A FIRST COURSE

IN

TOPOLOGY

A FIRST COURSE

IN

TOPOLOGY

MAT4002 Notebook

Lecturer

Prof. Daniel Wong

The Chinese University of Hongkong, Shenzhen

Tex Written By

Mr. Jie Wang

The Chinese University of Hongkong, Shenzhen

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments xv

Notations xvii

1 Week1 1

1.1 Monday for MAT3040 1

1.1.1 Introduction to Advanced Linear Algebra 1

1.1.2 Vector Spaces 2

1.2 Monday for MAT3006 5

1.2.1 Overview on uniform convergence 5

1.2.2 Introduction to MAT3006 6

1.2.3 Metric Spaces 7

1.3 Monday for MAT4002 11

1.3.1 Introduction to Topology 11

1.3.2 Metric Spaces 12

1.4 Wednesday for MAT3040 15

1.4.1 Review 15

1.4.2 Spanning Set 15

1.4.3 Linear Independence and Basis 17

1.5 Wednesday for MAT3006 21

1.5.1 Convergence of Sequences 21

1.5.2 Continuity 25

1.5.3 Open and Closed Sets 26

1.6 Wednesday for MAT4002 28

1.6.1 Forget about metric 28

1.6.2 Topological Spaces 31

1.6.3 Closed Subsets 32

2 Week2 35

2.1 Monday for MAT3040 35

2.1.1 Basis and Dimension 35

2.1.2 Operations on a vector space 38

2.2 Monday for MAT3006 41

2.2.1 Remark on Open and Closed Set 41

2.2.2 Boundary, Closure, and Interior 44

2.3 Monday for MAT4002 47

2.3.1 Convergence in topological space 47

2.3.2 Interior, Closure, Boundary 49

2.4 Wednesday for MAT3040 53

2.4.1 Remark on Direct Sum 53

2.4.2 Linear Transformation 54

2.5 Wednesday for MAT3006 61

2.5.1 Compactness 61

2.5.2 Completeness 66

2.6 Wednesday for MAT4002 68

2.6.1 Remark on Closure 68

2.6.2 Functions on Topological Space 70

2.6.3 Subspace Topology 72

2.6.4 Basis (Base) of a topology 74

3 Week3 75

3.1 Monday for MAT3040 75

3.1.1 Remarks on Isomorphism 75

3.1.2 Change of Basis and Matrix Representation 76

3.2 Monday for MAT3006 84
3.2.1 Remarks on Completeness 84
3.2.2 Contraction Mapping Theorem 85
3.2.3 Picard Lindelof Theorem 89
3.3 Monday for MAT4002 90
3.3.1 Remarks on Basis and Homeomorphism 90
3.3.2 Product Space 93
3.4 Wednesday for MAT3040 95
3.4.1 Remarks for the Change of Basis 95
3.5 Wednesday for MAT3006 102
3.5.1 Remarks on Contraction 102
3.5.2 Picard-Lindelof Theorem 102
3.6 Wednesday for MAT4002 107
3.6.1 Remarks on product space 107
3.6.2 Properties of Topological Spaces 110
4 Week4 113
4.1 Monday for MAT3040 113
4.1.1 Quotient Spaces 113
4.1.2 First Isomorphism Theorem 116
4.2 Monday for MAT3006 119
4.2.1 Generalization into System of ODEs 119
4.2.2 Stone-Weierstrass Theorem 121
4.3 Monday for MAT4002 125
4.3.1 Hausdorffness 125
4.3.2 Connectedness 126
4.4 Wednesday for MAT3040 130
4.4.1 Dual Space 135

4.5 Wednesday for MAT3006 138

4.5.1 Stone-Weierstrass Theorem . . . 139

4.6 Wednesday for MAT4002 144

4.6.1 Remark on Connectedness . . . 144

4.6.2 Compactness . . . 146

5 Week5 . . . 149

5.1 Monday for MAT3040 149

5.1.1 Remarks on Dual Space . . . 150

5.1.2 Annihilators . . . 152

5.2 Monday for MAT3006 156

5.2.1 Stone-Weierstrass Theorem in $\mathbb{C}$ . . . 157

5.2.2 Baire Category Theorem . . . 159

5.3 Monday for MAT4002 161

5.3.1 Continuous Functions on Compact Space . . . 161

5.4 Wednesday for MAT3040 164

5.4.1 Adjoint Map . . . 165

5.4.2 Relationship between Annihilator and dual of quotient spaces . . . 168

5.5 Wednesday for MAT3006 169

5.5.1 Remarks on Baire Category Theorem . . . 169

5.5.2 Compact subsets of $C[a,b]$ . . . 171

5.6 Wednesday for MAT4002 173

5.6.1 Remarks on Compactness . . . 173

5.6.2 Quotient Spaces . . . 174

6 Week6 . . . 179

6.1 Monday for MAT3040 179

6.1.1 Polynomials . . . 179

6.2 Monday for MAT3006 182

6.2.1 Compactness in Functional Space 182

6.2.2 An Application of Ascoli-Arzela Theorem 183

6.3 Monday for MAT4002 184

6.3.1 Quotient Topology 184

6.3.2 Properties in quotient spaces 186

6.4 Wednesday for MAT3040 191

6.4.1 Eigenvalues & Eigenvectors 194

6.5 Wednesday for MAT4002 197

6.5.1 Remarks on Compactness 197

7 Week7 199

7.1 Monday for MAT3040 199

7.1.1 Minimal Polynomial 199

7.1.2 Minimal Polynomial of a vector 204

7.2 Monday for MAT3006 206

7.2.1 Remarks on the outer measure 206

7.3 Monday for MAT4002 211

7.3.1 Quotient Map 211

7.3.2 Simplicial Complex 212

7.4 Wednesday for MAT3040 217

7.4.1 Cayley-Hamiton Theorem 217

7.5 Wednesday for MAT4002 224

7.5.1 Remarks on Triangulation 224

7.5.2 Simplicial Subcomplex 226

7.5.3 Some properties of simplicial complex 228

8 Week8 231

8.1 Monday for MAT3040 231

8.1.1 Cayley-Hamiton Theorem 233

8.1.2 Primary Decomposition Theorem 236

8.2 Monday for MAT3006 238

8.2.1 Remarks for Outer Measure 238

8.2.2 Lebesgue Measurable 239

8.3 Monday for MAT4002 244

8.3.1 Quotient Map 244

8.3.2 Simplicial Complex 245

8.4 Wednesday for MAT3006 248

8.4.1 Remarks on Lebesgue Measurability 248

8.4.2 Measures In Probability Theory 250

8.5 Wednesday for MAT4002 253

8.5.1 Homotopy 255

9 Week9 257

9.1 Monday for MAT3040 257

9.1.1 Remarks on Primary Decomposition Theorem 257

9.2 Monday for MAT3006 263

9.2.1 Measurable Functions 263

9.3 Monday for MAT4002 268

9.3.1 Remarks on Homotopy 268

9.4 Wednesday for MAT3040 274

9.4.1 Jordan Normal Form 274

9.4.2 Inner Product Spaces 280

9.5 Wednesday for MAT3006 282

9.5.1 Remarks on Measurable function 282

9.5.2 Lebesgue Integration 283

9.6 Wednesday for MAT4002 287

9.6.1 Simplicial Approximation Theorem 287

10 Week10 293

10.1 Monday for MAT3040 293

10.1.1 Inner Product Space 293

10.1.2 Dual spaces 296

10.2 Monday for MAT3006 299

10.2.1 Remarks on Markov Inequality 299

10.2.2 Properties of Lebesgue Integration 299

10.3 Monday for MAT4002 302

10.3.1 Group Presentations 304

10.4 Wednesday for MAT3040 305

10.4.1 Orthogonal Complement 305

10.4.2 Adjoint Map 308

10.5 Wednesday for MAT3006 311

10.5.1 Consequences of MCT 314

10.6 Wednesday for MAT4002 317

10.6.1 Reviewing On Groups 317

10.6.2 Free Groups 319

10.6.3 Relations on Free Groups 321

11 Week11 323

11.1 Monday for MAT3040 323

11.1.1 Self-Adjoint Operator 323

11.1.2 Orthononal/Unitary Operators 326

11.2 Monday for MAT3006 328

11.2.1 Consequences of MCT I 328

11.2.2 MCT II . . . 331

11.3 Monday for MAT4002 333

11.3.1 Cayley Graph for finitely presented groups . . . 335

11.3.2 Fundamental Group . . . 337

11.4 Wednesday for MAT3040 340

11.4.1 Unitary Operator . . . 340

11.4.2 Normal Operators . . . 343

11.5 Wednesday for MAT3006 345

11.5.1 Properties of Lebesgue Integrable Functions . . . 346

11.6 Wednesday for MAT4002 351

11.6.1 The fundamental group . . . 351

12 Week12 . . . 359

12.1 Monday for MAT3040 359

12.1.1 Remarks on Normal Operator . . . 359

12.1.2 Tensor Product . . . 363

12.2 Monday for MAT3006 365

12.2.1 Remarks on MCT . . . 365

12.2.2 Dominated Convergence Theorem . . . 368

12.3 Monday for MAT4002 372

12.3.1 Some basic results on $\pi_1(X,b)$ . . . 375

12.4 Wednesday for MAT3040 377

12.4.1 Introduction to Tensor Product . . . 377

12.5 Wednesday for MAT3006 381

12.5.1 Riemann Integration & Lebesgue Integration . . . 381

12.5.2 Continuous Parameter DCT . . . 383

12.6 Wednesday for MAT4002 386

12.6.1 Groups & Simplicial Complices . . . 386

13 Week13 391

13.1 Monday for MAT3040 391

13.1.1 Basis of $V \otimes W$ 393

13.1.2 Tensor Product of Linear Transformation 397

13.2 Monday for MAT3006 398

13.2.1 Double Integral 399

13.3 Monday for MAT4002 403

13.3.1 Isomorphsim between Edge Loop Group and the Fundamental Group 403

13.4 Wednesday for MAT3040 409

13.4.1 Tensor Product for Linear Transformations 409

13.5 Wednesday for MAT3006 415

13.5.1 Fubini's and Tonell's Theorem 415

13.6 Wednesday for MAT4002 420

13.6.1 Applications on the isomorphism of fundamental group 420

14 Week14 427

14.1 Monday for MAT3040 427

14.1.1 Multilinear Tensor Product 427

14.1.2 Exterior Power 430

14.2 Monday for MAT3006 432

14.2.1 Tonelli's and Fubini's Theorem 432

14.3 Monday for MAT4002 436

14.3.1 Fundamental group of a Graph 436

15 Week15 441

15.1 Monday for MAT3040 441

15.1.1 More on Exterior Power 441

15.1.2 Determinant 443

15.2 Monday for MAT3006 **448**

15.2.1 Applications on the Tonell's and Fubini's Theorem 448

15.2.2 Final Review 452

15.3 Monday for MAT4002 **458**

15.3.1 The Selfert-Van Kampen Theorem 459

Acknowledgments

This book is taken notes from the MAT4002 in spring semester, 2019. These lecture notes were taken and compiled in LaTeX by Jie Wang, an undergraduate student in spring 2019. The tex writter would like to thank Prof. Daniel Wong and some students for their detailed and valuable comments and suggestions, which significantly improved the quality of this notebook. Students taking this course may use the notes as part of their reading and reference materials. This version of the lecture notes were revised and extended for many times, but may still contain many mistakes and typos, including English grammatical and spelling errors, in the notes. It would be greatly appreciated if those students, who will use the notes as their reading or reference material, tell any mistakes and typos to Jie Wang for improving this notebook.

Notations and Conventions

$(X, \mathcal{T})$	Topological space
$X \cong Y$	The space X is homeomorphic to space Y
$G \cong H$	The group G is isomorphic to group H
p_X	Project mapping
$X \times Y$	Product Topology
$X/\sim$	Quotient Topology related to the topologcial space X and the equivalence class $\sim$
S^n	The n-sphere $\{\boldsymbol{x} \in \mathbb{R}^{n+1} \mid \\|\boldsymbol{x}\\| = 1\}$
D^n	The n-disk $\{\boldsymbol{x} \in \mathbb{R}^n \mid \\|\boldsymbol{x}\\| \leq 1\}$
$E^\circ, \partial E, \overline{E}$	The interior, boundary, closure of E
$\mathbb{T}^2$	The torus in $\mathbb{R}^3$
Δ^n	The n-simplex
$i : A \hookrightarrow X$	Inclusion mapping from $A \subseteq X$ to X
$K = (V, \Sigma)$	(Abstract) Simplicial Complex
$\lvert K \rvert$	Topological realization of the simplicial complex K
$\langle X \mid R \rangle$	The presentation of a group
$H : f \overset{H}{\simeq} g$	f and g are homotopic, where H denotes the homotopy
$X \simeq Y$	The space X and Y are homotopy equivalent
$\pi_1(X, x)$	The fundamental group of X w.r.t. the basepoint $x \in X$
$E(K, b)$	The edge loop group of the space K w.r.t. the basepoint b
f_*	The induced homomorphism $f_* : \pi_1(X, x) \to \pi_1(Y, y)$ for $f : X \to Y$

Chapter 1

Week1

1.1. Monday for MAT3040

1.1.1. Introduction to Advanced Linear Algebra

Advanced Linear Algebra is one of the most important course in MATH major, with pre-request MAT2040. This course will offer the really linear algebra knowledge.

What the content will be covered?.

- In MAT2040 we have studied the space $\mathbb{R}^n$; while in MAT3040 we will study the general vector space V.
- In MAT2040 we have studied the *linear transformation* between Euclidean spaces, i.e., $T : \mathbb{R}^n \to \mathbb{R}^m$; while in MAT3040 we will study the linear transformation from vector spaces to vector spaces: $T : V \to W$
- In MAT2040 we have studied the eigenvalues of $n \times n$ matrix $\boldsymbol{A}$; while in MAT3040 we will study the eigenvalues of a **linear operator** $T : V \to V$.
- In MAT2040 we have studied the dot product $\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i$; while in MAT3040 we will study the **inner product** $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$.

Why do we do the generalization?. We are studying many other spaces, e.g., $C(\mathbb{R})$ is called the space of all functions on $\mathbb{R}$, $C^\infty(\mathbb{R})$ is called the space of all infinitely differentiable functions on $\mathbb{R}$, $\mathbb{R}[x]$ is the space of polynomials of one-variable.

■ **Example 1.1** 1. Consider the Laplace equation $\Delta f = 0$ with linear operator Δ:

$$\Delta : C^\infty(\mathbb{R}^3) \to C^\infty(\mathbb{R}^3) \quad f \mapsto (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})f$$

The solution to the PDE $\Delta f = 0$ is the 0-eigenspace of Δ.

2. Consider the Schrödinger equation $\hat{H}f = Ef$ with the linear operator

$$\hat{H} : C^\infty(\mathbb{R}^3) \to C^\infty(\mathbb{R}^3), \quad f \to \left[\frac{-\hbar^2}{2\mu}\nabla^2 + V(x,y,z)\right] f$$

Solving the equation $\hat{H}f = Ef$ is equivalent to finding the eigenvectors of $\hat{H}$. In fact, the eigenvalues of $\hat{H}$ are **discrete**.

■

1.1.2. Vector Spaces

Definition 1.1 [Vector Space] A **vector space** over a field $\mathbb{F}$ (in particular, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) is a set of objects V equipped with vector addiction and scalar multiplication such that

1. the vector addiction $+$ is closed with the rules:

   (a) **Commutativity**: $\forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V$, $\boldsymbol{v}_1 + \boldsymbol{v}_2 = \boldsymbol{v}_2 + \boldsymbol{v}_1$.

   (b) **Associativity**: $\boldsymbol{v}_1 + (\boldsymbol{v}_2 + \boldsymbol{v}_3) = (\boldsymbol{v}_1 + \boldsymbol{v}_2) + \boldsymbol{v}_3$.

   (c) **Addictive Identity**: $\exists \mathbf{0} \in V$ such that $\mathbf{0} + \boldsymbol{v} = \boldsymbol{v}$, $\forall \boldsymbol{v} \in V$.

2. the **scalar multiplication** is closed with the rules:

   (a) **Distributive**: $\alpha(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \alpha\boldsymbol{v}_1 + \alpha\boldsymbol{v}_2, \forall \alpha \in \mathbb{F}$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$

   (b) **Distributive**: $(\alpha_1 + \alpha_2)\boldsymbol{v} = \alpha_1\boldsymbol{v} + \alpha_2\boldsymbol{v}$

   (c) **Compatibility**: $a(b\boldsymbol{v}) = (ab)\boldsymbol{v}$ for $\forall a, b \in \mathbb{F}$ and $\boldsymbol{b} \in V$.

   (d) $0\boldsymbol{v} = \mathbf{0}$, $1\boldsymbol{v} = \boldsymbol{v}$.

■

Here we study several examples of vector spaces:

■ **Example 1.2** For $V = \mathbb{F}^n$, we can define

1. Addictive Identity:
$$\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$
2. Scalar Multiplication:
$$\alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}$$
3. Vector Addiction:
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

■

■ **Example 1.3** 1. It is clear that the set $V = M_{n\times n}(\mathbb{F})$ (the set of all $m \times n$ matrices) is a vector space as well.

2. The set $V = C(\mathbb{R})$ is a vector space:
   (a) Vector Addiction:
   $$(f+g)(x) = f(x) + g(x), \forall f, g \in V$$
   (b) Scalar Multiplication:
   $$(\alpha f)(x) = \alpha f(x), \forall \alpha \in \mathbb{R}, f \in V$$
   (c) Addictive Identity is a zero function, i.e., $\mathbf{0}(x) = 0$ for all $x \in \mathbb{R}$.

■

Definition 1.2 A sub-collection $W \subseteq V$ of a vector space V is called a **vector subspace of** V if W itself forms a vector space, denoted by $W \le V$.

■ **Example 1.4**

1. For $V = \mathbb{R}^3$, we claim that $W = \{(x, y, 0) \mid x, y \in \mathbb{R}\} \le V$
2. $W = \{(x, y, 1) \mid x, y \in \mathbb{R}\}$ is not the vector subspace of V.

Proposition 1.1 $W \subseteq V$ is a **vector subspace** of V iff for $\forall \boldsymbol{w}_1, \boldsymbol{w}_2 \in W$, we have $\alpha \boldsymbol{w}_1 + \beta \boldsymbol{w}_2 \in W$, for $\forall \alpha, \beta \in \mathbb{F}$.

■ **Example 1.5**

1. For $V = M_{n \times n}(\mathbb{F})$, the subspace $W = \{A \in V \mid \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}\} \le V$
2. For $V = C^{\infty}(\mathbb{R})$, define $W = \{f \in V \mid \frac{\mathrm{d}^2}{\mathrm{d}x^2} f + f = 0\} \le V$. For $f, g \in W$, we have

$$(\alpha f + \beta g)'' = \alpha f'' + \beta g'' = \alpha(-f) + \beta(-g) = -(\alpha f + \beta g),$$

which implies $(\alpha f + \beta g)'' + (\alpha f + \beta g) = 0$.

1.2. Monday for MAT3006

1.2.1. Overview on uniform convergence

Definition 1.3 [Convergence] Let $f_n(x)$ be a sequence of functions on an interval $I = [a,b]$. Then $f_n(x)$ converges **pointwise** to $f(x)$ (i.e., $f_n(x_0) \to f(x_0)$) for $\forall x_0 \in I$, if

$$\forall \varepsilon > 0, \exists N_{x_0,\varepsilon} \text{ such that } |f_n(x_0) - f(x_0)| < \varepsilon, \forall n \ge N_{x_0,\varepsilon}$$

We say $f_n(x)$ converges uniformly to $f(x)$, (i.e., $f_n(x) \rightrightarrows f(x)$) for $\forall x_0 \in I$, if

$$\forall \varepsilon > 0, \exists N_{\varepsilon} \text{ such that } |f_n(x_0) - f(x_0)| < \varepsilon, \forall n \ge N_{\varepsilon}$$

■

■ **Example 1.6** It is clear that the function $f_n(x) = \frac{n}{1+nx}$ converges pointwise into $f(x) = \frac{1}{x}$ on $[0,\infty)$, and it is uniformly convergent on $[1,\infty)$. ■

Proposition 1.2 If $\{f_n\}$ is a sequence of continuous functions on I, and $f_n(x) \rightrightarrows f(x)$, then the following results hold:

1. $f(x)$ is continuous on I.
2. f is (Riemann) integrable with $\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x$.
3. Suppose furthermore that $f_n(x)$ is **continuously differentiable**, and $f_n'(x) \rightrightarrows g(x)$, then $f(x)$ is differentiable, with $f_n'(x) \to f'(x)$.

We can put the discussions above into the content of series, i.e., $f_n(x) = \sum_{k=1}^n S_k(x)$.

Proposition 1.3 If $S_k(x)$ is continuous for $\forall k$, and $\sum_{k=1}^n S_k \rightrightarrows \sum_{k=1}^\infty S_k$, then

1. $\sum_{k=1}^\infty S_k(x)$ is continuous,
2. The series $\sum_{k=1}^\infty S_k$ is (Riemann) integrable, with $\sum_{k=1}^\infty \int_a^b S_k(x)\,\mathrm{d}x = \int_a^b \sum_{k=1}^\infty S_k(x)\,\mathrm{d}x$
3. If $\sum_{k=1}^n S_k$ is continuously differentiable, and the derivative of which is uniform

convergent, then the series $\sum_{k=1}^{\infty} S_k$ is differentiable, with

$$\left(\sum_{k=1}^{\infty} S_k(x)\right)' = \sum_{k=1}^{\infty} S_k'(x)$$

Then we can discuss the properties for a special kind of series, say power series.

Proposition 1.4 Suppose the power series $f(x) = \sum_{k=1}^{\infty} a_k x^k$ has radius of convergence R, then

1. $\sum_{k=1}^{n} a_k x^k \rightrightarrows f(x)$ for any $[-L, L]$ with $L < R$.
2. The function $f(x)$ is continuous on $(-R, R)$, and moreover, is differentiable and (Riemann) integrable on $[-L, L]$ with $L < R$:

$$\int_0^x f(t)\,\mathrm{d}t = \sum_{k=1}^{\infty} \frac{a_k}{k+1} x^{k+1}$$
$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$$

1.2.2. Introduction to MAT3006

What are we going to do.

1. (a) Generalize our study of (sequence, series, functions) on $\mathbb{R}^n$ into a metric space.
   (b) We will study spaces outside $\mathbb{R}^n$.

   Remark:

   - For (a), different metric may yield different kind of convergence of sequences. For (b), one important example we will study is $X = C[a, b]$ (all continuous functions defined on $[a, b]$.) We will generalize X into $C_b(E)$, which means the set of bounded continuous functions defined on $E \subseteq \mathbb{R}^n$.
   - The insights of analysis is to find a **unified** theory to study sequences/series on a metric space X, e.g., $X = \mathbb{R}^n, C[a, b]$. In particular, for $C[a, b]$, we will see that

- most functions in $C[a,b]$ are nowhere differentiable. (repeat part of content in MAT2006)
- We will prove the existence and uniqueness of ODEs.
- the set $\text{poly}[a,b]$ (the set of polynomials on $[a,b]$) is dense in $C[a,b]$. (analogy: $\mathbb{Q} \subseteq \mathbb{R}$ is dense)

2. Introduction to the Lebesgue Integration.

   For convergence of integration $\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)$, we need the pre-conditions (a) $f_n(x)$ is continuous, and (b) $f_n(x) \rightrightarrows f(x)$. The natural question is that can we relax these conditions to

   (a) $f_n(x)$ is integrable?

   (b) $f_n(x) \to f(x)$ pointwisely?

   The answer is yes, by using the tool of Lebesgue integration. If $f_n(x) \to f(x)$ and $f_n(x)$ is Lebesgue integrable, then $\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x$, which is so called the **dominated convergence**.

1.2.3. Metric Spaces

We will study the **length** of an element, or the **distance** between two elements in an arbitrary set X. First let's discuss the length defined on a well-structured set, say vector space.

Definition 1.4 [Normed Space] Let X be a vector space. A **norm** on X is a function $\|\cdot\| : X \to \mathbb{R}$ such that

1. $\|\boldsymbol{x}\| \ge 0$ for $\forall \boldsymbol{x} \in X$, with equality iff $\boldsymbol{x} = \boldsymbol{0}$
2. $\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|$, for $\forall \alpha \in \mathbb{R}$ and $\boldsymbol{x} \in X$.
3. $\|\boldsymbol{x} + \boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ (triangular inequality)

Any vector space equipped with $\|\cdot\|$ is called a **normed space**. ▪

Example 1.7 1. For $X = \mathbb{R}^n$, define

$$\|\boldsymbol{x}\|_2 = \left(\sum_{i=1}^n x_i^2\right)^{1/2} \quad \text{(Euclidean Norm)}$$

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \quad (p\text{-norm})$$

2. For $X = C[a,b]$, define

$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$$

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{1/p}$$

Exercise: check the norm defined above are well-defined.

Here we can define the distance in an arbitrary set:

Definition 1.5 A set X is a **metric space** with metric (X,d) if there exists a (distance) function $d : X \times X \to \mathbb{R}$ such that

1. $d(\boldsymbol{x},\boldsymbol{y}) \geq 0$ for $\forall \boldsymbol{x},\boldsymbol{y} \in X$, with equality iff $\boldsymbol{x} = \boldsymbol{y}$.
2. $d(\boldsymbol{x},\boldsymbol{y}) = d(\boldsymbol{y},\boldsymbol{x})$.
3. $d(\boldsymbol{x},\boldsymbol{z}) \leq d(\boldsymbol{x},\boldsymbol{y}) + d(\boldsymbol{y},\boldsymbol{z})$.

Example 1.8 1. If X is a normed space, then define $d(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x}-\boldsymbol{y}\|$, which is so called the metric induced from the norm $\|\cdot\|$.

2. Let X be any (non-empty) set with $\boldsymbol{x},\boldsymbol{y} \in X$, the discrete metric is given by:

$$d(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} 0, & \text{if } x = y \\ 1, & \text{if } x \neq y \end{cases}$$

Exercise: check the metric space defined above are well-defined.

R We will mostly study metric spaces whose metrics come from the norm of a normed space. Adopting the infinite norm discussed in Example (1.7), we can define a metric on $C[a,b]$ by

$$d_\infty(f,g) = \|f-g\|_\infty := \max_{x\in[a,b]} |f(x)-g(x)|$$

which is the correct metric to study the uniform convergence for $\{f_n\} \subseteq C[a,b]$.

Definition 1.6 Let (X,d) be a metric space. An **open ball** centered at $\boldsymbol{x} \in X$ of radius r is the set

$$B_r(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid d(\boldsymbol{x},\boldsymbol{y}) < r\}.$$

■ **Example 1.9** 1. For $X = \mathbb{R}^2$, we can draw the $B_1(\mathbf{0})$ with respect to the metrics d_1, d_2:

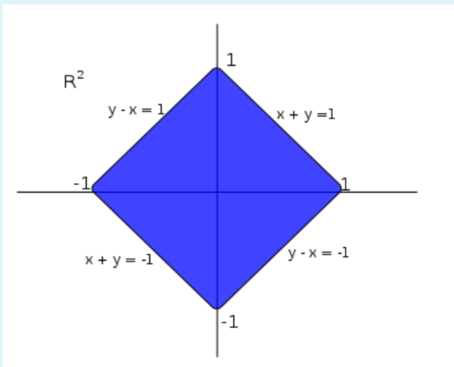


Figure 1.1: $B_1(\mathbf{0})$ w.r.t. the metric d_1

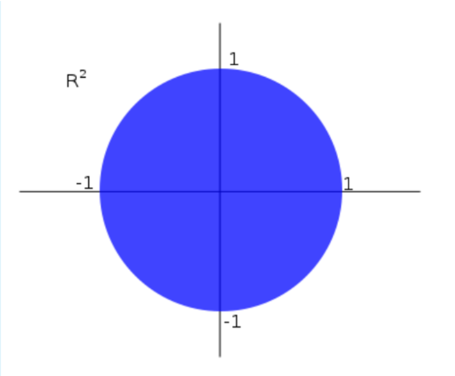
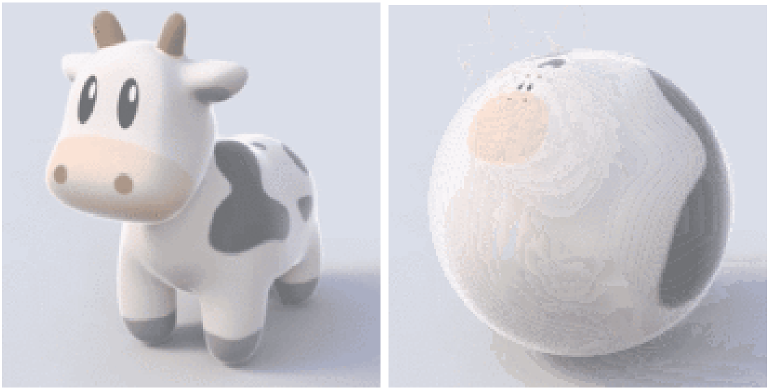


Figure 1.2: $B_1(\mathbf{0})$ w.r.t. the metric d_2

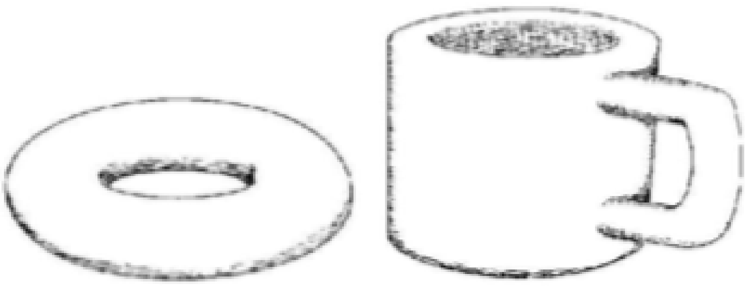
1.3. Monday for MAT4002

1.3.1. Introduction to Topology

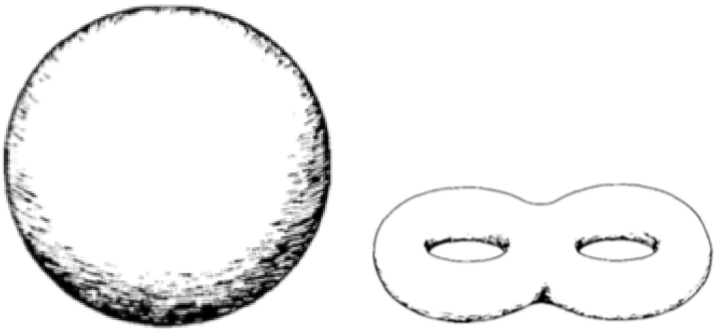
We will study global properties of a geometric object, i.e., *the distrance between 2 points in an object is totally ignored*. For example, the objects shown below are essentially invariant under a certain kind of transformation:



Another example is that the coffee cup and the donut have the same topology:



However, the two objects below have the intrinsically different topologies:



In this course, we will study the phenomenon described above mathematically.

1.3.2. Metric Spaces

In order to ingnore about the distances, we need to learn about distances first.

Definition 1.7 [Metric Space] Metric space is a set X where one can measure distance between any two objects in X.

Specifically speaking, a metric space X is a non-empty set endowed with a function (distance function) $d : X \times X \to \mathbb{R}$ such that

1. $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$ for $\forall \boldsymbol{x}, \boldsymbol{y} \in X$ with equality iff $\boldsymbol{x} = \boldsymbol{y}$
2. $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$
3. $d(\boldsymbol{x}, \boldsymbol{z}) \leq d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z})$ (triangular inequality)

■

■ **Example 1.10**

1. Let $X = \mathbb{R}^n$, with

$$d_2(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

$$d_\infty(\boldsymbol{x}, \boldsymbol{y}) = \max_{i=1,\dots,n} |x_i - y_i|$$

2. Let X be any set, and define the discrete metric

$$d(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0, & \text{if } x = y \\ 1, & \text{if } x \neq y \end{cases}$$

Homework: Show that (1) and (2) defines a metric. ■

Definition 1.8 [Open Ball] An **open ball** of radius r centered at $\boldsymbol{x} \in X$ is the set

$$B_r(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid d(\boldsymbol{x}, \boldsymbol{y}) < r\}$$

■

■ **Example 1.11** 1. The set $B_1(0,0)$ defines an open ball under the metric $(X = \mathbb{R}^2, d_2)$, or the metric $(X = \mathbb{R}^2, d_\infty)$. The corresponding diagram is shown below:

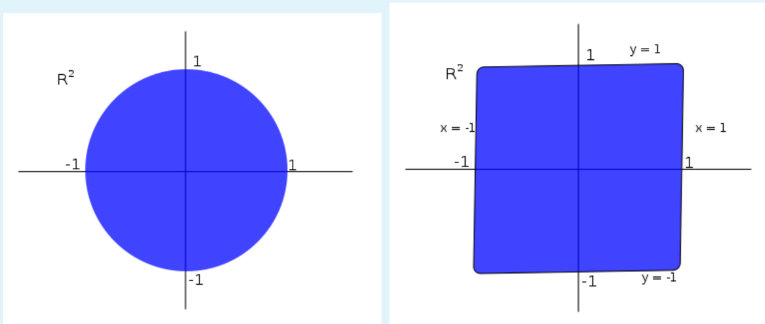


Figure 1.3: Left: under the metric $(X = \mathbb{R}^2, d_2)$; Right: under the metric $(X = \mathbb{R}^2, d_\infty)$

2. Under the metric $(X = \mathbb{R}^2, \text{discrete metric})$, the set $B_1(0,0)$ is one single point, also defines an open ball.

■

Definition 1.9 [Open Set] Let X be a metric space, $U \subseteq X$ is an open set in X if $\forall u \in U$, there exists $\epsilon_u > 0$ such that $B_{\epsilon_u}(u) \subseteq U$. ■

Definition 1.10 The **topology** induced from (X, d) is the collection of all open sets in (X, d), denoted as the symbol $\mathcal{T}$. ■

Proposition 1.5 All open balls $B_r(\boldsymbol{x})$ are open in (X, d).

Proof. Consider the example $X = \mathbb{R}$ with metric d_2. Therefore $B_r(x) = (x - r, x + r)$. Take $\boldsymbol{y} \in B_r(\boldsymbol{x})$ such that $d(\boldsymbol{x}, \boldsymbol{y}) = q < r$ and consider $B_{(r-q)/2}(\boldsymbol{y})$: for all $z \in B_{(r-q)/2}(\boldsymbol{y})$, we have

$$d(\boldsymbol{x}, \boldsymbol{z}) \le d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z}) < q + \frac{r-q}{2} < r,$$

which implies $\boldsymbol{z} \in B_r(x)$. ■

Proposition 1.6 Let $(X, \boldsymbol{d})$ be a metric space, and $\mathcal{T}$ is the topology induced from (X, d), then

1. let the set $\{G_\alpha \mid \alpha \in \mathcal{A}\}$ be a collection of (uncountable) open sets, i.e., $G_\alpha \in \mathcal{T}$,

then $\bigcup_{\alpha \in \mathcal{A}} G_\alpha \in \mathcal{T}$.

2. let $G_1, \ldots, G_n \in \mathcal{T}$, then $\bigcap_{i=1}^n G_i \in \mathcal{T}$. The finite intersection of open sets is open.

Proof. 1. Take $x \in \bigcup_{\alpha \in \mathcal{A}} G_\alpha$, then $x \in G_\beta$ for some $\beta \in \mathcal{A}$. Since G_β is open, there exists $\epsilon_x > 0$ s.t.

$$B_{\epsilon_x}(x) \subseteq G_\beta \subseteq \bigcup_{\alpha \in \mathcal{A}} G_\alpha$$

2. Take $x \in \bigcap_{i=1}^n G_i$, i.e., $x \in G_i$ for $i = 1, \ldots, n$, i.e., there exists $\epsilon_i > 0$ such that $B_{\epsilon_i}(x) \subseteq G_i$ for $i = 1, \ldots, n$. Take $\epsilon = \min\{\epsilon_1, \ldots, \epsilon_n\}$, which implies

$$B_\epsilon(x) \subseteq B_{\epsilon_i}(x) \subseteq G_i, \forall i$$

which implies $B_\epsilon(x) \subseteq \bigcap_{i=1}^n G_i$

■

Exercise.

1. let $\mathcal{T}_2, \mathcal{T}_\infty$ be topologies induced from the metrices d_2, d_∞ in $\mathbb{R}^2$. Show that $J_2 = J_\infty$, i.e., every open set in $(\mathbb{R}^2, d_2)$ is open in $(\mathbb{R}^2, d_\infty)$, and every open set in $(\mathbb{R}^2, d_\infty)$ is open in $(\mathbb{R}_2, d_2)$.
2. Let $\mathcal{T}$ be the topology induced from the discrete metric (X, d_{discrete}). What is $\mathcal{T}$?

1.4. Wednesday for MAT3040

1.4.1. Review

1. Vector Space: e.g., $\mathbb{R}, M_{n\times n}(\mathbb{R}), C(\mathbb{R}^n), \mathbb{R}[x]$.
2. Vector Subspace: $W \le V$, e.g.,
   (a) $V = \mathbb{R}^2$, the set $W := \mathbb{R}^2_+$ is not a vector subspace since W is not closed under scalar multiplication;
   (b) the set $W = \mathbb{R}^2_+ \bigcup \mathbb{R}^2_-$ is not a vector subspace since it is not closed under addition.
   (c) For $V = \mathbb{M}_{3\times 3}(\mathbb{R})$, the set of invertible 3×3 matrices is not a vector subspace, since we cannot define zero vector inside.
   (d) Exercise: How about the set of all singular matrices? Answer: it is not a vector subspace since the vector addition does not necessarily hold.

1.4.2. Spanning Set

Definition 1.11 [Span] Let V be a vector space over $\mathbb{F}$:

1. A linear combination of a subset S in V is of the form

$$\sum_{i=1}^{n}\alpha_i\boldsymbol{s}_i, \quad \alpha_i\in\mathbb{F}, \boldsymbol{s}_i\in S$$

   Note that the summation should be finite.

2. The **span** of a subset $S\subseteq V$ is

$$\operatorname{span}(S)=\left\{\sum_{i=1}^{n}\alpha_i\boldsymbol{s}_i \,\middle|\, \alpha_i\in\mathbb{F}, \boldsymbol{s}_i\in S\right\}$$

3. S is a spanning set of V, or say S spans V, if

$$\operatorname{span}(S)=V.$$

■ **Example 1.12** For $V = \mathbb{R}[x]$, define the set

$$S = \{1, x^2, x^4, \ldots, x^6\},$$

then $2 + x^4 + \pi x^{106} \in \text{span}(S)$, while the series $1 + x^2 + x^4 + \cdots \notin \text{span}(S)$.

It is clear that $\text{span}(S) \neq V$, but S is the spanning set of $W = \{p \in V \mid p(x) = p(-x)\}$.

■ **Example 1.13** For $V = M_{3\times 3}(\mathbb{R})$, let $W_1 = \{\boldsymbol{A} \in V \mid \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}\}$ and $W_2 = \{\boldsymbol{B} \in V \mid \boldsymbol{B}^{\mathrm{T}} = -\boldsymbol{B}\}$ (the set of skew-symmetric matrices) be two vector subspaces. Define the set

$$\boldsymbol{S} := W_1 \bigcup W_2$$

Exercise: $\boldsymbol{S}$ spans V.

Proposition 1.7 Let S be a subset in a vector space V.

1. $S \subseteq \text{span}(S)$
2. $\text{span}(S) = \text{span}(\text{span}(S))$
3. If $\boldsymbol{w} \in \text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} \setminus \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$, then

$$\boldsymbol{v}_1 \in \text{span}\{\boldsymbol{w}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \setminus \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$$

Proof. 1. For each $\boldsymbol{s} \in S$, we have

$$\boldsymbol{s} = 1 \cdot \boldsymbol{s} \in \text{span}(S)$$

2. From (1), it's clear that $\text{span}(S) \subseteq \text{span}(\text{span}(S))$, and therefore suffices to show $\text{span}(\text{span}(S)) \subseteq \text{span}(S)$:

Pick $\boldsymbol{v} = \sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i \in \text{span}(\text{span}(S))$, where $\boldsymbol{v}_i \in \text{span}(S)$. Rewrite

$$\boldsymbol{v}_i = \sum_{j=1}^{n_i} \beta_{ij} \boldsymbol{s}_j, \quad \boldsymbol{s}_j \in S,$$

which implies

$$\begin{aligned}\boldsymbol{v} &= \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n_i} \beta_{ij} \boldsymbol{s}_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n_i} (\alpha_i \beta_{ij}) \boldsymbol{s}_j,\end{aligned}$$

i.e., $\boldsymbol{v}$ is the finite combination of elements in S, whcih implies $\boldsymbol{v} \in \text{span}(S)$.

3. By hypothesis, $\boldsymbol{w} = \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n$ with $\alpha_1 \neq 0$, which implies

$$\boldsymbol{v}_1 = -\frac{\alpha_2}{\alpha_1} \boldsymbol{v}_2 + \cdots + \left(-\frac{1}{\alpha_1} \boldsymbol{w}\right)$$

which implies $\boldsymbol{v}_1 \in \text{span}\{\boldsymbol{w}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. It suffices to show $\boldsymbol{v}_1 \notin \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. Suppose on the contrary that $\boldsymbol{v}_1 \in \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. It's clear that $\text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} = \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. (left as exercise). Therefore,

$$\emptyset = \text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} \setminus \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\},$$

which is a contradiction.

■

1.4.3. Linear Independence and Basis

Definition 1.12 [Linear Independence] Let S be a (not necessarily finite) subset of V. Then S is **linearly independent** (l.i.) on V if for any finite subset $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_k\}$ in S,

$$\sum_{i=1}^{k} \alpha_i \boldsymbol{s}_i = 0 \Longleftrightarrow \alpha_i = 0, \forall i$$

■ **Example 1.14** For $V = C(\mathbb{R})$,

1. let $S_1 = \{\sin x, \cos x\}$, which is l.i., since

$$\alpha \sin x + \beta \cos x = \mathbf{0}(\text{means zero function})$$

   Taking $x = 0$ both sides leads to $\beta = 0$; taking $x = \frac{\pi}{2}$ both sides leads to $\alpha = 0$.

2. let $S_2 = \{\sin^2 x, \cos^2 x, 1\}$, which is linearly dependent, since

$$1 \cdot \sin^2 x + 1 \cdot \cos^2 x + (-1) \cdot 1 = 0, \forall x$$

3. Exercise: For $V = \mathbb{R}[x]$, let $S = \{1, x, x^2, x^3, \dots, \}$, which is l.i.:

   Pick $x^{k_1}, \dots, x^{k_n} \in S$ with $k_1 < \cdots < k_n$. Consider that the euqation

$$\alpha_1 x^{k_1} + \cdots + \alpha_n x^{k_n} = \mathbf{0}$$

   holds for all x, and try to solve for $\alpha_1, \dots, \alpha_n$ (one way is differentation.)

■

Definition 1.13 [Basis] A subset S is a **basis** of V if

(a) S spans V;

(b) S is l.i.

■

■ **Example 1.15** 1. For $V = \mathbb{R}^n$, $S = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is a basis of V

2. For $V = \mathbb{R}[x]$, $S = \{1, x, x^2, \dots\}$ is a basis of V

3. For $V = M_{2\times 2}(\mathbb{R})$,

$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

is a basis of V

R Note that there can be many basis for a vector space V.

Proposition 1.8 Let $V = \text{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$, then there exists a subset of $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$, which is a basis of V.

Proof. If $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ is l.i., the proof is complete.

Suppose not, then $\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_m \boldsymbol{v}_m = \mathbf{0}$ has a non-trivial solution. w.l.o.g., $\alpha_1 \neq 0$, which implies

$$\boldsymbol{v}_1 = -\frac{\alpha_2}{\alpha_1} \boldsymbol{v}_2 + \cdots + \left(\frac{\alpha_m}{\alpha_1} \right) \boldsymbol{v}_m \implies \boldsymbol{v}_1 \in \text{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_m\}$$

By the proof in (c), Proposition (1.7),

$$\text{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\} = \text{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_m\},$$

which implies $V = \text{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_m\}$.

Continuse this argument finitely many times to guarantee that $\{\boldsymbol{v}_i, \boldsymbol{v}_{i+1}, \dots, \boldsymbol{v}_m\}$ is l.i., and spans V. The proof is complete. ∎

Corollary 1.1 If $V = \text{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ (i.e., V is finitely generated), then V has a basis. (The same holds for non-finitely generated V).

Proposition 1.9 If $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ is a basis of V, then every $\boldsymbol{v} \in V$ can be expressed uniquely as

$$\boldsymbol{v} = \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n$$

Proof. Since $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ spans V, so $\boldsymbol{v} \in V$ can be written as

$$\boldsymbol{v} = \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n \tag{1.1}$$

Suppose further that

$$\boldsymbol{v} = \beta_1 \boldsymbol{v}_1 + \cdots + \beta_n \boldsymbol{v}_n, \tag{1.2}$$

it suffices to show that $\alpha_i = \beta_i$ for $\forall i$:

Subtracting (1.1) into (1.2) leads to

$$(\alpha_1 - \beta_1)\boldsymbol{v}_1 + \cdots + (\alpha_n - \beta_n)\boldsymbol{v}_n = 0.$$

By the hypothesis of linear independence, we have $\alpha_i - \beta_i = 0$ for $\forall i$, i.e., $\alpha_i = \beta_i$. ■

1.5. Wednesday for MAT3006

Reviewing.

- Normed Space: a norm on a vector space
- Metric Space
- Open Ball

1.5.1. Convergence of Sequences

Since $\mathbb{R}^n$ and $C[a,b]$ are both metric spaces, we can study the convergence in $\mathbb{R}^n$ and the functions defined on $[a,b]$ at the same time.

Definition 1.14 [Convergence] Let (X,d) be a metric space. A sequence $\{x_n\}$ in X is **convergent** to x if $\forall \varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$d(x_n, x) < \varepsilon, \forall n \geq N.$$

We can denote the convergence by

$$x_n \to x, \quad \text{or} \quad \lim_{n\to\infty} x_n = x, \quad \text{or} \quad \lim_{n\to\infty} d(x_n, x) = 0$$

Proposition 1.10 If the limit of $\{x_n\}$ exists, then it is unique.

R Note that the proposition above does not necessarily hold for topology spaces.

Proof. Suppose $x_n \to x$ and $x_n \to y$, which implies

$$0 \leq d(x,y) \leq d(x,x_n) + d(x_n,y), \forall n$$

Taking the limit $n \to \infty$ both sides, we imply $d(x,y) = 0$, i.e., $x = y$. ■

■ **Example 1.16** 1. Consider the metric space $(\mathbb{R}^k, d_\infty)$ and study the convergence

$$\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{x} \Longleftrightarrow \lim_{n\to\infty}\left(\max_{i=1\ldots,k}|x_{n_i} - x_i|\right) = 0$$

$$\Longleftrightarrow \lim_{n\to\infty}|x_{n_i} - x_i| = 0, \forall i = 1,\ldots,k$$

$$\Longleftrightarrow \lim_{n\to\infty} x_{n_i} = x_i,$$

i.e., the convergence defined in d_∞ is the same as the convergence defined in d_2.

2. Consider the convergence in the metric space $(C[a,b], d_\infty)$:

$$\lim_{n\to\infty} f_n = f \Longleftrightarrow \lim_{n\to\infty}\left(\max_{[a,b]}|f_n(x) - f(x)|\right) = 0$$

$$\Longleftrightarrow \forall\varepsilon > 0, \forall x \in [a,b], \exists N_\varepsilon \text{ such that } |f_n(x) - f(x)| < \varepsilon, \forall n \geq N_\varepsilon$$

which is equivalent to the uniform convergence of functions, i.e., the convergence defined in d_2.

■

Definition 1.15 [Equivalent metrics] Let d and ρ be metrics on X.

1. We say ρ is **stronger** than d (or d is **weaker** than ρ) if

$$\exists K > 0 \text{ such that } d(x,y) \leq K\rho(x,y), \forall x,y \in X$$

2. The metrics d and ρ are equivalent if there exists $K_1, K_2 > 0$ such that

$$d(x,y) \leq K_1\rho(x,y) \leq K_2 d(x,y)$$

■

R The strongerness of ρ than d is depicted in the graph below:

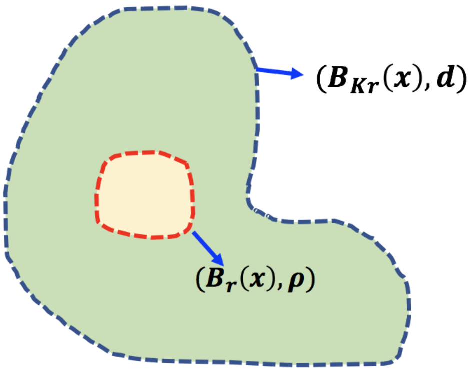


Figure 1.4: The open ball $(B_r(x), \rho)$ is contained by the open ball $(B_{Kr}(x), d)$

For each $x \in X$, consider the open ball $(B_r(x), \rho)$ and the open ball $(B_{Kr}(x), d)$:

$$B_r(x) = \{y \mid \rho(x, y) < r\}, \quad B_{Kr}(x) = \{z \mid d(x, z) < Kr\}.$$

For $y \in (B_r(x), \rho)$, we have $d(x, y) < K\rho(x, y) < Kr$, which implies $y \in (B_{Kr}(x), d)$, i.e, $(B_r(x), \rho) \subseteq (B_{Kr}(x), d)$ for any $x \in X$ and $r > 0$.

■ **Example 1.17** 1. d_1, d_2, d_∞ in $\mathbb{R}^n$ are equivalent

$$d_1(\boldsymbol{x}, \boldsymbol{y}) \leq d_\infty(\boldsymbol{x}, \boldsymbol{y}) \leq n d_1(\boldsymbol{x}, \boldsymbol{y})$$

$$d_2(\boldsymbol{x}, \boldsymbol{y}) \leq d_\infty(\boldsymbol{x}, \boldsymbol{y}) \leq \sqrt{n} d_2(\boldsymbol{x}, \boldsymbol{y})$$

We use two relation depiected in the figure below to explain these two inequalities:

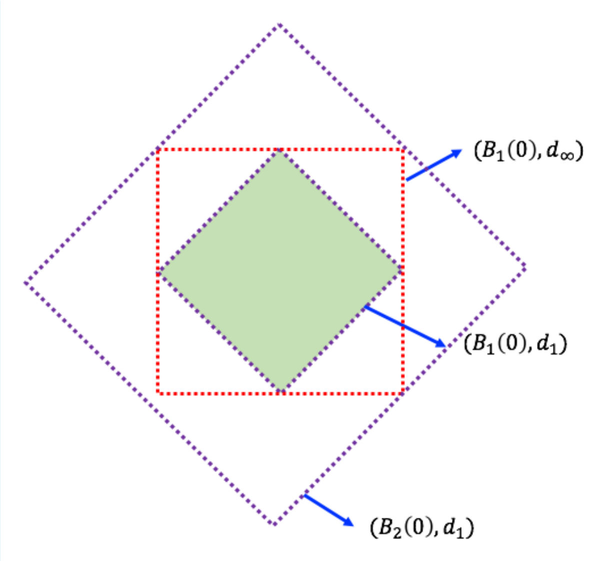


Figure 1.5: The diagram for the relation $(B_1(x), d_1) \subseteq (B_\infty(x), d_\infty) \subseteq (B_2(x), d_1)$ on $\mathbb{R}^2$

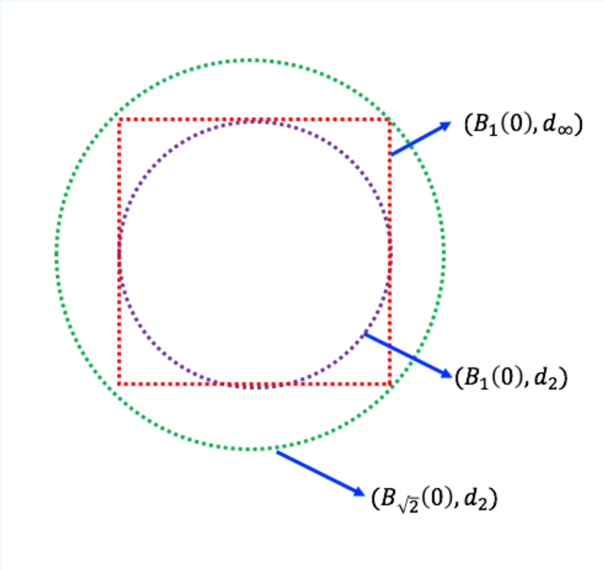


Figure 1.6: The diagram for the relation $(B_1(x), d_2) \subseteq (B_\infty(x), d_\infty) \subseteq (B_{\sqrt{2}}(x), d_2)$ on $\mathbb{R}^2$

It's easy to conclude the simple generalization for example (1.16):

Proposition 1.11 If d and ρ are equivalent, then

$$\lim_{n\to\infty} d(x_n, x) = 0 \Longleftrightarrow \lim_{n\to\infty} \rho(x_n, x) = 0$$

Note that this does not necessarily hold for topology spaces.

2. Consider d_1, d_∞ in $C[a,b]$:

$$d_1(f,g) := \int_a^b |f-g|\,\mathrm{d}x \le \int_a^b \sup_{[a,b]}|f-g|\,\mathrm{d}x = (b-a)d_\infty(f,g),$$

i.e., d_∞ is stronger than d_1. Question: Are they equivalent? **No**.

Justification. Consider $f_n(x) = n^2x^n(1-x)$ for $x \in [0,1]$. Check that

$$\lim_{n\to\infty} d_1(f_n(x),0) = 0, \quad \text{but } d_\infty(f_n(x),0) \to \infty$$

The peak of f_n may go to infinite, while the integration converges to zero, i.e., there is no $K > 0$ such that $d_\infty(f_n,0) < Kd_1(f_n,0), \forall n \in \mathbb{N}$. ■

We will discuss this topic at Lebsegue integration again.

1.5.2. Continuity

Definition 1.16 [Continuity] Let $f : (X,d) \to (Y,d)$ be a function and $x_0 \in X$. Then f is continuous at x_0 if $\forall \varepsilon > 0$, there exists $\delta > 0$ such that

$$d(x,x_0) < \delta \implies \rho(f(x), f(x_0)) < \varepsilon$$

The function f is continuous in X if f is continuous for all $x_0 \in X$.

Proposition 1.12 The function f is continuous at x if and only if for all $\{x_n\} \to x$ under d, $f(x_n) \to f(x)$ under ρ.

Proof. *Necessity:* Given $\varepsilon > 0$, by continuity,

$$d(x,x') < \delta \implies \rho(f(x'), f(x)) < \varepsilon. \tag{1.3}$$

Consider the sequence $\{x_n\} \to x$, then there exists N such that $d(x_n,x) < \delta$ for $\forall n \ge N$. By applying (1.3), $\rho(f(x_n), f(x)) < \varepsilon$ for $\forall n \ge N$, i.e., $f(x_n) \to f(x)$.

Sufficiency: Assume that f is not continuous at x, then there exists ε_0 such that for

$\delta_n = \frac{1}{n}$, there exists x_n such that

$$d(x_n, x) < \delta_n, \text{ but } \rho(f(x_n), f(x)) > \varepsilon_0.$$

Then $\{x_n\} \to x$ by our construction, while $\{f(x_n)\}$ does not converge to $f(x)$, which is a contradiction. ∎

Corollary 1.2 If the function $f : (X, d) \to (Y, \rho)$ is continuous at x, the function $g : (Y, \rho) \to (Z, m)$ is continuous at $f(x)$, then $g \circ f : (X, d) \to (Z, m)$ is continuous at x.

Proof. Note that

$$\{x_n\} \to x \overset{(a)}{\Longrightarrow} \{f(x_n)\} \to f(x) \overset{(b)}{\Longrightarrow} \{g(f(x_n))\} \to g(f(x)) \overset{(c)}{\Longrightarrow} g \circ f \text{ is continuous at } x.$$

where $(a), (b), (c)$ are all by proposition (1.12). ∎

1.5.3. Open and Closed Sets

We have open/closed intervals in $\mathbb{R}$, and they are important in some theorems (e.g, continuous functions bring closed intervals to closed intervals).

Definition 1.17 [Open] Let (X, d) be a metric space. A set $U \subseteq X$ is open if for each $x \in U$, there exists $\rho_x > 0$ such that $B_{\rho_x}(x) \subseteq U$. The empty set $\emptyset$ is defined to be open. ▪

▪ **Example 1.18** Let $(\mathbb{R}, d_2 \text{ or } d_\infty)$ be a metric space. The set $U = (a, b)$ is open. ▪

Proposition 1.13
1. Let (X, d) be a metric space. Then all open balls $B_r(x)$ are open
2. All open sets in X can be written as a union of open balls.

Proof.
1. Let $y \in B_r(x)$, i.e., $d(x, y) := q < r$. Consider the open ball $B_{(r-q)/2}(y)$. It suffices to show $B_{(r-q)/2}(y) \subseteq B_r(x)$. For any $z \in B_{(r-q)/2}(y)$,

$$d(x, z) \le d(x, y) + d(y, z) < q + \frac{r-q}{2} = \frac{r+q}{2} < r.$$

The proof is complete.

2. Let $U \subseteq X$ be open, i.e., for $\forall x \in U$, there exists $\varepsilon_x > 0$ such that $B_{\varepsilon_x}(x) \subseteq U$. Therefore

$$\{x\} \subseteq B_{\varepsilon_x}(x) \subseteq U, \forall x \in U$$

which implies

$$U = \bigcup_{x \in U} \{x\} \subseteq \bigcup_{x \in U} B_{\varepsilon_x}(x) \subseteq U,$$

i.e., $U = \bigcup_{x \in U} B_{\varepsilon_x}(x)$.

■

1.6. Wednesday for MAT4002

Reviewing.

- Metric Space (X,d)
- Open balls and open sets (note that the emoty set $\emptyset$ is open)
- Define the collection of open sets in X, say $\mathcal{T}$ is the topology.

Exercise.

1. Show that the $\mathcal{T}_2$ under $(X=\mathbb{R}^2,d_2)$ and $\mathcal{T}_\infty$ under $(X=\mathbb{R}^2,d_\infty)$ are the same.

   Ideas. Follow the procedure below:

   An open ball in d_2-metric is open in d_∞;

   Any open set in d_2-metric is open in d_∞;

   Switch d_2 and d_∞. ■

2. Describe the topology $\mathcal{T}_{\text{discrete}}$ under the metric space $(X=\mathbb{R}^2,d_{\text{discrete}})$.

   Outlines. Note that $\{x\}=B_{1/2}(x)$ is an open set.

   For any subset $W\subseteq\mathbb{R}^2$, $W=\bigcup_{w\in W}\{w\}$ is open.

   Therefore $\mathcal{T}_{\text{discrete}}$ is all subsets of $\mathbb{R}^2$. ■

1.6.1. Forget about metric

Next, we will try to define closedness, compactness, etc., without using the tool of metric:

Definition 1.18 [closed] A subset $V\subseteq X$ is **closed** if $X\setminus V$ is **open**.

Example 1.19 Under the metric space $(\mathbb{R},d_1)$,

$$\mathbb{R}\setminus[b,a]=(a,\infty)\bigcup(-\infty,b)\text{ is open}\implies[b,a]\text{ is closed}$$

Proposition 1.14 Let X be a metric space.

1. $\emptyset, X$ is closed in X
2. If F_α is closed in X, so is $\bigcap_{\alpha \in A} F_\alpha$.
3. If $F_1, \ldots, F_k$ is closed, so is $\bigcup_{i=1}^k F_i$.

Proof. 1. Note that X is open in X, which implies $\emptyset = X \setminus X$ is closed in X; Similarly, $\emptyset$ is open in X, which implies $X = X \setminus \emptyset$ is closed in X;

2. The set F_α is closed implies there exists open $U_\alpha \subseteq X$ such that $F_\alpha = X \setminus U_\alpha$. By De Morgan's Law,

$$\bigcap_{\alpha \in A} F_\alpha = \bigcap_{\alpha \in A} (X \setminus U_\alpha) = X \setminus (\bigcup_{\alpha \in A} U_\alpha).$$

By part (a) in proposition (1.6), the set $\bigcup_{\alpha \in A} U_\alpha$ is openm which implies $\bigcap_{\alpha \in A} F_\alpha$ is closed.

3. The result follows from part (b) in proposition (1.6) by taking complements.

■

We illustrate examples where open set is used to define convergence and continuity.

1. Convergence of sequences:

Definition 1.19 [Convergence] Let (X, d) be a metric space, then $\{x_n\} \to x$ means

$$\forall \varepsilon > 0, \exists N \text{ such that } d(x_n, x) < \varepsilon, \forall n \geq N.$$

We will study the convergence by using open sets instead of metric.

Proposition 1.15 Let X be a metric space, then $\{x_n\} \to x$ if and only if for $\forall$ open set $U \ni x$, there exists N such that $x_n \in U$ for $\forall n \geq N$.

Proof. *Necessity*: Since $U \ni x$ is open, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq U$. Since $\{x_n\} \to x$, there exists N such that $d(x_n, x) < \varepsilon$, i.e., $x_n \in B_\varepsilon(x) \subseteq U$ for $\forall n \geq N$.

Sufficiency: Let $\varepsilon > 0$ be given. Take the open set $U = B_\varepsilon(x) \ni x$, then there exists N such that $x_n \in U = B_\varepsilon(x)$ for $\forall n \geq N$, i.e., $d(x_n, x) < \varepsilon$, $\forall n \geq N$.

■

2. Continuity:

Definition 1.20 [Continuity] Let (X, d) and (Y, ρ) be given metric spaces. Then $f : X \to Y$ is continuous at $x_0 \in X$ if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } d(x, x_0) < \delta \implies \rho(f(x), f(x_0)) < \varepsilon.$$

The function f is continuous on X if f is continous for all $x_0 \in X$.

We can get rid of metrics to study continuity:

Proposition 1.16 (a) The function f is continuous at x if and only if for all open $U \ni f(x)$, there exists $\delta > 0$ such that the set $B(x, \delta) \subseteq f^{-1}(U)$.

(b) The function f is continuous on X if and only if $f^{-1}(U)$ is open in X for each open set $U \subseteq Y$.

During the proof we will apply a small lemma:

Proposition 1.17 f is continuous at x if and only if for all $\{x_n\} \to x$, we have $\{f(x_n)\} \to f(x)$.

Proof. (a) *Necessity*:

Due to the openness of $U \ni f(x)$, there exists a ball $B(f(x), \varepsilon) \subseteq U$.

Due to the continuity of f at x, there exists $\delta > 0$ such that $d(x, x') < \delta$ implies $d(f(x), f(x')) < \varepsilon$, which implies

$$f(B(x, \delta)) \subseteq B(f(x), \varepsilon) \subseteq U,$$

which implies $B(x, \delta) \subseteq f^{-1}(U)$.

Sufficiency:

Let $\{x_n\} \to x$. It suffices to show $\{f(x_n)\} \to f(x)$. For each open $U \ni f(x)$, by hypothesis, there exists $\delta > 0$ such that $B_\delta(x) \subseteq f^{-1}(U)$.

Since $\{x_n\} \to x$, there exists N such that

$$x_n \in B_\delta(x) \subseteq f^{-1}(U), \forall n \geq N \implies f(x_n) \in U, \forall n \geq N$$

Let $\varepsilon > 0$ be given, and then construct the $U = B_\varepsilon(f(x))$. The argument above shows that $f(x_n) \in B_\varepsilon(f(x))$ for $\forall n \geq N$, which implies $\rho(f(x_n), f(x)) < \varepsilon$, i.e., $\{f(x_n)\} \to f(x)$.

(b) For the forward direction, it suffices to show that each point x of $f^{-1}(U)$ is an interior point of $f^{-1}(U)$, which is shown by part (a); the converse follows trivially by applying (a).

■

R As illustracted above, convergence, continuity, (and compactness) can be defined by using open sets $\mathcal{T}$ only.

1.6.2. Topological Spaces

Definition 1.21 A **topological space** $(X, \mathcal{T})$ consists of a (non-empty) set X, and a family of subsets of X ("open sets" $\mathcal{T}$) such that

1. $\emptyset, X \in \mathcal{T}$
2. $U, V \in \mathcal{T}$ implies $U \bigcap V \in \mathcal{T}$
3. If $U_\alpha \in \mathcal{T}$ for all $\alpha \in \mathcal{A}$, then $\bigcup_{\alpha \in \mathcal{A}} U_\alpha \in \mathcal{T}$.

The elements in $\mathcal{T}$ are called **open subsets** of X. The $\mathcal{T}$ is called a **topology** on X.

Example 1.20 1. Let (X, d) be any metric space, and

$$\mathcal{T} = \{\text{all open subsets of } X\}$$

It's clear that $\mathcal{T}$ is a topology on X.

2. Define the discrete topology

$$\mathcal{T}_{\text{dis}} = \{\text{all subsets of } X\}$$

It's clear that $\mathcal{T}_{\text{dis}}$ is a topology on X, (which also comes from the discrete metric (X, d_{discrete})).

R We say $(X, \mathcal{T})$ is induced from a metric (X, d) (or it is **metrizable**) if $\mathcal{T}$ is the faimly of open subsets in (X, d).

3. Consider the indiscrete topology $(X, \mathcal{T}_{\text{indis}})$, where X contains more than one element:

$$\mathcal{T}_{\text{indis}} = \{\emptyset, X\}.$$

Question: is $(X, \mathcal{T}_{\text{indis}})$ metrizable? No. For any metric d defined on X, let x, y be distinct points in X, and then $\varepsilon := d(x, y) > 0$, hence $B_{\frac{1}{2}\varepsilon}(x)$ is a open set belonging to the corresponding induced topology. Since $x \in B_{\frac{1}{2}\varepsilon}(x)$ and $y \notin B_{\frac{1}{2}\varepsilon}(y)$, we conclude that $B_{\frac{1}{2}\varepsilon}(x)$ is neither $\emptyset$ nor X, i.e., the topology induced by any metric d is not the indiscrete topology.

4. Consider the cofinite topology $(X, \mathcal{T}_{\text{cofin}})$:

$$\mathcal{T}_{\text{cofin}} = \{U \mid X \setminus U \text{ is a finite set}\} \bigcup \{\emptyset\}$$

Question: is $(X, \mathcal{T}_{\text{cofin}})$ metrizable?

Definition 1.22 [Equivalence] Two metric spaces are **topologically equivalent** if they give rise to the same topology.

▪ **Example 1.21** Metrics d_1, d_2, d_∞ in $\mathbb{R}^n$ are topologically equivalent.

1.6.3. Closed Subsets

Definition 1.23 [Closed] Let $(X,\mathcal{T})$ be a topology space. Then $V \subseteq X$ is **closed** if $X \setminus V \in \mathcal{T}$

■ **Example 1.22** Under the topology space $(\mathbb{R},\mathcal{T}_{\text{usual}})$, $(b,\infty)\bigcup(-\infty,a) \in \mathcal{T}$. Therefore,

$$[a,b] = \mathbb{R} \setminus \left((b,\infty)\bigcup(-\infty,a)\right)$$

is closed in $\mathbb{R}$ under usual topology.

