$ is written nm.
It is now an exercise to show all the remaining algebraic properties of $\mathbb{N}$, namely, that 1 is neutral, $mn = nm$, $m(np) = (mn)p$, the distributive laws, cancellation and invariance, as described in §1.2.4.[∗]

[∗]: Check!

6.6. Finite sets

Finally we are in good shape to introduce rigorously the concept of infinite and finite sets. We take the opportunity to give the concept of finite sets a deep cleaning.

6.6.1. Definition of finite, infinite. A set S is called *infinite* if and only if

$$\text{there is a map } \phi : S \to S \text{ such that } \phi \text{ is injective but is not surjective.} \tag{6.6.1}$$

We say that S is *finite* when it is not infinite.

6.6.2. Remark (rigorous and practical finiteness are the same). The main objective of this section is to show that for a set to be finite according to Definition 6.6.1 is equivalent to being finite according to the more intuitive characterisation of finite sets given in §5.1.1.

The main reason behind the Definition 6.6.1 of finite and infinite sets is that it allows us to proceed (in our theoretical thread of set theory) without needing the concept of number. In other words the concepts of finite and infinite sets are more primitive than those of number, which is why we call the Axiom of Infinity.

A comparison between Definition 6.6.1 and the Pigeonhole Principle 6.7.9, whose proof is also an objective of this section, makes it clear that a finite set is basically one that satisfies one direction the Pigeonhole Principle. We will show later that the second direction is also satisfied.

6.6.3. Theorem (Pigeonhole Principle (version 1)). *A set S is finite if and only if it satisfies the Pigeonhole Principle: for any map $\sigma : S \to S$ we have*

$$\sigma \textit{ injective } \Rightarrow \sigma \textit{ surjective.} \tag{6.6.1}$$

Proof This is simply a consequence of the definition of finite set in §6.6.1 as a set which is not infinite. Any map $\sigma : S \to S$ which is injective must therefore be surjective. □

6.6.4. Corollary (finite invertibility criterion). *If a set S is finite and $\sigma : S \to S$ is a map on S, then the following are equivalent*

(a) σ is injective
(b) σ is bijective
(c) σ is invertible.

Proof By the definitions of bijective and invertible map and the Inverse Function Theorem 5.3.22, the only nontrivial implication, which may not hold only for finite sets, is (a) ⇒ (b), but this is a direct consequence of the Pigeonhole Principle (Theorem 6.6.3). □

6.6.5. Lemma (finite set injection into finite sets). *A set S is finite if and only if for some finite set R there is an injection from S into R.*

Proof One direction of this equivalence is trivial: if S is finite, then take R to be S and for the injection pick the identity map on S, id_S.

Conversely suppose R is finite and let $\rho : S \to R$ be an injection and let $\rho^{-1} : \mathrm{Img}\,\rho \to S$ be the inverse of ρ on its image. To show that S is finite it is enough to show that each injection from S into itself must be sujective, i.e., have image all of S. So let us consider $\phi : S \to S$ injective, and let us show that ϕ is surjective. Build the map $\psi : R \to R$ as follows:

- ⋆ if $r \in \mathrm{Img}\,\rho$, let $s = \rho^{-1}(r)$ (uniquely defined by injectivity of ρ as in §6.3.14) and define
$$\psi(r) := \rho(\phi(s)) = \rho \circ \phi \circ \rho^{-1}(r), \tag{6.6.1}$$
- ⋆ if $r \in R \setminus \mathrm{Img}\,\rho$ define
$$\psi(r) := r. \tag{6.6.2}$$

We thus obtain a map $\psi : R \to R$ which is injective and satisfies $\psi(\operatorname{Img}\rho) \subseteq \operatorname{Img}\rho$.[∗] [∗]: Check! But R's being finite and the Pigeonhole Principle 6.6.3 imply that ψ is surjective. Therefore, by the invariance of surjectivity under composition, we have that ϕ, which coincides with $\rho^{-1} \circ \psi \circ \rho$, is surjective, as desired. □

6.7. Cardinality of finite sets

In this section we will characterise the notion of finite sets by connecting to the natural numbers and "counting" the sets. We will thus intrdouce the notion of cardinality and counting maps rigorously as promised in Chapter 5.

6.7.1. Lemma (finite set injection into numbers). *A set S is finite if and only if there exists an $n \in \mathbb{N}_0$ and an injection $\phi : S \to n$ (where we recall the shorthand $n := [0 \dots n-1]$).*
Proof □

6.7.2. Theorem (finite counting). *A set S is finite if and only if there exists an $n \in \mathbb{N}_0$ and a bijection, called a counting map, $s : [1 \dots n] \to S$, whereby $S = \{s_1, \dots, s_n\}$ with $i \neq j \Rightarrow s_i \neq s_j$. In this case, we say that the set counts to n. (This same statement is valid if the set $[1 \dots n]$ is replaced by the set $n = [0 \dots n-1]$.)*
Proof Consider the set

$$M := \left\{ m \in \mathbb{N}_0 : \exists\, \psi : S \to m \text{ injective} \right\}. \tag{6.7.1}$$

By Lemma 6.7.1 we know that $\varnothing \neq M \subseteq \mathbb{N}_0$, so by the well-orderging principle 2.3.5 for $\mathbb{N}_0$, M must have a minimum, let $n := \min M$. Since $n \in M$, we know that there exists a injective map $\phi : S \to [1 \dots n]$; we will prove that ϕ is surjective, so in fact bijective, and in view of the invertibility of bijective maps consider its inverse for the counting map $s := \phi^{-1} : [1 \dots n] \to S$.
Suppose ϕ is not surjective, then there is a $k \in [1 \dots n]$ for which $\phi(x) \neq k$, for any $x \in S$. We distinguish two cases:

Case 1. If $k = n$, then the map

$$\begin{array}{rccc} \hat{\phi}: & S & \to & [1 \dots n-1] \\ & x & \mapsto & \hat{\phi}(x) = \phi(x) \end{array} \tag{6.7.2}$$

is well-defined and, since ϕ is injective, $\hat{\phi}$ will also be injective. But this means that $n-1 \in M$ which contrasts with $n = \min M$.

Case 2. If $k < n$, then we define a "rewired" map $\psi(x) = \tau(\phi(x))$ where τ is the transposition of k with n on $[1 \dots n]$, i.e., $\tau : [1 \dots n] \to [1 \dots n]$ such that

$$\tau(i) := \begin{cases} n & \text{if } i = k, \\ k & \text{if } i = n, \\ i & \text{otherwise.} \end{cases} \tag{6.7.3}$$

[∗]: Check! A map such as τ is called a *transposition* of $[1 \dots n]$ and is bijective[∗], so that τ^{-1} exists and

$$\phi = \tau^{-1} \circ \psi. \tag{6.7.4}$$

The map ψ is injective, because ϕ and τ are so; it therefore satisfies the same conditions as ϕ in Case 1, and this leads also to a contradiction.

So ϕ is surjective and hence bijective. □

6.7.3. Remark (uniqueness of cardinality). The n in Theorem 6.7.2 is unique, in that, if for a (necessarily finite) set S there are $n, m \in \mathbb{N}_0$ and bijections $s : n \to S$, $r : m \to S$, then $n = m$. (But the map s itself may not necessarily equal r.) Indeed, without loss of generality, we may suppose $n < m$, which implies that $n, m \in M$, with M as in the proof of Theorem 6.7.2. Hence $m \neq \min M$. But a map $\sigma : S \to [1 \dots k]$ is bijective if and only if $k = \min M$.

6.7.4. Theorem (finite set cardinality). *A set S is finite if and only if there exists a unique $n \in \mathbb{N}_0$ for which there are (not necessarily unique) bijective maps $\sigma[1 \dots n]S$.*
Proof. □

6.7.5. Definition of finite cardinality. If S is finite, its *cardinality* is defined as the unique n be the integer of which in Theorem 6.7.4.

6.7.6. Proposition. *For each $n \in \mathbb{N}_0$ the following are true*

(i) the initial segment of $\mathbb{N}$, $[1 \dots n]$ is finite and has cardinality n.
(ii) n itself (which is by definition equal to $n-1 \cup \{n-1\} = [0 \dots n-1]$) is a finite set of cardinality n.

Proof See Problem 6.X.8. □

6.7.7. Natural numbers as model finite sets. A consequence of the above facts is that as far as counting finite sets goes, natural numbers are a sufficient copy of the those sets.

6.7.8. Problem. *Let $A = [1 \dots n]$, $B = [1 \dots m]$ and $f : A \to B$. Show the following:*
(a) If f is injective then $n \leq m$.
(b) If f is surjective then $n \geq m$.
(c) If f is bijective then $n = m$.
Hint. Use the inverse-image map $f^{-1} : \wp(B) \to \wp(A)$ to partition A into a disjoint union and then use the inclusion-exclusion principle to count the elements of the union thus formed.

6.7.9. Theorem (Pigeonhole Principle (version 2)). *Let S be a finite and $f : S \to S$; the function f is a surjection if and only if it is an injection.*
Proof Suppose $S = [1 \dots n]$ first.
Suppose f is surjective, we want to show that it is also injective. Suppose, by contradiction, that f is not injective, i.e., there exists $z \in S$ such that

$$\# f^{-1}\{z\} \geq 2. \tag{6.7.1}$$

Since S is finite and a subset of a finite set is finite, we have that $f^{-1}\{z\} \subseteq S$ is finite and taking its cardinality makes sense. But this implies

$$\begin{aligned} & n = \#S && \text{(by assumption)} \\ \big(\text{by basic properties of functions}\big) \quad & = \#\bigcup_{i\in S} f^{-1}\{i\} \\ \big(\text{ mutually disjoint union §5.5.6 }\big) \quad & = \sum_{i=1}^{n} \#f^{-1}\{i\} \\ & = \#f\{z\} + \sum_{i\in S\smallsetminus\{z\}} f^{-1}\{i\} \\ \big(\text{by (6.7.1) and surjectivity of } f\big) \quad & \geq 2 + (n-1) \\ & = n+1. \lightning \end{aligned} \tag{6.7.2}$$

So f must be injective.
Viceversa, suppose, also by contradiction, that f is injective but not surjective, i.e., there exists $y \in B$ such that $f^{-1}\{y\} = \varnothing$, then, proceeding as above we have

$$\begin{aligned} n & = \sum_{i=1}^{n} \#f^{-1}\{i\} \\ & = \#f^{-1}\{y\} + \sum_{i\in S\smallsetminus\{y\}} \#f^{-1}\{i\} \\ & = \sum_{i\in S\smallsetminus\{y\}} \#f^{-1}\{i\} \\ \big(\text{by } f\text{'s injectivity}\big) \quad & \leq n-1 \lightning\ . \end{aligned} \tag{6.7.3}$$

Now let S be a finite set, by the finite counting principle there is a counting map σ : $[\to \ldots 1]\, nS$ (that is, σ is a one to one correspondence, i.e., a bijection. □

6.8. Permutations

In §5.6.3 we have encountered permutations as a way of counting certain operations on a finite set. We now consolidate the concept of permutation as a bijection on a finite set.

6.8.1. Definition of (finite) permutation. Consider a finite set X, a (finite) *permutation* on X is a bijection from X into X.

6.8.2. Definition of symmetric group. We denote the set of all possible permutations of a set X, by $\operatorname{Sym}(X)$. This set can be shown to be a group when endowed with the operation of map composition $\circ$ (see § 6.8.4); it is thus known as the *symmetric group* of X. If $X = [1\ldots n]$ (or $[0\ldots n-1]$) we write $\operatorname{Sym}(n)$ for $\operatorname{Sym}([1\ldots n])$

6.8.3. Notation and examples. Consider the set $X = \{1,2,3\}$. The permutations of this set can be described with an array consisting of two rows and 3 columns. For example

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \Leftrightarrow \sigma(1) = 3,\ \sigma(2) = 2,\ \sigma(3) = 1. \tag{6.8.1}$$

The upper row is X, seen as the Dom σ while the lower row is, still X, but ordered consistenly so as the above element is its counterimage. Note that this same permutation is also described by

$$\sigma = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix} \tag{6.8.2}$$

Composition of permutations can be performed using this notation as follows, say

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix} \text{ and } \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} \tag{6.8.3}$$

then

$$\sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}, \tag{6.8.4}$$

where we compose from right-to-left, as with other maps.

6.8.4. Proposition (the symmetric group is a…group). *Recalling the definition of a group from Ch. 1§1.2.3 and considering the composition of maps operation* $\circ$*, introduced in §5.3.21 (see also §6.3.2), the set* $\operatorname{Sym}(X)$ *endowed with the operation* $\circ$ *constitutes a group.*

Proof First of all, noting that bijectivity is invariant with respect to composition it follows that $\operatorname{Sym}(X)$ is closed with respect to composition. Composition is associative and the identity map, being trivially bijective, is an element of $\operatorname{Sym}(X)$. Finally, since bijective maps are invertible, it follows that $\operatorname{Sym}(X)$ is a group. □

6.8.5. Proposition (symmetric group cardinality). *For any finite set* X *we have*

$$\#\operatorname{Sym}(X) = (\#X)!. \tag{6.8.1}$$

Proof We have already given the informal "cherry-picking" argument for this proof in §5.6.2. A more formal proof can be done by induction on $n := \#X$ and is left as an exercise. □

6.8.6. Definition of transposition, cycle. Let $I = [1\dots n]$ with $\#I(=n) \geq 2$, a *transposition* (also known as *swap*) is a permutation $\tau \in \operatorname{Sym}(n)$ that exchanges two different elements of I, say $j \neq k$, and leaves all the remaining is (if any) unchanged, i.e., for each $i = 1, \dots, n$

$$\tau(i) = \begin{cases} k & \text{if } i = j, \\ j & \text{if } i = k, \\ i & \text{if } j \neq i \neq k. \end{cases} \tag{6.8.1}$$

A transposition that swaps element j with k is denoted $\langle j \quad k \rangle$.

A *cycle* of length l is a permutation κ such that for a given sequence of mutually distinct elements $i_1, \dots, i_l \in I$, different, we have

$$\kappa(i_j) = \begin{cases} i_{j+1} & \text{for } j = [1\dots l-1] \\ i_1 & \text{for } j = l, \end{cases} \tag{6.8.2}$$

while $\kappa(i) = i$ if $i \neq i_j$ for all $j = 1, \dots, l$. Cycles are so important that they have a special notation, in *cycle notation* the cycle defined by (6.8.2) is denoted

$$\langle i_1 \quad \dots \quad i_l \rangle. \tag{6.8.3}$$

A cycle of length 1 is the identity. Transpositions coincide with the cycles of length 2. The set of elements $i_1, \dots, i_l$ covered by the cycle κ is known as the *orbit* of cycle κ.

6.8.7. Example (cycles).

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 2 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 2 & 4 & 3 \end{pmatrix} \tag{6.8.1}$$

6.8.8. Lemma (cycle decomposition of permutations). *Let X be a finite set and $n := \#X$. Each permutation $\sigma \in \operatorname{Sym}(X)$ is a product of cycles. Namely, for each $\sigma \in \operatorname{Sym}(X)$, there exist disjoint cycles $\kappa_1, \dots, \kappa_s \in \operatorname{Sym}(X)$, of lengths $l_1, \dots, l_s$, for some $s \leq n$ such that*

$$\sigma = \kappa_s \circ \cdots \circ \kappa_1 \text{ and } n = l_1 + \cdots + l_s. \tag{6.8.1}$$

Proof Proceed by induction on n. For $n = 1$ (the base case) the only permutation is the identity which is itself a cycle of length 1 and thus a product of cycles. Suppose now $n \geq 2$, then consider an iterative sequence $i_0, \dots, i_n$ obtained by applying σ successively up to n times:

$$\begin{aligned} i_0 &:= n, \\ i_1 &:= \sigma(n), \\ i_2 &:= \sigma\circ^2(n) := \sigma \circ \sigma(n) \\ &\dots, \\ i_n &:= \sigma\circ^n(n) := \sigma \circ \cdots \circ \sigma(n). \end{aligned} \tag{6.8.2}$$

Since the i_js are all elements of $[1 \dots n]$ and they number $n+1$, by the pigeonhole principle, two of them must coincide, say $i_j = i_k$ for some $k > j$ and without loss of generality we may assume $i_p \neq i_q$ for all p, q integers such that $j \leq p < q \leq k$. Then for $l := k - j$ we have $n = i_l = \sigma\circ^l(n)$ and all $i_1, \dots, i_l$ mutually distinct, it follows that $\kappa := \begin{pmatrix} i_1 & \dots & i_l \end{pmatrix}$ is a cycle that ends in n. Restricting σ to σ' on the remaining set $X' := X \smallsetminus \{i_1, \dots, i_l\}$ results in $\sigma' \in \operatorname{Sym}(X')$. But X', which has cardinality $n' := n - l$ smaller than n, is either empty ($n' = 0$) and σ consists of one cycle, or, by the inductive hypothesis, σ' must be decomposable as a product of $s - 1$ cycles, for some $s \geq 2$:

$$\sigma' = \kappa_{s-1} \circ \cdots \circ \kappa_1 \text{ and } n' = l_1 + \cdots + l_{s-1}. \tag{6.8.3}$$

Denoting $\kappa_s := \kappa$ and $l_s := l$, we obtain the result. □

6.8.9. Example. The proof of Lemma 6.8.8 provides an algorithm to decompose a premutation into cycles by "following orbits until exhaustion". For example, let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix} \tag{6.8.1}$$

then starting from 5 we obtain the sequence $\sigma(5) = 3$ and $\sigma(3) = 5$, which closes the first cycle as $\begin{pmatrix} 3 & 5 \end{pmatrix}$ and then starting from 4 we get $\sigma(4) = 1$, $\sigma(1) = 2$, $\sigma(2) = 4$, closing another cycle again as $\begin{pmatrix} 1 & 2 & 4 \end{pmatrix}$. We thus have

$$\sigma = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \circ \begin{pmatrix} 3 & 5 \end{pmatrix} \tag{6.8.2}$$

6.8.10. Lemma (transposition decomposition of cycles). *A cycle of length l can be decomposed into $l-1$ transpositions.*

Proof It is enough to note that

$$\left\langle i_1 \quad \dots \quad i_l \right\rangle == \left\langle i_1 \quad i_2 \right\rangle \circ \left\langle i_2 \quad i_3 \right\rangle \circ \cdots \circ \left\langle i_{l-1} \quad i_l \right\rangle. \tag{6.8.1}$$

(A proof of this can be done by induction on l.) □

6.8.11. Theorem (decomposition of permutation into transpositions). *If X is a finite set with $n := \#X$, each permutation $\sigma \in \operatorname{Sym}(X)$ can be written as the composition of n transpositions at most.*

Exercises and problems on functions

Problem 6.X.1 (basic properties of functions)**.** Let $f : A \to B$ be a function.

(a) Prove that
$$X \subseteq Y \subseteq A \Rightarrow f(X) \subseteq f(Y). \tag{6.X.1.1}$$
Show with an example that we may not replace the inclusion with strict inclusions, i.e., that the *following is false*
$$X \subsetneq Y \subseteq A \Rightarrow f(X) \subsetneq f(Y). \tag{6.X.1.2}$$

(b) Prove that
$$Y \subseteq Z \subseteq B \Rightarrow f^{-1}(Y) \subseteq f^{-1}(Z). \tag{6.X.1.3}$$
Prove that the same is true with strict inclusions.

(c) Prove that
$$\forall\, Y \subseteq B : f(f^{-1}(Y)) = Y. \tag{6.X.1.4}$$

(d) Prove that
$$\forall\, X \subseteq A : f^{-1}(f(X)) \supseteq X, \tag{6.X.1.5}$$
but that the equality may fail in general.

(e) Prove that
$$\forall\, X, Y \subseteq B : f^{-1}(X \cap Y) = f^{-1}(X) \cap f^{-1}(Y). \tag{6.X.1.6}$$

(f) Prove that
$$\forall\, X, Y \subseteq A : f(X \cap Y) \subseteq f(X) \cap f(Y), \tag{6.X.1.7}$$
but that the equality may fail in general.

(g) Prove that
$$\forall\, X, Y \subseteq B : f^{-1}(X \cup Y) = f^{-1}(X) \cup f^{-1}(Y). \tag{6.X.1.8}$$

(h) Prove that
$$\forall\, X, Y \subseteq A : f(X \cup Y) = f(X) \cup f(Y). \tag{6.X.1.9}$$

Exercise 6.X.2 (invariance of intersection under injection)**.** Suppose f is injective, prove that, for all $X, Y \subseteq \operatorname{Dom} f$
$$f(X \cap Y) = f(X) \cap f(Y). \tag{6.X.2.1}$$

Hint. Recall that one inclusion is valid for all functions in view of (6.X.1.7).

Problem 6.X.3 (natural partition of a function's domain)**.** Let $f : A \to B$ be a function, and consider the collection of subsets ot A
$$\mathscr{C} := \left\{ f^{-1}(y) : \ y \in B \right\}. \tag{6.X.3.1}$$

Prove that

(i) the collection $\mathscr{C}$ is mutually disjoint: i.e.,
$$\forall\, X, Y \in \mathscr{C} : X \neq Y \ \Rightarrow\ X \cap Y = \varnothing, \tag{6.X.3.2}$$

(ii) the union of all members of $\mathscr{C}$ gives A: i.e.,
$$\bigcup_{X \in \mathscr{C}} X = A. \tag{6.X.3.3}$$

A collection satisfying these properties is called a partition of the set A.

Problem 6.X.4 (invariance of distinguished functions under composition). (a) Prove the following invariance under composition theorem.

Suppose that $\phi : Q \to R$ is composable with $\psi : S \to T$ (i.e., $R \subseteq S$).

(i) If ϕ and ψ are injective then $\psi \circ \phi$ is injective. This property is known as invariance of injectivity under composition.

(ii) If ϕ and ψ are surjective and $R = S$ then $\psi \circ \phi$ is surjective. This property is known as invariance of surjectivity under composition.

(iii) If ϕ and ψ are bijective and $R = S$ then $\psi \circ \phi$ is bijective. This property is known as invariance of bijectivity under composition.

(b) Show with a counterexample, that $R = S$ is essential in the second statement.

Problem 6.X.5 (counting with distinguished functions). Let $A = [1 \dots n]$, $B = [1 \dots m]$ and $f : A \to B$. Show the following:

(a) If f is injective then $n \leq m$.

(b) If f is surjective then $n \geq m$.

(c) If f is bijective then $n = m$.

Hint. Use the inverse-image map $f^{-1} : \wp(B) \to \wp(A)$ to partition A into a disjoint union and then use the inclusion-exclusion principle to count the elements of the union thus formed.

Exercise 6.X.6 (addition on $\mathbb{N}$). Let $\mathbb{N}$ be the set of natural numbers as defined by Peano's Axioms and let $+$ be the addition on $\mathbb{N}$ defined in 6.5.1.

(a) Prove the associative law for $+$ on $\mathbb{N}$.

(b) Show that $\forall a \in \mathbb{N} : a + 1 = 1 + a$.

(c) Deduce the commutative law for $\mathbb{N}$ from the previous two facts.

Exercise 6.X.7 (cancellation for addition in $\mathbb{N}$). Let $a, b \in \mathbb{N}$, show that the following are equivalent:

(i) $a \leq b$,

(ii) $\forall c \in \mathbb{N} : \quad a + c \leq b + c$,

(iii) $\exists c \in \mathbb{N} : \quad a + c \leq b + c$.

Hint. You will find it easy to prove, in the following order: (i) ⇒ (ii), (ii) ⇒ (iii), (iii) ⇒ (i). (Why is this enough?)

Problem 6.X.8 (integer segments are finite). Recall the definition of *integer segment*

$$[m \dots n] := \{k \in \mathbb{Z} : \ m \leq k \leq n\}. \tag{6.X.8.1}$$

Also recall the definition of *finite set* as one, say X, that satisfied the Pigeonhole Principle, i.e., any injection on X must also be a surjection on X.

(a) Prove the following result from first principles.

For each $n \in \mathbb{N}_0$ the following are true

(i) the initial segment of $\mathbb{N}$, $[1 \dots n]$ is finite and has cardinality n.

(ii) n itself (which is by definition equal to $n - 1 \cup \{n - 1\} = [0 \dots n - 1]$) is a finite set of cardinality n.

(b) Deduce the following result. The segment $[m \dots n]$ is finite and has cardinality $n - m + 1$ when $n \geq m$ and 0 otherwise.

Problem 6.X.9 (inclusion–exclusion principle (disjoint case))**.** Let A and B be finite sets, $m := \#A$, $n := \#B$, such that $A \cap B = \varnothing$ then $A \cup B$ is a finite set and

$$\#(A \cup B) = \#A + \#B. \tag{6.X.9.1}$$

Deduce the baby case of the Inclusion–Exclusion Principle

Given two finite sets A, B we have

$$\#(A \cup B) = \#A + \#B - \#(A \cap B). \tag{6.X.9.2}$$

CHAPTER 7

Relations and algebraic extensions

Les mathématiciens n'étudient pas des objets, mais des relations entre les objets ; il leur est donc indifférent de remplacer ces objets par d'autres, pourvu que les relations ne changent pas.

Mathematicians do not study objects, but the relations between objects; to them it is a matter of indifference if these objects are replaced by others, provided that the relations do not change.

— Henri Poincaré, "La Science et l'hypothèse." (Science and hypothesis.)

7.1. Relations

7.1.1. Definition of relation. Given a set S, a *binary relation* ρ on S, is a graph from S into itself, i.e., $\rho \subseteq S \times S$. Although graphs are very useful in representing relationship between elements of a set S into a (possibly different) set T, the word *relation*, in most mathematical fields, is usually reserved to graphs where $S = T$. Recalling the definition of graph given in §4.8.5, a binary relation on S is therefore simply a subset of $S \times S$, the cartesian square of S. A *ternary relation* on S is defined as a subset of $S \times S \times S := S^3$, the cartesian cube, of S, and similarly, for any $n \in \mathbb{N}_0$, an n*-ary relation* on S is a subset of S^n. So a *unary relation* on S is, trivially, a subset of S and the only 0-ary relation on a set S is the empty relation $\varnothing$ (which the only subset of $S^0 = \{\varnothing\}$).
In many contexts such as graph theory and computer science a binary relation is also known as a *directed graph* on the set S. In mathematics, almost unanimously, one says *relation* when referring to a binary relation and so we shall do.
When we are talking about a given relation, say ρ on a set S, and two elements $x, y \in S$ are in the relation ρ, we mean that $(x, y) \in \rho$. It is customary to denote this situation by $x \rho y$ instead of $(x, y) \in \rho$. And when $(x, y) \notin \rho$ we write $x \not\rho\ y$.

7.1.2. Example ("divisor of" as a relation). Consider the relation $\rfloor$ on the set $\mathbb{N}$ whereby

$$n \rfloor m :\Leftrightarrow n \rfloor m. \tag{7.1.1}$$

We may sketch some elements of this relation in a table

$$\begin{gathered} 1 \rfloor 1, 1 \rfloor 2, 1 \rfloor 3, 1 \rfloor 4, \dots \\ 2 \rfloor 2, 2 \rfloor 4, 2 \rfloor 6, 2 \rfloor 8, \dots \\ 3 \rfloor 6, \dots \end{gathered} \tag{7.1.2}$$

In fact, it is futile to try to find all elements of $\rfloor$: it has infintely many ones. Just like sets, relations can be handled better with properties that describe them, rather than lists.

7.1.3. Example (coprime-with and cofactored-with). Consider for two integers a, b,

$$a \perp b :\Leftrightarrow (a, b) \text{ is a coprime pair,} \tag{7.1.1}$$

which is equivalent to say $\operatorname{hcf}\{a, b\} = 1$. Then $\perp$ is clearly a relation on $\mathbb{N}$ or $\mathbb{Z}$, we call it the *coprime-with relation.* Also the negation of $\perp$,

$$a \, \beta \, b :\Leftrightarrow \operatorname{hcf}\{a, b\} > 1, \tag{7.1.2}$$

defines β as a relation on $\mathbb{N}$ (or $\mathbb{Z}$). We say that β is the *complementary* of $\perp$ because

$$x \, \beta \, y :\Leftrightarrow x \not\perp y. \tag{7.1.3}$$

(Viewed as subsets of $\mathbb{N}^2$ or $\mathbb{Z}^2$, $\perp$ and β are effectively the complementary of each other.) A good name for β would be the *cofactored-with relation.*

7.1.4. Example (geometric relations). In a geometric context, consider the set $\mathscr{L}$ of all lines in the plane. We define the *orthogonality relation* (also denoted $\perp$) by

$$\forall \, l, k \in \mathscr{L} : l \perp k :\Leftrightarrow l \text{ is perpendicular to } k. \tag{7.1.1}$$

Another possible relation on $\mathscr{L}$ is *parallelism* defined by

$$\forall \, l, k \in \mathscr{L} : l \, || \, k :\Leftrightarrow l \text{ is parallel to } k. \tag{7.1.2}$$

7.1.5. Example (logical relations). Consider a set $\mathscr{P}$ of logical propositions. Then α is a relation on $\mathscr{P}$, meaning

$$(P \, \alpha \, Q) \text{ if and only if } (P \Leftrightarrow Q). \tag{7.1.1}$$

Another relation is β where

$$\big(P \, \beta \, Q\big) \text{ if and only if } (P \Rightarrow Q). \tag{7.1.2}$$

We will see later that α is an equivalence relation on $\mathscr{P}$ while β is an order relation. We could have used $\Leftrightarrow$ instead of α, but that would have been confusing at first approach. Equally, there is no need to use β, strictly speaking. In the following, we will use the symbol $\Leftrightarrow$ and $\Rightarrow$.

7.1.6. Example (square root). Consider the set $\mathbb{C}$ of all complex numbers. Define the relation

$$z \, \sigma \, w :\Leftrightarrow w^2 = z. \tag{7.1.1}$$

We have $\sqrt{2} \, \sigma \, 2$, but also $-\sqrt{2} \, \sigma \, 2$, $i \, \sigma \, -1, -i \, \sigma \, -1$. So σ is not a function on $\mathbb{C}$. This example shows that relations (and graphs) on a set S are more general objects than functions on S.

7.1.7. Example (ordering in $\mathbb{R}$). Consider the set $\mathbb{R}$, then $\leq$, $<$, $\geq$ and $>$ are all relations on $\mathbb{R}$. If $\mathbb{E}$ is a subset of $\mathbb{R}$, the “restriction” of these relations to $\mathbb{E}$ are also relations. As you may be suspecting, these relations are used quite extensively throughout mathematics, it is a good idea to keep an eye on them.

7.2. Special relations

Some relations, with certain properties, are more important than the rest in mathematics. We list here some of the properties that make certain relations more “special” than others.

7.2.1. Definition of reflexive property. A relation ρ on a set S is *reflexive* if and only if

$$\forall\, x \in S : x \,\rho\, x. \tag{7.2.1}$$

That is, each element of S is in relation ρ with itself.

7.2.2. Example (reflexive relations). Equality "=" on a set A is clearly a reflexive relation, as $x = x$ is true for all $x \in A$. Other reflexive relations are $\leq$, $\geq$, $\rfloor$ on $\mathbb{Z}$. Relations that are not reflexive in $\mathbb{Z}$ are $<$ and $>$. In $\mathscr{L}$ (the set of all lines in the plane) the parallelism $||$ is a reflexive relation (by convention a line is parallel to itself), whereas orthgonality $\perp$ is not reflexive (because the angle is not $\pi/2$). The square root relation σ from 7.1.6 is not reflexive: for example $2\,\sigma\,2$ is not true.

7.2.3. Remark (geometric interpretation). If we sketch a relation ρ on a set A on the *cartesian square* $A \times A$, we see that ρ is a reflexive relation if and only if the *diagonal set*

$$\varDelta(A \times A) := \big\{(x, y) \in A \times A :\ x = y\big\} = \{(x, x) :\ x \in A\} \tag{7.2.1}$$

is a subset of ρ. Note that the diagonal set, is nothing but the *identity,* as well as *equality,* on S.

7.2.4. Definition of symmetric property. A relation ρ on S is called *symmetric* if and only if

$$x \,\rho\, y \Rightarrow y \,\rho\, x. \tag{7.2.1}$$

7.2.5. Example (symmetric relations). Equality is symmetric on any set A. On $\mathbb{Z}$, none of $\leq$, $<$, $\geq$, $>$, or $\rfloor$ is symmetric. On the other hand co-primeness and cofactorness from Example 7.1.3 are both symmetric relations. On the set $\mathscr{L}$ of all straight lines in the plane (§7.1.4), both $\perp$ and $||$ are symmetric relations. The "square root" relation σ from 7.1.6 is not symmetric.

7.2.6. Definition of transitive property. Let ρ be a relation on a set S, we say that ρ is *transitive* if and only if

$$x \,\rho\, y \text{ and } y \,\rho\, z \Rightarrow x \,\rho\, z. \tag{7.2.1}$$

7.2.7. Example. Among the relations introduced in § 7.1, we have that the equality, =, on a set A is transitive. Also the all the usual orderings $<$, $\leq$, $\geq$, $>$ on $\mathbb{R}$ (or a subset thereof) are transitive. The dividing relation $\rfloor$ in $\mathbb{Z}$ is transitive.The relation $||$ on $\mathscr{L}$ from §7.1.4 is transitive, but $\perp$ is not (can you see why?). With reference to Example 7.1.3, the coprime relation, γ, is not transitive (example: $8\,\gamma\,15$ and $15\,\gamma\,16$, but $8 \not\gamma 16$). Also β is not transitive (write a counterexample on the margin to convince yourself).

7.2.8. Example (inclusion relation). Let A be a set, denote by $\wp(A)$ its power set. The inclusion of sets, $\subseteq$, is a relation on $\wp(A)$, indeed for any $X, Y \in \wp(A)$ (i.e., $X, Y \subseteq A$), it makes sense to ask whether $X \subseteq Y$ or not. It is worth checking that $\subseteq$ is reflexive and transitive, and, with a counterexample, to show that it is *not* symmetric.[∗]

[∗]: Check!

7.2.9. Example (strict inclusion relation). Another relation in $\wp(A)$, that is closely related (no pun intended) to $\subseteq$, is the *strict inclusion* $\subsetneq$ defined by

$$X \subsetneq Y :\Leftrightarrow X \subseteq Y \text{ and } X \neq Y, \forall X, Y \in \wp(A). \qquad (7.2.1)$$

[∗]: Check! Check that $\subsetneq$ is transitive, but not reflexive nor symmetric.[∗]

7.2.10. Example (equipotency). Of course, equality of sets is a relation on $\wp(A)$, and it clearly satisfies all the special relation properties listed so far. Another relation on $\wp(A)$ that is not as "boring" as equality is *equipotency*: we say, for two subsets $X, Y \subseteq A$, that X is *equipotent* to Y, and we write $X \rightleftarrows Y$ if and only if

$$\exists\, \phi : X \to Y : \phi \text{ is bijective.} \qquad (7.2.1)$$

For example, if $A = \mathbb{N}$ then the set of even numbers E and the set of odd numbers O are equipotent: a bijection $\phi : O \to E$ is given by

$$\phi(x) = x + 1, \text{ for } x \in O \qquad (7.2.2)$$

whose inverse is seen to be

$$\phi^{-1}(x) = x - 1, \text{ for } x \in E. \qquad (7.2.3)$$

Note that also E is equipotent to $\mathbb{N}$ and O is also equipotent to $\mathbb{N}$: build a one-to-one correspondence for each of the pairs $(\mathbb{N}, E)$ and $(\mathbb{N}, O)$. Show that the set of prime
[∗]: Check! numbers $P \subseteq \mathbb{N}$ is also equipotent to $\mathbb{N}$.[∗]
[∗]: Check! It is a good exercise to show that $\rightleftarrows$ is reflexive, symmetric and transitive.[∗]
An important property of equipotency is the following result.

Theorem. *If $X, Y \subseteq A$, A a given set, and X (or Y) is finite then $X \rightleftarrows Y$, if and only if Y (resp. X) is finite and they have the same cardinality.*

We shall prove this result in the next chapter.

7.2.11. Definition of antisymmetric property. A relation ρ on a set S is called *antisymmetric* if and only if

$$x \,\rho\, y \Rightarrow x \not\rho\, y. \qquad (7.2.1)$$

7.2.12. Example (antisymmetric relations). Of the relation so far introduced, those that are antisymmetric are $<$ and $>$ on $\mathbb{R}$. Also the *strict inclusion* $\subsetneq$ in $\wp(A)$, for a given set A, is antisymmetric.
Some relations who are definitely *not* antisymmetric are $=$ (on a set A), co-primeness and co-factorness, the geometric relations $\parallel$ and $\perp$ and the square root relation σ on $\mathbb{C}$ from Example 7.1.6.
Some relations that *barely fail antisymmetry* are $\leq$ and $\geq$ on $\mathbb{R}$ (or one of its subsets), $\rfloor$ in $\mathbb{N}$ and $\subseteq$ in $\wp(A)$ (for any given A).[1] This failure motivates the following definition.

[1] Our definition of ordering does make $\rfloor$ an ordering on $\mathbb{N}$, but not on $\mathbb{Z}$.

7.2.13. Definition of weak antisymmetry. A relation ρ on a set S is called *weakly antisymmetric* if and only if

$$x \,\rho\, y \text{ and } y \,\rho\, x \Rightarrow x = y. \tag{7.2.1}$$

7.2.14. Remark. If ρ is antisymmetric then ρ is weakly antisymmetric.[∗] But the converse is not true in general (provide a counterexample).[∗]

[∗]: Check!
[∗]: Check!

7.2.15. Exercise (antisymmetry and weak antisymmetry).

(a) *Show that if ρ is an antisymmetric relation on a set S then ρ must be weakly antisymmetric.*
(b) *Show, by using a counterexample, that the converse is not true, i.e., find a relation that is weakly antisymmetric but not antisymmetric.*

7.2.16. Remark. If ρ is symmetric then ρ is not antisymmetric.[∗] But the converse is not true in general (provide a counterexample).[∗]

[∗]: Check!
[∗]: Check!

7.2.17. Example (weakly antisymmetric relations). As observed, all antisymmetric relations seen so far are weakly antisymmetric. Let us check those who barely fail to be antisymmetric: the orderings $\leq$ and $\geq$ are weakly antisymmetric on $\mathbb{R}$, and so is $\subseteq$ on $\wp(A)$, for any given set A. As for $\rfloor$ it is weakly antisymmetric on as a relation on $\mathbb{N}$, but not as a relation on $\mathbb{Z}$ (so the "domain" of a relation matters when talking about its properties!).

7.3. Ordering relations

We have seen that ordering relationss like $<$, $\leq$, $\geq$, $>$ on $\mathbb{R}$ are special relations, in the sense that they satisfy some of the properties listed in §7.2. We make it more precise what we mean by an ordering (relation) in this section. As a word of warning, theories of orderings in the literature usually suffer from terminology disorder (pardon the pun). As a result the terminology we employ here may differ in some other books (including the textbook!). So please make sure you understand what we are talking about before you start solving an exercise.[2]

7.3.1. Definition of ordering. A relation ρ on S is an *ordering* if and only if

(1) ρ is transitive,
(2) ρ weakly antisymmetric,

[2] There will be no trick questions at the exam regarding definitions of orderings.

7.3.2. Example.
(a) Equality $=$ is an ordering (a trivial one and not very useful one indeed) on a set A.

(b) The relations $<$, $\leq$, $>$, $\geq$, are orderings on $\mathbb{R}$.

(c) The relations $\subseteq$, $\subsetneq$, $\supseteq$ and $\supsetneq$ are orderings on $\wp(A)$, for any given set A.

(d) The relation $\mid$ over the natural numbers $\mathbb{N}$ is an ordering. Note that $\mid$ is not an ordering over the integers $\mathbb{Z}$, as it fails to be weakly antisymmetric.

7.3.3. Definition of weak and strict orderings. Let ρ be an ordering on S.
(a) If ρ is also reflexive then we call it *nonstrict* (or *partial*) *ordering.*

(b) If ρ is (strongly) antisymmetric then we call it a *strict ordering.*
The terminology here introduced is justified by Proposition 7.3.4.

7.3.4. Proposition (a nonstrict ordering is not a strict ordering). *An ordering ρ on a set S cannot be partial (nonstrict) and strict simultaneously.*
Proof Exercise. □

7.3.5. Example.
(a) The equality on a set A is a partial ordering, but it is not a strict ordering (by Proposition 7.3.4).

(b) The orderings $\leq$ and $\geq$ are partial orderings on $\mathbb{R}$. The orderings $<$ and $>$ are strict on $\mathbb{R}$.

(c) The ordering $\mid$ on $\mathbb{N}$ is partial because $n \mid n$, for all $n \in \mathbb{N}$.

(d) The orderings $\subseteq$ and $\supseteq$ are partial orderings.

7.3.6. Proposition (strict orderings $\rightleftarrows$ partial orderings). *There exists a one-to-one correspondence between the set of partial orderings on S and the set of strict orderings on S. Namely:*
(a) if $\prec$ is a strict ordering on S, and we define

$$x \preceq y :\Longleftrightarrow x \prec y \text{ or } x = y, \text{ for } x, y \in S, \tag{7.3.1}$$

then $\preceq$ is a partial ordering on S;

(b) if $\preceq$ is a strict ordering on S, and we define

$$x \prec y :\Longleftrightarrow x \preceq y \text{ and } x \neq y, \text{ for } x, y \in S, \tag{7.3.2}$$

then $\prec$ is a strict ordering on S.

7.3.7. Definition of linear ordering. Let ρ be an ordering on S, we say that it is a *linear ordering*[3] if and only if any two different elements can be "compared" by ρ. In symbols, this property is written as

$$\forall\, x, y \in S : x \neq y \text{ and } x \not\rho\; y \Rightarrow x \,\rho\, y. \tag{7.3.1}$$

[3] Some authors use "total ordering" to refer to linear ordering. Since partial ordering may very well be linear (e.g., $\leq$ in $\mathbb{R}$), those authors are led to consider *total partial orderings.*

7.3.8. Definition of quasi-ordering. This definition is far from standard but it is quite handy. A relation ρ on a set S is called a *quasi-ordering* if and only if

(i) ρ is reflexive,

(ii) ρ is transitive.

7.3.9. Example. The relation $\rfloor$ on $\mathbb{Z}$ is a quasi-ordering which is not an ordering. It is reflexive and transitive but it is not weakly antisymmetric: e.g., $17 \rfloor -17$ and $-17 \rfloor 17$ but $17 \neq -17$.

7.3.10. Exercise (quasi-oredering quotienting). *Suppose $\prec$ is a quasi-ordering on a set S. Define the relation $\approx$ by*

$$x \approx y :\Leftrightarrow x \prec y \text{ and } y \prec x. \tag{7.3.1}$$

(a) Prove that $\approx$ is an equivalence relation.

(b) Prove that $\prec$ is compatible with $\approx$, i.e.,

$$x \approx x', y \approx y' \Rightarrow \left(x \prec y \Leftrightarrow x' \prec y'\right). \tag{7.3.2}$$

(c) Consider the set $\Sigma = S/\approx$. Prove that

$$X \preceq Y :\Leftrightarrow \left(x \in X, y \in Y \Rightarrow x \prec y\right) \tag{7.3.3}$$

defines a partial ordering relation on Σ.

7.3.11. Exercise (injection is quasi-ordering). *Let A be a given set, for each $X, Y \subseteq A$ we define*

$$X \hookrightarrow Y :\Leftrightarrow \exists\, \phi : X \to Y : \phi \text{ is injective.} \tag{7.3.1}$$

(a) Prove that $\hookrightarrow$ is a quasi-ordering on $\wp(A)$.

(b) What is the difference between $\hookrightarrow$ and $\subseteq$ on $\wp(A)$, if any?

(c) Explain what the sentence "$\subseteq$ is a subordinate relation to $\hookrightarrow$ " means.

7.4. Equivalence relations

Another type of special relation is constituted by equivalence relations. This is a very important type of relations as it allows many interesting "constructions" in mathematics. We will use equivalence relations to build number system extensions.

7.4.1. Definition of equivalence relations. A relation ρ on a set S is called an *equivalence relation* (or mmore briefly an *equivalence*) if

(i) ρ is reflexive,

(ii) ρ is symmetric,

(iii) ρ is transitive.

7.4.2. Example (modular congruences in $\mathbb{Z}$). Let $p \in \mathbb{N}$ and take $m, n \in \mathbb{Z}$. We say that m is *equal to n modulo p* or *n is p-congruent to m* if and only if $m-n$ is a multiple of p. In symbols:

$$\big(m = n\,(\operatorname{mod} p)\big) :\Leftrightarrow \exists\, k \in \mathbb{Z} : m - n = pk. \tag{7.4.1}$$

Sometimes we also write this as $m \equiv_p n$.

Exercise. *Show that for a fixed $p \in \mathbb{Z}^+$, congruence modulo p is an equivalence relation on $\mathbb{Z}$.*

Solution.

(i) Congruence modulo p is reflexive, indeed, for each $m \in \mathbb{Z}$ we have $m - m = 0 = kp$, with $k := 0$, thus $m = m\,(\operatorname{mod} p)$.