R It is important to say that V is **closed in** X**.** You need to specify the underlying the space X.

Chapter 2

Week2

2.1. Monday for MAT3040

Reviewing.

1. Linear Combination and Span
2. Linear Independence
3. Basis: a set of vectors $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$ is called a **basis** for V if $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$ is linearly independent, and $V = \text{span}\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$.

   Lemma: Given $V = \text{span}\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$, we can find a basis for this set. Here V is said to be **finitely generated**.
4. Lemma: The vector $\boldsymbol{w} \in \text{span}\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\} \setminus \text{span}\{\boldsymbol{v}_2,\dots,\boldsymbol{v}_n\}$ implies that

$$\boldsymbol{v}_1 \in \text{span}\{\boldsymbol{w},\boldsymbol{v}_2,\dots,\boldsymbol{v}_n\} \setminus \text{span}\{\boldsymbol{v}_2,\dots,\boldsymbol{v}_n\}$$

2.1.1. Basis and Dimension

Theorem 2.1 Let V be a finitely generated vector space. Suppose $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_m\}$ and $\{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n\}$ are two basis of V. Then $m = n$. (where m is called the **dimension**)

Proof. Suppose on the contrary that $m \neq n$. Without loss of generality (w.l.o.g.), assume that $m < n$. Let $\boldsymbol{v}_1 = \alpha_1\boldsymbol{w}_1 + \cdots + \alpha_n\boldsymbol{w}_n$, with some $\alpha_i \neq 0$. w.l.o.g., assume $\alpha_1 \neq 0$. Therefore,

$$\boldsymbol{v}_1 \in \text{span}\{\boldsymbol{w}_1,\boldsymbol{w}_2,\dots,\boldsymbol{w}_n\} \setminus \text{span}\{\boldsymbol{w}_2,\dots,\boldsymbol{w}_n\} \tag{2.1}$$

which implies that $\boldsymbol{w}_1 \in \text{span}\{\boldsymbol{v}_1,\boldsymbol{w}_2,\dots,\boldsymbol{w}_n\} \setminus \text{span}\{\boldsymbol{w}_2,\dots,\boldsymbol{w}_n\}$.

Then we claim that $\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$ is a basis of V:

1. Note that $\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$ is a spannning set:

$$\begin{aligned}\boldsymbol{w}_1 \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} &\implies \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} \subseteq \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} \\ &\implies \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} \subseteq \operatorname{span}\{\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}\} \subseteq \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}\end{aligned}$$

   Since $V = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$, we have $\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} = V$.

2. Then we show the linear independence of $\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$. Consider the equation

$$\beta_1 \boldsymbol{v}_1 + \beta_2 \boldsymbol{v}_2 + \cdots + \beta_n \boldsymbol{w}_n = \mathbf{0}$$

   (a) When $\beta_1 \neq 0$, we imply

$$\boldsymbol{v}_1 = \left(-\frac{\beta_2}{\beta_1}\right) \boldsymbol{w}_2 + \cdots + \left(-\frac{\beta_n}{\beta_1}\right) \boldsymbol{w}_n \in \operatorname{span}\{\boldsymbol{w}_2, \dots, \boldsymbol{w}_n\},$$

   which contradicts (2.1).

   (b) When $\beta_1 = 0$, then $\beta_2 \boldsymbol{w}_2 + \cdots + \beta_n \boldsymbol{w}_n = \mathbf{0}$, which implies $\beta_2 = \cdots = \beta_n = 0$, due to the independence of $\{\boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$.

Therefore, $\boldsymbol{v}_2 \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$, i.e.,

$$\boldsymbol{v}_2 = \gamma_1 \boldsymbol{v}_1 + \cdots + \gamma_n \boldsymbol{v}_n,$$

where $\gamma_2, \dots, \gamma_n$ cannot be all zeros, since otherwise $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ are linearly dependent, i.e., $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ cannot form a basis. w.l.o.g., assume $\gamma_2 \neq 0$, which implies

$$\boldsymbol{w}_2 \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}_3, \dots, \boldsymbol{w}_n\} \setminus \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_3, \dots, \boldsymbol{w}_n\}.$$

Following the simlar argument above, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}_3, \dots, \boldsymbol{w}_n\}$ forms a basis of V.

Continuing the argument above, we imply $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m, \boldsymbol{w}_{m+1}, \dots, \boldsymbol{w}_n\}$ is a basis of

V. Since $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ is a basis as well, we imply

$$\boldsymbol{w}_{m+1} = \delta_1 \boldsymbol{v}_1 + \cdots + \delta_m \boldsymbol{v}_m$$

for some $\delta_i \in \mathbb{F}$, i.e., $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m, \boldsymbol{w}_{m+1}\}$ is linearly dependent, which is a contradction.

■

▪ **Example 2.1** A vector space may have more than one basis.

Suppose $V = \mathbb{F}^n$, it is clear that $\dim(V) = n$, and

$$\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\} \text{ is a basis of } V, \text{ where } \boldsymbol{e}_i \text{ denotes a unit vector.}$$

There could be other basis of V, such as

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \right\}$$

Actually, the columns of any invertible $n \times n$ matrix forms a basis of V. ▪

▪ **Example 2.2** Suppose $V = M_{m \times n}(\mathbb{R})$, we claim that $\dim(V) = mn$:

$$\left\{ E_{ij} \,\middle|\, \begin{matrix} 1 \le i \le m \\ 1 \le j \le n \end{matrix} \right\} \text{ is a basis of } V,$$

where E_{ij} is $m \times n$ matrix with 1 at (i,j)-th entry, and 0s at the remaining entries. ▪

▪ **Example 2.3** Suppose $V = \{\text{all polynomials of degree} \le \text{n}\}$, then $\dim(V) = n + 1$. ▪

▪ **Example 2.4** Supppose $V = \{\boldsymbol{A} \in M_{n \times n}(\mathbb{R}) \mid \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}\}$, then $\dim(V) = \frac{n(n+1)}{2}$. ▪

■ **Example 2.5** Let $W = \{\boldsymbol{B} \in M_{n\times n}(\mathbb{R}) \mid \boldsymbol{B}^{\mathrm{T}} = -\boldsymbol{B}\}$, then $\dim(V) = \frac{n(n-1)}{2}$. ■

R Sometimes it should be classified the field $\mathbb{F}$ for the scalar multiplication to define a vector space. Conside the example below:

1. Let $V = \mathbb{C}$, then $\dim(\mathbb{C}) = 1$ for the scalar multiplication defined under the field $\mathbb{C}$.
2. Let $V = \text{span}\{1, i\} = \mathbb{C}$, then $\dim(\mathbb{C}) = 2$ for the scalar multiplication defined under the field $\mathbb{R}$, since all $z \in V$ can be written as $z = a + bi$, $\forall a, b \in \mathbb{R}$.
3. Therefore, to aviod confusion, it is safe to write

$$\dim_{\mathbb{C}}(\mathbb{C}) = 1, \quad \dim_{\mathbb{R}}(\mathbb{C}) = 2.$$

2.1.2. Operations on a vector space

Note that the basis for a vector space is characterized as the **maximal linearly independent set**.

Theorem 2.2 — **Basis Extension.** Let V be a finite dimensional vector space, and $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ be a linearly independent set on V, Then we can extend it to the basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n\}$ of V.

Proof.

- Suppose $\dim(V) = n > k$, and $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ is a basis of V. Consider the set $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\} \bigcup \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$, which is linearly dependent, i.e.,

$$\alpha_1 \boldsymbol{w}_1 + \dots + \alpha_n \boldsymbol{w}_n + \beta_1 \boldsymbol{v}_1 + \dots + \beta_k \boldsymbol{v}_k = \boldsymbol{0},$$

with some $\alpha_i \neq 0$, since otherwise this equation will only have trivial solution. w.l.o.g., assume $\alpha_1 \neq 0$.

- Therefore, consider the set $\{\boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} \bigcup \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$. We keep removing elements from $\{\boldsymbol{w}_2,\dots,\boldsymbol{w}_n\}$ until we first get the set

$$S\bigcup\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\},$$

with $S\subseteq\{\boldsymbol{w}_1,\boldsymbol{w}_2,\dots,\boldsymbol{w}_n\}$ and $S\bigcup\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$ is linearly independent, i.e., S is a maximal subset of $\{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n\}$ such that $S\bigcup\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$ is linearly independent.

- Rewrite $S=\{\boldsymbol{v}_{k+1},\dots,\boldsymbol{v}_m\}$ and therefore $S'=\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k,\boldsymbol{v}_{k+1},\dots,\boldsymbol{v}_m\}$ are linearly independent. It suffices to show S' spans V.
  - Indeed, for all $\boldsymbol{w}_i\in\{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n\}$, $\boldsymbol{w}_i\in\operatorname{span}(S')$, since otherwise the equation

    $$\alpha\boldsymbol{w}_i+\beta_1\boldsymbol{v}_1+\cdots+\beta_m\boldsymbol{v}_m=\mathbf{0}\implies\alpha=0,$$

    which implies that $\beta_1\boldsymbol{v}_1+\cdots+\beta_m\boldsymbol{v}_m=\mathbf{0}$ admits only trivial solution, i.e.,

    $$\{\boldsymbol{w}_i\}\bigcup S'=\{\boldsymbol{w}_i\}\bigcup S\bigcup\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}\text{ is linearly independent,}$$

    which violetes the maximality of S.

  Therefore, all $\{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n\}\subseteq\operatorname{span}(S')$, which implies $\operatorname{span}(S')=V$.

  Therefore, S' is a basis of V.

■

R Start with a spanning set, we keep removing something to form a basis; start with independent set, we keep adding something to form a basis.

In other words, the basis is both the minimal spanning set, and the maximal linearly independent set.

Definition 2.1 [Direct Sum] Let W_1,W_2 be two vector subspaces of V, then

1. $W_1\bigcap W_2:=\{\boldsymbol{w}\in V\mid\boldsymbol{w}\in W_1,\text{ and }\boldsymbol{w}\in W_2\}$
2. $W_1+W_2:=\{\boldsymbol{w}_1+\boldsymbol{w}_2\mid\boldsymbol{w}_i\in W_i\}$

3. If furthermore that $W_1 \bigcap W_2 = \{\mathbf{0}\}$, then $W_1 + W_2$ is denoted as $W_1 \oplus W_2$, which is called **direct sum**.

Proposition 2.1 $W_1 \bigcap W_2$ and $W_1 + W_2$ are vector subspaces of V.

2.2. Monday for MAT3006

Reviewing.

1. Continuous functions: the function f is continuous is equivalent to say for $\forall x_n \to x$, we have $f(x_n) \to f(x)$.
2. Open sets: Let (X, d) be a metric space. A set $U \subseteq X$ is open if for each $x \in U$, there exists $\rho_x > 0$ such that $B_{\rho_x}(x) \subseteq U$.

(R) Unless stated otherwise, we assume that

$$C[a,b] \longleftrightarrow (C[a,b], d_\infty)$$

$$\mathbb{R}^n \longleftrightarrow (\mathbb{R}^n, d_2)$$

2.2.1. Remark on Open and Closed Set

Example 2.6 Let $X = C[a,b]$, show that the set

$$U := \{f \in X \mid f(x) > 0, \forall x \in [a,b]\} \quad \text{is open.}$$

Take a point $f \in U$, then

$$\inf_{[a,b]} f(x) = m > 0.$$

Consider the ball $B_{m/2}(f)$, and for $\forall g \in B_{m/2}(f)$,

$$\begin{aligned}
|g(x)| &\ge |f(x)| - |f(x) - g(x)| \\
&\ge \inf_{[a,b]} |f(x)| - \sup_{[a,b]} |f(x) - g(x)| \\
&\ge m - \frac{m}{2} \\
&= \frac{m}{2} > 0, \ \forall x \in [a,b]
\end{aligned}$$

Therefore, we imply $g \in U$, i.e., $B_{m/2}(f) \subseteq U$, i.e., U is open in X. ■

Proposition 2.2 Let (X,d) be a metric space. Then

1. $\emptyset, X$ are open in X
2. If $\{U_\alpha \mid \alpha \in \mathcal{A}\}$ are open in X, then $\bigcup_{\alpha\in\mathcal{A}}$ is also open in X
3. If $U_1,\dots,U_n$ are open in X, then $\bigcap_{i=1}^{n} U_i$ are open in X

R Note that $\bigcap_{i=1}^{\infty} U_i$ is not necessarily open if all U_i's are all open:

$$\bigcap_{i=1}^{\infty}\left(-\frac{1}{i},1+\frac{1}{i}\right)=[0,1]$$

Definition 2.2 [Closed] The closed set in metric space (X,d) are the complement of open sets in X, i.e., any closed set in X is of the form $V = X \setminus U$, where U is oepn. ■

For example, in $\mathbb{R}$,

$$[a,b]=\mathbb{R}\setminus\{(-\infty,a)\bigcup(b,\infty)\}$$

Proposition 2.3

1. $\emptyset, X$ are closed in X
2. If $\{V_\alpha \mid \alpha \in \mathcal{A}\}$ are closed subsets in X, then $\bigcap_{\alpha\in\mathcal{A}} V_\alpha$ is also closed in X
3. If $V_1,\dots,V_n$ are closed in X, then $\bigcup_{i=1}^{n} V_i$ is also closed in X.

R Whenever you say U is open or V is closed, you need to specify the underlying space, e.g.,

Wrong :U is open

Right :U is open in X

Proposition 2.4 The following two statements are equivalent:

1. The set V is closed in metric space (X,d).
2. If the sequence $\{v_n\}$ in V converges to x, then $x \in V$

Proof. Necessity.

Suppose on the contrary that $\{v_n\} \to x \notin V$. Since $X \setminus V \ni x$ is open, there exists an open ball $B_\varepsilon(x) \subseteq X \setminus V$.

Due to the convergence of sequence, there exists N such that $d(v_n, x) < \varepsilon$ for $\forall n \geq N$, i.e., $v_n \in B_\varepsilon(x)$, i.e., $v_n \notin V$, which contradicts to $\{v_n\} \subseteq V$.

Sufficiency.

Suppose on the contrary that V is not closed in X, i.e., $X \setminus V$ is not open, i.e., there exists $x \notin V$ such that for all open $U \ni x$, $U \bigcap V \neq \emptyset$. In particular, take

$$U_n = B_{1/n}(x), \implies \exists v_n \in B_{1/n}(x) \bigcap V,$$

i.e., $\{v_n\} \to x$ but $x \notin V$, which is a contradiction. ■

Proposition 2.5 Given two metric space (X,d) and (Y,ρ), the following statements are equivalent:

1. A function $f : (X,d) \to (Y,\rho)$ is continuous on X
2. For $\forall U \subseteq Y$ open in Y, $f^{-1}(U)$ is open in X.
3. For $\forall V \subseteq Y$ closed in Y, $f^{-1}(V)$ is closed in X.

■ **Example 2.7** The mapping $\Psi : C[a,b] \to \mathbb{R}$ is defined as:

$$f \mapsto f(c)$$

where Ψ is called a **functional**.

Show that Ψ is continuous by using d_∞ metric on $C[a,b]$:

1. Any open set in $\mathbb{R}$ can be written as countably union of open disjoint intervals, and therefore suffices to consider the pre-image $\Psi^{-1}(a,b) = \{f \mid f(c) \in (a,b)\}$. Following the similar idea in Example (2.6), it is clear that $\Psi^{-1}(a,b)$ is open in $(C[a,b], d_\infty)$. Therefore, Ψ is continuous.
2. Another way is to apply definition.

■

We now study open sets in a subspace $(Y, d_Y) \subseteq (X, d_X)$, i.e.,

$$d_Y(y_1, y_2) := d_X(y_1, y_2), \quad \forall y_1, y_2 \in Y.$$

Therefore, the open ball is defined as

$$\begin{aligned}
B_\varepsilon^Y(y) &= \{y' \in Y \mid d_Y(y, y') < \varepsilon\} \\
&= \{y' \in Y \mid d_X(y, y') < \varepsilon\} \\
&= \{y' \in X \mid d_X(y, y') < \varepsilon, y' \in Y\} \\
&= B_\varepsilon^X(y) \bigcap Y
\end{aligned}$$

Proposition 2.6 All open sets in the subspace $(Y, d_Y) \subseteq (X, d_X)$ are of the form $U \bigcap Y$, where U is open in X.

Corollary 2.1 For the subspace $(Y, d_Y) \subseteq (X, d_X)$, the mapping $i : (Y, d_Y) \to (X, d_X)$ with $i(y) = y, \forall y \in Y$ is continuous.

Proof. $i^{-1}(U) = U \bigcap Y$ for any subset $U \subseteq X$. The results follows from proposition (2.5). ■

(R) It's important to specify the underlying space to describe an open set.

For example, the interval $[0, \frac{1}{2})$ is not open in $\mathbb{R}$, while $[0, \frac{1}{2})$ is open in $[0, 1]$, since

$$[0, \frac{1}{2}) = (-\frac{1}{2}, \frac{1}{2}) \bigcap [0, 1].$$

2.2.2. Boundary, Closure, and Interior

Definition 2.3 Let (X, d) be a metric space, then

1. A point x is a **boundary point** of $S \subseteq X$ (denoted as $x \in \partial S$) if for any open $U \ni x$, then both $U \bigcap S, U \setminus S$ are non-empty.

(one can replace U by $B_{1/n}(x)$, with $n = 1,2,\dots$)

2. The **closure** of S is defined as $\overline{S} = S \bigcup \partial S$.

3. A point x is an **interior point** of S (denoted as $x \in S^\circ$) if there $\exists U \ni x$ open such that $U \subseteq S$. We use S° to denote the set of interior points.

Proposition 2.7 1. The closure of S can be equivalently defined as

$$\overline{S} = \bigcap \{C \in X \mid C \text{ is closed and } C \supseteq S\}$$

Therefore, $\overline{S}$ is the smallest closed set containing S.

(Note that $C = X$ is a closed set containing S, and hence $\overline{S}$ is well-defined.)

2. The interior set of S can be equivalently defined as

$$S^\circ = \bigcup \{U \subseteq X \mid U \text{ is open and } U \subseteq S\}$$

Therefore, S° is the largest open set contained in S.

Example 2.8 For $S = [0, \frac{1}{2}] \subseteq X$, we have

1. $\partial S = \{0, \frac{1}{2}\}$
2. $\overline{S} = [0, \frac{1}{2}]$
3. $S^\circ = (0, \frac{1}{2})$

Proof. 1. (a) Firstly, we show that $\overline{S}$ is closed, i.e., $X \setminus \overline{S}$ is open.

- Take $x \notin \overline{S}$. Since $x \notin \partial S$, there $\exists B_r(x) \ni x$ such that

$$B_r(x) \bigcap S, \quad \text{or} \quad B_r(x) \setminus S \text{ is } \emptyset.$$

- Since $x \notin S$, the set $B_r(x) \setminus S$ is not empty. Therefore, $B_r(x) \bigcap S = \emptyset$.
- It's clear that $B_{r/2}(x) \bigcap S = \emptyset$. We claim that $B_{r/2}(x) \bigcap \overline{S}$ is empty.

Suppose on the contrary that

$$y \in B_{r/2}(x) \bigcap \partial S,$$

which implies that $B_{r/2}(y) \bigcap S \neq \emptyset$. Therefore,

$$B_{r/2}(y) \subseteq B_r(x) \implies B_r(x) \bigcap S \supseteq B_{r/2}(y) \bigcap S \neq \emptyset,$$

which is a contradiction.

Therefore, $x \in X \setminus \overline{S}$ implies $B_{r/2}(x) \bigcap \overline{S} = \emptyset$, i.e., $X \setminus \overline{S}$ is open, i.e., $\overline{S}$ is closed.

(b) Secondly, we show that $\overline{S} \subseteq C$, for any closed $C \supseteq S$, i.e., suffices to show $\partial S \subseteq C$.

Take $x \in \partial S$, which implies that $B_\varepsilon(x) \cap S$ is non-empty for any $\varepsilon > 0$. Therefore, construct a sequence

$$x_n \in B_{1/n}(x) \bigcap S.$$

Here $\{x_n\}$ is a sequence in $S \subseteq C$ converging to x, which implies $x \in C$, due to the closeness of C in X.

Combining (a) and (b), the result follows naturally.

2. Exercise. Show that

$$S^\circ = S \setminus \partial S = X \setminus (\overline{X \setminus S}).$$

Then it's clear that S° is open, and contained in S.

■

The next lecture we will talk about compactness and sequential compactness.

2.3. Monday for MAT4002

Reviewing.

1. Topological Space $(X,\mathcal{J})$: a special class of topological space is that induced from metric space (X,d):

$$(X,\mathcal{T}), \quad \text{with } \mathcal{T} = \{\text{all open sets in } (X,d)\}$$

2. Closed Sets $(X\setminus U)$ with U open.

Proposition 2.8 Let $(X,\mathcal{T})$ be a topological space,

1. $\emptyset, X$ are closed in X
2. V_1,V_2 closed in X implies that $V_1\bigcup V_2$ closed in X
3. $\{V_\alpha \mid \alpha\in\mathcal{A}\}$ closed in X implies that $\bigcap_{\alpha\in\mathcal{A}} V_\alpha$ closed in X

Proof. Applying the De Morgan's Law

$$(X\setminus\bigcup_{i\in I} U_i) = \bigcap_{i\in I}(X\setminus U_i)$$

∎

2.3.1. Convergence in topological space

Definition 2.4 [Convergence] A sequence $\{x_n\}$ of a topological space $(X,\mathcal{T})$ converges to $x\in X$ if $\forall U\ni x$ is open, there $\exists N$ such that $x_n\in U, \forall n\geq N$.

■ **Example 2.9** 1. The topology for the space $(X=\mathbb{R}^n,d_2)\to(X,\mathcal{T})$ (i.e., a topological space induced from meric space $(X=\mathbb{R}^n,d_2)$) is called a **usual topology** on $\mathbb{R}^n$.

When I say $\mathbb{R}^n$ (or subset of $\mathbb{R}^n$) is a topological space, it is equipeed with usual topology.

Convergence of sequence in $(\mathbb{R}^n,\mathcal{T})$ is the usual convergence in analysis.

For $\mathbb{R}^n$ or metric space, the limit of sequence (if exists) is unique.

2. Consider the topological space $(X, \mathcal{T}_{\text{indiscrete}})$. Take any sequence $\{x_n\}$ in X, it is convergent to any $x \in X$. Indeed, for $\forall U \ni x$ open, $U = X$. Therefore,

$$x_n \in U(= X), \forall n \geq 1.$$

3. Consider the topological space $(X, \mathcal{T}_{\text{cofinite}})$, where X is infinite. Consider $\{x_n\}$ is a sequence satisfying $m \neq n$ implies $x_m \neq x_n$. Then $\{x_n\}$ is convergent to any $x \in X$. (Question: how to define openness for $\mathcal{T}_{\text{cofinite}}$ and $\mathcal{T}_{\text{indiscrete}}$)?
4. Consider the topological space $(X, \mathcal{T}_{\text{discrete}})$, the sequence $\{x_n\} \to x$ is equivalent to say $x_n = x$ for all sufficiently large n.

R The limit of sequences may not be unique. The reason is that "$\mathcal{T}$ is not big enough". We will give a criterion to make sure the limit is unique in the future. (Hausdorff)

Proposition 2.9 If $F \subseteq (X, \mathcal{T})$ is closed, then for any convergent sequence $\{x_n\}$ in F, the limit(s) are also in F.

Proof. Let $\{x_n\}$ be a sequence in F with limit $x \in X$. Suppose on the contrary that $x \notin F$ (i.e., $x \in X \setminus F$ that is open). There exists N such that

$$x_n \in X \setminus F, \forall n \geq N,$$

i.e., $x_n \notin F$, which is a contradiction. ■

R The converse may not be true. If the $(X, \mathcal{T})$ is metrizable, the converse holds.

Counter-example: Consider the co-countable topological space $(X = \mathbb{R}, \mathcal{T}_{\text{co-co}})$, where

$$\mathcal{T}_{\text{co-co}} = \{U \mid X \setminus U \text{ is a countable set}\} \bigcup \{\emptyset\},$$

and X is uncontable. Then note that $F = [0,1] \subsetneq X$ is an un-countable set, and under co-countable topology, $F \supseteq \{x_n\} \to x$ implies $x_n = x \in F$ for all n. It's clear that $X \setminus F \notin \mathcal{T}_{\text{co-co}}$, i.e., F is not closed.

2.3.2. Interior, Closure, Boundary

Definition 2.5 Let $(X, \mathcal{T})$ be a topological space, and $A \subseteq X$ a subset.

1. The **interior** of A is

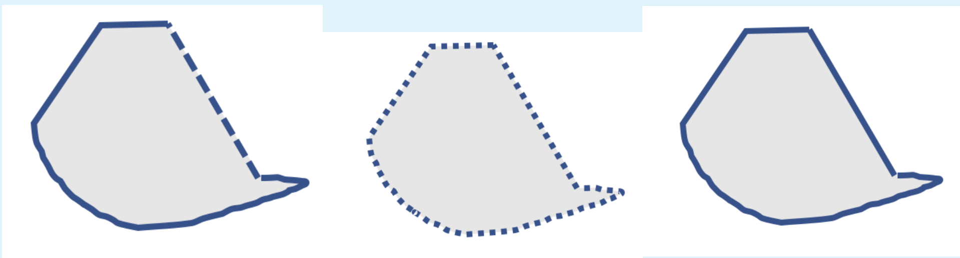
$$A^{\circ} = \bigcup_{U \subseteq A, U \text{ is open}} U$$

2. The **closure** of A is

$$\overline{A} = \bigcap_{A \subseteq V, V \text{ is closed}} V$$

If $\overline{A} = X$, we say that A is dense in X.

The graph illustration of the definition above is as follows:



(a) Illustration of A (b) Illustration of A° (c) Illustration of $\overline{A}$

Figure 2.1: Graph Illustrations

■ **Example 2.10** 1. For $[a,b) \subseteq \mathbb{R}$, we have:

$$[a,b)^{\circ} = (a,b), \quad \overline{[a,b)} = [a,b]$$

2. For $X = \mathbb{R}$, $\mathbb{Q}^{\circ} = \emptyset$ and $\overline{\mathbb{Q}} = \mathbb{R}$.

3. Consider the discrete topology $(X, \mathcal{T}_{\text{discrete}})$, we have

$$S^\circ = S, \quad \overline{S} = S$$

The insights behind the definition (2.5) is as follows

Proposition 2.10 1. A° is the largest open subset of X contained in A; $\overline{A}$ is the smallest closed subset of X containing A.

2. If $A \subseteq B$, then $A^\circ \subseteq B$ and $\overline{A} \subseteq \overline{B}$
3. A is open in X is equivalent to say $A^\circ = A$; A is closed in X is equivalent to say $\overline{A} = A$.

Example 2.11 Let (X, d) be a metric space. What's the closure of an open ball $B_r(x)$? The direct intuition is to define the closed ball

$$\bar{B}_r(x) = \{y \in X \mid d(x, y) \le r\}.$$

Question: is $\bar{B}_r(x) = \overline{B_r(x)}$?

1. Since $\bar{B}_r(x)$ is a closed subset of X, and $B_r(x) \subseteq \bar{B}_r(x)$, we imply that

$$\overline{B_r(x)} \subseteq \bar{B}_r(x)$$

2. However, we may find an example such that $\overline{B_r(x)}$ is a proper subset of $\bar{B}_r(x)$: Consider the discrete metric space (X, d_{discrete}) and for $\forall x \in X$,

$$B_1(x) = \{x\} \implies \overline{B_1(x)} = \{x\}, \quad \bar{B}_1(x) = X$$

The equality $\bar{B}_r(x) = \overline{B_r(x)}$ holds when (X, d) is a normed space.

Here is another characterization of $\overline{A}$:

Proposition 2.11

$$\overline{A} = \{x \in X \mid \forall \text{open } U \ni x, U \bigcap A \neq \emptyset\}$$

Proof. Define

$$S = \{x \in X \mid \forall \text{open } U \ni x, U \bigcap A \neq \emptyset\}$$

It suffices to show that $\overline{A} = S$.

1. First show that S is closed:

   $$X \setminus S = \{x \in X \mid \exists U_x \ni x \text{ open s.t. } U_x \bigcap A = \emptyset\}$$

   Take $x \in X \setminus S$, we imply there exists open $U_x \ni x$ such that $U_x \bigcap A = \emptyset$. We claim $U_x \subseteq X \setminus S$:

   - For $\forall y \in U_x$, note that $U_x \ni y$ that is open, such that $U_x \bigcap A = \emptyset$. Therefore, $y \in X \setminus S$.

   Therefore, we have $x \in U_x \subseteq X \setminus S$ for any $\forall x \in X \setminus S$.

   Note that

   $$X \setminus S = \bigcup_{x \in X \setminus S} \{x\} \subseteq \bigcup_{x \in X \setminus S} U_x \subseteq X \setminus S,$$

   which implies $X \setminus S = \bigcup_{x \in X \setminus S} U_x$ is open, i.e., S is closed in X.

2. By definition, it is clear that $A \subseteq S$:

   $$\forall a \in A, \forall \text{open } U \ni a, U \bigcap A \supseteq \{a\} \neq \emptyset \implies a \in S.$$

   Therefore, $\overline{A} \subseteq \overline{S} = S$.

3. Suppose on the contrary that there exists $y \in S \setminus \overline{A}$.

   Since $y \notin \overline{A}$, by definition, there exists $F \supseteq A$ closed such that $y \notin F$.

   Therefore, $y \in X \setminus F$ that is open, and

   $$(X \setminus F) \bigcap A \subseteq (X \setminus A) \bigcap A = \emptyset \implies y \notin S,$$

   which is a contradiction. Therefore, $S = \overline{A}$.

■

Definition 2.6 [accumulation point] Let $A \subseteq X$ be a subset in a topological space. We call $x \in X$ are an **accumulation point** (**limit point**) of A if

$$\forall U \subseteq X \text{ open s.t. } U \ni x, (U \setminus \{x\}) \bigcap A \neq \emptyset.$$

The set of accumulation points of A is denoted as A'

Proposition 2.12 $\overline{A} = A \bigcup A'$.

Proof. This proposition directly follows from Proposition (2.11) and the definition of A′. ■

2.4. Wednesday for MAT3040

Reviewing.

- Basis, Dimension
- Basis Extension
- $W_1 \bigcap W_2 = \emptyset$ implies $W_1 \oplus W_2 = W_1 + W_2$ (Direct Sum).

2.4.1. Remark on Direct Sum

Proposition 2.13 The set $W_1 + W_2 = W_1 \oplus W_2$ iff any $\boldsymbol{w} \in W_1 + W_2$ can be uniquely expressed as

$$\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2,$$

where $\boldsymbol{w}_i \in W_i$ for $i = 1, 2$.

R We can also define addiction among finite set of vector spaces $\{W_1, \dots, W_k\}$. If $\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k = \boldsymbol{0}$ implies $\boldsymbol{w}_i = 0, \forall i$, then we can write $W_1 + \cdots + W_k$ as

$$W_1 \oplus \cdots \oplus W_k$$

Proposition 2.14 — **Complementation.** Let $W \le V$ be a vector subspace of a fintie dimension vector space V. Then there exists $W' \le V$ such that

$$W \oplus W' = V.$$

Proof. It's clear that $\dim(W) := k \le n := \dim(V)$. Suppose $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is a basis of W.

By the basis extension proposition, we can extend it into $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n\}$, which is a basis of V.

Therefore, we take $W' = \operatorname{span}\{\boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n\}$, which follows that

1. $W + W' = V$: $\forall \boldsymbol{v} \in V$ has the form

$$\boldsymbol{v} = (\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_k \boldsymbol{v}_k) + (\alpha_{k+1} \boldsymbol{v}_{k+1} + \cdots + \alpha_n \boldsymbol{v}_n),$$

   where $\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_k \boldsymbol{v}_k \in W$ and $\alpha_{k+1} \boldsymbol{v}_{k+1} + \cdots + \alpha_n \boldsymbol{v}_n \in W'$.

2. $W \bigcap W' = \{\mathbf{0}\}$: Suppose $\boldsymbol{v} \in W \bigcap W'$, i.e.,

$$\begin{aligned}\boldsymbol{v} &= \left(\beta_1 \boldsymbol{v}_1 + \cdots + \beta_k \boldsymbol{v}_k\right) + (0\boldsymbol{v}_{k+1} + \cdots + 0\boldsymbol{v}_n) \in W \\ &= (0\boldsymbol{v}_1 + \cdots + 0\boldsymbol{v}_k) + \left(\beta_{k+1} \boldsymbol{v}_{k+1} + \cdots + \beta_n \boldsymbol{v}_n\right) \in W'.\end{aligned}$$

   By the uniqueness of coordinates, we imply $\beta_1 = \cdots = \beta_n = 0$, i.e., $\boldsymbol{v} = \mathbf{0}$.

Therefore, we conclude that $W \oplus W' = V$. ■

2.4.2. Linear Transformation

Definition 2.7 [Linear Transformation] Let V, W be vector spaces. Then $T : V \to W$ is a **linear transformation** if

$$T(\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2) = \alpha T(\boldsymbol{v}_1) + \beta T(\boldsymbol{v}_2),$$

for $\forall \alpha, \beta \in \mathbb{F}$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$.

Proposition 2.15

1. Suppose that $S : V \to W$ and $T : W \to U$ are linear transformations, then so is $T \circ S : V \to U$.
2. For any linear transformation $T : V \to W$, we have

$$T(\mathbf{0}_V) = \mathbf{0}_W$$

Proof. Simply apply the definition of the linear transformation. ■

■ **Example 2.12** 1. The transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ defined as $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ (where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$) is a linear transformation.

2. The transformation $T : \mathbb{R}[x] \to \mathbb{R}[x]$ defined as

$$p(x) \mapsto T(p(x)) = p'(x), \quad p(x) \mapsto T(p(x)) = \int_0^x p(t)\,\mathrm{d}t$$

is a linear transformation

3. The transformation $T : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ defined as

$$\boldsymbol{A} \mapsto \operatorname{trace}(\boldsymbol{A}) := \sum_{i=1}^{n} a_{ii}$$

is a linear transformation.

However, the transformation

$$\boldsymbol{A} \mapsto \det(\boldsymbol{A})$$

is not a linear transformation.

■

Definition 2.8 [Kernel/Image] Let $T : V \to W$ be a linear transfomation.

1. The **kernel** of T is

$$\ker(T) = T^{-1}(\mathbf{0}) = \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \mathbf{0}\}$$

2. The **image** (or range) of T is

$$\operatorname{Im}(T) = T(\boldsymbol{v}) = \{T(\boldsymbol{v}) \in W \mid \boldsymbol{v} \in V\}$$

■

■ **Example 2.13** 1. Let $T:\mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, then

$$\ker(T) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\} = \text{Null}(\boldsymbol{A}) \qquad \text{Null Space}$$

and

$$\text{Im}(T) = \{\boldsymbol{A}\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{R}^n\} = \text{Col}(\boldsymbol{A}) = \text{span}\{\text{columns of } \boldsymbol{A}\} \qquad \text{Column Space}$$

2. For $T(p(x)) = p'(x)$, $\ker(T) = \{\text{constant polynomials}\}$ and $\text{Im}(T) = \mathbb{R}[x]$.

Proposition 2.16 The kernel or image for a linear transformation $T: V \to W$ also forms a vector subspace:

$$\ker(T) \le V, \quad \text{Im}(T) \le W$$

Proof. For $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \ker(T)$, we imply

$$T(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) = \boldsymbol{0},$$

which implies $\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2 \in \ker(T)$.

The remaining proof follows similarly. ■

Definition 2.9 [Rank/Nullity] Let V, W be finite dimensional vector spaces and $T: V \to W$ a linear transformation. Then we define

$$\text{rank}(T) = \dim(\text{im}(T))$$
$$\text{nullity}(T) = \dim(\ker(T))$$

R Let

$$\text{Hom}_{\mathbb{F}}(V, W) = \{\text{all linear transformations } T: V \to W\},$$

and we can define the addiction and scalar multiplication to make it a vector space:

1. For $T, S \in \mathrm{Hom}_{\mathbb{F}}(V, W)$, define

$$(T+S)(\boldsymbol{v}) = T(\boldsymbol{v}) + S(\boldsymbol{v}),$$

   which implies $T + S \in \mathrm{Hom}_{\mathbb{F}}(V, W)$.

2. Also, define

$$(\gamma T)(\boldsymbol{v}) = \gamma T(\boldsymbol{v}), \qquad \text{for } \forall \gamma \in \mathbb{F},$$

   which implies $\gamma T \in \mathrm{Hom}_{\mathbb{F}}(V, W)$.

In particular, if $V = \mathbb{R}^n, W = \mathbb{R}^m$, then

$$\mathrm{Hom}_{\mathbb{F}}(V, W) = M_{m\times n}(\mathbb{R}).$$

Proposition 2.17 If $\dim(V) = n, \dim(W) = m$, then $\dim(\mathrm{Hom}_{\mathbb{F}}(V, W)) = mn$.

Proposition 2.18 There are anternative characterizations for the injectivity and surjectivity of lienar transformation T:

1. The linear transformation T is injective if and only if

$$\ker(T) = 0, \Longleftrightarrow \mathrm{nullity}(T) = 0.$$

2. The linear transformation T is surjective if and only if

$$\mathrm{im}(T) = W, \Longleftrightarrow \mathrm{rank}(T) = \dim(W).$$

3. If T is bijective, then T^{-1} is a linear transformation.

Proof. 1. (a) For the forward direction of (1),

$$\boldsymbol{x} \in \ker(T) \implies T(\boldsymbol{x}) = 0 = T(\boldsymbol{0}) \implies \boldsymbol{x} = \boldsymbol{0}$$

(b) For the reverse direction of (1),

$$T(\boldsymbol{x}) = T(\boldsymbol{y}) \implies T(\boldsymbol{x}-\boldsymbol{y}) = \boldsymbol{0} \implies \boldsymbol{x}-\boldsymbol{y} \in \ker(T) = \boldsymbol{0} \implies \boldsymbol{x} = \boldsymbol{y}$$

2. The proof follows similar idea in (1).
3. Let $T^{-1}: W \to V$. For all $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$, there exists $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ such that $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$, i.e., $T^{-1}(\boldsymbol{w}_i) = \boldsymbol{v}_i$ $i = 1,2$.

   Consider the mapping

$$\begin{aligned} T(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) &= \alpha T(\boldsymbol{v}_1) + \beta T(\boldsymbol{v}_2) \\ &= \alpha\boldsymbol{w}_1 + \beta\boldsymbol{w}_2, \end{aligned}$$

   which implies $\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2 = T^{-1}(\alpha\boldsymbol{w}_1 + \beta\boldsymbol{w}_2)$, i.e.,

$$\alpha T^{-1}(\boldsymbol{w}_1) + \beta T^{-1}(\boldsymbol{w}_2) = T^{-1}(\alpha\boldsymbol{w}_1 + \beta\boldsymbol{w}_2).$$

■

Definition 2.10 [isomorphism] We say that the vector subspaces V and W are isomorphic if there exists a bijective linear transfomation $T: V \to W$. ($V \cong W$)

This mapping T is called an **isomorphism** from V to W.

R If $\dim(V) = \dim(W) = n < \infty$, then $V \cong W$:

Take $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}, \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ as basis of V and W, respectively. Then one can construct $T: V \to W$ satisfying $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$ for $\forall i$ as follows:

$$T(\alpha_1\boldsymbol{v}_1 + \dots + \alpha_n\boldsymbol{v}_n) = \alpha_n\boldsymbol{w}_1 + \dots + \alpha_n\boldsymbol{w}_n \;\; \forall \alpha_i \in \mathbb{F}$$

It's clear that our constructed T is a linear transformation.

R $V \cong W$ doesn't imply any linear transformations $T: V \to W$ is an isomorphism. e.g., $T(\boldsymbol{v}) = \boldsymbol{0}$ is not an isomorphic if $W \neq \{\boldsymbol{0}\}$.

Theorem 2.3 — **Rank-Nullity Theorem.** Let $T : V \to W$ be a linear transformation with $\dim(V) < \infty$. Then

$$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

Proof. Since $\ker(T) \le V$, by proposition (2.14), there exists $V_1 \le V$ such that

$$V = \ker(T) \oplus V_1.$$

1. Consider the transformation $T|_{V_1} : V_1 \to T(V_1)$, which is an isomorphism, since:
   - Surjectivity is immediate
   - For $\boldsymbol{v} \in \ker(T|_{V_1})$,
     $$T(\boldsymbol{v}) = \mathbf{0} \implies \boldsymbol{v} \in \ker(T),$$
     which implies $\boldsymbol{v} = \mathbf{0}$ since $\boldsymbol{v} \in \ker(T) \cap V_1 = 0$, i.e., the injectivity follows.

   Therefore, $\dim(V_1) = \dim(T(V_1))$.
2. Secondly, given an isomorphism T from X to Y with $\dim(X) < \infty$, then $\dim(X) = \dim(T(X))$. The reason follows from assignment 1 questions (8-9):
   $$\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\} \text{ is a basis of } X \implies \{T(\boldsymbol{v}_1), \dots, T(\boldsymbol{v}_k)\} \text{ is a basis of } Y$$
3. Note that $T(V_1) = T(V) = \text{im}(T)$, since:
   - for $\forall \boldsymbol{v} \in V$, $\boldsymbol{v} = \boldsymbol{v}_k + \boldsymbol{v}_1$, where $\boldsymbol{v}_k \in \ker(T), \boldsymbol{v}_1 \in V_1$, which implies
     $$T(\boldsymbol{v}) = T(\boldsymbol{v}_k) + T(\boldsymbol{v}_1) = \mathbf{0} + T(\boldsymbol{v}_1),$$
     i.e., $T(V) \subseteq T(V_1) \subseteq T(V)$, i.e., $T(V) = T(V_1)$.
4. We can show that $\dim(V) = \dim(\ker(T)) + \dim(V_1)$: Let $\{v_1, \dots, v_k\}$ be a basis of $\ker(T)$, and $\{v_{k+1}, \dots, v_n\}$ be a basis of V_1, then by the proof of complementation

proposition (2.14), we imply $\{v_1, \ldots, v_n\}$ is a basis of V, i.e., $\dim(V) = n = k + (n - k) = \dim(\ker(T)) + \dim(V_1)$.

Therefore, we imply

$$\begin{aligned}
\dim(V) &= \dim(\ker(T)) + \dim(V_1) \\
&= \text{nullity}(T) + \dim(T(V_1)) \\
&= \text{nullity}(T) + \dim(T(V)) \\
&= \text{nullity}(T) + \dim(\text{im}(T)) \\
&= \text{nullity}(T) + \text{rank}(T).
\end{aligned}$$

■

2.5. Wednesday for MAT3006

2.5.1. Compactness

This lecture will talk about the generalization of closeness and boundedness property in $\mathbb{R}^n$. First let's review some simple definitions:

Definition 2.11 [Compact] Let (X,d) be a metric space, and $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ a collection of open sets.

1. $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ is called an **open cover** of $E \subseteq X$ if $E \subseteq \cup_{\alpha\in\mathcal{A}} U_\alpha$
2. A **finite subcover** of $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ is a finite sub-collection $\{U_{\alpha_1},\dots,U_{\alpha_n}\} \subseteq \{U_\alpha\}$ covering E.
3. The set $E \subseteq X$ is **compact** if every open cover of E has a finite subcover.

A well-known result is talked in MAT2006:

Theorem 2.4 — Heine-Borel Theorem. The set $E \subseteq \mathbb{R}^n$ is **compact** if and only if E is closed and bounded.

However, there's a notion of sequentially compact, and we haven't identify its gap and relation with compactness.

Definition 2.12 [Sequentially Compact] Let (X,d) be a metric space. Then $E \subseteq X$ is **sequentially compact** if every sequence in E has a convergent subsequence with limit in E.

A well-known result is talked in MAT2006:

Theorem 2.5 — Bolzano-Weierstrass Theorem. The set $E \subseteq \mathbb{R}^n$ is closed and bounded if and only if E is sequentially compact.

Actually, the definitions of comapctness and the sequential compactness are equivalent under a metric space.

Theorem 2.6 Let (X, d) be a metric space, then $E \subseteq X$ is compact if and only if E is sequentially compact.

Proof. *Necessity*

Suppose $\{x_n\}$ is a sequence in E, it suffices to show it has a convergent subsequence. Consider the tail of $\{x_n\}$, say

$$F_n = \overline{\{x_k \mid k \geq n\}} \implies F_1 \supseteq F_2 \supseteq \cdots.$$

- Note that $\cap_{i=1}^{\infty} F_i \neq \emptyset$. Assume not, then we imply $\cup_{i=1}^{\infty}(E \setminus F_i) = E$, i.e., $\{E \setminus F_i\}_{i=1}^{\infty}$ a open cover of E. By the compactness of E, we imply there exists a finite subcover of E:
$$E = \bigcup_{j=1}^{r}(E \setminus F_{i_j}) \implies \bigcap_{j=1}^{r} F_{i_j} = \emptyset \implies F_{i_r} = \emptyset,$$
which is a contradiction, and there must exist an element $x \in \cap_{n=1}^{\infty} F_i$.

- For any $n \geq 1$ and $x \in \cap_{n=1}^{\infty} F_i$, either $x \in \{x_k \mid k \geq n\}$ or $x \in \partial\{x_k \mid k \geq n\}$. In both cases, the open ball $B_\varepsilon(x)$ must intersect with the n-th tail of the sequence $\{x_n\}$ for any $\varepsilon > 0$:
$$B_\varepsilon(x) \cap \{x_k \mid k \geq n\} \neq \emptyset, \ \forall \varepsilon > 0.$$
Therefore, construct $x_{n_1} \in B_1(x) \cap \{x_k \mid k \geq 1\}$ and for $r > 1$,
$$x_{n_r} \in B_{1/r}(x) \cap \{x_k \mid k \geq n_{r-1} + 1\}.$$
Therefore, the subsequence $x_{n_r} \to x$ as $r \to \infty$. The proof for necessity is complete.

Sufficiency

Firstly, let's assume the claim below hold (which will be shown later):

Proposition 2.19 If $E \subseteq X$ is sequentially compact, then for any$\varepsilon > 0$, there exists finitely many open balls, say $\{B_\varepsilon(x_1), \ldots, B_\varepsilon(x_n)\}$, covering E.

Suppose on the contrary that there exists an open cover $\{U_\alpha\}_{\alpha \in \mathcal{A}}$ of E, that has no finite subcover.

- By proposition (2.19), for $n \geq 1$, there are finitely many balls of radius $1/n$ covering E. Due to our assumption, there exists a open ball $B_{1/n}(y_n)$ such that $B_{1/n}(y_n) \cap E$ cannot be covered by finitely many members in $\{U_\alpha\}_{\alpha \in \mathcal{A}}$.

- Pick $x_n \in B_{1/n}(y_n)$ to form a sequence. Due to the sequential compactness of E, there exists a subsequence $\{x_{n_j}\} \to x$ for some $x \in E$.

- Since $\{U_\alpha\}_{\alpha \in \mathcal{A}}$ covers E, there exists a U_β containing x. Since U_β is open and the radius of $B_{1/n_j}(y_{n_j})$ tends to 0, we imply that, for sufficiently large n_j, the set $B_{1/n_j}(y_{n_j}) \cap E$ is contained in U_β.

  In oteher words, U_β forms a **single** subcover of $B_{1/n}(y) \cap E$, which contradicts to our choice of $B_{1/n_j}(y_{n_j}) \cap E$. The proof for sufficiency is complete.

∎

Proof for proposition (2.19). Pick $B_\varepsilon(x_1)$ for some $x_1 \in E$. Suppose $E \setminus B_\varepsilon(x_1) \neq \emptyset$. We can find $x_2 \notin B_\varepsilon(x_1)$ such that $d(x_2, x_1) \geq \varepsilon$.

Suppose $E \setminus (B_\varepsilon(x_1) \bigcup B_\varepsilon(x_2))$ is non-empty, then we can find $x_3 \notin B_\varepsilon(x_1) \bigcup B_\varepsilon(x_2)$ so that $d(x_j, x_3) \geq \varepsilon$, $j = 1, 2$.

Keeping this procedure, we obtain a sequence $\{x_n\}$ in E such that

$$E \setminus \bigcup_{j=1}^{n} B_\varepsilon(x_j) \neq \emptyset, \qquad \text{and} \qquad d(x_j, x_n) \geq \varepsilon, j = 1, 2, \dots, n-1.$$

By the sequential compactness of E, there exists $\{x_{n_j}\}$ and $x \in E$ so that $x_{n_j} \to x$ as $j \to \infty$. But then $d(x_{n_j}, x_{n_k}) < d(x_{n_j}, x) + d(x_{n_k}, x) \to 0$, which contradicts that $d(x_j, x_n) \geq \varepsilon$ for $\forall j < n$.

Therefore, one must have $E \setminus \bigcup_{j=1}^{N} B_\varepsilon(x_j) = \emptyset$ for some finite N.

The proof is complete. ∎

R

1. Given the condition metric space,

$$\text{Sequential Compactness} \iff \text{Compactness}$$

2. Given the condition metric space, we will show that

$$\text{Compactness} \implies \text{Closed and Bounded}$$

However, the converse may not necessarily hold. Given the condition the metric space is $\mathbb{R}^n$, then

$$\text{Compactness} \iff \text{Closed and Bounded}$$

Proposition 2.20 Let (X,d) be a metric space. Then $E \subseteq X$ is compact implies that E is closed and bounded.

We say a set E if bounded if there exists $K \geq 0$ such that

$$d(e_1,e_2) < K, \quad \forall e_1,e_2 \in E$$

Proof. 1. Let $\{x_n\}$ be a convergent sequence in E. By sequential compactness, $\{x_{n_j}\} \to x$ for some $x \in E$. By the uniqueness of limits, under metric space, $\{x_n\} \to x$ for $x \in E$. The closeness is shown

2. Take $x \in E$ and consider the open cover $\bigcup_{n=1}^{\infty} B_n(x)$ of E. By compactness,

$$E \subseteq \bigcup_{i=1}^{k} B_{n_i}(x) = B_{n_k}(x),$$

which implies that for any $y,z \in E$,

$$d(y,z) \leq d(y,x) + d(x,z) \leq n_k + n_k = 2n_k.$$

The boundness is shown.

■

Here we raise several examples to show that the coverse does not necessarily hold under a metric space.

■ **Example 2.14** Given the metric space $C[0,1]$ and a set $E = \{f \in C[0,1] \mid 0 \le f(x) \le 1\}$. Notice that E is closed and bounded:

- $E = \cap_{x\in[0,1]}\Psi_x^{-1}([0,1])$, where $\Psi_x(f) = f(x)$, which implies that E is closed.
- Note that $E \subseteq B_2(\mathbf{0}) = \{f \mid |f| < 2\}$, i.e., E is bounded.

However, E may not be compact. Consider a sequence $\{f_n\}$ with

$$f_n(x) = \begin{cases} nx, & 0 \le x \le \frac{1}{n} \\ 1, & \frac{1}{n} \le x \le 1 \end{cases}$$

Suppose on the contrary that E is sequentially compact, therefore there exists a subsequence $\{f_{n_k}\} \to f$ under d_∞ metric, which implies, $\{f_{n_k}\}$ uniformly converges to f. By the definition of $f_n(x)$, we imply

$$f(x) = \begin{cases} 0, & x = 0 \\ 1, & x \in (0,1] \end{cases}$$

However, since d_∞ indicates uniform convergence, the limit for $\{f_{n_k}\}$, say f, must be continuous, which is a contradiction. ■

Theorem 2.7 Let the set E be compact in (X,d) and the function $f : (X,d) \to (Y,\rho)$ is continuous. Then $f(E)$ is compact in Y.

Note that the technique to show compactness by using the sequential compactness is very useful. However, this technique only applies to the metric space, but fail in general topological spaces.

Proof. Let $\{y_n\} = \{f(x_n)\}$ be any sequence in $f(E)$.

- By the compactness of X, $\{x_n\}$ has a convergent subsequence $\{x_{n_r}\} \to x$ as $r \to \infty$.

- Therefore, $\{y_{n_r}\} := \{f(x_{n_r})\} \to f(x)$ by the continuity of f.
- Therefore, $f(E)$ is sequentially compact, i.e., compact.

■

R The Theorem (2.7) is a generalization of the statement that *a continuous function on $\mathbb{R}^n$ admits its minimum and maximum*. Note that such an extreme value property no longer holds for arbitrary closed, bounded sets in a general metric space, but it continues to hold when the sets are strengthened to compact ones.

Another characterization of compactness in $C[a,b]$ is shown in the Ascoli-Arzela Theorem (see Theorem (14.1) in MAT2006 Notebook).

2.5.2. Completeness

Definition 2.13 [Complete] Let (X,d) be metric space.

1. A sequence $\{x_n\}$ in (X,d) is a **Cauchy sequence** if for every $\varepsilon > 0$, there exists some N such that $d(x_n,x_m) < \varepsilon$ for all $n,m \ge N$.
2. A subset $E \subseteq X$ is said to be **complete** if every Cauchy sequence in E is convergent.

■ **Example 2.15** The set $X = C[a,b]$ is complete:

- Suppose $\{f_n\}$ is Cauchy in $C[a,b]$, i.e., $\{f_n(x)\}$ is Cauchy in $\mathbb{R}$ for $\forall x \in [a,b]$.
- By the completeness of $\mathbb{R}$, the sequence $f_n(x) \to f(x)$ for some $f(x) \in \mathbb{R}, \forall x \in [a,b]$. It suffices to show $f_n \to f$ uniformly:
  - For fixed $\varepsilon > 0$, there exists $N > 0$ such that

$$d_\infty(f_n, f_{n+k}) < \frac{\varepsilon}{2}, \qquad \forall n \ge N, k \in \mathbb{N}$$

which implies that for $\forall x \in [a,b]$, $\forall n \geq N, k \in \mathbb{N}$,

$$|f_n(x) - f_{n+k}(x)| < \frac{\varepsilon}{2} \implies \lim_{k \to \infty} |f_n(x) - f_{n+k}(x)| \leq \frac{\varepsilon}{2}$$

Therefore, we imply

$$|f_n(x) - f(x)| = \lim_{k \to \infty} |f_n(x) - f_{n+k}(x)| \leq \frac{\varepsilon}{2} < \varepsilon, \qquad \forall n \geq N, x \in [a,b]$$

The proof is complete.

■

2.6. Wednesday for MAT4002

Reviewing.

1. Interior, Closure:

$$\overline{A} = \{x \mid \forall U \ni x \text{ open}, U \bigcap A \neq \emptyset\}$$

2. Accumulation points

2.6.1. Remark on Closure

Definition 2.14 [Sequential Closure] Let A_S be the set of limits of any convergent sequence in A, then A_S is called the **sequential closure** of A.

Definition 2.15 [Accumulation/Cluster Points] The set of accumulation (limit) points is defined as

$$A' = \{x \mid \forall U \ni x \text{ open }, (U \setminus \{x\}) \bigcap A \neq \emptyset\}$$

R

1. (a) There exists some point in A but not in A':

$$A = \{1, 2, 3, \dots, n, \dots\}$$

   Then any point in A is not in A'

   (b) There also exists some point in A' but not in A:

$$A = \{\frac{1}{n} \mid n \geq 1\}$$

   Then the point 0 is in A' but not in A.

2. The closure $\overline{A} = A \bigcup A'$.
3. The size of the sequentical closure A_S is between A and $\overline{A}$, i.e., $A \subseteq A_S \subseteq \overline{A}$:

It's clear that $A \subseteq A_S$, since the sequence $\{a_n := a\}$ is convergent to a for $\forall a \in A$.

For all $a \in A_S$, we have $\{a_n\} \to a$. Then for any open $U \ni a$, there exists N such that $\{a_N, a_{N+1}, \dots\} \subseteq U \bigcap A \neq \emptyset$. Therefore, $a \in \overline{A}$, i.e., $A_S \subseteq \overline{A}$.

Question: Is $A_S = \overline{A}$?

Proposition 2.21 Let (X, d) be a metric space, then $A_S = \bar{A}$.

Proof. Let $a \in \overline{A}$, then there exists $a_n \in B_{1/n}(a) \bigcap A$, which implies $\{a_n\} \to a$, i.e., $a \in A_S$. ∎

R If $(X, \mathcal{T})$ is metrizable, then $A_S = \overline{A}$. The same goes for first countable topological spaces. However, A_S is a proper subset of $\overline{A}$ in general:

Let $A \subseteq X$ be the set of continuous functions, where $X = \mathbb{R}^{\mathbb{R}}$ denotes the set of all real-valued functions on $\mathbb{R}$, with the topology of pointwise convergence.

Then $A_S = B_1$, the set of all functions of first Baire-Category on $\mathbb{R}$; and $[A_S]_S = B_2$, the set of all functions of second Baire-Category on $\mathbb{R}$. Since $B_1 \neq B_2$, we have $[A_S]_S = A_S$. Note that $\overline{\overline{A}} = \overline{A}$. We conclude that A_S cannot equal to $\overline{A}$, since the sequential closure operator cannot be idemotenet.

Definition 2.16 [Boundary] The **boundary** of $\boldsymbol{A}$ is defined as

$$\partial \boldsymbol{A} = \overline{A} \setminus A^\circ$$

Proposition 2.22 Let $(X, \mathcal{T})$ be a topological space with $A, B \subseteq X$.

$$\overline{X \setminus A} = X \setminus A^\circ, \quad (X \setminus B)^\circ = X \setminus \overline{B} \quad \partial A = \overline{A} \bigcap (\overline{X \setminus A})$$

Proof.

$$X \setminus A^\circ = X \setminus \left(\bigcup_{U \text{ is open}, U \subseteq A} U \right) \tag{2.2a}$$

$$= \bigcap_{U \text{ is open}, U \subseteq A} (X \setminus U) \tag{2.2b}$$

$$= \bigcap_{V \text{ is closed}, F \supseteq X \setminus A} F \tag{2.2c}$$

$$= \overline{X \setminus A} \tag{2.2d}$$

Denoting $X \setminus A$ by B, we obtain:

$$(X \setminus B)^\circ = A^\circ \tag{2.3a}$$

$$= X \setminus (X \setminus A^\circ) \tag{2.3b}$$

$$= X \setminus \overline{X \setminus A} \tag{2.3c}$$

$$= X \setminus \overline{B} \tag{2.3d}$$

By definition of ∂A,

$$\partial A = \overline{A} \setminus A^\circ \tag{2.4a}$$

$$= \overline{A} \bigcap (X \setminus A^\circ) \tag{2.4b}$$

$$= \overline{A} \bigcap (\overline{X \setminus A}) \tag{2.4c}$$

■

2.6.2. Functions on Topological Space

Definition 2.17 [Continuous] Let $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ be a map. Then the function f is continuous, if

$$U \in \mathcal{T}_Y \implies f^{-1}(U) \in \mathcal{T}_X$$

■ **Example 2.16**

1. The identity map $\text{id} : (X, \mathcal{T}) \to (X, \mathcal{T})$ defined as $x \mapsto x$ is continuous
2. The identity map $\text{id} : (X, \mathcal{T}_{\text{discrete}}) \to (X, \mathcal{T}_{\text{indiscrete}})$ defined as $x \mapsto x$ is continuous. Since $\text{id}^{-1}(\emptyset) = \emptyset$ and $\text{id}^{-1}(X) = X$
3. The identity map $\text{id} : (X, \mathcal{T}_{\text{indiscrete}}) \to (X, \mathcal{T}_{\text{discrete}})$ defined as $x \mapsto x$ is not continuous.

Proposition 2.23 If $f : X \to Y$, and $g : Y \to Z$ be continuous, then $g \circ f$ is continuous

Proof. For given $U \in \mathcal{T}_Z$, we imply

$$g^{-1}(U) \in \mathcal{T}_Y \implies f^{-1}(g^{-1}(U)) \in \mathcal{T}_X,$$

i.e., $(g \circ f)^{-1}(U) \in \mathcal{T}_X$ ■

Proposition 2.24 Suppose $f : X \to Y$ is continuous between two topological spaces. Then $\{x_n\} \to x$ implies $\{f(x_n)\} \to f(x)$.

Proof. Take open $U \ni f(x)$, which implies $f^{-1}(U) \ni x$. Since $f^{-1}(U)$ is open, we imply there exists N such that

$$\{x_n \mid n \geq N\} \subseteq f^{-1}(U),$$

i.e., $\{f(x_n) \mid n \geq N\} \subseteq U$ ■

We use the notion of Homeomorphism to describe the equivalence between two topological spaces.

Definition 2.18 [Homeomorphism] A **homeomorphism** between spaces topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ is a bijection

$$f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y),$$

such that both f and f^{-1} are continuous.

2.6.3. Subspace Topology

Definition 2.19 Let $A \subseteq X$ be a non-empty set. The **subspace topology** of A is defined as:

1. $\mathcal{T}_A := \{U \bigcap A \mid U \in \mathcal{T}_A\}$
2. The **coarsest** topology on A such that the **inclusion map**

$$i : (A, \mathcal{T}_A) \to (X, \mathcal{T}_X), \quad i(x) = x$$

   is continuous.

   (We say the topology $\mathcal{T}_1$ is **coarser** than $\mathcal{T}_2$, or $\mathcal{T}_2$ is **finer** than $\mathcal{T}_1$, if $\mathcal{T}_1 \subseteq \mathcal{T}_2$ e.g., $\mathcal{T}_{\text{discrete}}$ is the finest topology, and $\mathcal{T}_{\text{indiscrete}}$ is coarsest topology.)
3. The (**unique**) topology such that for any $(Y, \mathcal{T}_Y)$,

$$f : (Y, \mathcal{T}_Y) \to (A, \mathcal{T}_A)$$

   is continuous iff $i \circ f : (Y, \mathcal{T}_Y) \to (X, \mathcal{T}_X)$ (where i is the inclusion map) is continuous.

Proposition 2.25 The definition (1) and (2) in (2.19) are equivalent.

Outline. The proof is by applying

$$i^{-1}(S) = S \bigcap A, \quad \forall S$$

∎

Example 2.17 Let all English and numerical letters be subset of $\mathbb{R}^2$:

P,6

The homeomorphism can be constrcuted between these two English letters.

Proposition 2.26 The definition (2) and (3) in (2.19) are equivalent.

Proof. Necessity.

- For $\forall U \in \mathcal{T}_X$, consider that

$$(i \circ f)^{-1}(U) = f^{-1}(i^{-1}(U)) = f^{-1}(U \bigcap A)$$

since $U \bigcap A \in \mathcal{T}_A$ and f is continuous, we imply $(i \circ f)^{-1}(U) \in \mathcal{T}_Y$

- For $\forall U' \in \mathcal{T}_A$, we have $U' = U \bigcap A$ for some $U \in \mathcal{T}_X$. Therefore,

$$f^{-1}(U') = f^{-1}(U \bigcap A) = f^{-1}(i^{-1}(U)) = (i \circ f)^{-1}(U) \in \mathcal{T}_Y.$$

The sufficiency is left as exercise. ■

Proposition 2.27 1. The definition (1) in (2.19) does define a topology of A

2. Closed sets of A under subspace topology are of the form $V \bigcap A$, where V is closed in X

Proposition 2.28 Suppose $(A, \mathcal{T}_A) \subseteq (X, \mathcal{T}_X)$ is a subspace topology, and $B \subseteq A \subseteq X$. Then

1. $\bar{B}^A = \bar{B}^X \bigcap A$.
2. $B^{\circ A} \supseteq B^{\circ X}$

Proof. By proposition (2.27), $\bar{B}^X \bigcap A$ is closed in A, and $\bar{B}^X \bigcap A \supset B$, which implies

$$\bar{B}^A \subseteq \bar{B}^X \bigcap A$$

Note that $\bar{B}^A \supset B$ is closed in A, which implies $\bar{B}^A = V \bigcap A \subseteq V$, where V is closed in X. Therefore,

$$\bar{B}^X \subseteq V \implies \bar{B}^X \bigcap A \subseteq V \bigcap A = \bar{B}^A$$

Therefore, $\bar{B}^A = \bar{B}^X \subseteq V$

■

Can we have $B^{\circ X} = B^{\circ A}$?

2.6.4. Basis (Base) of a topology

Roughly speaking, a basis of a topology is a family of "generators" of the topology.

Definition 2.20 Let $(X, \mathcal{T})$ be a topological space. A family of subsets $\mathcal{B}$ in X is a **basis** for $\mathcal{T}$ if

1. $\mathcal{B} \subseteq \mathcal{T}$, i.e., everything in $\mathcal{B}$ is open
2. Every $U \in \mathcal{T}$ can be written as union of elements in $\mathcal{B}$.

■ **Example 2.18** 1. $\mathcal{B} = \mathcal{T}$ is a basis.

2. For $X = \mathbb{R}^n$,

$$\mathcal{B} = \{B_r(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathbb{Q}^n, r \in \mathbb{Q} \bigcap (0, \infty)\}$$

Exercise: every $(a, b) = \bigcup_{i \in I}(p_i, q_i)$ for $p_i, q_i \in \mathbb{Q}$.

Therefore, $\mathcal{B}$ is countable.

Proposition 2.29 If $(X, \mathcal{T})$ has a countable basis, e.g., $\mathbb{R}^n$, then $(X, \mathcal{T})$ has a second-countable space.

Chapter 3

Week3

3.1. Monday for MAT3040

Reviewing.

1. Complementation. Suppose $\dim(V) = n < \infty$, then $W \le V$ implies that there exists W' such that
$$W \oplus W' = V.$$
2. Given the linear transformation $T : V \to W$, define the set $\ker(T)$ and $\mathrm{Im}(T)$.
3. Isomorphism of vector spaces: $T : V \cong W$
4. Rank-Nullity Theorem

3.1.1. Remarks on Isomorphism

Proposition 3.1 If $T : V \to W$ is an isomorphism, then

1. the set $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is linearly independent in V if and only if $\{T\boldsymbol{v}_1, \dots, T\boldsymbol{v}_k\}$ is linearly independent.
2. The same goes if we replace the linearly independence by spans.
3. If $\dim(V) = n$, then $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ forms a basis of V if and only if $\{T\boldsymbol{v}_1, \dots, T\boldsymbol{v}_n\}$ forms a basis of W. In particular, $\dim(V) = \dim(W)$.
4. Two vector spaces with finite dimensions are isomorphic if and only if they have the same dimension:

Proof. It suffices to show the reverse direction. Let $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ be two basis of V, W, respectively. Define the linear transformation $T : V \to W$ by

$$T(a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n) = a_1\boldsymbol{w}_1 + \cdots + a_n\boldsymbol{w}_n$$

Then T is surjective since $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ spans W; T is injective since $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ is linearly independent. ■

3.1.2. Change of Basis and Matrix Representation

Definition 3.1 [Coordinate Vector] Let V be a finite dimensional vector space and $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ an **ordered** basis of V. Any vector $\boldsymbol{v} \in V$ can be uniquely written as

$$\boldsymbol{v} = \alpha_1\boldsymbol{v}_1 + \cdots + \alpha_n\boldsymbol{v}_n,$$

Therefore we define the map $[\cdot]_{\mathcal{B}} : V \to \mathbb{F}^n$, which maps any vector in $\boldsymbol{v}$ into its **coordinate vector**:

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

R Note that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and $\{\boldsymbol{v}_2, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ are distinct ordered basis.