(ii) Congruence modulo p is symmetric. Suppose $m = n(\operatorname{mod} p)$, then $m-n = pl$ for some $l \in \mathbb{Z}$. It follows tht $n - m = pk$ with $k := -l$, which means that $n = m\,(\operatorname{mod} p)$.

(iii) Congruence modulo p is transitive. Suppose $m = n\,(\operatorname{mod} p)$ and $n = q\,(\operatorname{mod} p)$, then we have $m - n = pl$ and $n - q = pj$ for some $l, j \in \mathbb{Z}$. It follows that

$$m - q = m - n + n - q = pl + pj = pk, \tag{7.4.2}$$

with $k := l + j$.

To make this example more concrete, let $p = 5$, then:

$$(10 = 5\,(\operatorname{mod} 5)) \text{ is true,} \tag{7.4.3}$$

$$(11 = 5\,(\operatorname{mod} 5)) \text{ is false,} \tag{7.4.4}$$

$$(17 = 12\,(\operatorname{mod} 5)) \text{ is true,} \tag{7.4.5}$$

$$(-18 = 12\,(\operatorname{mod} 5)) \text{ is true}. \tag{7.4.6}$$

Proposition. *Let $p \in \mathbb{N}$, for any two numbers $m, n \in \mathbb{Z}$, denote by $(q_m, r_m) = \operatorname{div}(m, p)$ and $(q_n, r_n) = \operatorname{div}(n, p)$ (where* div *is the Euclidean Division defined in Ch.1§2.4.1) we have that*

$$m = n\,(\operatorname{mod} p) \Leftrightarrow r_m = r_n; \tag{7.4.7}$$

meaning that two elements are equal modulo p if and only if their remainders after
[∗]: Check! *division by p are equal.[∗]*

7.4.3. Example (parallel lines in plane). Relation $||$ defined in Example 7.1.4 is an
[∗]: Check! equivalence relation on $\mathscr{L}$.[∗]

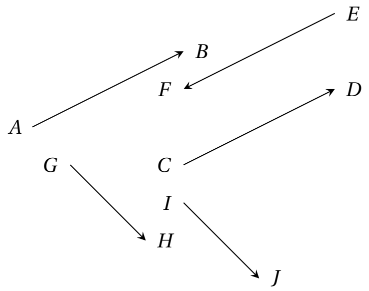
7.4.4. Example (equipollent segments in plane). Let Π be the set of all points in the plane. For $(A, B) \in \Pi \times \Pi$ by $\overrightarrow{AB}$ the *oriented segment* with beginning A and ending B. We say $\overrightarrow{AB}$ is *equipollent* or *(equal modulo translations)* to $\overrightarrow{CD}$, and we write $\overrightarrow{AB} \equiv \overrightarrow{CD}$, if and only if

(i) the segments have the same length $\left|\overrightarrow{AB}\right| = \left|\overrightarrow{CD}\right|$;

(ii) the segments lie on two parallel lines $\overrightarrow{AB} \parallel \overrightarrow{CD}$;

(iii) $\overrightarrow{AB}$ and $\overrightarrow{CD}$ have the same orientation.

Check that congruence of segments is an equivalence relation.



In the picture we have

- ⋆ $\overrightarrow{AB} \equiv \overrightarrow{CD} \equiv \overrightarrow{FE}$, but
- ⋆ $\overrightarrow{AB} \not\equiv \overrightarrow{EF}, \overrightarrow{GH}, \overrightarrow{IJ}$,
- ⋆ $\overrightarrow{GH} \equiv \overrightarrow{IJ}$,
- ⋆ etc.

7.4.5. Definition of equivalence class, residue class, congruence class. Let $\sim$ be an equivalence on S and let $x \in S$. An *equivalence class* for $\sim$ is a nonempty subset C of S, such that

$$x, y \in C \iff x \sim y. \tag{7.4.1}$$

The *equivalence class represented* (or *generated*) by x is the set

$$[x]_\sim = \left\{ y \in S : \ y \sim x \right\}, \tag{7.4.2}$$

which can be seen to be an equivalence class according to (7.4.1).
When there is no room for confusion (for example, if only one equivalence is involved in the discussion) the subscript is usually omitted and we write $[x]$ for the equivalence class represented by x.
For historical reasons, synonyms of equivalence class are *residue class* or *congruence class.*

7.4.6. Example (congruence modulo 5). In $\mathbb{Z}$, consider the equality modulo 5 (also known as congruence modulo 5), denoted $(\text{mod}\, 5)$. Then

$$[3] = \{\ldots, -12, -7, -2, 3, 8, 13, \ldots\}; \tag{7.4.1}$$

$$[4] \ni 4, -1, 9, -6, 14, -11; \tag{7.4.2}$$

$$[0] = [5] = [10]; \tag{7.4.3}$$

$$[1] = [-4]. \tag{7.4.4}$$

7.4.7. Proposition (properties of equivalence classes). *Let $\sim$ be an equivalence relation on S. The following hold true:*

(a) An equivalence class is non-empty:

$$\forall\, x \in S : \varnothing \subsetneq [x] \subseteq S. \tag{7.4.1}$$

(b) Equivalence classes are mutually disjoint:

$$x \sim y \iff [x] = [y] \iff [x] \cup [y] \neq \varnothing. \tag{7.4.2}$$

(c) The union of all equivalence classes equals S:

$$\bigcup_{x \in S} [x] = S. \tag{7.4.3}$$

Proof

(i) Fix an arbitrary $x \in S$. By the reflexive property of $\sim$, $x \sim x$. So (by definition of $[x]$) $x \in [x]$. By definition, $[x] \subseteq S$.

(ii) $[\Rightarrow]$: Fix $x, y \in S$ and assume that $x \sim y$. Noting that

$$\begin{aligned} u \in [x] &\Rightarrow u \sim x \\ &\Rightarrow u \sim y \qquad (x \sim y \text{ and } \sim \text{ transitive}) \\ &\Rightarrow u \in [y], \end{aligned} \tag{7.4.4}$$

it follows $[x] \subseteq [y]$. Similarly $[y] \subseteq [x]$, and thus $[x] = [y]$.

$[\Leftarrow]$: Suppose $[y] \cap [x] \neq \emptyset$, then there exists $z \in [x]$ and $z \in [y]$. This implies that $z \sim x$ and $z \sim y$, which, by the symmetric and transitive properties of $\sim$, imply that $x \sim y$.

(iii) For any $y \in S$, we have that $y \in [y]$ and thus $y \in \bigcup_{x \in S}[x]$; thus $S \subseteq \bigcup_{x \in S}[x]$. Conversely, by definition we have $[x] \subseteq S$, for any $x \in S$, so $\bigcup_{x \in S}[x] = S$.

□

7.4.8. Definition of partition of a set. Given a set S, a collection $\mathscr{S}$ of subsets of S is called a *partition* of S if and only if all the following hold:

(i) the collection is mutually disjoint:

$$\forall X, Y \in \mathscr{S} : X = Y \text{ or } X \cap Y = \emptyset, \tag{7.4.1}$$

which can be stated as the following equivalent condition

$$\forall X, Y \in \mathscr{S} : X \neq Y \Rightarrow X \cap Y = \emptyset; \tag{7.4.2}$$

(ii) the union of the collection equals the set S:

$$\bigcup_{X \in \mathscr{S}} X = S. \tag{7.4.3}$$

7.4.9. Example (partitions). Let $S = \{1, 2, 3, 4, 5, 6\}$. The following are partitions of S:

$$\{\{1,2,3\},\{4,5,6\}\}, \quad \{\{1\},\{2\},\{3\},\{4\},\{5,6\}\}, \quad \{\{1,2,3,4,5,6\}\}. \tag{7.4.1}$$

However $\{\{1,2\},\{2,3\},\{4\},\{5,6\}\}$ is *not* a partition of S, as the element 2 appears more than once.
Also $\{\{1,2\},\{3,4,5\}\}$ is not a partition of S because the union of all elements of $\mathscr{P}$ is not equal to S.

7.4.10. Remark (partition language). Result §7.4.7 could be restated by saying that each equivalence relation $\sim$ on a set S defines a partition given by the collection of all equivalence classes of $\sim$.

7.4.11. Example (citizenship as an equivalence). Consider

$$S = \left\{\text{Living people with exactly one citizenship}\right\} \tag{7.4.1}$$

and define

$$x \sim y :\Leftrightarrow x \text{ and } y \text{ are citizens of the same country,} \tag{7.4.2}$$

for all $x, y \in S$.
The relation $\sim$ is easily seen to be reflexive, symmetric and transitive. It is thus an equivalence on S. Each equivalence class consists of the (single-citizenship) nationals of a given county, and the equivalence classes partition the set S.

7.4.12. Example (congruence modulo $p \in \mathbb{N}$ in $\mathbb{Z}$). Let us consider now a more theoretical situation. On the set $\mathbb{Z}$, consider the relation

$$x \equiv_4 y :\Leftrightarrow x - y \text{ is a multiple of } 4. \tag{7.4.1}$$

This relation is also denoted as

$$x = y \,(\text{mod}\, 4), \tag{7.4.2}$$

which is pronounced "x equals (or is congruent to) y modulo 4".
This relation, called *congruence* (mod 4), is an equivalence. Indeed:
It is reflexive: fix $x \in \mathbb{Z}$, then $x - x = 0 = 4 \times 0$, so $x = x \,(\text{mod}\, 4)$.
It is symmetric: fix $x, y \in \mathbb{Z}$ such that $x = y \,(\text{mod}\, 4)$, then $x - y = 4k$ for some $k \in \mathbb{Z}$. Thus $y - x = 4(-k)$, which implies $y = x \,(\text{mod}\, 4)$.
It is transitive: fix $x, y, z \in \mathbb{Z}$ such that $x = y\,(\text{mod}\, 4)$ and $y = z\,(\text{mod}\, 4)$; then $x-y = 4k$ and $y - z = 4l$ for some $k, l \in \mathbb{Z}$. It follows that

$$x - z = (x - y) + (y - z) = 4(k + l), \tag{7.4.3}$$

hence $x = z \,(\text{mod}\, 4)$.
Let us have a look at the equivalence classes of (mod 4). We can express them in a concise way

$$\begin{aligned}
[0] &= \{0, 4, 8, 12, \ldots\} \cup \{-4, -8, -12, \ldots\} = \{4k : k \in \mathbb{Z}\}, \\
[1] &= \{1, 5, 9, \ldots\} \cup \{-3, -7, -11, \ldots\} = \{4k + 1 : k \in \mathbb{Z}\}, \\
[2] &= \{4k + 2 : k \in \mathbb{Z}\} = \{\ldots, -6, -2, 2, 6, \ldots\} \\
[3] &= \{4k + 3 : k \in \mathbb{Z}\} = \{\ldots, -5, -1, 3, 7, \ldots\} \\
[4] &= \{4k + 4 : k \in \mathbb{Z}\} \\
&= \{4k : k \in \mathbb{Z}\} = [0].
\end{aligned} \tag{7.4.4}$$

Indeed, for each $x = 0\,(\text{mod}\, 4)$ we have $[x] = [0]$, and similarly

$$[1] = [5] = [9] = [-3] = [-7], \ [2] = [6] = \text{etc.} \tag{7.4.5}$$

It looks as if there are only 4 equivalence classes for (mod 4), given by

$$[0], [1], [2], \text{ and } [3]. \tag{7.4.6}$$

We show now that this is indeed true. It is enough to show that for any $n \in \mathbb{Z}$ we have

$$\exists\, i \in [0 \ldots 3] : n \in [i]. \tag{7.4.7}$$

By the Euclidean Division, (2.4.1) in Chapter 2, we know that there exist $q, r \in \mathbb{Z}$ such that

$$n = 4q + r \text{ and } 0 \le r \le 3. \tag{7.4.8}$$

This means that $n - r = 4q$, i.e., $n = r \,(\text{mod}\, 4)$. Taking $i := r$ relation (7.4.7) becomes true.
It follows that $\mathbb{Z} = \bigcup_{i \in [0 \ldots 3]} [i]$.
The same type of equivalence can be formed by replacing 4 by any fixed number $p \in \mathbb{N}$; in this case one talks about the *congruence* $(\text{mod}\, p)$. For $p = 1$ we get the identity, which is a strivial equivalence relation. The set formed by the equivalence classes can be endowed with operations and it plays an important row in Algebra.

7.4.13. Definition of representative, transversal, quotient set. Given an equivalence $\sim$ on S, fix an equivalence class X. A *representative* of X is any $x \in S$ such that $[x] = X$.
A *transversal* is a set $T \subseteq S$ such that

(1) $\forall X$ equivalence class : $\exists x \in T : [x] = X$,
(2) $\forall x, y \in T : [x] = [y] \Rightarrow x = y$.

The set of all equivalence classes (which forms a partition of S by Theorem is called the *quotient set* of S with respect to $\sim$ and is denoted by $S/\sim$.

7.4.14. Example (representative, transversal, quotient set in politics). Consider set E of European people with only one citizenship can be endowed with the relation $\sim$ defined in §7.4.11, then $\sim$ is an equivalence on E. Furthermore:

- each *equivalence class* is formed by all the citizens (with no more than a dual citizenship) of the individual countries,
- the EU parliament is a set of representatives, but *not a transversal* (becasue more than one representative of a particular country may be present),
- the EU commission is a transversal (currently each country has exactly one commissioner), say
$$T = \{\text{Austrian Com., Belgian Com., Cypriot Com., etc.}\}, \tag{7.4.1}$$
- the quotient set is
$$E/\sim = \{\text{Austria, Belgium, Cyprus, etc.}\}, \tag{7.4.2}$$
where each country name stands for the set of citizens (with only one citizenship) of that country.

We note that the transversal and the quotient set look very much the same. Indeed, they are often considered as the "same" set in view of the following result. When two sets behave in the same identical way for all purposes, we say they are *isomorphic.*

7.4.15. Theorem (natural map and transversal-quotient isomorphism). *Let $\sim$ be an equivalence relation on a set S, the natural map (also known as canonical map) ϕ induced by $\sim$ on S, defined by*

$$\begin{array}{rccc} \phi: & S & \to & S/\sim \\ & x & \mapsto & [x] \end{array}, \tag{7.4.1}$$

is surjective. Furthermore, for each transveral T of $\sim$, the restriction of ϕ to T, defined by

$$\begin{array}{rccc} \phi\big|_T: & T & \to & S/\sim \\ & x & \mapsto & \phi(x) \end{array}, \tag{7.4.2}$$

is a bijection; such a map is called a transversal-quotient isomorphism. In particular the quotient set can be represented by any transversal:

$$S/\sim = \{[x] : x \in T\}. \tag{7.4.3}$$

Proof The proof is left as an exercise. □

7.4.16. Example (quotient of $\equiv_n$ in $\mathbb{Z}$). Consider the case $S = \mathbb{Z}$ and $\sim$ given by $\pmod 4$, defined in 7.4.12. Then the quotient set is

$$\mathbb{Z}/\equiv_4 = \{[0],[1],[2],[3]\} \tag{7.4.1}$$

A stranger, but equally valid, choice may be

$$\{[-14],[3],[0],[-3]\}, \tag{7.4.2}$$

or even

$$\{[-14],[2],[3],[0],[-3]\}, \tag{7.4.3}$$

but this one has some redundancy as $[2] = [-14]$.
Both $\{0,1,2,3\}$ and $\{-14,3,0,-3\}$ are transversals and the first two representations of the quotient set are not redundant. In fact these sets are transversals. On the other hand the set $\{-14,2,3,0,-3\}$ is not a transversal, because it has two elements, -14 and 2, which are equivalent.
Note that there are inifinitely many choices for transversals. For some "aesthetic" reason, which needs not be explained, the set $\{0,1,2,3\}$ is usuall taken to be the main (or canonical) transversal. In fact, most mathematicians will think of this set "being" $\mathbb{Z}/\equiv_4$, also denoted as $\mathbb{Z}/\pmod 4$, $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/4$, $\mathbb{Z}_4$ or simply by 4, owing to the tranversal-quotient isomorphism between this set and $\{0,1,2,3\}$ (the v-NUMBER 4).
The same arguments, including definitions and notations, of this example apply, with the appropriate modifications, to any congruence $\pmod n$, with $n \in \mathbb{N}$ replacing the 4.

7.4.17. Proposition (partition-induced equivalences). *Let S be a set and $\mathscr{P}$ a partition of S. The relation $\sim$ induced by $\mathscr{P}$ on S, defined by*

$$x \sim y :\Leftrightarrow \exists\, C \in \mathscr{P} : x \in C \text{ and } y \in C, \tag{7.4.1}$$

is an equivalence on S.
Proof The proof, which should not be too difficult, is left as an exercise. □

7.4.18. Example (sign equality is an equivalence on $\mathbb{R}$). A possible partition of $\mathbb{R}$ is by sign, i.e., $\{\mathbb{R}^+, \mathbb{R}^-, \{0\}\}$. It is not hard to see that this forms a partition of $\mathbb{R}$, as the sets are disjoint and any number in $\mathbb{R}$ is either positive, negative or 0.
The equivalence induced by this partition, is the one where two elements are equivalent if and only if they have the same sign.

7.4.19. Example (same argument as an equivalence on $\mathbb{C}$). The set $\mathbb{C}$ of complex numbers, viewed as the plane, can be partitioned into "rays" (straight half-lines) emanating from the origin. Let $\sim$ denote the equivalence induced by this geometric partition.
Two different numbers z and w are called equivalent if and only if they are non-zero and lie on the same ray, while the element 0 forms a partition element $\{0\}$ on its own and is equivalent only to itself.
Algebraically this can be translated as follows

$$z \sim w \Leftrightarrow \exists\, \lambda \in \mathbb{R}_{0+} : z = \lambda w. \tag{7.4.1}$$

Given a number $z = x + \mathrm{i}\, y \in \mathbb{C}$, we recall that its modulus (or absolute value) is that number $|z| \in \mathbb{R}_{0+}$ which satisfies

$$|z|^2 := x^2 + y^2. \tag{7.4.2}$$

Using this definition we see that

$$z \sim w \iff z = 0 = w \text{ or } \frac{z}{|z|} = \frac{w}{|w|}. \tag{7.4.3}$$

The second condition means, geometrically, that there exists a complex number α on the circle of radius one such that

$$z = |z|\alpha \text{ and } w = |w|\alpha. \tag{7.4.4}$$

We say that z and w have the same argument.[4]
Denote by $\mathbb{C}^* := \mathbb{C} \smallsetminus \{0\}$ and $\mathscr{C}$ the unit circle in the plane $\mathbb{C}$, the natural map of $\sim$ on $\mathbb{C}^*$ is the function

$$\begin{array}{ccc} \mathbb{C}^* & \to & \mathscr{C} \\ z & \mapsto & {}^{z}/_{|z|} \end{array}, \tag{7.4.5}$$

and the corresponding isomorphism is

$$\begin{array}{ccc} \mathbb{C}^*/\sim & \to & \mathscr{C} \\ [z] & \mapsto & {}^{z}/_{|z|} \end{array}, \tag{7.4.6}$$

which means that $\mathbb{C}^*/\sim$ can be thought as the unit circle.

7.4.20. Definition of compatible operations and relations. Let $\sim$ be an equivalence relation on a set S

(i) A *relation* σ on S is *compatible* with the equivalence $\sim$ if and only if

$$x\sigma y, x \sim x' \text{ and } y \sim y' \Rightarrow x'\sigma y'. \tag{7.4.1}$$

(ii) An *operation* on S,

$$* : S \times S \to S \tag{7.4.2}$$

is *compatible* with $\sim$ if and only if

$$x \sim x' \text{ and } y \sim y' \Rightarrow x * y \sim x' * y'. \tag{7.4.3}$$

7.4.21. Example (algebraic operations and congruence). The algebraic operations + and × are compatible with the congruence modulo p on $\mathbb{Z}$, for each fixed $p \in \mathbb{N}$. Indeed, suppose $x = x' \pmod n$ and $y = y' \pmod n$ then

$$x - x' = kn \text{ and } y - y' = ln, \text{ for some } k, l \in \mathbb{Z}. \tag{7.4.1}$$

Thus

$$(x + y) - (x' + y') = (k + l)n, \tag{7.4.2}$$

which means that

$$x + y = x' + y' \pmod n. \tag{7.4.3}$$

Therefore + is compatible with the congruence $\pmod n$.
Similarly we have

$$xy - x'y' = x(y - y') + (x - x')y' = (xl + ky')n, \tag{7.4.4}$$

which implies

$$xy = x'y' \pmod n. \tag{7.4.5}$$

Therefore × is also compatible with the congruence $\pmod n$.

[4] Technically speaking, the argument is that number $\theta \in [0, 2\pi)$ such that $z/|z| = \exp(\mathrm{i}\theta) = \cos\theta + \mathrm{i}\sin\theta$. An important result in Geometry is that such a theta exists for all $z \in \mathbb{C}^*$

7.4.22. Example (a multiplicative but not additive equivalence on $\mathbb{R}$). Consider the equivalence $\sim$ on $\mathbb{R}$ given by the sign equality (see Example 7.4.18).
Addition is *not compatible with* $\sim$: indeed, for $1 \sim 2$ and $-2 \sim -1$, but

$$1+(-2) = -1 \not\sim 1 = 2+(-1). \tag{7.4.1}$$

However, multiplication is compatible with $\sim$. Given $x \sim y$ and $x' \sim y'$, i.e.,

$$\operatorname{sign} x = \operatorname{sign} x' \text{ and } \operatorname{sign} y = \operatorname{sign} y'. \tag{7.4.2}$$

Hence, using the basic fact that

$$\forall\, a, b \in \mathbb{R} : \operatorname{sign}(ab) = \operatorname{sign} a \operatorname{sign} b, \tag{7.4.3}$$

we obtain

$$\operatorname{sign}(xy) = \operatorname{sign} x \operatorname{sign} y = \operatorname{sign} x' \operatorname{sign} y' = \operatorname{sign}(x'y'). \tag{7.4.4}$$

Finally, note that the ordering $\leq$ in $\mathbb{R}$ is compatible with $\sim$.

7.4.23. Proposition (quotient (also known as induced) operation). *Suppose $\sim$ is an equivalence on the set S and that $*$ is an operation on S that is compatible with $\sim$. Then the following rule*

$$\begin{array}{rccl} *: & (S/\sim)\times(S/\sim) & \to & S/\sim \\ & (X,Y) & \mapsto & X*Y := [x*y] \textit{ where } [x]=X \textit{ and } [y]=Y \end{array} \tag{7.4.1}$$

defines an operation $S/\sim$. This operation is called the quotient operation of $*$ through $\sim$ (or the induced operation by $*$ on $S/\sim$) and is usually denoted by the same symbol as $*$.

7.4.24. Example (plus and times tables in modular arithmetic). The set $\{0,1,2,3\}$ can be endowed with an addition and multiplication, induced by the corresponding operation in $\mathbb{Z}$. The restulting addition and multiplication tables in $\{0,1,2,3\}$ are given by

$$\begin{array}{c|cccc} + & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 1 & 2 & 3 \\ 1 & 1 & 2 & 3 & 0 \\ 2 & 2 & 3 & 0 & 1 \\ 3 & 3 & 0 & 1 & 2 \end{array} \quad \text{and} \quad \begin{array}{c|cccc} \times & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 3 \\ 2 & 0 & 2 & 0 & 2 \\ 3 & 0 & 3 & 2 & 1 \end{array}. \tag{7.4.1}$$

Note that though some properties of these operations carry over to 4, not all of them do. For example,

$$a \neq 0 \text{ and } ab = 0 \Rightarrow b = 0 \tag{7.4.2}$$

which is true for $a, b \in \mathbb{Z}$ is no longer true in 4 with the quotient multiplication, because $2 \times 2 = 0$ therein.

7.4.25. Exercise (plus and times tables in modular arithmetic). *Write down the multiplication table for $\mathbb{Z}/(\operatorname{mod} n)$, with $n = 2,3,5,6,7,8,9$. What do you observe? How about 10.*

7.4.26. Proposition (quotient relation). *Suppose that a $\sim$ is an equivalence relation on a set S and that σ is a relation that is compatible with $\sim$; denote by $\boldsymbol{\sigma}$ the quotient relation induced by σ on S/σ. If σ is one of the following:*

(a) reflexive
(b) symmetric
(c) antisymmetric
(d) weakly antisymmetric
(e) transitive

then so is $\boldsymbol{\sigma}$.

Proof The proof is left as an exercise. □

7.4.27. Example (sign equality as equivalence). Consider the set $\Sigma := \{-1,0,1\}$ which is isomorphic to $\mathbb{R}/\sim$, where $\sim$ is given by the sign equality:

$$x \sim y \iff \big(x y > 0 \text{ or } x = 0 = y\big). \tag{7.4.1}$$

[∗]: Check! Since $\leq$ and $\times$ in $\mathbb{R}$ are compatible with $\sim$,[∗] it follows that the set Σ can be endowed with a product $\times_\Sigma$ and and ordering $\leq_\Sigma$ induced by the multiplication and ordering of $\mathbb{R}$. Namely, we have $-1 \leq_\Sigma 0 \leq_\Sigma 1$ and

$$\begin{array}{r|rrr} \times_\Sigma & -1 & 0 & 1 \\ \hline -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 \end{array} \tag{7.4.2}$$

7.4.28. Exercise (complex multiplication and geometry). *Omit if you do not know yet complex numbers. Consider the set $\mathbb{C}^* := \mathbb{C} \smallsetminus \{0\}$ and the argument-equality equivalence $\sim$ relation given in 7.4.19. Denote by $\mathscr{C}$ the unit circle in $\mathbb{C}$; assume it is known that the circle $\mathscr{C}$ and the quotient set $\mathbb{C}^*/\sim$ are in a one-to-one correpsondence (bijection). Show that multiplication is compatible with $\sim$. Give a geometric description of the quotient multiplication on $\mathscr{C}$.*

Exercises and problems on relations

Exercise 7.X.1 (general relations). On the set $\mathscr{L}$ of all straight lines in the plane consider the relation

$$l \mathbin{\triangle} m :\Leftrightarrow l \text{ and } m \text{ form an (unoriented) angle of } \pi/3 \text{ radiants (i.e., 60 degrees).} \tag{7.X.1.1}$$

(a) State the condition that defines $\triangle$ to be reflexive. Is $\triangle$ reflexive? Explain.
(b) State the condition that defines $\triangle$ to be symmetric. Is $\triangle$ symmetric? Explain.
(c) Show (with a picture) that there are some lines x, y, $z \in \mathscr{L}$ such that

$$x \mathbin{\triangle} y,\ y \mathbin{\triangle} z \text{ and } x \mathbin{\triangle} z. \tag{7.X.1.2}$$

(d) State the condition that defines $\triangle$ to be transitive. Is $\triangle$ transitive? Explain.

Exercise 7.X.2 (ordering relations). (a) Let S be a set and ρ a relation on S. Give the conditions on ρ, explaining any terminology, that make it an *ordering relation* on S.

(b) Give an example of an ordering relation.
Prove your claim.

Exercise 7.X.3 (equivalence relations). (a) Let S be a non-empty set and ρ be a relation on S. When do we say that the relation ρ is an *equivalence* on S? Explain your terminology.

(b) Give an example of an equivalence relation, *different from the equality relation "$=$"*, on $\mathbb{Z}$.

Prove your claim.

Exercise 7.X.4 (partial orderings and equivalence). (a) Show that *the power set of a given set is partially ordered by inclusion.*

Hint. In symbols, consider a set X, take $Z := \wp(X)$ and $\varrho = \subseteq$, and show that ϱ is a partial ordering on Z.

(b) Show that $\subseteq$ is weakly antisymmetric.

(c) Show that $\subsetneq$ is a strict ordering of Z.

(d) For a mental picture of this situation, consider $X = \{a_1, a_2, a_3\}$ where $a_i \neq a_j$ for $i \neq j$. Draw a big diagram of the power set, $\wp(X)$, by representing each subsets with a point and labeling it properl. It is better to put the subsets with same cardinality in the same "column" as follows

$$\begin{array}{cccc}
 & \{a_1\} & \{a_1, a_2\} & \\
\varnothing & \{a_2\} & \{a_1, a_3\} & X \\
 & \{a_3\} & \{a_2, a_3\} &
\end{array}$$

Complete this diagram by drawing arrows that relate any two subsets $U, V \in \wp(X)$ for which $U \subsetneq V$. You will realise that some of these arrows are "consequences" of other arrows (by transitivity). Use different colours to represent different "levels" of arrows.

(e) Describe in words, how you can create from each partial ordering a strict ordering and viceversa.

Exercise 7.X.5 ("divides" orders $\mathbb{Z}$)**.** On the set of integers, $\mathbb{Z}$, consider the following relation

$$n \rfloor m :\Leftrightarrow \exists k \in \mathbb{Z} \colon m = kn. \tag{7.X.5.1}$$

(a) Show that $\rfloor$ is a quasi-ordering of $\mathbb{Z}$.

(b) Show that $\rfloor$ is neither weakly antisymmetric nor antisymmetric.

(c) Define the following relation on $\mathbb{Z}$

$$n \equiv m :\Leftrightarrow (n \rfloor m \text{ and } n \rfloor m). \tag{7.X.5.2}$$

Show that $\equiv$ is an equivalence relation on $\mathbb{Z}$.

(d) Describe the equivalence classes $[1]$ and $[3]$.

(e) Show that $\rfloor$ is *compatible* with $\equiv$, i.e.,

$$n \rfloor m \Leftrightarrow (\forall p, q \in \mathbb{Z} \colon p \equiv n \text{ and } q \equiv m \Rightarrow p \rfloor q). \tag{7.X.5.3}$$

(f) Consider the quotient set $\mathcal{N} = \mathbb{Z}/\equiv$ and show that δ given by

$$[n]\delta[m] :\Leftrightarrow n \rfloor m, \tag{7.X.5.4}$$

makes sense as a relation on $\mathcal{N}$.

Hint. You want to show that the definition of δ is independent of the choice of the representatives m and n.

(g) Can you think of a set that you know, also partially ordered by "divide", that "looks like" $(\mathcal{N}, \delta)$?

Exercise 7.X.6 (rational numbers)**.** (a) Given $q, r \in \mathbb{Q}$, show that there exists an infinite sequence of rational numbers (q_n), all different, such that $q < q_n < r$.

(b) State the Archimedean Property for the set of real numbers $\mathbb{R}$.

(c) Using the Archimedean Property show that for each $x \in \mathbb{R}^+$ there exists $n \in \mathbb{N}$ such that $x > 1/n$.

Problem 7.X.7 (extension of $\mathbb{Z}$ to $\mathbb{Q}$)**.** Following the footsteps of the exercises on the extension of $\mathbb{N}$ to $\mathbb{Z}$ (see Lecture Notes), design two similar exercises (and solve them) aimed at settling the question of solving the multiplication equation. Namely, peform the following steps.

(a) Identify a problem with the equations of the type

$$\text{Given } a, b \in \mathbb{Z},\ a \neq 0, \text{ find } x \in \mathbb{Z} : ax = b. \tag{7.X.7.1}$$

(b) What do you suggest to do, intuitively, in order to solve this problem.

(c) Draw up a step-by-step strategy in order to build a set that replaces $\mathbb{Z}$ in (7.X.7.1) and such that the equation becomes solvable.

(d) Explain why you think your strategy is going to work and highlight the most interesting parts of the proof.

Exercise 7.X.8 (antisymmetry and weak antisymmetry)**.** (a) Show that if ρ is an antisymmetric relation on a set S then ρ must be weakly antisymmetric.

(b) Show, by using a counterexample, that the converse is not true, i.e., find a relation that is weakly antisymmetric but not antisymmetric.

Exercise 7.X.9 (partial orderings cannot be strict). Prove the following statement. *An ordering ρ on a set S cannot be partial (nonstrict) and strict simultaneously.*

Exercise 7.X.10 (quasi-orderings and quotient). Suppose $\prec$ is a quasi-ordering on a set S. Define the relation $\approx$ by

$$x \approx y :\Leftrightarrow\ x \prec y \text{ and } y \prec x. \qquad (7.X.10.1)$$

(a) Prove that $\approx$ is an equivalence relation.

(b) Prove that $\prec$ is *compatible* with $\approx$, i.e.,

$$x \approx x', y \approx y' \Rightarrow \big(x \prec y \Leftrightarrow x' \prec y'\big). \qquad (7.X.10.2)$$

(c) Consider the set $\Sigma = S/\approx$. Prove that

$$X \preceq Y :\Leftrightarrow \big(x \in X, y \in Y \Rightarrow\ x \prec y\big) \qquad (7.X.10.3)$$

defines a partial ordering relation on Σ.

Exercise 7.X.11 (injection is a quasi-ordering). Let A be a given set, for each $X, Y \subseteq A$ we define

$$X \hookrightarrow Y :\Leftrightarrow\ \exists\, \phi : X \to Y : \phi \text{ is injective.} \qquad (7.X.11.1)$$

(a) Prove that $\hookrightarrow$ is a quasi-ordering on $\wp(A)$.

(b) What is the difference between $\hookrightarrow$ and $\subseteq$ on $\wp(A)$, if any?

(c) Explain what the sentence "$\subseteq$ is a subordinate relation to $\hookrightarrow$" means.

Exercise 7.X.12 (complex multiplication and geometry). *You need to know basics of complex numbers for this problem.*
Consider the set $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$ and the argument-equality equivalence $\sim$ relation given in 7.4.19. Denote by $\mathscr{C}$ the unit circle in $\mathbb{C}$; assume it is known that the circle $\mathscr{C}$ and the quotient set $\mathbb{C}^*/\sim$ are in a one-to-one correpsondence (bijection). Show that multiplication is compatible with $\sim$. Give a geometric description of the quotient multiplication on $\mathscr{C}$.

Problem 7.X.13. Prove the following statement.

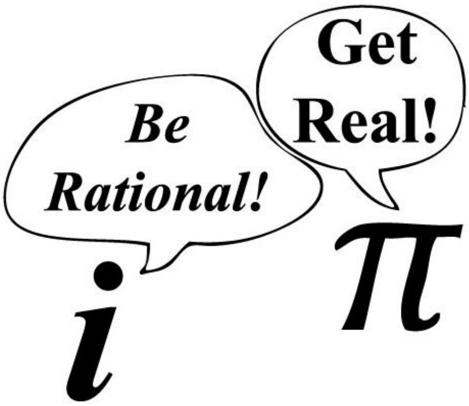
Let S be a set and $\mathscr{P}$ a partition of S. The relation $\sim$ induced by $\mathscr{P}$ on S, defined by

$$x \sim y :\Leftrightarrow\ \exists\, C \in \mathscr{P} : x \in C \text{ and } y \in C, \qquad (7.X.13.1)$$

is an equivalence on S.

CHAPTER 8

Real Numbers



The set of real numbers, $\mathbb{R}$, with whom most of us are already familiar from high-school mathematics satisfies some properties that characterise it (with respect to other sets of numbers). While it is possible to “build” $\mathbb{R}$, starting for example from $\mathbb{Q}$ and using *Dedekind cuts* or *Cauchy sequences* we prefer to approach $\mathbb{R}$ axiomatically by assuming it exists with the given properties. In a second moment, if needed, we will be returning to check how $\mathbb{R}$ can be built from scratch.

8.1. Why real numbers?

The reason for studying real numbers is one of convenience. While integers $\mathbb{Z}$ and rational numbers (also known as fractions of integers) $\mathbb{Q}$ are within the grasp of anyone with basic numeracy, measuring space and time imposes in a very natural way the need for (positive) real numbers but created a crisis in early number scholars (e.g., the Pythagoreans). An early example, provided by Pythagoras, is the length diagonal of a square of side 1, if there is to be a number, say s that measures the diagonal's length it must (thanks to Pythagoras's theorem) satisfisy

$$s^2 = 1^2 + 1^2 = 2. \tag{8.1.1}$$

It is however known, that there is no $s \in \mathbb{Q}$ that would satisfy (8.1.1). This caused consternation among early philosophers, who tried to build Mathematics on solid foundations and made it seem an impossible task; according to the legend the Pythagoreans guarded this lack of square root fact very jealously as a secret, punishing any transgressor with death. On a less violent note, the question can be easily solved by requiring (in our modern terms) the existence of a larger set of numbers $\mathbb{R}$ that includes $\mathbb{Q}$ as a proper subset, where numbers like s can exist.

Incidentally, while the Pythagoreans had their almost theological concerns about the reality of numbers, archeological findings (e.g., clay tablet *YBC7289*) show that in a neighbouring region of the planet, Babylonia, some had already established a numerical algorithm able to capture the value of s with very high accuracy. Of course, since they used fractions to represent numbers, this meant that the algorithm would require infinitely many iterations to yield the exact result (i.e., never), but at each iteration the algorithm would get “closer” to s.

The first rigorous construction of $\mathbb{R}$ was provided by *Richard Dedekind* by using what we know now as a *Dedekind cut* of $\mathbb{Q}$. Meanwhile, an earlier than Dedekind, real numbers were approached (literally) by using sequences of rational numbers and a rigorous approach to them can be founded on the concept of *Cauchy sequence* of rational numbers, popularised by nineteenth century mathematician *Augustin Cauchy.*

8.2. Ordered fields

8.2.1. Definition of (algebraic) field. A set $\mathbb{F}$, equipped with $+$ and $\times$, is called a field if and only if it satisfies the following properties:

- ⋆ $(\mathbb{F},+)$ is an *Abelian group*

$$(x+y)+z=x+(y+z), \tag{8.2.1}$$
$$\exists\, 0_{\mathbb{F}}\in\mathbb{F}: x+0_{\mathbb{F}}=x, \tag{8.2.2}$$
$$\forall\, x\in\mathbb{F}: \exists\, x'\in\mathbb{F}: x+x'=0_{\mathbb{F}}, \tag{8.2.3}$$
$$x+y=y+x; \tag{8.2.4}$$

  we write 0 (instead of $0_{\mathbb{F}}$) for the additive neutral element and call it *zero* and $-x$ for the *additive inverse* (or *opposite*) of x.

- ⋆ $(\mathbb{F}\smallsetminus\{0\},\times)$ is an Abelian group

$$(x\times y)\times z=x\times(y\times z) \tag{8.2.5}$$
$$\exists\, 1_{\mathbb{F}}\in\mathbb{F}: x\times 1_{\mathbb{F}}=x, \tag{8.2.6}$$
$$\forall\, x\in\mathbb{F}: \exists\, x'\in\mathbb{F}: x\times x'=1, \tag{8.2.7}$$
$$x\times y=y\times x \tag{8.2.8}$$

  we omit the × sign for multiplication when one or both of the factors are variables:

$$xy:=x\times y, \tag{8.2.9}$$

  we write 1 (instead of $1_{\mathbb{F}}$) and say *one* or *unit* or *identity* for the multiplicative neutral element, and we write $\frac{1}{x}$, $1/x$ or x^{-1} for the *multiplicative inverse* (also known as *reciprocal*) (the motivation behind the notation x^{-1} will be clear when we look at powers and exponentials);

- ⋆ $\times$ is distributive with respect to $+$, i.e.,

$$x(y+z)=xy+xz, \tag{8.2.10}$$

  where × has *operational precendence* over + to avoid overbracketting and write concisely the right hand side:

$$xy+xz:=(x\times y)+(x\times z); \tag{8.2.11}$$

8.2.2. Definition of ordered field. A field $(\mathbb{F},+,\times,<)$ is said to be an *ordered field* if the ordered set $(\mathbb{F},<)$ is a linearly ordered set

- ⋆ $(\mathbb{F},<)$ is a *strongly ordered set,* that is, for all $x,y,z\in\mathbb{F}$ we have

$$x<y \Rightarrow y\not< x \tag{8.2.1}$$

  and

$$x<y \text{ and } y<z \Rightarrow x<z; \tag{8.2.2}$$

- ⋆ the order $<$ makes $\mathbb{F}$ a *linearly ordered set* (also known as *totally ordered set*), in that it satisfies *trichotomy*, which the the property that for any pair of elements $x, y \in \mathbb{F}$ exactly one of the following must occur:

$$x < y,\ x > y \text{ or } x = y, \tag{8.2.3}$$

(where $x > y$ means $y < x$);

and furthermore the order $<$ is *compatible with the field structure* of $\mathbb{F}$, meaning that

- ⋆ the order $<$ is *compatible with respect to addition* $+$:

$$a < b \Rightarrow a + c < b + c, \tag{8.2.4}$$

- ⋆ the order $<$ is *compatible with respect to multiplication* $\times$:

$$c > 0 \Rightarrow (a < b \Rightarrow ac < bc), \tag{8.2.5}$$

- ⋆ $\mathbb{F}$ is nontrivial and *positively oriented*:

$$0 < 1. \tag{8.2.6}$$

8.2.3. Proposition (ordering's compatibility with product rules). *The following are equivalent*

(a) $a < b$
(b) $ac < bc$ *for all* $c > 0$
(c) $ac < bc$ *for some* $c > 0$

Proof See Problem 8.X.17. □

8.2.4. Proposition (product and ordering properties). *Show, from first principles (i.e., the axioms of ordered field), the following.*

Let $\mathbb{F}$ *be an ordered field that is positively oriented (i.e.,* $1 > 0$*) and denote by* -1 *the opposite (also known as additive inverse) of the unit* 1*. Then the following properties hold:*

(a) *For any* $a \in \mathbb{F}$ *we have* $-a = (-1)a$.
(b) *If* $a > 0$ *then* $-a < 0$.
(c) *If* $a < b$ *then* $-a > -b$.
(d) *If* $a, b, c \in \mathbb{F}$, $a < b$ *and* $ac < bc$ *then* $c > 0$.
(e) *If* $a, b, c \in \mathbb{F}$, $a < b$ *and* $ac > bc$ *then* $c < 0$.

Proof See Problem 8.X.18. □

8.2.5. Definition of nonstrictly ordered field. Recalling that any strictly ordered set $(S, \prec)$ can be turned into a nonstrictly ordered set $(S, \preceq)$ by defining

$$x \preceq y \Leftrightarrow x \prec y \text{ or } x = y. \tag{8.2.1}$$

It turns out $\preceq$ is a nonstrict ordering, i.e., reflexive, weakly antisymmetric and transitive if and only if $\prec$ is a strict ordering. Therefore any ordered field $(\mathbb{F}, +, \times, <)$ has a nonstrict ordering $\leq$ defined

$$x \leq y \Leftrightarrow x < y \text{ or } x = y. \tag{8.2.2}$$

From now on we shall be using $<$ or $\leq$ an ordered field $\mathbb{F}$ as connected this way on without further apology.

8.2.6. Definition of positive and negative. Given an element x of an ordered $\mathbb{F}$ we say that x is

- ⋆ (nonstrictly) *positive* if $x \geq 0$
- ⋆ *strictly positive* if $x > 0$
- ⋆ (nonstrictly) *negative* if $x \leq 0$
- ⋆ *strictly negative* if $x < 0$.

Many texts, especially North American ones, employ "positive" (resp. "negative") where we say "strictly positive" (resp. "strictly negative") and "nonnegative" (resp. "nonpositive") where we say "(nonstrictly) positive" (resp. "(nonstrictly) negative"). When confusion may arise the safest bet is to use the universally agreed upon symbols $\leq 0, \geq 0, < 0, > 0$.

We use the following notation for subsets of $\mathbb{F}$

$$\begin{aligned}\mathbb{F}_{0+} &:= \{x \in \mathbb{F} : \ x \geq 0\}\\ \mathbb{F}_{+} &:= \{x \in \mathbb{F} : \ x > 0\}\\ \mathbb{F}_{0-} &:= \{x \in \mathbb{F} : \ x \leq 0\}\\ \mathbb{F}_{-} &:= \{x \in \mathbb{F} : \ x < 0\}.\end{aligned} \tag{8.2.1}$$

8.2.7. Proposition (sign invariance under reciprocation). *If $x \in \mathbb{F}_+$, for an ordered field $\mathbb{F}$ then ${}^1\!/\!_x \in \mathbb{F}_+$.*

Proof Take $b = x$, $a = 0$ and $c = {}^1\!/\!_x$ in 8.X.18d. Since $x > 0$ and $x({}^1\!/\!_x) = 1 > 0 = 0({}^1\!/\!_x)$, it follows that ${}^1\!/\!_x > 0$. □

8.2.8. Definition of absolute value, modulus. Given an ordered field $\mathbb{F}$, the *absolute value* (also known as *modulus*) of an $x \in \mathbb{F}$

$$|x| := \begin{cases} x \text{ if } x \geq 0 \\ -x \text{ if } x < 0. \end{cases} \tag{8.2.1}$$

The function $\mathbb{F} \ni x \longmapsto |x| \in \mathbb{F}_{0+}$.