■ **Example 3.1** Given $V = M_{2\times 2}(\mathbb{F})$ and the ordered basis

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \right\}$$

Any matrix has the coordinate vector w.r.t. $\mathcal{B}$, i.e.,

$$\left[\begin{pmatrix}1 & 4\\ 2 & 3\end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix}1\\4\\2\\3\end{pmatrix}$$

However, if given another ordered basis

$$\mathcal{B}_1 = \left\{\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, \right\},$$

the matrix may have the different coordinate vector w.r.t. $\mathcal{B}_1$:

$$\left[\begin{pmatrix}1 & 4\\ 2 & 3\end{pmatrix}\right]_{\mathcal{B}_1} = \begin{pmatrix}4\\1\\2\\3\end{pmatrix}$$

■

Theorem 3.1 The mapping $[\cdot]_{\mathcal{B}} : V \to \mathbb{F}^n$ is an isomorphism.

Proof. 1. First show the operator $[\cdot]_{\mathcal{B}}$ is well-defined, i.e., the same input gives the same output. Suppose that

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} \quad [\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha_n'\end{pmatrix},$$

then we imply

$$\begin{aligned}\boldsymbol{v} &= \alpha_1\boldsymbol{v}_1 + \cdots + \alpha_n\boldsymbol{v}_n\\ &= \alpha_1'\boldsymbol{v}_1 + \cdots + \alpha_n'\boldsymbol{v}_n.\end{aligned}$$

By the uniqueness of coordinates, we imply $\alpha_i = \alpha_i'$ for $i = 1,\dots,n$.

2. It's clear that the operator $[\cdot]_{\mathcal{B}}$ is a linear transformation, i.e.,

$$[p\boldsymbol{v} + q\boldsymbol{w}]_{\mathcal{B}} = p[\boldsymbol{v}]_{\mathcal{B}} + q[\boldsymbol{w}]_{\mathcal{B}} \quad \forall p,q \in \mathbb{F}$$

3. The operator $[\cdot]_B$ is surjective:

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \implies \boldsymbol{v} = 0\boldsymbol{v}_1 + \cdots + 0\boldsymbol{v}_n = \mathbf{0}.$$

4. The injective is clear, i.e., $[\boldsymbol{v}]_{\mathcal{B}} = [\boldsymbol{w}]_{\mathcal{B}}$ implies $\boldsymbol{v} = \boldsymbol{w}$.

Therefore, $[\cdot]_B$ is an isomorphism. ■

We can use the Theorem (3.1) to simplify computations in vector spaces:

■ **Example 3.2** Given a vector sapce $V = P_3[x]$ and its basis $B = \{1,x,x^2,x^3\}$.

To check if the set $\{1+x^2, 3-x^3, x-x^3\}$ is linearly independent, by part (1) in Proposition (3.1) and Theorem (3.1), it suffices to check whether the corresponding coordinate vectors

$$\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}\right\}$$

is linearly independent, i.e., do Gaussian Elimination and check the number of pivots. ■

Here gives rise to the question: if $\mathcal{B}_1, \mathcal{B}_2$ form two basis of V, then how are $[\boldsymbol{v}]_{\mathcal{B}_1}, [\boldsymbol{v}]_{\mathcal{B}_2}$ related to each other?

Here we consider an easy example first:

■ **Example 3.3** Consider $V = \mathbb{R}^n$ and its basis $\mathcal{B}_1 = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. For any $\boldsymbol{v} \in V$,

$$\boldsymbol{v} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_n \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n \implies [\boldsymbol{v}]_{\mathcal{B}_1} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Also, we can construct a different basis of V:

$$\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\},$$

which gives a different coordinate vector of $\boldsymbol{v}$:

$$[\boldsymbol{v}]_{\mathcal{B}_2} = \begin{pmatrix} \alpha_1 - \alpha_2 \\ \alpha_2 - \alpha_3 \\ \vdots \\ \alpha_{n-1} - \alpha_n \\ \alpha_n \end{pmatrix}$$

■

Proposition 3.2 — **Change of Basis.** Let $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\mathcal{A}' = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ be two ordered basis of a vector space V. Define the **change of basis** matrix from $\mathcal{A}$ to $\mathcal{A}'$, say $C_{\mathcal{A}',\mathcal{A}} := [\alpha_{ij}]$, where

$$\boldsymbol{v}_j = \sum_{i=1}^{m} \alpha_{ij} \boldsymbol{w}_i$$

Then for any vector $\boldsymbol{v} \in V$, the *change of basis amounts to left-multiplying the change of basis matrix*:

$$C_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}]_A = [\boldsymbol{v}]_{A'} \tag{3.1}$$

Define matrix $C_{\mathcal{A},\mathcal{A}'} := [\beta_{ij}]$, where

$$\boldsymbol{w}_j = \sum_{i=1}^{n} \beta_{ij} \boldsymbol{v}_i$$

Then we imply that

$$(C_{\mathcal{A},\mathcal{A}'})^{-1} = C_{\mathcal{A}',\mathcal{A}}$$

Proof. 1. First show (3.1) holds for $\boldsymbol{v} = \boldsymbol{v}_j,\ j = 1,\dots,n$:

$$\text{LHS of (3.1)} = [\alpha_{ij}]\boldsymbol{e}_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

$$\text{RHS of (3.1)} = [\boldsymbol{v}_j]_{\mathcal{A}'} = \left[\sum_{i=1}^{n} \alpha_i \boldsymbol{w}_i\right]_{\mathcal{A}'} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

Therefore,

$$C_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}_j]_{\mathcal{A}} = [\boldsymbol{v}_j]_{\mathcal{A}'}, \quad \forall j = 1,\dots,n. \tag{3.2}$$

2. Then for any $\boldsymbol{v} \in V$, we imply $\boldsymbol{v} = r_1\boldsymbol{v}_1 + \cdots + r_n\boldsymbol{v}_n$, which implies that

$$\begin{aligned}
C_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}]_{\mathcal{A}} &= C_{\mathcal{A}',\mathcal{A}}[r_1\boldsymbol{v}_1 + \cdots + r_n\boldsymbol{v}_n]_{\mathcal{A}} && \text{(3.3a)} \\
&= C_{\mathcal{A}',\mathcal{A}}(r_1[\boldsymbol{v}_1]_A + \cdots + r_n[\boldsymbol{v}_n]_{\mathcal{A}}) && \text{(3.3b)} \\
&= \sum_{j=1}^{n} r_j C_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}_j]_{\mathcal{A}} && \text{(3.3c)} \\
&= \sum_{j=1}^{n} r_j [\boldsymbol{v}_j]_{\mathcal{A}'} && \text{(3.3d)} \\
&= \left[\sum_{j=1}^{n} r_j \boldsymbol{v}_j\right]_{\mathcal{A}'} && \text{(3.3e)} \\
&= [\boldsymbol{v}]_{\mathcal{A}'} && \text{(3.3f)}
\end{aligned}$$

where (3.3a) and (3.3e) is by applying the lineaity of $[\cdot]_{\mathcal{A}}$ and $[\cdot]_{\mathcal{A}'}$; (3.3d) is by applying the result (3.13). Therefore (3.1) is shown for $\forall \boldsymbol{v} \in V$.

3. Now we show that $(C_{\mathcal{A}\mathcal{A}'}C_{\mathcal{A}'\mathcal{A}}) = \boldsymbol{I}_n$. Note that

$$\begin{aligned} \boldsymbol{v}_j &= \sum_{i=1}^{n} \alpha_{ij} \boldsymbol{w}_i \\ &= \sum_{i=1}^{n} \alpha_{ij} \sum_{k=1}^{n} \beta_{ki} \boldsymbol{v}_k \\ &= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \beta_{ki} \alpha_{ij} \right) \boldsymbol{v}_i \end{aligned}$$

By the uniqueness of coordinates, we imply

$$\left(\sum_{i=1}^{n} \beta_{ki} \alpha_{ij} \right) = \delta_{jk} := \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

By the matrix multiplication, the (k, j)-th entry for $C_{\mathcal{A}\mathcal{A}'}C_{\mathcal{A}'\mathcal{A}}$ is

$$[C_{\mathcal{A}\mathcal{A}'}C_{\mathcal{A}'\mathcal{A}}]_{kj} = \left(\sum_{i=1}^{n} \beta_{ki} \alpha_{ij} \right) = \delta_{jk} \implies (C_{\mathcal{A}\mathcal{A}'}C_{\mathcal{A}'\mathcal{A}}) = \boldsymbol{I}_n$$

Noew, suppose

$$\begin{aligned} \boldsymbol{v}_j &= \sum_{i=1}^{n} \alpha_{ij} \boldsymbol{w}_i \\ &= \sum_{i=1}^{n} \alpha_{ij} \sum_{k=1}^{n} \beta_{ki} \boldsymbol{v}_k \\ &= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \beta_{ki} \alpha_{ij} \right) \boldsymbol{v}_i \end{aligned}$$

By the uniqueness of coordinates, we imply

$$\left(\sum_{i=1}^{n} \beta_{ki} \alpha_{ij} \right) = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

where

$$\left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij}\right) = (C_{AA'}C_{A'A}).$$

Therefore, $(C_{AA'}C_{A'A}) = \boldsymbol{I}_n$. ■

■ **Example 3.4** Back to Example (3.3), write $\mathcal{B}_1, \mathcal{B}_2$ as

$$\mathcal{B}_1 = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}, \quad \mathcal{B}_2 = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$$

and therefore $\boldsymbol{w}_i = \boldsymbol{e}_1 + \cdots + \boldsymbol{e}_i$. The change of basis matrix is given by

$$C_{\mathcal{B}_1,\mathcal{B}_2} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

which implies that for $\boldsymbol{v}$ in the example,

$$C_{\mathcal{B}_1,\mathcal{B}_2}[\boldsymbol{v}]_{\mathcal{B}_2} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 - \alpha_2 \\ \vdots \\ \alpha_{n-1} - \alpha_n \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = [\boldsymbol{v}]_{\mathcal{B}_1}$$

■

Definition 3.2 Let $T : V \longrightarrow W$ be a linear transformation, and

$$\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}, \quad \mathcal{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$$

be basis of V and W, respectively. The **matrix representation** of T with respect to (w.r.t.) $\mathcal{A}$ and $\mathcal{B}$ is defined as $(T)_{\mathcal{B}\mathcal{A}} := (\alpha_{ij}) \in M_{m\times m}(\mathbb{F})$, where

$$T(\boldsymbol{v}_j) = \sum_{i=1}^{m} \alpha_{ij}\boldsymbol{w}_i$$

3.2. Monday for MAT3006

Reviewing.

1. Compactness/Sequential Compactness:
   - Equivalence for metric space
   - Stronger than closed and bounded
2. Completeness:
   - The metric space (E,d) is complete if every Cauchy sequence on E is convergent.
   - $\mathbb{P}[a,b] \subseteq C[a,b]$ is not complete:

$$f_N(x) = \sum_{n=0}^{N}(-1)^n \frac{x^{2n}}{(2n)!} \to \cos x,$$

     while $\cos x \notin \mathcal{P}[a,b]$.

3.2.1. Remarks on Completeness

Proposition 3.3 Let (X,d) be a metric space.

1. If X is complete and $E \subseteq X$ is closed, then E is complete.
2. If $E \subseteq X$ is complete, then E is closed in X.
3. If $E \subseteq X$ is compact, then E is complete.

Proof. 1. Every Cauchy sequence $\{e_n\}$ in $E \subseteq X$ is also a Cauchy sequence in E.
   Therefore we imply $\{e_n\} \to x \in X$, due to the completeness of X.
   Due to the closedness of E, the limit $x \in E$, i.e., E is complete.
2. Consider any convergent sequence $\{e_n\}$ in E, with some limit $x \in X$.
   We imply $\{e_n\}$ is Cauchy and thus $\{e_n\} \to e \in E$, due to the completeness of E.
   By the uniqueness of limits, we must have $x = z \in E$, i.e., E is closed.
3. Consider a Cauchy sequence $\{e_n\}$ in E. There exists a subsequence $\{e_{n_j}\} \to e \in E$, due to the sequential compactness of E.

It follows that for large n and j,

$$d(e_n, e) \overset{\text{(a)}}{\leq} d(e_n, e_{n_j}) + d(e_{n_j}, e) \overset{\text{(b)}}{<} \varepsilon$$

where (a) is due to triangle inequality and (b) is due to the Cauchy property of $\{e_n\}$ and the convergence of $\{e_{n_j}\}$.

Therefore, we imply $\{e_n\} \to e \in E$, i.e., E is complete.

■

R Given any metric space that may not be necessarily complete, we can make the union of it with another space to make it complete, e.g., just like the completion from $\mathbb{Q}$ to $\mathbb{R}$.

3.2.2. Contraction Mapping Theorem

The motivation of the contraction mapping theorem comes from solving an equation $f(x)$. More precisely, such a problem can be turned into a problem for fixed points, i.e., it suffices to find the fixed points for $g(x)$, with $g(x) = f(x) + x$.

Definition 3.3 Let (X, d) be a metric space. A map $T : (X, d) \to (X, d)$ is a **contraction** if there exists a constant $\tau \in (0, 1)$ such that

$$d(T(x), T(y)) < \tau \cdot d(x, y), \quad \forall x, y \in X$$

A point x is called a fixed point of T if $T(x) = x$.

R All contractions are continuous: Given any convergence sequence $\{x_n\} \to x$, for $\varepsilon > 0$, take N such that $d(x_n, x) < \frac{\varepsilon}{\tau}$ for $n \geq N$. It suffices to show the convergence of $\{T(x_n)\}$:

$$d(T(x_n), T(x)) < \tau \cdot T(x_n, x) < \tau \cdot \frac{\varepsilon}{\tau} = \varepsilon.$$

Therefore, the contraction is Lipschitz continuous with Lipschitz constant τ.

Theorem 3.2 — Contraction Mapping Theorem / Banach Fixed Point Theorem. Every contraction T in a **complete** metric space X has a unique fixed point.

■ **Example 3.5**

1. The mapping $f(x) = x + 1$ is not a contraction in $X = \mathbb{R}$, and it has no fixed point.
2. Consider an in-complete space $X = (0, 1)$ and a contraction $f(x) = \frac{x+1}{2}$. It doesn't admit a fixed point on X as well.

■

Proof. Pick any $x_0 \in X$, and define a sequence recursively by setting $x_{n+1} = T(x_n)$ for $n \geq 0$.

1. Firstly show that the sequence $\{x_n\}$ is Cauchy.

   We can upper bound the term $d(T^n(x_0), T^{n-1}(x_0))$:

$$d(T^n(x_0), T^{n-1}(x_0)) \leq \tau d(T^{n-1}(x_0), T^{n-2}(x_0)) \leq \cdots \leq \tau^{n-1} d(T(x_0), x_0) \tag{3.4}$$

   Therefore for any $n \geq m$, where m is going to be specified later,

$$\begin{aligned} d(x_n, x_m) &= d(T^n(x_0), T^m(x_0)) && \text{(3.5a)} \\ &\leq \tau d(T^{n-1}(x_0), T^{m-1}(x_0)) \leq \cdots \leq \tau^m d(T^{n-m}(x_0), x_0) && \text{(3.5b)} \\ &\leq \tau^m \sum_{j=1}^{n-m} \tau^{n-m-j} d(T(x_0), x_0) && \text{(3.5c)} \\ &< \frac{\tau^m}{1-\tau} d(T(x_0), x_0) && \text{(3.5d)} \\ &\leq \varepsilon && \text{(3.5e)} \end{aligned}$$

   where (3.5b) is by repeatedly applying contraction property of d; (3.5c) is by applying the triangle inequality and (3.4); (3.5e) is by choosing sufficiently large m such that $\frac{\tau^m}{1-\tau} d(T(x_0), x_0) < \varepsilon$.

   Therefore, $\{x_n\}$ is Cauchy. By the completeness of X, we imply $\{x_n\} \to x \in X$.

2. Therefore, we imply

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n) = T(x),$$

i.e., x is a fixed point of T.

Now we show the uniqueness of the fixed point. Suppose that there is another fixed point $y \in X$, then

$$d(x,y) = d(T(x),T(y)) < \tau \cdot d(x,y) \implies d(x,y) < \tau d(x,y), \quad \tau \in (0,1),$$

and we conclude that $d(x,y) = 0$, i.e., $x = y$.

■

■ **Example 3.6** [Convergence of Newton's Method] The Newton's method aims to find the root of $f(x)$ by applying the iteration

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x)}$$

Suppose r is a root for f, the pre-assumption for the convergence of Newton's method is:

1. $f'(r) \neq 0$
2. $f \in C^2$ on some neighborhood of r

Proof. 1. We first show that there exists $[r-\varepsilon, r+\varepsilon]$ such that the mapping

$$\begin{aligned} T: &\quad \mathbb{R} \to \mathbb{R}, \\ \text{with} &\quad x \mapsto x - \tfrac{f(x)}{f'(x)} \end{aligned}$$

satisfies

$$|T'(x)| < 1, \quad \forall x \in [r-\varepsilon, r+\varepsilon]. \tag{3.6}$$

Note that $T'(x) = \frac{f(x)}{[f'(x)]^2} f''(x)$, and we define $h(x) = |T'(x)|$.

It's clear that $h(r) = 0$ and $h(x)$ is continuous, which implies

$$r \in h^{-1}((-1,1)) \implies B_\rho(r) \subseteq h^{-1}((-1,1)) \text{ for some } \rho > 0.$$

Or equivalently, $h((r-\rho, r+\rho)) \subseteq (-1,1)$. Take $\varepsilon = \frac{\rho}{2}$, and the result is obvious.

2. Moreover, for any $x, y \in [r-\varepsilon, r+\varepsilon]$,

$$\begin{aligned} d(T(x), T(y)) &:= |T(x) - T(y)| && \text{(3.7a)} \\ &= |T'(\xi)||x-y| && \text{(3.7b)} \\ &\le \max_{\xi \in [r-\varepsilon, r+\varepsilon]} |T'(\xi)||x-y| && \text{(3.7c)} \\ &< m \cdot |x-y| && \text{(3.7d)} \end{aligned}$$

where (3.7b) is by applying MVT, and ξ is some point in $[r-\varepsilon, r+\varepsilon]$; we assume that $\max_{\xi \in [r-\varepsilon, r+\varepsilon]} |T'(\xi)| < m$ for some $m < 1$ in (3.7d).

3. Note that (2) is not enough to show that T is a contraction. We further need to show that $T(x) \in [r-\varepsilon, r+\varepsilon]$ provided that $x \in [r-\varepsilon, r+\varepsilon]$:

$$|T(x) - r| = |T(x) - T(r)| = |T'(s)||x-r| \le \sup_{[r-\varepsilon, r+\varepsilon]} |T'(s)||x-r| < |x-r|.$$

Combining (2) and (3), we imply T is a contraction on $[r-\varepsilon, r+\varepsilon]$. By applying the contraction mapping theorem, there exists a unique fixed point near $[r-\varepsilon, r+\varepsilon]$:

$$x - \frac{f(x)}{f'(x)} = x \implies \frac{f(x)}{f'(x)} = 0 \implies f(x) = 0,$$

i.e., we obtain a root $x = r$. ■

Summary: if we use Newton's method on any point between $[r-\varepsilon, r+\varepsilon]$ where $f(r) = 0$ and ε is sufficiently small, then we will eventually get close to r.

3.2.3. Picard Lindelof Theorem

We will use Banach fixed point theorem to show the existence and uniqueness of the solution of ODE

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x, y(x)) \\ y(x_0) = y_0 \end{cases} \qquad \textbf{Initial Value Problem, IVP} \tag{3.8}$$

■ **Example 3.7** Consider the IVP

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x^2 y^{1/5} \\ y(x_0) = c > 0 \end{cases} \implies y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}$$

which can be solved by the separation of variables:

$$c > 0 \implies y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}.$$

However, when $c = 0$, the ODE does not have a unique solution. One can verify that y_1, y_2 given below are both solutions of this ODE:

$$y_1 = (\frac{4x^3}{15})^{5/4}, \qquad y_2 = 0$$

This example shows that even when f is very nice, the IVP may not have unique solution. The Picard-Lindelof theorem will give a clean condition on f ensuring the unique solvability of the IVP (3.8). ■

3.3. Monday for MAT4002

3.3.1. Remarks on Basis and Homeomorphism

Reviewing.

1. $A \subseteq A_S \subseteq \overline{A}$, where A_S is sequential closure and $\overline{A}$ denotes closure.
2. Subspace topology.
3. Homeomorphism. Consider the mapping $f : X \to Y$ with the topogical space X, Y shown below, with the standard topology, the question is whether f is continuous?

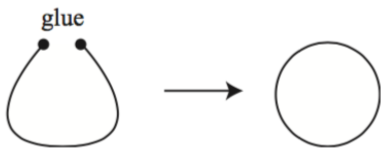


Figure 3.1: Diagram for mapping f

The answer is no, since the left in (3.1) can be isomorphically mapped into $(0,1)$; the right can be isomorphically mapped into $[0,1]$, and the mapping $(0,1) \to [0,1]$ cannot be isomorphism:

Proof. Assume otherwise the mapping $g : (0,1) \to [0,1]$ is isomorphism, and therefore $f^{-1}(U)$ is open for any open set U in the space $[0,1]$.
Construct $U = (1-\delta, 1]$ for $\delta \le 1$, and therfore $f^{-1}((1-\delta,1])$ is open, and therfore for the point $x = f^{-1}(1)$, there exists $\varepsilon > 0$ such that

$$B_\varepsilon(x) \subseteq f^{-1}((1-\delta,1]) \implies [x-\varepsilon, x) \subseteq f^{-1}((1-\delta,1)), \text{ and } (x, x+\varepsilon] \subseteq f^{-1}((1-\delta,1)).$$

which implies that there exists a, b such that $[x-\varepsilon, x) = f^{-1}((a,1))$ and $(x, x+\varepsilon] = f^{-1}((b,1))$, i.e., $f^{-1}((a,b) \cap (b,1))$ admits into two values in $[x-\varepsilon, x)$ and $(x, x+\varepsilon]$, which is a contradiction. ∎

4. Basis of a topology $\mathcal{B} \subseteq (X, \mathcal{T})$ is a collection of open sets in the space such that the whole space can be recovered, or equivalently

(a) $\mathcal{B} \subseteq \mathcal{T}$

(b) Every set in $\mathcal{T}$ can expressed as a union of sets in $\mathcal{B}$

Example: Let $\mathbb{R}^n$ be equipped with usual topology, then

$$B = \{B_q(x) \mid x \in \mathbb{Q}^n, q \in \mathbb{Q}^+\} \text{ is a basis of } \mathbb{R}^n.$$

It suffices to show $U \subseteq \mathbb{R}^n$ can be written as

$$U = U_{x \in \mathbb{Q}} B_{q_x}(x)$$

Proposition 3.4 Let X, Y be topological spaces, and $\mathcal{B}$ a basis for topology on Y. Then

$$f : X \to Y \text{ is continuous} \iff f^{-1}(B) \text{ is open in X, } \forall B \in \mathcal{B}$$

Therefore checking $f^{-1}(U)$ is open for all $U \in \mathcal{T}_Y$ suffices to checking $f^{-1}(N)$ is open for all $B \in \mathcal{B}$.

Proof. The forward direction follows from the fact $B \subseteq \mathcal{T}_Y$.

To show the reverse direction, let $U \in \mathcal{T}_Y$, then $U = \bigcup_{i \in I} B_i$, where $B_i \in \mathcal{B}$, which implies

$$f^{-1}(U) = f^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f^{-1}(B_i)$$

which is open in X by our hypothesis. ■

Corollary 3.1 Let $f : X \to Y$ be a bijection. Suppose there is a basis $\mathcal{B}_X$ of $\mathcal{T}_X$ such that $\{f(B) \mid B \in \mathcal{B}_X\}$ forms a basis of $\mathcal{T}_Y$. Then $X \cong Y$.

Proof. Suppose $W \in \mathcal{T}_Y$, then by our hypothesis,

$$W = \bigcup_{i \in I} f(B_i),\ B_i \in \mathcal{B}_X \implies f^{-1}(W) = \bigcup_{i \in I} B_i \in \mathcal{T}_X,$$

which implies f is continuous.

Suppose $U \in \mathcal{T}_X$, then

$$U = \bigcup_{i \in I} B_i \implies f(U) = \bigcup_{i \in I} f(B_i) \in \mathcal{T}_Y \implies [f^{-1}]^{-1}(U) \in \mathcal{T}_Y,$$

i.e., f^{-1} is continuous. ■

Question: *how to recognise whether a family of subsets is a basis for some given topology?*

Proposition 3.5 Let X be a set, $\mathcal{B}$ is a collection of subsets satisfying

1. X is a union of sets in $\mathcal{B}$, i.e., every $x \in X$ lies in some $B_x \in \mathcal{B}$
2. The intersection $B_1 \cap B_2$ for $\forall B_1, B_2 \in \mathcal{B}$ is a union of sets in $\mathcal{B}$, i.e., for each $B_1, B_2 \in \mathcal{B}$, and $x \in B_1 \cap B_2$, then there exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

Then the collection of subsets $\mathcal{T}_{\mathcal{B}}$, formed by taking any union of sets in $\mathcal{B}$, is a topology, and $\mathcal{B}$ is a basis for $\mathcal{T}_{B}$.

Proof. 1. $\emptyset \in \mathcal{T}_{\mathcal{B}}$ (taking nothing from $\mathcal{B}$); for $x \in X, B_x \in \mathcal{B}$, by hypothesis (1),

$$X = \bigcup_{x \in X} B_x \in \mathcal{T}_{\mathcal{B}}$$

2. Suppose $T_1, T_2 \in \mathcal{T}_{\mathcal{B}}$. Let $x \in T_1 \cap T_2$, where T_i is a union of subsets in $\mathcal{B}$. Therefore,

$$\begin{cases} x \in B_1 \subseteq T_1, & B_1 \in \mathcal{B} \\ x \in B_2 \subseteq T_2, & B_2 \in \mathcal{B} \end{cases}$$

which implies $x \in B_1 \cap B_2$, i.e., $x \in B_x \subseteq B_1 \cap B_2$ for some $B_x \in \mathcal{B}$. Therefore,

$$\bigcup_{x \in B_1 \cap B_2} \{x\} \subseteq \bigcup_{x \in B_1 \cap B_2} B_x \subseteq B_1 \cap B_2,$$

i.e., $B_1 \cap B_2 = \bigcup_{x \in B_1 \cap B_2} B_x$, i.e., $B_1 \cap B_2 \in \mathcal{T}_{\mathcal{B}}$.

3. The property that $\mathcal{T}_{\mathcal{B}}$ is closed under union operations can be checked directly.

The proof is complete. ■

3.3.2. Product Space

Now we discuss how to construct new topological spaces out of given ones is by taking Cartesian products:

Definition 3.4 Let $(X,\mathcal{T}_X),(Y,\mathcal{T}_Y)$ be topological spaces. Consider the family of subsets in $X\times Y$:

$$\mathcal{B}_{X\times Y}=\{U\times V\mid U\in\mathcal{T}_X,V\in\mathcal{T}_y\}$$

This $\mathcal{B}_{X\times Y}$ forms a basis of a topology on $X\times Y$. The induced topology from $\mathcal{B}_{X\times Y}$ is called **product topology**.

For example, for $X=\mathbb{R},Y=\mathbb{R}$, the elements in $\mathcal{B}_{X\times Y}$ are rectangles.

Proof for well-definedness in definition (3.4). We apply proposition (3.5) to check whether $B_{X\times Y}$ forms a basis:

1. For any $(x,y)\in X\times Y$, we imply $x\in X,y\in Y$. Note that $X\in\mathcal{T}_X,Y\in\mathcal{T}_Y$, we imply $(x,y)\in X\times Y\in\mathcal{B}_{X\times Y}$.
2. Suppose $U_1\times V_1,U_2\times V_2\in\mathcal{B}_{X\times Y}$, then

$$(U_1\times V_1)\cap(U_2\times V_2)=(U_1\cap U_2)\times(V_1\cap V_2),$$

where $U_1\cap U_2\in\mathcal{T}_X,V_1\cap V_2\in\mathcal{T}_Y$. Therefore, $(U_1\times V_1)\cap(U_2\times V_2)\in\mathcal{B}_{X\times Y}$.

■

R However, the product topology may not necessarily become the largest topology in the space $X\times Y$. Consider $X=\mathbb{R},Y=\mathbb{R}$, the open set in the space $X\times Y$ may not necessarily be rectangles. However, all elements in $\mathcal{B}_{X\times Y}$ are rectangles.

Example 3.8 The space $\mathbb{R}\times\mathbb{R}$ is homeomorphic to $\mathbb{R}^2$, where the product topology is defined on $\mathbb{R}\times\mathbb{R}$ and the standard topology is defined on $\mathbb{R}^2$:

Construct the function $f:\mathbb{R}\times\mathbb{R}\to\mathbb{R}^2$ with $(a,b)\to(a,b)$.

Obviously, $f : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}^2$ is a bijection.

Take the basis of the topology on $\mathbb{R}$ as open intervals,

$$B_X = \{(a,b) \mid a < b \text{ in } \mathbb{R}\}$$

Therefore, one can verify that the set $\mathcal{B} := \{(a,b) \times (c,d) \mid a < b, c < d\}$ forms a basis for the product topology, and

$$\{f(B) \mid B \in \mathcal{B}\} = \{(a,b) \times (c,d) \mid a < b, c < d\}$$

forms a basis of the usual topology in $\mathbb{R}^2$.

By Corollary (3.1), we imply $\mathbb{R} \times \mathbb{R} \cong \mathbb{R}^2$. ■

We also raise an example on the homeomorphism related to product spaces:

■ **Example 3.9** Let $S^1 = \{(\cos x, \sin x \mid x \in [0, 2\pi])\}$ be a unit circle on $\mathbb{R}^2$.

Consider $f : S^1 \times (0, \infty) \longrightarrow \mathbb{R}^2 \setminus \{\mathbf{0}\}$ defined as

$$f(\cos x, \sin x, r) \mapsto (r\cos x, r\sin x)$$

It's clear that f is a bijection, and f is continuous. Moreover, the inverse $g := f^{-1}$ is defined as

$$g(a,b) = (\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}}, \sqrt{a^2+b^2})$$

which is continuous as well. Therefore, the $f : \mathcal{S}^1 \times (0, \infty) \longrightarrow \mathbb{R}^2 \setminus \{\mathbf{0}\}$ is a homeomorphism.

■

3.4. Wednesday for MAT3040

3.4.1. Remarks for the Change of Basis

Reviewing.

- $[\cdot]_{\mathcal{A}} : V \to \mathbb{F}^n$ denotes coordinate vector mapping
- Change of Basis matrix: $C_{\mathcal{A}',\mathcal{A}}$
- $T : V \to W$, $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\boldsymbol{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$.

  $\mathrm{Hom}_{\mathbb{F}}(V,W) \to M_{m\times n}(\mathbb{F})$

■ **Example 3.10** Let $V = \mathbb{P}_3[x]$ and $\mathcal{A} = \{1, x, x^2, x^3\}$.

Let $T : V \to V$ defined as $p(x) \mapsto p'(x)$:

$$\begin{cases} T(1) = 0\cdot 1 + 0\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x) = 1\cdot 1 + 0\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x^2) = 0\cdot 1 + 2\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x^3) = 0\cdot 1 + 0\cdot x + 3\cdot x^2 + 0\cdot x^3 \end{cases}$$

We can define the change of basis matrix for a linear transformation T as well, w.r.t. $\mathcal{A}$ and $\mathcal{A}$:

$$C_{\mathcal{A},\mathcal{A}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Also, we can define a different basis $\mathcal{A}' = \{x^3, x^2, x, 1\}$ for the output space for T, say $T : V_{\mathcal{A}} \to V_{\mathcal{A}'}$:

$$(T)_{\mathcal{A},\mathcal{A}'} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Our observation is that the corresponding coordinate vectors before and after linear transformation admits a matrix multiplication:

$$(2x^2 + 4x^3) \xrightarrow{T} ((4x + 12x^2))$$

$$(2x^2 + 4x^3)_{\mathcal{A}} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 4 \end{pmatrix} \qquad (4x + 12x^2)_{\mathcal{A}} = \begin{pmatrix} 0 \\ 4 \\ 12 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 12 \\ 0 \end{pmatrix}$$

$$C_{\mathcal{A}\mathcal{A}} \cdot (2x^2 + 4x^3)_{\mathcal{A}} = (4x + 12x^2)_{\mathcal{A}}$$

■

Theorem 3.3 — Matrix Representation. Let $T : V \to W$ be a linear transformation of finite dimensional vector sapces. Let $\mathcal{A}, \mathcal{B}$ the ordered basis of V, W, respectively. Then the following diagram holds:

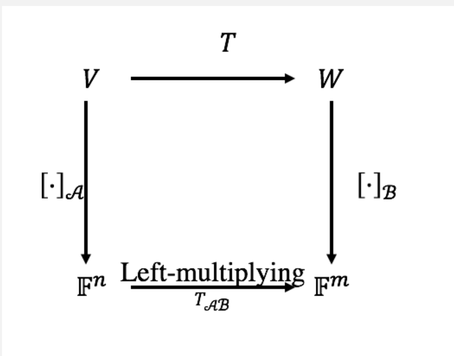


Figure 3.2: Diagram for the matrix reprentation, where $n := \dim(V)$ and $m := \dim(W)$

namely, for any $\boldsymbol{v} \in V$,

$$(T)_{\mathcal{B},\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} = (T\boldsymbol{v})_{\mathcal{B}}$$

Therefore, we can compute $T\boldsymbol{v}$ by matrix multiplication.

Therefore, linear transformation corresponds to coordinate matrix multiplication.

Proof. Suppose $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\mathcal{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$. The proof of this theorem follows the same procedure of that in Theorem (3.1)

1. We show this result for $\boldsymbol{v} = \boldsymbol{v}_j$ first:

$$\text{LHS} = [\alpha_{ij}]\boldsymbol{e}_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

$$\text{RHS} = (T\boldsymbol{v}_j)_{\mathcal{B}} = \left(\sum_{i=1}^{m} \alpha_{ij}\boldsymbol{w}_i\right)_{\mathcal{B}} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

2. Then we show the theorem holds for any $\boldsymbol{v} := \sum_{j=1}^{n} r_j \boldsymbol{v}_j$ in V:

$$
\begin{aligned}
(T)_{\mathcal{B}\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} &= (T)_{\mathcal{B}\mathcal{A}}\left(\sum_{j=1}^{n} r_j \boldsymbol{v}_j\right)_{\mathcal{A}} && \text{(3.9a)}\\
&= (T)_{\mathcal{B}\mathcal{A}}\left(\sum_{j=1}^{n} r_j (\boldsymbol{v}_j)_{\mathcal{A}}\right) && \text{(3.9b)}\\
&= \sum_{j=1}^{n} r_j (T)_{\mathcal{B}\mathcal{A}} (\boldsymbol{v}_j)_{\mathcal{A}} && \text{(3.9c)}\\
&= \sum_{j=1}^{n} r_j (T\boldsymbol{v}_j)_{\mathcal{B}} && \text{(3.9d)}\\
&= \left(\sum_{j=1}^{n} r_j (T\boldsymbol{v}_j)\right)_{\mathcal{B}} && \text{(3.9e)}\\
&= \left[T(\sum_{j=1}^{n} r_j \boldsymbol{v}_j)\right]_{\mathcal{B}} && \text{(3.9f)}\\
&= (T\boldsymbol{v})_{\mathcal{B}} && \text{(3.9g)}
\end{aligned}
$$

The justification for (3.9a) is similar to that shown in Theorem (3.1). The proof is complete.

■

R Consider a special case for Theorem (3.3), i.e., $T = \text{id}$ and $\mathcal{A}, \mathcal{A}'$ are two ordered basis for the input and output space, respectively. Then the result in Theorem (3.3) implies

$$C_{\mathcal{A}',\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} = (\boldsymbol{v})_{\mathcal{A}'}$$

i.e., the matrix representation theorem (3.3) is a general case for the change of basis theorem (3.1)

Proposition 3.6 — **Functoriality.** Suppose V, W, U are finite dimensional vector spaces, and let $\mathcal{A}, \mathcal{B}, C$ be the ordered basis for V, W, U, respectively. Suppose that

$$T : V \to W, \quad S : W \to U$$

are given two linear transformations, then

$$(S \circ T)_{C,\mathcal{A}} = (S)_{C,\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}$$

Composition of linear transformation corresponds to the multiplication of change of basis matrices.

Proof. Suppose the ordered basis $\mathcal{A} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$, $\mathcal{B} = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$, $C = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$. By defintion of change of basis matrices,

$$\begin{aligned} T(\boldsymbol{v}_j) &= \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \boldsymbol{w}_i \\ S(\boldsymbol{w}_i) &= \sum_k (S_{C,\mathcal{B}})_{ki} \boldsymbol{u}_k \end{aligned}$$

We start from the j-th column of $(S \circ T)_{C,\mathcal{A}}$ for $j = 1, \ldots, n$, namely

$$\begin{aligned} (S \circ T)_{C,\mathcal{A}}(\boldsymbol{v}_j)_{\mathcal{A}} &= (S \circ T(\boldsymbol{v}_j))_C && \text{(3.10a)} \\ &= \left[S \circ \left(\sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \boldsymbol{w}_i \right) \right]_C && \text{(3.10b)} \\ &= \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \, (S(\boldsymbol{w}_i))_C && \text{(3.10c)} \\ &= \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \left(\sum_k (S_{C,\mathcal{B}})_{ki} \boldsymbol{u}_k \right)_C && \text{(3.10d)} \\ &= \sum_k \sum_i (S_{C,\mathcal{B}})_{ki} (T_{\mathcal{B},\mathcal{A}})_{ij} (\boldsymbol{u}_k)_C && \text{(3.10e)} \\ &= \sum_k (S_{C,\mathcal{B}} T_{\mathcal{B},\mathcal{A}})_{kj} (\boldsymbol{u}_k)_C && \text{(3.10f)} \\ &= \sum_k (S_{C,\mathcal{B}} T_{\mathcal{B},\mathcal{A}})_{kj} \boldsymbol{e}_k && \text{(3.10g)} \\ &= j\text{-th column of } [S_{C\mathcal{B}} T_{\mathcal{B},\mathcal{A}}] && \text{(3.10h)} \end{aligned}$$

where (3.10a) is by the result in theorem (3.3); (3.10b) and (3.10d) follows from definitions of $T(\boldsymbol{v}_j)$ and $S(\boldsymbol{w}_i)$; (3.10c) and (3.10e) follows from the linearity of C; (3.10f) follows from the matrix multiplication definition; (3.10g) is because $(\boldsymbol{u}_k)_C = \boldsymbol{e}_k$.

Therefore, $(S \circ T)_{C\mathcal{A}}$ and $(S_{C,\mathcal{B}})(T_{\mathcal{B},\mathcal{A}})$ share the same j-th column, and thus equal to each other. ∎

Corollary 3.2 Suppose that S and T are two identity mappings $V \longrightarrow V$, and consider $(S)_{\mathcal{A}'\mathcal{A}}$ and $(T)_{\mathcal{A},\mathcal{A}'}$ in proposition (3.6), then

$$(S \circ T)_{\mathcal{A}',\mathcal{A}'} = (S)_{\mathcal{A}'\mathcal{A}}(T)_{\mathcal{A},\mathcal{A}'}$$

Therefore,

$$\text{Identity matrix} = C_{\mathcal{A}',\mathcal{A}}C_{\mathcal{A},\mathcal{A}'}$$

Proposition 3.7 Let $T : V \longrightarrow W$ with $\dim(V) = n, \dim(W) = m$, and let

- $\mathcal{A}, \mathcal{A}'$ be ordered basis of V
- $\mathcal{B}, \mathcal{B}'$ be ordered basis of W

then the change of basis matrices admit the relation

$$(T)_{\mathcal{B}',\mathcal{A}'} = C_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}C_{\mathcal{A}\mathcal{A}'} \tag{3.11}$$

Here note that $(T)_{\mathcal{B}',\mathcal{A}'}, (T)_{\mathcal{B},\mathcal{A}} \in \mathbb{F}^{m \times n}$; $C_{\mathcal{B}',\mathcal{B}} \in \mathbb{F}^{m \times m}$; and $C_{\mathcal{A}\mathcal{A}'} \in \mathbb{F}^{n \times n}$.

Proof. Let $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}, \mathcal{A}' = \{\boldsymbol{v}'_1, \dots, \boldsymbol{v}'_n\}$. Consider simplifying the j-th column for the LHS and RHS of (3.11) and showing they are equal:

$$\begin{aligned}
\text{LHS} &= (T)_{\mathcal{B}',\mathcal{A}'}\boldsymbol{e}_j \\
&= (T)_{\mathcal{B}',\mathcal{A}'}(\boldsymbol{v}'_j)_{\mathcal{A}'} \\
&= (T\boldsymbol{v}'_j)_{\mathcal{B}'}
\end{aligned}$$

$$
\begin{aligned}
\text{RHS} &= C_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}C_{\mathcal{A}\mathcal{A}'}\boldsymbol{e}_j \\
&= C_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}C_{\mathcal{A}\mathcal{A}'}(\boldsymbol{v}'_j)_{\mathcal{A}'} \\
&= C_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}(\boldsymbol{v}'_j)_{\mathcal{A}} \\
&= C_{\mathcal{B}',\mathcal{B}}(T\boldsymbol{v}'_j)_{\mathcal{B}} \\
&= (T\boldsymbol{v}'_j)_{\mathcal{B}'}
\end{aligned}
$$

■

R Let $T : V \to V$ be a linear operator with $\mathcal{A}, \mathcal{A}'$ being two ordered basisof V, then

$$(T)_{\mathcal{A}'\mathcal{A}'} = C_{\mathcal{A}',\mathcal{A}}(T)_{\mathcal{A}\mathcal{A}}C_{\mathcal{A},\mathcal{A}'} = (C_{\mathcal{A},\mathcal{A}'})^{-1}(T)_{\mathcal{A}\mathcal{A}}C_{\mathcal{A},\mathcal{A}'}$$

Therefore, the change of basis matrices $(T)_{\mathcal{A}'\mathcal{A}'}$ and $(T)_{\mathcal{A}\mathcal{A}}$ are similar to each other, which means they share the same eigenvalues, determinant, trace.

Therefore, two similar matrices cooresponds to same linear transformation using different basis.

3.5. Wednesday for MAT3006

3.5.1. Remarks on Contraction

Reviewing.

- Suppose $E \subseteq X$ with X being complete, then E is closed in X iff E is complete
- Suppose $E \subseteq X$, then E is closed in X if E is complete.
- Contraction Mapping Theorem

3.5.2. Picard-Lindelof Theorem

Consider solving the the initival value problem given below

$$\begin{cases} \frac{\mathrm{d}y}{\mathrm{d}x} = f(x,y) \\ y(\alpha) = \beta \end{cases} \implies y(x) = \beta + \int_{\alpha}^{x} f(t, y(t))\,\mathrm{d}t \tag{3.12}$$

Definition 3.5 Let $R = [\alpha - a, \alpha + a] \times [\beta - b, \beta + b]$. Then the function $f(x,y)$ satisfies the **Lipschitz condition** on R if there exists $L > 0$ such that

$$|f(x,y_1) - f(x,y_2)| < L \cdot |y_1 - y_2|, \quad \forall (x,y_i) \in R \tag{3.13}$$

The smallest number $L^* = \inf\{L \mid \text{The relation (3.13) holds for } L\}$ is called the **Lipschitz constant** for f. ▪

▪ **Example 3.11** A C^1-function f(x,y) in a rectangle automatically satisfies the Lipschitz condition:

$$f(x,y_1) - f(x,y_2) \overset{\text{Appling MVT}}{=} \frac{\partial f}{\partial y}(x,\tilde{y})(y_1 - y_2)$$

Since $\frac{\partial f}{\partial y}$ is continuous on R and thus bounded, we imply

$$|f(x,y_1) - f(x,y_2)| < L \cdot |y_1 - y_2|, \quad \forall (x,y_i) \in R$$

where

$$L = \max\left\{ \text{abs}\left(\frac{\partial f}{\partial y}\right) \middle| (x,y) \in R \right\}$$

Theorem 3.4 — **Picard-Lindelof Theorem (existence part).** Suppose $f \in C(R)$ be such that f satisfies the Lipschitz condition, then there exists $a'' \in (0,a]$ such that (**??**) is solvable with $y(x) \in C([\alpha - a'', \alpha + a''])$.

Proof. Consider the complete metric space

$$X = \{y(x) \in C([\alpha - a, \alpha + a]) \mid \beta - b \le y(x) \le \beta + b\},$$

with the mapping $T : X \to X$ defined as

$$(Ty)(x) = \beta + \int_\alpha^x f(t, y(t))\,\mathrm{d}t$$

It suffices to show that T is a contraction, but here we need to estrict a a smaller number as follows:

1. Well-definedness of T: Take $M := \sup\{f(x,y) \mid (x,y) \in R\}$ and construct $a' = \min\{b/M, a\}$. Consider the complete matric space

$$X = \{y(x) \in C([\alpha - a', \alpha + a']) \mid \beta - b \le y(x) \le \beta + b\}$$

which implies that

$$|(Ty)(x) - \beta| \le \left| \int_\alpha^x f(t, y(t))\,\mathrm{d}t \right| \le M|x - \alpha| \le Ma' \le b,$$

i.e., $T(X) \subseteq X$, and therefore $T : X \to X$ is well-defined.

2. Contraction for T: Construct $a'' \in \min\{a', \frac{1}{2L^*}\}$, where L^* is the Lipschitz constant for f. and consider the complete metric space

$$X = \{y(x) \in C([\alpha - a'', \alpha + a'']) \mid \beta - b \le y(x) \le \beta + b\}$$

Therefore for $\forall x \in [\alpha - a'', \alpha + a'']$ and the mapping $T : X \to X$,

$$\begin{aligned}
|[T(y_1) - T(y_2)](x)| &\le \left| \int_\alpha^x [f(t, y_2(t)) - f(t, y_1(t))] \, \mathrm{d}t \right| \\
&\le \int_\alpha^x |f(t, y_2) - f(t, y_1)| \, \mathrm{d}t \le \int_\alpha^x L^* |y_2(t) - y_1(t)| \, \mathrm{d}t \\
&\le L^* |x - \alpha| \sup |y_2(t) - y_1(t)| \le L^* a'' d_\infty(y_2, y_1) \le \frac{1}{2} d_\infty(y_2, y_1)
\end{aligned}$$

Therefore, we imply $d_\infty(Ty_2, Ty_1) \le \frac{1}{2} d_\infty(y_2, y_1)$, i.e., T is a contraction.

Applying contraction mapping theorem, there exists $y(x) \in X$ such that $Ty = y$, i.e.,

$$y = \beta + \int_\alpha^x f(t, y(t)) \, \mathrm{d}t$$

Thus y is a solution for the IVP (3.12). ■

R On $[\alpha - a'', \alpha + a'']$, we can solve the IVP (3.12) by recursively applying T:

$$\begin{aligned}
y_0(x) &= \beta, \qquad \forall x \in [\alpha - a'', \alpha + a''] \\
y_1 &= T(y_0) = \beta + \int_\alpha^x f(t, \beta) \, \mathrm{d}t \\
y_2 &= T(y_1) \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

By studying (3.12) on different rectangles, we are able to show the uniqueness of our solution:

Proposition 3.8 Suppose f satisfies the Lipschitz conditon, and y_1, y_2 are two solutions of (3.12), where y_1 is defined on $x \in I_1$, and y_2 is defined on $x \in I_2$. Suppose $I_1 \cap I_2 \neq \emptyset$ and y_1, y_2 share the same initial value condition $y(\alpha) = \beta$. Then $y_1(x) = y_2(x)$ on $I_1 \cap I_2$.

Proof. Suppose $I_1 \cap I_2 = [p, q]$ and let $z := \sup\{x \mid y_1 \equiv y_2 \text{ on } [\alpha, x]\}$. It suffices to show $z = q$. Now suppose on the contrary that $z < q$, and consider the subtraction $|y_1 - y_2|$:

$$y_i = \beta + \int_\alpha^x f(t, y_i)\,\mathrm{d}t \implies |y_1 - y_2| = \left| \int_z^x f(t, y_1) - f(t, y_2)\,\mathrm{d}t \right|.$$

Construct an interval $I^* = [z - \frac{1}{2L^*}, z + \frac{1}{2L^*}] \cap [p, q]$, and let $x_m = \arg\max_{x \in I^*} |y_1(x) - y_2(x)|$, which implies for $\forall x \in I^*$,

$$\begin{aligned}
|y_1(x) - y_2(x)| &= \left| \int_z^x f(t, y_1) - f(t, y_2)\,\mathrm{d}t \right| \\
&\le \int_z^x |f(t, y_1(t)) - f(t, y_2(t))|\,\mathrm{d}t \\
&\le L^* \int_z^x |y_1(x) - y_2(x)|\,\mathrm{d}t \\
&\le L^* |x - z| |y_1(x_m) - y_2(x_m)| \\
&\le \frac{1}{2} |y_1(x_m) - y_2(x_m)|.
\end{aligned}$$

Taking $x = x_m$, we imply $y_1 \equiv y_2$ for $\forall x \in I^*$, which contradicts the maximality of z. ∎

Combining Theorem (3.4) and proposition (3.8), we imply the existence of a unique "maximal" solution for the IVP (3.12), i.e., the unique solution is defined on a maximal interval.

Corollary 3.3 Let $U \subseteq \mathbb{R}^2$ be an open set such that $f(x, y)$ satisfies the Lipschitz condition for any $[a, b] \times [c, d] \subseteq U$, then there exists x_m and x_M in $\overline{\mathbb{R}}$ such that

- The IVP (3.12) admits a solution $y(x)$ for $x \in (x_m, x_M)$, and if y^* is another solution of (3.12) on some interval $I \subseteq (x_m, x_M)$, then $y \equiv y^*$ on I.
- Therefore $y(x)$ is maximally defined; and $y(x)$ is unique.

■ **Example 3.12** Consider the IVP

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x^2 y^{1/5} \\ y(0) = C \end{cases} \implies \frac{\partial f}{\partial y} = \frac{x^2}{5y^{4/5}}.$$

- Taking $U = \mathbb{R} \times (0, \infty)$ implies $y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}$, defined on $(\sqrt[3]{-15/4c^{4/5}}, \infty)$.
- When $c = 0$, $f(x, y)$ does not satisfy the Lipschitz condition. The uniqueness of solution does not hold.

■

3.6. Wednesday for MAT4002

3.6.1. Remarks on product space

Reviewing.

- Product Topology: For topological space $(X, \mathcal{T}_X)$ and $(Y, \mathcal{Y})$, define the basis

$$\mathcal{B}_{X\times Y} = \{U \times V \mid U \in \mathcal{T}_X, V \in \mathcal{T}_Y\}$$

  and the family of union of subsets in $\mathcal{B}_{X\times Y}$ forms a product topology.

Proposition 3.9 a ring torus is homeomorphic to the Cartesian product of two circles, say $S^1 \times S^1 \cong T$.

Proof. Define a mapping $f : [0,2\pi] \times [0,2\pi] \to T$ as

$$f(\theta,\phi) = \Big((R + r\cos\theta)\cos\phi, \quad (R + r\cos\theta)\sin\phi, \quad r\sin\theta\Big)$$

Define $i : T \to \mathbb{R}^3$, we imply

$$i \circ f : [0,2\pi] \times [0,2\pi] \to \mathbb{R}^3 \text{ is continuous}$$

Therefore we imply $f : [0,2\pi] \times [0,2\pi] \to T$ is continuous. Together with the condition that

$$\begin{cases} f(0,y) = f(2\pi, y) \\ f(x,0) = f(x,2\pi) \end{cases}$$

we imply the function $f : S^1 \times S^1 \to T$ is continuous. We can also show it is bijective. We can also show f^{-1} is continuous. ∎

Proposition 3.10 1. Let $X \times Y$ be endowed with product topology. The projection mappings defined as

$$p_X : X \times Y \to X, \text{ with } p_X(x, y) = x$$

$$p_Y : X \times Y \to Y, \text{ with } p_Y(x, y) = y$$

are continuous.

2. (an equivalent definition for product topology) The product topology is the **coarest topology** on $X \times Y$ such that p_X and p_Y are both continuous.
3. (an equivalent definition for product topology) Let Z be a topological space, then the product topology is the unique topology that the red and the blue line in the diagram commutes:

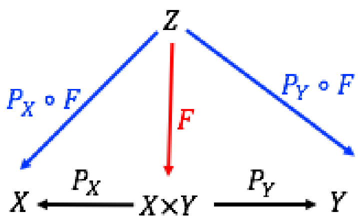


Figure 3.3: Diagram summarizing the statement (*)

namely,

the mapping $F : Z \to X \times Y$ is continuous iff both $P_X \circ F : Z \to X$ and $P_Y \circ F : Z \to Y$ are continuous. (*)

Proof. 1. For any open U, we imply $p_X^{-1}(U) = U \times Y \in \mathcal{B}_{X\times Y} \subseteq \mathcal{T}_{X\times Y}$, i.e., $p_X^{-1}(U)$ is open. The same goes for p_Y.

2. It suffices to show any topology $\mathcal{T}$ that meets the condition in (2) must contain $\mathcal{T}_{\text{product}}$. We imply that for $\forall U \in \mathcal{T}_X, V \in \mathcal{T}_Y$,

$$\begin{cases} p_X^{-1}(U) = U \times X \in \mathcal{T} \\ p_Y^{-1}(V) = X \times V \in \mathcal{T} \end{cases} \implies (U \times Y) \cap (X \times V) = (U \cap X) \times (Y \cap V) = U \times V \in \mathcal{T},$$

which implies $\mathcal{B}_{X\times Y} \subseteq \mathcal{T}$. Since $\mathcal{T}$ is closed for union operation on subsets, we imply $\mathcal{T}_{\text{product topology}} \subseteq \mathcal{T}$.

3. (a) Firstly show that $\mathcal{T}_{\text{product}}$ satisfies (*).

   - For the forward direction, by (1) we imply both $p_X \circ F$ and $p_Y \circ F$ are continuous, since the composition of continuous functions are continuous as well.
   - For the reverse direction, for $\forall U \times \mathcal{T}_X, V \in \mathcal{T}_Y$,

$$F^{-1}(U \times V) = (p_X \circ F)^{-1}(X) \cap (p_Y \circ F)^{-1}(Y),$$

     which is open due to the continuity of $p_X \circ F$ and $p_Y \circ F$.

   (b) Then we show the uniqueness of $\mathcal{T}_{\text{product}}$. Let $\mathcal{T}$ be another topology $X \times Y$ satisfying (*).

   - Take $Z = (X \times Y, \mathcal{T})$, and consider the identity mapping $F = \text{id} : Z \to Z$, which is continuous. Therefore $p_X \circ \text{id}$ and $p_Y \circ \text{id}$ are continuous, i.e., p_X and p_Y are continuous. By (2) we imply $\mathcal{T}_{\text{product}} \subseteq \mathcal{T}$.
   - Take $Z = (X \times Y, \mathcal{T}_{\text{product}})$, and consider the identity mapping $F = \text{id} : Z \to Z$. Note that $p_X \circ F = p_X$ and $p_Y \circ F = p_Y$, which is continuous by (1). Therefore, the identity mapping $F : (X \times Y, \mathcal{T}_{\text{product}}) \to (X \times Y, \mathcal{T})$ is continuous, which implies

$$U = \text{id}^{-1}(U) \subseteq \mathcal{T}_{\text{product}} \text{ for } \forall U \in \mathcal{T},$$

     i.e., $\mathcal{T} \subseteq \mathcal{T}_{\text{product}}$.

   The proof is complete.

■

Definition 3.6 [Disjoint Union] Let $X \times Y$ be two topological spaces, then the **disjoint union** of X and Y is

$$X \coprod Y := (X \times \{0\}) \cup (Y \times \{1\})$$

R

1. We define that U is open in $X \coprod Y$ if

   (a) $U \cap (X \times \{0\})$ is open in $X \times \{0\}$; and

   (b) $U \cap (Y \times \{1\})$ is open in $Y \times \{1\}$.

   We also need to show the well-definedness for this definition.

2. S is open in $X \coprod Y$ iff S can be expressed as

$$S = (U \times \{0\}) \cup (V \times \{1\})$$

   where $U \subseteq X$ is open and $V \subseteq Y$ is open.

3.6.2. Properties of Topological Spaces

3.6.2.1. Hausdorff Property

Definition 3.7 [First Separation Axiom] A topological space X satisfies the **first separation axiom** if for any two distinct points $x \neq y \in X$, there exists open $U \ni x$ but not including y.

Proposition 3.11 A topological space X has first separation property if and only if for $\forall x \in X$, $\{x\}$ is closed in X.

Proof. *Sufficiency.* Suppose that $x \neq y$, then construct $U := X \setminus \{y\}$, which is a open set that contains x but not includes y.

Necessity. Take any $x \in X$, then for $\forall y \neq x$, there exists $y \in U_y$ that is open and $x \notin U_y$. Thus

$$\{y\} \subseteq U_y \subseteq X \setminus \{x\}$$

which implies

$$\bigcup_{y \in X \setminus \{x\}} \{y\} \subseteq \bigcup_{y \in X \setminus \{x\}} U_y \subseteq X \setminus \{x\},$$

i.e., $X \setminus \{x\} = \bigcup_{y \in X \setminus \{x\}} U_y$ is open in X, i.e., $\{x\}$ is closed in X. ∎

Definition 3.8 [Second separation Axiom] A topological space satisfies the **second separation axiom** (or X is Hausdorff) if for all $x \neq y$ in X, there exists open sets U, V such that

$$x \in U, \quad y \in V, \quad U \cap V = \emptyset$$

■

■ **Example 3.13** All metrizable topological spaces are Hausdorff.

Suppose $d(x,y) = r > 0$, then take $B_{r/2}(x)$ and $B_{r/2}(y)$ ■

■ **Example 3.14** Note that a topological space that is **first separable** may not necessarily be **second separable**:

Consider $\mathcal{T}_{\text{co-finite}}$, then X is first separable but not Hausdorff:

Suppose on the contrary that for given $x \neq y$, there exists open sets U, V such that $x \in U, y \in V$, and

$$U \cap V = \emptyset \implies X = X \setminus (U \cap V) = (X \setminus U) \cup (X \setminus V),$$

implying that the union of two finite sets equals X, which is infinite, which is a contradiction.

■

Chapter 4

Week4

4.1. Monday for MAT3040

4.1.1. Quotient Spaces

Now we aim to divide a big **vector space** into many pieces of slices.

- For example, the Cartesian plane can be expressed as union of set of vertical lines as follows:

$$\mathbb{R}^2 = \bigcup_{m\in\mathbb{R}} \left\{ \begin{pmatrix} m \\ 0 \end{pmatrix} + \operatorname{span}\{(0,1)\} \right\}$$

- Another example is that the set of integers can be expressed as union of three sets:

$$\mathbb{Z} = Z_1 \cup Z_2 \cup Z_3,$$

where Z_i is the set of integers z such that $z \bmod 3 = i$.

Definition 4.1 [Coset] Let V be a vector space and $W \leq V$. For any element $\boldsymbol{v} \in V$, the **(right) coset** determined by $\boldsymbol{v}$ is the set

$$\boldsymbol{v} + W := \{\boldsymbol{v} + \boldsymbol{w} \mid \boldsymbol{w} \in W\}$$

For example, consider $V = \mathbb{R}^3$ and $W = \operatorname{span}\{(1,2,0)\}$. Then the coset determined by $\boldsymbol{v} = (5,6,-3)$ can be written as

$$\boldsymbol{v} + W = \{(5+t, 6+2t, -3) \mid t \in \mathbb{R}\}$$

It's interesting that the coset determined by $\boldsymbol{v}' = \{(4,4,-3)\}$ is exactly the same as the coset shown above:

$$\boldsymbol{v}' + W = \{(4+t, 4+2t, -3) \mid t \in \mathbb{R}\} = \boldsymbol{v} + W.$$

Therefore, write the exact expression of $\boldsymbol{v} + W$ may sometimes become tedious and hard to check the equivalence. We say $\boldsymbol{v}$ is a **representative** of a coset $\boldsymbol{v} + W$.

Proposition 4.1 Two cosets are the same iff the subtraction for the corresponding representatives is in W, i.e.,

$$\boldsymbol{v}_1 + W = \boldsymbol{v}_2 + W \iff \boldsymbol{v}_1 - \boldsymbol{v}_2 \in W$$

Proof. *Necessity.* Suppose that $\boldsymbol{v}_1 + W = \boldsymbol{v}_2 + W$, then $\boldsymbol{v}_1 + \boldsymbol{w}_1 = \boldsymbol{v}_2 + \boldsymbol{w}_2$ for some $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$, which implies

$$\boldsymbol{v}_1 - \boldsymbol{v}_2 = \boldsymbol{w}_2 - \boldsymbol{w}_1 \in W$$

Sufficiency. Suppose that $\boldsymbol{v}_1 - \boldsymbol{v}_2 = \boldsymbol{w} \in W$. It suffices to show $\boldsymbol{v}_1 + W \subseteq \boldsymbol{v}_2 + W$. For any $\boldsymbol{v}_1 + \boldsymbol{w}' \in \boldsymbol{v}_1 + W$, this element can be expressed as

$$\boldsymbol{v}_1 + \boldsymbol{w}' = (\boldsymbol{v}_2 + \boldsymbol{w}) + \boldsymbol{w}' = \boldsymbol{v}_2 + \underbrace{(\boldsymbol{w} + \boldsymbol{w}')}_{\text{belong to } W} \in \boldsymbol{v}_2 + W.$$

Therefore, $\boldsymbol{v}_1 + W \subseteq \boldsymbol{v}_2 + W$. Similarly we can show that $\boldsymbol{v}_2 + W \subseteq \boldsymbol{v}_1 + W$. ■

Exercise: Two cosets with representatives $\boldsymbol{v}_1, \boldsymbol{v}_2$ have no intersection iff $\boldsymbol{v}_1 - \boldsymbol{v}_2 \notin W$.

Definition 4.2 [Quotient Space] The **quotient space** of V by the subspace W, is the collection of all cosets $\boldsymbol{v} + W$, denoted by V/W.

To make the quotient space a vector space structure, we define the addition and scalar multiplication on V/W by:

$$(\boldsymbol{v}_1 + W) + (\boldsymbol{v}_2 + W) := (\boldsymbol{v}_1 + \boldsymbol{v}_2) + W$$
$$\alpha \cdot (\boldsymbol{v} + W) := (\alpha \cdot \boldsymbol{v}) + W$$

For example, consider $V = \mathbb{R}^2$ and $W = \text{span}\{(0,1)\}$. Then note that

$$\left(\begin{pmatrix}1\\0\end{pmatrix} + W\right) + \left(\begin{pmatrix}2\\0\end{pmatrix} + W\right) = \left(\begin{pmatrix}3\\0\end{pmatrix} + W\right)$$
$$\pi \cdot \left(\begin{pmatrix}1\\0\end{pmatrix} + W\right) = \left(\begin{pmatrix}\pi\\0\end{pmatrix} + W\right)$$

Proposition 4.2 The addition and scalar multiplication is well-defined.

Proof.

1. Suppose that
$$\begin{cases} \boldsymbol{v}_1 + W = \boldsymbol{v}_1' + W \\ \boldsymbol{v}_2 + W = \boldsymbol{v}_2' + W \end{cases}, \tag{4.1}$$
and we need to show that $(\boldsymbol{v}_1 + \boldsymbol{v}_2) + W = (\boldsymbol{v}_1' + \boldsymbol{v}_2') + W$.
From (4.1) and proposition (4.1), we imply
$$\boldsymbol{v}_1 - \boldsymbol{v}_1' \in W, \quad \boldsymbol{v}_2 - \boldsymbol{v}_2' \in W$$
which implies
$$(\boldsymbol{v}_1 - \boldsymbol{v}_1') + (\boldsymbol{v}_2 - \boldsymbol{v}_2') = (\boldsymbol{v}_1 + \boldsymbol{v}_2) - (\boldsymbol{v}_1' + \boldsymbol{v}_2') \in W$$
By proposition (4.1) again we imply $(\boldsymbol{v}_1 + \boldsymbol{v}_2) + W = (\boldsymbol{v}_1' + \boldsymbol{v}_2') + W$
2. For scalar multiplication, similarly, we can show that $\boldsymbol{v}_1 + W = \boldsymbol{v}_1' + W$ implies $\alpha\boldsymbol{v}_1 + W = \alpha\boldsymbol{v}_1' + W$ for all $\alpha \in \mathbb{F}$.

■

Proposition 4.3 The canonical projection mapping

$$\begin{aligned} \pi_W : V &\to V/W, \\ \boldsymbol{v} &\mapsto \boldsymbol{v} + W, \end{aligned}$$

is a **surjective linear transformation** with $\ker(\pi_W) = W$.

Proof. 1. First we show that $\ker(\pi_W) = W$:

$$\pi_W(\boldsymbol{v}) = 0 \implies \boldsymbol{v} + W = \mathbf{0}_{V/W} \implies \boldsymbol{v} + W = \mathbf{0} + W \implies \boldsymbol{v} = (\boldsymbol{v} - \mathbf{0}) \in W$$

Here note that the zero element in the quotient space V/W is the coset with representative $\mathbf{0}$.

2. For any $\boldsymbol{v}_0 + W \in V/W$, we can construct $\boldsymbol{v}_0 \in V$ such that $\pi_W(\boldsymbol{v}_0) = \boldsymbol{v}_0 + W$. Therefore the mapping π_W is surjective.
3. To show the mapping π_W is a linear transformation, note that

$$\begin{aligned} \pi_W(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) &= (\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) + W \\ &= (\alpha\boldsymbol{v}_1 + W) + (\beta\boldsymbol{v}_2 + W) \\ &= \alpha(\boldsymbol{v}_1 + W) + \beta(\boldsymbol{v}_2 + W) \\ &= \alpha\pi_W(\boldsymbol{v}_1) + \beta\pi_W(\boldsymbol{v}_2) \end{aligned}$$

■

4.1.2. First Isomorphism Theorem

The key of linear algebra is to solve the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. The general step for solving this linear system is as follows:

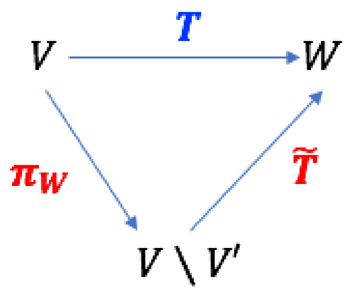
1. Find the solution set for $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$, i.e., the set $\ker(\boldsymbol{A})$
2. Find a particular solution $\boldsymbol{x}_0$ such that $\boldsymbol{A}\boldsymbol{x}_0 = \boldsymbol{b}$.