8.2.9. Proposition (properties of the absolute value). *Let $\mathbb{F}$ be an ordered field. The absolute value on $\mathbb{F}$ satisfies the following properties:*

- ⋆ *positivity*

$$|a| \geq 0 \textit{ for each } a \in \mathbb{F} \tag{8.2.1}$$

- ⋆ *definiteness*

$$|a| = 0 \Rightarrow a = 0 \tag{8.2.2}$$

- ⋆ *homogeneity*

$$|ab| = |a||b|, \tag{8.2.3}$$

- ⋆ *subadditivity (also known as triangle inequality)*

$$|a + b| \leq |a| + |b|. \tag{8.2.4}$$

8.2.10. Definition of interval. Let $\mathbb{F}$ be an ordered field we define the *open interval* in $\mathbb{F}$ with end-points a and $b \in \mathbb{F}$ to be the set

$$\mathbb{F} \cap (a, b) := \{x \in \mathbb{F} : \ a < x < b\} \tag{8.2.1}$$

and the *closed interval* in $\mathbb{F}$ with end-points a and $b \in \mathbb{F}$ to be the set

$$\mathbb{F} \cap [a, b] : \{x \in \mathbb{F} : \ a \leq x \leq b\}. \tag{8.2.2}$$

We also define the *open–closed interval* to be

$$\mathbb{F} \cap (a, b] := \{x \in \mathbb{F} : \ a < x \leq b\} \tag{8.2.3}$$

and the *closed-open interval* to be

$$\mathbb{F} \cap [a, b) := \{x \in \mathbb{F} : \ a \leq x < b\}. \tag{8.2.4}$$

We also use the following *open half-bounded intervals* (also known as *open half-lines*)

$$\mathbb{F} \cap (a, \infty) := \{x \in \mathbb{F} : \ x > a\} \text{ and } \mathbb{F} \cap (-\infty, b) := \{x \in \mathbb{F} : \ x < b\} \tag{8.2.5}$$

of which we already know $\mathbb{F}_+ = \mathbb{F} \cap (0, \infty)$ and $\mathbb{F}_{0+} = \mathbb{F} \cap [0, \infty]$, and the following *closed half-bounded intervals* (also known as closed half-lines)

$$\mathbb{F} \cap [a, \infty) := \{x \in \mathbb{F} : \ x \geq a\} \text{ and } \mathbb{F} \cap (-\infty, b] := \{x \in \mathbb{F} : \ x \leq b\} \tag{8.2.6}$$

8.2.11. Proposition (an ordered field is infinite). *An ordered field $\mathbb{F}$ is infinite. Namely, there is a subset of $\mathbb{F}$ that is isomorphic to $\mathbb{N}_0$, i.e., there exists a map $\phi : \mathbb{N}_0 \to \mathbb{F}$ such that for any $m, n \in \mathbb{N}_0$ we have*

$$\phi(n + m) = \phi(n) + \phi(m) \tag{8.2.1}$$

$$\phi(nm) = \phi(n)\phi(m) \tag{8.2.2}$$

$$n \leq m \Rightarrow \phi(n) \leq \phi(m). \tag{8.2.3}$$

Proof *Omit this proof at first reading.*
Let us indicate, in this proof, the neutral elements of $\mathbb{F}$ by $0_\mathbb{F}$ and $1_\mathbb{F}$, while 0 and 1 indicate the zero and unit in $\mathbb{N}_0$. Consider the map $\phi : \mathbb{N}_0 \to \mathbb{F}$ defined by

$$\phi(0) := 0_\mathbb{F}, \text{ and } \phi(n + 1) := \phi(n) + 1_\mathbb{F} \text{ for } n \in \mathbb{N}. \tag{8.2.4}$$

We claim that ϕ is injective. We note first that for all $n \in \mathbb{N}$ we have $\phi(n) > 0$, this can be seen by induction on n: for $n = 1$ we have $\phi(1) = 1_\mathbb{F} > 0_\mathbb{F}$, and for an integer $n > 1$ we have $\phi(n) = \phi(n - 1) + 1_\mathbb{F} > \phi(n - 1) > 0$ by the inductive hypothesis. Again by induction on n we can show that addition $+$ is invariant under the map ϕ, i.e.,

$$\phi(m + n) = \phi(m) + \phi(n) \tag{8.2.5}$$

for any $m, n \in \mathbb{N}_0$. For $m \in \mathbb{N}_0$ and $n = 0$, (8.2.5) is satisfied since $\phi(0) = 0_\mathbb{F}$ and this proves the base case. Assuming the inductive hypothesis that $\phi(m + n - 1) = \phi(m) + \phi(n - 1)$, we deduce

$$\phi(m+n) = \phi(m+(n-1)+1) = \phi(m+n-1) + 1_\mathbb{F} = \phi(m) + \phi(n-1) + 1_\mathbb{F} = \phi(m) + \phi(n). \tag{8.2.6}$$

To prove injectivity, suppose $\phi(n) = \phi(m)$ for some $n, m \in \mathbb{N}_0$ we want to show $n = m$. Without loss of generality, we may suppose that $m \leq n$ (otherwise exchange them), and we proceed by induction on n. If $n = 0$ then the integer m satisfies $0 \leq m \leq n = 0$

and hence $m = 0$. If integer $n \geq 1$, since $m \leq n$, we may write $n = m + d$, $d \geq 0$, and, by invariance of $+$ under ϕ, we get

$$\phi(n) = \phi(m+d) = \phi(m) + \phi(d) = \phi(m), \tag{8.2.7}$$

whence $\phi(d) = 0$, so $d = 0$ and $m = n$. We have thus checked ϕ is injective; since $\operatorname{Dom}\phi = \mathbb{N}_0$ and $\operatorname{Cod}\phi = \mathbb{F}$, it follows that $\mathbb{F}$ is infinite. Furthermore, the map ϕ esablishes a one-to-one correspondence that conserves addition between $\mathbb{N}_0$ and the image set $N_0 := \phi(\mathbb{N}_0)$. To finish the proof and show that ϕ is an isomorphism between $\mathbb{N}_0$ and N_0, we still have to show that

$$\phi(mn) = \phi(m)\phi(n). \tag{8.2.8}$$

But again, this can be done by induction and the definition of ϕ. □

8.2.12. Characteristic of a field. *May be omitted at first reading*
In technical jargon, the proof of §8.2.11 shows that an ordered field has characteristic 0. Where the *characteristic* of a field $\mathbb{F}$ (or a unitary ring,[1] for that matter) is defined as

$$\max\{0, \min\{n \in \mathbb{N} : n1_{\mathbb{F}} = 0\}\}. \tag{8.2.1}$$

where $n1_{\mathbb{F}}$ is the n-th additive power defined by

$$nx := x+^n := \underbrace{x+\cdots+x}_{n \text{ times}} \text{ for any } x \in \mathbb{F}. \tag{8.2.2}$$

Note that the additive power defines (on any ring or field[2]) the concept of multiplication by an arbitrary integer (it can be extended to negative integers by definining $(-1)x := -x$ for any element of the ring or field. Thanks to this, we could have defined the set N_0, built in the proof of 8.2.11, as

$$N_0 := \{n1_{\mathbb{F}} : n \in \mathbb{N}_0\}. \tag{8.2.3}$$

Also the homomorphism ϕ built in the proof, turns out to be a good way to compute the characteristic of a ring, by looking at its image.

8.3. Bounded sets, extremums and best bounds

Analysis is for many the art of bounds. We develop the basic concepts about bounds, in ordered fields, in this section, by reviewing some known concepts such as maximum and minimum, which go under the common name of extremum and then look at least upper bounds and smallest lower bound.

8.3.1. Definition of upper and lower bound in an ordered field. Given a subset S of an ordered field $\mathbb{F}$ and $x \in \mathbb{F}$ we say that x is an *upper bound* (respectively *lower bound*) of S if and only if

$$y \in S \Rightarrow y \leq x. \tag{8.3.1}$$

In words an upper bound, resp. lower bound, on a set S is simply any number that is bigger, resp. smaller, than all the numbers in S.

[1] A unitary ring is a ring with unit. E.g., $\mathbb{Z}$ is a ring (of characteristic 0). Also $\mathbb{Z}/n\mathbb{Z}$, for $n \in \mathbb{N}$, is a ring (of charactersitic n).

[2] Fields are a special type of rings.

8.3.2. Example (upper bound). Consider the set $S = \{1/n : n \in \mathbb{N}\}$ then 1, 2, and any element $x \in \mathbb{Q}$, such that $x \geq 1$ is an upper bound on S.

8.3.3. Exercise (upper bound's transitivity). *Suppose $x \in \mathbb{F}$, $S \subseteq \mathbb{F}$ and x is an upper (lower) bound on S then any $y \in \mathbb{F}$ greater (smaller) than x is also an upper (lower) bound on S.*

8.3.4. Definition of extremum. Let S be a linearly (or totally) ordered set, e.g., S is the subset of an ordered field $\mathbb{F}$. Recall the definition of $\max S$, the *maximum* of S, already given in §2.5.1 as the largest element of S, if any (as many sets may not have a maximum).

Same for $\min S$: the *minimum* of S, this is the smallest element of the set S, if any (like for a maximum, many linearly ordered sets may not have a minimum).

An *extremum* of S is defined as an element of S that is either a maximum or a minimum.

8.3.5. Exercise (maximum is an upper bound).

(i) Show the following statement:

> *The maximum* $\max S$*, whenever it exists, is an upper bound on S.*

Note however, with counterexamples, that

(ii) An upper bound on $S \subseteq \mathbb{F}$ is not necessarily a maximum of S.

(iii) Sets that are bounded above need not have a maximum.

Similar facts hold for minimum, $\geq$ and lower bound replacing maximum, $\leq$ and upper bound, like-with-like.

8.3.6. Definition of least upper bound and greatest lower bound. Let $\mathbb{F}$ be an archimedean field and let $S \in \mathbb{F}$ we define the *least upper bound* (also known as *supermum*) and *greatest lower bound* (also known as *infimum*), respectively as

$$\begin{aligned} \overset{\mathbb{F}}{\sup} S &:= \min\left\{y \in \mathbb{F} : y \text{ upper bound on } S\right\}, \\ \text{and } \overset{\mathbb{F}}{\inf} S &:= \max\left\{y \in \mathbb{F} : y \text{ lower bound on } S\right\}. \end{aligned} \tag{8.3.1}$$

If the field $\mathbb{F}$ is known from the context we omit the superscript on the sup and inf symbols. In particular, we never write $\mathbb{R}$, as that is the de facto default set which makes the concepts of supremum and infimum work best, as we shall see in 8.6.

8.3.7. Exercise (basic facts about supremum and infimum).

(a) Argue that by the definition of minimum and maximum, that if a $\sup^{\mathbb{F}}$ *(or* $\inf^{\mathbb{F}}$*) exists then it is unique.*

(b) Prove that $\beta = \sup S$ is equivalent to the following two properties

  (a) if $x \in S$ then $x \leq \beta$,

  (b) if λ is an upper bound on S, then $\lambda \geq \beta$.

(c) Show the following.

> *Let S be a subset of an archimedean field $\mathbb{F}$ $\beta := \sup S$ if and only if*
>
> $$\forall\, \varepsilon \in \mathbb{F}_+ : \exists\, x \in S : \beta < x + \varepsilon, \tag{8.3.1}$$

this property is known as the property (of $\sup S$*) of being the greatest closure point (of the set S).*

Likewise, $\alpha := \inf S$ *if and only if*

$$\forall\, \varepsilon \in \mathbb{F}_+ : \exists\, x \in S : \alpha > x - \varepsilon, \tag{8.3.2}$$

i.e., the infimum is characterised as the lowest closure point.

8.3.8. Proposition (uniqueness of the least upper bound). *Let* $\mathbb{F}$ *be an archimedean field (e.g.,* $\mathbb{F} = \mathbb{Q}$ *or* $\mathbb{R}$*). Each subset* S *of* $\mathbb{F}$ *has at most one least upper bound.*
Proof Suppose x_1 and x_2 are upper bounds of the set S, we want to show $x_1 = x_2$. Since x_1 is an upper bound on S, and x_2 a least upper bound we have $x_2 \le x_1$. Exchanging roles we have $x_1 \le x_2$, and thus $x_1 = x_2$. □

8.3.9. Definition of archimedean field. An ordered field $(\mathbb{F}, +, \times, <)$, which has a subset that is isomorphic to $(\mathbb{N}_0, +, \times, <)$, which we may thus identify with $\mathbb{N}_0$, is called *archimedean* if and only if it satisfies

$$\forall\, x \in \mathbb{F} : \exists\, N \in \mathbb{N}_0 : x \le N. \tag{8.3.1}$$

8.3.10. Lemma (vanishing constant). *Let* $\mathbb{F}$ *be an archimedean field and let* $x \in \mathbb{F}$ *such that* $0 \le x < \varepsilon$ *for any* $\varepsilon \in \mathbb{F}_+$ *then* $x = 0$.
Proof We prove the contrapositive. Suppose $x > 0$ then $1/x > 0$ and by the archimedean property we know that there exists $N \in \mathbb{N}$ such that $N \ge 1/x$, therefore $\varepsilon := \frac{1}{N} \le x$ and $\varepsilon = 1/N \in \mathbb{F}_+$. □

8.4. Rational numbers

8.4.1. Definition of $\mathbb{Q}$, the set of rational numbers. Recall first that $\mathbb{Q}$ is defined as the set of fractions of the form n/m, with $n \in \mathbb{Z}$ and $m \in \mathbb{N}$, such that

$$\frac{n}{m} \le \frac{l}{k} \iff nk \le lm, \tag{8.4.1}$$

$$\frac{n}{m} + \frac{l}{k} = \frac{nk + lm}{mk} \tag{8.4.2}$$

and

$$\frac{n}{m} \times \frac{l}{k} = \frac{nl}{mk} \tag{8.4.3}$$

for all $n, l \in \mathbb{Z}$ and $m, k \in \mathbb{N}$.
In rigrous terms $\mathbb{Q} = \mathbb{Z} \times \mathbb{N} / \equiv$ where the equivalence relations is given by

$$(n, m) \equiv (l, k) \iff nk = lk, \tag{8.4.4}$$

and n/m is a notation for the equivalence class generated by (n, m). It is possible to show that $\mathbb{Q}$, thus defined is a field that contains all integers with the same operations.

8.4.2. Corollary (rationals are embedded in all ordered fields). *If $\mathbb{F}$ is an ordered field, then there exits a subset Q of $\mathbb{F}$ that is isomorphic to $\mathbb{Q}$. Namely, there is an injective map $\psi : \mathbb{Q} \to \mathbb{F}$ such that*

$$\psi(q+p) = \psi(q) + \psi(p) \tag{8.4.1}$$

$$\psi(qp) = \psi(q)\psi(p) \tag{8.4.2}$$

$$q \le p \Rightarrow \psi(q) \le \psi(p). \tag{8.4.3}$$

When these three properties are verified we call the map ψ a homomorphism. Because ψ is injective, it is bijective onto its image and and we call it an isomorphism.
Proof From 8.2.11 we know that there exists an injective map $\phi : \mathbb{N}_0 \to \mathbb{F}$ that preserves algebraic operations and ordering. Consider the map ϕ and define

$$\psi(q) = \begin{cases} \phi(q) & \text{if } q \in \mathbb{N}_0 \\ -\phi(-q) & \text{if } q \in \mathbb{Z}_- \\ {}^{\phi(n)}/_{\phi(m)} & \text{if } q = {}^{n}/_{m}. \end{cases} \tag{8.4.4}$$

It is an exercise to show ψ satisfies the three properties (8.4.1)–(8.4.3) for an ordered field homomorphism and from that deduce that ψ is injective. □

8.4.3. Proposition (rationals are archimedean). *The set $\mathbb{Q}$ is an archimedean field.*
Proof We already know that $\mathbb{Q}$ is an ordered field which contains $\mathbb{N}_0$ (strictly speaking a copy thereof) as a subset. The subset of $\mathbb{Q}$ given by equals that of fractions ${}^{n}/_{1}$ with $n \in \mathbb{N}_0$. All we need to show is that for each $q \in \mathbb{Q}$ there exists $N \in \mathbb{N}_0$ such that $q \le N$. Suppose $q = {}^{n}/_{m}$, $n \in \mathbb{Z}$, $m \in \mathbb{N}$, if $n < 0$ then picking $N = 0$, we have $q \le N$. If $n \ge 0$ then let $N := n$, since $m \in \mathbb{N}$ we have

$$n \le mn = mN, \tag{8.4.1}$$

and hence

$$\frac{n}{m} \le \frac{N}{1} = N. \tag{8.4.2}$$

□

8.4.4. Example (extremums and best bounds in $\mathbb{Q}$).

- ⋆ $\sup^{\mathbb{Q}}(\mathbb{Q} \cap [0,1]) = 1$, $\max(\mathbb{Q} \cap [0,1]) = 1$,
- ⋆ $\sup^{\mathbb{Q}}(\mathbb{Q} \cap (0,1)) = 1$, $\max(\mathbb{Q} \cap (0,1))$ does not exist,
- ⋆ $\sup^{\mathbb{Q}}\{2 + {}^{1}/_{x} : x \in \mathbb{Q} \cap (0,1)\} = 2$, $\max\{2 + {}^{1}/_{x} : x \in \mathbb{Q} \cap (0,1)\}$ does not exist,
- ⋆ $\inf^{\mathbb{Q}}\{{}^{1}/_{n} : n \in \mathbb{N}\} = 0$, $\min\{{}^{1}/_{n} : n \in \mathbb{N}\}$ does not exist.

8.4.5. Least upper bound (supermum) vs. greatest element (maximum). If $S \in \mathbb{F}$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{Q}$, then the concept of maximum and supremum are related *but not the same.* In fact, the concept of supremum is weaker (or more general, if you prefer), in that if S has a maximum, $\max S$, then S has also a supremum and they coincide. For example, the closed rational interval

$$I := \mathbb{Q} \cap [a,b] := \{x \in \mathbb{Q} : a \le x \le b\} \tag{8.4.1}$$

for any *rational* $a, b \in \mathbb{Q}$, has both a maximum and minimum

$$\max(\mathbb{Q} \cap [a,b]) = b \text{ and } \min(\mathbb{Q} \cap [a,b]) = a, \tag{8.4.2}$$

and these coincide with $\sup I$ and $\inf I$. On the other hand the open interval

8.4.6. Exercise. *Prove that if* $S \in \mathbb{F}$*, where* $\mathbb{F}$ *is an ordered field, has a maximum* $\max S$ *(resp. minimum* $\min S$*) then* S *has also a supremum* $\sup S$ *(resp. infimum* $\inf S$*) and* $\sup S = \max S$ *(resp.* $\inf S = \min S$*).*

8.5. Sequences in archimedean fields

The definitions and properties we give here work for any archimedean field, which include $\mathbb{Q}$ (which we know) and $\mathbb{R}$ which will be introduced later. Let $\mathbb{F}$ be an ordered archimedean field, define $\mathbb{F}_+$ to the the set of all positive elements of $\mathbb{F}$, i.e., all $x \in \mathbb{F}$ such that $x > 0$.

8.5.1. Definition of $\mathbb{F}$-valued sequence and sequence in $\mathbb{F}$. Let $\mathbb{F}$ be an archimedean field, a function whose domain is $\mathbb{N}$ and that takes values in $\mathbb{F}$ is a called a $\mathbb{F}$*-valued sequence*, or a *sequence in* $\mathbb{F}$. In symbols we write $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$, where $a_n \in \mathbb{F}$ for each $n \in \mathbb{N}$. Sometimes the set of indices is extended to $\mathbb{N}_0$ or restricted to an infinite segment of integers starting at an $n_1 \in \mathbb{Z}$, other than 1; since all definitions can be modified accordingly to include such sequences we will use $\mathbb{N}$ as the set of indices.

The operations of the field $\mathbb{F}$, which work for elements of $\mathbb{F}$, can be inherited by sequences. In particular, given two sequences $\boldsymbol{a}$ and $\boldsymbol{b}$ we define their sum to be the sequence $\boldsymbol{c}$ where $c_n := a_n + b_n$ for each $n \in \mathbb{N}$, we write $\boldsymbol{a}+\boldsymbol{b}$ for such a sequence. In symbols

$$(a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} := (a_n + b_n)_{n\in\mathbb{N}}. \tag{8.5.1}$$

Similarly the product of the two sequences $\boldsymbol{a}$ and $\boldsymbol{b}$ is the sequence

$$\boldsymbol{a}\boldsymbol{b} = (a_n)_{n\in\mathbb{N}}(b_n)_{n\in\mathbb{N}} := (a_n b_n)_{n\in\mathbb{N}}. \tag{8.5.2}$$

Also the ordering relations $<$ and $\leq$ are inherited by sequences by definining

$$\boldsymbol{a} < \boldsymbol{b} :\Leftrightarrow (a_n)_{n\in\mathbb{N}} < (b_n)_{n\in\mathbb{N}} :\Leftrightarrow (\forall\, n \in \mathbb{N} : a_n < b_n), \tag{8.5.3}$$

and

$$\boldsymbol{a} \leq \boldsymbol{b} :\Leftrightarrow (a_n)_{n\in\mathbb{N}} \leq (b_n)_{n\in\mathbb{N}} :\Leftrightarrow (\forall\, n \in \mathbb{N} : a_n \leq b_n). \tag{8.5.4}$$

8.5.2. Definition of convergence to 0 and vanishing sequences. Let $\mathbb{F}$ be an archimedean field. We say that a sequence $\boldsymbol{a}$ in $\mathbb{F}$ *converges to* 0 or *vanishes* if and only if for any positive $\varepsilon \in \mathbb{F}$ there exists an integer $N \geq 1$ such that for any $n \geq N$ we have that $|a_n| < \varepsilon$. In symbols we may write this statement as

$$\text{for each } \varepsilon > 0 : \exists\, N \in \mathbb{N} : n \geq N \Rightarrow |a_n| < \varepsilon. \tag{8.5.1}$$

8.5.3. Definition of convergence, convergent sequences and limit. Let $\mathbb{F}$ be an archimedean field. We say that a sequence $\boldsymbol{a}$ in $\mathbb{F}$ converges to a point $\hat{a} \in \mathbb{F}$, called a limit of the sequence $\boldsymbol{a}$, if and only if the sequence $\boldsymbol{e}$ of *errors* $e_n := a_n - \hat{a}$ vanishes as $n \to \infty$.

As we will shortly prove (see 8.5.4) a sequence $\boldsymbol{a}$ can have no more than one limit (if any). Thanks to this property, called *uniqueness of limit,* a convergent sequence $(a_n)_{n\in\mathbb{N}}$ identifies a single limit denoted $\lim_{n\to\infty} a_n$. Often, when the limit has a short name, say $\hat{a}$, e.g., we also write

$$x_n \to \hat{x} \text{ as } n \to \infty \tag{8.5.1}$$

and we say that x_n *tends to* (or *converges to*) $\hat{x}$ as n tends to ∞.
Note that equivalently, the sequence $\boldsymbol{a}$ converges to $\hat{a}$ if and only if

$$\text{for each } \varepsilon \in \mathbb{F}_+ : \exists N \in \mathbb{N} : n \geq N \Rightarrow |a_n - \hat{a}| < \varepsilon. \tag{8.5.2}$$

8.5.4. Proposition (uniqueness of limit (in archimedean fields)). *If $\boldsymbol{a}$ converges, in $\mathbb{F}$, to the limits $\hat{a}$ and $\hat{b}$, both in $\mathbb{F}$, then $\hat{a} = \hat{b}$.*

Proof Fix an arbitrary $\varepsilon > 0$ in $\mathbb{F}$. Then there exist $N_{\hat{a}}, N_{\hat{b}}$ (corresponding to each of the limits and $\varepsilon/2$) such that

$$n \geq N_x \Rightarrow |x - a_n| < \varepsilon/2 \tag{8.5.1}$$

for each $x = \hat{a}, \hat{b}$. It follows that for $N := \max\left\{N_{\hat{a}}, N_{\hat{b}}\right\}$

$$\left|\hat{a} - \hat{b}\right| \leq \left|\hat{a} - a_N + a_N - \hat{b}\right| \leq |\hat{a} - a_N| + \left|a_N - \hat{b}\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{8.5.2}$$

Since $\varepsilon > 0$ is arbitrary, by the Vanishing Constant Lemma 8.3.10, we conclude that $\left|\hat{a} - \hat{b}\right| = 0$, i.e., $\hat{a} = \hat{b}$. □

8.5.5. Definition of bounded above, bounded below and bounded sequence. The sequence $\boldsymbol{x}$ is called *bounded above* (resp. *below*) if and only if the set of its values $\{x_n\}_{n\in\mathbb{N}}$ is bounded above (resp. below), i.e., there exists a number $M \in \mathbb{F}$ such that

$$x_k \leq M \text{ for each } k \in \mathbb{N}, \tag{8.5.1}$$

(resp.

$$x_k \geq M \text{ for each } k \in \mathbb{N}). \tag{8.5.2}$$

The sequence $\boldsymbol{x}$ is called a *bounded sequence* if and only if it is bounded both above and below.

8.5.6. Proposition (bounded means absolutely bounded). *A sequence $\boldsymbol{x}$ in an ordered field $\mathbb{F}$ is bounded if and only if the sequence of absolute values $(|a_n|)_{n\in\mathbb{N}}$ is bounded above.*

Proof Suppose $\boldsymbol{x}$ is bounded, then there exist $M_\sharp$ and $M_\flat$ such that

$$M_\flat \leq x_n \leq M_\sharp \text{ for each } n \in \mathbb{N}. \tag{8.5.1}$$

Taking $M := \max\left\{|M_\flat|, \left|M_\sharp\right|\right\}$ and using the properties of the absolute value we obtain

$$|x_n| \leq M \text{ for each } n \in \mathbb{N}. \tag{8.5.2}$$

Conversely supposing (8.5.2) then taking $M_\flat = -M$ and $M_\sharp = M$ we see that (8.5.1) is satisfied. □

8.5.7. Definition of divergence to infinity and infinite limits. Let $\mathbb{F}$ be an archimedean field (e.g., $\mathbb{F} = \mathbb{Q}$ or $\mathbb{R}$). A sequence $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$ in $\mathbb{F}$ is said to *diverge to (positive) infinity,* or to have an *(plus or positive) infinite limit,* if and only if for each $\lambda \in \mathbb{F}$ there exists $N \in \mathbb{N}$ such that

$$n \geq N \Rightarrow a_n \geq \lambda. \tag{8.5.1}$$

In this case we write[3]

$$\lim_{n\to\infty} a_n = \infty \qquad \left(\text{ or } +\infty \text{ when needed}\right). \tag{8.5.2}$$

[3] By convention ∞ is the same as $+\infty$.

Similarly we say that a *sequence diverges to minus (or negative) infinity* if for each $\lambda \in \mathbb{R}$ there exists $N \in \mathbb{N}$ for which

$$n \geq N \Rightarrow a_n \leq \lambda. \tag{8.5.3}$$

We say that a sequence $\boldsymbol{a}$ *diverges in absolute value* if and only if the sequence $(|a_n|)_{n\in\mathbb{N}}$ diverges to ∞.

8.5.8. Proposition (comparison of divergent sequences). *Suppose $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$ and $\boldsymbol{b} = (b_n)_{n\in\mathbb{N}}$ are two sequences such that $\boldsymbol{b}$ ultimately dominate $\boldsymbol{a}$, i.e., that for some $N \in \mathbb{N}$*

$$n \geq N \Rightarrow a_n \leq b_n, \tag{8.5.1}$$

then

- ⋆ *if $\boldsymbol{a}$ diverges to infinity so does $\boldsymbol{b}$,*
- ⋆ *if $\boldsymbol{b}$ diverges to minus infinifty so does $\boldsymbol{a}$,*

Proof See Problem 8.X.19. □

8.5.9. Lemma (Cauchy's necessary criterion). *In an ordered field, a convergent sequence $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$ must satisfy the following Cauchy sequence property*

$$\forall\, \varepsilon > 0 : \exists\, N \in \mathbb{N} : (n \geq N \text{ and } k \geq 1) \Rightarrow |a_{n+k} - a_n| < \varepsilon. \tag{8.5.1}$$

Proof Let $\hat{a} := \lim_{n\to\infty} a$. Fix $\varepsilon > 0$, by convergence of a_n to $\hat{a}$, as $n \to \infty$, we have that for some $N \in \mathbb{N}$

$$n \geq N \Rightarrow |a_n - \hat{a}| < \varepsilon/2. \tag{8.5.2}$$

Therefore for $n \geq N$ and $k \geq 0$, since $n + k \geq n$ we have by triangle inequality and (8.5.2) we have

$$|a_n - a_{n+k}| \leq |a_n - \hat{a}| + |\hat{a} - a_{n+k}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{8.5.3}$$

□

8.5.10. Definition of Cauchy sequence. A sequence $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$ in an odered field $\mathbb{F}$ is called a Cauchy sequence if and only if it satisfies the Cauchy sequence property (8.5.1).

8.5.11. Theorem (convergent sequences are bounded). *Let $\boldsymbol{a}$ be a sequence in an ordered field $\mathbb{F}$. If $\boldsymbol{a}$ is Cauchy then it must be bounded. In particular all convergent sequences are bounded.*
Proof See Problem 8.X.7 as to why a Cauchy sequence must be bounded. From the Necessary Cauchy Criterion, a convergent sequence must be Cauchy and therefore bounded.
(It is possible to prove the second statement without using Cauchy sequences. You should try this.) □

8.5.12. Theorem (algebra of limits). *If $\boldsymbol{a}$ and $\boldsymbol{b}$ are convergent sequences in $\mathbb{F}$ then*

(a) the sum of the two sequences, which is defined as the sequence $\boldsymbol{c}$, where $c_n := a_n + b_n$ for each $n \in \mathbb{N}$, converges and

$$\lim_{n\to\infty} [a_n + b_n] = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \tag{8.5.1}$$

(b) the product of the two sequences, which is defined as the sequence $\boldsymbol{c}$, where $c_n := a_n b_n$ for each $n \in \mathbb{N}$, converges and

$$\lim_{n\to\infty} [a_n b_n] = \lim_{n\to\infty} a_n \lim_{n\to\infty} b_n. \tag{8.5.2}$$

(c) if $a_n \neq 0$ for all $n \geq 1$ and $\lim_{n\to\infty} a_n \neq 0$ then the sequence of reciprocals $\boldsymbol{c}$, where $c_n := {}^1/_{a_n}$ for each $n \in \mathbb{N}$, converges and

$$\lim_{n\to\infty} \left[\frac{1}{a_n}\right] = \frac{1}{\lim_{n\to\infty} a_n}. \tag{8.5.3}$$

Proof

(a) See Problem 8.X.6.

(b) Fix $\varepsilon > 0$, since $\boldsymbol{a}$ converges to $\hat{a}$ we know that for a large enough N_1 we have

$$n \geq N_1 \Rightarrow |a_n - \hat{a}| \leq \frac{\varepsilon}{\hat{b}}. \tag{8.5.4}$$

Similarly we know that there exists N_2 for which

$$n \geq N_2 \Rightarrow \left|b_n - \hat{b}\right| \leq \frac{\varepsilon}{2}. \tag{8.5.5}$$

Then for any $n \geq N$ we have

$$\left|a_n b_n - \hat{a}\hat{b}\right| = |a_n b_n - \hat{a} b_n| + \left|\hat{a} b_n - \hat{a}\hat{b}\right| \leq \beta \frac{\varepsilon}{2\beta} + |\hat{a}| \frac{\varepsilon}{2|\hat{a}|} = \varepsilon. \tag{8.5.6}$$

□

8.5.13. Lemma (squeezing). *Given three sequences $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$, $\boldsymbol{b} = (b_n)_{n\in\mathbb{N}}$ and $\boldsymbol{c} = (c_n)_{n\in\mathbb{N}}$ in $\mathbb{F}$, such that $a_n \leq c_n \leq b_n$ for all $n \in \mathbb{N}$ and both $\boldsymbol{a}$ and $\boldsymbol{b}$ converge to the same limit $\hat{l} = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$, then also $\boldsymbol{c}$ converges to this limit:*

$$\lim_{n\to\infty} c_n = \hat{l}. \tag{8.5.1}$$

Proof Note first by the algebra of limits that $b_n - a_n \to \infty$, by the triangle inequality and the fact that $a_n \leq c_n \leq b_n$ we have

$$\left|c_n - \hat{l}\right| \leq \left|a_n - \hat{l}\right| + |c_n - a_n| = \left|a_n - \hat{l}\right| + (c_n - a_n) \leq \left|a_n - \hat{l}\right| + (b_n - a_n). \tag{8.5.2}$$

□

8.5.14. Remark. The assumption that $a_n \neq 0$ can be removed provided the sequence of reciprocals $1/\boldsymbol{a}$ is considered from a certain index N, possibly strictly larger than 1, as long at the assunmption that $\lim_{n\to\infty} a \neq 0$ is kept.

8.6. Real numbers

A quick way to introduce real numbers is to assume their existence with their properties, in effect $\mathbb{R}$ can be seen as the smallest complete superset of $\mathbb{Q}$, whereby *complete* means on of the following (equivalent) properties

(1) each set that is bounded above must have a least upper bound
(2) each Cauchy sequence must converge.

In detail, this means that $\mathbb{R}$ can be introduced axiomatically as an ordered archimedean field, which contains $\mathbb{Q}$, but that is complete.

8.6.1. Incompleteness of $\mathbb{Q}$. The main shortcoming of rational numbers is that they are incomplete. We leave this to you, reader, in the form of a guided problem.

Problem (Heron's algorithm and incompleteness of rationals)**.** *Consider the recursive sequence defined by $x_0 \in \mathbb{Q}$ with $x_0 \geq 2$ and*

$$x_{k+1} := \frac{1}{2}\left(\frac{2}{x_k} + x_k\right) \text{ for } k \geq 0. \tag{8.6.1}$$

(a) Show the following properties

$$0 \leq x_{k+1} \leq x_k \text{ for all } k \geq 0, \tag{8.6.2}$$

$$0 \leq x_{k+1}^2 - 2 \leq \frac{1}{8}\left(x_k^2 - 2\right)^2. \tag{8.6.3}$$

(b) Deduce that $\left(x_k^2\right)_{k\in\mathbb{N}_0}$ converges with limit 2, i.e., $x_k^2 \to 2$ as $k \to \infty$.
(c) Deduce that $\left(x_k^2\right)_{k\in\mathbb{N}_0}$ is a bounded sequence.
(d) Deduce that $(x_k)_{k\in\mathbb{N}_0}$ is a Cauchy sequence in $\mathbb{Q}$.
(e) Prove that $(x_k)_{k\in\mathbb{N}_0}$ does not converge in $\mathbb{Q}$.

8.6.2. Axiomatic definition of real numbers. There exists a set denoted $\mathbb{R}$, with operations $+$, $\times$ that make $(\mathbb{R}, +, \times)$ a field, an order relation $<$ which makes $(\mathbb{R}, +, \times, <)$ an archimedean field and *furthermore* $\mathbb{R}$ it satisfies the following.

Completeness Axiom.

Each subset $S \subseteq \mathbb{R}$ which is not empty and bounded above, must have a least upper bound in $\mathbb{R}$.

It is possible to show

(1) $\mathbb{R}$ contains an exact copy of $\mathbb{Q}$ (with the same operations $+$, $\times$ and ordering $<$), i.e., $\mathbb{Q} \subseteq \mathbb{R}$ up to an order-preserving field isomorphism,
(2) $\mathbb{Q} \neq \mathbb{R}$, i.e., there are elements of $\mathbb{R}$ that do not belong to $\mathbb{Q}$,
(3) $\mathbb{R}$ can be constructed from $\mathbb{Q}$ (which means that it being an axiom is not necessary).

8.6.3. Rational numbers are real. We now show that the field $\mathbb{R}$ contains a copy of $\mathbb{Q}$.
Since $\mathbb{R}$ is a field, it has a zero and a unit, call them $0_\mathbb{R}$ and $1_\mathbb{R}$. We see that $1_\mathbb{R} > 0_\mathbb{R}$. By compatibility of $<$ with $+$ in $\mathbb{R}$, the element $1_\mathbb{R} + 1_\mathbb{R} =: 2_\mathbb{R}$ is stricIty bigger than $1_\mathbb{R}$.

8.6.4. Some reals are irrational. A number that is real but not rational, is called an *irrational number*. In fact, *almost all*[4] real numbers are irrational.
It is enough to show that there is a number $s \in \mathbb{R}$ such that $s^2 = 2$. Since we know that there can be no such number in $\mathbb{Q}$, it follows that s is irrational. To do this, consider the set

$$X := \left\{ x \in \mathbb{R} : \ x^2 \leq 2 \right\}. \tag{8.6.1}$$

The set X is bounded above, this can be seen by noting first that if $0 \leq a \leq b$ then $a^2 \leq b^2$.

8.7. Power and root

Completness of $\mathbb{R}$ will now allow us to settle the long-standing question of whether $\sqrt{2}$, $\sqrt{x}$, and more generally $\sqrt[n]{x}$, exists for x and n is some reasonable domains. We do this by intrducing the inverse function of the power function. We have already used the integer arithmetic power function, for example, in describing the cardinality of the power of a finite set 5.1.4 or when we studied the Binomial Theorem 5.4; in this section we merely formalise our intuitive knowledge by recalling the definition of integer arithemtic power in 8.7.1 and then prove in (5.2.1) that each positive real number must have a square root, i.e., that $\sqrt{\cdot} : \mathbb{R}_{0+} \to \mathbb{R}_{0+}$, informally introduced (5.2.1) is well defined. We will do this in this section.

8.7.1. Definition of (arithmetic) power function. Let $n \in \mathbb{N}_0$ we define the n-th (*arithmetic*)integer arithmetic power *power* of a number $x \in \mathbb{R}$ as

$$x^n := \begin{cases} 1 & \text{if } n = 0 \\ x x^{n-1} & \text{if } n \in \mathbb{N}. \end{cases} \tag{8.7.1}$$

The function

$$\mathbb{R} \ni x \mapsto x^n \in \mathbb{R} \tag{8.7.2}$$

is called the *n-th power function*. We will use the notations $\cdot^n$ or pow^n to indicate the n-th power function.
The 2nd power (function) is called *square* and the 3rd power (function) is called *cube*. Note that the 1st power function is simply the identity.

8.7.2. Remark. Note that if $x = 0$ in the definition of power we have

$$0^n := \begin{cases} 1 \text{ if } n = 0 \\ 0 \text{ if } n \geq 1. \end{cases} \tag{8.7.1}$$

There is nothing wrong with defining $0^0 = 1$, in fact, this is quite useful.

8.7.3. Proposition (square root of positive reals). *Let $r \in \mathbb{R}_{0+}$, there exists a unique number $s \in \mathbb{R}_{0+}$ such that*

$$s^2 = r. \tag{8.7.1}$$

Proof The proof is sketched as part of the guided problem 8.7.4. □

[4] We use "almost all" in a sense that can be made fully rigorous by using *measure theory*, a branch of analysis related to integrals and probability. In these notes we will look at Cantor's diagonal argument, which shows that real numbers cannot be "counted".

8.7.4. Problem (square root of positive reals). *Prove the following result*

Let $r \in \mathbb{R}_{0+}$, there exists a unique number $s \in \mathbb{R}_{0+}$ such that

$$s^2 = r. \tag{8.7.1}$$

Hint. Fix a number $r \in \mathbb{R}_{0+}$

(i) Consider the set $S : \left\{x \in \mathbb{R} : \; x^2 \leq r\right\}$ and show that S is nonempty and bounded-above.

(ii) Use the Completeness Axiom to conclude that there exists $\sup S \in \mathbb{R}$ *and give this number a name, say s.*

(iii) Show that it cannot be $s^2 < r$.

(iv) Show that it cannot be $s^2 > r$ and conclude that

$$s^2 = r. \tag{8.7.2}$$

This shows existence.

(v) Prove uniqueness by supposing that for some other $t \in \mathbb{R}_{0+}$ we have $t^2 = r$ and deduce $t = s$.

8.7.5. Proposition (root of an integer power). *Let $n \in \mathbb{N}$ and $x \in \mathbb{R}_+$ then there exists a unique $s \in \mathbb{R}_+$ such that*

$$s^n = x. \tag{8.7.1}$$

If n is odd, the statement can be extended to the include $x \in \mathbb{R}$ and $s \in \mathbb{R}$.
Proof We leave the proof of this result as Problem. □

8.7.6. Definition of n-th root function. For $n \in \mathbb{N}$, denoting by $D = \mathbb{R}$ if n is odd or $D = \mathbb{R}_{0+}$ if n is even, the function $\sqrt[n]{\cdot} = \text{pow}^{1/n} : D \to \mathbb{R}$ is defined for each x by $y := \sqrt[n]{\cdot} = x^{1/n} \in \mathbb{R}$ such that $y^n = x$; such a y is called an n-th *root* of x, with the exception of $n = 2$ when we say *square root*, and $n = 3$ when we say *cubic root*. Thanks to Proposition 8.7.5, the definition of n-th root is well posed. The notations $\sqrt[n]{\cdot}$ and $\text{pow}^{1/n}$ are interchangeable, but for tradition reasons we usually prefer the former.

8.7.7. Integer power and root laws. For any $n \in \mathbb{N}$ and $x \in \mathbb{R}_{0+}$ we have

$$\left(x^n\right)^{1/n} = \left(x^{1/n}\right)^n = x. \tag{8.7.1}$$

If n is odd the admissible domain for x can be extended to $\mathbb{R}$. If n is even then we have

$$\left(x^n\right)^{1/n} = |x|. \tag{8.7.2}$$

8.7.8. Definition. The root functions allow to define the q-th power function for any $q \in \mathbb{Q}$, namely if $q = {}^n/_m$, for $n \in \mathbb{Z}$ and $m \in \mathbb{N}$ then

$$x^{-1} := {}^1/_x \text{ and } x^{n/m} := \left(x^n\right)^{1/m}, \text{ for } n \in \mathbb{Z} \text{ and } m \in \mathbb{N}. \tag{8.7.1}$$

8.7.9. Proposition (power laws). *For all $q, r \in \mathbb{Q}$ and $x, y \in \mathbb{R}_+$, we have*

$$\big(xy\big)^q = x^q y^q, \tag{8.7.1}$$

$$x^{q+r} = x^q x^r, \tag{8.7.2}$$

$$x^{qr} = (x^q)^r. \tag{8.7.3}$$

Proof We leave the proof of these facts as an exercise. As general hint, prove first the case where q and r are integers and then the case where there are of the form $1/m$ and $1/n$ for some $m, n \in \mathbb{N}$. □

What is not so easy to extend the domain of q (and r) to be all of $\mathbb{R}$, not just $\mathbb{Q}$. We will do this later.

8.8. Montone sequences

8.8.1. Definition of monontone, increasing, and decreasing sequences. Let $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. We say that $(a_n)_{n\in\mathbb{N}}$ is *increasing* (resp. *decreasing*) if and only if

$$\begin{aligned} a_n &\leq a_{n+1} \text{ for all } n \geq 1,\\ a_n &\geq a_{n+1} \text{ for all } n \geq 1. \end{aligned} \tag{8.8.1}$$

We say the sequence $\boldsymbol{a}$ is *ultimately increasing* (resp. *decreasing*) if and only if text for some $n_0 \geq 0$

$$\begin{aligned} a_n &\leq a_{n+1} \text{ for all } n \geq n_0,\\ a_n &\geq a_{n+1} \text{ for all } n \geq n_0. \end{aligned} \tag{8.8.2}$$

All convergence-related behaviour of sequences, depends only on their ultimate properties, so most of our results can be adapted to cover the case of ultimately.

8.8.2. Definition of best bounds of a sequence. Let $\boldsymbol{a}$ be a sequence in an ordered field $\mathbb{F}$ we say that $\boldsymbol{a}$ has a *supremum* (resp. *infimum*) if and only $A := \{a_n\}_{n\in\mathbb{N}}$ has a supremum (resp. infimum). Often, when talking about a sequence, to avoid introducing A, we use the notation $\sup_{n\in\mathbb{N}} a_n$ (resp. $\inf_{n\in\mathbb{N}} a_n$) to indicate $\sup A$ (resp. $\inf A$).

8.8.3. Theorem (monotone sequence). *Let $\boldsymbol{a}$ be an ultimately increasing sequence in $\mathbb{R}$. If $\boldsymbol{a}$ is bounded above, then it has a supremum in $\mathbb{R}$, it converges and*

$$\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n. \tag{8.8.1}$$

Likewise if $\boldsymbol{a}$ is ultimately decreasing sequence in $\mathbb{R}$ and $\boldsymbol{a}$ is bounded below, then it has an infimum in $\mathbb{R}$, it converges and

$$\lim_{n\to\infty} a_n - \inf_{n\in\mathbb{N}} a_n. \tag{8.8.2}$$

Proof We provide the proof for increasing sequence and let you, the reader, adapt it to the case of ultimately increasing sequences. First thing we note that A is bounded above and in view of $\mathbb{R}$ completeness A has a supremum, call it $\alpha \in \mathbb{R}$. Now we need to show that $a_n \to \alpha$. For this recall the following *closure point property* which characterises the supremum of the set A

$$\forall\, \varepsilon > 0 : \exists\, x \in A : x \leq \sup A < x + \varepsilon. \tag{8.8.3}$$

In our context this is equivalent to

$$\forall\, \varepsilon > 0 : \exists\, N \in A : \alpha < a_N + \varepsilon. \tag{8.8.4}$$

And by the increasing nature of $\boldsymbol{a}$ we can rewrite

$$\forall\, \varepsilon > 0 : \exists\, N \in A : n \geq N \Rightarrow \alpha < a_n + \varepsilon. \tag{8.8.5}$$

Noting that $a_n \leq \alpha$ for all $n \in \mathbb{N}$ and some algebraic manipulations allow us to conclude that

$$\forall\, \varepsilon > 0 : \exists\, N \in A : n \geq N \Rightarrow |a_n - \alpha| < \varepsilon, \tag{8.8.6}$$

which means that $a_n \to \infty$ as $n \to \infty$. □

8.8.4. Remark (completeness is essential for Monotone Sequence Theorem). The assumption that $\boldsymbol{a}$ is in $\mathbb{R}$ is crucial for Theorem 8.8.3 to be valid. Indeed, the conclusion is not valid in $\mathbb{Q}$: for example the sequence $a_0 = 2$ and $a_{n+1} = (2/a_n + a_n)/2$ is monotone and bounded below, but does not converge in $\mathbb{Q}$.