Then the general solution set to this linear system is $\boldsymbol{x}_0 + \ker(\boldsymbol{A})$, which is a coset in the space $\mathbb{R}^n/\ker(\boldsymbol{A})$. Therefore, to solve the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ suffices to study the quotient space $\mathbb{R}^n/\ker(\boldsymbol{A})$:

Proposition 4.4 — Universal Property I. Suppose that $T : V \to W$ is a linear transformation, and that $V' \leq \ker(T)$. Then the mapping

$$\begin{aligned} \tilde{T} : V/V' &\to W \\ \boldsymbol{v} + V' &\mapsto T(\boldsymbol{v}) \end{aligned}$$

is a well-defined linear transformation. As a result, the diagram below commutes:



In other words, we have $T = \tilde{T} \circ \pi_W$.

Proof. First we show the well-definedness. Suppose that $\boldsymbol{v}_1 + V' = \boldsymbol{v}_2 + V'$ and suffices to show $\tilde{T}(\boldsymbol{v}_1 + V') = \tilde{T}(\boldsymbol{v}_2 + V')$, i.e., $T(\boldsymbol{v}_1) = T(\boldsymbol{v}_2)$. By proposition (4.1), we imply

$$\boldsymbol{v}_1 - \boldsymbol{v}_2 \in V' \leq \ker(T) \implies T(\boldsymbol{v}_1 - \boldsymbol{v}_2) = \boldsymbol{0} \implies T(\boldsymbol{v}_1) - T(\boldsymbol{v}_2) = \boldsymbol{0}.$$

Then we show $\tilde{(T)}$ is a linear transformation:

$$\begin{aligned} \tilde{T}(\alpha(\boldsymbol{v}_1 + V') + \beta(\boldsymbol{v}_2 + V')) &= \tilde{T}((\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) + V') \\ &= T(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) \\ &= \alpha T(\boldsymbol{v}_1) + \beta T(\boldsymbol{v}_2) \\ &= \alpha\tilde{T}(\boldsymbol{v}_1 + V') + \beta\tilde{T}(\boldsymbol{v}_2 + V') \end{aligned}$$

■

Actually, if we let $V' = \ker(T)$, the mapping $\tilde{T} : V/V' \to T(V)$ forms an isomorphism, In particular, if further T is surjective, then $T(V) = W$, i.e., the mapping $\tilde{T} : V/V' \to W$ forms an isomorphism.

Theorem 4.1 — **First Isomorphism Theorem.** Let $T : V \to W$ be a surjective linear transformation. Then the mapping

$$\begin{aligned} \tilde{T} : V/\ker(T) &\to W \\ \boldsymbol{v} + \ker(T) &\mapsto T(\boldsymbol{v}) \end{aligned}$$

is an isomorphism.

Proof. *Injectivity.* Suppose that $\tilde{T}(\boldsymbol{v}_1 + \ker(T)) = \tilde{T}(\boldsymbol{v}_2 + \ker(T))$, then we imply

$$T(\boldsymbol{v}_1) = T(\boldsymbol{v}_2) \implies T(\boldsymbol{v}_1 - \boldsymbol{v}_2) = \mathbf{0}_W \implies \boldsymbol{v}_1 - \boldsymbol{v}_2 \in \ker(T),$$

i.e., $\boldsymbol{v}_1 + \ker(T) = \boldsymbol{v}_2 + \ker(T)$.

Surjectivity. For $\boldsymbol{w} \in W$, due to the surjectivity of T, we can find a $\boldsymbol{v}_0$ such that $T(\boldsymbol{v}_0) = \boldsymbol{w}$. Therefore, we can construct a set $\boldsymbol{v}_0 + \ker(T)$ such that

$$\tilde{T}(\boldsymbol{v}_0 + \ker(T)) = \boldsymbol{w}.$$

■

4.2. Monday for MAT3006

Our first quiz will be held on next Wednesday.

Reviewing.

- Picard Lindelof Theorem on ODEs. e.g., consider

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{x}{1-y}, \ (x,y) \in G := (-\infty,\infty) \times (-\infty,1) \\ y(0) = 2 \end{cases}$$

  Since $f \in C^1(G)$ satisfies the Lipschitz condition on some closed ball of the point (x_0, y_0), the setting for Picard Lindelof Theorem is satisfied, and the solution is uniquely given by:

$$y = 1 + \sqrt{1-x^2}, \ -1 < x < 1.$$

  Therefore, the maximal interval of existence is given by $(-1,1)$. In order to restrict G to be open to construct a closed ball of (x_0, y_0), we need the initial condition $y(0) \neq 1$.

4.2.1. Generalization into System of ODEs

Formal Setting of System of ODEs. Consider the system of ODEs

$$\begin{cases} y_1'(x) = f_1(x, y_1(x), \dots, y_n(x)) \\ \quad \vdots \\ y_n'(x) = f_n(x, y_1(x), \dots, y_n(x)) \end{cases} \qquad \begin{cases} y_1(\alpha) = \beta_1 \\ \quad \vdots \\ y_n(x) = \beta_n \end{cases}$$

It's convenient to denote

$$\boldsymbol{y}(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix} \in C(\mathbb{R}, \mathbb{R}^n), \quad \boldsymbol{f}(x, \boldsymbol{y}) = \begin{pmatrix} f_1(x, \boldsymbol{y}) \\ \vdots \\ f_n(x, \boldsymbol{y}) \end{pmatrix}, \quad \boldsymbol{\beta} := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

Here the notation $C(X,Y)$ denotes the set of bounded continuous mapping from X to Y. Therefore we can express the system of ODE as a compact form:

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{f}(x,\boldsymbol{y}) \\ \boldsymbol{y}(\alpha) = \boldsymbol{\beta} \end{cases}$$

Generalization of Picard Lindelof Theorem. Consider the rectangle

$$S = \{(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}\times\mathbb{R}^n \mid \alpha - a \le x \le \alpha + a, \beta_i - b_i \le y_i \le \beta_i + b_i, i = 1,\dots,n\}$$

Suppose that

- $\|\boldsymbol{f}(x,\boldsymbol{y})\| \le M, \forall (x,\boldsymbol{y}) \in S$

- $\|\boldsymbol{f}(x,\boldsymbol{y}) - \boldsymbol{f}(x,\boldsymbol{y}')\| \le L \cdot \|\boldsymbol{y} - \boldsymbol{y}'\|$ for $\forall x \in [\alpha - a, \alpha + a]$

Then consider the complete metric space

$$X = \{\boldsymbol{y} \in C([\alpha - a, \alpha + a], \mathbb{R}^n) \mid \beta_i - b_i \le y_i(x) \le \beta_i + b_i\}$$

(Verification of completeness: if Y is complete, then $C(X,Y)$ is complete.) Under this setting, the similar argument gives the Picard-Lindelof for system of ODEs.

Higher Order ODEs. Note that there is a standard way to transform the ODE with higher order derivatives into a system of first order ODEs. Suppose we want to solve the initival value problem

$$\begin{cases} y^{(m)} = f(x,y,y',\dots,y^{(m-1)}) \\ y(\alpha) = \beta_0,\ y'(\alpha) = \beta_1,\dots,y^{(m-1)}(\alpha) = \beta_{m-1} \end{cases}$$

We can define the variables

$$\begin{pmatrix} y_{m-1}(x) \\ \vdots \\ y_1(x) \\ y_0(x) \end{pmatrix} = \begin{pmatrix} y^{(m-1)}(x) \\ \vdots \\ y'(x) \\ y(x) \end{pmatrix}$$

which gives an equivalent system of ODE:

$$\begin{cases} y'_{m-1} = f(x, y_0, \dots, y_{m-1}) \\ y'_{m-2} = y_{m-1} \\ \quad\vdots \\ \quad y'_0 = y_1 \end{cases}, \text{ with } \begin{cases} y_{m-1}(\alpha) = \beta_{m-1} \\ y_{m-2}(\alpha) = \beta_{m-2} \\ \quad\vdots \\ \quad y_0(\alpha) = \beta_0 \end{cases}$$

4.2.2. Stone-Weierstrass Theorem

Under the compact metric space X, the goal is to approximate **any** functions in $C(X)$. For example, under $X = [a, b]$, one can apply Taylor polynomials $p_n(x)$ to approximate differentiable functions:

$$\|f(x) - p_n(x)\|_\infty < \varepsilon, \text{ for large } n.$$

To formally describe the phenomenon for the approximation of **any** functions in $C(X)$, we need to describe the set of approximate functions, which usually obtains a common property:

Definition 4.3 [Algebra] A subset $\mathcal{A} \subseteq C(X)$ (where X is a general space) is an **algebra** if the following holds:

- If $f_1, f_2 \in \mathcal{A}$, then $\alpha f_1 + \beta f_2 \in \mathcal{A}$
- If $f_1, f_2 \in \mathcal{A}$, then $f_1 \cdot f_2 \in \mathcal{A}$

■ **Example 4.1** 1. $\mathcal{A} = C(X)$ is an algebra.

2. $X = [a,b]$, then $\mathcal{A} = P[a,b] = \{\text{All polynomials } p(x)\}$ is an algebra. ■

The goal is to approximate any $f \in C(X)$ by $p \in \mathcal{A}$, i.e., for $\forall f \in C(X)$, there exists $p \in \mathcal{A}$ such that

$$\|f - p\|_\infty < \varepsilon, \ \forall \varepsilon > 0.$$

In other words, we aim to find an algebra $\mathcal{A} \subseteq C(X)$ such that $\overline{\mathcal{A}} = C(X)$, i.e., $\mathcal{A}$ is dense in $C(X)$.

Theorem 4.2 — **Weierstrass Approximation.** $\mathcal{P}[a,b]$ is dense in $C[a,b]$.

Proof. Consider any function $f \in C[0,1]$. By rescaling, assume that $f \in C[0,1]$. By subtracting a linear function $\ell(x)$, assume that $f(0) = f(1) = 0$. Then we extend $f(x)$ into $\mathbb{R}$ by setting $f(x) = 0, \forall x \notin [0,1]$.

- **Step 1: Construction of approximate function**: Consider the **Landaus kernel function**

$$Q_n(x) = \begin{cases} c_n \cdot (1-x^2)^n, & -1 \le x \le 1 \\ 0, & |x| > 1 \end{cases}$$

where c_n is chosen such that $\int Q_n(x)\,\mathrm{d}x = 1$. Then construct the approximation of f by defining

$$p_n(x) := Q_n * f = \int_{-1}^{1} f(x+t) Q_n(t)\,\mathrm{d}t$$

The intuition behind this construction is that as $n \to \infty$, $Q_n(x) \to \delta(x)$, where

$$\delta(x) = \begin{cases} \infty, & x = 0 \\ 0, & x \neq 0 \end{cases} \implies \int_{-1}^{1} f(x+t)\delta(t)\,\mathrm{d}t = f(x).$$

Step 2: Argue that $p_n(x) \in \mathcal{P}[a,b]$: Now it's clear that

$$\begin{aligned}
p_n(x) &= \int_{-1}^{1} f(x+t)Q_n(t)\,\mathrm{d}t &\quad (4.2a)\\
&= \int_{-x}^{1-x} f(x+t)Q_n(t)\,\mathrm{d}t &\quad (4.2b)\\
&= \int_{-1}^{1} f(u)\cdot Q_n(u-x)\,\mathrm{d}u &\quad (4.2c)\\
&= \int_{-1}^{1} f(u)\cdot (1-(u-x)^2)^n\,\mathrm{d}u, &\quad (4.2d)
\end{aligned}$$

where (4.2b) is because that $f = 0$, for $x \notin [0,1]$ and $Q_n = 0$ for $|x| > 1$; (4.2c) is by change of variables; and (4.2d) is by substitution of $Q_n(x)$. Therefore, p_n is still a polynomial of x.

- **Step 3: Construct an upper bound on c_n:** It's clear that

$$\begin{aligned}
c_n^{-1} &= \int_{-1}^{1} (1-x^2)^n\,\mathrm{d}x\\
&= 2\int_{0}^{1} (1-x^2)^n\,\mathrm{d}x\\
&\ge 2\int_{0}^{1} (1-nx^2)\,\mathrm{d}x\\
&\ge 2\int_{0}^{1/\sqrt{n}} (1-nx^2)\,\mathrm{d}x\\
&= 2\left(\frac{1}{\sqrt{n}} - \frac{1}{3\sqrt{n}}\right) > \frac{1}{\sqrt{n}}
\end{aligned}$$

and therefore $c_n < \sqrt{n}$. As a result, for any fixed $\delta \in (0,1)$, we imply

$$Q_n(x) \le \sqrt{n}(1-\delta^2)^n, \qquad \forall x \in [\delta, 1],$$

which implies $Q_n(x) \to 0$ uniformly on $[\delta, 1]$.

- **Step 4: Show that $\|p_n - f\|_\infty \to 0$.** Since f is continuous, for given $\varepsilon > 0$, there exists $\delta \in (0,1)$ such that

$$|f(x) - f(y)| < \varepsilon, \quad \text{when } |x-y| < \delta, x, y \in [0,1].$$

Therefore, for any $x \in [0,1]$, and for sufficiently large n,

$$
\begin{aligned}
|p_n(x) - f(x)| &= \left| \int_{-1}^{1} f(x+t) Q_n(t) - \int_{-1}^{1} f(x) Q_n(t)\,\mathrm{d}t \right| && \text{(4.3a)} \\
&\le \int_{-1}^{1} |f(x+t) - f(x)| Q_n(t)\,\mathrm{d}t && \text{(4.3b)} \\
&\le 2M \int_{-1}^{-\delta} Q_n(t)\,\mathrm{d}t + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} Q_n(t)\,\mathrm{d}t + 2M \int_{\delta}^{1} Q_n(t)\,\mathrm{d}t && \text{(4.3c)} \\
&\le 4M\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2} && \text{(4.3d)} \\
&\le \varepsilon && \text{(4.3e)}
\end{aligned}
$$

where (4.3c) is by separating the integrand into three parts, and then upper bounding $|f(x+t) - f(x)|$ by $2M := 2\sup_x |f(x)|$ for the integrand $t \in [-1, \delta) \cup (\delta, 1]$, and upper bounding $|f(x+t) - f(x)|$ by $\frac{\varepsilon}{2}$ due to the continuity of f for the integrand $t \in [\delta, \delta]$; (4.3e) is by choosing n sufficiently enough to make $4M\sqrt{n}(1-\delta^2)^n$ sufficiently small.

Therefore $\|p_n - f\|_\infty = \max_{x \in [0,1]} |p_n(x) - f(x)| < \varepsilon$ for large n. The proof is complete.

■

4.3. Monday for MAT4002

There will be a quiz next Monday. The scope is everything before CNY holiday. There will be one question with four parts for 40 minutes.

4.3.1. Hausdorffness

Reviewing. A topological space $(X,\mathcal{T})$ is said to be **Hausdorff** (or satisfy the second separtion property), if given any distinct points $x,y\in X$, there exist disjoint open sets U,V such that $U\ni x$ and $V\ni y$.

Proposition 4.5 If the topological space $(X,\mathcal{T})$ is Hausdorff, then all sequences $\{x_n\}$ in X has at most one limit.

Proof. Suppose on the contrary that

$$\{x_n\}\to a,\quad \{x_n\}\to b,\text{ with }a\neq b$$

By separation property, there exists $U,V\in\mathcal{T}$ and $U\cap V=\emptyset$ such that $U\ni a$ and $V\ni b$.

By tje openness of U, there exists N such that $\{x_N,x_{N+1},\dots\}\subseteq U$, since $\{x_n\}\to a\in U$. Similarly, there exists M such that $\{x_M,x_{M+1},\dots\}\subseteq V$. Take $K=\max\{M,N\}+1$, then $\emptyset\neq U\cap V\ni x_K$, which is a contradiction. ■

Proposition 4.6 Let X,Y be Hausdorff spaces. Then $X\times Y$ is Hausdorff with product topology.

Proof. Suppose that $(x_1,y_1)\neq(x_2,y_2)$ in $X\times Y$. Then $x_1\neq x_2$ or $y_1\neq y_2$. w.l.o.g., assume that $x_1\neq x_2$, then there exists U,V open in X such that $x_1\in U,x_2\in V$ with $U\cap V=\emptyset$.

Therefore, we imply $(U\times Y),(V\times Y)\in\mathcal{T}_{X\times Y}$, and

$$(U\times Y)\cap(V\times Y)=(U\cap V)\cap Y=\emptyset$$

with $(x_1,y_1)\in U\times Y,(x_2,y_2)\in V\times Y$, i.e., $X\times Y$ is Hausdorff with product topology. ■

The same argument applies if the second separation property is replaced by first separation property.

Proposition 4.7 If $f : X \to Y$ is an injective continuous mapping, then Y is Hausdorff implies X is Hausdorff.

Proof. Suppose that Y satisfies the second separation property. For given $a \neq b$ in X, we imply $f(a) \neq f(b)$ in Y. Therefore, there exists $U \ni f(a), V \ni f(b)$ with $U \cap V = \emptyset$. It follows that

$$a \in f^{-1}(U), b \in f^{-1}(V), \quad f^{-1}(U) \cap f^{-1}(V) = f^{-1}(U \cap V) = \emptyset,$$

i.e., X is Hausdorff. ∎

Corollary 4.1 If $f : X \to Y$ is homeomorphic, then X is Hausdorff iff Y is Hausdorff, i.e., Hausdorffness is a topological property (i.e., a property that is preserved under homeomorphism).

4.3.2. Connectedness

Definition 4.4 [Connected] The topological space $(X, \mathcal{T})$ is **disconnected** if there are open $U, V \in \mathcal{T}$ such that

$$U \neq \emptyset, V \neq \emptyset, \quad U \cap V = \emptyset, \quad U \cup V = X. \tag{4.4}$$

If no such $U, V \in \mathcal{T}$ exist, then X is **connected**.

Proposition 4.8 Let $(X, \mathcal{T})$ be topological spaces. TFAE (i.e., the followings are equivalent):

1. X is connected
2. The **only** subset of X which are both open and closed are $\emptyset$ and X
3. Any continuous function $f : X \to \{0,1\}$ ($\{0,1\}$ is equipped with discrete topology) is a constant function.

Proof. (1) implies (2): Suppose that $U \subseteq X$ is both open and closed. Then $U, X \setminus U$ are both open and disjoint, and $U \cup (X \setminus U) = X$. By connnectedness, either $U = \emptyset$ or $X \setminus U = \emptyset$. Therefore, $U = \emptyset$ or X.

(2) implies (3): Note that $U = f^{-1}(\{0\})$ and $V = f^{-1}(\{1\})$ are open disjoint sets in X satisfying $U \cup V = X$. By the connectedness of X, either $(U, V) = (X, \emptyset)$ or $(V, U) = (\emptyset, X)$. In either case, we imply f is a constant function.

(3) implies (2): Suppose that $U \subseteq X$ is both open and closed. Construct the mapping

$$f(x) = \begin{cases} 0, & x \in U \\ 1, & x \in X \setminus U \end{cases}$$

It's clear that f is continuous, and therefore $f(x) = 0$ or 1. Therefore $U = \emptyset$ or X.

(2) implies (1): Suppose on the contrary that there exists open U, V such that (4.4) holds. By (4.4), we imply $U = X \setminus V$ is closed as well. Since $U \neq \emptyset$ and $U = \emptyset$ or X, we imply $U = X$, which implies $V = \emptyset$, which is a contradiction. ■

Corollary 4.2 The interval $[a, b] \subseteq \mathbb{R}$ is connnected

Proof. Suppose on the contrary that there exists continuous function $f : [a, b] \to \{0, 1\}$ that takes 2 values. Construct the mapping $\tilde{f} : [a, b] \to \mathbb{R}$

$$\begin{aligned} \tilde{f} : [a, b] &\xrightarrow{f} \{0, 1\} \xrightarrow{i} \mathbb{R}, \\ \text{with} \quad \tilde{f} &= i \circ f. \end{aligned}$$

Note that $\{0, 1\} \subseteq \mathbb{R}$ denotes the subspace topology, we imply the inclusion mapping $i : \{0, 1\} \to \mathbb{R}$ with $s \mapsto s$ is continuous. The composition of continuous mappings is continuous as well, i.e., $\tilde{f}$ is continuous.

Since the function f can take two values, there exists $p, q \in [a, b]$ such that $\tilde{f}(p) = i \circ f(p) = 0$ and $\tilde{f}(q) = i \circ f(q) = 1$. By intermediate value theorem, there exists $r \in [a, b]$ such that $\tilde{f}(r) = i \circ f(r) = 1/2$, which implies $f(r) = \frac{1}{2}$, which is a contradiction. ■

Definition 4.5 [Connected subset] A non-empty subset $S \subseteq X$ is **connected** if S with the subspace topology is connected

Equivalently, $S \subseteq X$ is connected if, whenever U, V are open in X such that $S \subseteq U \cup V$, and $(U \cap V) \cap S = \emptyset$, one can imply either $U \cap S = \emptyset$ or $V \cap S = \emptyset$.

Proposition 4.9 If $f : X \to Y$ is continuous mapping, and the subset $A \subseteq X$ is connected, then $f(A)$ is connected. In other words, the continuous image of a connected set is connected.

Proof. Suppose that $U, V \subseteq Y$ is open such that

$$f(A) \subseteq U \cup V, \quad (U \cap V) \cap f(A) = \emptyset.$$

Therefore we imply

$$A \subseteq f^{-1}(U) \cup f^{-1}(V), \quad (f^{-1}(U) \cap A) \cap (f^{-1}(V) \cap A) = \emptyset$$

By connectedness of A, either $f^{-1}(U) \cap A = \emptyset$ or $f^{-1}(V) \cap A = \emptyset$. Therefore, $f(A) \cap U = \emptyset$ or $f(A) \cap V = \emptyset$, i.e., $f(A)$ is connected. ∎

Proposition 4.10 If $\{A_i\}_{i \in I}$ are connected and $A_i \cap A_j \neq \emptyset$ for $\forall i, j \in I$, then the set $\bigcup_{i \in I} A_i$ is connected.

Proof. Suppose the function $f : \cup_{i \in I} A_i \to \{0,1\}$ is a continuous map. Then we imply that its restriction $f|_{A_i} = f \circ i : A_i \to \{0,1\}$ is continuous for all $i \in I$. Thus $f|_{A_i}$ is a constant for all $i \in I$. Due to the non-empty intersection of A_i, A_j for $\forall i, j \in I$, we imply f is constant. ∎

Proposition 4.11 If X, Y are connnected, then $X \times Y$ is connected using product topology.

Proof. It's clear that $X \times \{y_0\}$ is connected in $X \times Y$ for fixed y_0; and $\{x_0\} \times Y$ is connected for fixed x_0.

Therefore, for fixed $y_0 \in Y$, construct $B = X \times \{y_0\}$ and $C_x = \{x\} \times Y$, which follows that

$$B \cap C_x = \{(x, y_0)\} \neq \emptyset, \forall x \in X \implies B \cup \left\{ \bigcup_{x \in X} C_x \right\} = X \times Y \text{ is connected.}$$

■

Definition 4.6 [Path Connectes] Let $(X, \mathcal{T})$ be a topological space.

1. A path connecting 2 points $x, y \in X$ is a continuous function $\tau : [0,1] \to X$ with $\tau(0) = x, \tau(1) = y$.
2. X is path-connected if any 2 points in X can be connected by a path.
3. The set $A \subseteq X$ is path-connected, if A sastisfies the condition using subspace topology.

   Or equivalently, A is path-connected if for any 2 points in X, there exists a continuous $t : [0,1] \to X$ with $t(x) \in A$ for any x, connecting the 2 points.

4.4. Wednesday for MAT3040

Reviewing.

- Quotient Space:
$$V/W = \{\boldsymbol{v} + W \mid \boldsymbol{v} \in V\}$$
The elements in V/W are cosets. Note that V/W does not mean a subset of V.

- Define the canonical projection mapping
$$\begin{aligned} \pi_W : V &\to V/W, \\ \text{with} \quad \boldsymbol{v} &\mapsto \boldsymbol{v} + W, \end{aligned}$$
then we imply π_W is a surjective linear transformation with $\ker(\pi_W) = W$.

  If $\dim(V) < \infty$, then by Rank-Nullity Theorem (2.3), we imply that
$$\dim(V) = \dim(W) + \dim(V/W),$$
i.e., $\dim(V/W) = \dim(V) - \dim(W)$.

- **(Universal Property I)** Every linear transformation $T : V \to W$ with $V' \leq \ker(T)$ can be descended to the composition of the canonical projection mapping $\pi_{V'}$ and the mapping
$$\begin{aligned} \tilde{T} : V/V' &\to W \\ \text{with} \quad \boldsymbol{v} + V' &\mapsto T(\boldsymbol{v}). \end{aligned}$$
In other words, the diagram (2.1) commutes:

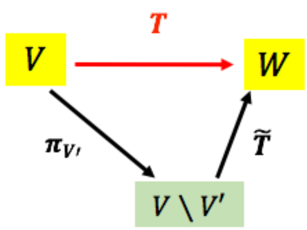


Diagram (2.1)

In other words, the mapping starting from either the black or red line gives the same result, i.e., $T(\boldsymbol{v}) = \tilde{T} \circ \pi_{V'}(\boldsymbol{v}) = \tilde{T}(\boldsymbol{v} + V')$ for any $\boldsymbol{v} \in V$.

- **(First Isomorphism Theorem)** Under the setting of Universal Property I (UPI), if T is a surjective linear transformation with $V' = \ker(T)$, then the $\tilde{T}$ is an isomorphism.

■ **Example 4.2** Suppose that $U, W \le V$ with $U \cap W = \{\mathbf{0}\}$, then define the mapping

$$\begin{aligned} &\phi : U \oplus W \to U \\ \text{with} \quad &\phi(\boldsymbol{u} + \boldsymbol{w}) = \boldsymbol{u} \end{aligned}$$

(R) Exercise: if $U, W \le V$ but $U \cap W \neq \{\mathbf{0}\}$, then the mapping

$$\begin{aligned} &\phi : U + W \to U \\ \text{with} \quad &\boldsymbol{u} + \boldsymbol{w} \mapsto \boldsymbol{u} \end{aligned} \qquad \text{is not well-defined:}$$

Suppose that $\mathbf{0} \neq \boldsymbol{v} \in U \cap W$ and for any $\boldsymbol{u} \in U, \boldsymbol{w} \in W$, we construct

$$\boldsymbol{u}' = \boldsymbol{u} - \boldsymbol{v} \in U, \quad \boldsymbol{w}' = \boldsymbol{w} + \boldsymbol{v} \in V \quad \Longrightarrow \phi(\boldsymbol{u}' + \boldsymbol{w}') = \boldsymbol{u} - \boldsymbol{v}$$

Therefore we get $\boldsymbol{u} + \boldsymbol{w} = \boldsymbol{u}' + \boldsymbol{w}'$ but $\phi(\boldsymbol{u} + \boldsymbol{w}) \neq \phi(\boldsymbol{u}' + \boldsymbol{w}')$.

Back to the situation $U \cap W = \{\mathbf{0}\}$, then it's clear that $\phi : U \oplus W \to U$ is surjective linear transformation with $\ker(\phi) = W$. Therefore, construct the new mapping

$$\begin{aligned} &\tilde{\phi} : U \oplus W / W \to U \\ \text{with} \quad &\boldsymbol{u} + \boldsymbol{w} + W \mapsto \phi(\boldsymbol{u} + \boldsymbol{w}) \end{aligned}$$

We imply $\tilde{\phi}$ is an isomorphism by First Isomorphism Theorem. ■

Now we study the generalized quotients, which is defined to satisfy the generalized version of universal property I.

Definition 4.7 [Universal Property for Quotients] Let V be a vector space and $V' \leq V$. Consider the collection of linear transformations

$$\text{Obj} = \left\{ T : V \to W \;\middle|\; \begin{array}{l} T \text{ is a linear transformation} \\ V' \leq \ker(T) \end{array} \right\}$$

(For example, $\pi_{V'} : V \to V/V'$ is an element from the set Obj.)

An element $(\phi : V \to U) \in \text{Obj}$ is said to satisfy the **universal property** if it satisfies the following:

> Given any element $(T : V \to W) \in \text{Obj}$, we can extend the transformation ϕ with a **uniquely existing** $\tilde{T} : U \to W$ so that the diagram (2.2) commutes:

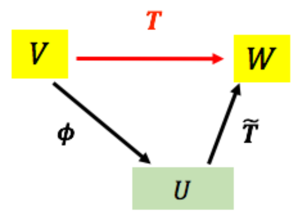


Diagram (2.2)

> Or equivalently, for given $(T : V \to W) \in \text{Obj}$, there exists the **unique** mapping $\tilde{T} : U \to W$ such that $T = \tilde{T} \circ \phi$.

Theorem 4.3 — **Universal Property II.**
1. The mapping $(\pi_{V'} : V \to V/V') \in \text{Obj}$ is a universal object, i.e., it satisfies the universal property.
2. If $(\phi : V \to U)$ is a universal object, then $U \cong V/V'$, i.e., there is intrinsically "one" element in the set of universal objects.

Proof. 1. Consider any linear transformation $T : V \to W$ such that $V' \leq \ker(T)$, then define (construct) the same $\tilde{T} : V/V' \to W$ as that in UPI. Therefore, for given T, applying the result of UPI, we imply $T = \tilde{T} \circ \pi_{V'}$, i.e., $\pi_{V'}$ satisfies the diagram (2.2).

To show the uniqueness of $\tilde{T}$, suppose there exists $\tilde{S} : V/V' \to W$ such that the diagram (2.3) commutes.

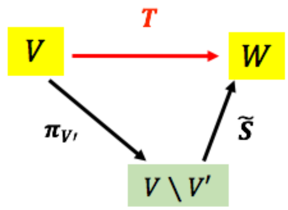


Diagram (2.3)

It suffices to show the mapping $\tilde{S} = \tilde{T}$: for any $\boldsymbol{v} + V' \in V/V'$, we have

$$\tilde{S}(\boldsymbol{v} + V') := \tilde{S} \circ \pi_{V'}(\boldsymbol{v}) = T(\boldsymbol{v}),$$

where the first equality is due to the surjectivity of $\pi_{V'}$. By the result of UPI, $T(\boldsymbol{v}) = \tilde{T}(\boldsymbol{v} + V')$. Therefore $\tilde{T}(\boldsymbol{v} + V') = \tilde{S}(\boldsymbol{v} + V')$ for all $\boldsymbol{v} + V' \in V/V'$. The proof is complete.

2. Suppose that $(\phi : V \to U)$ satisfies the universal property. In particular, the following two diagrams hold:

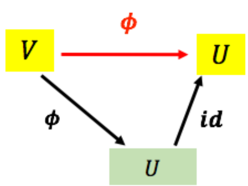


Diagram (2.4)

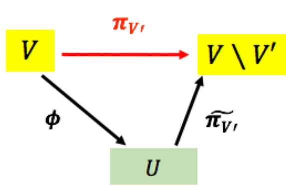


Diagram (2.5)

Since $(\pi_{V'})$ satisfies the universal property, in particular, the following two diagrams hold:

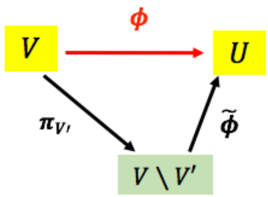


Diagram (2.6)

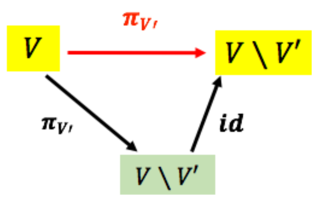


Diagram (2.7)

Then we claim that: Combining Diagram (2.5) and (2.6), we imply the diagram (2.8):

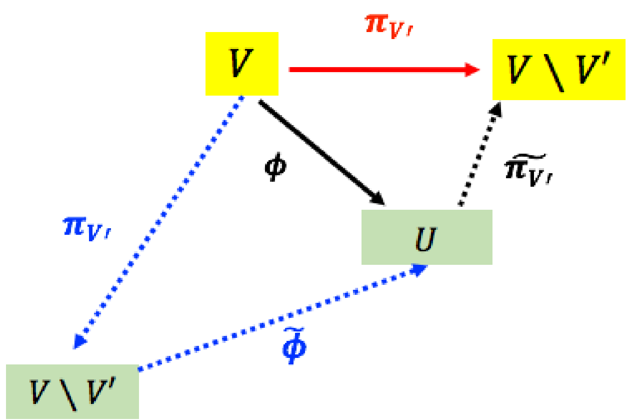


Diagram (2.8)

Graph Description: Note that this diagram commutes, i.e., the mapping starting from either the red line or the dash line gives the same result, i.e., $\pi_{V'} = \tilde{\pi}_{V'} \circ \tilde{\phi} \circ \pi_{V'}$. Comparing Diagram (2.7) and Diagram (2.8), we have $\tilde{\pi}_{V'} \circ \tilde{\phi} = id$, by the **uniqueness** of the universal object.

Therefore, $\tilde{\pi}_{V'} \circ \tilde{\phi} = id$ implies $\tilde{\pi}_{V'}$ is surjective and $\tilde{\phi}$ is injective.

Also, combining Diagram (2.6) and (2.5), we imply diagram (2.9):

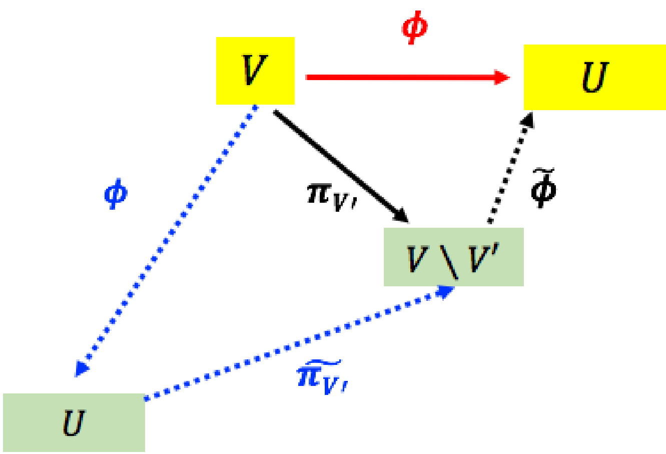


Diagram (2.9)

Graph Description: Note that this diagram commutes, i.e., the mapping starting from either the red line or the dash line gives the same result, i.e., $\phi = \tilde{\phi} \circ \tilde{\pi}_{V'} \circ \phi$. Comparing Diagram (2.9) and Diagram (2.4), we have $\tilde{\phi} \circ \tilde{\pi}_{V'} = id$, by the **uniqueness** of the universal object

Therefore, $\tilde{\phi} \circ \tilde{\pi}_{V'} = id$ implies $\tilde{\phi}$ is surjective and $\tilde{\pi}_{V'}$ is injective.

Therefore, both $\tilde{\phi} : U \to V/V'$ and $\tilde{\pi}_{V'} : V/V' \to U$ are bijective, i.e., $U \cong V/V'$.

The proof is complete.

■

4.4.1. Dual Space

Definition 4.8 Let V be a vector space over a field $\mathbb{F}$. The **dual vector space** V^* is defined as

$$\begin{aligned} V^* &= \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F}) \\ &= \{f : V \to \mathbb{F} \mid f \text{ is a linear transformation}\} \end{aligned}$$

■ **Example 4.3** 1. Consider $V = \mathbb{R}^n$ and define $\phi_i : V \to \mathbb{R}$ as the i-th component of input:

$$\phi_i \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_i,$$

Then we imply $\phi_i \in V^*$. On the contrary, $\phi_i^2 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_i^2$ is not in V^*

2. Consider $V = \mathbb{F}[x]$ and define $\phi : V \to \mathbb{F}$ as:

$$\phi(p(x)) = p(1),$$

It's clear that $\phi \in V^*$:

$$\begin{aligned} \phi(ap(x) + bq(x)) &= ap(1) + bq(1) \\ &= a\phi(p(x)) + b\phi(q(x)) \end{aligned}$$

3. Also, $\psi : V \to \mathbb{F}$ by $\psi(p(x)) = \int_0^1 p(x)\,\mathrm{d}x$ is in V^*.
4. Also, for $V = M_{n\times n}(\mathbb{F})$, the mapping $\mathrm{tr} : V \to \mathbb{F}$ by $\mathrm{tr}(M) = \sum_{i=1}^n M_{ii}$ is in V^*. However, the $\det : V \to \mathbb{F}$ is not in V^*

■

Definition 4.9 Let V be a vector space, with basis $B = \{v_i \mid i \in I\}$ (I can be finite or countable, or uncountable). Define

$$B^* = \{f_i : V \to \mathbb{F} \mid i \in I\},$$

where f_i's are defined on the basis B:

$$f_i(v_j) = \delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

Then we extend f_i's linearly, i.e., for $\sum_{j=1}^{N} \alpha_j v_j \in V$,

$$f_i(\sum_{j=1}^{N} \alpha_j v_j) = \sum_{i=1}^{N} \alpha_j f_i(v_j).$$

It's clear that $f_i \in V^*$ is well-defined. ■

Our question is that whether the B^* can be the basis of V^*?

4.5. Wednesday for MAT3006

The quiz will be held on Wednesday.

Reviewing. Let's go through the proof for Weierstrass Theorem quickly.

- Study $Q_n(x) = c_n(1-x^2)^n$ and construct the approximate function

$$p_n(x) = \int_{-1}^{1} Q_n(t) f(x+t)\,\mathrm{d}t$$

- Show that

$$\begin{aligned}
|p_n(x) - f(x)| &\le \int_{-1}^{1} |f(x+t) - f(x)| Q_n(t)\,\mathrm{d}t \\
&= \left(\int_{\delta}^{1} + \int_{\delta}^{-\delta} + \int_{-1}^{-\delta}\right) |f(x+t) - f(x)| Q_n(t)\,\mathrm{d}t \\
&\le 4M\sqrt{n}(1-\delta^2)^n + \int_{\delta}^{-\delta} |f(x+t) - f(x)| Q_n(t)\,\mathrm{d}t \\
&\le 4M\sqrt{n}(1-\delta^2)^n + \varepsilon \cdot \int_{\delta}^{-\delta} Q_n(t)\,\mathrm{d}t \\
&\le 4M\sqrt{n}(1-\delta^2)^n + \varepsilon
\end{aligned}$$

  Therefore, $\|p_n - f\|_\infty \to 0$ as $n \to \infty$.

- Generalization for $\forall g \in \mathcal{C}[0,1]$: Recall that we have assumed $f(0) = f(1) = 0$. Now consider the general case, say

$$g(0) = a, \quad g(1) = b.$$

  Consider $f(x) := g(x) - l(x)$, where l is the line segment from $(0,a)$ to $(1,b)$. Then we imply $|f(x) - p_n(x)| < \varepsilon$, i.e.,

$$|g(x) - (p_n(x) + l(x))| < \varepsilon, \quad \forall x.$$

- Generlization for $\forall h \in \mathcal{C}[a,b]$: Recall that we have restrict f is continuous on $[0,1]$. For any $h \in \mathcal{C}[a,b]$, define $g(x) = h((b-a)x + a)$ for $x \in [0,1]$. Therefore,

$g \in C[0,1]$, i.e., $|g(y) - p_n(y)| < \varepsilon, \forall y \in [0,1]$, which implies

$$|h((b-a)y+a) - p_n(y)| < \varepsilon, \quad \forall y \in [0,1]$$

Applying change of variables with $x = (b-a)y + a$, we imply

$$\left|h(x) - p_n\left(\frac{x-a}{b-a}\right)\right| < \varepsilon, \quad \forall x \in [a,b],$$

where $p_n(\cdot)$ is a polynomial function.

4.5.1. Stone-Weierstrass Theorem

The motivation is to generalize the Weierstrass approximation into the space $C(X)$, where (X,d) is a general compact space. Here $C(X) := \{f : X \to \mathbb{R} \text{ is continuous}\}$. Note that

- $C(X)$ has a norm:
$$\|f\|_\infty := \sup\{f(x) \mid x \in X\}$$
This is well-defined, since $f(X) \subseteq \mathbb{R}$ is compact, i.e., closed and bounded.
- $(C(X), d_\infty)$ is complete. The proof follows similarly from the proof that $C[a,b]$ is complete (see Example (2.15)).

R If X is not compact, then the norm $\|\cdot\|_\infty$ is **not** well-defined on $C(X)$, but this norm is still well-defined on the space

$$C_b(X) = \{f : X \to \mathbb{R} \mid f \text{ is continous and bounded}\}.$$

If X is compact, then $C(X) = C_b(X)$.

Definition 4.10 [Separation Property] Let (X,d) be any metric space, and $\mathcal{A} \subseteq C_b(X)$ is an algebra (closed under linear combination and pointwise product), then

1. $\mathcal{A}$ is said to be equipped with the **separation property** if for any $x_1 \neq x_2 \in X$, there exists $f \in \mathcal{A}$ such that $f(x_1) \neq f(x_2)$
2. $\mathcal{A}$ is said to be equipped with the **nonvanishing property** if for any $x \in X$, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$.

■

■ **Example 4.4** Suppose that $X := S^1 := \{e^{i\theta} \mid \theta \in [0,2\pi]\} \subseteq \mathbb{C} \cong \mathbb{R}^2$, and consider the algebra

$$\mathcal{A} = \langle g \rangle := \text{span}\{1, g, g^2, \dots\}$$

Define $g : S^1 \to \mathbb{R}$ as $g(e^{i\theta}) = \cos\theta$. Note that

1. $\mathcal{A}$ does not satisfy the separation property: take $e^{i\theta}, e^{i(2\pi-\theta)}$
2. However, $\mathcal{A}$ satisfies the nonvanishing property. Consider the special element of $\mathcal{A}$: $f \equiv 1$.

■

Theorem 4.4 — **Stone-Weierstrass Theorem.** Let (X,d) be a compact space, and $\mathcal{A} \subseteq C(X)$ is an algebra. Then $\overline{A} = C(X)$ iff A satisfies both the **nonvanishing** and **separation** property.

Before going through the proof, we establish two lemmas below:

Proposition 4.12 If both f,g belong to the algebra $\mathcal{A}$, then $\max\{f,g\} \in \mathcal{A}$ and $\min\{f,g\} \in \overline{\mathcal{A}}$.

Proof. Since

$$\begin{aligned}
\max\{f,g\} &= \frac{1}{2}(f+g) + \frac{1}{2}|f-g| \\
\min\{f,g\} &= \frac{1}{2}(f+g) - \frac{1}{2}|f-g|,
\end{aligned}$$

it suffices to show $|h| \in \overline{\mathcal{A}}$ given that $h \in \mathcal{A}$.

Let $M = \max\{|h(x)| \mid x \in X\}$. Consider the function (w.r.t. t) $|t| \in C[-M, M]$. By Weierstrass approximation, there exists a polynomial p such that $||t| - p(t)| < \varepsilon$, which implies

$$||h(x)| - p(h(t))| < \varepsilon.$$

Note that $p(h(t))$ is a polynomial of $h(t)$, and therefore an element from the algebra $\mathcal{A}$. Therefore, $|h|$ can be approximated by some element from $\mathcal{A}$, i.e., $|h| \in \overline{\mathcal{A}}$. ■

Proposition 4.13 Let $\mathcal{A} \subseteq C(X)$ be an algebra satisfying the separation property and non-vanishing property. Then for all $x_1 \neq x_2 \in X$, and any $\alpha, \beta \in \mathbb{R}$, there exists $f \in \mathcal{A}$ such that

$$\begin{cases} f(x_1) = \alpha \\ f(x_2) = \beta \end{cases}$$

Proof. By separation property, there exists $h \in \mathcal{A}$ such that $h(x_1) \neq h(x_2)$.

1. We claim that we can construct a new h such that

$$h(x_1) \neq h(x_2), \quad h(x_1) \neq 0, \quad h(x_2) \neq 0 \tag{4.5}$$

   (a) If both $h(x_1), h(x_2) \neq 0$, we have done.
   (b) If not, suppose $h(x_1) = 0$. By non-vanishing property, there eixsts $p \in \mathcal{A}$ such that $p(x_1) \neq 0$. Then some linear transformation of h and p will do the trick. (hint: construct t such that $h \leftarrow h + t \cdot p$ gives the desired result.)

2. Now suppose the requirement (4.5) is met. Consider the function

$$f(x) = ah(x) + bh^2(x) \in \mathcal{A},$$

   where a, b are two parameters to be determined.

   Indeed, it suffices to find a, b such that $f(x_1) = \alpha, f(x_2) = \beta$, or equivalently, solve the linear system

$$f(x_1) = ah(x_1) + bh^2(x_1) = \alpha$$
$$f(x_2) = ah(x_2) + bh^2(x_2) = \beta$$

Since the determinant of the linear system is not equal to 0, a, b can be clearly found.

The proof is complete. ■

Necessity part of the proof. Given that $\mathcal{A}$ has separation and non-vanishing, we aim to show $\overline{\mathcal{A}} = C(X)$.

1. Take any $f \in C(X)$. By proposition (4.13), for any $x, y \in X$, there exists $\phi_{x,y} \in \mathcal{A}$ such that
$$\begin{cases} \phi_{x,y}(x) = f(x) \\ \phi_{x,y}(y) = f(y) \end{cases}.$$
Construct the open set $U_{x,y} = (f - \phi_{x,y})^{-1}((-\varepsilon, \varepsilon))$, i.e.,
$$U_{x,y} = \{t \in X \mid \phi_{x,y}(t) - \varepsilon < f(t) < \phi_{x,y}(t) + \varepsilon\}.$$

2. It's clear that $x, y \in U_{x,y}$. For fixed $y \in X$, the collection $\{U_{x,y}\}_{x \in X}$ forms an open cover of X. By the compactness of X, there exists the finite subcover
$$\{U_{x_1,y}, \dots, U_{x_N,y}\} \supseteq X.$$
By proposition (4.12), the function $\phi_y := \max\{\phi_{x_1,y}, \dots, \phi_{x_N,y}\} \in \overline{\mathcal{A}}$. Furthermore, for $\forall x \in X$, we imply there exists some $U_{x_i,y} \ni x$, i.e.,
$$f(x) < \phi_{x_i,y}(x) + \varepsilon \implies f(x) < \phi_y(x) + \varepsilon, \ \forall x \in X.$$

3. Also, consider $V_y = \cap_{i=1}^N U_{x_i,y}$, which is the open set containing y, and $\{V_y\}_{y \in X}$

covers X (why?). Note that for any $x \in V_y$, we imply $x \in U_{x_i,y}, \forall i$, i.e.,

$$\phi_{x_i,y}(x) - \varepsilon < f(x), \quad \forall i \implies \phi_y(x) - \varepsilon < f(x),\ \forall x \in V_y.$$

By the compactness of X again, we take finite subcover $\{V_{y_j}\}_{j=1}^{M}$ and define

$$\phi(x) := \min\{\phi_{y_1}(x), \dots, \phi_{y_M}(x)\} \in \overline{\mathcal{A}}.$$

Therefore, for any $x \in X$ we imply $x \in V_{y_m}$, i.e.,

$$\phi_{y_m}(x) - \varepsilon < f(x) \implies \phi(x) - \varepsilon < f(x) \tag{4.6}$$

4. Also, from (2) we have obtained $f(x) < \phi_y(x) + \varepsilon$ for $\forall y \in X$. In particular,

$$f(x) < \phi_{y_m}(x) + \varepsilon,\ \forall m = 1, \dots, M \tag{4.7}$$

Combining (4.6) and (4.7), we imply $|\phi(x) - f(x)| < \varepsilon$.

Therefore, we have constructed a function $\phi \in \overline{\mathcal{A}}$ such that $|\phi(x) - f(x)| < \varepsilon$, which implies $f \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$. The proof is complete. ■

4.6. Wednesday for MAT4002

There will be a quiz on Monday.

Reviewing.

- Connectedness / Path-Connectedness

4.6.1. Remark on Connectedness

Proposition 4.14 All path connected spaces X are connected.

Proof. Fix any $x \in X$, for all $y \in X$, there exists a continuous mapping $p_y : [0,1] \to X$ such that

$$p_y(0) = x, \quad p_y(1) = y.$$

Consider $C_y = p_y([0,1])$, which is connected, due to proposition (4.9).

Note that $\{C_y\}_{y \in X}$ is a collection of connected sets, and for any $y, y' \in X$, $C_y \cap C_{y'} \ni \{x\}$ is non-empty. Applying proposition (4.10), we imply $X = \cup_{y \in X} C_y$ is connected. ■

■ **Example 4.5** 1. Exercise: if $A \subset B \subset \overline{A}$, then A is connected implies B is connected. (Hint: $U \cap A = \emptyset$ implies $U \cap \overline{A} = \emptyset$ for all open sets U in X.)

Proof. Suppose B is not connected, i.e., for any open U, V such that $B \subseteq U \cup V$ and $(U \cap V) \cap B = \emptyset$, we imply $U \cap B \neq \emptyset$ and $V \cap B \neq \emptyset$, and therefore

$$U \cap \overline{A} \neq \emptyset, \quad V \cap \overline{A} \neq \emptyset$$

which implies

$$U \cap A \neq \emptyset, \quad V \cap A \neq \emptyset$$

which contradicts to the connectedness of A. ■

2. The converse of proposition (4.14) may not be necessarily true. Consider the so-called **Topologist's comb** example:

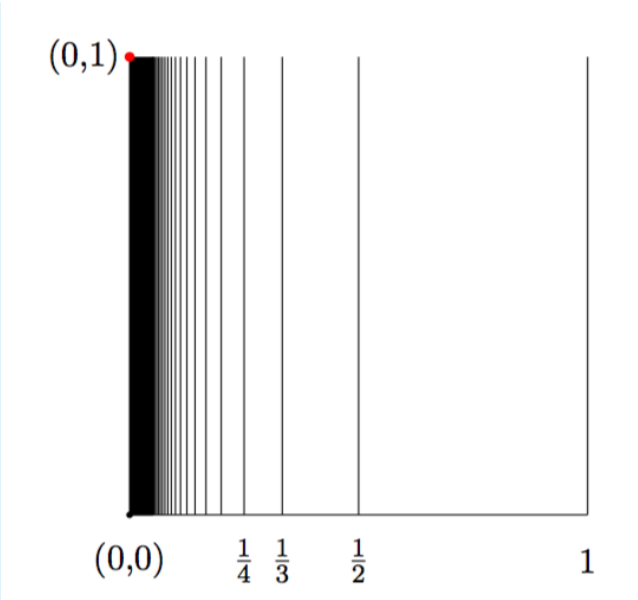


Figure 4.1: Connected space X but not path-connected

Here we construct a connected space $X \subseteq \mathbb{R}^2$ but not path-connected shown in Fig (4.1), i.e., the union of the interval $[0,1]$ together with vertical line segments from $(1/n, 0)$ to $(1/n, 1)$ and the single point $(0,1)$.

$$X = ([0,1] \times \{0\}) \cup \bigcup_{n \geq 1} (\{1/n\} \times [0,1]) \cup (0,1).$$

(a) Firstly, X is not path-connected. We show that there is no path in X links $(0,1)$ to any other point, i.e., for continuous mapping $p : [0,1] \to X$ with $p(0) = (0,1)$, we may imply $p(t) = (0,1)$ for any t.

Define

$$A = \{t \in [0,1] \mid p(t) = (0,1)\}.$$

We claim that $A = [0,1]$, i.e., suffices to show A is both open and closed in $[0,1]$:

i. The set $A = p^{-1}((0,1))$ is nonempty and closed, since the pre-image of a closed set is closed as well.

ii. The set A is open: choose $t_0 \in A$. By continuity of p, there exists $\delta > 0$ such that

$$\|p(t) - (0,1)\| = \|p(t) - p(t_0)\| < \frac{1}{2}, \quad t \in [0,1] \cap (t_0 - \delta, t_0 + \delta).$$

Since there is no point on the x-axis with the distance $1/2$ to the point $(0,1)$, we imply $p(t)$ is not on the x-axis when $t \in [0,1] \cap (t_0 - \delta, t_0 + \delta)$. Therefore, the x-coordinate of $p(t)$ is either 0 or of the form $1/n$.

It suffices to show the open interval $I := [0,1] \cap (t_0 - \delta, t_0 + \delta)$ is in A. Define the composite function $f = x \circ p : I \to \mathbb{R}$, where the mapping $x : \mathbb{R}^2 \to \mathbb{R}$ is defined as $(a,b) \mapsto a$. Note that I is connected, we imply $f(I)$ is connected, and $f(I)$ belongs to $\{0\} \cup \{1/n\}$.

The only nonempty connected subset of $\{0\} \cup \{1/n\}$ is a single point (left as exercise), and therefore $f(I)$ is a single point. Since $f(t_0) = 0$,we imply $f(I) = \{0\}$, i.e., $I \subseteq A$. Therefore A is open.

■

4.6.2. Compactness

Compact set in X is used to generalize "closed and bounded" in $\mathbb{R}^n$.

Definition 4.11 Let $(X, \mathcal{T})$ be a topological space. A collection $\mathcal{U} = \{U_i \mid i \in I\}$ of open sets is an open cover of X if

$$X = \bigcup_{i \in I} U_i$$

A subcover of $\mathcal{U}$ is a subfamily

$$\mathcal{U}' = \{U_j \mid j \in J\}, \quad J \subseteq I$$

such that $\bigcup_{j \in J} U_j = X$.

If J has finitely many elements, we say $\mathcal{U}'$ is a finite subcover of X.

We say X is compact if any open cover of X has a finite subcover. ■

R If $A \subseteq X$ has a subspace topology. then A is compact iff for any open collection of open sets (in X) $\{U_i\}$ such that $A \subseteq \bigcup_{i \in I} U_i$, there exists a fintie subcover $A \subseteq \bigcup_{k=1}^{n} U_{i_k}$.

Proposition 4.15 Let X be a topological space. The followings are equivalent:

1. The space X is compact
2. If $\{V_i \mid i \in I\}$ is a collection of closed subsets in X such that
$$\bigcap_{j \in J} V_j \neq \emptyset, \quad \text{for all finite } J \subseteq I,$$
then $\cap_{i \in I} V_i \neq \emptyset$.

Compactness is an **intrisical** property, i.e., we do not need to worry about which underlying space for this definition.

▪ **Example 4.6** 1. $X \subseteq \mathbb{R}^n$ is compact iff X is closed and bounded. (Heine-Borel)

2. Let $K \subseteq \mathbb{R}^n$ be compact, then define the set
$$C(K) = \{\text{all continuous mapping } f : K \to \mathbb{R}\}$$
Note that the d_∞ metric associated with $C(K)$, say $\|f\|_\infty = \sup_{k \in K} f(k)$, is well-defined.

Under the metric space $(C(K), d_\infty)$, any $\mathcal{J} \subseteq C(K)$ is compact, if and only if $\mathcal{J}$ is closed, bounded, and equi-continuous. (Aresul-Ascoli)

Therefore, we can see that the compactness is not equivalent to the closedness together with boundedness. ▪

Proposition 4.16 Let X be a compact space, then all closed subset $A \subseteq X$ are compact.

Proof. Let $\{V_i \mid i \in I\}$ be a collection of closed subsets in A such that
$$\cap_{j \in J} V_j \neq \emptyset, \quad \text{for any finite } J \subseteq I.$$
As A is closed in X, we imply V_j is closed in X.

Due to the compactness of X and proposition (4.15), we imply

$$\cap_{i\in I} V_i \neq \emptyset$$

By the reverse direction of proposition (4.15), we imply A is compact. ∎

R Now consider the reverse direction of proposition (4.16), i.e., are all compact subsets $K \subseteq X$ closed in X?

In general, the converse does not hold. Note that $K = \{x\}$ is compact for any topology X. However, there are some topologies such that $\{x\}$ is closed.

In order to obtain the converse of proposition (4.16), we need to obtain another **separation axiom**:

Proposition 4.17 Let X be Hausdorff, $K \subseteq X$ be compact, and $x \in X \setminus K$. Then there exists open $U, V \subseteq X$ such that $U \cap V = \emptyset$ and

$$U \cap V = \emptyset, \quad K \subseteq U, \quad x \in V.$$

Proof. Let $k \in K$, then by Hausdorffness, there exists open $U_k \ni k, V_k \ni x$ such that $U_k \cap V_k = \emptyset$. Therefore, $\{U_k\}_{k\in K}$ forms an open cover of K. By compactness of K, $\{U_{k_i}\}_{i=1}^n$ covers K. Constructing the set

$$U := \bigcup_{i=1}^n U_{k_i}, \quad V = \bigcap_{i=1}^n V_{k_i}$$

gives the desired result. ∎

By making use of this separation axiom, we obtain the converse of proposition (4.16):

Corollary 4.3 All compact K in Hausdorff X is closed.

Proof. For $\forall x \in X \setminus K$, by proposition (4.17) there exists open V such that $x \in V \subseteq X \setminus K$, and therefore $X \setminus K$ is open. ∎

Chapter 5

Week5

5.1. Monday for MAT3040

Reviewing.

- Dual space: the set of linear transformations from V to $\mathbb{F}$, denoted as $\mathrm{Hom}(V,\mathbb{F})$.
- Suppose $B = \{\boldsymbol{v}_i \mid i \in I\}$ is the basis of V, define $B^* = \{f_i \mid i \in I\}$ by

$$f_i(\boldsymbol{v}_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

  Actually, the above recipe uniquely defines a linear transformation $f_i : V \to \mathbb{F}$: For any $\boldsymbol{v} \in V$, it can be written as $\boldsymbol{v} = \sum_{i\in I} \alpha_i \boldsymbol{v}_i$, and therefore

$$f_i(\boldsymbol{v}) = f_i(\sum_{i\in I} \alpha_i \boldsymbol{v}_i) = \sum_{i\in I} \alpha_i f_i(\boldsymbol{v}_i).$$

Example 5.1 Consider $V = \mathbb{R}^n, B = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. Then we imply $B^* = \{\phi_i\}_{i=1}^n$, where ϕ_i is the mapping $V \to \mathbb{R}$ defined by

$$\phi_i \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \phi(x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n) = \sum_{j=1}^n x_j \phi_i(\boldsymbol{e}_j) = x_i$$

5.1.1. Remarks on Dual Space

Proposition 5.1
1. B^* is always lienarly independent, i.e., any finite subset of B^* is linearly independent.
2. If V has finite dimension, then B^* is a basis of V^*.

Proof.
1. Suppose that
$$\alpha_1 f_{i_1} + \alpha_2 f_{i_2} + \cdots + \alpha_k f_{i_k} = \mathbf{0}_{V^*}.$$
In particular, let the input of these linear transformations be $\boldsymbol{v}_{i_1}$, we imply
$$\begin{aligned}\alpha_1 f_{i_1}(\boldsymbol{v}_{i_1}) + \alpha_2 f_{i_2}(\boldsymbol{v}_{i_1}) + \cdots + \alpha_k f_{i_k}(\boldsymbol{v}_{i_1}) = \mathbf{0}(\boldsymbol{v}_{i_1}) &\equiv \mathbf{0}\\ &= \alpha_1 \cdot 1 + \cdots + 0\\ &= \alpha_1\end{aligned}$$
Applying the same trick, one can show that $\alpha_2 = \cdots = \alpha_k = 0$. Therefore, $\{f_{i_1}, \ldots, f_{i_k}\}$ is linearly independent.
2. Suppose that $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ and $B^* = \{f_1, \ldots, f_n\}$. For any $f \in V^*$, construct the linear transformation
$$g := \sum_{i=1}^{n} f(\boldsymbol{v}_i) \cdot f_i \in \operatorname{span}\{B^*\}.$$
It follows that for $j = 1, 2, \ldots, n$,
$$g(\boldsymbol{v}_j) = \sum_{i=1}^{n} f(\boldsymbol{v}_i) \cdot f_i(\boldsymbol{v}_j) = f(\boldsymbol{v}_j).$$
It's clear that $g(\boldsymbol{v}) = f(\boldsymbol{v})$ for all $\boldsymbol{v} \in V$, i.e., $f \equiv g \in \operatorname{span}(B^*)$. Therefore B^* spans V^*, i.e., forms a basis of V^*.

∎

Corollary 5.1 If $\dim(V) = n$, then $\dim(V^*) = n$.

Proof. It's eay to show the mapping defined as

$$\begin{aligned} & V \to V^* \\ \text{with} \quad & \boldsymbol{v}_i \mapsto f_i \end{aligned}$$

is an isomorphism from $V \to V^*$. Note that this constructed isomorphism depends on **the choice of basis** B in V. (We say this is not a **natural isomorphism**.) ■

R The part 2 for proposition (5.1) does not hold for V with infinite dimension. The reason is that the spanning set is defined with **finite** linear combinations. Check the example below for a counter-example.

▪ **Example 5.2** Suppose that $V = \mathbb{F}[x]$, and $B^* = \{1, x, x^2, \dots,\}$ forms a basis of V. We imply that $B^* = \{\phi_0, \phi_1, \phi_2, \dots,\}$, where ϕ_i is the mapping defined as

$$\phi_i(x^j) = \begin{cases} 1, & i = j \\ 0, & \text{otherwise} \end{cases}$$

Consider a special element $\phi \in V^*$ with $f(p(x)) = p(1)$:

$$\phi(1) = 1, \quad \phi(x) = 1, \quad \phi(x^2) = 1, \quad \cdots \quad \phi(x^n) = 1, \quad \forall n \in \mathbb{N}.$$

If following the proof in proposition (5.1), we expect that

$$g := \sum_{n=0}^{\infty} \phi(x^n)\phi_n = \sum_{n=0}^{\infty} \phi_n \in \text{span}\{B^*\},$$

which is a contradiction, since $\text{span}\{B^*\}$ consists of finite sum of ϕ_i's only. ▪

R Therefore, if V is not finite-dimensional, we can say the cardinality of V is strictly less than the cardinality of V^*.

Any subspace of a given vector space has some gap. Now we want to describe this gap formally from the perspective of the dual space.

5.1.2. Annihilators

Definition 5.1 Let V be a vector space, $S \subseteq V$ be a subset. The **annihilator** of S is defined as

$$\text{Ann}(S) = \{f \in V^* \mid f(s) = 0, \forall s \in S\}$$

■ **Example 5.3** Consider $V = \mathbb{R}^4$, $B = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_4\}$. Let $B^* = \{f_1, \ldots, f_4\}$, $S = \{\boldsymbol{e}_3, \boldsymbol{e}_4\}$.

- Then $f_1 \in \text{Ann}(S)$, since

$$f_1(\boldsymbol{e}_3) = 0, \quad f_1(\boldsymbol{e}_4) = 0$$

Indeed, any $a \cdot f_1 + b \cdot f_2 \in V^*$ is in $\text{Ann}(S)$.

Proposition 5.2
1. The set $\text{Ann}(S)$ is a vector subspace of V^*
2. The mapping $\text{Ann}(\cdot)$ is **inclusion-reversing**, i.e., if $W_1 \subseteq W_2 \subseteq V$, then

$$\text{Ann}(W_1) \supseteq \text{Ann}(W_2)$$

3. The mapping $\text{Ann}(\cdot)$ is **idempotent**, i.e., $\text{Ann}(S) = \text{Ann}(\text{span}(S))$.
4. If V has finite dimension, and $W \leq V$, then $\text{Ann}(W)$ fills in the gap, i.e.,

$$\dim(W) + \dim(\text{Ann}(W)) = \dim(V)$$

Proof.
1. Suppose that $f, g \in \text{Ann}(S)$, i.e., $f(s) = g(s) = 0, \forall s \in S$. It's clear that $(af + bg) \in \text{Ann}(S)$.
2. Suppose that $f \in \text{Ann}(W_2)$, we imply $f(\boldsymbol{w}) = 0$ for any $\boldsymbol{w} \in W_2$. Therefore, $f(\boldsymbol{w}_1) = 0$ for any $\boldsymbol{w}_1 \in W_1 \subseteq W_2$, i.e., $f \in \text{Ann}(W_1)$.
3. Note that $S \subseteq \text{span}(S)$. Therefore we imply $\text{Ann}(S) \supseteq \text{Ann}(\text{span}(S))$ from part (b). It suffices to show $\text{Ann}(S) \subseteq \text{Ann}(\text{span}(S))$:

For any $f \in \operatorname{Ann}(S)$ and any $\sum_{i=1}^{n} k_i \boldsymbol{s}_i \in \operatorname{span}(S)$, we imply

$$\begin{aligned} f\left(\sum_{i=1}^{n} k_i \boldsymbol{s}_i\right) &= \sum_{i=1}^{n} k_i f(\boldsymbol{s}_i) \\ &= \sum_{i=1}^{n} k_i \cdot 0 \\ &= 0, \end{aligned}$$

i.e., $f \in \operatorname{Ann}(\operatorname{span}(S))$.

4. Let $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ be a basis of W. By basis extension, we construct a basis of V:

$$B = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n\}.$$

Let $B^* = \{f_1, \dots, f_k, f_{k+1}, \dots, f_n\}$ be a basis of V^*. We claim that $\{f_{k+1}, \dots, f_n\}$ is a basis of $\operatorname{Ann}(W)$:

- Firstly, f_j's are the elements in $\operatorname{Ann}(W)$ for $j = k+1, \dots, n$, since for any $\boldsymbol{w} = \sum_{i=1}^{k} \alpha_i(\boldsymbol{v}_i) \in W$, we have

$$\begin{aligned} f_j(\boldsymbol{w}) &= \sum_{i=1}^{k} \alpha_i f_j(\boldsymbol{v}_i) \\ &= \sum_{i=1}^{k} \alpha_i \cdot 0 \\ &= 0, \quad j = k+1, k+2, \dots, n \end{aligned}$$

- Secondly, the set $\{f_{k+1}, \dots, f_n\}$ is linearly independent, since the set $B^* = \{f_1, \dots, f_n\}$ is linearly independent.

- Thirdly, $\{f_{k+1}, \dots, f_n\}$ spans $\operatorname{Ann}(W)$: for any $g \in \operatorname{Ann}(W) \subseteq V^*$, it can be expressed as $g = \sum_{i=1}^n \beta_i f_i$. It follows that

$$\begin{aligned} g(\boldsymbol{v}_1) &= \sum_{i=1}^n \beta_i f_i(\boldsymbol{v}_1) = 0 \implies \beta_1 = 0 \\ &\vdots \\ g(\boldsymbol{v}_k) &= \sum_{i=1}^n \beta_i f_i(\boldsymbol{v}_k) = 0 \implies \beta_k = 0 \end{aligned}$$

Substituting $\beta_1 = \cdots = \beta_k = 0$ into $g = \sum_{i=1}^n \beta_i f_i$, we imply

$$g = \beta_{k+1} f_{k+1} + \cdots + \beta_n f_n \in \operatorname{span}\{f_{k+1}, \dots, f_n\}.$$

Therefore, $\{f_{k+1}, \dots, f_n\}$ forms a basis for $\operatorname{Ann}(W)$, i.e., $\dim(\operatorname{Ann}(W)) = n - k$.

$\blacksquare$

R Let $W \le V$, where V has finite dimension, recall that we have obtained two relations below:

$$\begin{aligned} \dim(\operatorname{Ann}(W)) &= \dim(V) - \dim(W) \\ \dim((V/W)^*) &= \dim(V/W) = \dim(V) - \dim(W) \end{aligned}$$

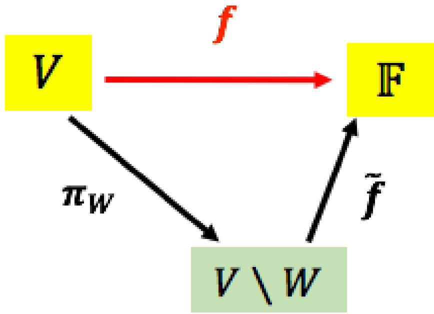
Therefore, $\dim((V/W)^*) = \dim(\operatorname{Ann}(W))$, i.e.,

$$(V/W)^* \cong \operatorname{Ann}(W).$$

The question is that can we construct an isomorphism explicitly? We claim that the mapping defined below is an isomorphism:

$$\begin{aligned} &\operatorname{Ann}(W) \to (V/W)^* \\ \text{with} \quad &f \mapsto \tilde{f}, \end{aligned}$$

where $\tilde{f} : V/W \to \mathbb{F}$ is constructed from the **universal property I**, i.e., given the mapping $f \in \text{Ann}(W)$, since $W \leq \ker(f)$, there exists $\tilde{f} : V/W \to \mathbb{F}$ such that the diagram below commutes:



i.e., $\tilde{f}(\boldsymbol{v} + W) = f(\boldsymbol{v})$.

5.2. Monday for MAT3006

Our first quiz will be held on this Wednesday.

Reviewing. We have shown that the algebra $\mathcal{A} \subseteq C(X)$ with separation, non-vanishing property implies $\overline{\mathcal{A}} = C(X)$.

Now we show that if $\overline{\mathcal{A}} = C(X)$, then the algebra $\mathcal{A}$ has separation, non-vanishing property:

1. Suppose on the contrary that $\mathcal{A}$ is not separating, i.e., there exists $x_1, x_2 \in X$ such that $\phi(x_1) = \phi(x_2)$, $\forall \phi \in \mathcal{A}$.
   By the defintion of closure, it's clear that for given $S \subseteq (X,d)$, $\forall x \in \overline{S}$, there exists a sequence $\{S_n\}$ in S such that $S_n \to x$.
   Construct $f \in C(X)$ defined by $f(x) = d(x, x_1)$. It follows that
   $$f(x_1) = 0, \quad f(x_2) = d(x_2, x_1) := k > 0$$
   Now we claim that $f \notin \overline{\mathcal{A}}$, since otherwise there exists $\{\phi_n\}$ in $\mathcal{A}$ such that $\phi_n \to f$, i.e.,
   $$\phi_n(x_1) \to f(x_1), \quad \phi_n(x_2) \to f(x_2), \quad \phi_n(x_1) = \phi_n(x_2), \forall n,$$
   i.e., $0 = f(x_1) = f(x_2) > 0$.
2. Suppose on the contrary that $\mathcal{A}$ is not non-vanishing, i.e., there exists some $x_0 \in X$ such that $\phi(x_0) = 0, \forall \phi \in \mathcal{A}$. Construct $g \in C(X)$ defined by $g(x) = d(x, x_0) + 1$. Following the similar idea, we can show that there does not exist $\phi_n \in \mathcal{A}$ such that $\phi_n \to g$, i.e., $g \notin \overline{\mathcal{A}}$, which is a contradiction.

■ **Example 5.4** 1. Let $X \subseteq \mathbb{R}^n$ be a compact space. Then the polynomial ring
$$\mathbb{R}[x_1, \dots, x_n] = \{\text{Polynomials in } n \text{ variables with coefficients in } \mathbb{R}\}$$
forms a dense set in $C(X)$.

It's clear that the set $\mathbb{R}[x_1, \ldots, x_n]$ satisfies the separating and non-vanishing property. For the special case $n = 1$ and $X = [a, b]$, we get the Weierstrass Approximation Theorem.

2. In particular, when $X = S^1 \subseteq \mathbb{R}^2$, we imply $\mathbb{R}[x, y]$ is dense in $C(S^1)$.

5.2.1. Stone-Weierstrass Theorem in $\mathbb{C}$

Consider the circle $S^1 \subseteq \mathbb{C}$ and the mappings

$$\begin{array}{rlrl} & c : S^1 \to \mathbb{R} & & s : S^1 \to \mathbb{R} \\ \text{with} & e^{i\theta} \to \cos\theta & \qquad \text{with} & e^{i\theta} \to \sin\theta \end{array}$$

are both continuous.

The algebra formed by s and c is given by

$$\mathcal{J} := \langle c, s \rangle = \text{span}\{\cos^m\theta \sin^n\theta \mid m, n \in \mathbb{N}\}$$

1. The $\mathcal{J}$ satisfies both separating and non-vanishing property, which implies $\overline{\mathcal{J}} = C(S^1)$.
2. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a continuous, 2π-periodic mapping. It's easy to construct a continuous mapping $\tilde{f} : S^1 \to \mathbb{R}$ such that the diagram below commutes:

$$\begin{tikzcd} \mathbb{R} \arrow[rr, "f"] \arrow[dr, "\theta \mapsto e^{i\theta}"'] & & \mathbb{R} \\ & S^1 \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

Or equivalently, $f(\theta) = \tilde{f}(e^{i\theta})$ for some $\tilde{f} \in C(S^1)$. Since $\overline{\mathcal{J}} = C(S^1)$, we can approximate $\tilde{f} \in C(S^1)$ by $\langle \cos\theta, \sin\theta \rangle$, which implies that the $f(\theta)$ can be approximated by

$$\sum_{m,n \in \mathbb{N}} a_{m,n} \cos^m \theta \sin^n \theta.$$

Since $\text{span}\{\cos^m \theta \sin^n \theta\}_{m,n\in\mathbb{N}} = \text{span}\{\cos(m\theta), \sin(n\theta), 1\}_{m,n\in\mathbb{N}}$, we imply $f(\theta)$ can be approximated by

$$\sum_{m,n \in \mathbb{N}} a_m \cos(m\theta) + b_n \sin(n\theta).$$

Or equivalently, for any $\varepsilon > 0$, there exists $N > 0$ and $a_m, a_n \in \mathbb{R}$ such that

$$\left| f(\theta) - \left(a_0 + \sum_{m=1}^{N} a_m \cos(m\theta) + \sum_{n=1}^{N} b_n \sin(n\theta) \right) \right| < \varepsilon, \quad \forall \theta \in [0, 2\pi]. \tag{5.1}$$

R The natural question is that do we have the following equation hold:

$$f(\theta) = a_0 + \sum_{m=1}^{\infty} a_m \cos(m\theta) + \sum_{n=1}^{\infty} b_n \sin(n\theta) \tag{5.2}$$

It seems that Eq.(5.2) above is equivalent to the expression in (5.4). However, unlike the Taylor expansion, the values of a_m, a_n, M, N may change once we switch the number $\varepsilon > 0$.

Therefore, Eq.(5.2) does not hold for most functions, but only for some functions with nice structure.

Fourier Analysis. Given the condition that the Eq.(5.2) holds. How can we get the values of a_m and b_n? The way is to take "inner product" between $f(\theta)$ and trigonometric functions. For example, by taking the inner product with $\cos(k\theta)$ for

Eq.(5.2) both sides, we have

$$\begin{aligned}\int_{-\pi}^{\pi} f(\theta)\cos(k\theta)\,\mathrm{d}\theta &= \frac{a_0}{2}\int_{-\pi}^{\pi}\cos(k\theta)\,\mathrm{d}\theta \\ &+ \sum_{m=1}^{\infty} a_m \int_{-\pi}^{\pi}\cos(m\theta)\cos(k\theta)\,\mathrm{d}\theta + \sum_{m=1}^{\infty} b_n \int_{-\pi}^{\pi}\sin(n\theta)\cos(k\theta)\,\mathrm{d}\theta \\ &= \pi\cdot a_k\end{aligned}$$

Following the same trick, we obtain:

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos(k\theta)\,\mathrm{d}\theta \\ b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin(k\theta)\,\mathrm{d}\theta \end{aligned} \tag{5.3}$$

Naturally, we define the fourier expansion for general $f(\theta)$, even though we don't verify whether (5.2) holds or not:

$$g_N(\theta) = \frac{a_0}{2} + \sum_{n=1}^{N} a_m\cos(m\theta) + \sum_{n=1}^{N} b_n\sin(n\theta),$$

where the term a_m and b_n follow the definition in (5.3). The natural question is that whether $g_N(\theta)\to f(\theta)$ as $N\to\infty$?

5.2.2. Baire Category Theorem

Motivation. The set $\mathcal{P}[a,b]\subseteq C[a,b]$ is dense by Weierstrass Approximation. However, it is not "abundant" in $C[a,b]$, just like $\mathbb{Q}\subseteq\mathbb{R}$ is dense in $\mathbb{R}$. (Every $r\in\mathbb{R}$ is a limit of a sequence in $\mathbb{Q}$)

The set $\mathbb{Q}$ is countable yet $\mathbb{R}\setminus\mathbb{Q}$ is uncountable, i.e., there are many more holes in $\mathbb{R}\setminus\mathbb{Q}$.

Definition 5.2 [Nowhere Dense] A subset $S\subseteq(X,d)$ is **nowhere dense** if $\overline{S}$ does not contain any open ball, i.e.,

$$X\setminus\overline{S} \text{ is dense in } X$$

For example, a single point is nowhere dense.

Theorem 5.1 Let $\{E_i\}_{i=1}^{\infty}$ be a collection of nowhere dense sets in a complete metric space (X, d). Then the set

$$\bigcup_{i=1}^{\infty} \overline{E_i}$$

also does not contain any open ball.

Proof. I have no time to review and modify the proof during the lecture. Therefore, we encourage the reader to go through the proof in the note

> W,Ni & J. Wang (January, 2019). Lecture Notes for MAT2006. Retrieved from `https://walterbabyrudin.github.io/information/information.html`

Of course, I will also add the proof in this note during this week. ∎

5.3. Monday for MAT4002

5.3.1. Continuous Functions on Compact Space

Proposition 5.3 Let $f: X \to Y$ be continuous function on topological spaces, with $A \subseteq X$ compact. Then $f(A) \subseteq Y$ is compact.

Proof. Let $\{U_i \mid i \in I\}$ be an open cover of $f(A)$, i.e.,

$$f(A) \subseteq \bigcup_{i \in I} U_i, \quad U_i \in \mathcal{T}_Y$$

It follows that $\{f^{-1}(U_i) \mid i \in I\}$ is an open cover of A:

$$A \subseteq f^{-1}\left(\bigcup_{i \in I} U_i\right) = \bigcup_{i \in I} f^{-1}(U_i)$$

By the compactness of A, there exists finite subcover of A:

$$A \subseteq \bigcup_{k=1}^{n} f^{-1}(U_{i_k}),$$

which implies the constructed finite subcover of $f(A)$:

$$\begin{aligned} f(A) &\subseteq f(\cup_{k=1}^{n} f^{-1}(U_{i_k})) \\ &= \bigcup_{k=1}^{n} U_{i_k} \end{aligned}$$

■

Corollary 5.2
1. Suppose that X is compact, and the mapping $f: X \to \mathbb{R}$ is continuous, then $f(X)$ is closed and bounded, i.e., there exists $m, M \in X$ such that $f(m) \le f(x) \le f(M)$, $\forall x \in X$.
2. Suppose moreover that X is connected, then

$$f(X) = [f(m), f(M)].$$

Theorem 5.2 The space X, Y are compact iff $X \times Y$ is compact under product topology.

Proof. 1. *Sufficiency:* Given that $X \times Y$ is compact, consider the projection mapping (which is continuous):

$$\begin{cases} P_X : X \times Y \to X \\ P_Y : X \times Y \to Y \end{cases}$$

By applying proposition (5.3), $P_X(X \times Y) = X, P_Y(X \times Y) = Y$ are both compact.

2. *Necessity:* Suppose that $\{W_i\}_{i \in I}$ is an open cover of $X \times Y$. Each open set W_i can be written as:

$$W_i = \bigcup_{j \in \mathcal{J}_i} U_{ij} \times V_{ij}, \quad U_{ij} \in \mathcal{T}_X, V_{ij} \in \mathcal{T}_Y.$$

It follows that

$$X \times Y = \bigcup_{(i,j) \in K} U_{ij} \times V_{ij}, \quad K = \{(i,j) \mid i \in I, j \in \mathcal{J}_i\}$$

Therefore, it suffices to show $\{U_{ij} \times V_{ij} \mid (i,j) \in K\}$ has a finite subcover of $X \times Y$.

- Note that $X \times \{y\} \subseteq \bigcup_{(i,j) \in K} U_{ij} \times V_{ij}$ is compact for each $y \in Y$, which implies there exists finite $S_y \in K$ such that

$$X \times \{y\} \subseteq \bigcup_{s \in S_y} U_s \times V_s$$

- w.l.o.g., assume that $y \in V_s, \forall s \in S_y$, since we can remove the $U_s \times V_s$ such that $y \notin V_s$. Define the set $V_y := \cap_{s \in S_y} V_s$, which is an open set containing y. We imply $\{V_y\}_{y \in Y}$ forms an open cover of Y. By the compactness of Y,

$$\{V_{y_1}, \dots, V_{y_m}\}$$

forms a finite subcover of Y.

- For each $\ell = 1, \dots, m$,
$$X \times \{y_\ell\} \subseteq \bigcup_{s \in S_{y_\ell}} U_s \times V_s$$
Note that for any $(x, y) \in X \times Y$, there exists $\ell \in \{1, \dots, m\}$ such that $y \in V_{y\ell}$, i.e., $y \in V_s$ for $\forall s \in S_{y\ell}$. Therefore,
$$X \times Y = \bigcup_{\ell=1}^{m} \bigcup_{s \in S_{y_\ell}} U_s \times V_s$$

Now pick
$$I' = \{i \in I \mid (i, j) \in \cup_{\ell=1}^{m} S_{y_\ell}\},$$
we imply $X \times Y = \bigcup_{i' \in I'} W_i$ and I' is finite.

■

Theorem 5.3 Suppose that X is compact, Y is Hausdorff, $f : X \to Y$ is continuous, bijective, then f is a **homeomorphism**.

Proof. It suffices to show f^{-1} is continuous. Therefore, it suffices to show $(f^{-1})^{-1}(V)$ is closed, given that V is closed in X:

Let $V \subseteq X$ be closed. Then V is compact, which implies $f(V)$ is compact. Since $f(V) \subseteq Y$ is Hausdorff, we imply $f(V)$ is compact, i.e., $f(V)$ is closed. ■

5.4. Wednesday for MAT3040

There will be a quiz on next Monday.

$$\text{Scope :}\quad \text{From Week 1 up to (including) the definition of } B^*.$$

Reviewing.

1. If V is finite dimensional, and B a basis of V, then B^* is a basis of the dual space V^*.
2. Define the Annihilator $\text{Ann}(S) \le V^*$:
$$\text{Ann}(S) = \{f \in V^* \mid f(s) = 0, \forall s \in S\}$$
3. If V is finite dimensional, and $W \le V$, then $\text{Ann}(W)$ fills the gap, i.e.,
$$\dim(\text{Ann}(W)) = \dim(V) - \dim(W)$$
4. Define a map
$$\begin{aligned}\Phi: \quad & \text{Ann}(W) \to (V/W)^* \\ & f \mapsto \tilde{f}\end{aligned}$$
where $\tilde{f}$ is defined such that the diagram (5.1) below commutes

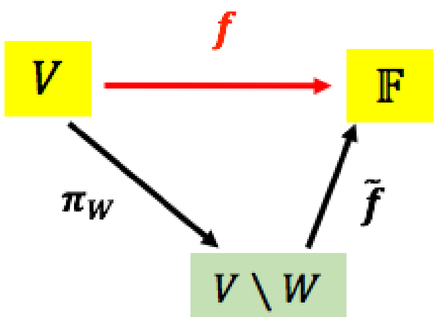


Figure 5.1: Construction of $\tilde{f}$

Or equivalently, $\tilde{f} : V/W \to \mathbb{F}$ is such that $\tilde{f}(\boldsymbol{v} + W) = f(\boldsymbol{v})$.

5.4.1. Adjoint Map

The natural question is that whether Φ is the isomorphism between $\text{Ann}(W)$ and $(V/W)^*$:

Proposition 5.4 Φ is a linear transformation, i.e.,

$$\Phi(af+bg) = a\cdot\Phi(f) + b\cdot\Phi(g).$$

Proof. Itt suffices to show that

$$\overline{af+bg} = a\overline{f} + b\overline{g}$$

■

Therefore, we need to answer whether Φ a bijective map. We will show this conjucture at the end of this lecture. The definition of Φ is **natural**, i.e., we do not need to specify any basis to define this Φ . However, as studied in Monday, the constructed isomorphism $V \to V^*$ with $\boldsymbol{v}_i \mapsto f_i$ is not natural.

Definition 5.3 [Adjoint Map] Let $T: V \to W$ be a linear transformation. Define the **adjoint** of T by

$$T^*: \quad W^* \to V^*$$

such that for any $f \in W^*$,

$$[T^*(f)](\boldsymbol{v}) := f(T(\boldsymbol{v})),\ \forall \boldsymbol{v} \in V.$$

R

1. In other words, $T^*(f) = f\circ T$, i.e., a linear transformation from V to $\mathbb{F}$, i.e., belongs to V^*.
2. Moreover, the mapping T^* itself is a linear transformation: For $f, g \in W^*$,

and $\forall \boldsymbol{v} \in V$,

$$
\begin{aligned}
[T^*(af+bg)](\boldsymbol{v}) &= (af+bg)[T(\boldsymbol{v})] \\
&= af(T(\boldsymbol{v})) + bg(T(\boldsymbol{v})) && \text{definition of } W^* \text{ as a vector space} \\
&= a[T^*(f)](\boldsymbol{v}) + b[T^*(g)](\boldsymbol{v}) \\
&= [aT^*(f) + bT^*(g)](\boldsymbol{v}) && \text{definition of } V^* \text{ as a vector space}
\end{aligned}
$$

Proposition 5.5 Let $T : V \to W$ be a linear transformation.

1. If T is **injective**, then T^* is **surjective**.
2. If T is **surjetive**, then T^* is **injective**.

This statement is quite intuitive, since T^* reverses the dual of output into the dual of input:

$$
T : V \to W
$$

$$
T^* : W^* \to V^*
$$

Proof. We only give a proof of (2), i.e., suffices to show $\ker(T) = \{\mathbf{0}\}$.

Consider any $g \in W^*$ such that $T^*(g) = \mathbf{0}_{V^*}$. It follows that

$$
[T^*(g)](\boldsymbol{v}) = \mathbf{0}_{V^*}(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in V. \Longleftrightarrow g(T(\boldsymbol{v})) = \mathbf{0}, \quad \forall \boldsymbol{v} \in V. \tag{5.4}
$$

To show $g = \mathbf{0}_{W^*}$, it suffices to show $g(\boldsymbol{w}) = \mathbf{0}$ for $\forall \boldsymbol{w} \in W$. For all $\boldsymbol{w} \in W$, by the surjectivity of T, there exists $\boldsymbol{v}' \in V$ such that

$$
\boldsymbol{w} = T(\boldsymbol{v}').
$$

By substituting $\boldsymbol{w}$ with $T(\boldsymbol{v}')$ and (5.4), we imply

$$
g(\boldsymbol{w}) = g(T(\boldsymbol{v}')) = \mathbf{0}.
$$

The proof is complete. ∎

Proposition 5.6 Let $T : V \to W$ be a linear transformation, and $\mathcal{A} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}, \mathcal{B} = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ be the bases of V and W, respectively. Let $\mathcal{A}^* = \{f_1, \ldots, f_n\}, \mathcal{B}^* = \{g_1, \ldots, g_m\}$ be bases of dual spaces V^* and W^*, respectively. Then $T^* : W^* \to V^*$ admits a matrix representation

$$(T^*)_{\mathcal{A}^*\mathcal{B}^*} = \text{transpose}((T)_{\mathcal{B}\mathcal{A}})$$

where $(T^*)_{\mathcal{A}^*\mathcal{B}^*} \in \mathbb{F}^{n\times m}$ and $(T)_{\mathcal{B}\mathcal{A}} \in \mathbb{F}^{m\times n}$

Proof. Let $(T)_{\mathcal{B}\mathcal{A}} = (\alpha_{ij})$ and $(T^*)_{\mathcal{A}^*\mathcal{B}^*} = (\beta_{ij})$. By definition of matrix representation,

$$T(\boldsymbol{v}_j) = \sum_{i=1}^m \alpha_{ij}\boldsymbol{w}_i, \qquad T^*(g_i) = \sum_{k=1}^n \beta_{ki} f_k \in V^*$$

As a result,

$$\begin{aligned}
[T^*(g_i)](\boldsymbol{v}_j) &= g_i(T(\boldsymbol{v}_j)) \\
&= g_i\left(\sum_{\ell=1}^m \alpha_{\ell j}\boldsymbol{w}_\ell\right) \\
&= \sum_{\ell=1}^m \alpha_{\ell j} g_i(\boldsymbol{w}_\ell) \\
&= \alpha_{ij}
\end{aligned}$$

and

$$\begin{aligned}
[T^*(g_i)](\boldsymbol{v}_j) &= \left(\sum_{k=1}^n \beta_{ki} f_k\right)(\boldsymbol{v}_j) \\
&= \sum_{k=1}^n \beta_{ki} f_k(\boldsymbol{v}_j) \\
&= \beta_{ji}
\end{aligned}$$

Therefore, $\beta_{ji} = \alpha_{ij}$. The proof is complete. ∎

5.4.2. Relationship between Annihilator and dual of quotient spaces

■ **Example 5.5** Consider the canonical projection mapping $\pi_W : V \to V/W$ with its **adjoint** mapping:

$$(\pi_W)^* : (V/W)^* \to V^*$$

The understanding of $(\pi_W)^*$ is as follows:

1. Take $h \in (V/W)^*$ and study $(\pi_W)^*(h) \in V^*$
2. Take $\boldsymbol{v} \in V$ and understand

$$[(\pi_W)^*(h)](\boldsymbol{v}) = h(\pi_W(\boldsymbol{v})) = h(\boldsymbol{v} + W)$$

(a) In particular, for all $\boldsymbol{w} \in W \leq V$, we have

$$[(\pi_W)^*(h)](\boldsymbol{w}) = h(\boldsymbol{w} + W) = h(\mathbf{0}_{V/W}) = \mathbf{0}_{\mathbb{F}}$$

Therefore,

$$(\pi_W)^*(h) \in \mathsf{Ann}(W).$$

i.e., $(\pi_W)^*$ is a mapping from $(V/W)^*$ to $\mathsf{Ann}(W)$.

(b) By proposition (5.5), π_W is surjective implies $(\pi_W)^*$ is injective.

Combining (a) and (b), it's clear that (i.e., left as homework problem)

$$\Phi \circ \pi_W^* = \mathrm{id}_{(V/W)^*} \text{ and } \pi_W^* \circ \Phi = \mathrm{id}_{\mathsf{Ann}(W)}$$

This relationship implies Φ is an isomorphism. ■

5.5. Wednesday for MAT3006

5.5.1. Remarks on Baire Category Theorem

Theorem 5.4 — Baire Category Theorem. If (X,d) is complete, and $E_i \subseteq X$ is nowhere dense for $i \in \mathbb{N}$, then

$$\bigcup_{i=1}^{\infty} \overline{E_i}$$

contains no open balls.

Definition 5.4 Let (X,d) be a complete metric space.

1. We say $S \subseteq X$ is **meager** if

$$S = \bigcup_{i=1}^{\infty} E_i, \qquad E_i \text{ is nowhere dense}$$

   In this case we say S is of **first category**.

2. $S' \subseteq X$ is **comeager** if

$$S' = X \setminus S, \qquad \text{where } S \text{ is meager}$$

For example, $\mathbb{Q} = \cup_{x \in \mathbb{Q}} \{x\}$ is megre; $\mathbb{R} \setminus \mathbb{Q}$ is comeager.

R

1. By the Baire Category Theorem, $\cup_{i=1}^{n} \bar{E}_i$ contains no open balls, i.e.,

$$S := \cup_{i=1}^{\infty} E_i \subseteq \cup_{i=1}^{\infty} \bar{E}_i$$

   contains no open balls.

2. S' is comeager implies S' is dense in X: for $\forall x \in X$ and $\forall n \in \mathbb{N}$, $B_{1/n}(x) \cap S'$ is non-empty, since otherwise $X \setminus S'$ contains a open ball, which is a contradiction. Therefore, $x \in \overline{S'}$.

Proposition 5.7 If a set S is meager, it cannot be comeager and vice versa.

Proof. Suppose on contrary that S is meager and comeager, then

$$S = \bigcup_{i=1}^{\infty} E_i, \quad E_i \text{ is nowhere dense}$$
$$X \setminus S = \bigcup_{j=1}^{\infty} F_j, \quad F_j \text{ is nowhere dense}$$

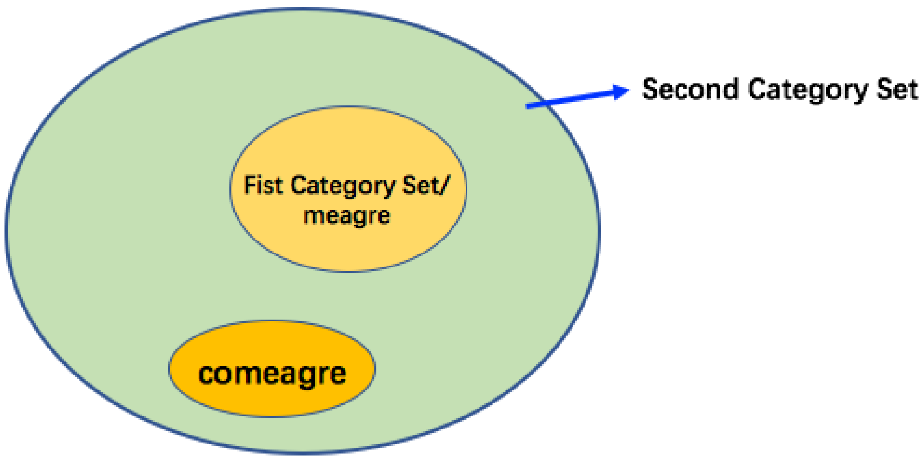
Therefore,

$$X = \bigcup_{i=1}^{\infty} E_i \cup \bigcup_{j=1}^{\infty} F_j$$

is a countable union of nowhere desne sets. By applying Baire Category Theorem, X has no open balls, which is a contradiction. ■

(R) We say $S \subseteq X$ is of **first category** if S is meager. Any subset that is not of first category is of **second category**. Therefore, comeager implies second category.

We illustrate the relationship above in the figure below:



Note that there are subsets that are **neither meager nor co-meager**.

■ **Example 5.6** 1. Here is another proof of $[0,1]$ is un-countable: Suppose on the contrary that $[0,1]$ is countable, then we imply

$$[0,1] = \bigcup_{n\in\mathbb{N}} \{x_n\}, \quad \text{for some } x_n.$$

Applying Baire Category Theorem (since $[0,1]$ is complete), $[0,1] = \cup_{n\in\mathbb{N}}\{x_n\}$ contains no open balls. However, the open ball $(0.5, 0.7) \subseteq [0,1]$, which is a contradiction.

2. The set $X := C[a,b]$ is complete.

   (a) The set of all nowhere differentiable functions is **of 2nd Category** in $C[a,b]$. (Check Theorem (4.1) in MAT2006) Actually, the set of all nowhere differentiable functions is **comeager**. The proof for this statement is omitted.

   (b) Due to the relationship

   $$\mathcal{P}[a,b] \subseteq C^{\infty}[a,b] \subseteq \{f : [a,b] \to \mathbb{R} \mid f \text{ is differentiable somewhere}\}$$

   and that the last subset is meager, we imply $\mathcal{P}[a,b]$ and $C^{\infty}[a,b]$ is meager.

■

5.5.2. Compact subsets of $C[a,b]$

Recall that for metrice spaces, the compactness implies closed and bounded, but in general the converse does not hold. We will study extra conditions to make subsets of $C[a,b]$ compact.

Definition 5.5 [(Uniformly) Bounded] The subset S in metric space $(C[a,b], d_{\infty})$ is **(uniformly) bounded** if there exists $M > 0$ such that

$$\sup_{f\in S} \|f\|_{\infty} = M$$

■

In next class, we will show that $K \subseteq C[a,b]$ is compact if and only if K is closed,(uniformly) bounded, and equi-continuous.

5.6. Wednesday for MAT4002

5.6.1. Remarks on Compactness

Theorem 5.5 X is compact, Y is Hausdorff, $f : X \to Y$ is continuous and bijective. Then X is **homeomorphic** to Y

Corollary 5.3 If X is compact, Y is Hausdorff, $f : X \to Y$ is injective and continous, then $f : X \to f(X)$ is homeomorphisc.

■ **Example 5.7** Here we give another proof for the fact that $S^1 \times S^1$ is homeomorphic to donut. Construct the mapping

$$\begin{aligned} f : \quad & S^1 \times S^1 \to \mathbb{R}^3 \\ \text{with} \quad & (e^{i\theta}, e^{i\phi}) \mapsto ((R + r\cos\theta)\cos\phi, (R + r\cos\theta)\sin\phi, r\sin\theta) \quad (R > r > 0) \end{aligned}$$

Note that:

- $X = S^1 \times S^1$ is compact, $\mathbb{R}^3$ is Hausdorff;
- f is continuous and injective.
- $f(S^1 \times S^1)$ is a "donut".

Therefore, we conclude that $S^1 \times S^1$ is homeomorphic to donut in $\mathbb{R}^3$. ■

Definition 5.6 [Sequential Compactness] A topological space X is **sequentially compact** if every sequence in X has a convergent sub-sequence. ■

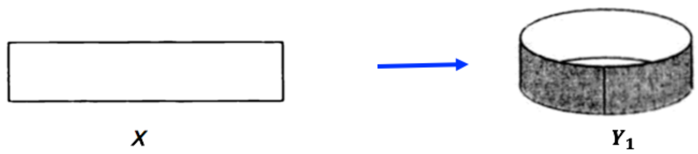
In $\mathbb{R}^n$, the compactness is equivalent to sequential compactness. The same goes for any metric space (X, d). (Check notes for MAT3006)

However, compactness and sequential compactness is different for topological spaces in general.

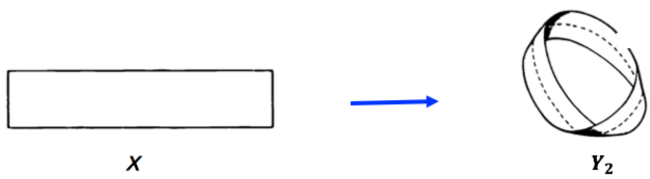
5.6.2. Quotient Spaces

Motivation. Just like product space and disjoint union, we give another way to construct new topological spaces from some old ones. This new way of construction is by gluing some special pieces from old topological spaces together.

Idea. Let $X = [0,1] \times [0,1]$ (just like a paper on a plane), we want to glue the leftmost edge with the rightmost edge to form a cylinder Y_1, as shown below:



If we give a half-twist to the strip before glue the ends together, we will get the **Moebius stripe** Y_2 shown below:



Interestingly, the first topology Y_1 has two sides, while the second has only one side.

5.6.2.1. Equivalence Relations and partitions

Definition 5.7 [Equivalence Relation] The **equivalence** relation on a set X is a relation $\sim$ such that

1. (Reflexive): $x \sim x, \forall x \in X$
2. (Symmetric): $x \sim y$ implies $y \sim x$
3. (Transitive): $x \sim y$ and $y \sim z$ implies $x \sim z$.

■ **Example 5.8**

1. Let $X = V$ be a vector space, and $W \le V$ be a vector subspace. Define $\boldsymbol{v}_1 \sim \boldsymbol{v}_2$ if $\boldsymbol{v}_1 - \boldsymbol{v}_2 \in W$.

   (The well-definedness is left as exercise).

2. (Mobius Stripe): Let $X = [0,1] \times [0,1]$. We define $(x_1, y_1) \sim (x_2, y_2)$ if

   - $x_1 = x_2, y_1 = y_2$; (e.g., $(0.5, 0.6) \sim (0.5, 0.6)$) or
   - $x_1 = 0, x_2 = 1$, and $y_1 = 1 - y_2$ (e.g., $(0, 1/4) \sim (1, 3/4)$)
   - $x_1 = 1, x_2 = 0$, and $y_1 = 1 - y_2$ (e.g., $(1, 3/4) \sim (0, 1/4)$)

■

Definition 5.8 [Partition] Let X be a nonempty set. A **partition** $\mathcal{P} = \{p_i \mid i \in I\}$ of X is a collection of subsets such that

1. $P_i \subseteq X$ is non-empty
2. $P_i \cap P_j = \emptyset$ if $i \neq j$
3. $\bigcup_{i \in I} P_i = X$

■

R Given a partition $\mathcal{P} = \{p_i \mid i \in I\}$, we can define an equivalence relation $\sim$ on X by setting

$$x \sim y \quad \text{whenever } x, y \in p_i, \text{ for some } i \in I$$

For example, if $X = [0,1] \times [0,1]$, then

$$X = \{(x,y)\}_{x \in (0,1), y \in [0,1]} \cup \{(1,y), (0, 1-y)\}_{y \in [0,1]}$$

gives a partition on X. This gives the same equivalence relation as in part (2) in example (5.8).

Conversely, given an equivalence relation $\sim$, we could form a corresponding partition of X. This kind of partition is called the equivalence class:

Definition 5.9 [Equivalence Class] Let X be a set with equivalence relation $\sim$. The equivalence class of an element $x \in X$ is

$$[x] := \{y \in X \mid x \sim y\}.$$

Proposition 5.8 The collection of all $[x]$ in $X/\sim$ gives a partition on X.

Consider the equivalence class defined in part (1) in example (5.8). The equivalence class has the form

$$[\boldsymbol{v}] = \{\boldsymbol{u} \in V \mid \boldsymbol{v} - \boldsymbol{u} \in W\} := \boldsymbol{v} + W.$$

Therefore, the equivalence class is a generalization of the **coset** in linear algebra. Similarly, we define the set of generalized cosets as **quotient space**.

Definition 5.10 The collection of all equivalence classes is called the **quotient space**, denoted as $X/\sim$, i.e.,

$$X/\sim = \{[x] \mid x \in X\}.$$

Example 5.9

1. Consider part (1) in example (5.8) again. The quotient space $V/\sim$ reduces to the V/W in linear algebra:

$$V/\sim = \{[\boldsymbol{v}] \mid \boldsymbol{v} \in V\} = \{\boldsymbol{v} + W \mid \boldsymbol{v} \in V\} = V/W.$$

2. Consider part (2) in example (5.8) again. Then $X/\sim$ essentially forms the **Mobius band**, e.g.,

$$[(1/2, 1/2)] = \{x \mid (1/2, 1/2) \sim x\} = \{(1/2, 1/2)\}$$

$$[(1, 3/4)] = \{x \mid x \sim (1, 3/4)\} = \{(1, 3/4), (0, 1/4)\}$$

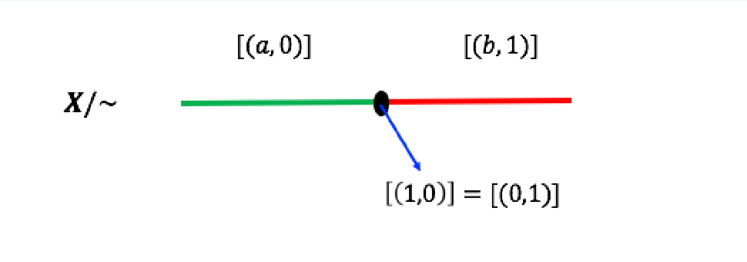
Example 5.10 Consider $X = [0,1] \sqcup [0,1]$, i.e.,

$$X = ([0,1] \times \{0\}) \cup ([0,1] \times \{1\})$$

Take a partition on X by

$$\{(a,0)\}_{0 \le a < 1} \cup \{(b,1)\}_{0 < b \le 1} \cup \{(1,0),(0,1)\}$$

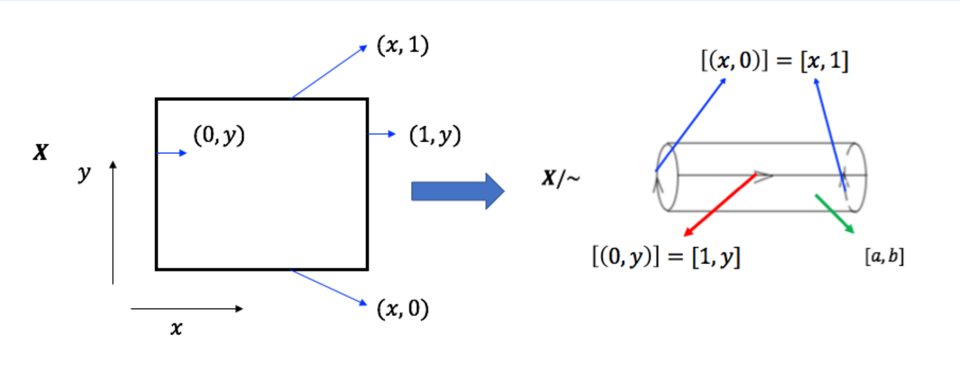
As a result, the corresponding quotient space is plotted below:



Example 5.11 Comes from $X = [0,1] \times [0,1]$ with partition

$$\{(a,b)\}_{0<a<1;0<b<1} \cup \{(x,0),(x,1)\}_{0 \le x \le 1} \cup \{(0,y),(1,y)\}_{0<y<1}$$

The corresponding quotient space is plotted below:



Proposition 5.9 Let $(X, \mathcal{T})$ be topological space, with the equivalence relation. Define the canonical projection map

$$\begin{aligned} p: \quad & X \to X/\sim \\ \text{with} \quad & x \mapsto [x] \end{aligned}$$

Define a collection of subsets $\tilde{\mathcal{T}}$ on $X/\sim$ by:

$$U \subseteq X/\sim \text{ is in } \tilde{\mathcal{T}} \text{ if } p^{-1}(U) \text{ is in } \mathcal{T}.$$

Then $\tilde{\mathcal{T}}$ is a topology for $X/\sim$, called **quotient topology**.

Chapter 6

Week6

6.1. Monday for MAT3040

6.1.1. Polynomials

We recall some useful properties of polynomial before studying eigenvalues/eigenvectors.

Definition 6.1 [Polynomial]

1. A polynomial over $\mathbb{F}$ has the form

$$p(z) = a_m z^m + \cdots + a_1 z + a_0, \quad (a_m \neq 0).$$

Here $a_m z^m$ is called the **leading term** of $p(z)$; m is called the degree; a_m is called the **leading coefficient**; $a_m, \cdots, a_0$ are called the coefficients of this polynomial.

2. A polynomial over $\mathbb{F}$ is monic if its leading coefficient is $1_{\mathbb{F}}$.
3. A polynomial $p(z) \in \mathbb{F}[z]$ is **irreducible** if for any $a(z), b(z) \in \mathbb{F}[z]$,

$$p(z) = a(z)b(z) \implies \text{either } a(z) \text{ or } b(z) \text{ is a constant polynomial.}$$

Otherwise $p(z)$ is **reducible**.

■ **Example 6.1** For example, the polynomial $p(x) = x^2 + 1$ is irreducible over $\mathbb{R}$; but $p(x) = (x - i)(x + i)$ is **reducible** over $\mathbb{C}$. ■

Theorem 6.1 — **Division Theorem.** For all $p, q \in \mathbb{F}[z]$ such that $p \neq 0$, there exists unique $s, r \in \mathbb{F}[x]$ satisfying $\deg(r) < \deg(f)$, such that

$$p(z) = s(z) \cdot q(z) + r(z).$$

Here $r(z)$ is called the **remainder**.

■ **Example 6.2** Given $p(x) = x^4 + 1$ and $q(x) = x^2 + 1$, the junior school knowledge tells us that uniquely

$$x^4 + 1 = (x^2 - 1)(x^2 + 1) + 2.$$

■

Theorem 6.2 — **Root Theorem.** For $p(x) \in \mathbb{F}[x]$, and $\lambda \in \mathbb{F}$, $x - \lambda$ divides p if and only if $p(\lambda) = 0$.

Proof.

1. If $(x - \lambda)$ divides p, then $p = (x - \lambda)q$ for some $q \in \mathbb{F}[x]$. Thus clearly $p(\lambda) = 0$.
2. For the other direction, suppose that $p(\lambda) = 0$. By division theorem, there exists $s, r \in \mathbb{F}[x]$ such that

$$p = (x - \lambda)s + r \quad \text{with } \deg(r) < \deg(x - \lambda) = 1. \tag{6.1}$$

Therefore, the polynomial r must be constant.

Substituting λ into x both sides in (6.1), we have

$$0 = p(\lambda) = 0 \cdot s + r \implies r = 0.$$

Therefore, $p = (x - \lambda) \cdot s$, i.e., $(x - \lambda)$ divides p.

6.2. Monday for MAT3006

6.2.1. Compactness in Functional Space

In functional space, previous study have shown that closedness and boundedness is not equivalent to compactness. We need the equi-continuity to rescue the situation:

Definition 6.2 [Equi-continuity] Let $X \subseteq \mathbb{R}^n$. A subset $\mathcal{T} \subseteq C(X)$ is called **equi-continuous** if for any $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $d(x,y) < \delta, x, y \in X$

$$d_\infty(f(x), f(y)) < \varepsilon, \forall f \in \mathcal{T}$$

■ **Example 6.3** 1. Let $\mathcal{T}$ be a collection of Lipschitz continuous functions with the same Lipschitz constant L, i.e., $\forall f \in \mathcal{T}$, $|f(x) - f(y)| < L|x - y|$ for $\forall x, y \in X$. It's clear that $\mathcal{T}$ is equi-continuous.

2. Let $\mathcal{T} \subseteq C[a,b]$ be such that

$$\sup_{x \in [a,b]} |f'(x)| < M, \quad \forall f \in \mathcal{T},$$

then for any $\forall x, y \in [a,b]$, we imply $|f(y) - f(x)| = |f'(\xi)||y - x|$ for some $\xi \in [a,b]$. Therefore,

$$|f(y) - f(x)| < M|y - x|, \quad \forall f \in \mathcal{T},$$

i.e., $\mathcal{T}$ reduces to the space studied in (1) with Lipschitz constant M, thus is equi-continuous.

Theorem 6.3 Let $K \subseteq \mathbb{R}^n$ be a compact set, and $\mathcal{T} \subseteq C(K)$. Then $\mathcal{T}$ is **compact** if and only if $\mathcal{T}$ is **closed**, **uniformly bounded**, and **equicontinuous**.

Proof. To be added. ■

Corollary 6.1 Let $K \subseteq \mathbb{R}^n$ be compact, and $\{f_n\}$ be a sequence of uniformly bounded, equi-continuous functions on K. Then $\{f_n\}$ has the **Bolzano-Weierstrass property**, i.e., it has a convergent subsequence.

Proof. To be added. ∎

6.2.2. An Application of Ascoli-Arzela Theorem

The Ascoli-Arzela Theorem has a novel application on the ODE. Consider the IVP problem again:

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x,y) \\ y(\alpha) = \beta \end{cases} \tag{6.2}$$

where f is continuous on a rectangle R containing (α, β). Now we show the existence of Picard-Lindelof Theorem without the Lipschitz condition:

Theorem 6.4 — Cauchy-Peano Theorem. Consider the problem (6.3). Then there exists a solution of this ODE on some rectangle $R' \subseteq R$.

Proof. To be added. ∎

6.3. Monday for MAT4002

6.3.1. Quotient Topology

Now given a topologcal space X and an equivalence relation $\sim$ on it, our goal is to construct a topology on the space $X/\sim$.

Proposition 6.1 Suppose $(X,\mathcal{T})$ is a topological space, and $\sim$ is an equivalene relation on X. Define the canonical projection map:

$$\begin{aligned} p: \quad & X \to X/\sim \\ \text{with} \quad & x \to [x] \end{aligned},$$

which assigns each point $x \in X$ into the equivalence class $[x]$. Then define a family of subsets $\tilde{\mathcal{T}}$ on $X/\sim$ by:

$$\tilde{U} \subseteq X/\sim \text{ is in } \tilde{\mathcal{T}} \text{ if } p^{-1}(\tilde{U}) \text{ is in } \mathcal{T}$$

Then $\tilde{\mathcal{T}}$ is a topology for $X/\sim$, called the **quotient topology**, and $(X/\sim,\tilde{\mathcal{T}})$ is called the quotient space, and $p: X \to X/\sim$ is called the **natural map**.

Proof. 1. $p^{-1}(X/\sim) = X \in \mathcal{T}$ and $p^{-1}(\emptyset) = \emptyset \in \mathcal{T}$, which implies $X/\sim \in \tilde{\mathcal{T}}$ and $\emptyset \in \tilde{\mathcal{T}}$.

2. Suppose that $\tilde{U},\tilde{V} \in \tilde{\mathcal{T}}$, then we imply

$$p^{-1}(\tilde{U}),p^{-1}(\tilde{V}) \in \mathcal{T} \implies p^{-1}(\tilde{U}\cap\tilde{V}) \in \mathcal{T},$$

i.e., $\tilde{U}\cap\tilde{V} \in \tilde{\mathcal{T}}$.

3. Following the similar argument in (2), and the relation

$$p^{-1}\left(\bigcup \tilde{U}_i\right) = \bigcup p^{-1}(\tilde{U}_i),$$

we conclude that $\tilde{T}$ is closed under countably union.

The proof is complete. ■

R

1. The proposition (6.1) claims that $\tilde{U}$ is open in $X/\sim$ iff $p^{-1}(\tilde{U})$ is open in X. The general question is that, does $p(U)$ is open in $X/\sim$, given that U is open in X? This may not necessarily hold. (See example (6.4)) In general $p^{-1}(p(U))$ is strictly larger than U, and may not be necessarily open in X, even when U is open.
2. By definition, we can show that p is continuous.

To fill the gap on the question shown in the remark, we consider the notion of the open mapping:

Definition 6.3 [Open Mapping] A function $f : X \to Y$ between two topological spaces is an **open mapping** if for each open U in X, $f(U)$ is open in Y.

R From the remark above, we can see that:

1. Not every continuous mapping is an open mapping
2. The canonical projection mapping p is not necessarily be an open mapping.

Example 6.4 1. The mapping $p : [0,1] \times [0,1] \to ([0,1] \times [0,1])/\sim$ sending the square to the Mobius band M is not an open mapping:

Consider the open ball $U = B_{1/2}((0,0))$ in $[0,1] \times [0,1]$. Note that $p(U)$ is open in M iff $p^{-1}(p(U))$ is open in $[0,1] \times [0,1]$. We can calculate $p^{-1}(p(U))$ explicitly:

$$p^{-1}(p(U)) = U \cup \{(1, y) \mid 1/2 \le y \le 1\},$$

which is not open.

6.3.2. Properties in quotient spaces

6.3.2.1. Closedness on $X/\sim$

Proposition 6.2 A subset $\tilde{V}$ is closed in the quotient space $X/\sim$ iff $p^1(\tilde{V})$ is closed in X, where $p : X \to X/\sim$ denotes the canonical projection mapping.

Proof. It follows from the fact that

$$p^{-1}\left((X/\sim)\setminus\tilde{V}\right) = X\setminus p^{-1}(\tilde{V})$$

■

6.3.2.2. Isomorphism on $X/\sim$

The quotient space can be used to study other type of spaces:

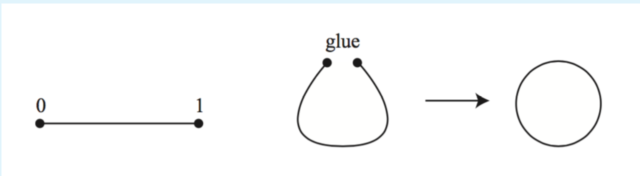
▪ **Example 6.5** Consider $X = [0,1]$. We define $x_1 \sim x_2$ if:

$$x_1 = 0, x_2 = 1, \quad \text{or} \quad x_1 = 1, x_2 = 0$$

In other words, the partition on X is given by:

$$X = \{0,1\} \cup \Big(\bigcup_{x\in(0,1)}\{x\}\Big)$$

The quotient space seems "glue" the endpoints of the interval $[0,1]$ together, shown in the figure below:



It is intuitive that the constructed quotient space should be homeomorphic to a circle S^1. We will give a formal proof on this fact. ▪

Proposition 6.3 Let X and Z be topological spaces, and $\sim$ an equivalence relation on X. Let $g : X/\sim \to Z$ be a function, and $p : X \to X/\sim$ is a projection mapping The mapping g is continuous if and only if $g \circ p : X \to Z$ is continuous.

Proof.
1. *Necessity*. Suppose that g is continuous. It's clear that p is continuos, i.e, $g \circ p : X \to Z$ is continuous.
2. *Sufficiency*. Suppose that $g \circ p : X \to Z$ is continuous. Given any open U in Z, we imply $(g \circ p)^{-1}(U) = p^{-1}g^{-1}(U)$ is open in X. By definition of the quotient topology, we imply $g^{-1}(U)$ is open in $X/\sim$. Therefore, g is continuous.

■

R This useful lemma can be generalized into the case for generlized canonical projection mapping, called quotient mapping.

> **Definition 6.4** [Quotient mapping] A map $p : X \to Y$ between topological spaces is a **quotient mapping** if
>
> 1. p is **surjective**; and
> 2. p is continuous;
> 3. For any $U \subseteq Y$ such that $p^{-1}(U)$ is open in X, we imply U is open in Y.

The canonical projection map is clearly a quotient map. Actually, a stronger version of proposition (6.3) follows:

Proposition 6.4 Suppose that $p : X \to Y$ is a quotient map and that $g : Y \to Z$ is any mapping to another space Z. Then g is continuous iff $g \circ p$ is continuous.

Proof. The proof follows similarly as in proposition (6.3). ■

Now we give a formal proof of the conclusion in the example (6.5):

Proof. Define the mapping

$$\begin{aligned} f: &\quad [0,1] \to S^1 \\ \text{with} &\quad t \mapsto (\cos 2\pi t, \sin 2\pi t). \end{aligned}$$

Since $f(0) = f(1)$, the function f induces a well-defined function

$$\begin{aligned} g: &\quad [0,1]/\sim \longrightarrow S^1 \\ \text{with} &\quad [t] \mapsto f(t) \end{aligned}$$

such that $f = g \circ p$, where p denotes the canonical projection mapping. Note that f is continuous. By proposition (6.3), we imply g is continuous. Furthermore,

1. Since $[0,1]$ is compact and p is continuous, we imply $p([0,1]) = [0,1]/\sim$ is compact
2. S^1 is Hausdorff
3. g is a bijection

By applying theorem(5.3), we conclude that g is a homeomorphism, i.e., $[0,1]/\sim$ and S^1 are homeomorphic.

∎

The argument in the proof can be generalized into the proposition below:

Proposition 6.5 Let $f: X \to Y$ be a surjective continuous mapping between topologcial spaces. Let $\sim$ be the equivalence relation on X defined by the partition $\{f^{-1}(y) \mid y \in Y\}$ (i.e., $f(x) = (x')$ iff $x \sim x'$). If X is compact and Y is Hausdorff, then $X/\sim$ and Y are homeomorphic.

R The proposition (6.5) is a pattern of argument we should use several times. In order to show $X/\sim$ and Y are homeomorphic, we should think up a surjective continuous mapping $f: X \to Y$ "with respect to the identifications", i.e., $f(x_1) = f(x_2)$ whenever $x_1 \sim x_2$. Therefore f will induce a well-defined function $g: X/\sim \longrightarrow Y$ such that $f = g \circ f$. Then checking the conditions in theorem(5.3) leads to the desired results.

Torus. We now study the torus in more detail.

1. Consider $X = [0,1] \times [0,1]$ and define $(s_1, t_1) \sim (s_2, t_2)$ if one of the following holds:

   - $s_1 = s_2$ and $t_1 = t_2$;
   - $\{s_1, s_2\} = \{0,1\}$, $t_1 = t_2$;
   - $\{t_1, t_2\} = \{0,1\}$ and $s_1 = s_2$;
   - $\{s_1, s_2\} = \{0,1\}$, $\{t_1, t_2\} = \{0,1\}$

   The corresponding quotient space $([0,1] \times [0,1])/\sim$ is hoemomorhpic to the 2-dimension torus $\mathbb{T}^2$.

   Proof. Define the mapping $f : [0,1] \times [0,1] \to \mathbb{T}^2$ as $(t_1, t_2) \mapsto (e^{2\pi i t_1}, e^{2\pi i t_2})$.

   (a) f is surjective, which also implies $\mathbb{T}^2 = f([0,1] \times [0,1])$ is compact.

   (b) $\mathbb{T}^2$ is Hausdorff

   (c) It's clear that $(s_1, t_1) \sim (s_2, t_2)$ implies $f(s_1, t_1) = f(s_2, t_2)$. Conversely, suppose

   $$e^{2\pi i s_1} = e^{2\pi i s_2}, \quad e^{2\pi i t_1} = e^{2\pi i t_2}$$

   By the familiar property of e^{ix}, we imply either $t_1 = t_2$ or $\{t_1, t_2\} = \{0,1\}$; and either $s_1 = s_2$ or $\{s_1, s_2\} = \{0,1\}$

   By applying proposition (6.5), we conclude that $([0,1] \times [0,1])/\sim$ is homeomorphic to $\mathbb{T}^2$. ■

2. Consider the closed disk $\mathbb{D}^2 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 1\}$, and defube $(x_1, y_1) \sim (x_2, y_2)$ if one of the following holds:

   - $x_1 = x_2$ and $y_1 = y_2$;
   - (x_1, y_1) and (x_2, y_2) are in the boundary circle $\mathbb{S}^1$

   The corresponding quotient space $\mathbb{D}^2/\sim$ is hoemomorhpic to the 2-dimension sphere $\mathbb{S}^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$.

Proof. Define the mapping

$$\begin{aligned}
f: \quad & \mathbb{D}^2 \to \mathbb{S}^2 \\
\text{with} \quad & (0,0) \mapsto (0,0,1) \\
& (x,y) \mapsto \left(\frac{x}{\sqrt{x^2+y^2}} \sin(\pi\sqrt{x^2+y^2}), \frac{y}{\sqrt{x^2+y^2}} \sin(\pi\sqrt{x^2+y^2}), \cos(\pi\sqrt{x^2+y^2}) \right)
\end{aligned}$$

It's easy to check the conditions in proposition (6.5), and we conclude that $\mathbb{D}^2/\sim$ is hoemomorhpic to $\mathbb{S}^2$ ■

6.4. Wednesday for MAT3040

Reviewing: Root Theorem: $p(\lambda) = 0$ iff $(x - \lambda)$ divdes $p(x)$.

Corollary 6.2 A polynomial with degree n has at most n roots counting multiplicity.

For example, the polynomial $(x-3)^2$ has one root $x = 3$ with multiplicity 2. When counting multiplicity, we say the polynomial $(x-3)^2$ has two roots.

Definition 6.5 [Algebraically Closed] A field $\mathbb{F}$ is called **algebraically closed** if every non-constant polynomial $p(x) \in \mathbb{F}[x]$ has a root $\lambda \in \mathbb{F}$.

Theorem 6.5 — Fundamental Theroem of Algebra. The set of complex numbers $\mathbb{C}$ is algebraically closed.

Proof. One way is by complex analysis; Another way is by the topology on $\mathbb{C} \setminus \{0\}$. ∎

R By induction, we can show that every polynomial with degree n on algebraically closed field $\mathbb{F}$ has **exactly** n roots, counting multiplicity. Therefore, for any $p(x)$ on algebraically closed field $\mathbb{F}$,

$$p(x) = c(x - \lambda_1)\cdots(x - \lambda_n) \tag{6.3}$$

for $c, \lambda_1, \dots, \lambda_n \in \mathbb{F}$.

The polynomials on general field $\mathbb{F}$ may not necessarily be factorized as in (6.3) , but still admit unique factorization property:

Theorem 6.6 — Unique Factorization. Every $f(x) = a_n x^n + \cdots + a_0$ in $\mathbb{F}[x]$ can be factorized as

$$f(x) = a_n [p_1(x)]^{e_1} \cdots [p_k(x)]^{e_k}$$

where p_i's are **monic, irreducible,distinct**. Furthermore, this expression is unique up to the permutation of factors.

Definition 6.6 [Factor] If $p(x) = q(x)s(x)$ with $p, q, s \in \mathbb{F}[x]$, then we say

- $p(x)$ is **divisible** by $s(x)$;
- $s(x)$ is a **factor** of $p(x)$;
- $s(x) | p(x)$
- $s(x)$ **divides** $p(x)$
- $p(x)$ is **multiple** of $s(x)$

Definition 6.7 [Common Factor]

1. The polynomial $g(x)$ is said to be a **common factor** of $f_1, \dots, f_k \in \mathbb{F}[x]$ if

$$g | f_i, i = 1, \dots, k$$

2. The polynomial $g(x)$ is said to be a **greatest common divisor** of $f_1, \dots, f_k$ if
   - g is **monic**.
   - g is common factor of $f_1, \dots, f_k$
   - g is of largest possible (maximal) degree.

R

- $\gcd(f_1, \dots, f_k) = \gcd(\gcd(f_1, f_2), f_3, \dots, f_k) = \gcd(\gcd(f_1, f_2, f_3), \dots, f_k)$
- $\gcd(f_1, \dots, f_k)$ is unique.
- If $\gcd(f_1, \dots, f_k) = 1$, we say $f_1, \dots, f_k$ is **relatively prime**
- Polynomials $f_1, \dots, f_k$ are relatively prime does not necessarily mean $\gcd(f_i, f_j) = 1$ for any $i \neq j$.

Counter-example: Let $a_1, \ldots, a_n$ distinct irreducible polynomials, and

$$f_i(x) = a_1(x) \cdots \hat{a}_i(x) \cdots a_n(x) := a_1 \cdots a_{i-1} a_{i+1} \cdots a_n,$$

then $\gcd(f_1, \ldots, f_n) = 1$, but $\gcd(f_i, f_j) = a_1 \cdots \hat{a}_i \cdots \hat{a}_j \cdots a_n$, which does not necessarily equal to 1.

■ **Example 6.6** The $\gcd(f_1, f_2)$ is easy to compute for factorized polynomials. For example, let $f_1(x) = (x^2 + x + 1)^3 (x-3)^2 x^4$ and $f_2(x) = (x^2+1)(x-3)^4 x^2$ in $\mathbb{R}[x]$, then

$$\gcd(f_1, f_2) = (x-3)^2 x^2$$

■

The question is how to find $\gcd(f_1, f_2)$ for given un-factorized polynomials?

Theorem 6.7 — **Bezout.** Let $g = \gcd(f_1, f_2)$, then there exists $r_1, r_2 \in \mathbb{F}[x]$ such that

$$g(x) = r_1(x) f_1(x) + r_2(x) f_2(x)$$

More generally, $g = \gcd(f_1, \ldots, f_k)$ implies there exists $r_1, \ldots, r_k$ such that

$$g = r_1 f_1 + \cdots + r_k f_k$$

The derivation of r_i's is by applying **Euclidean algorithm**. For example, given $x^3 + 6x + 7$ and $x^2 + 3x + 2$, we imply

$$x^3 + 6x + 7 - (x-3)(x^2 + 3x + 2) = 13x + 13$$

and

$$x^2 + 3x + 2 - \frac{x+2}{13}(13x + 13) = 0$$

Therefore, $\gcd(x^3 + 6x + 7, x^2 + 3x + 2) = \gcd(x^2 + 3x + 2, 13x + 13) = x + 2$.

6.4.1. Eigenvalues & Eigenvectors

Definition 6.8 [Eigenvalues] Let $T: V \to V$ be a linear operator.

1. We say $\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}$ is an eigenvector of T with eigenvalue λ if $T(\boldsymbol{v}) = \lambda \boldsymbol{v}$;
2. Or equivalently, $\boldsymbol{v} \in \ker(T - \lambda I)$, the λ-eigenspace of T. Here the mapping $I: V \to V$ denotes identity map, i.e., $I(\boldsymbol{v}) = \boldsymbol{v}, \forall \boldsymbol{v} \in V$.

Definition 6.9 A vector $\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}$ is a **generalized eigenvector** of T with **generalized eigenvalue** λ if $\boldsymbol{v} \in \ker((T - \lambda I)^k)$ for some $k \in \mathbb{N}^+$.

Note that an eigenvector is a generalized eigenvector of T; while the converse does not necessarily hold.

Example 6.7 Consider the linear transformation $A: \mathbb{R}^2 \to \mathbb{R}^2$ with

$$\begin{aligned} A: &\quad \mathbb{R}^2 \to \mathbb{R}^2 \\ \text{with} &\quad \boldsymbol{x} \to \boldsymbol{A}\boldsymbol{x} \\ \text{where} &\quad \boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

1. Note that $[1,0]^{\mathrm{T}}$ is an eigenvector with eigenvalue 1, since

$$A\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

2. However, $[0,1]^{\mathrm{T}}$ is not an eigenvector, since

$$A\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Note that

$$(A-I)^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (A-I)^3 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and therefore

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \ker(A-I)^2,$$

i.e., a generalized eigenvector with generalized eigenvalue 1.

■

■ **Example 6.8** Consider $V = C^\infty(\mathbb{R})$, which is a set of all infinitely differentiable functions.

Define the linear operator $T : V \to V$ as $T(f) = f''$. Then the (-1)-eigenspace of T has $f \in V$ satisfying

$$f'' = -f$$

From ODE course, we imply $\{\sin x, \cos x\}$ forms a basis of (-1)-eigenspace. ■

Assumption. From now on, we assume V has finite dimension by default.

Definition 6.10 [Determinant] Let $T : V \to V$ be a linear operator. The **determinant** of T is given by

$$\det(T) = \det((T)_{\mathcal{A},\mathcal{A}})$$

where $\mathcal{A}$ is some basis of V. ■

R Assume we have complete knowledge about $\det(M)$ for matrices for now. The determinant is well-defined, i.e., independent of the choice of basis $\mathcal{A}$. For another basis $\mathcal{B}$, we imply

$$\det(T_{\mathcal{B},\mathcal{B}}) = \det(C_{\mathcal{B},\mathcal{A}} T_{\mathcal{A},\mathcal{A}} C_{\mathcal{A},\mathcal{B}}) = \det(C_{\mathcal{B},\mathcal{A}}) \det(T_{\mathcal{A},\mathcal{A}}) \det(C_{\mathcal{A},\mathcal{B}}) = \det(T_{\mathcal{A},\mathcal{A}})$$

Definition 6.11 [characteristic polynomial] The **characteristic polynomial** $\mathcal{X}_T(x)$ of $T : V \longrightarrow V$ is defined as

$$\mathcal{X}_T(x) = \det((T)_{\mathcal{A},\mathcal{A}} - xI)$$

for any basis $\mathcal{A}$

In the next few lectures, we will study

- Cayley-Hamilton Theorem
- Jordan Canonical Form

These theorems can be stated using matrices, and they both hold up to change of basis. We have a unified statement of these theorem using vecotor space rather than $\mathbb{R}^n$.

6.5. Wednesday for MAT4002

6.5.1. Remarks on Compactness

The image of the quotient mapping is intrinsically the same as the quotient topology:

Proposition 6.6 Suppose that $q : X \to Y$ is a quotient map and that $\sim$ is the equivalence relation on X corresponding to the partition $\{q^{-1}(y) : y \in Y\}$. Then $X/\sim$ and Y are homeomorphic.

Proof. Construct the mapping

$$\begin{aligned} g: \quad & X/\sim \longrightarrow Y \\ \text{with} \quad & g([x]) = q(x). \end{aligned}$$

It suffices to show g is a homeomorphism:

1. The mapping g is surjective since q is surjective. The mapping g is clearly a injective.
2. Construct the canonical projection mapping $p : X \to X/\sim$. Therefore $g \circ p = q$ is continuous.

■

Chapter 7

Week7

7.1. Monday for MAT3040

Reviewing. Define the characteristic polynomial for an linear operator T:

$$\mathcal{X}_T(x) = \det((T)_{\mathcal{A},\mathcal{A}} - x\boldsymbol{I})$$

We will use the notation "$I/\boldsymbol{I}$" in two different occasions:

1. I denotes the identity transformation from V to V with $I(\boldsymbol{v}) = \boldsymbol{v}, \forall \boldsymbol{v} \in V$
2. $\boldsymbol{I}$ denotes the identity matrix $(I)_{\mathcal{A},\mathcal{A}}$, defined based on any basis $\mathcal{A}$.

7.1.1. Minimal Polynomial

Definition 7.1 [Linear Operator Induced From Polynomial] Let $f(x) := a_m x^m + \cdots + a_0$ be a polynomial in $\mathbb{F}[x]$, and $T : V \to V$ be a linear operator. Then the mapping

$$f(T) = a_m T^m + \cdots + a_1 T + a_0 I : \quad V \to V,$$

is called a linear operator induced from the polynomial $f(x)$.

R

1. The composition of linear operators is not abelian, e.g., in general $S \circ T = T \circ S$ does not hold. The reason follows similarly from the fact that square-matrix multiplication is not abelian in general.

2. However, we always have $f(T)T = Tf(T)$, where $f(T)$ is a linear operator induced from the polynomial $f(x)$:

Proof. We can show that $T^nT = TT^n, \forall n$ by induction. Suppose that $f(x) = \sum_i a_i x^i$, which follows that

$$f(T)T = \sum_i a_i T^i T = \sum_i a_i T T^i = T \sum_i a_i T^i = Tf(T).$$

■

3. We can generalize the statement in (2) into the fact that the composition of linear operators induced from polynomials is abelian, i.e.,

$$f(T)g(T) = g(T)f(T)$$

for any polynomials $f(x), g(x)$.

Definition 7.2 [Minimal Polynomial] Let $T : V \to V$ be a linear operator. The **minimal polynomial** $m_T(x)$ is a **nonzero monic polynomial** of least (minimal) degree such that

$$m_T(T) = \mathbf{0}_{V \to V}.$$

where $\mathbf{0}_{V \to V}$ denotes the zero vector in $\text{Hom}(V, V)$. ■

■ **Example 7.1** 1. Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then $\boldsymbol{A}$ defines a linear operator:

$$\begin{aligned} A : \quad & \mathbb{F}^2 \to \mathbb{F}^2 \\ \text{with} \quad & \boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} \end{aligned}$$

Here $\mathcal{X}_A(x) = (x-1)^2$ and $\boldsymbol{A} - \boldsymbol{I} = \boldsymbol{0}$, which gives $m_A(x) = x - 1$.