8.8.5. Example (Napier–Bernoulli–Euler constant). Consider the sequence $\boldsymbol{e} = (e_n)_{n\in\mathbb{N}}$ given by

$$e_n := (1 + {}^1\!/\!_n)^n. \tag{8.8.1}$$

Show that the sequence is increasing and bounded above. Deduce that it converges. Its limit is an important constant in mathematics and the sciences, know as the constant of Napier or Euler. It was first described by one of the Bernoulli.

Solution. To show monotonicity we use the binomial theorem to expand

$$(1 + {}^1\!/\!_n)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \text{ and } (1 + {}^1\!/\!_{n+1})^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} \frac{1}{(n+1)^k}. \tag{8.8.2}$$

Using basic properties of the binomial coefficients and the descending factorial notation

$$k^{\underline{n}} := n(n-1)\cdots(n-k+1) = \frac{n!}{k!}, \tag{8.8.3}$$

we have

$$\begin{aligned}\binom{n}{k}\frac{1}{n^k} &= \frac{n!}{k!(n-k)!} \times \frac{1}{n^k} = \frac{1}{k!} \times \frac{k^{\underline{n}}}{n^k} \\ &= \frac{1}{k!}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-k+1}{n}\right) \\ &< \frac{1}{k!}\left(\frac{n+1}{n+1}\right)\left(\frac{n}{n+1}\right)\cdots\left(\frac{n-k+2}{n+1}\right) \\ &= \binom{n+1}{k}\frac{1}{n^k}.\end{aligned} \tag{8.8.4}$$

This proves that

$$e_n = \left(1 + \frac{1}{n}\right)^n < \left(1 + \frac{1}{n+1}\right)^{n+1} = e_{n+1} \text{ for each } n \in \mathbb{N}, \tag{8.8.5}$$

i.e., that the sequence $(e_n)_{n\in\mathbb{N}}$ is strictly increasing.
Next we show that the sequence $(e_n)_{n\in\mathbb{N}}$ is bounded above.

8.9. Rational and constructive approaches to reals

So far we have axiomatically accepted the real numbers. We show here that the axioms of real numbers are not necessary and that the existence of $\mathbb{R}$ can be based on that of rational numbers and the ZF axioms of set-theory. Various constructions of real numbers exist. The two most popular alternatives are:

(a) defining real numbers as *rational cuts* (also known as Dedekind cuts), or

(b) defining real numbers are equivalence classes of Cauchy sequences of rational numbers.[5]

8.9.1. Dedekind cuts. A Dedkind cut of the rationals $\mathbb{Q}$ is a pair (L, U) of subsets thereof such that

$$x \in L \text{ and } y \in U \Rightarrow x < y, \tag{8.9.1}$$

$$L \cup U = \mathbb{Q} \setminus S, \text{ with } S := \begin{cases} \{q\} \text{ for some } q \in \mathbb{Q} \\ \varnothing. \end{cases} \tag{8.9.2}$$

In words L lies below U, and the pair L, U either misses only one point of $\mathbb{Q}$ or none. In the first case (one point q missed), we say that (L, U) is a *rational cut* (identified with the missed number q), in the second case (no point missed) we call it an *irrational cut.* It is then possible to define on the set $\mathfrak{D}$ of all Dedekind cuts of $\mathbb{Q}$:

- ⋆ an ordering relation via

$$(L_1, U_1) <_{\mathfrak{D}} (L_2, U_2) \Leftrightarrow L_1 \subsetneq L_2, \tag{8.9.3}$$

- ⋆ an addition,

$$(L_1, U_1) +_{\mathfrak{D}} (L_2, U_2) = (L_1 \dot{+} L_2, U_1 \dot{+} U_2) \tag{8.9.4}$$

where for two sets $X, Y \in \mathbb{Q}$ their *pointwise addition* is

$$X \dot{+} Y := \{x + y : x \in X \text{ and } y \in Y\} \tag{8.9.5}$$

- ⋆ a multiplication when we have $U_i \subseteq \mathbb{Q}_{0+}$ for both $i = 1, 2$,

$$(L_1, U_1) \times_{\mathfrak{D}} (L_2, U_2) := (L, U_1 \cdot U_2) \tag{8.9.6}$$

where for two sets $X, Y \subseteq \mathbb{Q}$ their *pointwise product* is

$$X \cdot Y := \{xy : x \in X \text{ and } y \in Y\} \tag{8.9.7}$$

and

$$L := (L_1 \cap \mathbb{Q}_{0+}) \cdot (L_2 \cap \mathbb{Q}_{0+}) \cup \mathbb{Q}_{-}. \tag{8.9.8}$$

8.9.2. Theorem (Dedekind cuts). *The set* $(\mathfrak{D}, +_{\mathfrak{D}}, \times_{\mathfrak{D}}, <_{\mathfrak{D}})$ *of all Dedekind cuts of* $\mathbb{Q}$ *with the corresponding operations is a complete archimedean field and thus a constructible copy of* $\mathbb{R}$*.*

Proof Omitted. □

[5] This is one of the reason we developped the concepts of convergent and Cauchy sequences to fit $\mathbb{Q}$.

8.9.3. Cauchy sequences of rational numbers. An alternative approach to reals, avoiding Dedekind cuts, can taken by considering the set $\mathfrak{R} \subseteq \mathbb{Q}^{\mathbb{N}}$ of all Cauchy sequences of rational numbers with the usual sum and multiplication of sequences and ordering

$$\boldsymbol{a} \preceq \boldsymbol{b} :\Longleftrightarrow 0 \leq b_n - a_n \text{ for all } n \text{ greater than some } N \in \mathbb{N}. \tag{8.9.1}$$

Then after introducing the relation

$$\boldsymbol{a} \equiv \boldsymbol{b} :\Longleftrightarrow a_n - b_n \to 0 \text{ as } n \to \infty, \tag{8.9.2}$$

and checking that $\equiv$ is an equivalence relation compatible with ordering, addition and multiplication in $\mathfrak{R}$, it is possible to show that the quotient set $\mathfrak{R}/\equiv$ with the induced ordering and algebra is a complete archimedean field, therefore a constructible copy of $\mathbb{R}$.

Exercises and problems on real numbers

Exercise 8.X.1 (real and rational numbers). (a) Given $q, r \in \mathbb{Q}$, show that there exists an infinite sequence of rational numbers (q_n), all different, such that $q < q_n < r$.

(b) State the Archimedean Property for the set of rational numbers $\mathbb{Q}$.

(c) Using the Archimedean Property show that for each $x \in \mathbb{Q}$ with $x > 0$, there exists $n \in \mathbb{N}$ such that $x > 1/n$.

(d) Deduce that the sequence $({}^1/_n)_{n\in\mathbb{N}}$ converges to 0 in $\mathbb{Q}$.

Problem 8.X.2 (irrational algebra). Let x be a real number. Show that if x^2 is irrational, then so is x. Deduce that $\sqrt{2}+\sqrt{3}$ is irrational.

Problem 8.X.3 (Heron's algorithm and incompleteness of rationals). Consider the recursive sequence defined by $x_0 \in \mathbb{Q}$ with $x_0 \geq 2$ and

$$x_{k+1} := \frac{1}{2}\left(\frac{2}{x_k} + x_k\right) \text{ for } k \geq 0. \tag{8.X.3.1}$$

(a) Show the following properties

$$0 \leq x_{k+1} \leq x_k \text{ for all } k \geq 0, \tag{8.X.3.2}$$

$$0 \leq x_{k+1}^2 - 2 \leq \frac{1}{8}\left(x_k^2 - 2\right)^2. \tag{8.X.3.3}$$

(b) Deduce that $\left(x_k^2\right)_{k\in\mathbb{N}_0}$ converges with limit 2, i.e., $x_k^2 \to 2$ as $k \to \infty$.

(c) Deduce that $\left(x_k^2\right)_{k\in\mathbb{N}_0}$ is a bounded sequence.

(d) Deduce that $(x_k)_{k\in\mathbb{N}_0}$ is a Cauchy sequence in $\mathbb{Q}$.

(e) Prove that $(x_k)_{k\in\mathbb{N}_0}$ does not converge in $\mathbb{Q}$.

Exercise 8.X.4 (rational numbers). (a) Given $q, r \in \mathbb{Q}$, show that there exists an infinite sequence of rational numbers (q_n), all different, such that $q < q_n < r$.

(b) State the Archimedean Property for the set of real numbers $\mathbb{R}$.

(c) Using the Archimedean Property show that for each $x \in \mathbb{R}^+$ there exists $n \in \mathbb{N}$ such that $x > 1/n$.

Problem 8.X.5 (vanishing sequences in archimedean fields). Let $\mathbb{F}$ be one of the fields $\mathbb{Q}$ or $\mathbb{R}$. Explain what we mean by saying that $\mathbb{F}$ *satisfies the archimedean property.*

Prove that the following sequences all converge to 0 in $\mathbb{F}$:

$$a_n := \frac{1}{n} \text{ for all } n \in \mathbb{N} \tag{8.X.5.1}$$

$$x_n := \frac{1}{n^2} \text{ for all } n \in \mathbb{N} \tag{8.X.5.2}$$

$$y_n := \frac{1}{n^2+n} \text{ for all } n \in \mathbb{N} \tag{8.X.5.3}$$

$$z_n := \frac{1}{n^p} \text{ for all } n \in \mathbb{N} \text{ and some fixed } p \geq 1. \tag{8.X.5.4}$$

Problem 8.X.6 (sum of limits is limit of sum). Suppose $\mathbb{F} = \mathbb{Q}$ or $\mathbb{R}$. Prove the following:

Given two convergent sequences $\boldsymbol{x} = (x_k)_{k\in\mathbb{N}}$, $\boldsymbol{y} = \left(y_k\right)_{k\in\mathbb{N}}$, in $\mathbb{F}$, then

$$\lim_{k\to\infty} \left[x_k + y_k\right] = \lim_{k\to\infty} x_k + \lim_{k\to\infty} y_k. \tag{8.X.6.1}$$

Problem 8.X.7 (Cauchy sequences are bounded)**.** Show the following statement.

Any Cauchy sequence must be bounded; namely let (X,ρ) be a metric space and $\boldsymbol{x} = (x_n)_{n\in\mathbb{N}}$ a Cauchy squence, then $\boldsymbol{x}$ is bounded, in that there exist $r \in \mathbb{R}_{0+}$, $a \in X$ such that

$$\rho(a, x_n) < r \text{ for all } n \in \mathbb{N}. \tag{8.X.7.1}$$

Big Fat Note. If you do not know (yet) what a metric space is then replace X with $\mathbb{F} = \mathbb{Q}$ or $\mathbb{R}$, and $\rho(x, y)$ with $\left|x - y\right|$ for any two $x, y \in \mathbb{F}$.

Exercise 8.X.8. Based on the definition of rational numbers $\mathbb{Q}$, as an ordered archimedean field, show that for $a, b, c \in \mathbb{Q}$ and $a < b$ we have

$$c < 0 \Leftrightarrow ac > bc. \tag{8.X.8.1}$$

Exercise 8.X.9. Let $a, b \in \mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$, show that

$$a \leq b \Leftrightarrow a^2 \leq b^2. \tag{8.X.9.1}$$

Show with an example that the same is not valid if we assumed $a, b \in \mathbb{R}$, not necessarily positive.

Problem 8.X.10 (charaterisations of supremum and infimum)**.** (a) Argue that by the definition of minimum and maximum, that if a $\sup^{\mathbb{F}}$ (or $\inf^{\mathbb{F}}$) exists then it is unique.

(b) Prove that $\beta = \sup S$ is equivalent to the following two properties
   (a) if $x \in S$ then $x \leq \beta$,
   (b) if λ is an upper bound on S, then $\lambda \geq \beta$.

(c) Show the following.

Let S be a subset of an archimedean field $\mathbb{F}$ $\beta := \sup S$ if and only if

$$\forall\, \varepsilon \in \mathbb{F}_+ : \exists\, x \in S : \beta < x + \varepsilon, \tag{8.X.10.1}$$

this property is known as the property (of $\sup S$) of being the greatest closure point (of the set S).

Likewise, $\alpha := \inf S$ if and only if

$$\forall\, \varepsilon \in \mathbb{F}_+ : \exists\, x \in S : \alpha > x - \varepsilon, \tag{8.X.10.2}$$

i.e., the infimum is characterised as the lowest closure point.

Problem 8.X.11 (limits of monotone sequences)**.** Show that if $(x_n)_{n\in\mathbb{N}}$ is an increasing sequence then

$$\lim_{n\to\infty} x_n = \sup_{n\in\mathbb{N}} x_n \tag{8.X.11.1}$$

Distinguish the cases when $(x_n)_{n\in\mathbb{N}}$ is bounded and not.

Exercise 8.X.12 (upper bound transitivity)**.** Suppose $x \in \mathbb{F}$, $S \subseteq \mathbb{F}$ and x is an upper (lower) bound on S then any $y \in \mathbb{F}$ greater (smaller) than x is also an upper (lower) bound on S.

Exercise 8.X.13 (maximum is an upper bound)**.** Recall the definition of $\max S$, the maximum of a linearly (totally) ordered set $S \subseteq \mathbb{F}$, is an m such that

$$m \in S \text{ and } m \geq x \; \forall \, x \in S. \tag{8.X.13.1}$$

(i) Show the following statement:

The maximum $\max S$*, whenever it exists, is an upper bound on* S*.*

Note however, with counterexamples, that

(ii) An upper bound on $S \subseteq \mathbb{F}$ is not necessarily a maximum of S.

(iii) Sets that are bounded above need not have a maximum.

Problem 8.X.14 (inequalities and intervals in $\mathbb{R}$)**.** Describe the following sets as union of open, half-open or closed intervals.

$$S := \left\{ x \in \mathbb{R} : \frac{x+1}{2x-3} \leq 1 \right\}, \tag{8.X.14.1}$$

$$T := \left\{ x \in \mathbb{R} : \frac{3x+2}{x-1} < 1 \right\}, \tag{8.X.14.2}$$

$$U := \left\{ x \in \mathbb{R} : \; x^2 \leq 2 \right\}. \tag{8.X.14.3}$$

For U, you may assume the existence of a function $\sqrt{\cdot} : \mathbb{R}_{0+} \to \mathbb{R}_{0+}$ such that $\left(\sqrt{x}\right)^2 = x = \sqrt{x^2}$ for each $x \in \mathbb{R}_{0+}$.

Problem 8.X.15 (square root function)**.** Prove the following result

Let $r \in \mathbb{R}_{0+}$*, there exists a unique number* $s \in \mathbb{R}_{0+}$ *such that*

$$s^2 = r. \tag{8.X.15.1}$$

Hint. Fix a number $r \in \mathbb{R}_{0+}$

(i) Consider the set $S : \left\{ x \in \mathbb{R} : \; x^2 \leq r \right\}$ and show that S is nonempty and bounded-above.

(ii) Use the Completeness Axiom to conclude that there exists $\sup S \in \mathbb{R}$ and give this number a name, say s.

(iii) Show that it cannot be $s^2 < r$.

(iv) Show that it cannot be $s^2 > r$ and conclude that

$$s^2 = r. \tag{8.X.15.2}$$

This shows existence.

(v) Prove uniqueness by supposing that for some other $t \in \mathbb{R}_{0+}$ we have $t^2 = r$ and deduce $t = s$.

Problem 8.X.16 (quadratic equations)**.** Given $a \in \mathbb{R}$ consider *solving for* $x \in \mathbb{R}$ the equation

$$x^2 = a. \tag{8.X.16.1}$$

Show that

(i) $a > 0$ if and only if (8.X.16.1) has two distinct solutions $\sqrt{a}$ and $-\sqrt{a}$.

(ii) $a = 0$ if and only if (8.X.16.1) has a unique solution 0.

(iii) $a < 0$ if and only if (8.X.16.1) has no solution.

Problem 8.X.17 (ordering's compatibility with product rules)**.** Using the axioms of ordered field, show the following statement.

The following are equivalent

(a) $a < b$
(b) $a c < b c$ *for all* $c > 0$
(c) $a c < b c$ *for some* $c > 0$

Problem 8.X.18 (product and ordering properties)**.** Show, from first principles (i.e., the axioms of ordered field), the following.

Let $\mathbb{F}$ *be an ordered field that is positively oriented (i.e.,* $1 > 0$*) and denote by* -1 *the opposite (also known as additive inverse) of the unit* 1*. Then the following properties hold:*

(a) *For any* $a \in \mathbb{F}$ *we have* $-a = (-1)a$*.*
(b) *If* $a > 0$ *then* $-a < 0$*.*
(c) *If* $a < b$ *then* $-a > -b$*.*
(d) *If* $a, b, c \in \mathbb{F}$*,* $a < b$ *and* $a c < b c$ *then* $c > 0$*.*
(e) *If* $a, b, c \in \mathbb{F}$*,* $a < b$ *and* $a c > b c$ *then* $c < 0$*.*

Problem 8.X.19 (comparison of divergent sequences)**.** Prove the following

Suppose $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$ *and* $\boldsymbol{b} = (b_n)_{n\in\mathbb{N}}$ *are two sequences such that* $\boldsymbol{b}$ *ultimately dominate* $\boldsymbol{a}$*, i.e., that for some* $N \in \mathbb{N}$

$$n \geq N \Rightarrow a_n \leq b_n, \tag{8.X.19.1}$$

then

- ⋆ *if* $\boldsymbol{a}$ *diverges to infinity so does* $\boldsymbol{b}$*,*
- ⋆ *if* $\boldsymbol{b}$ *diverges to minus infinifty so does* $\boldsymbol{a}$*,*

CHAPTER 9

Infinity

> "To infinity… and beyond!"
> – Buzz Lightyear, Toy Story franchise

In Chapter 6 we have seen how to introduce the set of natural numbers with zero, $\mathbb{N}_0$, arithmetic thereon, and the concept of finite sets and cardinality from basic set-theoretical principles. In this chapter we look at how these concepts can be generalised to infinite sets. This will allow us to "classify" the different types of infinity, for example, we learn that while $\natural$, $\mathbb{Z}$ and even $rationals$ have the same infinite cardinality, the real numbers $\mathbb{R}$ and integers $\mathbb{N}$ are not comparable.

9.1. Infinite sets

9.1.1. Lemma (characterisation of infinity via $\mathbb{N}_0$). *A set S is infinite if and only if there exists a map $\sigma : \mathbb{N}_0 \to S$ that is injective.*

Proof Suppose S is infinite then, by definition, there exists a map $\phi : S \to S$ that is injective but not surjective. This means that $S \neq \varnothing$ [$*$] and that there exists $s_0 \in S$ [$*$]: Check!
such that $s_0 \neq \phi(s)$ for any $s \in S$. Now define recursively (by using the Recursion Theorem 6.4.13) the map

$$\sigma(n) := \begin{cases} s_0 & \text{for } n = 0 \\ \phi(\sigma(n-1)) & \text{for } n \geq 1. \end{cases} \tag{9.1.1}$$

(In other words $\sigma(n) = \phi^{\circ n}(s_0)$ and $\phi^{\circ n} = \phi \circ \cdots \circ \phi$, where the "factor" ϕ appears n times as in Definition 6.3.16.) The map σ is well defined and it is injective, indeed suppose that for some $m, n \in \mathbb{N}_0$ we have $\sigma(m) = \sigma(n)$, which means $\phi^{\circ m}(s_0) = \phi^{\circ n}(s_0)$ and assuming without loss of generality[1] that $n \geq m$, we obtain

$$\phi^{\circ n-m}(s_0) = s_0. \tag{9.1.2}$$

If $n > m$ this identity would imply that $s_0 = \phi(s)$.

[1] The expression "without loss of generality" is used for an assumption when that assumption can be easily removed without having to change much of the proof.

Conversely suppose there is an injective map $\sigma : \mathbb{N}_0 \to S$, we want to show S is infinite. To see it, let us build $\phi : S \to S$ as follows

$$\phi(s) := \begin{cases} s & \text{if } s \notin \sigma(\mathbb{N}_0), \\ \sigma(\sigma^{-1}(s)+1) & \text{if } s \in \sigma(\mathbb{N}_0). \end{cases} \tag{9.1.3}$$

Since σ is injective, then the inverse function σ^{-1} is well-defined with domain $\sigma(\mathbb{N}_0)$. The map ϕ is injective, indeed if $\phi(s) = \phi(r)$, then either both s, r are elements of $\sigma(\mathbb{N}_0)$, in which case

$$\sigma(\sigma^{-1}(s)+1) = \phi(s) = \phi(r) = \sigma(\sigma^{-1}(r)+1) \tag{9.1.4}$$

and, by injectivity of σ, $\cdot + 1$ and σ^{-1} we get $s = r$. Otherwise, if s or r is not in $\sigma(\mathbb{N}_0)$, then, by definition of ϕ, $s = \phi(s) = \phi(r) = r$.
To conclude, we show next that ϕ is not surjective. Namely, let $s_0 := \sigma(0)$, and suppose $\phi(s) = s_0$ for some s, then either $s \notin \sigma(\mathbb{N}_0)$ or $s \in \sigma(\mathbb{N}_0)$). If $s \notin \sigma(\mathbb{N}_0)$, then $s_0 = \phi(s) = s$, but $s_0 = \sigma(0) \in \sigma(\mathbb{N}_0)$ a contradiction. In case $s \in \sigma(\mathbb{N}_0)$, we have

$$s_0 = \phi(s) = \sigma(\sigma^{-1}(s)+1), \tag{9.1.5}$$

from which, after applying σ^{-1} on both members and manipulating,

$$0 = \sigma(s_0) = \sigma^{-1}(s)+1, \tag{9.1.6}$$

which means that 0 is the successor the nonnegative integer $\sigma^{-1}(s)$, again a contradiction. So $\phi : S \to S$ is not surjective, yet surjective, hence S is infinite. □

9.1.2. Remark. In practical words, Lemma 9.1.1 says that a set is infinite if it contains a sequence indexed on $\mathbb{N}_0$ (or $\mathbb{N}$).