2. Let $\boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, which implies

$$\mathcal{X}_B(x) = (x-1)^2,$$

The question is that can we get the minimal polynomial with degree 1?

The answer is no, since $\boldsymbol{B} - k\boldsymbol{I} = \begin{pmatrix} 1-k & 1 \\ 0 & 1-k \end{pmatrix} \neq \boldsymbol{0}$.

In fact, $m_B(x) = (x-1)^2$, since

$$(\boldsymbol{B} - \boldsymbol{I})^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Two questions naturally arises:

1. Does $m_T(x)$ exist? If exists, is it unique?
2. What's the relationship between $m_T(x)$ and $\mathcal{X}_T(x)$?

Regarding to the first question, the minimal polynomial $m_T(x)$ may not exist, if V has infinite dimension:

Example 7.2 Consider $V = \mathbb{R}[x]$ and the mapping

$$\begin{aligned} T: \quad & V \to V \\ & p(x) \mapsto \int_0^x p(t)\,\mathrm{d}t \end{aligned}$$

In particular, $T(x^n) = \frac{1}{n+1}x^{n+1}$. Suppose $m_T(x)$ is with degree n, i.e.,

$$m_T(x) = x^n + \cdots + a_1 x + a_0,$$

then

$$m_T(T) = T^n + \cdots + a_0 I \text{ is a zero linear transformation}$$

It follows that

$$[m_T(T)](x) = \frac{1}{n!}x^n + a_{n-1}\frac{1}{(n-1)!}x^{n-1} + \cdots + a_1 x + a_0 = 0_{\mathbb{F}},$$

which is a contradiction since the coefficients of x^k is nonzero on LHS for $k = 1, \ldots, n$, but zero on the RHS. ■

Proposition 7.1 The minimal polynomial $m_T(x)$ always exists for $\dim(V) = n < \infty$.

Proof. It's clear that $\{I, T, \ldots, T^n, T^{n+1}, \cdots, T^{n^2}\} \subseteq \mathrm{Hom}(V,V)$. Since $\dim(\mathrm{Hom}(V,V)) = n^2$, we imply $\{I, T, \ldots, T^n, T^{n+1}, \cdots, T^{n^2}\}$ is linearly dependent, i.e., there exists a_i's that are not all zero such that

$$a_0 I + a_1 T + \cdots + a_{n^2} T^{n^2} = 0$$

i.e., there is a polynomial $g(x)$ of degree less than n^2 such that $g(T) = 0$.

The proof is complete. ■

Proposition 7.2 The minimal polynomial $m_T(x)$, if exists, then it exists uniquely.

Proof. Suppose f_1, f_2 are two distinct minimal polynomials with $\deg(f_1) = \deg(f_2)$. It follows that

- $\deg(f_1 - f_2) < \deg(f_1)$.
- $f_1 - f_2 \neq 0$
- $(f_1 - f_2)(T) = f_1(T) - f_2(T) = 0_{V \to V}$

By scaling $f_1 - f_2$, there is a monic polynomial g with lower degree satisfying $g(T) = 0$, which contradicts the definition for minimal polynomial. ■

Proposition 7.3 Suppose $f(x) \in \mathbb{F}[x]$ satisfying $f(T) = \mathbf{0}$, then

$$m_T(x) \mid f(x).$$

Proof. It's clear that $\deg(f) \geq \deg(m_T)$. The division algorithm gives

$$f(x) = q(x)m_T(x) + r(x).$$

Therefore, for any $\boldsymbol{v} \in V$

$$[r(T)](\boldsymbol{v}) = [f(T)](\boldsymbol{v}) - [q(T)m_T(T)](\boldsymbol{v}) = \mathbf{0}_V - q(T)\mathbf{0}_V = \mathbf{0}_V - \mathbf{0}_V = \mathbf{0}_V$$

Therefore, $r(T) = \mathbf{0}_{V \to V}$. By definition of minimal polynomial, we imply $r(x) \equiv 0$. ■

Proposition 7.4 If $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{F}^{n \times n}$ are similar to each other, then $m_A(x) = m_B(x)$.

Proof. Suppose that $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, and that

$$m_A(x) = x^k + \cdots + a_1 x + a_0, \quad m_B(x) = x^\ell + \cdots + b_0.$$

It follows that

$$\begin{aligned} m_A(\boldsymbol{B}) &= \boldsymbol{B}^k + \cdots + a_0 I \\ &= \boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P} + \cdots + a_0\boldsymbol{P}^{-1}\boldsymbol{P} \\ &= \boldsymbol{P}^{-1}(\boldsymbol{A}^k + \cdots + a_0\boldsymbol{I})\boldsymbol{P} \\ &= \boldsymbol{P}^{-1}(m_A(\boldsymbol{A}))\boldsymbol{P} \end{aligned}$$

Therefore, $m_A(\boldsymbol{B}) = \mathbf{0}$ since $m_A(\boldsymbol{A}) = \mathbf{0}$. By proposition (7.3), we imply $m_B(x) \mid m_A(x)$. Similarly, $m_A(x) \mid m_B(x)$. Since $m_A(x)$ and $m_B(x)$ are monic, we imply $m_A(x) = m_B(x)$.

■

R Proposition (7.4) claims that the minimal polynomial is **similarity-invariant**; actually, the characteristic polynomial is **similarity-invariant** as well.

Assumption. We will asssume V has finite dimension from now on. Now we study the vanishing of a single vector $\boldsymbol{v} \in V$.

Notation. The $m_T(x)$ is a nonzero monic poylnomial of least degree such that

$$m_T(T) = \mathbf{0}_{V \to V}.$$

7.1.2. Minimal Polynomial of a vector

Definition 7.3 [Minimal Polynomial of a vector] Similar to the minimal polynomial, we define the **minimal polynomial of a vector $\boldsymbol{v}$ relative to** T, say $m_{T,\boldsymbol{v}}(x)$, as the monic polynomial of least degree such that

$$m_{T,\boldsymbol{v}}(T)(\boldsymbol{v}) = 0$$

The existence of minimal polynomial of a vector is due to the existence of minimal polynomial; the uniqueness follows similarly as in proposition (7.2).

Proposition 7.5 Let $T : V \to V$ be a linear operator and $\boldsymbol{v} \in V$. The degree of the minimal polynomial of a vector is upper bounded by:

$$\deg(m_{T,\boldsymbol{v}}(x)) \le \dim(V).$$

Proof. It's clear that $\{\boldsymbol{v}, T\boldsymbol{v}, \dots, T^n\boldsymbol{v}\} \subseteq V$ and the proof follows similarly as in proposition (7.1). ∎

Similar to the division property in proposition (7.3), we have the division proprty for minimal polynomial of a vector:

Proposition 7.6 Suppose $f(x) \in \mathbb{F}[x]$ satisfying $f(T)(\boldsymbol{v}) = \mathbf{0}_V$, then

$$m_{T,\boldsymbol{v}}(x) \mid f(x).$$

In particular, $m_{T,\boldsymbol{v}} \mid m_T(x)$.

Proof. The proof follows similarly as in proposition (7.3). ∎

Proposition 7.7 Suppose that $m_{T,\boldsymbol{v}}(x) = f_1(x)f_2(x)$, where f_1, f_2 are both monic. Let $\boldsymbol{w} = f_1(T)\boldsymbol{v}$, then

$$m_{T,\boldsymbol{w}}(x) = f_2(x)$$

Proof. 1.

$$f_2(T)\boldsymbol{w} = f_2(T)f_1(T)\boldsymbol{v} = m_{T,\boldsymbol{v}}(T)\boldsymbol{v} = \mathbf{0}$$

By the proposition (7.3), we imply $m_{T,\boldsymbol{w}}|f_2$.

2. On the other hand,

$$\mathbf{0} = m_{T,\boldsymbol{w}}(T)(\boldsymbol{w}) = m_{T,\boldsymbol{w}}(T)f_1(T)\boldsymbol{v} = f_1(T)m_{T,\boldsymbol{w}}(T)\boldsymbol{v},$$

which implies that $m_{T,\boldsymbol{v}}(x) \mid f_1(x)m_{T,\boldsymbol{w}}(x)$,, i.e.,

$$f_1 \cdot f_2 \mid f_1 \cdot m_{T,\boldsymbol{w}} \implies f_2 \mid m_{T,\boldsymbol{w}}.$$

The proof is complete.

■

7.2. Monday for MAT3006

Our first mid-term will be held on this Wednesday.

Reviewing. In last lecture, we mainly talk about

- The extended real line
- Definition for limsup and liminf
- For interval I of the form (a,b), $[a,b)$, $(a,b]$ or $[a,b]$, we define

$$m(I) := b - a$$

- We constructed a kind of function to measure the length of a given subset $E \subseteq \mathbb{R}$:

$$m^*(E) = \inf\left\{\sum_{n=1}^{\infty} m(I_n) \middle| E \subseteq \bigcup_{n=1}^{\infty} I_n,\ I_n \text{ are open intervals}\right\}$$

which is called the **outer measure**.

7.2.1. Remarks on the outer measure

Proposition 7.8

1. $m^*(\phi) = 0$, $m^*(\{x\}) = 0$.
2. $m^*(E + x) = m^*(E)$
3. $m^*(I) = b - a$, where I denotes any interval with endpoints a or b.
4. If $A \subseteq B$, then $m^*(A) \le m^*(B)$
5. $m^*(kE) = |k| m^*(E)$
6. $m^*(\cup_{m=1}^{\infty} E_n) \le \sum_{n=1}^{\infty} m^*(E_n)$ for subsets $E_n \subseteq \mathbb{R}$

R The trick in the proof to show $x \le y$ is by the argument $x \le y + \varepsilon, \forall \varepsilon > 0$.

(1),(2),(5) is clear. (4) is by one-line argument:

Suppose that $B \subseteq \cup_{n=1}^{\infty} I_n$, then $A \subseteq \cup_{n=1}^{\infty} I_n$.

Proof for (3). Consider $m^*([a,b])$ first. The proof for $m^*([a,b]) \leq b-a$ is by explicitly constructing a sequence of open intervals:

$$[a,b] \subseteq (a-\frac{\varepsilon}{2}, b+\frac{\varepsilon}{2}) \cup (a,a) \cup \cdots$$

It follows that

$$\begin{aligned} m^*([a,b]) &\leq m((a-\frac{\varepsilon}{2}, b+\frac{\varepsilon}{2})) + 0 + \cdots + 0 \\ &= (b-a) + \varepsilon, \ \forall \varepsilon > 0 \end{aligned}$$

In particular, $m^*([a,b]) \leq b-a$.

Conversely, the proof for $b-a \leq m^*([a,b])$ is by implicitly constructing a sequence of open interval via the infimum. For all $\varepsilon > 0$, there exists $I_n,\ n \in \mathbb{N}$ such that

$$[a,b] \subseteq \cup_{n=1}^{\infty} I_n, \qquad \sum_{n=1}^{\infty} m(I_n) \leq m^*([a,b]) + \varepsilon.$$

By Heine-Borel Theorem, there exists finite subcover $[a,b] \subseteq \cup_{n=1}^{k} I_n$. Let $I_n = (\alpha_n, \beta_n)$, consider $\alpha := \min\{\alpha_n \mid a \in I_n\}$ and $\beta := \max\{\beta_n \mid b \in I_n\}$. Then we imply

$$[a,b] \subseteq (\alpha, \beta) \subseteq \cup_{n=1}^{k} I_n.$$

It's clear that $\beta - \alpha \leq \sum_{n=1}^{k} m(I_n)$, which follows that

$$b - a \leq \beta - \alpha \leq \sum_{n=1}^{k} m(I_n) \leq \sum_{n=1}^{\infty} m(I_n) \leq m^*([a,b]) + \varepsilon$$

The proof is complete.

The other cases of (3) follows similarly. For example, $m^*((a,b))$ can be lower bounded as:

$$m^*((a,b)) + \varepsilon \geq m^*([a+\frac{\varepsilon}{2}, b-\frac{\varepsilon}{2}]) + \varepsilon = b-a$$

■

Proof for (6). The case for which $m^*(E_n) = \infty$ for some n is trivial, since both sides clearly equal to infinite. Consider the case where $m^*(E_n) < \infty$ only.

By definition, for each E_n we can find $\{I_{n,k}\}_{k=1}^{\infty}$ such that

$$E_n \subseteq \cup_{k=1}^{\infty} I_{n,k}, \quad \sum_{k=1}^{\infty} m(I_{n,k}) \le m^*(E_n) + \frac{\varepsilon}{2^n}.$$

It follows that

- $\cup_{n=1}^{\infty} \cup_{k=1}^{\infty} I_{n,k}$ is a countable open cover of $\cup_{n=1}^{\infty} E_n$, i.e.,
$$m^*(\cup_{n=1}^{\infty} E_n) \le \sum_{n,k} m(I_{n,k})$$
- $$\sum_{n,k} m(I_{n,k}) \le \sum_{n=1}^{\infty} m^*(E_n) + \varepsilon$$

The proof is complete. ■

The natural question is that when does the equality in (6) holds? We will study it in next week.

Definition 7.4 [Null Set] The set $E \subseteq \mathbb{R}$ is a **null set** if $m^*(E) = 0$.

Null sets are the set of points which we can "ignore" when consider the length for sets.

Corollary 7.1
1. If E is null, so is any subset $E' \subseteq E$
2. If E_n is null for all $n \in \mathbb{E}$, so is $\cup_{n=1}^{\infty} E_n$
3. All countable subsets of $\mathbb{R}$ are null.

Proof. (1) follows from (4) in proposition (7.8); (2) follows from (6) in proposition (7.8); (3) follows from (1) and (6) in proposition (7.8). ■

In the remaining of this lecture let's discuss two interesting questions:

1. Are there any uncountable null sets?
2. Both "null" and "meagre" is small. Is null = meagre?

The classic example, cantor set is meagre, null, and uncountable:

■ **Example 7.3** [Cantor Set] Starting from the interval $C_0 = [0,1]$, one delete the open middle third $(1/3, 2/3)$ from C_0, leaving two line segments:

$$C_1 = [0,1/3] \cup [2/3,1].$$

Next, the open middle third of each of these remaining segments is deleted, leaving four line segments:

$$C_2 = [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1].$$

Continuing this process infinitely, and define $C = \cap_{n=1}^{\infty} C_n$.

1. The cantor set C is null, since $C \subseteq C_n$ for all n, i.e.,

$$m^*(C) \le m^*(C_n) = (2/3)^n, \ \forall n \implies m^*(C) = 0.$$

2. The cantor set C is uncountable: every element in C can be expressed uniquely in ternary expression, i.e., only use 0,1,2 as digits. Suppose on the contrary that C is countable, i.e., $C = \{c_n\}_{n\in\mathbb{N}}$. Then construct a new number such that $c \notin \{c_n\}_{n\in\mathbb{N}}$ by diagonal argument.
3. C is nowhere dense, i.e., C is meagre:
   (a) Firstly, C is closed, since intersection of closed sets is closed.
   (b) Suppose on the contrary that $(\alpha,\beta) \subseteq C$ for some open interval (α,β), then $(\alpha,\beta) \subseteq C_n = \sqcup_{k=1}^{2^n}[a_{n,k}, b_{n,k}]$ for all n. Therefore, for any fixed n, $(\alpha,\beta) \subseteq [a_{n,k}, b_{n,k}]$ for some k, which implies

$$\beta - \alpha < b_{n,k} - a_{n,k} = \frac{1}{3^n}, \ \forall n \in \mathbb{N}$$

   Therefore, $\beta - \alpha = 0$, which is a contradiction.

■

R However, the answer for the second question is no. There exists a mergre set S with $m^*(S) = \infty$; and also a null set that is co-meagre. The construction of these examples are left as exercise.

The outer measure m^* is a special measure of the length of a given subset. Now we define the generalized measure of length:

Definition 7.5 [Measure] A meaasure of length for all subsets in $\mathbb{R}$ is a function m satisfying

1. $m(\emptyset) = m(\{x\}) = 0$
2. $m(\{a,b\}) = b - a$
3. $m(A + x) = m(A), \forall x \in \mathbb{R}$
4. If $A \subseteq B$, then $m(A) \leq m(B)$
5. $m(kA) = |k| m(A)$
6. If $E_i \cap E_j = \emptyset, \forall i \neq j$, then

$$\sum_{i=1}^{\infty} m(E_i) = m(\cup_{i=1}^{\infty} E_i)$$

Question: m^* satisfies (1) to (5), does m^* satisfies (6) for any subsets? In other words, is outer measure the special case of the definition of measure?

Answer: no.

7.3. Monday for MAT4002

7.3.1. Quotient Map

Definition 7.6 [Quotient Map] A mapping $q : X \to Y$ between topological spaces is a **quotient map** if

1. q is surjective
2. For any $U \subseteq Y$, U is open iff $q^{-1}(U)$ is open.

R

1. The canonical projection mapping $p : X \to X/\sim$ is a quotient mapping
2. We say f is an open mapping if U is open in X implies $f(U)$ is open in Y. Note that a continuous open mapping satisfies condition (2) in definition (7.6).

In proposition (6.5) we show the homeomorphism between $X/\sim$ and Y given the compactness of X and Hausdorffness of Y. Now we show the homeomorphism by replacing these conditions with the quotient mapping q:

Proposition 7.9 Suppose $q : X \to Y$ is a quotient map, and that $\sim$ is an equivalence relation on X given by the partition $\{q^{-1}(y) \mid y \in Y\}$. Then $X/\sim$ and Y are **homeomorphic**.

Proof. Construct the mapping

$$\begin{aligned} h: &\quad X/\sim \to Y \\ \text{with} &\quad h([x]) = q(x) \end{aligned}$$

Note that:

1. The mapping h is well-defined and injective.
2. Surjective is easy to shown.

3. The quotient mapping $q := h \circ p$, by definition, is continuous. By applying proposition (6.4), h is continuous.

It suffices to show h^{-1} is continuous:

- For any open $\tilde{U} \subseteq X/\sim$, it suffices to show $h(\tilde{U})$ is open in Y.

  Note that
$$q^{-1}(h(\tilde{U})) = p^{-1}h^{-1}(h(\tilde{U})) = p^{-1}(\tilde{U}),$$
  which is open by the definition of quotient topology (check proposition (6.1)).

  Therefore, $h(\tilde{U})$ is open by (2) in definition (7.6).

■

■ **Example 7.4** The $\mathbb{R}/\mathbb{Z}$ is homeomorphic to the unit circle S^1:

Define the mapping
$$\begin{aligned} q: \quad \mathbb{R} &\to S^1 \\ x &\mapsto e^{2\pi i x} \end{aligned}$$

It's clear that

1. q is a continuous open mapping (why?)
2. q is surjective

Therefore, $\mathbb{R}/\sim \cong S^1$, provided that $x \sim y$ iff $q(x) = q(y)$, i.e., $x - y \in \mathbb{Z}$. Therefore,
$$\mathbb{R}/\mathbb{Z} \cong S^1$$

7.3.2. Simplicial Complex

Combinatorics is the slums of topology. — J. H. C. Whitehead

The idea is to build some new spaces from some "fundamental" objects. The combinatorialists often study topology by the combinatorics of these fundamental objects. First we define what are the "fundamental" objects:

Definition 7.7 [n-simplex] The standard n-simplex is the set

$$\Delta^n = \left\{ (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \middle| x_i \geq 0, \forall i \text{ and } \sum_{i=1}^{n+1} x_i = 1 \right\}$$

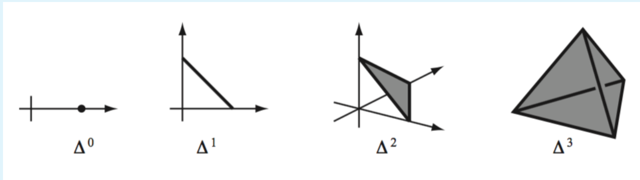


Figure 7.1: Simplices on $\mathbb{R}^2$ are the triangles, so you may consider simplexes as the "triangles" in general spaces

1. The non-negative integer n is the **dimension** of this simplex
2. Its **vertices**, denoted as $V(\Delta^n)$, are those points $(x_1, \ldots, x_{n+1})$ in Δ^n such that $x_i = 1$ for some i.
3. For each given non-empty $\mathcal{A} \subseteq \{1, \ldots, n+1\}$, its **face** is defined as

$$\{(x_1, \ldots, x_{n+1}) \in \Delta^n \mid x_i = 0, \ \forall i \notin \mathcal{A}\}$$

   In particular, Δ^n is a face of itself
4. The **inside** of Δ^n is

$$\text{inside}(\Delta^n) := \{(x_1, \ldots, x_{n+1}) \in \Delta^n \mid x_i > 0, \forall i\}$$

   In particular, the inside of Δ^0 is Δ^0.

Definition 7.8 [Face Inclusion] A face inclusion of Δ^m into Δ^n $(m < n)$ is a function $\Delta^m \to \Delta^n$ which comes from the restriction of an **injective linear** map $f : \mathbb{R}^{m+1} \to \mathbb{R}^{n+1}$ that maps vertices in Δ^m into vertices in Δ^n.

For example, the linear transformation $f : \mathbb{R}^2 \to \mathbb{R}^3$ defined below is a face inclusion:

$$f(1,0) = (0,1,0), \quad f(0,1) = (0,0,1).$$

R Any injection mapping from $\{1,\dots,m+1\} \to \{1,\dots,n+1\}$ gives a face inclusion $\Delta^m \to \Delta^n$, and vice versa.

Motivation. Now we build new spaces by making use of simplices. This new space is called the **abstract complex**. If a simplex is a part of the complex, so are all its faces.

Definition 7.9 [Abstract Simplicial Complex] An (abstract) **simplicial complex** is a pair $K = (V, \Sigma)$, where V is a set of vertices and Σ is a collection of non-empty finite subsets of V (simplices) such that

1. For any $\boldsymbol{v} \in V$, the 1-element set $\{\boldsymbol{v}\}$ is in Σ
2. If σ is an element of Σ, then so is any non-empty subset of σ.

For example, if $V = \{1,2,3,4\}$, then

$$\Sigma = \{\{1\},\{2\},\{3\},\{4\},\{1,3,4\},\{2,4\},\{1,3\},\{3,4\},\{1,4\}\}$$

We can associate to an abstract simplicial complex K a topological space $|K|$, which is called its **geometric realization**:

Definition 7.10 [Topological Realization] The **topological realization** of $K = (V, \Sigma)$ is a topological space $|K|$ (or denoted as $|(V,\Sigma)|$), where

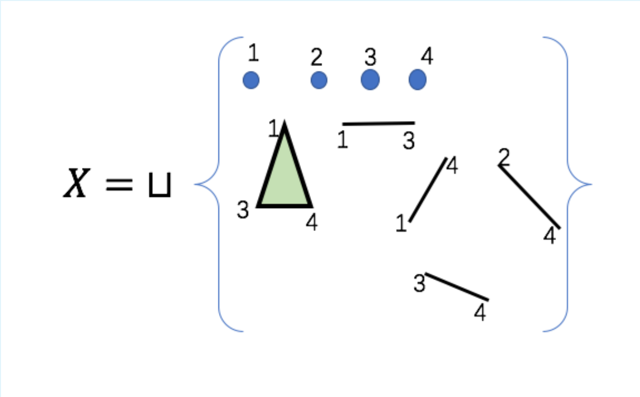
1. For each $\sigma \in \Sigma$ with $|\sigma| = n+1$, take a copy of n-simplex and denote it as Δ_σ
2. Whenever $\sigma \subset \tau \in \Sigma$, identify Δ_σ with a face of Δ_τ through face inclusion.

R Or equivalently, $|K|$ is a quotient space of the disjoint union

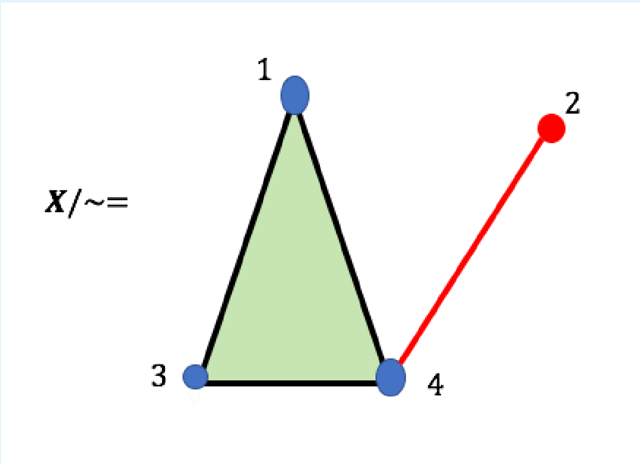
$$\coprod_{\sigma\in\Sigma}\sigma$$

by the equivalence relation which identifies a point $y\in\sigma$ with its image under the face inclusion $\sigma\to\tau$, for any $\sigma\subset\tau$.

■ **Example 7.5** Take



As a result,

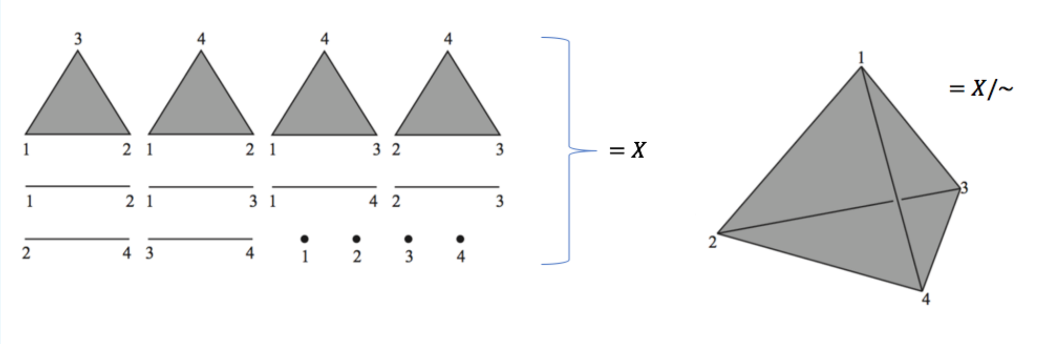


■

■ **Example 7.6** Take $V = \{1,2,3,4\}$ and

$$\Sigma = \{\text{all subsets of } V \text{ except } V\}$$

As shown in the figure below, $|(V,\Sigma)| = \Delta^3$:



■

Definition 7.11 [Triangulation] A **triangulation** of a topological space X is a simplicial complex $K = (V,\Sigma)$ together with a choice of homeomorphism $|K| \to X$. ■

■ **Example 7.7** The triangulation of $S^1 \times S^1$ can be realized by using nine vertices given below:

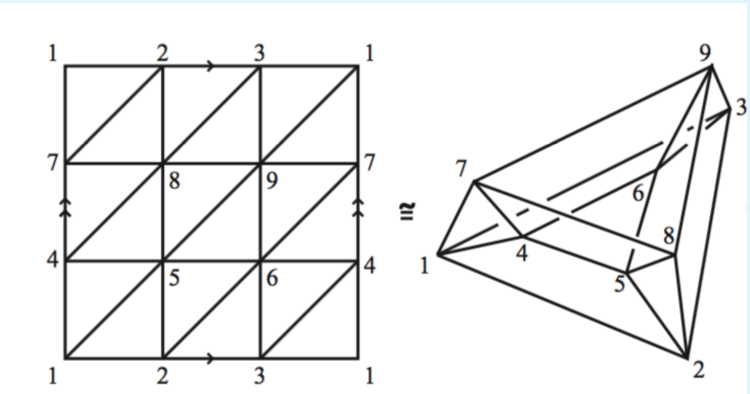


Figure 7.2: The quotient space $|K| := X/\sim$

(Try to identify X) ■

7.4. Wednesday for MAT3040

Reviewing.

- Given the polynomial $f(x) \in \mathbb{F}[x]$, we extend it into the linear operator $f(T): V \to V$.
- The minimal polynomial $m_T(x)$ is defined to be the polynomial with least degree such that
$$m_T(T) = \mathbf{0}_{V\to V},$$
i.e., $[m_T(T)]\boldsymbol{v} = 0_{\boldsymbol{V}}, \forall \boldsymbol{v} \in V$.
- The minimial polynomial of a vector $\boldsymbol{v}$ relative to T is defined to be the polynomial $m_{T,\boldsymbol{v}}(x)$ with the least degree such that
$$m_{T,\boldsymbol{v}}(T)(\boldsymbol{v}) = 0$$
- If $f(T) = \mathbf{0}_{V\to V}$, then we imply $m_T(x) \mid f(x)$. If $[g(T)](\boldsymbol{w}) = 0_V$, following the similar argument, we imply $m_{T,\boldsymbol{w}}(x) \mid g(x)$.
- In particular, $m_T(T)\boldsymbol{w} = \mathbf{0}$, which implies $m_{T,\boldsymbol{w}}(x) \mid m_T(x)$.

7.4.1. Cayley-Hamiton Theorem

Let's raise an motivative example first:

■ **Example 7.8** Consider the matrix and its induced mapping $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. It has the characteristic polynomial
$$\mathcal{X}_A = (x-1)(x-2).$$

- Note that $m_A(x)$ cannot be with degree one, since otherwise $m_A(x) = x - k$ with some k, and

$$m_A(\boldsymbol{A}) = \boldsymbol{A} - k\boldsymbol{I} = \begin{pmatrix} 1-k & 0 \\ 0 & 2-k \end{pmatrix} \neq \boldsymbol{0}, \quad \forall k,$$

which is a contradiction.

- However, one can verify that the $m_A(x)$ is with degree 2:

$$m_A(x) = (x-1)(x-2).$$

- The minimial polynomial with eigenvectors can be with degree 1:

$$\boldsymbol{w} = [0,1]^{\mathrm{T}} \implies (A-2I)\boldsymbol{w} = \boldsymbol{0} \implies m_{A,\boldsymbol{w}}(x) = x-2$$

R More generally, given an eigen-pair $(\lambda, \boldsymbol{v})$, the minimal polynomial of an $\boldsymbol{v}$ has the explicit form

$$m_{T,\boldsymbol{v}}(x) = (x-\lambda) \implies (x-\lambda) \mid m_T(x)$$

Now we want to relate the characterstic polynomial $m_T(x)$ with $\mathcal{X}_T(x)$. Suppose that

$$\mathcal{X}_T(x) = (x-\lambda_1)^{e_1} \cdots (x-\lambda_k)^{e_k} \in \mathbb{F}[x]. \tag{7.1}$$

Then we imply

- λ_i is an eigenvalue of T;
- $(x-\lambda_i) \mid m_T(x)$;

which implies that $(x-\lambda_1)\cdots(x-\lambda_k) \mid m_T(x)$.

Furthermore, (a). does $m_T(x)$ possess other factors, e.g., does there exist $\mu \neq \lambda_i, i = 1,\dots,k$ such that $(x-\mu) \mid m_T(x)$? (b). does $(x-\lambda_i)^{f_i} \mid m_T(x)$ when $f_i > e_i$?

The answer is no for both question (a) and (b).

Theorem 7.1 — Cayley-Hamilton. $m_T(x) \mid \mathcal{X}_T(x)$. In particular, $\mathcal{X}_T(T) = \mathbf{0}$.

The nice equality in (7.1) does not necessarily hold. Sometimes $\mathcal{X}_T(x)$ cannot be factorized into linear factors in $\mathbb{F}[x]$, e.g., $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ in $\mathbb{R}$.

However, for every $f(x) \in \mathbb{F}[x]$, we can extend $\mathbb{F}$ into the algebraically closed set $\overline{F} \supseteq \mathbb{F}$ such that

$$f(x) = (x-\lambda_1)^{e_1} \cdots (x-\lambda_k)^{e_k}$$

where $\lambda_i \in \overline{\mathbb{F}}$.

For example, for $f(x) = x^2 + 1 \in \mathbb{R}[x]$, we can extend $\mathbb{R}$ into $\mathbb{C}$ to obtain

$$f(x) = (x+i)(x-i).$$

Therefore, the general proof outline for the Cayley-Hamilton Theorem is as follows:

- Consider the case where $m_T(x), \mathcal{X}_T(x)$ are both in $\overline{F}[x]$
- Show that $m_T(x) \mid \mathcal{X}_T(x)$ under $\overline{F}[x]$.

Before the proof, let's study the invariant subspaces, which leads to the decomposition of charactersitc polynomial:

Assumption. From now on, we assume that V is finite dimensional by default.

Definition 7.12 [Invariant Subspace] An **invariant subspace** of a linear operator $T : V \to V$ is a subspace $W \leq V$ that is preserved by T, i.e., $T(W) \subseteq W$. We also call W as T-invariant. ■

(R) If $W \leq V$ is T-invariant, then the restriction of the linear operator $T : V \to V$ induces the linear operator

$$T\mid_W : W \to W.$$

■ **Example 7.9** 1. V itself is T-invariant.

2. For the eigenvalue λ, the associated λ-eigenspace $U = \ker(T - \lambda I)$ is T-invariant.

3. More generally, $U = \ker(g(T))$ is T-invariant for any polynomial g:

   If $\boldsymbol{v} \in \ker(g(T))$, i.e., $g(T)\boldsymbol{v} = \boldsymbol{0}$, it suffices to show $T(\boldsymbol{v}) \in \ker(g(T))$:

$$\begin{aligned} g(T)[T(\boldsymbol{v})] &= (a_m T^m + \cdots + a_0 I)[T(\boldsymbol{v})] \\ &= (a_m T \circ T^m + \cdots + a_1 T \circ T + a_0 T \circ I)(\boldsymbol{v}) \\ &= T[g(T)\boldsymbol{v}] = T(\boldsymbol{0}) = \boldsymbol{0} \end{aligned}$$

4. For $\boldsymbol{v} \in \ker(T - \lambda I)$, $U = \operatorname{span}\{\boldsymbol{v}\}$ is T-invariant.

■

Proposition 7.10 Suppose that $T : V \to V$ is a linear transformation and $W \leq V$ is T-invariant, then we construct the subspace mapping and the recipe mapping

$$\begin{aligned} T\,|_W \colon &\quad W \to W \\ \text{with} &\quad \boldsymbol{w} \mapsto T(\boldsymbol{w}) \end{aligned} \tag{7.2a}$$

$$\begin{aligned} \tilde{T} : &\quad V/W \to V/W \\ \text{with} &\quad \boldsymbol{v} + W \mapsto T(\boldsymbol{v}) + W \end{aligned} \tag{7.2b}$$

(Here the well-definedness of the recipe mapping $\tilde{T}$ is shown in Hw2, Exercise 4),

which leads to the decomposition of the charactersitic polynomial:

$$\mathcal{X}_T(x) = \mathcal{X}_{T|_W}(x)\mathcal{X}_{\tilde{T}}(x).$$

Proof. Suppose $\mathcal{C} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is a basis of W, and extend it into the basis of V, denoted as

$$\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n\}$$

Therefore, $\overline{\mathcal{B}} = \{\boldsymbol{v}_{k+1} + W, \ldots, \boldsymbol{v}_n + W\}$ is a basis of V/W. By Homework 2, Question 5, the representation $(T)_{\mathcal{B},\mathcal{B}}$ can be written as the block matrix

$$(T)_{\mathcal{B},\mathcal{B}} = \begin{pmatrix} (T|_W)_{\mathcal{C},\mathcal{C}} & \times \\ \mathbf{0} & (\tilde{T})_{\overline{\mathcal{B}},\overline{\mathcal{B}}} \end{pmatrix}_{(k+(n-k))\times(k+(n-k))}$$

Therefore, the characteristic polynomial of T can be calculated as:

$$\begin{aligned} \mathcal{X}_T(x) &= \det((T)_{\mathcal{B},\mathcal{B}} - xI) \\ &= \det((T|_U)_{\mathcal{C},\mathcal{C}} - xI) \cdot \det((\tilde{T})_{\overline{\mathcal{B}},\overline{\mathcal{B}}} - xI) \end{aligned}$$

■

Proposition 7.11 Suppose that

$$\mathcal{X}_T(x) = (x - \lambda_1)\cdots(x - \lambda_n)$$

where λ_i's are not necessarily distinct. Then there exists a basis of V, say $\mathcal{A}$, such that

$$(T)_{\mathcal{A},\mathcal{A}} = \begin{pmatrix} \lambda_1 & \times & \times & \times \\ 0 & \lambda_2 & \cdots & \times \\ 0 & \cdots & \ddots & \times \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Proof. The proof is by induction on n, i.e., suppose the results hold for all vector spaces with dimension no more than $n-1$, and we aim to show this result holds for dimension n.

1. **Step 1**: Argue that there exists the associated eigenvector $\boldsymbol{v}$ of λ_1 under the linear operator T.

   Consider any basis $\mathcal{M}$, by MAT2040, there exists associated eigenvector of λ_1, say $\boldsymbol{y} \in \mathbb{C}^n$ such that

   $$(T)_{\mathcal{M},\mathcal{M}} \cdot \boldsymbol{y} = \lambda_1 \boldsymbol{y}$$

   Since the operator $(\cdot)_{\mathcal{M}} : V \to \mathbb{C}^n$ is an isomorphism, there exists $\boldsymbol{v} \in V \setminus \{\mathbf{0}\}$ such that $(\boldsymbol{v})_{\mathcal{M}} = \boldsymbol{y}$. It follows that

$$(T)_{\mathcal{M},\mathcal{M}}(\boldsymbol{v})_{\mathcal{M}} = \lambda_1(\boldsymbol{v})_{\mathcal{M}} \implies (T\boldsymbol{v})_{\mathcal{M}} = (\lambda_1\boldsymbol{v})_{\mathcal{M}} \implies T\boldsymbol{v} = \lambda_1\boldsymbol{v}$$

2. **Step 2**: Dimensionality reduction of $\mathcal{X}_T(x)$: Construct $W = \text{span}\{\boldsymbol{v}\}$, which is T-invariant. By the proof of proposition (7.11), we have $\tilde{T} : V/W \to V/W$ admits the characteristic polynomial

$$\mathcal{X}_{\tilde{T}}(x) = (x - \lambda_2)\cdots(x - \lambda_n)$$

3. **Step 3:** Applying the induction, there exists basis $\overline{C}$ of V/W, i.e.,

$$\overline{C} = \{\boldsymbol{w}_2 + W, \dots, \boldsymbol{w}_n + W\}$$

such that

$$(\tilde{T})_{\overline{C},\overline{C}} = \begin{pmatrix} \lambda_2 & \times & \times & \times \\ 0 & \lambda_3 & \cdots & \times \\ 0 & \cdots & \ddots & \times \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

4. **Step 4:** Therefore, we construct the set $\mathcal{A} := \{\boldsymbol{v}, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$. We claim that
   - $\mathcal{A}$ is a basis of V (left as exercise in Hw2, Question 2)
   - $$(T)_{\mathcal{A},\mathcal{A}} = \begin{pmatrix} \lambda_1 & \times \\ \mathbf{0} & (\tilde{T})_{\overline{C},\overline{C}} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \times & \times & \times \\ 0 & \lambda_2 & \cdots & \times \\ 0 & \cdots & \ddots & \times \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

     (This statement is also left as exercise in Hw2, Question 5.)

■

Proposition 7.12 Suppose that $\mathcal{X}_T(x) = (x - \lambda_1)\cdots(x - \lambda_n)$, then $\mathcal{X}_T(T) = \mathbf{0}$.

R One special case is that $\boldsymbol{A} = \text{diag}(\lambda_1, \dots, \lambda_n)$. The results for proposition (7.12) gives

$$(A - \lambda_1 \boldsymbol{I}) \cdots (A - \lambda_n \boldsymbol{I}) \text{ is a zero matrix}$$

7.5. Wednesday for MAT4002

7.5.1. Remarks on Triangulation

Consider the simplical complex $K = (V, \Sigma)$ with

$$V = \{1,2,3,4,\dots,9\}, \quad \Sigma = \begin{cases} 9 \text{ subsets with } 1 \text{ element} \\ 27 \text{ subsets with } 2 \text{ elements} \\ 18 \text{ subsets with } 3 \text{ elements} \end{cases}$$

We start to build the topological realization of K with 9 **0**-simplicies, 27 **1**-simplicies, and 18 **2**-simplicies. The identification of them is as follows:

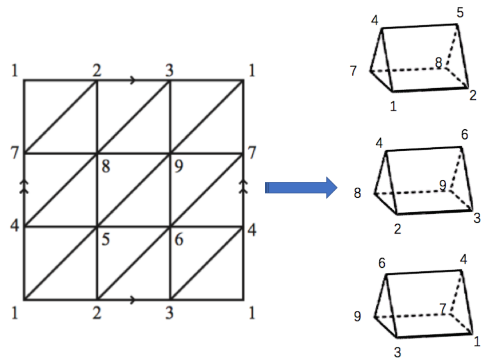


Figure 7.3: Step 1: Identify 3 columns separately, i.e., identify $\{1,7,4,1,2,8,5,2\}$, $\{2,8,5,2,3,9,6,3\}$, and $\{3,9,6,3,1,7,4,1\}$.

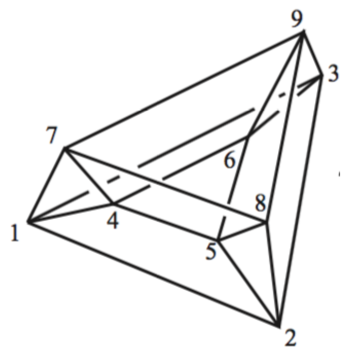
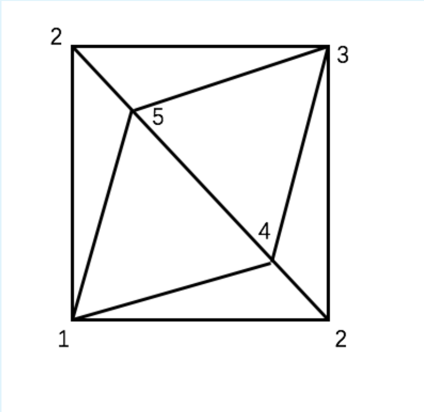


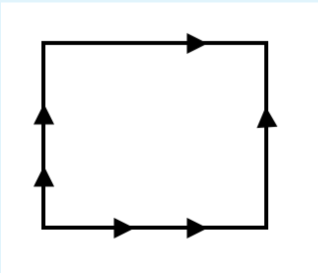
Figure 7.4: Step 2: "gluing" these three prisms in the figure above together.

Question: why K is homeomorphic to the torus?

■ **Example 7.10** Consider the simplicial complex (V,Σ) described below:

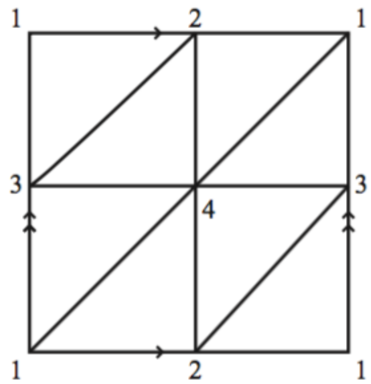


The $|(V,\Sigma)|$ is homeomorphism to the quotient space S^1 plotted below

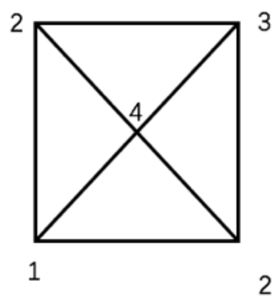


■

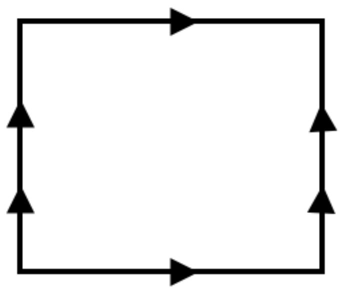
Furthermore, can we build a triangulation of the tours using fewer simplices? The answer is no. Consider the figure below: at the bottom edge of this square, there are two 1-simplicies lablled $\{1,2\}$, which cannot happen in a tours.



Interesting question: does the triangulation of the Fig. (7.1a) below leads to S^2?



$(7.1a)$: The simplicial complex (V,Σ)



$(7.1b)$: Quotient Space of S^2

Answer: No. Since the 2-simplex $\Delta_{\{2,3,4\}}$ appears twice in the Fig. (7.1a), the triangulation of this figure means that we need to stick the top triangle and the right triangle together, which contradicts to the structure of the quotient space S^2 shown in Fig. (7.1b).

The simplicial complex gives us another way to study X, i.e., it suffices to study (V,Σ) such that $|(V,\Sigma)| \cong X$. The question is that can we distinguish $X = S^1 \times S^1$ and $Y = S^2$? In other words, can we distinguish the difference of corresponding topological realizations?

Theorem 7.2 — Euler's Formula. Suppose that $|(V_1,\Sigma_1)| \cong |(V_2,\Sigma_2)|$, then

$$\begin{aligned}&\sum_{i=1}^{\infty}(-1)^i\,(\text{number of subsets in } \Sigma_1 \text{ with } (i+1)\text{-element})\\ =&\sum_{i=1}^{\infty}(-1)^i\,(\text{number of subsets in } \Sigma_2 \text{ with } (i+1)\text{-element})\end{aligned}$$

From previous examples we can see that $\mathcal{X}(S^2) = 5-9+6 = 2$ and $\mathcal{X}(S^1 \times S^1) = 9-27+18 = 0$, which implies

$$S^2 \ncong S^1 \times S^1.$$

7.5.2. Simplicial Subcomplex

Definition 7.13 [Simplicial Subcomplex] A subcomplex of a simplicial complex $K = (V, \Sigma)$ is a simplicial complex $K' = (V', \Sigma')$ such that

$$V' \subseteq V, \quad \Sigma' \subseteq \Sigma$$

Proposition 7.13 Suppose K' is subcomplex of K, then $|K'|$ is closed in $|K|$.

Proof. Suppose that D is the disjoint union of all the simplicial complex forming $|K|$. (note that the number of component in D is $|\Sigma|$)

Consider the canonical projection mapping $D \to |K|$. Observe that $p^{-1}(|K'|)$ precisely equals to $\coprod_{\sigma' \in \Sigma'} \sigma'$, which is closed in D. By definition of quotient topology, $|K'|$ is also closed. ■

Definition 7.14 [Subcomplex spanned by vertices] Let $K = (V, \Sigma)$ be a simplicial complex and $V' \subseteq V$. Then the subcomplex spanned by V' is (V', Σ') such that

- V' denotes the vertex set.
- the simplices Σ' is given by

$$\{\sigma \in \Sigma \mid \sigma \subseteq V'\}$$

Definition 7.15 [Link and Star] Let $(V, \Sigma) = K$ be simplicial complex

- The **link** of $\boldsymbol{v} \in V$, denoted as $\mathrm{lk}(\boldsymbol{v})$ is the sub-complex with
  - vertex set

$$\{\boldsymbol{w} \in V \setminus \{\boldsymbol{v}\} \mid \{\boldsymbol{v}, \boldsymbol{w}\} \in \Sigma\}$$

  - simplicies

$$\{\sigma \in \Sigma \mid \boldsymbol{v} \notin \sigma \text{ and } \sigma \cup \{\boldsymbol{v}\} \in \Sigma\}$$

- The star of $\boldsymbol{v}$ (denoted as $\mathrm{st}(\boldsymbol{v})$) is

$$\bigcup\{\mathrm{inside}(\sigma) \mid \sigma \in \Sigma, \boldsymbol{v} \in \sigma\}$$

Proposition 7.14 $\mathrm{st}(\boldsymbol{v})$ is open and $\boldsymbol{v} \in \mathrm{st}(\boldsymbol{v})$.

Proof. Omitted. ■

In fact, $|K| \setminus \mathrm{st}(\boldsymbol{v})$ is the simplicial subcomplex spanned by V.

7.5.3. Some properties of simplicial complex

Proposition 7.15 Suppose that $K = (V, \Sigma)$, where V is finite. Then $|K|$ is compact.

Proof. The mapping $p : D \to |K|$ is a canonical projection mapping, which is continuous; and D (the finite disjoint union of Δ_σ's) is compact.

Therefore, $p(D) = |K|$ is compact. ■

Proposition 7.16 For any simplicial complex $K = (V, \Sigma)$, where V is finite, there is a continuous injection

$$f : |K| \to \mathbb{R}^n \quad \text{for some } n$$

Proof. Let $K' = (V, \Sigma')$, where $\Sigma' =$ power set of V. Then

$$|K'| = \Delta^{|V|-1} \subseteq \mathbb{R}^{|V|}$$

Consider the inclusion

$$i : |K| \to |K'|$$

which comes from the following:

1. Consider the $D := \coprod_{\sigma \in \Sigma} \Delta_\sigma$ and $D' = \coprod_{\sigma' \in \Sigma'} \Delta_{\sigma'}$ in (V, Σ) and (V, Σ')
2. Construct the mapping $\tilde{i} : D \hookrightarrow D' \xrightarrow{p'} |K|$.

3. The mapping $\tilde{i}$ descends to $i : D/\sim \longrightarrow |K'|$ (try to write down the detailed mapping), which is continuous and injective.

Therefore, $|K| \hookrightarrow |K'|$, i.e., $|K| \hookrightarrow \mathbb{R}^n$. The proof is complete.

∎

Chapter 8

Week8

8.1. Monday for MAT3040

Reviewing.

- If $\mathcal{X}_T(x) = (x-\lambda_1)\cdots(x-\lambda_n)$, then

$$(T)_{\mathcal{A},\mathcal{A}} = \begin{pmatrix} \lambda_1 & \times & \times & \times \\ 0 & \lambda_2 & \cdots & \times \\ 0 & \cdots & \ddots & \times \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

for some basis $\mathcal{A}$. In other words, T is **triangularizable** with the diagonal entries $\lambda_1,\dots,\lambda_n$.

(R) I hope you appreciate this result. Consider the example below: In linear algebra we have studied that the matrix $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable, and the characteristic polynomial is given by

$$\mathcal{X}_A(x) = (x-1)^2.$$

However, the theorem above claims that $\boldsymbol{A}$ is *triangularizable*, with diagonal entries 1 and 1. The diagonalization of $\boldsymbol{A}$ only uses the eigenvector of $\boldsymbol{A}$, but the 1-eigenspace has only 1 dimension. Fortunately, the triangularization gives a rescue such that we can make use of the generalized eigenvector $(0,1)^{\mathrm{T}}$

(but not an eigenvector) of $\boldsymbol{A}$ by considering the mapping below:

$$U = \operatorname{span}\left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}$$
$$\bar{A}: \quad V/U \to V/U$$

Here $(0,1)^{\mathrm{T}} + U$ is an eigenvector of $\bar{A}$, with eigenvalue 1.

Theorem 8.1 The linear operator T is triangularizable with diagonal entries $(\lambda_1, \dots, \lambda_n)$ if and only if

$$\mathcal{X}_T = (x - \lambda_1) \cdots (x - \lambda_n)$$

Proof. It suffices to show only the sufficiency. Suppose that there exists basis $\mathcal{A}$ such that

$$(T)_{\mathcal{A},\mathcal{A}} = \begin{pmatrix} \lambda_1 & \times & \times & \times \\ 0 & \lambda_2 & \cdots & \times \\ 0 & \cdots & \ddots & \times \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Then we compute the characteristic polynomial directly:

$$\begin{aligned} \mathcal{X}_T(x) &= \det[(xI - T)_{\mathcal{A},\mathcal{A}}] \\ &= \det \begin{pmatrix} x - \lambda_1 & \times & \times & \times \\ 0 & x - \lambda_2 & \cdots & \times \\ 0 & \cdots & \ddots & \times \\ 0 & 0 & \cdots & x - \lambda_n \end{pmatrix} \\ &= (x - \lambda_1) \cdots (x - \lambda_n) \end{aligned}$$

■

8.1.1. Cayley-Hamiton Theorem

Proposition 8.1 — **A Useful Lemma.** Suppose that $\mathcal{X}_T(x) = (x-\lambda_1)\cdots(x-\lambda_n)$, then $\mathcal{X}_T(T) = 0$.

Proof. Since $\mathcal{X}_T(x) = (x-\lambda_1)\cdots(x-\lambda_n)$, we imply T is triangularizable under some basis $\mathcal{A}$. Note that

- $T \mapsto (T)_{\mathcal{A},\mathcal{A}}$ is an isomorphism between $\mathrm{Hom}(V,V)$ and $M_{n\times n}(\mathbb{F})$,
- $(\underbrace{T\circ T\circ\cdots\circ T}_{m \text{ times}})_{\mathcal{A},\mathcal{A}} = [(T)_{\mathcal{A},\mathcal{A}}]^m$, for any m,

It suffices to show $\mathcal{X}_T((T)_{\mathcal{A},\mathcal{A}})$ is the zero matrix (why?):

$$\mathcal{X}_T((T)_{\mathcal{A},\mathcal{A}}) = ((T)_{\mathcal{A},\mathcal{A}} - \lambda_1\boldsymbol{I})\cdots((T)_{\mathcal{A},\mathcal{A}} - \lambda_n\boldsymbol{I}).$$

Observe the matrix multiplication

$$((T)_{\mathcal{A},\mathcal{A}} - \lambda_i\boldsymbol{I})\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1-\lambda_i & \times & \times & \times \\ 0 & \lambda_2-\lambda_i & \cdots & \times \\ 0 & \cdots & \ddots & \times \\ 0 & 0 & \cdots & \lambda_n-\lambda_i \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathrm{span}\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_{i-1}\}$$

Therefore, for any $\boldsymbol{v}\in V$,

$$((T)_{\mathcal{A},\mathcal{A}} - \lambda_n\boldsymbol{I})\boldsymbol{v} \in \mathrm{span}\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_{n-1}\}.$$

Applying the same trick, we conclude that

$$((T)_{\mathcal{A},\mathcal{A}} - \lambda_1\boldsymbol{I})\cdots((T)_{\mathcal{A},\mathcal{A}} - \lambda_n\boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}, \quad \forall \boldsymbol{v}\in V,$$

i.e., $\mathcal{X}_T((T)_{\mathcal{A},\mathcal{A}}) = ((T)_{\mathcal{A},\mathcal{A}} - \lambda_1\boldsymbol{I})\cdots((T)_{\mathcal{A},\mathcal{A}} - \lambda_n\boldsymbol{I})$ is a zero matrix. ∎

Now we are ready to give a proof for the Cayley-Hamiton Theorem:

Proof. Suppose that $\mathcal{X}_T(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{F}[x]$. By considering algebrically closed field $\overline{\mathbb{F}} \supseteq \mathbb{F}$, we imply

$$\begin{aligned}
\mathcal{X}_T(x) &= x^n + a_{n-1}x^{n-1} + \cdots + a_0 && (8.1a)\\
&= (x-\lambda_1)\cdots(x-\lambda_n), \quad \lambda_i \in \overline{\mathbb{F}} && (8.1b)
\end{aligned}$$

By applying proposition (8.1), we imply $\mathcal{X}_T(T) = 0$, where the coefficients in the formula $\mathcal{X}_T(T) = 0$ w.r.t. T are in $\overline{\mathbb{F}}$.

Then we argue that these coefficients are essentially in $\mathbb{F}$. Expand the whole map of $\mathcal{X}_T(T)$:

$$\begin{aligned}
\mathcal{X}_T(T) &= (T-\lambda_1 I)\cdots(T-\lambda_n I) && (8.2a)\\
&= T^n - (\lambda_1 + \cdots + \lambda_n)T^{n-1} + \cdots + (-1)^n\lambda_1\cdots\lambda_n I && (8.2b)\\
&= T^n + a_{n-1}T^{n-1} + \cdots + a_0 I && (8.2c)
\end{aligned}$$

where the derivation of (8.2c) is because that the polynomial coefficients for (8.1a) and (8.1b) are all identical.

Therefore, we conclude that $\mathcal{X}_T(T) = 0$, under the field $\mathbb{F}$. ■

Corollary 8.1 $m_T(x) \mid \mathcal{X}_T(x)$. More precisely, if

$$\mathcal{X}_T(x) = [p_1(x)]^{e_1}\cdots[p_k(x)]^{e_k}, \ e_i > 0, \forall i$$

where p_i's are distinct, monic, and irreducible polynomials. Then

$$m_T(x) = [p_1(x)]^{f_1}\cdots[p_k(x)]^{f_k}, \text{ for some } 0 < f_i \le e_i, \forall i$$

Proof. The statement $m_T(x) \mid \mathcal{X}_T(x)$ is from Cayley-Hamiton Theorem. Therefore, $0 \le f_i \le e_i, \forall i$. Suppose on the contrary that $f_i = 0$ for some i. w.l.o.g., $i = 1$.

It's clear that $\gcd(p_1, p_j) = 1$ for $\forall j \ne 1$, which implies

$$a(x)p_1(x) + b(x)p_j(x) = 1, \text{ for some } a(x), b(x) \in \mathbb{F}[x].$$

Considering the field extension $\overline{\mathbb{F}} \supseteq \mathbb{F}$, we have $p_1(x) = (x-\mu_1)\cdots(x-\mu_\ell)$. For any root μ_m of p_1, $m = 1,\dots,\ell$, we have

$$a(\mu_m)p_1(\mu_m) + b(\mu_m)p_j(\mu_m) = 1 \implies b(\mu_m)p_j(\mu_m) = 1 \implies p_j(\mu_m) \neq 0,$$

i.e., μ_m is not a root of p_j, $\forall j \neq 1$.

Therefore, μ_m is a root of $\mathcal{X}_T(x)$, but not a root of $m_T(x)$. Then μ_m is an eigenvalue of T, e.g., $T\boldsymbol{v} = \mu_m\boldsymbol{v}$ for some $\boldsymbol{v} \neq \mathbf{0}$. Recall that $m_{T,\boldsymbol{v}} = x - \mu_m$, we imply $m_{T,\boldsymbol{v}} = x - \mu_m \mid m_T(x)$, which is a contradiction. ■

■ **Example 8.1** We can use Corollary (8.1), a stronger version of Cayley-Hamiltion Theorem to determine the minimal polynomials:

1. For matrix $\boldsymbol{A} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$, we imply $\mathcal{X}_A(x) = (x^2+x+1)^1$. Since x^2+x+1 is irreducible in $\mathbb{R}$, we have $m_A(x) = x^2+x+1$.
2. For matrix
$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},$$
we imply $\mathcal{X}_A(x) = (x-1)^2(x-2)^2$.
By Corollary (8.1), we imply both $(x-1)$ and $(x-2)$ should be roots of $m_T(x)$, i.e., $m_A(x)$ may have the four options:
$$(x-1)^2(x-2)^2, \text{ or}$$
$$(x-1)(x-2)^2, \text{ or}$$
$$(x-1)^2(x-2), \text{ or}$$
$$(x-1)(x-2).$$
By trial and error, one sees that $m_A(x) = (x-1)^2(x-2)$.

8.1.2. Primary Decomposition Theorem

We know that not every linear operator is diagonalizable, but diagonalization has some nice properties:

Definition 8.1 [diagonalizable] The linear operator $T : V \to V$ is diagonalizable over $\mathbb{F}$ if and only if there exists a basis $\mathcal{A}$ of V such that

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n),$$

where λ_i's are not necessarily distinct.

Proposition 8.2 If the linear operator $T : V \to V$ is diagonalizable, then

$$m_T(x) = (x - \mu_1) \cdots (x - \mu_k),$$

where μ_i's are **distinct**.

Proof. Suppose T is diagonalizable, then there exists a basis $\mathcal{A}$ of V such that

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\mu_1, \dots, \mu_1, \mu_2, \dots, \mu_2, \dots, \mu_k, \dots, \mu_k)$$

It's clear that $((T)_{\mathcal{A},\mathcal{A}} - \mu_1 \boldsymbol{I}) \cdots ((T)_{\mathcal{A},\mathcal{A}} - \mu_k \boldsymbol{I}) = \boldsymbol{0}$, i.e., $m_T(x) \mid (x - \mu_1) \cdots (x - \mu_k)$.

Then we show the minimality of $(x - \mu_1) \cdots (x - \mu_k)$. In particular, if $(x - \mu_i)$ is omitted for any $1 \le i \le k$, then it's easy to show

$$(T_{\mathcal{A},\mathcal{A}} - \mu_1 \boldsymbol{I}) \cdots (T_{\mathcal{A},\mathcal{A}} - \mu_{i-1} \boldsymbol{I})(T_{\mathcal{A},\mathcal{A}} - \mu_{i+1} \boldsymbol{I}) \cdots (T_{\mathcal{A},\mathcal{A}} - \mu_k \boldsymbol{I}) \ne \boldsymbol{0},$$

since all μ_i's are distinct. Therefore, $m_T(x)$ will not divide $(x - \mu_1) \cdots (x - \mu_{i-1})(x - \mu_{i+1}) \cdots (x - \mu_k)$ for any i, i.e.,

$$m_T(x) = (x - \mu_1) \cdots (x - \mu_k)$$

■

R The converse of proposition (8.2) is also true, which is a special case for the Primary Decomposition Theorem.

Theorem 8.2 — Primary Decomposition Theorem. Let $T : V \longrightarrow V$ be a linear operator with

$$m_T(x) = [p_1(x)]^{e_1} \cdots [p_k(x)]^{e_k},$$

where p_i's are distinct, monic, and irreducible polynomials. Let $V_i = \ker([p_i(x)]^{e_i}) \leq V, i = 1, \ldots, k$, then

1. Each V_i is T-invariant ($T(V_i) \leq V_i$)
2. $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$
3. Consider $T|_{V_i}: V_i \longrightarrow V_i$, then

$$m_{T|V_i}(x) = [p_i(x)]^{e_i}$$

8.2. Monday for MAT3006

Reviewing. We define the **outer** measure of a subset $E \subseteq \mathbb{R}$ to be

$$m^*(E) = \inf\left\{\sum_{n=1}^{\infty} m(I_n) \middle| E \subseteq \bigcup_{n=1}^{\infty} I_n,\ I_n\text{'s are open intervals}\right\}$$

One Special Property of Outer Measure:

$$m^*(\cup_{n=1}^{\infty} E_n) \le \sum_{n=1}^{\infty} m^*(E_n)$$

8.2.1. Remarks for Outer Measure

We want to make a special hyphothesis become true: If E_n's are disjoint, then

$$m^*(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m^*(E_n) \tag{8.3}$$

However, (8.3) does not necessary hold for a sequence of disjoint subsets $\{E_n\}$. One counter-example is shown in Example (8.2).

■ **Example 8.2** [Vitali Set] Suppose that $A \subseteq [0,1]$ satisfies the following properties:

- For any $x \in \mathbb{R}$, there exists $q \in \mathbb{Q}$ such that $x + q \in A$.
- If $x, y \in A$ such that $x \neq y$, then $x - y \notin \mathbb{Q}$

In other words, the group $\mathbb{R}$ is partitioned into the cosets of its additive subgroup $\mathbb{Q}$, and the properties above say that A contains exactly one member of each coset of $\mathbb{Q}$. The existence of such A relies on the Axiom of Choice. Moreover, we imply:

- $[0,1] \subseteq \bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A - q)$: since $\forall x \in [0,1]$, there exists $q \in \mathbb{Q}$ s.t. $x + q \in A$, which implies $x \in A - q$. Moreover, we can bound the possible region of q:

$$0 \le x + q \le 1 \implies -x \le q \le 1 - x \implies -1 \le q \le 1$$

- $\bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A-q)\subseteq[-1,2]$: elements in $A-q$ are of the form $x-q, x\in[0,1], q\in[-1,1]$, and therefore $x-q\in[-1,2]$.
- The sets $(A-q)$ are disjoint as q varies, i.e., $(A-q_1)\cap(A-q_2)=\emptyset, \forall q_1\neq q_2\in[-1,1]\cap\mathbb{Q}$: Suppose on the contrary that there exists $y\in(A-q_1)\cap(A-q_2)$, which follows

$$y+q_1, y+q_2\in A, y+q_1\neq y+q_2 \implies (y+q_1)-(y+q_2)=q_1-q_2\notin\mathbb{Q}$$

Suppose on the contrary that (8.3) holds for $\{A-q \mid \forall q\in[-1,1]\cap\mathbb{Q}\}$, then

$$m^*\left(\bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A-q)\right)=\sum_{q\in[-1,1]\cap\mathbb{Q}}m^*(A-q)=\sum_{q\in[-1,1]\cap\mathbb{Q}}m^*(A), \tag{8.4}$$

where the second equality is because that $m^*(A-q)=m^*(A),\forall q$. However,

$$1=m^*([0,1])\leq m^*\left(\bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A-q)\right)\leq m^*([-1,2])=3 \tag{8.5}$$

From (8.4) we derive the $m^*\left(\bigcup_{q\in[-1,1]\cap\mathbb{Q}}(A-q)\right)$ can either be 0 or ∞, which is a contradiction. ■

8.2.2. Lebesgue Measurable

Therefore, (8.5) does not hold for some bad subsets of $\mathbb{R}$, which are sets cannot be explicitly described. Let's focus on sets with good behaviour only:

Definition 8.2 [Carathedory Property] A subset $E\subseteq\mathbb{R}$ is **measurable** if

$$m^*(A)=m^*(A\cap E)+m^*(A\setminus E) \tag{8.6}$$

for all subsets $A\subseteq\mathbb{R}$, where E is not assumed to be in A, i.e., $A\setminus E:=A\cap E^c$. ■

R To argue whether (8.6) holds, we essentially suffice to verify the inequality $m^*(A)\geq m^*(A\cap E)+m^*(A\setminus E)$. There are many other equivalent definitions for measurable set $E \subseteq \mathbb{R}$:

1. For any $\varepsilon > 0$, there exists open set $U \supseteq E$ such that

$$m^*(U \setminus E) \leq \varepsilon$$

2. Its outer and inner measures are equal:

$$m^*(E) = m_*(E) := \sup \left\{ \sum_{n=1}^{\infty} m(I_n) \,\middle|\, \bigcup_{n=1}^{\infty} I_n \subseteq E,\ I_n\text{'s are compact, and disjoint subsets} \right\}$$

Note that the inner measure m_* admits the inequality

$$m_*(\cup_{n=1}^{\infty} E_n) \geq \sum_{n=1}^{\infty} m_*(E_n), \text{ for disjoint } E_n$$

R If $E \subseteq \mathbb{R}$, then for all $B \supseteq E$, we have

$$m^*(B) = m^*(B \cap E) + m^*(B \setminus E) = m^*(E) + m^*(B \setminus E): \tag{8.7}$$

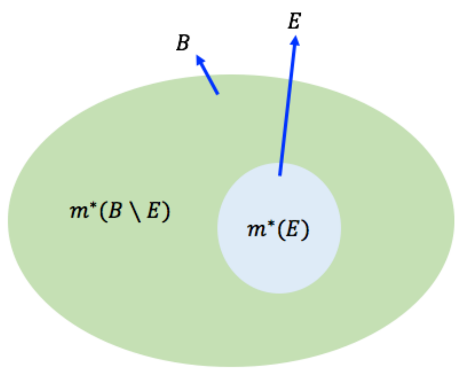


Figure 8.1: Illustration for the useful equality (8.7)

Proposition 8.3 1. If $E \subseteq \mathbb{R}$ is null, then E is measurable

2. If I is any interval, then I is measurable

3. If E is measurable, then $E^c := \mathbb{R} \setminus E$ is measurable

4. If E is measurable, then both $\cup_{i=1}^{n} E_i$ and $\cap_{i=1}^{n} E_i$ are measurable

Proof. 1. For any subsets A,

$$\begin{cases} m^*(A \cap E) = 0 \\ m^*(A \cap E^c) \le m^*(A) \end{cases} \implies m^*(A) \ge m^*(A \cap E) + m^*(A \cap E^c).$$

2. Take $I = [a, b]$. For all $A \subseteq \mathbb{R}$,

   - take $\{I_n\}$ such that $A \subseteq \cup_{n=1}^{\infty} I_n$ and

$$\sum_{n=1}^{\infty} m^*(I_n) \le m^*(A) + \varepsilon \tag{8.8}$$

   - Note that the $m^*(A \cap I)$ can be upper bounded:

$$A \cap I \subseteq \cup_{n=1}^{\infty} (I_n \cap I) \implies m^*(A \cap I) \le \sum_{n=1}^{\infty} m^*(I_n \cap [a, b])$$

     Similarly, $m^*(A \cap I^c)$ can be upper bounded:

$$A \cap I^c \subseteq \cup_{n=1}^{\infty} I_n \cap ((-\infty, a) \cup (b, \infty)) = \left(\bigcup_{n=1}^{\infty} I_n \cap (-\infty, a) \right) \cup \left(\bigcup_{n=1}^{\infty} I_n \cap (b, \infty) \right),$$

     i.e.,

$$m^*(A \cap I^c) \le \sum_{n=1}^{\infty} m^*(I_n \cap (-\infty, a)) + m^*(I_n \cap (b, \infty))$$

   - Therefore,

$$\begin{aligned} m^*(A \cap I) + m^*(A \cap I^c) &\le \sum_{n=1}^{\infty} m^*(I_n \cap (-\infty, a)) + m^*(I_n \cap [a, b]) + m^*(I_n \cap (b, \infty)) \\ &= \sum_{n=1}^{\infty} m^*(I_n \cap (-\infty, \infty)) = \sum_{n=1}^{\infty} m^*(I_n) \\ &\le m^*(A) + \varepsilon, \end{aligned}$$

     i.e., $m^*(A \cap I) + m^*(A \cap I^c) \le m^*(A)$.

3. Part (3) is trivial.

4. Part (4) is by induction on n: suppose that

   - E_i is measurable for $i = 1, \ldots, k+1$
   - $E = \cup_{i=1}^{k} E_i$ is measurable

   By the measurablitiy of E_{k+1},

$$m^*(A \cap E^c) = m^*(A \cap E^c \cap E_{k+1}) + m^*(A \cap E^c \cap E_{k+1}^c) \tag{8.9}$$

   By the measurablitiy of E,

$$\begin{aligned} m^*(A) &\geq m^*(A \cap E) + m^*(A \cap E^c) \\ &\geq [m^*(A \cap E) + m^*(A \cap E^c \cap E_{k+1})] + m^*(A \cap E^c \cap E_{k+1}^c) \end{aligned} \tag{8.10}$$

   It's easy to show

$$E \cup (E^c \cap E_{k+1}) = E \cup E_{k+1},$$

   which implies

$$\begin{aligned} m^*(A \cap (E \cup E_{k+1})) &= m^*(A \cap (E \cup (E^c \cap E_{k+1}))) \\ &= m^*((A \cap E) \cup (A \cap (E^c \cap E_{k+1}))) \\ &\leq m^*(A \cap E) + m^*(A \cap (E^c \cap E_{k+1})) \end{aligned} \tag{8.11}$$

   Substituting (8.11) into (8.10) gives

$$m^*(A) \geq m^*(A \cap (E \cup E_{k+1})) + m^*(A \cap (E \cup E_{k+1})^c),$$

   i.e., $E \cup E_{k+1}$ is measurable as well.

   By the equality

$$\mathbb{R} \setminus \left(\bigcup_{i=1}^{n} E_i \right) = \bigcup_{i=1}^{n} (\mathbb{R} \setminus E_i),$$

   and the result in part (3), one can show $\cap_{i=1}^{n} E_i$ is measurable as well.

∎

Proposition 8.4 If E_i is measurable, then $\cup_{i=1}^{\infty} E_i$ is measurable. Moreover, if E_i's are disjoint, then

$$m^*(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} m^*(E_i)$$

R Note that $m^*(A) \neq 0$ for Vitali set A: suppose contrary that $m^*(A) = 0$, i.e., A is null set. Since countably null set is also measurable, together with (8.4), we imply

$$m^*\left(\bigcup_{q \in [-1,1] \cap \mathbb{Q}} (A - q)\right) = 0,$$

which contradicts to (8.5).

Notations.

1. We will write $m(E) = m^*(E)$ for all measurable sets $E \subseteq \mathbb{R}$, and therefore

$$m(\{a,b\}) = m^*(\{a,b\}) = b - a$$

2. The sets E satisfying

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c)$$

   are called **Lebesgue** measurable in some other textbooks.

8.3. Monday for MAT4002

8.3.1. Quotient Map

Definition 8.3 [Quotient Map] A map $q : X \to Y$ between topological spaces is a **quotient map** if

1. q is surjective
2. For any $U \subseteq Y$, U is open iff $q^{-1}(U)$ is open.

For example,

1. the mapping $p : X \to X/\sim$ is a quotient map
2. For f is continuous and maps open sets to open sets. Then f satisfies (2).

Proposition 8.5 Suppose $q : X \to Y$ is a quotient map, and $\sim$ is an equivalence relation given by the partition $\{q^{-1}(y) \mid y \in Y\}$ of X. Then X and Y are **homeomorphic**.

Proof. Let $h : X/\sim \to Y$ defined as $h([x]) = q(x)$.

1. The mapping h is well-defined and injective.
2. Surjective is easy to shown
3. $q = h \circ p$, by (2), is continuous. Therefore, h is continuous.

It suffices to show h^{-1} is continuous: For any open $\tilde{U} \subseteq X/\sim$, it suffices to show $h(\tilde{U})$ is open in Y. Note that

$$q^{-1}[h(\tilde{U})] = p^{-1}h^{-1}(h(\tilde{U})) = p^{-1}(\tilde{U}),$$

which is open by definition of quotient topology. By (2), $h(\tilde{U})$ is open ■

■ **Example 8.3** Define the mapping

$$\begin{aligned} q: \quad & \mathbb{R} \to S^1 \\ & x \mapsto e^{2\pi i x} \end{aligned}$$

Note that q open balls to open balls. (Check what $q(a,b)$ is)

$q^{-1}(e^{2\pi i x}) = \{x + z \mid z \in \mathbb{Z}\}$. Therefore,

$$\mathbb{R}/\!\sim\, \cong S^1,$$

where $x \sim y$ iff $x - y \in \mathbb{Z}$, i.e..,

$$\mathbb{R}/\mathbb{Z} \cong S^1$$

■

8.3.2. Simplicial Complex

The idea is to build some new spaces from some fundamental objects. Then we can use the combinatorics of these fundamental objects to study topology.

Definition 8.4 [n-simplex] The standard n-simplex is the set

$$\Delta^n = \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_i \ge 0, \sum_i x_i = 1\}$$

1. vertices of Δ^n are the points on Δ^n with $x_i = 1$ for some i.
2. face of Δ^n : For $\mathcal{A} \subseteq \{1, \dots, n+1\}$, a face is given by

$$\{(x_1, \dots, x_{n+1}) \in \Delta^n \mid x_i = 0, \ \forall i \notin \mathcal{A}\}$$

3. inside of Δ^n is

$$\{(x_1, \dots, x_{n+1}) \in \Delta^n \mid x_i > 0, \forall i\}$$

(Inside of Δ^0 is Δ^0)

Definition 8.5 [face inclusion] A face inclusion of Δ^m into Δ^n $(m < n)$ is a function $\Delta^m \to \Delta^n$ which comes from the restriction of an **injective linear** map

$$f : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1},$$

which maps vertices into vertices.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Any injection from $\{1,\dots,m+1\} \to \{1,\dots,n+1\}$ gives a face inclusion $\Delta^m \to \Delta^n$, and vice versa.

Definition 8.6 [Simplicial Complex] An (abstract) **simplicial complex** is a pair (V,Σ), where V is a set of vertices and Σ is a collection of non-empty finite subsets of V (simplices) such that

1. $\forall \boldsymbol{v} \in V$, $\{\boldsymbol{v}\} \in \Sigma$
2. If $\sigma \in \Sigma$, then any non-empty subset of σ must lie in Σ.

   e.g., $V = \{1,2,3,4\}$, then

$$\Sigma = \{\{1\},\{2\},\{3\},\{4\},\{1,3,4\},\{2,4\},\{1,3\},\{3,4\},\{1,4\}\}$$

Definition 8.7 [Topological Realization] The **topological realization** of $K = (V,\Sigma)$ is a topological space $|K|$ (or denoted as $|(V,\Sigma)|$), where

1. For each $\sigma \in \Sigma$ with $|\sigma| = n+1$, take a copy of n-simplex and denote it as Δ_σ
2. Whenever $\sigma \subset \tau \in \Sigma$, identify Δ_σ with a face of Δ_τ through face inclusion.

■ **Example 8.4** Take

■ **Example 8.5** Take $V = \{1,2,3,4\}$ and

$$\Sigma = \{\text{all subsets of } V \text{ except } V\}$$

■ **Example 8.6** Take $V = \{1,\dots,n+1\}$ and

$$\Sigma = \{\text{All subsets of } V\}$$

Then $|(V,\Sigma)| = \Delta^n$.

Definition 8.8 [Triangulation] A **triangulation** of a topological spcae X is a simplicial complex $(V,\Sigma) = K$ with a choice of homeomorphism $|(V,\Sigma)| \to X$.

8.4. Wednesday for MAT3006

Reviewing.

- All null sets are measurable

- If $E \subseteq \mathbb{R}$ is measurable, then $E^c := R \setminus E$ is measurable.

- E_i is measurable implies $\cup_{i=1}^n E_i$ is measurable.

8.4.1. Remarks on Lebesgue Measurability

Proposition 8.6 If E_i is measurable for $\forall i \in \mathbb{N}$, then so is $\cup_{i=1}^{\infty} E_i$. Moreover, if further E_i's are pairwise disjoint, then

$$m^*\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m^*(E_i)$$

Proof. • Consider the case where E_i's are measurable, pairwise disjoint first. For all subsets $A \subseteq \mathbb{R}$, and all $n \in \mathbb{N}$,

$$m^*(A) = m^*(A \cap (\cup_{i=1}^n E_i)) + m^*(A \cap (\cup_{i=1}^n E_i)^c) \tag{8.12a}$$

$$= [m^*(A \cap (\cup_{i=1}^n E_i) \cap E_n) + m^*(A \cap (\cup_{i=1}^n E_i) \cap E_n^c)] + m^*(A \cap (\cup_{i=1}^n E_i)^c) \tag{8.12b}$$

$$= [m^*(A \cap E_n) + m^*(A \cap (\cup_{i=1}^{n-1} E_i)] + m^*(A \cap (\cup_{i=1}^n E_i)^c) \tag{8.12c}$$

where (8.12a) is by the measurability of $\cup_{i=1}^n E_i$; (8.12b) is by the measurability of E_n; (8.12c) is by direct calculation.

Proceeding these trick similarly, we obtain:

$$m^*(A) = [m^*(A \cap E_n) + m^*(A \cap (\cup_{i=1}^{n} E_i)^c)] + m^*(A \cap (\cup_{i=1}^{n-1} E_i)] \tag{8.13a}$$

$$= \sum_{\ell=1}^{n} m^*(A \cap E_\ell) + m^*(A \cap (\cup_{i=1}^{\ell} E_i)^c) \tag{8.13b}$$

$$\geq \sum_{\ell=1}^{n} m^*(A \cap E_\ell) + m^*(A \cap (\cup_{i=1}^{\infty} E_i)^c) \tag{8.13c}$$

for any $n \in \mathbb{N}$, where (8.13c) is by lower bounding $(\cup_{i=1}^{\ell} E_i)^c \supseteq (\cup_{i=1}^{\infty} E_i)^c$. Taking $n \to \infty$ in (8.13c), we imply

$$m^*(A) \geq \sum_{\ell=1}^{\infty} m^*(A \cap E_\ell) + m^*(A \cap (\cup_{i=1}^{\infty} E_i)^c) \tag{8.13d}$$

$$\geq m^*(\cup_{i=1}^{\infty}(A \cap E_i)) + m^*(A \cap (\cup_{i=1}^{\infty} E_i)^c) \tag{8.13e}$$

$$= m^*(A \cap (\cup_{i=1}^{\infty} E_i)) + m^*(A \cap (\cup_{i=1}^{\infty} E_i)^c) \tag{8.13f}$$

where (8.13e) is by the countable sub-addictivity of m^*. Therefore, $\cup_{i=1}^{\infty} E_i$ is measurable.

- Moreover, taking $A = \cup_{i=1}^{\infty} E_i$ in (8.13d) gives

$$m^*(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} m^*(E_i) + m^*(\emptyset) = \sum_{i=1}^{\infty} m^*(E_i) + 0.$$

- Now suppose that E_i's are measurable but not necessarily pairwise disjoint. We need to show $\bigcup_{i=1}^{\infty} E_i$ is measurable. The way is to construct the disjoint sequence of sets first:

$$\begin{cases} F_1 = E_1, \\ F_{k+1} = E_k \setminus \left(\cup_{i=1}^{k} E_i\right), \ \forall k > 1 \end{cases} \implies \cup_{i=1}^{\infty} F_i = \cup_{i=1}^{\infty} E_i$$

It's clear that F_i's are pairwise disjoint and measurable, which implies $\cup_{i=1}^{\infty} E_i = \cup_{i=1}^{\infty} F_i$ is measrable. The proof is complete. ■

Notations. We denote $\mathcal{M}$ as the collection of all **(Lebesgue) measurable** subsets of $\mathbb{R}$, and

$$m(E) = m^*(E), \quad \forall E \in \mathcal{M}$$

8.4.2. Measures In Probability Theory

Definition 8.9 [σ-Algebra]

- Let Ω be any set, and $\mathbb{P}(\Omega)$ (**power set**) denotes the collection of all subsets of Ω
- A family of subsets of Ω, denoted as $\mathcal{T}$, is a σ-algebra if it satisfies
  1. $\emptyset, \Omega \in \mathcal{T}$
  2. If $E \in \mathcal{T}$, then $E^c \in \mathcal{T}$
  3. If $E_i \in \mathcal{T}$ for $\forall i \in \mathbb{N}$, then $\cup_{i=1}^{\infty} E_i \in \mathcal{T}$ (and therefore $\cap_{i=1}^{\infty} E_i \in \mathcal{T}$).

Definition 8.10 [Measure] A **measure** on a σ-algerba $(\Omega, \mathcal{T})$ is a function $\mu : \mathcal{T} \to [0, \infty]$ such that

- $\mu(\emptyset) = 0$
- $\mu(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$ whenever E_i's are pairwise disjoint in $\mathcal{T}$.

As a result, $(\Omega, \mathcal{T}, \mu)$ is called a **measurable space**.

Example 8.7 1. Let Ω be any set, $\mathcal{T} = \mathcal{M}$, and $\mu(E) = |E|$ (the number of elements in E). Then $(\Omega, \mathbb{P}(\Omega), \mu)$ is a measure space, and μ is called a counting measure on Ω.

Definition 8.11 [Borel σ-algebra] Let $\boldsymbol{B}$ be a collection of all intervals in $\mathbb{R}$. Then there is a **unique** σ-algebra $\mathcal{B}$ of $\mathbb{R}$, such that

1. $\boldsymbol{B} \subseteq \mathcal{B}$
2. For all σ-algebra $\mathcal{T}$ containing $\boldsymbol{B}$, we have $\mathcal{B} \subseteq \mathcal{T}$

This $\mathcal{B}$ is called a **Borel σ-algebra**

R

1. In particular, $C_i \in \mathcal{B}$ implies $\cup_{i=1}^{\infty} C_i$ and $\cap_{i=1}^{\infty} C_i \in \mathcal{B}$.
2. $\mathcal{B} \subseteq \mathcal{M}$, since $\boldsymbol{B} \subseteq \mathcal{M}$ and $\mathcal{M}$ is a σ-algebra.
3. However, $\mathcal{M}$ and $\mathcal{B}$ are not equal. The element $C \in \mathcal{B}$ is called **Borel measurable subsets**

Definition 8.12 [complete] Let $(\Omega, \mathcal{T}, \mu)$ be a measurable space. Then we say it is **complete** if for any $E \in \mathcal{T}$ with $\mu(E) = 0$, $N \subseteq E$ implies $N \in \mathcal{T}$. (and therefore $\mu(N) = 0$)

Example 8.8
1. $(\mathbb{R}, \mu, m^*)$ is complete.

   Reason: if $m^*(E) = 0$, then $m^*(N) = 0$, $\forall N \subseteq E$
2. $(\mathbb{R}, \mu, m)$ is complete.

   Reason: the same as in (1)
3. However, $(\mathbb{R}, \mathcal{B}, m\,|_{\mathcal{B}})$ is not complete. (left as exercise)

Then we study the difference between $\mathcal{B}$ and $\mathcal{M}$:

Definition 8.13 [Completion] Let $(\Omega, \mathcal{T}, \mu)$ be measurable space. The **completion** of $(\Omega, \mathcal{T}, \mu)$ with respect to μ is the smallest complete σ-algebra containing $\mathcal{T}$, denoted as $\overline{\mathcal{T}}$. More precisely,

$$\overline{\mathcal{T}} = \{G \cup N \mid G \in \mathcal{T}, N \subseteq F \in \mathcal{T}, \text{with } \mu(F) = 0\}$$

e.g., take $G = \emptyset \in \mathcal{T}$. For all $F \in \mathcal{T}$ such that $\mu(F) = 0$, $N \subseteq F$ implies $N \in \overline{\mathcal{T}}$.

R If further define $\overline{\mu} : \overline{\mathcal{T}} \to [0,\infty]$ by

$$\overline{\mu}(G \cup N) = \mu(G),$$

then $(\Omega, \overline{\mathcal{T}}, \overline{\mu})$ is a measurable space.

Theorem 8.3 The completion of $(\mathbb{R}, \mathcal{B}, m|_{\mathcal{B}})$ is $(\mathbb{R}, \mathcal{M}, m)$

R Another completion of $(\mathbb{R}, \mu, m)$ is as follows:

Define $\ell(\{a,b\}) = b - a$ for all intervals $\{a,b\} \in \boldsymbol{B}$ Then by Caratheodory extension theorem, we can extend $\ell : \boldsymbol{B} \to [0,\infty]$ to $\ell : \mathcal{B} \to [0,\infty]$.

Complete $\ell : \mathcal{B} \to [0,\infty]$ to $\bar{\ell} : \mathcal{M} \to [0,\infty]$. Then $\bar{\ell} = m$ as in our course.

8.5. Wednesday for MAT4002

Reviewing. We can construct a continuous injection from $|K|$ to $|K'|$, where $K = (V, \Sigma)$ is a simplicial complex, and $K' = (V', \Sigma')$ is its subcomplex:

Let $D_\Sigma := \coprod_{\sigma\in\Sigma} \sigma$ and $D_{\Sigma'} := \coprod_{\sigma'\in\Sigma'} \sigma'$, then $|K'| = D_{\Sigma'}/\sim_{\Sigma'}$ and $|K| = D_\Sigma/\sim_\Sigma$, which follows that

$$f : D_{\Sigma'} \to D_\Sigma \xrightarrow{P} D_\Sigma/\sim_\Sigma, \quad P \text{ denotes the canonical projection mapping}$$

The whole mapping f descends to a continuous mapping

$$\tilde{f} : D_{\Sigma'}/\sim_{\Sigma'} \to D_\Sigma/\sim_\Sigma$$

The $\tilde{f}$ is injective since

$$x \sim_{\Sigma'} y \iff i(x) \sim_\Sigma i(y), \qquad \forall x, y \in D_\Sigma, \tag{8.14}$$

where i denotes the inclusion mapping.

Another way is to consider the inclusion $i : |K'| \to |K|$, which is continuous and injective as well. Note that $i(|K'|)$ is closed in $|K|$.

Proposition 8.7 For each $K = (V, \Sigma)$, and finite V, there is a continuous injection $g : |K| \hookrightarrow \mathbb{R}^n$ for some n.

Proof. Consider $K^p := (V, \Sigma^p)$, where Σ^p is the power set of V. Therefore, $|K^p| = \Delta^{|V|-1} \subseteq \mathbb{R}^{|V|}$, and K is a simplicial subcomplex of K^p, which follows that

$$l : |K'| \xrightarrow{i} |K^p| \xrightarrow{i} \mathbb{R}^{|V|}$$

The whole mapping l is an inclusion mapping from $|K'|$ to $\mathbb{R}^{|V|}$, which is continuous and injective. The proof is complete. ■

Proposition 8.8 — Hausdorff. If $K = (V, \Sigma)$ with fintie V, then $|K|$ is Hausdorff.

Proof. Let $g : |K| \xrightarrow{l} \mathbb{R}^n$. Consider the bijective $g : |K| \to g(|K|)$, which is continuous.

Sicne $|K|$ is compact, and $g(|K|) \subseteq \mathbb{R}^n$ is Hausdorff, we imply that $|K|$ and $g(|K|)$ are homeomorphic, i.e., $|K|$ is Hausdorff. ■

Definition 8.14 [Edge Path] An **edge path** of $K = (V, \Sigma)$ is a sequence of vertices $(v_1, \dots, v_n), v_i \in V$ such that $\{v_i, v_{i+1}\} \in \Sigma, \forall i$.

Proposition 8.9 — **Connectedness.** Let $K = (V, \Sigma)$ be a simplicial complex. TFAE:

1. $|K|$ is connected
2. $|K|$ is path-connected
3. Any 2 vertices in (V, Σ) can be joined by an edge path, i.e., for $\forall u, v \in V$, there exists $v_1, \dots, v_k \in V$ such that $(u, v_1, \dots, v_k, v)$ is an edge path.

Sketch of Proof (to be revised).

1. (3) implies (2): For every $x, y \in |K|$,

$$\begin{cases} x \in \Delta_{\sigma_1} \text{ for some } \sigma_1 \in \Sigma. \\ y \in \Delta_{\sigma_2} \text{ for some } \sigma_2 \in \Sigma. \end{cases}$$

   Take a path joining x to a vertex $v_1 \in \sigma_1$ and a path joining y to a vertex $v_2 \in \sigma_2$. By (3), we have a path joninig v_1 and v_2.
2. (1) implies (3): Suppose on the contrary that there is a vertex v not satisfying (3). Take V' as the set of vertexs that can be joined with v; and V'' as the set of vertexs that cannot be joinied with v.

   Then $V', V'' \neq \emptyset$. Consider K', K'' be simplicial subcomplexes of K, spanned by V' and V''. Then $|K'|, |K''|$ are disjoint, closed in $|K|$.

   $|K| = |K'| \cup |K''|$. If there exists $x \in |K| \setminus (|K'| \cup |K''|)$, then for any $\sigma \in \Sigma$ such that $x \in \Delta_\sigma$, we imply $\Delta_\sigma \not\subseteq |K'|$ or $|K''|$.

   Therefore, σ consists of vertices in both V' and V''. Then there is $v', v'' \in \sigma$ joining V' and V''.

   Therefore, there is no such x and hence $|K| = |K'| \cup |K''|$ is a disjoint union of two closed sets, i.e., not connected.

■

8.5.1. Homotopy

Yoneda's "philosophy". To understand an object X (in our focus, X denotes topological space), we should understand functions

$$f : A \to X, \quad \text{or} \quad g : X \to B$$

One special example is to let $B = \mathbb{R}$.

There are many type of continuous mappings from X to Y. We will group all these mappings into equivalence classes.

Definition 8.15 [Homotopy] A **Homotopy** between two continuous maps $f, g : X \to Y$ is a continuous map

$$H : X \times [0,1] \to Y$$

such that

$$H(x,0) = f(x), \quad H(x,1) = g(x)$$

If such H exists, we say f and g are **homotopic**, denoted as $f \simeq g$.

■ **Example 8.9** Let $Y \subseteq \mathbb{R}^2$ be a convex subset. Consider two continuous maps $f : X \to Y$ and $g : X \to Y$. They are always homotopic since we can define the homotopy

$$H(x,t) = tg(x) + (1-t)f(x)$$

Proposition 8.10 Homotopy is an equivalent relation.

Proof. 1. Let $f : X \to Y$ be any continuous map. Then $f \simeq f$: we can define a homotopy $H(x,t) = f(x), \forall 0 \le t \le 1$.

2. Suppose $f \simeq g$, i.e., H is a homotopy between f and g, then $g \simeq f$: Define the

mapping $H'(x,t) = H(x,1-t)$, then

$$H'(x,0) = g(x), \quad H'(x,1) = f(x)$$

3. Let $f,g,h : X \to Y$ be three continuous maps. If f and g are homotopic and g and h are homotopic, then f and h are homotopic:
   Let $H : X \times [0,1] \to Y$ be a continuous map such that

$$H(x,0) = f(x), H(x,1) = g(x);$$

   $K : X \times [0,1] \to Y$ be a continuous map such that

$$K(x,0) = g(x), K(x,1) = h(x).$$

   Define a function $J : X \times [0,1] \to Y$ by

$$J(x,t) = \begin{cases} H(x,2t), & 0 \le t \le 1/2 \\ K(x,2t-1), & 1/2 \le t \le 1 \end{cases}$$

   - J is continuous, since for all closed $V \subseteq Y$,

$$J^{-1}(V) = (J^{-1}(V) \cap (X \times [0,1/2])) \cup (J^{-1}(V) \cap (X \times [1/2,1])) = H^{-1}(V) \cup K^{-1}(V),$$

     and the closedness of $H^{-1}(V)$ and $K^{-1}(V)$ implies the closedness of $J^{-1}(V)$
   - Moreover, J has the property that $J(x,0) = H(x,0) = f(x)$, while $J(x,1) = K(x,1) = h(x)$.

■

R There are only one equivalence class in example (8.9). Actually, for given space X and Y, if any two continuous mapping are homotopic, then we imply there is only one equivalence class.

Chapter 9

Week9

9.1. Monday for MAT3040

Reviewing.

- $\mathcal{X}_T(x) = (x - \lambda_1)\cdots(x - \lambda_n)$ over $\mathbb{F}$ if and only if T is triangularizable over $\mathbb{F}$.
- $m_T(x) = (x - \mu_1)\cdots(x - \mu_k)$, where μ_i's are distinct over $\mathbb{F}$ if and only if T is diagonalizable over $\mathbb{F}$.

  The converse for this statement is the proposition (8.2). Let's focus on the proof for the forward direction.

9.1.1. Remarks on Primary Decomposition Theorem

Theorem 9.1 — **Primary Decomposition Theorem.** Let $T : V \to V$ be a linear operator with $\dim(V) < \infty$, and

$$m_T(x) = [p_1(x)]^{e_1} \cdots [p_k(x)]^{e_k}$$

where p_i's are distinct, monic, irreducible polynomials. Let $V_i = \ker(p_i(T)^{e_i})$, then

1. each V_i is T-invariant $(i.e., T(V_i) \le V_i)$
2. $V = V_1 \oplus \cdots \oplus V_k$
3. $T\mid_{V_i}$ has the minimal polynomial $p_i(x)^{e_i}$.

Proof. 1. (1) follows from part (2) for example (**??**).

2. Let $q_i(x) = [p_1(x)]^{e_1} \cdots \widehat{[p_i(x)]^{e_i}} \cdots [p_k(x)]^{e_k} := m_T(x)/[p_i(x)]^{e_i}$, then it is clear that

   (a) $\gcd(q_1, \dots, q_k) = 1$

   (b) $\gcd(q_i, p_i^{e_i}) = 1$

   (c) $q_i \cdot p_i^{e_i} = m_T$

   (d) If $i \neq j$, then $m_T(x) \mid q_i(x)q_j(x)$

   - By (a) and Bezout's Theorem (6.7), there exists polynomials $a_1, \dots, a_k$ such that

     $$a_1(x)q_1(x) + \cdots + a_k(x)q_k(x) = 1,$$

     which implies

     $$\underbrace{a_1(T)q_1(T)\boldsymbol{v}}_{\boldsymbol{v}_1} + \cdots + \underbrace{a_k(T)q_k(T)\boldsymbol{v}}_{\boldsymbol{v}_k} = \boldsymbol{v}$$

     Therefore, $\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k$ for our constructed $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$.

   - Note that

     $$p_i(T)^{e_i}\boldsymbol{v}_i = p_i(T)^{e_i}a_i(T)q_i(T)\boldsymbol{v} = a_i(T)[q_i(T)p_i(T)^{e_i}]\boldsymbol{v} = a_i(T)m_T(T)\boldsymbol{v} = \mathbf{0},$$

     whcih implies $\boldsymbol{v}_i \in \ker([p_i(T)]^{e_i}) := V_i$, and therefore

     $$V = V_1 + \cdots + V_k \tag{9.1}$$

   - To show that the summation in (9.3) is essentially the direct sum, consider

     $$\mathbf{0} = \boldsymbol{v}_1' + \cdots + \boldsymbol{v}_k', \quad \forall \boldsymbol{v}_i' \in V_i. \tag{9.2}$$

     By (a) and Bezout's Theorem (6.7), there exists $b_i(x), c_i(x)$ such that

     $$b_i(x)q_i(x) + c_i(x)p_i(x)^{e_i} = 1 \implies b_i(T)q_i(T) + c_i(T)p_i(T)^{e_i} = I,$$

     i.e.,

     $$b_i(T)q_i(T)\boldsymbol{v}_i' + c_i(T)p_i(T)^{e_i}\boldsymbol{v}_i' = b_i(T)q_i(T)\boldsymbol{v}_i' = \boldsymbol{v}_i'.$$

Appying the mapping $b_i(T)q_i(T)$ into equality (9.4) both sides, $i = 1,\dots,k$, we obtain

$$\mathbf{0} = b_i(T)q_i(T)\mathbf{0} = b_i(T)q_i(T)\boldsymbol{v}_1' + \cdots + b_i(T)q_i(T)\boldsymbol{v}_k'$$

Note that all terms on RHS vanish except for $b_i(T)q_i(T)\boldsymbol{v}_i' = \boldsymbol{v}_i'$, since $q_i(x) = [p_1(x)]^{e_1}\cdots\widehat{[p_i(x)]^{e_i}}\cdots[p_k(x)]^{e_k}$ and $\boldsymbol{v}_j' \in \ker([p_j(x)]^{e_j})$. Therefore, $\boldsymbol{v}_i' = 0$ for $i = 1,\dots,k$, i.e., $V = V_1 \oplus \cdots \oplus V_k$.

3. For any $\boldsymbol{v}_i \in V_i$, we have $p_i(T)^{e_i}\boldsymbol{v}_i = \mathbf{0}$, which implies $m_{T|V_i}(x) \mid p_i(x)^{e_i}$. Together with Corollary (8.1), $m_{T|v_i}(x) = p_i(x)^{f_i}$ for some $1 \le f_i \le e_i$.
Suppose on the contrary that there exists $f_i < e_i$ for some i, consider any $\boldsymbol{v} := \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k \in V$, and

$$p_1(T)^{f_1}\cdots p_k(T)^{f_k}\boldsymbol{v} = p_1(T)^{f_1}\cdots p_k(T)^{f_k}(\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k)$$

The term on the RHS vanishes since $p_j(T)^{f_j}\boldsymbol{v}_j = \mathbf{0}$, which implies

$$m_T \mid p_1^{f_1}\cdots p_k^{f_k},$$

but there exists i such that $e_i > f_i$, which is a contradiction.

■

Corollary 9.1 If $m_i(x) = (x-\mu_1)\cdots(x-\mu_k)$ over $\mathbb{F}$, where μ_i's are distinct, then T is diagonalizable over $\mathbb{F}$. (the converse actually also holds, see proposition (8.2))

Proof. By primary decomposition theorem,

$$V = \underbrace{\ker(T-\mu_1 I)}_{V_1} \oplus \cdots \oplus \underbrace{\ker(T-\mu_k I)}_{V_k}$$

Take B_i as a basis of V_i, an μ_i-eigenspace of T. Then $B := \cup_{i=1}^k B_i$ is a basis consisting of eigenvectors of T.

It's clear that $(T|_{V_i})_{\mathcal{B},\mathcal{B}} = \mathrm{diag}(\mu_i, \dots, \mu_i)$, and T is diagonalizable with

$$(T)_{\mathcal{B},\mathcal{B}} = \mathrm{diag}((T|_{V_1})_{\mathcal{B},\mathcal{B}}, \cdots, (T|_{V_k})_{\mathcal{B},\mathcal{B}}).$$

■

Corollary 9.2 [Spectral Decomposition] Suppose $T : V \to V$ is diagonalizable, then there exists a linear operator $p_i : V \to V$ for $1 \le i \le k$ such that

- $p_i^2 = p_i$ (idempotent)
- $p_i p_j = 0, \forall i \neq j$
- $\sum_{i=1}^k p_i = I$
- $p_i T = T p_i, \forall i$

and scalars $\mu_1, \dots, \mu_k$ such that

$$T = \mu_1 p_1 + \cdots + \mu_k p_k$$

Proof. Diagonlization of T is equivalent to say that $m_T(x) = (x - \mu_1) \cdots (x - \mu_k)$, where μ_i's are distinct. Construct

- $V_i := \ker(T - \mu_i I)$
- $p_i : V \to V$ given by $p_i = a_i(T) q_i(T)$ as in the proof of primary decomposition theorem

Then:

- $p_i T = T p_i$ is obvious
- $\sum_{i=1}^k p_i = \sum_{i=1}^k a_i(T) q_i(T) = I$
- $p_i p_j = a_i(T) a_j(T) q_i(T) q_j(T) := a_i(T) a_j(T) s(T) m_T(T) = \mathbf{0}$
- $p_i^2 = p_i(p_1 + \cdots + p_k) = p_i \cdot I = p_i$

For the last part, note that

- $p_i V \leq V_i, \forall i$: for $\forall \boldsymbol{v} \in V$,

$$(T - \mu_i I)p_i \boldsymbol{v} = (T - \mu_i I)a_i(T)q_i(T)\boldsymbol{v} = a_i(T)m_T(x)\boldsymbol{v} = \mathbf{0}$$

  Therefore, $p_i V \leq \ker(T - \mu_i I) = V_i$

- Now, for all $\boldsymbol{w} \in V$,

$$\begin{aligned} T\boldsymbol{w} &= T(p_1 + \cdots + p_k)\boldsymbol{w} \\ &= Tp_1\boldsymbol{w} + \cdots + Tp_k\boldsymbol{w} \\ &= (\mu_1 p_1)\boldsymbol{w} + \cdots + (\mu_k p_k)\boldsymbol{w} \end{aligned}$$

  and therefore $T = \mu_1 p_1 + \cdots + \mu_k p_k$

■

Organization of future two weeks. We are interested in under which condition does the T is diagonalizable. One special case is $T = A$, where $\boldsymbol{A}$ is a symmetric matrix. We will study normal operators, which includes the case for symmetric matrices.

Question: what happens if $m_T(x)$ contains repeated linear factors? We will spend the next whole class to show the Jordan Normal Form:

Theorem 9.2 — Jordan Normal Form. Let $\mathbb{F}$ be algebraically closed field such that every linear operator $T : V \to V$ has the form

$$m_T(x) = \prod_{i=1}^{k}(x - \lambda_i)^{e_i}$$

where λ_i's are distinct.

Then there exists basis $\mathcal{A}$ of V such that

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_k)$$

where

$$J_i = \begin{pmatrix} \mu & 1 & 0 & 0 \\ 0 & \mu & 1 & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu \end{pmatrix}$$

for some $\mu \in \{\lambda_1, \ldots, \lambda_k\}$

9.2. Monday for MAT3006

Reviewing.

- The collection of all Lebesgue measurable subsets, denoted by $\mathcal{M}$, is a σ-algebra
- Borel σ-algebra: the smallest σ-algebra containing all the intervals of $\mathbb{R}$:

  Well-definedness. Let $\mathcal{S} = \{$all σ-algebras containing all the inervals of $\mathbb{R}\}$. For example, $\mathbb{P}(R) \in \mathcal{S}$. Note that $\forall f_i \in \mathcal{S}$, $\cap_{i\in I}\mathcal{A}_i \in \mathcal{S}$.

  Then define $\mathcal{B} = \cap_{\mathcal{A}\in\mathcal{S}}\mathcal{A}$, which is the smallest σ-algebra containing all intervals. ■

  Furthermore, $\mathcal{M} \in \mathcal{S}$. Therefore, $\mathcal{B} \subseteq \mathcal{M}$ but they are not equal. (Check Royden's note on blackboard, there exists a counter-example A such that $A \in \mathcal{M}$ but $A \notin \mathcal{B}$.)

- The set $\mathcal{M}$ has a good property: If $N \in \mathcal{M}$ is null, then all $E \subseteq N$ are null sets, and therefore $E \in \mathcal{M}$.

  The probelm is that it is not necessary the case that $N' \in \mathcal{B}$ implies $m\mid_{\mathcal{B}}(N') = 0$, i.e., $E' \in \mathcal{B}, \forall E' \subseteq N'$ does not necessarily hold.

  (check back to the Roydon's counter-example)

- Therefore, we need the **completion process** of $\mathcal{B}$ to get $\mathcal{M}$.

9.2.1. Measurable Functions

Motivation. The Riemann integration divides the function into a grid of 1 (unit) squares, and then measure the altitude of the function at the center of each square. Therefore, the total "volume" of this function is 1 times the sum of the altitudes.

However, the Lebesgue integraion aims to study the vertical length of the function, and the total volumn of this function is 1 times the sum of the vertical lengths. Riemann integration

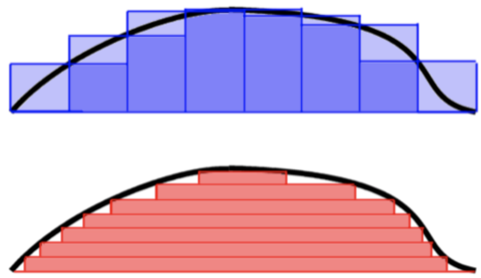


Figure 9.1: Riemann Integration (in blue) and Lebesgue Integration (in red)

Definition 9.1 [Measurable] Let $f : (\mathbb{R}, \mu, m) \to \mathbb{R}$ be a function. We say f is **(Lebesgue) measurable** if $f^{-1}(I) \in \mathcal{M}$ for all intervals $I \subseteq \mathbb{R}$.

Proposition 9.1 If $f : \mathbb{R} \to \mathbb{R}$ is continuous, then f is measurable.

The trick during the proof is to check only intervals of the form (a, ∞) instead of checking all intervals in $\mathbb{R}$.

Proof. 1. By continuity of f, $f^{-1}((a, \infty))$ is open in $\mathbb{R}$.

- Note that any open set U can be expressed as a countable union of open intervals: for given $q \in \mathbb{Q}$, define the set

$$I_q = \bigcup_{I \text{ is an open interval, } q \in I \subseteq U} I,$$

which is a union of non-disjoint open intervals, hence an open interval as well. We claim that $U \subseteq \cup_{q \in \mathbb{Q} \cap U} I_q$:

Consider any $x \in U$. When $x \in \mathbb{Q}$, the result is clear; otherwise there exists an open interval $(x - \varepsilon, x + \varepsilon) \subseteq U$. By the denseness of $\mathbb{Q}$, there exists $q \in \mathbb{Q}$ such that $q \in (x - \varepsilon, x + \varepsilon)$. By definition of I_q, $(x - \varepsilon, x + \varepsilon) \in I_q$. Therefore, $x \in I_q$.

The proof for this statement is complete.

Therefore,

$$f^{-1}((a, \infty)) = \cup_{i=1}^{\infty} U_i \in \mathcal{M}$$

since each open interval $U_i \in \mathcal{M}$

2. For other types of intervals, e.g., $[a,\infty)$, consider

$$\bigcap_{n=1}^{\infty}(a-\frac{1}{n},\infty)=[a,\infty),$$

which follows that

$$f^{-1}([a,\infty))=f^{-1}(\bigcap_{n=1}^{\infty}(a-\frac{1}{n},\infty))=\cap_{n=1}^{\infty}f^{-1}((a-\frac{1}{n},\infty))\in\mathcal{M}$$

The proof for other types of intervals needs similar reformulations of them:

$$f^{-1}((-\infty,a))=f^{-1}(\mathbb{R}\setminus[a,\infty))=\mathbb{R}\setminus f^{-1}([a,\infty))\in\mathcal{M}$$
$$f^{-1}((b,a))=f^{-1}((-\infty,a))\cap f^{-1}((b,\infty))\in\mathcal{M}$$

■

R

1. From the proof above we also find: the function f is measurable if and only if $f^{-1}((a,\infty))\in\mathcal{M}$, for $\forall a\in\mathbb{R}$.
2. Homework question: the function f is measurable if and only if $f^{-1}(B)\in\mathcal{M}$ for $\forall B\in\mathcal{B}$.

Proposition 9.2 1. Constant functions, and monotone functions are measurable

2. If $A\subseteq\mathbb{R}$ is measurable, then the characterstic function

$$\mathcal{X}_A(x):=\begin{cases}1, & \text{if } x\in A\\ 0, & \text{if } x\notin A\end{cases}$$

is measurable.

3. If f is measurable, h is continuous, then $h\circ f$ is continuous.
4. If f,g are measurable, then so is

$$f+g,\quad fg,\quad \max/\min(f,g),\quad |f|$$

Proof.
- (1) and (2) are easy to show.
- The proof for (3) is simply by applying the formula

$$(h \circ f)^{-1}((a,\infty)) = f^{-1}(h^{-1}(a,\infty)),$$

- The proof for the measurability of $f+g$ is by definition:

$$\begin{aligned}(f+g)^{-1}(a,\infty) &= \{x \mid f+g \in (a,\infty)\}\\ &= \cup_{q\in\mathbb{Z}}(\{x \mid f \in (q,\infty)\} \cap \{x \mid g \in (a-q,\infty)\})\\ &= \cup_{q\in\mathbb{Z}}(f^{-1}(q,\infty) \cap f^{-1}(a-q,\infty)) \in \mathcal{M}\end{aligned}$$

The measurability of $fg, |f|, \max/\min(f,g)$ are by the equalities

$$\begin{aligned} fg &= \frac{1}{4}[(f+g)^2 + (f-g)^2]\\ |f| &= h \circ f \quad h(x) = |x|\\ \max/\min(f,g) &= \frac{1}{2}(f+g \pm |f-g|)\end{aligned}$$

∎

R If both f,g are measurable, then $g \circ f$ is not necessarily measurable.

Definition 9.2 [Almost Everywhere] Let $f,g : (\mathbb{R},\mu,m) \to \mathbb{R}$. We say $f = g$ almost everywhere (a.e.) if $E := \{x \mid f(x) \ne g(x)\}$ is a null set.

More generally, we say $f(x)$satisfies a condition on (R,μ,m) a.e. if the set

$$\{x \mid f(x) \text{ does not satisfy the condition}\} \text{ is a null set.}$$

For example, the characteristic function $\mathcal{X}_{\mathbb{Q}}(x)$ is equal to zero function a.e.

The measurability ignores the null set.

Proposition 9.3 Suppose that f is measurable, and $g = f$ a.e., then g is measurable.

Proof. Note that

$$g^{-1}((a,\infty)) = \{x \mid g(x) \in (a,\infty), g(x) = f(x)\} \cup \{x \mid g(x) \in (a,\infty), g(x) \neq f(x)\}$$

where $\{x \mid g(x) \in (a,\infty), g(x) \neq f(x)\} \subseteq E$, i.e., belongs to $\mathcal{M}$; and

$$\{x \mid g(x) \in (a,\infty), g(x) = f(x)\} = f^{-1}((a,\infty)) \cap E^c \in \mathcal{M}.$$

■

R During the proof, we have used the fact that $N \subseteq E$ is measurable for all null set E.

Definition 9.3 [Measurable on extended real line] A function $f : \mathbb{R} \to [-\infty,\infty]$ is measurable if $f^{-1}(I) \in \mathcal{M}$ for all intervals $I \in [-\infty,\infty]$.

Following the similar idea of previous examples, it suffices to show that

$$f^{-1}((a,\infty]) \in \mathcal{M}, \quad \forall a \in \mathbb{R}.$$

Or equivalently,

$$f^{-1}(B) \in \mathcal{M}, \forall B \in \mathcal{B}, \text{ and } f^{-1}(\{-\infty\}), f^{-1}(\{\infty\}) \in \mathcal{M}.$$

Example:

$$f(x) = \begin{cases} \tan x & x \neq \dfrac{2n+1}{2}\pi, n \in \mathbb{Z} \\ \infty, & x = \dfrac{2n+1}{2}\pi, n \in \mathbb{Z} \end{cases}$$

is measurable.

9.3. Monday for MAT4002

Reviewing.

1. Homotopy: we denote the homotopic function pair as $f \simeq g$.
2. If $Y \subseteq \mathbb{R}^n$ is convex, then the set of continuous functions $f : X \to Y$ form a single equivalence class, i.e., $\{\text{continuous functions } f : X \to Y\}/\sim$ has only one element

9.3.1. Remarks on Homotopy

Proposition 9.4 Consider four continous mappings

$$W \xrightarrow{f} X, \quad X \xrightarrow{g} Y, \quad X \xrightarrow{h} Y, \quad Y \xrightarrow{k} Z.$$

If $g \simeq h$, then

$$g \circ f \simeq h \circ f, \quad k \circ g \simeq k \circ h$$

Proof. Suppose there exists the homotopy $H : g \simeq h$, then $k \circ H : X \times I \to Z$ gives the momotopy between $k \circ g$ and $k \circ h$.

Simiarly, $H \circ (f \times \text{id}_I) : W \times I \to Y$ gives the homotopy $g \circ f \simeq h \circ f$. ∎

Definition 9.4 [Homotopy Equivalence] Two topological spaces X and Y are **homotopy equivalent** if there are continuous maps $f : X \to Y$, and $g : Y \to X$ such that

$$g \circ f \simeq \text{id}_{X \to X}$$
$$f \circ g \simeq \text{id}_{Y \to Y},$$

which is denoted as $X \simeq Y$.

R

1. If $X \cong Y$ are homeomorphic, then they are homotopic equivalent.
2. The homotopy equivalence $X \simeq Y$ gives a bijection between $\{\phi : \text{continuous } W \to X\}/\sim$ and $\{\phi : \text{continuous } W \to Y\}/\sim$, for any given topological space W.

Proof. Since $X \simeq Y$, we can find $f : X \to Y$ and $g : Y \to X$ such that $f \circ g \simeq \mathrm{id}_Y$ and $g \circ f \simeq \mathrm{id}_X$. We construct a mapping

$$\begin{aligned} \phi : \quad & \{\phi : \text{continuous } W \to X\}/\sim \longrightarrow \{\phi : \text{continuous } W \to Y\}/\sim \\ \text{with} \quad & [\phi] \mapsto [f \circ \phi] \end{aligned}$$

ϕ is well-defined since $\phi_1 \sim \phi_2$ implies $f \circ \phi_1 \sim f \circ \phi_2$

Also, we can construct a mapping

$$\begin{aligned} \beta : \quad & \{\phi : \text{continuous } W \to Y\}/\sim \longrightarrow \{\phi : \text{continuous } W \to X\}/\sim \\ \text{with} \quad & [\psi] \mapsto [g \circ \phi] \end{aligned}$$

Similarly, β is well-defined.

Also, we can check that $\alpha \circ \beta = \mathrm{id}$ and $\beta \circ \alpha = \mathrm{id}$. For example,

$$\alpha \circ \beta[\psi] = [f \circ g \circ \psi] = [\psi],$$

where the last equality is because that $f \circ g \simeq \mathrm{id}_Y$. ■

3. The homotopy equivalence $X \simeq Y$ forms an equivalence relation between topological spaces

Compared with homeomorphism, some properties are lost when consider the homotopy equivalence.

Definition 9.5 [Contractible] The topological space X is **contractible** if it is homotopy equivalent to any point $\{c\}$.

R In other words, there exists continuous mappings f, g such that

$$\{c\} \xrightarrow{f} X \xrightarrow{g} \{c\},\ g \circ f \simeq \mathrm{id}_{\{c\}}$$
$$X \xrightarrow{g} \{c\} \xrightarrow{f} X,\ f \circ g \simeq \mathrm{id}_X$$

Note that $g \circ f \simeq \mathrm{id}_{\{c\}}$ follows naturally; and since $X \cong X$, we can find

f, g such that $f \circ g = c_y$ for some $y \in X$, where $c_y : X \to X$ is a constant function $c_y(x) = y, \forall x \in X$. Therefore, to check X is contractible, it suffices to check $c_y \simeq \mathrm{id}_X, \forall y \in X$.

Therefore, X is contractible if its identity map id_X is homotopic to any constant map $c_y, \forall y \in X$.

■

Proposition 9.5 The definition for contractible can be simplified further:

1. X is contractible if it is homotopy equivalent to some point $\{c\}$
2. X is contractible if the identity map id_X is homotopic to some constant map $c_y(x) = y$.

Proof. The only thing is to show that $c_y \simeq c_{y'}, \forall y, y' \in X$. By hw 3, X is path-connected, and therefore there exists continous $p(t)$ such that

$$p(0) = y, \quad p(1) = y'$$

Therefore, we construct the homotopy between c_y and $c_{y'}$ as follows:

$$H(x, t) = p(t).$$

■

■ **Example 9.1** 1. $X = \mathbb{R}^2$ is contractible:

It suffices to show that the mapping $f(\boldsymbol{x}) = \boldsymbol{x}, \forall \boldsymbol{x} \in \mathbb{R}^2$ is homotopic to the constant function $g(x) = (0,0), \forall x \in \mathbb{R}^2$, i.e., $g = c_{(0,0)}$.

Consider the continuous mapping $H(\boldsymbol{x}, t) = t f(\boldsymbol{x})$, with

$$H(\boldsymbol{x}, 0) = c_{(0,0)}, \quad H(\boldsymbol{x}, 1) = \mathrm{id}_X$$

Therefore, $c_{(0,0)} \simeq \mathrm{id}_X$. Since $c_{(0,0)} \simeq c_{\boldsymbol{y}}, \forall \boldsymbol{y} \in \mathbb{R}^2$, we imply $c_{\boldsymbol{y}} \simeq \mathrm{id}_X$ for any $\boldsymbol{y} \in \mathbb{R}^2$. Therefore, X is contractible.

More generally, any convex $X \subseteq \mathbb{R}^n$ is contractible.

(R) S^1 is not contractible, and we will see it in 3 weeks' time. In particular, we are not able to construct the continuous mapping

$$H : S^1 \times [0,1] \to S^1$$

such that

$$H(e^{2\pi ix},0) = e^{2\pi ix}, \quad H(e^{2\pi ix},1) = e^{2\pi i(0)} = 1$$

How about the mapping $H(e^{2\pi ix},t) = e^{2\pi ixt}$? Unfortunately, it is not well-defined, since

$$H(e^{2\pi i(1)},t) = e^{2\pi it} = H(e^{2\pi i(0)},t) = 1$$

and the equality is not true for $t \neq 0,1$.

Definition 9.6 [Homotopy Retract] Let $A \subseteq X$ and $i : A \hookrightarrow X$ be an inclusion. We say A is a **homotopy retract** of X if there exists continuous mapping $r : X \to A$ such that

$$r \circ i : A \hookrightarrow X \xrightarrow{r} A = \mathrm{id}_A$$

$$i \circ r : X \xrightarrow{r} A \hookrightarrow X \simeq \mathrm{id}_X$$

In particualr, $A \simeq X$.

■ **Example 9.2** The 1-sphere S^1 is a homotopy retract of Mobius band M.

Let $M = [0,1]^2/\sim$ and $S^1 = [0,1]/\sim$. Define the inclusion i and r as:

$$\begin{aligned} i : \quad & S^1 \hookrightarrow M \\ \text{with} \quad & [x] \mapsto [(x, \tfrac{1}{2})] \end{aligned}$$

$$\begin{aligned} r: \quad & M \to S^1 \\ \text{with} \quad & [(x,y)] \mapsto [x] \end{aligned}$$

As a result,

$$r \circ i = \mathrm{id}_{S^1}, \qquad i \circ r([(x,y)]) = [(x, 1/2)]$$

It suffices to show $i \circ r \simeq \mathrm{id}_M$, where $\mathrm{id}_M([(x,y)]) = [(x,y)]$.

Construct the continous mapping $H : M \times I \to M$ with

$$H([(x,y)],t) := [(x,(1-t)y + t/2)]$$

To show the well-definedness of H, we need to check

$$H([(0,y)],t) = H([(1,1-y)],t), \quad \forall y \in [0,1]$$

It's clear that H gives a homotopy between $i \circ r$ and id_M, i.e., $i \circ r \simeq \mathrm{id}_M$ ■

■ **Example 9.3** The $n-1$-sphere S^{n-1} is a homotopy retract of $\mathbb{R}^n \setminus \{\mathbf{0}\}$:

We have the inclusion $i : S^{n-1} \to \mathbb{R}^n \setminus \{0\}$ and

$$\begin{aligned} r: \quad & \mathbb{R}^n \setminus \{0\} \to S^{n-1} \\ \text{with} \quad & x \mapsto \frac{x}{\|x\|} \end{aligned}$$

Therefore, $r \circ i = \mathrm{id}_{S^{n-1}}$ and $i \circ r(x) = \frac{x}{\|x\|}$.

It suffices to show that $i \circ r \simeq \mathrm{id}_{\mathbb{R}^n \setminus \{0\}}$. Consider the homotopy $H(x,t) = t\boldsymbol{x} + (1-t)\boldsymbol{x}/\|\boldsymbol{x}\|$ such that

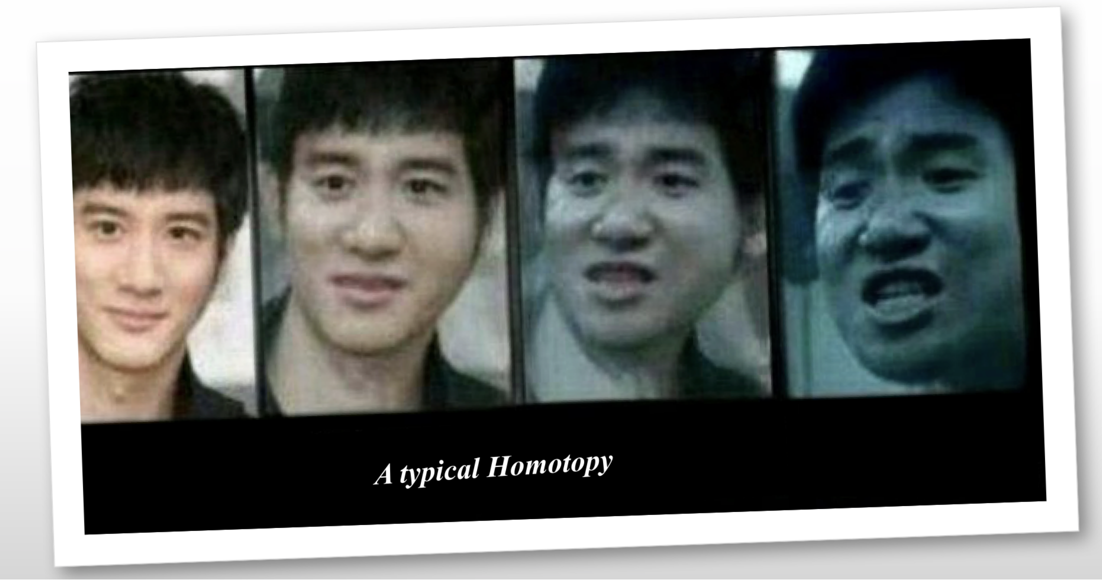
$$H(\boldsymbol{x},0) = i \circ r(\boldsymbol{x}), \quad H(\boldsymbol{x},1) = \boldsymbol{x} = \mathrm{id}(\boldsymbol{x})$$

To show the well-definedness of H, we need to check $H(x,t) \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ for all $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $t \in [0,1]$. ■

Definition 9.7 [Homotopic Relative] Let $A \subseteq X$ be topological spaces. We say $f, g : X \to$

Y are homotopic relative to A if there eixsts $H : X \times I \longrightarrow Y$ such that

$$\begin{cases} H(x,0) = f(x) \\ H(x,1) = g(x) \end{cases} \quad \text{and } H(a,t) = f(a) = g(a), \forall a \in A$$



9.4. Wednesday for MAT3040

9.4.1. Jordan Normal Form

Theorem 9.3 — Jordan Normal Form. Suppose that $T : V \to V$ has minimial polynomial

$$m_T(x) = \prod_{i=1}^{k} (x - \lambda_i)^{e_i},$$

then there exists a basis $\mathcal{A}$ such that

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(J_1, \dots, J_\ell),$$

where each block J_i is a square matrix of the form

$$J_i = \begin{bmatrix} \mu_i & 1 & & \\ & \mu_i & \ddots & \\ & & \ddots & 1 \\ & & & \mu_i \end{bmatrix}.$$

R By primary decomposition theorem,

$$V = V_1 \oplus \cdots \oplus V_k, \quad \text{where } V_i = \ker((T - \lambda_i I)^{e_i}),\ i = 1, \dots, k,$$

and each V_i is T-invariant.

We pick basis $\mathcal{B}_i$ for each subspace V_i, then $\mathcal{B} := \cup_{i=1}^{k} \mathcal{B}_i$ is a basis of V, and

$$(T)_{\mathcal{B},\mathcal{B}} = \begin{pmatrix} (T\mid_{V_1})_{\mathcal{B}_1,\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & (T\mid_{V_2})_{\mathcal{B}_2,\mathcal{B}_2} & \vdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \vdots & (T\mid_{V_k})_{\mathcal{B}_k,\mathcal{B}_k} \end{pmatrix}$$

with $m_{T|_{V_i}}(x) = (x - \lambda_i)^{e_i}$.

Therefore, it suffices to show the Jordan normal form holds for the linear operator

T with minimal polynomial $m_T(x) = (x-\lambda)^e$.

Firstly, we consider the case where the minimal polynomial has the form x^m:

Proposition 9.6 Suppose $T : V \to V$ is such that $m_T(x) = x^m$, then the theorem (9.3) holds, i.e., there exists a basis $\mathcal{A}$ such that

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_\ell),$$

where each block J_i is a square matrix of the form

$$J_i = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

Proof.

- Suppose that $m_T(x) = x^m$, then it is clear that

$$\{0\} := \ker(T^0) \le \ker(T) \le \ker(T^2) \le \cdots \le \ker(T^m) := V$$

  Furthermore, we have $\ker(T^{i-1}) \subsetneqq \ker(T^i)$ for $i = 1, \dots, m$: Note that $\ker(T^{m-1}) \subsetneqq \ker(T^m) := V$ due to the minimality of $m_T(x)$; and $\ker(T^{m-2}) \subsetneqq \ker(T^{m-1})$ since otherwise for any $\boldsymbol{x} \in \ker(T^m)$,

$$T^{m-1}(T\boldsymbol{x}) = \boldsymbol{0} \implies T\boldsymbol{x} \in \ker(T^{m-1}) = \ker(T^{m-2}) \implies T^{m-2}(T\boldsymbol{x}) = T^{m-1}(\boldsymbol{x}) = \boldsymbol{0},$$

  i.e., $\boldsymbol{x} \in \ker(T^{m-1})$, which contradicts to the fact that $\ker(T^{m-1}) \subsetneqq \ker(T^m)$. Proceeding this trick sequentially for $i = m, m-1, \dots, 1$, we proved the disired result.

- Then construct the quotient space $W_i = \ker(T^i)/\ker(T^{i-1})$ and define $\mathcal{B}_i'$ to be a basis of W_i:

$$\mathcal{B}_i' = \{a_1^i + \ker(T^{i-1}), \dots, a_{\ell_i}^i + \ker(T^{i-1})\}$$

  Construct $\mathcal{B}_i = \{a_1^i, \dots, a_{\ell_i}^i\}$, then we claim that $B := \cup_{i=1}^m \mathcal{B}_i$ forms a basis of V:

- First proof the case $m = 2$ first: let $U \le V$ ($\dim(V) < \infty$), and $\mathcal{B}_1 = \{a_1^1, \dots, a_{k_1}^1\}$ be a basis of U, and

$$\mathcal{B}_2' = \{a_1^2 + U, \dots, a_{k_2}^2 + U\}$$

be a basis of V/U. Then to show the statement suffices to show that

$$\bigcup_{i=1}^{2} \{a_1^i, \dots, a_{k_i}^i\} \text{ forms a basis of } V.$$

It's clear that $\cup_{i=1}^{2} \{a_1^i, \dots, a_{k_i}^i\}$ spans V. Furthermore, $\dim(V) = \dim(U) + \dim(V/U) = k_1 + k_2$, i.e., $\cup_{i=1}^{2} \{a_1^i, \dots, a_{k_i}^i\}$ contains correct amount of vectors. The proof is complete.

- This result can be extended from 2 to general m, thus the claim is shown.

• For $i < m$, consider the set $S_i = \{T(\boldsymbol{w}_j) + \ker(T^{i-1}) \mid \boldsymbol{w}_j \in B_{i+1}\}$. Note that

  - Since $T^{i+1}(\boldsymbol{w}_j) = \mathbf{0}$, $T^i(T(\boldsymbol{w}_j)) = \mathbf{0}$, we imply $T(\boldsymbol{w}_j) \in \ker(T^i)$, i.e., $S_i \subseteq W_i$.
  - The set S_i is linearly independent: consider the equation

$$\sum_j k_j(T(\boldsymbol{w}_j) + \ker(T^{i-1})) = \mathbf{0}_{W_i} \Longleftrightarrow T\left(\sum_j k_j \boldsymbol{w}_j\right) + \ker(T^{i-1}) = \mathbf{0}_{W_i}$$

i.e.,

$$T\left(\sum_j k_j \boldsymbol{w}_j\right) \in \ker(T^{i-1}) \Longleftrightarrow T^{i-1}(T(\sum_j k_j \boldsymbol{w}_j)) = \mathbf{0}_V,$$

i.e., $\sum_j k_j \boldsymbol{w}_j \in \ker(T^i)$, i.e.,

$$\sum_j k_j \boldsymbol{w}_j + \ker(T^i) = \mathbf{0}_{W_{i+1}} \Longleftrightarrow \sum_j k_j(\boldsymbol{w}_j + \ker(T^i)) = \mathbf{0}_{W_{i+1}}.$$

Since $\{\boldsymbol{w}_j + \ker(T^i),\ \forall j\}$ fomrs a basis of W_{i+1}, we imply $k_j = 0, \forall j$.

From $\mathcal{B}_{i+1}$ we construct S_i, which is linearly independent in W_i. Therefore, we imply $|T(\mathcal{B}_{i+1})| \le |\mathcal{B}_i|$ for $\forall i < m$ (why?).

• Now we start to construct a basis $\mathcal{A}$ of V:

– Start with $\mathcal{B}'_m := \{u_1^m + \ker(T^{m-1}), \dots, u_{\ell_m}^m + \ker(T^{m-1})\}$, and $\mathcal{B}_m = \{u_1^m, \dots, u_{\ell_m}^m\}$.

– By the previous result,

$$\{T(u_1^m) + \ker(T^{m-2}), \dots, T(u_{\ell_m}^m) + \ker(T^{m-2})\}$$

is linear independent in W_{m-1}. By basis extension, we get a basis $\mathcal{B}'_{m-1}$ of W_{m-1}, and let

$$\mathcal{B}_{m-1} = \{T(u_1^m), \dots, T(u_{\ell_m}^m)\} \cup \xi_{m-1}$$

where $\xi_{m-1} := \{u_1^{m-1}, \dots, u_{\ell_{m-1}}^{m-1}\}$

– Continue the process above to obtain $\mathcal{B}_{m-2}, \dots, \mathcal{B}_1$, and $\cup_{i=1}^m \mathcal{B}_i$ forms a basis of V:

$\mathcal{B}_1$	$\mathcal{B}_2$	$\cdots$	$\mathcal{B}_{m-1}$	$\mathcal{B}_m$
$\{T^{m-1}(u_1^m), \dots, T^{m-1}(u_{\ell_m}^m)\}$	$\{T^{m-2}(u_1^m), \dots, T^{m-2}(u_{\ell_m}^m)\}$	$\cdots$	$\{T(u_1^m), \dots, T(u_{\ell_m}^m)\}$	$\{u_1^m, \dots, u_{\ell_m}^m\}$
$\{T^{m-2}(u_1^{m-1}), \dots, T^{m-2}(u_{\ell_{m-1}}^{m-1})\}$	$\{T^{m-3}(u_1^{m-1}), \dots, T^{m-3}(u_{\ell_{m-1}}^{m-1})\}$	$\cdots$	$\{u_1^{m-1}, \dots, u_{\ell_{m-1}}^{m-1}\}$	
$\vdots$	$\vdots$	$\cdots$		
$\{T(u_1^2), \dots, T(u_{\ell_2}^2)\}$	$\{u_1^2, \dots, u_{\ell_2}^2)\}$			
$\{u_1^1, \dots, u_{\ell_1}^1)\}$				

- Now construct the ordered basis $\mathcal{A}$:

$$\mathcal{A}=\left\{\begin{array}{ccccc}
T^{m-1}(u_1^m) & \cdots & T^2(u_1^m) & T(u_1^m) & u_1^m \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
T^{m-1}(u_{\ell_m}^m) & \cdots & T^2(u_{\ell_m}^m) & T(u_{\ell_m}^m) & u_{\ell_m}^m \\
 & T^{m-2}(u_1^{m-1}) & \cdots & T(u_1^{m-1}) & u_1^{m-1} \\
 & \vdots & \ddots & \vdots & \vdots \\
 & T^{m-2}(u_{\ell_{m-1}}^{m-1}) & \cdots & T(u_{\ell_{m-1}}^{m-1}) & u_{\ell_{m-1}}^{m-1} \\
 & & \vdots & \ddots & \vdots \\
 & & & & u_1^1 \\
 & & & & \vdots \\
 & & & & u_{\ell_1}^1
\end{array}\right\}$$

- Then the diagonal entries of $(T)_{\mathcal{A},\mathcal{A}}$ should be all zero, since

$$T(T^{i-1}(u_j^i))=T^i(u_j^i)=0,\forall i=1,\dots,m,j=1,\dots,\ell_i,$$

and every entry on the superdiagonal is 1:

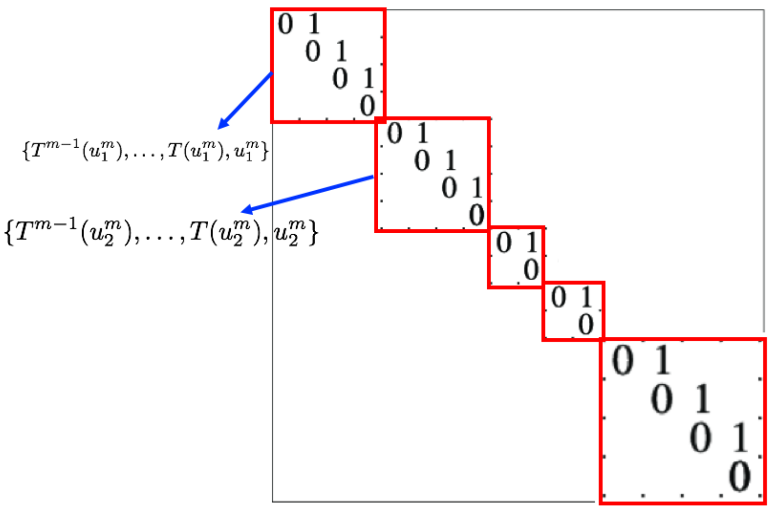


Figure 9.2: Illustration for $(T)_{\mathcal{A},\mathcal{A}}$

■

Then we consider the case where $m_T(x)=(x-\lambda)^e$:

Corollary 9.3 Suppose $T : V \longrightarrow V$ is such that $m_T(x) = (x - \lambda)^e$, then the theorem (9.3) holds, i.e., there exists a basis $\mathcal{A}$ such that

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(J_1, \dots, J_\ell),$$

where each block J_i is a square matrix of the form

$$J_i = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}.$$

Proof. Suppose that $m_T(x) = (x - \lambda)^e$. Consider the operator $U := T - \lambda I$, then $m_U(x) = x^e$.