9.1.3. Proposition ($\mathbb{N}_0$ and $\mathbb{N}$ are infinite). *Any set S satisfying Peano's Axioms 6.4.3 is infinite.*
Proof By the Peano's Axiom, we know that there exists $\sigma : S \to S$ that is injective. Also we have that $0 \notin \operatorname{Img} \sigma$, hence σ is not surjective. Therefore, according to Definition 6.6.1, the set $\mathbb{N}$ is infinite, because it has an injective map that is not surjective. □

9.2. Countable and uncountable sets

9.2.1. Definition of countable. A set X is called *countable* if and only if there exists a subset S of $\mathbb{N}$ and a bijection $\phi : S \to X$. If there is no such map then X is called uncountable. If X is countable it is either finite or infinite; in the second case we say that X is *numerable* and its cardinality is *aleph-zero*

$$\#X = \aleph_0. \tag{9.2.1}$$

9.2.2. Remark (finite sets are countable). It follows from the the definition that all finite sets are countable.

9.2.3. Proposition. *Let X and Y be two sets related by a one-to-one correspondence (bijective map) $\phi : X \leftrightarrows Y$, then*

(a) X is countable if and only if Y is countable,

(b) X is infinite if and onluy if Y is infinite.

Proof Exercise. □

Exercise 9.2.4 (subsets of countable sets are countable)**.** Prove that if X is countable and $Y \subseteq X$, then Y is countable.

Exercise 9.2.5 (supersets of uncountable sets are uncountable)**.** Prove that if X is uncountable and $X \subseteq Y$, then Y is uncountable.

Problem 9.2.6. Prove that the set of rational numbers $\mathbb{Q}$ is countable.

9.2.7. Theorem (Cantor's diagonal). *The interval* $[0,1)$ *in* $\mathbb{R}$ *is uncountable.*
Proof Consider the decimal expansion[2] of all numbers between 0 and 1 in the form

$$0.x_1x_2x_3\dots \tag{9.2.1}$$

with $x_i = 0,\dots,9$ for each $i \in \mathbb{N}$. Some numbers have two decimal representations, exactly those that have all 0s after a certain point (finite decimal expansion) and those that have all 9s after a certain point, for example

$$0.20000\dots = 0.19999\dots \tag{9.2.2}$$

To avoid overlap exclude the expansions of the form $x_i = 9$ for all $i \geq k$ for some $k \geq 1$ as these will yield the same number as $y_{k-1} = x_{k-1} + 1$ and $y_i = 0$ for all $i \geq k$ if $k > 1$ and $y_{k-1} \neq 9$, or 1 otherwise, but $1 \notin I$. Let $\sigma : \mathbb{N} \to [0,1)$ be injective, we will show that σ cannot be surjective, which will prove that $[0,1)$ is uncountable. To show that σ is not surjective it will be enough to find $\hat{x} \in [0,1)$ for which $\sigma(j) \neq \hat{x}$ for all $i \in \mathbb{N}$. Write $x^j := \sigma(j)$ and consider the expansion

$$x^j = 0.x^j_1x^j_2\dots. \tag{9.2.3}$$

Introduce the map $s : [0\dots9] \to [0\dots9]$

$$s(d) := \begin{cases} d+1 & \text{for } d = 0,\dots,7 \\ 0 & \text{for } d = 8 \text{ or } 9. \end{cases} \tag{9.2.4}$$

and define $\hat{x}$ via its decimal expansion where

$$\hat{x}_i := s(x^i_i) \text{ for each } i \in \mathbb{N}. \tag{9.2.5}$$

By construction the $\hat{x}_i$s cannot be all 9s from a point on (in fact, no one $\hat{x}_i$ is nine), and $\hat{x}_i \neq x^i_i$ for all $i \in \mathbb{N}$. Therefore the decimal expansion of $\hat{x}$ does not coincide with any of those in the sequence $\sigma(j)$, so by uniqueness of expansion (having excluded the ultimately-all-nines ones) we deduce that $\hat{x} \neq \sigma(j)$ for all j as required. □

9.3. The Axiom Schema of Replacement

Axiom 8 (schema of replacement)**.** *Suppose ϕ is a formula that defines a set $\phi(x)$ for any set x, then for any set X there exists a set Y such that ϕ provides the formula for a function with domain X and codomain Y.*

[2] Any other basis will work too, with appropriate modifcations.

9.4. The Axiom of Choice

Axiom 9 (choice)**.** *Given two sets* X, Y*,* $Y \neq \varnothing$*, and a function* $\Phi : X \to \wp(Y) \setminus \{\varnothing\}$*, then there exists a choice function* $\phi : X \to Y$ *such that* $\phi(x) \in \Phi(x)$ *for each* $x \in X$*.*

Exercises and problems on infinity

Problem 9.X.1 (partial inverses and distinguished functions)**.** Let $f : A \to B$ be a function and A and B both nonempty. Show that

(i) f is injective if and only if it has a post-inverse (also known as left-inverse) g, i.e., $g : B \to A$ and $g \circ f = \mathrm{id}_A$.

(ii) f is surjective if and only if it has a pre-inverse (also known as right-inverse) g, i.e., $g : B \to A$ and $f \circ g = \mathrm{id}_B$.

Hint. For this part you need the *Axiom of Choice* in the following form: Given two sets X, Y, $Y \neq \varnothing$, and a function $\Phi : X \to \wp(Y) \smallsetminus \{\varnothing\}$, then there exists a *choice function* $\phi : X \to Y$ such that $\phi(x) \in \Phi(x)$ for each $x \in X$.

What can you say about the uniqueness of the partial inverse in each case.

Problem 9.X.2 (invertibility of surjective maps and axiom of choice equivalence)**.** Prove that the following are equivalent:

(i) for any sets A, B and any function $f : A \to B$,

$$f \text{ surjective } \Rightarrow f \text{ has a pre-inverse} \tag{9.X.2.1}$$

where *pre-inverse* (also known as *right-inverse*) of f means a function $g : B \to A$ such that $f \circ g = \mathrm{id}_B$;

(ii) for any nonempty collection $\mathscr{C}$ of pair-wise disjoint sets there exists a set C such that $C \cap X$ is a singleton for each $X \in \mathscr{C}$.

Bibliography

[1] Martin Aigner and Günter M. Ziegler. *Proofs from The Book*. Fifth. Including illustrations by Karl H. Hofmann. Springer-Verlag, Berlin, 2014. ISBN: 978-3-662-44204-3 978-3-662-44205-0. (Visited on 12/12/2016).

[2] Mario Dolcher. *Elementi di analisi matematica (1-2)*. it. Google-Books-ID: 8EB-jPQAACAAJ. Lint Editoriale, 1991. ISBN: 978-88-85083-59-2.

[3] Hagen von Eitzen. *Axiom of Choice and Right Inverse*. Mathematics Stack Exchange. URL:http://math.stackexchange.com/q/463060 (version: 2013-08-08). http://math.stack
von-eitzen, 2013. eprint: `http://math.stackexchange.com/q/463060`. (Visited on 08/08/2013).

[4] Ronald L. Graham, Donald E. Knuth and Oren Patashnik. *Concrete mathematics*. Second. A foundation for computer science. Reading, MA: Addison-Wesley Publishing Company, 1994, pp. xiv+657. ISBN: 0-201-55802-5.

[5] Paul R. Halmos. *Naive set theory*. Reprint of the 1960 edition, Undergraduate Texts in Mathematics. New York: Springer-Verlag, 1974, pp. vii+104.

[6] John M. Howie. *Real analysis*. Springer Undergraduate Mathematics Series. London: Springer-Verlag London Ltd., 2001, pp. x+276. ISBN: 1-85233-314-6.

[7] D. L. Johnson. *Elements of logic via numbers and sets*. Springer Undergraduate Mathematics Series. London: Springer-Verlag London Ltd., 1998, pp. x+174. ISBN: 3-540-76123-3.

[8] Geoff Smith. *Introductory mathematics: algebra and analysis*. Springer Undergraduate Mathematics Series. London: Springer-Verlag London Ltd., 1998, pp. xiv+215. ISBN: 3-540-76178-0.

[9] Terence Tao. *Structure and Randomness in the Prime Numbers, UCLA*. youtube. UCLA. Jan. 2009. (Visited on 09/11/2016).

Index

0, 128
0-tuple, 82
1, 130
A-monomial, 13
$\mathbb{F}_{0+}$, 166
$\mathbb{F}_{+}$, 166
$\mathbb{Q}_{0+}$, 166
$\mathbb{Q}_{+}$, 166
$\mathbb{R}_{0+}$, 166
$\mathbb{R}_{+}$, 166
$\aleph_0$, 188
$\mathbb{F}$-valued sequence, 172
$\mathbb{Q}$, 13, 170
$\mathbb{Z}$, 6
$\exists$, 8
$\forall$, 8
∞, 92
$\mathbb{F}_{-}$, 166
$\mathbb{F}_{0-}$, 166
$\mathbb{Q}_{-}$, 166
$\mathbb{Q}_{0-}$, 166
$\mathbb{R}_{-}$, 166
$\mathbb{R}_{0-}$, 166
card A, 92
pow^n, 177
pow$^{1/n}$, 178
$\cdot^n$, 177
$\sqrt[n]{\cdot}$, 178
n choose k, 104
$\binom{n}{k}$, 104
rational cuts, 181
1.15, Ex. 1.21, 185

Abelian
  group, 9
Abelian group, 164
Abraham
  Fraenkel, 67
Abraham Fraenkel, 67
absolute value, 13, 166
absorbing, 73, 90
ad infinitum, 129
addition, *see* sum
  seesum, 132
additive inverse, 164
adjacency array, 86
adjacency matrix, 86
aleph-zero, 188
algebra, 95, 183
almost all, 177
alone, 67
analysis, 95, 169, 184
antisymmetric, 147, 160
antisymmetry, 147, 160
archimedean, 160, 170, 183
archimedean field, 170
argument, 6, 93, 117, 119
arithmetic, 141
arithmetic power, 177
ary relation, 143
associative, 8, 9
  operation, 6
associativity, 73, 90
atomic predicate, 68
Augustin Cauchy, 164
Axiom
  of Extension, 68
  Specification, 72
axiom, 82, 90
Axiom of Choice, 191
Axiom of Foundation, 129
Axiom of Regularity, 129
Axiom of Separation, 72
Axiom of Specification, 72
axiom schema, 72
Axiom Schema of Specification, 72

basic, 159
belongs, 68
best bounds of a sequence, 14, 179
bijection, 121
bijective, 120, 121
binary operation, 6
binary relation, 143
binomial, 114
binomial coefficient, 104

Binomial Theorem, 104
bound, 57, 169, 184
bounded, 13, 173, 184
bounded above, 13, 41, 173
bounded below, 13, 173
bounded sequence, 173
bounded set, 14
bounds, 168

cancellation, 26, 141
cancellation property, 26
canonical map, 154
Cantor, 67
  Georg, 67
cardinal number, 92
cardinality, 92, 112, 114, 135, 142
cartesian, 112
cartesian cube, 87
cartesian product, 85
cartesian square, 85, 87, 145
Cassini, 31
Cauchy, 176, 183, 184
Cauchy sequence, 164
Cauchy sequence property, 174
Cauchy sequences, 163
certifying, 50, 56
Ch.2, 57
characterisation, 64
characteristic, 168
Chinese remainder, 57
Chinese remainder theorem, 57
choice function, 190, 191
choose, 114
class, 16
closed interval, 167
closed-open interval, 167
closure point property, 179
codomain, 94, 117, 119
coefficient, 114
cofactored-with relation, 144
collection, 16, 74
combinations, 113
combinatorics, 113, 114, 141
commutative
  operation, 6
commutative ring, 11
commutativity, 73, 90
compatibility, 26
compatible, 156
compatible operation, 12, 156
compatible relation, 156
compatible with respect to addition, 165
compatible with respect to multiplication, 165
compatible with the field structure, 165
compatiblity of ordering with addition, 26
complement
  set, 77
complete, 176
Completeness Axiom, 13, 176
composable
  functions, 122
composed
  function, 122
composite, 34
  function, 122
  number, 34
composition, 103
  of functions, 103, 122
composition power, 126
comprehension
  axiom, 72
condition
  necessary, 64
  sufficient, 64
congruence, 153
congruence class, 12, 151
congruent modulo, 150
constant, 13
constructive, 114
contradiction, *see* proof, 34
contraposition, *see* proof, 35
convergence, 13, 14, 172, 176, 183
convergence to 0, 13, 172
convergent, 176, 183
convergent sequence, 13, 172
converges to, 173
converges to 0, 172
coprime, 50
coprime set, 50
coprime-with relation, 144
coprime pair, 50
correspondence, 93
count, 92, 93, 112, 114, 141
countable, 188
counterimage, 9, 120
counterimage set, 120
counting, 113, 141, 142
counting map, 134
criterion, 64
Cryptology, 34
cube, 29, 177
cubic, 13
cubic root, 178
cycle, 11, 137
cycle notation, 137

David Hilbert, 67
De Morgan Laws, 65
decreasing, 14, 179
Dedekind cut, 164, 181

Dedekind cuts, 163
definiteness, 166
degree, 13
dependent variable, 6
descending factorial notation, 180
diagonal set, 145
difference
  set, 77
digital, 113
Diophantine equations, 56
directed graph, 86, 143
discrete, 127
disjoint, 78, 142
disjoint sets, 7, 78
distinguished, 141
diverge to negative infinity, 174
diverge to infinity, 173
diverge to minus infinity, 174
divergence to infinity, 13, 173
diverges in absolute value, 174
divide, 55
divide-and-average, 176, 183
divides, 4, 33
division
  Euclidean, 39
divisor, 33, 55, 89
domain, 94, 117, *see* function, 119

element, 68
elementary proposition, 68
elements
  of a set, 68
equal
  sets, 68
equal modulo, 150
equality, 145
  Axiom, 68
equivalance class generated, 151
equivalence, 149, 161
equivalence class, 12, 151
equivalence class represented, 151
Erdős
  Pál/Paul, 33
Ernst Zermelo, 67
errors, 172
Euclid, 53, 57
Euclid’s algorithm, 43
Euclid’s theorem, 53, 57
Euclidean algorithm, 43
Euclidean division, 5, 39
existential quantifier, 64
expansion, 113
Extension
  Axiom of, 68
extension
  by description, 69
Extension Axiom, 69
extremum, 13, 169

factor, 9, 33
factorial, 52
family, 74
Fibonacci, 31
Fibonacci numbers, 20
Fibonacci sequence, 20
finite, 92, 132, 141, 142
finite permutation, 136
finite sequence, 7, 82
finite set, 141
formula, 22, 71
  univariate, 71
free variables, 72
function, 93, 117, 141, 184

Gödel
  Kurt, 67
gap, 57
geometry, 159
Georg Cantor, 67
Giuseppe Peano, 67
graph, 86
  directed, 86
greatest, 169, 184
greatest closure point, 169, 184
greatest common divisor, 42
greatest element, 41
greatest lower bound, 13, 169
group, 16

half-bounded interval, 167
half-line, 167
HCF, 50, 57
Heron, 176, 183
highest common factor, 42, 50, 57
Hilbert
  David, 67
homogeneity, 166

identity, 9, 96, 124, 145, 164
identity function, 10, 124
  seeidentity, 96
ifisornot, 65
image, 6, 93, 117, 119, 120
  of a function, 120
image set, 119
implication, 65
inclusion–exclusion, 142
increasing, 14, 179
independent variable, 6, 93, 119
indeterminate, 13, 114

index, 17
induction, 21, 29, 30
Induction Principle, 4, 21, 131
inductive, 130
inequality, 30, 185
inference, 66
infimum, 169, 179
  sequence, 179
infinimum, 169, 184
infinite, 34, 92, 132
infinite limit, 13, 173
infinity, 53, 57, 92, 130
  Axiom of, 130
infinite set, 188
infinity of primes, 53, 57
injection, 99, 121, 141
injective, 99, 120, 121, 141
input, 6, 93
integer, 5, 6, 89, 141
integer arithmetic power, 177
integer segment, 141
intersection, 73, 74, 90
interval, 185
invariance, 26
invariance of bijectivity under composition, 124, 141
invariance of inequalities, 26
invariance of injectivity under composition, 124, 141
invariance of surjectivity under composition, 124, 141
invariant, 167
inverse, 9, 125
Inverse Function
  Theorem, 103
inverse function, 10, 103, 125
invertible, 125
irrational, 5, 37, 183
irrational cut, 181
irrational number, 177

Kronecker
  Leopold, 5
Kurt Gödel, 67

largest element, 41
law, 6
least, 169, 184
least element, 5, 38
least upper bound, 13, 169
left-inverse, 125
length, 33, 82
  of a sequence, 82
limit, 13, 15, 172, 183
linear, 13
linear Diophantine equations, 56
linearly ordered set, 165
locus, 16
logic, 55, 66
logical inference, 63
logical operation, 62
long cartesian product, 87
lower, 57, 169, 184
lower bound, 13, 168
lowest closure point, 170, 184

map, 93, 141
mapping, 93
matrix addition, 11
matrix multiplication, 11
maximum, 41, 169
measure theory, 177
meet, 73
member, 68
metric space, 14
minimality, 73, 90
minimum, 5, 38, 169
modulus, 13, 166
monoid, 126
monomial, 13
monontone, 14, 179
monotone, 184
monotonicity, 73, 90, 184
monotonicity property, 45
multiple, 33, 89
multiplication, 9
multiplicative inverse, 164
multivariate, 14
multivariate formula, 72
mutually prime, 50

natural, 141
natural map, 154
natural numbers, 130
natural numbers with 0, 128
necessary condition, 64
negative, 166
  number, 25
neutral element, 8, 9
nonnegative integers, 130
nonstrict, 148, 161
nonstrict ordering, 148
nonstrictly negative, 166
nonstrictly ordered field, 12, 165
nonstrictly positive, 166
number, 5, 50, 53, 55, 57, 141, 160, 169, 183–185
  rational, 7
numbers, 30
numerable, 188

one, 9, 164
one-to-one, 121, 141
one-to-one correspondence, 92, 121
one-to-one into, 99
onetuple, 82
onto, 120, 141
open interval, 167
open–closed interval, 167
operand, 93
operation, 6
  compatible with an equivalence, 156
  logical, 62
operational precendence, 164
operator
  unary, 93
opposite, 164
orbit, 138
ordered, 82, 90
ordered set, 164
ordered tuple, 7, 82
ordering, 148, 149, 160, 161, 169, 183, 184
order relation, 26
ordinal, 92
orthogonality relation, 144
outcome, 93
output, 6, 93

pair
  coprime, 50
pairing, 82, 90
pairwise coprime, 50
paraadox
  Russell's, 59
parallelism, 144
parameters, 72
partial, 148, 161
partial inverse, 10, 125
partial ordering, 148
partition, 152
partition, of a set, 152
Peano
  Giuseppe, 67
Peano's Axioms, 10, 128
Peano's axioms, 128
permutation, 136
Pigeonhole Principle, 10, 135
placeholder, 114
PMI, *see* induction, 131
pointwise addition, 181
pointwise product, 181
polynomial, 14
positive, 166
  number, 25, 26
positive integers, 130
positively oriented, 165
positivity, 166
post-inverse, 125
power, 30, 177
  arithmetic, 177
power function, 177
pre-inverse, 125, 191
predicate, 22
prime, 34, 53, 57
  number, 34
prime gap, 57
Principle of Mathematical Induction, 10, 21, 131
Problem 1.38, 30
product, 9, 112
product of two sequences
  algebraic, 172
proof
  by contradition, 34
  by contraposition, 35
proposition, 60

quadratic, 13
quantifier, 64
  specification,existential, 64
  universal, 64
quartic, 13
quasi-ordering, 12, 149, 161
quintic, 13
quintuple, 82
quotient set, 12, 154
quotient-set, 149, 161

range, 120
rational, 160, 183
rational cut, 181, 287
rational number, 7, 13, 170
real, 169, 183–185
realisation, 93
reciprocal, 164
recursion, 20, 29
recursive
  definition, 20
recursive expression, 114
recursive sequence, 20
reflexive, 159
regular, 129
  set, 129
regularity, 129
relation, 143, 159, 161
  compatible with an equivalence, 156
relatively prime, 50
representative, 12, 154
residue class, 12, 151
Richard Dedekind, 164
right-inverse, 125, 191

root, 178
rule, 6
Russell
  Bertrand, 59
Russell’s Paradox, 6, 59

segment, 141
  of integers, 33
segment of integers, 33
separation
  axiom, 72
septuple, 82
sequence, 14, 17, 176, 179, 183, 184
sequence in, 172
set, 74, 82, 89, 90, 142
set of natural numbers, 130
set of natural numbers with zero, 130
set-valued functions, 86
sextuple, 82
singleton, 82, 90
solving for, 185
Specification
  Axiom, 72
specification, 72
  quantifier, 64
square, 177, 184
square root, 5, 36, 178
squares, 29
statement, 60
strict, 161
strict ordering, 148
strictly negative, 166
strictly positive, 166
strictly positive integers, 130
strongly ordered set, 164
subadditivity, 166
successor, 127, 130
  map, 129
successor map, 10, 128, 129
sufficient condition, 64
sum, 6, 29
  of two natural numbers, 132
sum of two sequences
  alebraic, 172
summation, 29
supermum, 169
supremum, 169, 179, 184
  sequence, 179
surjection, 120, 141
surjective, 120, 141
swap, 137
syllogism, 66
symmetric group, 11, 136
symmsetric, 159

table, 86
tends to, 173
term
  of a sequence, 17
ternary relation, 143
the set counts to n, 134
theory, 50, 53, 55, 57, 74, 82, 90, 142
to converge to, 173
to tend to, 173
totally ordered set, 165
transformation, 93
transitive, 159
transitivity, 169, 184
transitivity of implication, 66
transposition, 11, 134, 137
transversal, 12, 154
triangle inequality, 166
trichotomy, 165
truth table, 64
tuple, 82, 90

ultimately dominate, 174, 186
ultimately increasing, 179
ultimately vanishing, 53
UMBER, 129
unary operator, 7
unary operation, 7
unary operator, 93
unary relation, 143
undirected graph, 86
undirected graphs, 86
union, 78
union-of-member-and-bracket-enclosed rule, 129
unique factorisation theorem, 51
uniqueness of limit, 172
unit, 9, 130, 164
univariate, 14
universal quantifier, 64
upper, 169, 184
upper bound, 41, 168

v-NUMBERS, 8, 83
vacuous truth, 78
value, 6, 93, 117
vanish, 183
vanishes, 172
vanishing sequence, 13, 172
variable, 6, 13, 117
von Neumann
  John, 83
Von Neumann’s construction, 83

weakly antisymmetric, 147, 160
Well Ordering Principle, 5, 38

well-defined operator (or function), 36
without loss of generality, 187

YBC7289, 163

Zermelo
  Ernst, 67
Zermelo–Fraenkel–Choice, 68
zero, 128, 164
ZFC, 68