By applying proposition (9.6),

$$(U)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_\ell),$$

where

$$J_i = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

Or equivalently,

$$(T)_{\mathcal{A},\mathcal{A}} - \lambda(I)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_\ell)$$

i.e.,

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\boldsymbol{K}_1, \dots, \boldsymbol{K}_\ell),$$

where

$$K_i = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

■

(R) The Jordan Normal Form Theorem (9.3) follows from our arguments using the primary decomposition.

Corollary 9.4 Any matrix $A \in M_{n\times n}(\mathbb{C})$ is similar to a matrix of the Jordan normal form

$$\text{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_\ell).$$

9.4.2. Inner Product Spaces

Definition 9.8 [Bilinear] Let V be a vector space over $\mathbb{R}$. A bilinear form on V is a mapping

$$F : V \times V \to \mathbb{R}$$

satisfying

1. $F(\boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}) = F(\boldsymbol{u},\boldsymbol{w}) + F(\boldsymbol{v},\boldsymbol{w})$
2. $F(\boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}) = F(\boldsymbol{u},\boldsymbol{v}) + F(\boldsymbol{u},\boldsymbol{w})$
3. $F(\lambda\boldsymbol{u},\boldsymbol{v}) = \lambda F(\boldsymbol{u},\boldsymbol{v}) = F(\boldsymbol{u},\lambda\boldsymbol{v})$

We say

- F is symmetric if $F(\boldsymbol{u},\boldsymbol{v}) = F(\boldsymbol{v},\boldsymbol{u})$
- F is non-degenerate if $F(\boldsymbol{u},\boldsymbol{w}) = \boldsymbol{0}$ for $\forall \boldsymbol{u} \in V$ implies $\boldsymbol{w} = 0$

- F is positive definite if $F(\boldsymbol{v},\boldsymbol{v}) > 0$ for $\forall \boldsymbol{v} \neq \mathbf{0}$

R If F is positive-definite, then F is non-degenerate: Suppose that $F(\boldsymbol{v},\boldsymbol{v}) > 0, \forall \boldsymbol{v} \neq \mathbf{0}$. If we have $F(\boldsymbol{u},\boldsymbol{w}) = 0$ for any $\boldsymbol{u} \in V$, then in particular, when $\boldsymbol{u} = \boldsymbol{w}$, we imply $F(\boldsymbol{w},\boldsymbol{w}) = 0$. By positive-definiteness, $\boldsymbol{w} = \mathbf{0}$, i.e., F is non-degenerate.

9.5. Wednesday for MAT3006

9.5.1. Remarks on Measurable function

Proposition 9.7 Let f_n be a sequence of measurable functions $f_n : \mathbb{R} \to [-\infty, \infty]$. Then the functions

$$\sup_{n\in\mathbb{N}} f_n(x), \quad \inf_{n\in\mathbb{N}} f_n(x), \quad \lim_{n\to\infty}\sup f_n(x), \quad \lim_{n\to\infty}\inf f_n(x)$$

are measurable.

Proof.

- $$\begin{aligned}\left(\sup_{n\in\mathbb{N}} f_n\right)^{-1}((a,\infty]) &= \{x \in \mathbb{R} \mid \sup_n f_n(x) > a\} \\ &= \{x \in \mathbb{R} \mid f_n(x) > a \text{ for some } a\} \\ &= \bigcup_{n\in\mathbb{N}} f_n^{-1}((a,\infty])\end{aligned}$$

  which is measurable due to the measurability of f_n.

- The proof for the measurablitiy of $\inf_n f_n(x), \lim_{n\to\infty}\sup f_n(x), \lim_{n\to\infty}\inf f_n(x)$ is directly by applying the formula

$$\begin{aligned}\inf f_n(x) &= -(\sup(-f_n(x))) \\ \lim_{n\to\infty}\sup f_n(x) &= \lim_{m\to\infty}(\sup_{n\geq m} f_n(x)) = \inf_{m\in\mathbb{N}}(\sup_{n\geq m} f_n(x)) \\ \lim_{n\to\infty}\inf f_n(x) &= -\lim_{n\to\infty}\sup(-f_n(x))\end{aligned}$$

■

Corollary 9.5 If $\{f_n\}$ is measurable, and $f_n(x)$ converges to $f(x)$ pointwisely a.e., then f is measurable.

Proof. By proposition (9.3), w.l.o.g., $f_n(x)$ conveges to $f(x)$ pointwisely, which follows that

$$f(x) := \lim_{n\to\infty} f_n = \lim_{n\to\infty} \sup f_n(x)$$

i.e., f is measurable due to the measurability of $\lim_{n\to\infty} \sup f_n(x)$. ■

9.5.2. Lebesgue Integration

Definition 9.9 [Simple Function] A function $\phi : \mathbb{R} \to \mathbb{R}$ is **simple** if

- ϕ is measurable and
- $\{\phi(x) \mid x \in \mathbb{R}\}$ takes finitely many values.

More precisely, if the simple function ϕ takes distinct values $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$ on disjoint non-empty sets $A_1, \ldots, A_k \subseteq \mathbb{R}$, then

$$\phi = \sum_{i=1}^{k} \alpha_i \mathcal{X}_{A_i}$$

Note that A_i's are measurable since $\phi^{-1}(\{\alpha_i\}) = A_i$

R

1. All functions written in the form $\psi = \sum_{i=1}^{\ell} \beta_i \mathcal{X}_{B_i}$, where B_i's are measurable, are simple; All simple functions can be expressed as the form $\psi = \sum_{i=1}^{\ell} \beta_i \mathcal{X}_{B_i}$ (where B_i's are disjoint) uniquely, up to permutation of terms. This is called the canonical form.
2. If ϕ_1, ϕ_2 are simple, then so are

$$\phi_1 + \phi_2, \quad \phi_1 \cdot \phi_2, \quad \alpha \cdot \phi, \max(\phi_1, \phi_2), \quad h \circ \phi.$$

for all function h.

Definition 9.10 [Lebesgue integral for Simple Function] Given a simple function with the canonical form $\phi := \sum_{i=1}^{k} \alpha_i \mathcal{X}_{\mathcal{A}_i}$,

- The Lebesgue integral for ϕ (over $\mathbb{R}$) is

$$\int \phi \, \mathrm{d}m = \sum_{i=1}^{k} \alpha_i m(A_i),$$

- The Lebesgue integral for ϕ over a measurable set E is

$$\int_E \phi \, \mathrm{d}m = \int \phi \cdot \mathcal{X}_E \, \mathrm{d}m = \sum_{i=1}^{k} \alpha_i m(A_i \cap E)$$

Proposition 9.8 For any simple function $\phi = \sum_{i=1}^{\ell} \beta_i \mathcal{X}_{B_i}$, where B_i's are not necessarily disjoint, we still have

$$\begin{aligned} \int \phi \, \mathrm{d}m &= \sum_{i=1}^{\ell} \beta_i m(B_i), \\ \int (\phi + \psi) \, \mathrm{d}m &= \int \phi \, \mathrm{d}m + \int \psi \, \mathrm{d}m, \quad \text{where } \psi \text{ is another simple function,} \\ \int \phi \, \mathrm{d}m &\le \int \psi \, \mathrm{d}m, \qquad \text{provided that } \phi \le \psi. \end{aligned}$$

Proof. It suffices to show the first equality. w.l.o.g., suppose $\phi = \beta_1 \mathcal{X}_{B_1} + \beta_2 \mathcal{X}_{B_2}$, which can be reformulated as the canonical form:

$$\phi = (\beta_1 + \beta_2) \mathcal{X}_{B_1 \cap B_2} + \beta_1 \mathcal{X}_{B_1 \cap B_2^c} + \beta_2 \mathcal{X}_{B_1^c \cap B_2}$$

Then we can take the Lebesgue integration for ϕ:

$$\int \phi \, \mathrm{d}m = (\beta_1 + \beta_2) m(B_1 \cap B_2) + \beta_1 m(B_1 \cap B_2^c) + \beta_2 m(B_1^c \cap B_2),$$

which is equal to $\beta_1 m(B_1) + \beta_2 m(B_2)$ due to the caratheodory property (definition (8.2)) ■

Definition 9.11 [Lebesgue integral for Measurable Function] Let f be a measurable function $f:\mathbb{R}\to[0,\infty]$. Then the Lebesgue integral of f is given by:

$$\int f\,\mathrm{d}m = \sup\left\{\int \phi\,\mathrm{d}m \middle| 0\le\phi\le f, \phi \text{ is simple}\right\} \tag{9.3}$$

We say f is integrable if $\int f\,\mathrm{d}m < \infty$.

R

- It's not appropriate if we try to define the Lebesgue integral by

$$\int f\,\mathrm{d}m = \inf\left\{\int \phi\,\mathrm{d}m \middle| 0\le f\le\phi, \phi \text{ is simple}\right\} \tag{9.4}$$

  The problem is due to the function $f(x)=\frac{1}{\sqrt{x}}$ on $(0,1)$. Note that the function values can be arbitrarily large.

  Since a simple function takes only finitely many values, every simple function that is bounded below by f has to be infinite on a set of non-zero measure.

  Therefore, the integral using your suggested infimum definition would be ∞, whereas the usual Lebesgue integral would have a finite value.

- Also, one can try to define $\int f\,\mathrm{d}m$ for non-measurable function f. The problem is that

$$\int (f+g)\,\mathrm{d}m \ne \int f\,\mathrm{d}m + \int g\,\mathrm{d}m \text{ in general}$$

  We will see the detailed reason later.

Proposition 9.9

- The formula (9.3) and (9.4) matches with each other for any simple functions $\phi\ge 0$.
- For $\alpha\ge 0$,

$$\int \alpha f\,\mathrm{d}m = \alpha\int f\,\mathrm{d}m$$

- If $0 \le f \le g$, then
$$\int f \, \mathrm{d}m \le \int g \, \mathrm{d}m$$

Proof. omitted. ■

Proposition 9.10 — Markov Inequality. Suppose that $f : \mathbb{R} \to [0, \infty]$ is measurable, then

$$m(f^{-1}[\lambda, \infty]) \le \frac{1}{\lambda} \int f \, \mathrm{d}m$$

Corollary 9.6 If $f : \mathbb{R} \to [0, \infty]$ is integrable, then $m(f^{-1}\{\infty\}) = 0$, i.e., f is finite a.e.

Proof.

$$m(f^{-1}\{\infty\}) \le m(f^{-1}[\lambda, \infty]) \le \frac{1}{\lambda} \int f \, \mathrm{d}m, \ \forall \lambda \ge 0.$$

Since $\int f \, \mathrm{d}m$ is finite, we imply $\frac{1}{\lambda} \int f \, \mathrm{d}m$ can be arbitrarily small, i.e., $m(f^{-1}\{\infty\}) = 0$. ■

9.6. Wednesday for MAT4002

9.6.1. Simplicial Approximation Theorem

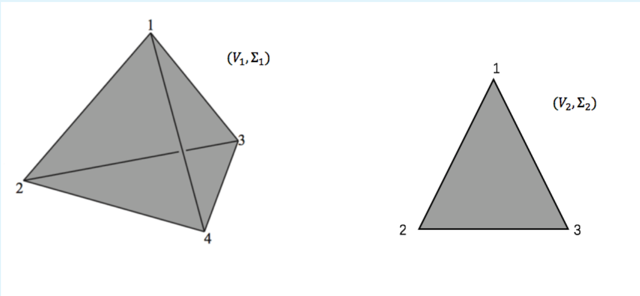
Aim: understand homotopy between simplicial complexes $f, g : |K| \to |L|$

Definition 9.12 [Simplicial Map] A **simplicial map** between $K_1 = (V_1, \Sigma_1)$ and $K_2 = (V_2, \Sigma_2)$ is a mapping $f : K_1 \to K_2$ such that

1. It maps vertexes to vertexes
2. It maps simplicies to simplicies, i.e.,

$$f(\sigma_1) \in \Sigma_2, \ \forall \sigma_1 \in \Sigma_1,$$

■ **Example 9.4** For instance, consider the simplicial complexes defined as follows:



In particular, $\{1,2,3,4\} \notin \Sigma_1$ and $\{1,2,3\} \in \Sigma_2$.

In this case, we can define the simplicial map as:

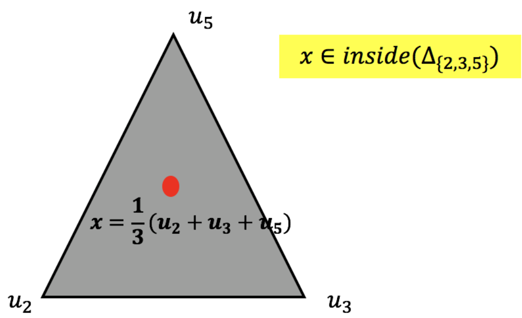
$$f(1) = 1, \quad f(2) = 2, \quad f(3) = 3, \quad f(4) = 3$$

In particular, $f(\{1,2,4\}) = \{1,2,3\} \in \Sigma_2$.

Now we want to define the simplicial map between the topological realizations. There are several observations:

Key Observations.

1. We have seen that each $|K| \subseteq \mathbb{R}^m$ for some m. In particular, $m = \#V - 1$.
2. Each point $x \in |K|$ lies uniquely on an inside of some Δ_σ,, where $\sigma \in \Sigma$.
3. Suppose that the vertices of K_1 are $V_1 = \{u_1, \ldots, u_n\} \subseteq \mathbb{R}^m$. Then every $\boldsymbol{x} \in K_1$ can be uniquely written as
$$\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i U_{\sigma_i}$$
with $\alpha_i > 0, \sum \alpha_i = 1$ and $\sigma = \{U_{\sigma_1}, \ldots, U_{\sigma_k}\}$ is the unqiue simplex where $x \in$ inside(Δ_σ).



4. Our simplicial map f maps V_1 to $V_2 = \{w_1, \ldots, w_p\} \subseteq \mathbb{R}^m$, so for each i, we have $f(\boldsymbol{u}_i) = \boldsymbol{w}_j$ for some $j \in \{1, \ldots, p\}$.

Definition 9.13 [Mapping induced from Simplicial Mapping] The simplicial map $f : K_1 \to K_2$ induces a mapping $|f| : |K_1| \to |K_2|$ between the topological realizations such that

1. It maps vertexes to vertexes, i.e., $|f|(v_1) = f(v_1), \forall v_1 \in V(K_1)$.
2. it is affine, i.e.,
$$|f|\left(\sum_{i=1}^{k} \alpha_i v_i\right) = \sum_{i=1}^{k} \alpha_i |f|(v_i)$$

R $|f| : |K_1| \to |K_2|$ is continuous.

Motivation. Suppose we are given a continuous map $|g| : |K| \to |L|$, we want to approximate $|g|$ by $|f|$, such that $f : K \to L$ is a simplicial map. In this case, f is an easier object to study compared with $|g|$.

We hope to find a mapping f such that $|f| \simeq |g|$. However, we cannot achieve this goal unless we subdivide K into smaller pieces:

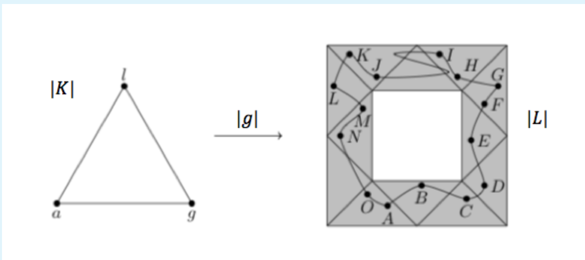
Definition 9.14 [Subdivision] Let K be a simplicial complex. A simplicial complex K' is called a **subdivision** of K if

1. Each simplex of K' is contained in a simplex of K
2. Each simplex of K equals the union of finitely many simplices of K'

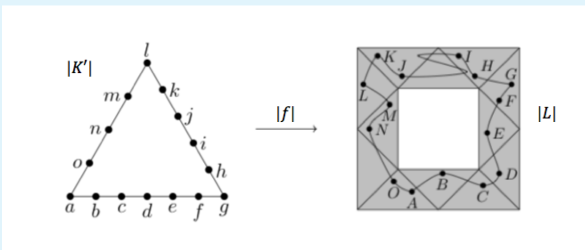
As a result, we can form an homeomorphism $h : |K'| \to |K|$ such that for each $\sigma' \in \Sigma_{K'}$, there exists $\sigma \in \Sigma_K$ satisfying

$$f(\Delta_{\sigma'}) \in \Delta_{\sigma}$$

■ **Example 9.5** Consider the mapping $|g| : |K| \to |L|$ given in the figure below:



Here we denote $|g|(a)$ by A and similarly for the other vertices. It's clear that we can not form a homeomorphism from $|K|$ to $|L|$. One remedy is to subdivide K into smaller pieces as follows:



In this case, it is clear that $|f| : |K'| \to |L|$ is a homeomorphism. ■

■ **Example 9.6** [Barycentric Subdivision] One typical subdivision is the Barycentric Subdivision:

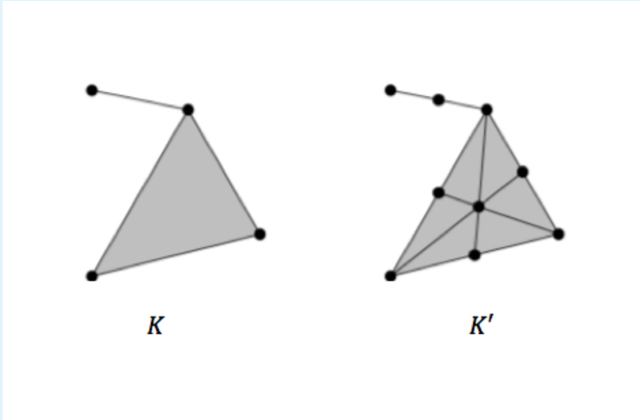


Figure 9.3: Right: the subdivision of K

■

R Suppose we have a matric on $|K|$. By subdivision, we can consider $|K'|$ such that for any $\sigma' \in \Sigma_{K'}$, any two points in $\Delta_{\sigma'}$ has a smaller distance.

The following result gives a criterion for the existence of a simplicial approximation for a mapping between topological realizations. For this we recall the notion of star. For a given simplicial complex K, define the star at a vertex v by

$$\text{star}(v) = \bigcup_{v \in \sigma} \sigma^\circ.$$

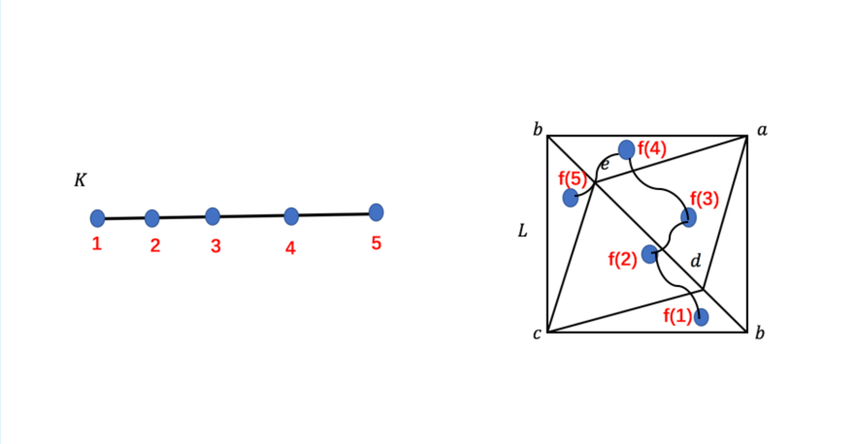
Proposition 9.11 Let $f : |K| \to |L|$ be a continuous mapping. Suppose that for each $v \in V_K$, there exists $g(v) \in V_L$ such that

$$f(\text{st}_K(v)) \subseteq \text{st}_L(g(v)),$$

then the mapping $g : V_K \to V_K$ gives $|g| \simeq f$.

In particular, g is called the **simplicial approximation** to f.

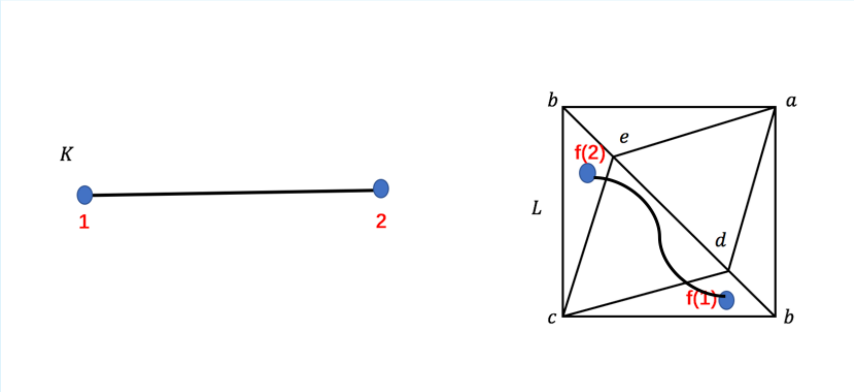
■ **Example 9.7** 1. First, we give an example of mapping f such that the assumption in proposition (9.11) is satisfied and therefore an simplicial approximation exists:



We could define the simplicial approximation g with

$$g(1) = b, g(2) = e, g(3) = e, g(4) = d, g(5) = d \text{ or } c$$

2. In the example below, the hypothesis of proposition (9.11) is not satisfied, so we cannot apply this proposition to construct a simplicial map.



■

Theorem 9.4 — Simplicial Approximation. Let K, L be simplicial complexes with V_K finite, and $f : |K| \to |L|$ be continuous. Then there eixsts a subdivision $|K'|$ of $|K|$ and a simplicial map g such that $|g| \simeq f$.

Chapter 10

Week10

10.1. Monday for MAT3040

10.1.1. Inner Product Space

- Symmetric: $F(\boldsymbol{u},\boldsymbol{w}) = F(\boldsymbol{w},\boldsymbol{u}), \forall \boldsymbol{u},\boldsymbol{w}$
- Non-degenerate: $F(\boldsymbol{u},\boldsymbol{w}) = 0, \forall \boldsymbol{w}$ implies $\boldsymbol{u} = \mathbf{0}$
- Positive definite: $F(\boldsymbol{v},\boldsymbol{v}) > 0, \forall \boldsymbol{v} \neq \mathbf{0}$

Classification. When we say V be a vector space over $\mathbb{F}$, we treat $\alpha \in \mathbb{F}$ as a scalar.

Definition 10.1 [Sesqui-linear Form] Let V be a vector space over $\mathbb{C}$. A **sesquilinear form** on V is a function $F : V \times V \to \mathbb{C}$ such that

1. $F(\boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}) = F(\boldsymbol{u},\boldsymbol{w}) + F(\boldsymbol{v},\boldsymbol{w})$
2. $F(\boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}) = F(\boldsymbol{u},\boldsymbol{v}) + F(\boldsymbol{u},\boldsymbol{w})$
3. $F(\overline{\lambda}\boldsymbol{v},\boldsymbol{w}) = F(\boldsymbol{v},\lambda\boldsymbol{w}) = \lambda F(\boldsymbol{v},\boldsymbol{w}), \forall \lambda \in \mathbb{C}$

In this case, we say F is **conjugate symmetric** if

$$F(\boldsymbol{v},\boldsymbol{w}) = \overline{F(\boldsymbol{w},\boldsymbol{v})}, \quad \forall \boldsymbol{v},\boldsymbol{w} \in V.$$

The definition for non-degenerateness, and positve definiteness is the same as that in bilinear form.

R In the sesquilinear form, why there is a $\bar{\lambda}$ shown in condition (3)?

Partial Answer: We want our F to be positive definite in many cases:

- Suppose that $F(\boldsymbol{v},\boldsymbol{v}) > 0$ and we do not have $\bar{\lambda}$ in sesquilinear form F, it follows that
$$F(i\boldsymbol{v},i\boldsymbol{v}) = i^2 F(\boldsymbol{v},\boldsymbol{v}) = -F(\boldsymbol{v},\boldsymbol{v}) < 0$$
As a result, there will be no positive bilinear form for vector space over $\mathbb{C}$.

Therefore, $\bar{\lambda}$ is essential to guarantee that we have a positive definite form on vector space over $\mathbb{C}$, i.e.,

$$F(i\boldsymbol{v},i\boldsymbol{v}) = \bar{i}iF(\boldsymbol{v},\boldsymbol{v}) = F(\boldsymbol{v},\boldsymbol{v})$$

▪ **Example 10.1** Consider $V = \mathbb{C}^n$, and a basic sesquilinear form is the Hermitian inner product:

$$F(\boldsymbol{v},\boldsymbol{u}) = \boldsymbol{v}^{\mathrm{H}}\boldsymbol{u} = \begin{pmatrix} \bar{v}_1 & \cdots & \bar{v}_n \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \sum_{i=1}^{n} \bar{v}_i w_i$$

In this case, we do not have symmetric property $F(\boldsymbol{v},\boldsymbol{w}) = F(\boldsymbol{w},\boldsymbol{v})$ any more, instead, we have the conjugate symmetric property $F(\boldsymbol{v},\boldsymbol{w}) = \overline{F(\boldsymbol{w},\boldsymbol{v})}$. ▪

Definition 10.2 [Inner Product] A real (complex) vector space V with a bilinear (sesquilinear) form with symmetric (conjugate symmetric) and positive definite property is called an **inner product** on V. Any vector space equipped with inner product is called an **inner product space**. ▪

Notation. We write $\langle\cdot,\cdot\rangle$ instead of $F(\cdot,\cdot)$ to denote inner product.

Definition 10.3 [Norm] The **norm** of a vector $\boldsymbol{v}$ is $\|\boldsymbol{v}\| = \sqrt{\langle\boldsymbol{v},\boldsymbol{v}\rangle}$. ▪

(R) As a result, $\|\alpha\boldsymbol{v}\| = \sqrt{\langle\alpha\boldsymbol{v},\alpha\boldsymbol{v}\rangle} = \sqrt{\bar{\alpha}\alpha\langle\boldsymbol{v},\boldsymbol{v}\rangle} = \sqrt{|\alpha|^2\langle\boldsymbol{v},\boldsymbol{v}\rangle} = |\alpha|\|\boldsymbol{v}\|$.

The norm is well-defined since $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ (positive definiteness of inner product).

Definition 10.4 [Orthogonal] We say a family of vectors $S = \{\boldsymbol{v}_i \mid i \in I\}$ is **orthogonal** if

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0, \ \forall i \neq j$$

If furthermore $\langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle = 1, \forall i$, then we say S is an **orthonormal** set. ■

R

1. The Cauchy-Scharwz inequality holds for inner product space:

$$|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\| \|\boldsymbol{v}\|, \ \forall \boldsymbol{u}, \boldsymbol{v} \in V.$$

Proof. The proof for $\langle \boldsymbol{u}, \boldsymbol{v} \rangle \in \mathbb{R}$ is the same as in MAT2040 course. Check Theorem (6.1) in the note

`https://walterbabyrudin.github.io/information/Notes/MAT2040.pdf`

However, for $\langle \boldsymbol{u}, \boldsymbol{v} \rangle \in \mathbb{C} \setminus \mathbb{R}$, we need the re-scaling technique:
Let $\boldsymbol{w} = \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \boldsymbol{u}$, then $\langle \boldsymbol{w}, \boldsymbol{v} \rangle \in \mathbb{R}$:

$$\langle \boldsymbol{w}, \boldsymbol{v} \rangle = \langle \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \boldsymbol{u}, \boldsymbol{v} \rangle = \overline{\left(\frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}}\right)} \langle \boldsymbol{u}, \boldsymbol{v} \rangle = \frac{1}{\langle \boldsymbol{u}, \boldsymbol{v} \rangle} \langle \boldsymbol{u}, \boldsymbol{v} \rangle = 1.$$

Applying the Cauchy-Scharwz inequality for $\langle \boldsymbol{w}, \boldsymbol{v} \rangle \in \mathbb{R}$ gives

$$\begin{aligned} \left| \langle \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \boldsymbol{u}, \boldsymbol{v} \rangle \right| &= |\langle \boldsymbol{w}, \boldsymbol{v} \rangle| \\ &\leq \|\boldsymbol{w}\| \|\boldsymbol{v}\| = \left\| \frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}} \boldsymbol{u} \right\| \|\boldsymbol{v}\| \end{aligned}$$

Or equivalently,

$$\left|\frac{1}{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}\right| |\langle \boldsymbol{u}, \boldsymbol{v}\rangle| \leq \left|\frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}}\right| \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

Since $\left|\frac{1}{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}\right| = \left|\frac{1}{\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}}\right|$, we imply

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

■

2. The triangle inequality also holds for inner product process:

$$\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$$

3. The Gram-Schmidt process holds for finite set of vectors: let $S = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ be (finite) linearly independent. Then we can construct an orthonormal set from S:

$$\boldsymbol{w}_1 = \boldsymbol{v}_1, \quad \boldsymbol{w}_{i+1} = \boldsymbol{v}_{i+1} - \frac{\langle \boldsymbol{v}_{i+1}, \boldsymbol{w}_1\rangle}{\|\boldsymbol{w}_1\|^2} - \frac{\langle \boldsymbol{v}_{i+1}, \boldsymbol{w}_2\rangle}{\|\boldsymbol{w}_2\|^2} - \cdots - \frac{\langle \boldsymbol{v}_{i+1}, \boldsymbol{w}_i\rangle}{\|\boldsymbol{w}_i\|^2}, \ i = 1, \dots, n-1$$

Then after normalization, we obtain the constructed orthonormal set. Consequently, every finite dimensional inner product space has an orthonormal basis.

10.1.2. Dual spaces

Theorem 10.1 — Riesz Representation. Consider the mapping

$$\begin{aligned} \phi: \quad & V \longrightarrow V^* \\ \text{with} \quad & \boldsymbol{v} \mapsto \phi_{\boldsymbol{v}} \\ \text{where} \quad & \phi_{\boldsymbol{v}}(w) = \langle \boldsymbol{v}, w\rangle, \ \forall w \in V \end{aligned}$$

Then the mapping ϕ is well-defined and it is an $\mathbb{R}$-linear transformation. Moreover, if V is finite dimensional, then ϕ is an isomorphism.

The $\mathbb{R}$-linear transformation $V \to V^*$ means that, when V, V^* are vector space over $\mathbb{R}$, the $\mathbb{R}$-linear transformation deduces into exactly the linear transformation.

R The $\mathbb{R}$-linear transformation $V \to V^*$ is **not** necessarily linear if V, V^* are vector spaces over $\mathbb{C}$.

However, we can transform a vector space over $\mathbb{C}$ into a vector space over $\mathbb{R}$:

- For example, suppose that $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ is a basis of V over $\mathbb{C}$, i.e.,

$$\boldsymbol{v} = \sum_{j=1}^{n} \alpha_j \boldsymbol{v}_j$$

where $\alpha_j = p_j + iq_j, \forall p_j, q_j \in \mathbb{R}$, then

$$\boldsymbol{v} = \sum_j p_j \boldsymbol{v}_j + \sum_j q_j (i\boldsymbol{v}_j),\ p_j, q_j \in \mathbb{R}$$

Therefore, $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n, i\boldsymbol{v}_1, \dots, i\boldsymbol{v}_n\}$ forms a basis of V over $\mathbb{R}$.

Note that $i\boldsymbol{v}_1$ cannot be considered as a linear combination of $\boldsymbol{v}_1$ over $\mathbb{R}$, but a linear combination of $\boldsymbol{v}_1$ over $\mathbb{C}$.

In particular, if $\phi : V \to V^*$ is a $\mathbb{R}$-linear transformation, then

$$\phi(i\boldsymbol{v}) \neq i\phi(\boldsymbol{v}), \text{ but } \phi(2\boldsymbol{v}) = 2\phi(\boldsymbol{v}).$$

Proof. 1. Well-definedness: We need to show $\phi_{\boldsymbol{v}} \in V^*$, i.e., for scalars a, b,

$$\phi_{\boldsymbol{v}}(a\boldsymbol{w}_1 + b\boldsymbol{w}_2) = \langle \boldsymbol{v}, a\boldsymbol{w}_1 + b\boldsymbol{w}_2 \rangle = a\langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle + b\langle \boldsymbol{v}, \boldsymbol{w}_2 \rangle = a\phi_{\boldsymbol{v}}(\boldsymbol{w}_1) + b\phi_{\boldsymbol{v}}(\boldsymbol{w}_2)$$

Therefore, $\phi_{\boldsymbol{v}} \in V^*$.

2. $\mathbb{R}$-linearity of ϕ: it suffices to show

$$\phi(c\boldsymbol{v}_1 + d\boldsymbol{v}_2) = c\phi(\boldsymbol{v}_1) + d\phi(\boldsymbol{v}_2), \quad \forall c, d \in \mathbb{R}, \boldsymbol{v}_1, \boldsymbol{v}_2 \in V.$$

For all $\boldsymbol{w} \in V$, we have

$$\phi_{c\boldsymbol{v}_1+d\boldsymbol{v}_2}(\boldsymbol{w}) = \langle c\boldsymbol{v}_1 + d\boldsymbol{v}_2, \boldsymbol{w}\rangle = c\langle \boldsymbol{v}_1, \boldsymbol{w}\rangle + d\langle \boldsymbol{v}_2, \boldsymbol{w}\rangle = c\phi_{\boldsymbol{v}_1}(\boldsymbol{w}) + d\phi_{\boldsymbol{v}_2}(\boldsymbol{w})$$

where the second equality holds because $c, d \in \mathbb{R}$.

Therefore,

$$\phi(c\boldsymbol{v}_1 + d\boldsymbol{v}_2) = c\phi(\boldsymbol{v}_1) + d\phi(\boldsymbol{v}_2).$$

■

10.2. Monday for MAT3006

10.2.1. Remarks on Markov Inequality

Proposition 10.1 — **Markov Inequality.** Suppose that $f : \mathbb{R} \to [0, \infty]$ is measurable, then

$$m(f^{-1}[\lambda, \infty]) \le \frac{1}{\lambda} \int f \, \mathrm{d}m, \ \forall \lambda > 0$$

Proof. Define the function

$$g := \lambda \mathcal{X}_{f^{-1}([\lambda,\infty])},$$

it follows that $g \le f$ globally. Applying proposition (9.9), we imply

$$\int g \, \mathrm{d}m \le \int f \, \mathrm{d}m \implies \lambda m(f^{-1}[\lambda, \infty]) \le \int f \, \mathrm{d}m.$$

■

Corollary 10.1 If $f : \mathbb{R} \to [0, \infty]$ is integrable, and $\int f \, \mathrm{d}m = 0$, then $f = 0$ a.e.

Proof. Consider that for any $\lambda > 0$,

$$0 \le m(f^{-1}[\lambda, \infty]) \le \frac{1}{\lambda} \int f \, \mathrm{d}m = 0.$$

Therefore, $m(\{x \mid f(x) \ne 0\}) = m(f^{-1}(0, \infty]) = 0$. ■

10.2.2. Properties of Lebesgue Integration

In this lecture, we will show several lemmas, which is very useful during the proof of monotone convergence theorem.

Proposition 10.2 If $f : \mathbb{R} \to [0, \infty]$ is such that $f = 0$ a.e., then $\int f \, \mathrm{d}m = 0$.

Proof. Any simple function $\psi \leq f$ must be 0 almost everywhere:

$$\phi = \sum_i \alpha_i \mathcal{X}_{A_i}, \alpha_i > 0, \cup_i A_i \text{ is null.}$$

Direct computation of the Lebesgue integral for this simple function ψ gives

$$\int f \, \mathrm{d}m = \sum_i \alpha_i m(A_i) = 0,$$

where the last equality is because that for each i, the set A_i is null. ■

R Given a non-negative integrable function f on a measurable set E, the integral $\int_E f \, \mathrm{d}m = 0$ if and only if $f = 0$ a.e. on E.

Proposition 10.3 If A, B are measurable, disjoint sets, then

$$\int_{A\cup B} f \, \mathrm{d}m = \int_A f \, \mathrm{d}m + \int_B f \, \mathrm{d}m$$

Proof. The key is to apply $f \cdot \mathcal{X}_{A\cup B} = f \cdot \mathcal{X}_A + f \cdot \mathcal{X}_B$ and

$$\int_E f \, \mathrm{d}m = \int f \cdot \mathcal{X}_E \, \mathrm{d}m, \text{ for any measurable } E.$$

■

Proposition 10.4 If $f : \mathbb{R} \to [0, \infty]$ is measurable, then there exists an increasing sequence of simple functions $\{\phi_n\}$ such that $\phi_n(x) \to f(x)$ pointwise.

Proof. For each $n \in \mathbb{N}$, we divide the interval $[0, 2^n] \subseteq [0, \infty]$ into 2^{2n} subintervals of width 2^{-n}:

$$I_{k,n} = (k2^{-n}, (k+1)2^{-n}], \quad k = 0, 1, \dots, 2^{2n} - 1.$$

Let $J_n = (2^n, \infty]$ be the remaining part of the range of f, and define

$$E_{k,n} = f^{-1}(I_{k,n}), \quad F_n = f^{-1}(J_n).$$

Then the sequence of simple functions are given by:

$$\phi_n = \sum_{k=0}^{2^n-1} k \cdot 2^{-n} \mathcal{X}_{E_{k,n}} + 2^n \mathcal{X}_{F_n}.$$

■

Proposition 10.5 — Fatou's Lemma. Let $\{F_n\}$ be a sequence of non-negative measurable functions, then

$$\liminf_{n\to\infty} \int f_n \, \mathrm{d}m \geq \int \left(\liminf_{n\to\infty} f_n \right) \mathrm{d}m$$

R The inequality in the Fatou's lemma could be strict, e.g., consider $f_n(x) = (n+1)x^n$ on $[0,1]$.

10.3. Monday for MAT4002

Proposition 10.6 — Simplicial Approximation Proposition. Let K and L be two simplifical complexes, and $f : |K| \to |L|$ be a continuous mapping. If there exists a simplicial mapping $g : K \to L$ such that $f(\text{st}_K(\boldsymbol{v})) \subseteq \text{st}_L(g(\boldsymbol{v})), \forall \boldsymbol{v} \in V(K)$, then

$$|g| \simeq f$$

Recall the definition

$$\text{st}_K(\boldsymbol{v}) = \bigcup \{\text{inside}(\sigma) : \sigma \text{ is a simplex of } |K| \text{ and } x \in \sigma\}$$

Proof.

- We first show a statement: Suppose that $\sigma = \{v_0, \dots, v_n\} \in \Sigma(K)$, and $x \in \text{inside}(\sigma) \subseteq |K|$. If $f(x) \in |L|$ lies in the inside of the (unique) simplex $\tau \in \Sigma_L$, (i.e., $f(x)$ can uniquely be expressed as $\sum_{u_i \in \tau} \beta_i u_i$, such that $\beta_i > 0, \forall i$ and $\sum_i \beta_i = 1$) then $g(v_0), \dots, g(v_n)$ are vertices of τ.

  By definition of inside(σ), $x = \sum_{i=0}^n \alpha_i v_i$ with $\alpha_i > 0$ and $\sum_{i=1}^n \alpha_i = 1$. Therefore, $x \in \text{st}_K(v_i)$ for $i = 1, \dots, n$, where

  $$\text{st}_K(v_i) := \left\{ a v_i + \sum_{j=1}^m b_j w_j \mid a > 0, b_j > 0, a + \sum_{j=1}^m b_j = 1, \{v_i, w_1, \dots, w_m\} \in \Sigma_K \right\}.$$

  Therefore, $f(x) \in \text{int}(\text{st}_K(v_i)) \subseteq \text{st}_L(g(v_i))$, which follows that

  $$f(x) = a g(v_i) + \sum_{j=1}^m b_j u_j, \text{ where } a > 0, b_j > 0, a + \sum_{j=1}^m b_j = 1, \ \{g(v_i), u_1, \dots, u_m\} \in \Sigma_L$$

  Comparing the above formula with our hypothesis on $f(x)$, $g(v_i)$ is a vertex of the simplex τ, $i = 1, \dots, n$. Moreover, $\{g(v_0), \dots, g(v_n)\}$ is a subset of τ, which is a face of τ, and therefore $\{g(v_0), \dots, g(v_n)\} \in \Sigma_L$.

- Therefore, the mapping $g : K \to L$ maps simplicies to simplicies, which is a simplicial mapping. We can construct a homotopy between f and $|g|$ as follows: Consider any $x \in |K|$, and let $\tau \in \Sigma_L$ be such that $f(x) \in \text{inside}(\tau)$. We write

$x = \sum_{i=0}^n \lambda_i v_i$ for some $\{v_0, \dots, v_n\} \in \Sigma_K$ and $\lambda_i > 0, \sum_{i=1}^n \lambda_i = 1$. Applying our claim,

$$|g|(x) = \sum_{i=0}^n \lambda_i g(v_i),$$

where $g(v_0), \dots, g(v_n)$ are all vertices of τ.

We can directly construct a homotopy between f and $|g|$. Before that, we need some reformulations. Since $f(x) \in \text{inside}(\tau)$, we let $f(x) = \sum_{i=0}^m \mu_i \tau_i$. Since $|g|(x) = \sum_{i=0}^n \lambda_i g(v_i) \in \text{inside}(\tau)$, we rewrite $|g|(x) = \sum_{i=0}^m \lambda_i' \tau_i$. (by adding some $\lambda_i' := 0$ if necessary) We define the map

$$\begin{aligned} H : \quad & |K| \times I \longrightarrow |L| \\ \text{with} \quad & (x,t) \mapsto \textstyle\sum_{i=0}^m t\lambda_i' + (1-t)\mu_i \end{aligned}$$

which follows that $f \simeq |g|$.

■

Theorem 10.2 — **Simplicial Approximation Theorem.** Let K, L be simplicial complexes with V_K finite, and $f : |K| \to |L|$ be continuous. Then there exists a subdivison $|K'|$ of $|K|$ together with a simplicial map g such that $|g| \simeq f$.

Here the way for constructing subdivison $|K'|$ is as follows. There exists a constant $\delta > 0$. As long as the coarseness of K' is less than δ, our constructed subdivision satisfies the condition.

Proof. The sets $\{\text{st}_L(w) \mid w \in V(L)\}$ forms an open cover of $|L|$, which implies $\{f^{-1}(\text{st}_L(w))\}$ forms an open cover of $|K|$. By compactness, there exists a finite subcover of $|K|$, denoted as

$$|K| \subseteq \bigcup_{i=1}^n f^{-1}(\text{st}_L(w_i))$$

There exists a small number $\delta > 0$ such that for any $x, y \in |K|$ with $d(x,y) < \delta$, $x, y \in f^{-1}(\text{st}_L(w_i))$ for some i. Then we construct a simplicial subdivision $|K'|$ of $|K|$ with coarseness less than δ, i.e., $\forall x, y \in \text{st}_{K'}(v)$, $d(x,y) < \delta$.

Therefore, $\text{st}_{K'}(v) \subseteq f^{-1}(\text{st}_L(w_i))$ for any $v \in V(K;)$ and some $w_i \in V(L)$, i.e., $f(\text{st}_{K'}(v)) \subseteq$

$\mathrm{st}_L(w_i)$.

Setting $g(v) = w_i$ and applying proposition (10.6) gives the desired result. ■

10.3.1. Group Presentations

Group is a highlight of our course, which interwises topology and algebra. I assume that most students have learnt abstract algebra course MAT3004, and encourage those without this knowledge to read the notes for group posted on blackboard.

10.4. Wednesday for MAT3040

Reviewing. Consider the mapping

$$\begin{aligned} \phi: \quad & V \to V^* \\ \text{with} \quad & \phi(\boldsymbol{v}) = \phi_{\boldsymbol{v}} \\ \text{where} \quad & \phi_{\boldsymbol{v}}(\boldsymbol{w}) = \langle \boldsymbol{v}, \boldsymbol{w} \rangle \end{aligned}$$

The Riesz Representation Theorem claims that

1. ϕ is a $\mathbb{R}$-linear transformation.
2. ϕ is injective.
3. If $\dim(V) < \infty$, then ϕ is an isomorphism.

Proof for Claim (2). Consider the equality $\phi(\boldsymbol{v}) = \phi_{\boldsymbol{v}} = 0_{V^*}$, which implies

$$\phi_{\boldsymbol{v}}(\boldsymbol{w}) = \langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0, \forall \boldsymbol{w} \in V$$

By the non-degenercy property, $\boldsymbol{v} = 0_{\boldsymbol{v}}$, i.e., ϕ is injective. ■

Proof for Claim (3). Since $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(V^*)$, and ϕ is injective as a $\mathbb{R}$-linear transformation, we imply ϕ is an isomorphism from V to V^*, where V, V^* are treated as vector spaces over $\mathbb{R}$. ■

10.4.1. Orthogonal Complement

Definition 10.5 [Orthogonal Complement] Let $U \leq V$ be a subspace of an inner product space. Then the **orthogonal complement** of U is

$$U^{\perp} = \{\boldsymbol{v} \in V \mid \langle \boldsymbol{v}, \boldsymbol{u} \rangle = 0, \forall \boldsymbol{u} \in U\}$$

The analysis for orthogonal complement for vector spaces over $\mathbb{C}$ is quite similar as what we have studied in MAT2040.

Proposition 10.7 1. $U^\perp$ is a subspace of V

2. $U \cap U^\perp = \{0\}$
3. $U_1 \subseteq U_2$ implies $U_2^\perp \leq U_1^\perp$.

Proof. 1. Suppose that $\boldsymbol{v}_1, \boldsymbol{v}_2 \in U^\perp$, where $a, b \in K$ ($K = \mathbb{C}$ or $\mathbb{R}$), then for all $\boldsymbol{u} \in U$,

$$\begin{aligned} \langle a\boldsymbol{v}_1 + b\boldsymbol{v}_2, \boldsymbol{u}\rangle &= \bar{a}\langle \boldsymbol{v}_1, \boldsymbol{u}\rangle + \bar{b}\langle \boldsymbol{v}_2, \boldsymbol{u}\rangle \\ &= \bar{a} \cdot 0 + \bar{b} \cdot 0 = 0 \end{aligned}$$

Therefore, $a\boldsymbol{v}_1 + b\boldsymbol{v}_2 \in U^\perp$.

2. Suppose that $\boldsymbol{u} \in U \cap U^\perp$, then we imply $\langle \boldsymbol{u}, \boldsymbol{u}\rangle = 0$. By the positive-definiteness of inner product, $\boldsymbol{u} = \mathbf{0}$.
3. The statement (3) is easy.

■

Proposition 10.8 1. If $\dim(V) < \infty$ and $U \leq V$, then $V = U \oplus U^\perp$

2. If $U, W \leq V$, then

$$\begin{aligned} (U + W)^\perp &= U^\perp \cap W^\perp \\ (U \cap W)^\perp &\supseteq U^\perp + W^\perp \\ (U^\perp)^\perp &\supseteq U \end{aligned}$$

Moreover, if $\dim(V) < \infty$, then these are equalities.

Proof. 1. Suppose that $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ forms a basis for U, and by basis extension, we obtain $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n\}$ is a basis for V.

By Gram-Schmidt Process, any finite basis induces an orthonormal basis.

Therefore, suppose that $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$ forms an orthonormal basis for U, and $\{\boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n\}$ forms an orthonormal basis for $U^\perp$.

It's easy to show $V = U + U^\perp$ using orthonormal basis.

2. (a) The reverse part $(U + W)^\perp \supseteq U^\perp \cap W^\perp$ is trivial; for the forward part, sup-

pose $\boldsymbol{v} \in (U+W)^{\perp}$, then

$$\langle \boldsymbol{v}, \boldsymbol{u}+\boldsymbol{w} \rangle = 0, \ \forall \boldsymbol{u} \in U, \ \boldsymbol{w} \in W$$

Taking $\boldsymbol{u} \equiv \mathbf{0}$ in the equality above gives $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$, i.e., $\boldsymbol{v} \in U^{\perp}$. Similarly, $\boldsymbol{v} \in W^{\perp}$.

(b) Follow the similar argument as in (2a). If $\dim(V) < \infty$, then write down the orthonormal basis for $U^{\perp} + W^{\perp}$ and $(U \cap W)^{\perp}$.

(c) Follow the similar argument as in (2a). If $\dim(V) < \infty$, then

$$V = U^{\perp} \oplus (U^{\perp})^{\perp} = U \oplus U^{\perp}.$$

Therefore, $(U^{\perp})^{\perp} = U$.

∎

Proposition 10.9 The mapping $\phi : V \to V^*$ maps $U^{\perp} \leq V$ injectively to $\text{Ann}(U) \leq V^*$. If $\dim(V) < \infty$, then $U^{\perp} \cong \text{Ann}(U)$ as $\mathbb{R}$-vector spaces

Proof. The injectivity of ϕ has been shown at the beginning of this lecture. For any $\boldsymbol{v} \in U^{\perp}$, we imply $\phi_{\boldsymbol{v}}(\boldsymbol{u}) = 0, \forall \boldsymbol{u} \in U$, i.e., $\phi_{\boldsymbol{v}} \in \text{Ann}(U)$.

Therefore, $\phi(U^{\perp}) \leq \text{Ann}(U)$.

Provided that $\dim(V) < \infty$, by (1) in proposition (10.8),

$$\dim(U) + \dim(U^{\perp}) = \dim(V)$$

Since $\dim(U) + \dim(\text{Ann}(U)) = \dim(V)$, we imply $\dim(U^{\perp}) = \dim(\text{Ann}(U))$.

Moreover,

$$\phi : U^{\perp} \to \text{Ann}(U)$$

is an isomorphism between $\mathbb{R}$-vector spaces $U^{\perp}$ and $\text{Ann}(U)$. ∎

10.4.2. Adjoint Map

Motivation. Then we study the induced mapping based on a given linear operator T, denoted as T'. This induced mapping essentially plays the similar role as taking the Hermitian for a complex matrix.

Notation. Previously we have studied the **adjoint** of $T : V \to W$, denoted as $T^* : W^* \to V^*$. However, from now on, we use the same terminology but with different meaning. If $T : V \to V$ is a linear operator, then the **adjoint** of T is the linear operator $T' : V \to V$ defined as follows.

Definition 10.6 [Adjoint] Let $T : V \to V$ be a linear operator between inner product spaces. The **adjoint** of T is defined as $T' : V \to V$ satisfying

$$\langle T'(\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle, \ \forall \boldsymbol{w} \in V \tag{10.1}$$

Proposition 10.10 If $\dim(V) < \infty$, then T' exists, and it is unique. Moreove, T' is a linear map.

Proof. Fix any $\boldsymbol{v} \in V$. Consider the mapping

$$\alpha_{\boldsymbol{v}} : \boldsymbol{w} \xrightarrow{T} T(\boldsymbol{w}) \xrightarrow{\phi_{\boldsymbol{v}}} \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle$$

This is a linear transformation from V to $\mathbb{F}$, i.e., $\alpha_{\boldsymbol{v}} \in V^*$

By Riesz representation theorem, ϕ is an isomorphism from V to V^*. Therefore, for any $\alpha_{\boldsymbol{v}} \in V^*$, there exists a vector $T'(\boldsymbol{v}) \in V$ such that

$$\phi(T'(\boldsymbol{v})) = \alpha_{\boldsymbol{v}} \in V^*$$

Or equivalently, $\phi_{T'(\boldsymbol{v})}(\boldsymbol{w}) = \alpha_{\boldsymbol{v}}(\boldsymbol{w}), \forall \boldsymbol{w} \in V$, i.e., $\langle T'(\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle$.

Therefore, from $\boldsymbol{v}$ we have constructed $T'(\boldsymbol{v})$ satisfying (10.1). Now define $T' : V \to V$ by $\boldsymbol{v} \mapsto T'(\boldsymbol{v})$.

- Since the choice of $T'(\boldsymbol{v})$ is unique by the injectivity of ϕ, T' is well-defined.
- Now we show T' is a linear transformation: Let $\boldsymbol{v}_1,\boldsymbol{v}_2 \in V, a,b \in K$. For all $\boldsymbol{w} \in V$, we have

$$\begin{aligned}\langle T'(a\boldsymbol{v}_1 + b\boldsymbol{v}_2),\boldsymbol{w}\rangle &= \langle a\boldsymbol{v}_1 + b\boldsymbol{v}_2, T(\boldsymbol{w})\rangle \\ &= \bar{a}\langle \boldsymbol{v}_1, T(\boldsymbol{w})\rangle + \bar{b}\langle \boldsymbol{v}_2, T(\boldsymbol{w})\rangle \\ &= \bar{a}\langle T'(\boldsymbol{v}_1),\boldsymbol{w}\rangle + \bar{b}\langle T'(\boldsymbol{v}_2),\boldsymbol{w}\rangle \\ &= \langle aT'(\boldsymbol{v}_1) + bT'(\boldsymbol{v}_2),\boldsymbol{w}\rangle\end{aligned}$$

  Therfore,

$$\langle T'(a\boldsymbol{v}_1 + b\boldsymbol{v}_2) - [aT'(\boldsymbol{v}_1) + bT'(\boldsymbol{v}_2)],\boldsymbol{w}\rangle = 0, \ \forall \boldsymbol{w} \in V$$

  By the non-degeneracy of inner product,

$$T'(a\boldsymbol{v}_1 + b\boldsymbol{v}_2) - [aT'(\boldsymbol{v}_1) + bT'(\boldsymbol{v}_2)] = \mathbf{0},$$

  i.e., $T'(a\boldsymbol{v}_1 + b\boldsymbol{v}_2) = aT'(\boldsymbol{v}_1) + bT'(\boldsymbol{v}_2)$

■

■ **Example 10.2** Let $V = \mathbb{R}^n$, $\langle\cdot,\cdot\rangle$ as the usual inner product. Consider the matrix-multiplication mapping

$$\begin{aligned}T: \quad & V \to V \\ & T(\boldsymbol{v}) = A\boldsymbol{v}\end{aligned}$$

Then $\langle T'(\boldsymbol{v}),\boldsymbol{w}\rangle = \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle$ implies

$$\begin{aligned}(T'(\boldsymbol{v}))^{\mathrm{T}}\boldsymbol{w} &= \langle \boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}\rangle \\ &= \boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{w} \\ &= (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{w}\end{aligned}$$

Therfore, $T'(\boldsymbol{v}) = A^{\mathrm{T}}\boldsymbol{v}$. ■

Proposition 10.11 Let $T : V \to V$ be a linear transformation, V a inner product space. Suppose that $\mathcal{B} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is an orthonormal basis of V, then

$$(T')_{\mathcal{B},\mathcal{B}} = \overline{((T)_{\mathcal{B},\mathcal{B}})^{\mathrm{T}}}$$

Proof. Suppose that $(T)_{\mathcal{B},\mathcal{B}} = (a_{ij})$, where $T(\boldsymbol{e}_j) = \sum_{k=1}^n a_{kj}\boldsymbol{e}_k$, then

$$\begin{aligned}
\langle \boldsymbol{e}_i, T(\boldsymbol{e}_j)\rangle &= \langle \boldsymbol{e}_i, \sum_{k=1}^n a_{kj}\boldsymbol{e}_k\rangle \\
&= \sum_{k=1}^n a_{kj}\langle \boldsymbol{e}_i, \boldsymbol{e}_k\rangle \\
&= a_{ij}
\end{aligned}$$

Also, suppose $(T')_{\mathcal{B},\mathcal{B}} = (b_{ij})$, we imply $T'(\boldsymbol{e}_j) = \sum_{k=1}^n b_{ij}\boldsymbol{e}_k$, which follows that

$$\langle \boldsymbol{e}_i, T'(\boldsymbol{e}_j)\rangle = b_{ij} \implies \overline{\langle T'(\boldsymbol{e}_j), \boldsymbol{e}_i\rangle} = b_{ij} \implies \overline{\langle \boldsymbol{e}_j, T(\boldsymbol{e}_i)\rangle} = b_{ij},$$

i.e., $\overline{a_{ji}} = b_{ij}$. ∎

R Proposition (10.11) does not hold if $\mathcal{B}$ is not an orthonormal basis.

10.5. Wednesday for MAT3006

Proposition 10.12 — **Fatou's Lemma.** Suppose $\{f_n\}$ is a sequence of measurable, nonnegative functions.

$$\liminf_{n\to\infty} \int f_n \,\mathrm{d}m \ge \int \liminf_{n\to\infty}(f_n)\,\mathrm{d}m$$

Proof. Define $g_n(x) := \inf_{k\ge n} f_k(x)$ and

$$f(x) = \liminf_{n\to\infty} f_n(x) = \lim_{n\to\infty}\left(\inf_{k\ge n} f_k(x)\right) := \lim_{n\to\infty} g_n(x)$$

To study the integral $\int f\,\mathrm{d}m$, we will only focus on $f(x)$ on $E \subseteq \mathbb{R}$, where $f(x) > 0, \forall x \in E$.

It suffices to show that $\int_E \phi\,\mathrm{d}m \le \lim_{n\to\infty}\inf \int_E f_n\,\mathrm{d}m$ for all simple ϕ satisfying $0 \le \phi(x) \le f(x), \forall x \in E$. (Then taking supremum both sides leads to the desired result.)

1. Construct the simple function ϕ' on E such that

   $$\phi'(x) = \begin{cases} \phi(x) - \varepsilon, & \text{if } \phi(x) > 0 \\ 0, & \text{if } \phi(x) = 0 \end{cases}$$

   in which we pick ε small enough such that $\phi(x) - \varepsilon \ge 0$.

   As a result, $\phi' < f, \forall x \in E$ (why?).

2. Note that $g_n(x)$ is monotone increasing with n, and therefore convergent to $f(x)$. Consider $A_n := \{x \in E \mid \phi'(x) \le g_n(x)\}$, which follows that

   (a) $A_n \subseteq A_{n+1}$

   (b) $\cup_{n=1}^{\infty} A_n = E$ (We do need ϕ' is strictly less than f to obtain this condition).

   Therefore, for any $k \ge n$,

   $$\int_{A_n} \phi'\,\mathrm{d}m \le \int_{A_n} g_n\,\mathrm{d}m \le \int_{A_n} f_k\,\mathrm{d}m,$$

which implies $\int_{A_n} \phi' \, \mathrm{d}m \leq \int_E f_k \, \mathrm{d}m$ since $f_k \mathcal{X}_{A_n} \leq f_k \mathcal{X}_E$. Or equivalently,

$$\int_{A_n} \phi' \, \mathrm{d}m \leq \inf_{k \geq n} \int_E f_k \, \mathrm{d}m \tag{10.2}$$

3. Taking limits $n \to \infty$ both sides for (10.2):
   - For LHS, suppose that $\phi' = \sum_i \alpha_i \mathcal{X}_{c_i}$, then $\int_{A_n} \phi' \, \mathrm{d}m = \sum_i \alpha_i m(c_i \cap A_n)$, which follows that

   $$\lim_{n\to\infty} \int_{A_n} \phi' \, \mathrm{d}m = \sum_i \alpha_i \lim_{n\to\infty} m(c_i \cap A_n) = \sum_i \alpha_i m(c_i) = \int_E \phi' \, \mathrm{d}m$$

   - The limit of RHS equals $\lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m$, and therefore

   $$\int_E \phi' \, \mathrm{d}m \leq \lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m$$

   Note that the goal is to show $\int_E \phi \, \mathrm{d}m \leq \lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m$, and therefore we need to evaluate ϕ' in terms of ϕ.

4. (a) Consider the case where $m(\phi^{-1}(0,\infty)) = P < \infty$, then

   $$\int_E \phi' \, \mathrm{d}m = \int_E \phi \, \mathrm{d}m - \varepsilon \cdot P \leq \lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m,$$

   for all small $\varepsilon > 0$. Then the desired result holds.

   (b) Consider the case where $m(\phi^{-1}(0,\infty)) = \infty$, and we write the canonical form $\phi = \sum \alpha_i \mathcal{X}_{c_i}$ with $\alpha_i > 0$. Define $C = \cup_i c_i$ such that $m(c) = \infty$.

   Construct the simple function $\phi' = a\mathcal{X}_C$, where $a := \frac{1}{2} \min\{\alpha_i\}$, which implies
   - $\phi' \leq \phi$
   - $\int_E \phi' \, \mathrm{d}m = am(c) = \infty$, which follows that $\int_E \phi \, \mathrm{d}m = \infty$.

   Our goal is to show $\lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m = \infty$.

   Consider $B_n = \{x \in E \mid g_n(x) > a\}$, then $\cup B_n = E, B_n \subseteq B_{n+1}$.

   Observe the inequality

   $$\int_{C \cap B_n} a \, \mathrm{d}m \leq \int_{B_n} a \, \mathrm{d}m \leq \int_{B_n} g_n \, \mathrm{d}m \leq \inf_{k \geq n} \int_E f_n \, \mathrm{d}m$$

Taking $n \to \infty$ both sides. For LHS, by definition of B_n, the limit equals $\int_C a \, \mathrm{d}m = \int \phi' \, \mathrm{d}m = \infty$; and the limit of RHS equals to $\lim_{n\to\infty} \inf \int_E f_n \, \mathrm{d}m$, i.e.,

$$\liminf_{n\to\infty} \int_E f_n \, \mathrm{d}m = \infty$$

■

Theorem 10.3 — Monotone Convergence Theorem I. Let $\{f_n\}$ be a sequence of non-negative measurable functions, with

- $f_n(x)$ being monotone increasing
- $f_n(x) \to f(x)$ pointwisely

Then we have

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}m = \int \left(\lim_{n\to\infty} f_n\right) \mathrm{d}m := \int f \, \mathrm{d}m$$

Proof.

- On the one hand, for all $n \in \mathbb{N}$, we have

$$f_n \le f \implies \int f_n \, \mathrm{d}m \le \int f \, \mathrm{d}m \implies \limsup_{n\to\infty} \int f_n \, \mathrm{d}m \le \int f \, \mathrm{d}m$$

- On the other hand, applying the Fatou's lemma,

$$\int f \, \mathrm{d}m := \int \left(\liminf_{n\to\infty} f_n\right) \mathrm{d}m \le \liminf_{n\to\infty} \int f_n \, \mathrm{d}m$$

  Togehter with the previous inequality, we imply

$$\limsup_{n\to\infty} \int f_n \, \mathrm{d}m \le \int f \, \mathrm{d}m \le \liminf_{n\to\infty} \int f_n$$

Therefore, all inequalities above are equalities, and the limit exists since limsup and liminf coincides. Moreover,

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}m = \int f \, \mathrm{d}m.$$

■

From MCT I, the Lebesgue integral $\int f \, \mathrm{d}m$ can be computed as follows:

- Construct simple functions $\phi_n \leq \phi_{n+1}$ with $\phi_n \to f$
- Evaluate $\int \phi_n \, \mathrm{d}m$ and then $\int f \, \mathrm{d}m = \lim_{n\to\infty} \int \phi_n \, \mathrm{d}m$

10.5.1. Consequences of MCT

Proposition 10.13 The Lebesgue integral is finitely addictive for measurable non-negative functions. In other words, suppose f, g are measurable and nonnegative, then

$$\int f \, \mathrm{d}m + \int g \, \mathrm{d}m = \int (f+g) \, \mathrm{d}m$$

Proof. Suppose we have simple increasing functions $\{\phi_n\}$ and $\{\psi_n\}$ such that $\phi_n \to f$ and $\psi_n \to f$. Then

$$\int (f+g) \, \mathrm{d}m = \lim_{n\to\infty} \int (\phi_n + \psi_n) \, \mathrm{d}m \tag{10.3a}$$

$$= \lim_{n\to\infty} \int \phi_n \, \mathrm{d}m + \lim_{n\to\infty} \int \psi_n \, \mathrm{d}m \tag{10.3b}$$

$$= \int f \, \mathrm{d}m + \int g \, \mathrm{d}m \tag{10.3c}$$

where (10.3a) and (10.3c) is by applying MCT I; and (10.3b) is by definition of simple function. ■

Corollary 10.2 The Lebesgue integral is linear defined for measurable, nonnegative functions. In other words, suppose f, g are measurable and nonnegative, then

$$\int (af + bg) \, \mathrm{d}m = a \int f \, \mathrm{d}m + b \int g \, \mathrm{d}m,$$

for any $a, b \geq 0$.

Proposition 10.14 The Lebesgue integral for non-negative continuous function on a bounded closed interval coincides with the Riemann integral. In other words, let f be a non-negative continuous function on $[a, b]$. then

$$\int_{[a,b]} f \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x.$$

We will extend this result into all proper Riemann integrable functions on $[a, b]$ soon.

Proof. Let ϕ_n be the simple function giving the Riemann lower sum of $f(x)$ with 2^n equal subintervals:

$$\phi_n(x) = \sum_{k=1}^{2^n} \left(\min_{y \in \bar{I}_k} f(y) \right) \mathcal{X}_{I_k}, \text{ where } I_k = [a + (b-a)\frac{k-1}{2^n}, a + (b-a)\frac{k}{2^n}]$$

- $\phi_n(x) \geq 0$ is monotone increasing (that's the reason we should divde intervals into 2^n pieces instead of n pieces)

- $\phi_n(x) \to f(x)$ pointwisely: for any $x \in [a, b]$ and $\varepsilon > 0$, by (uniform) continuity of f, there exists $\delta > 0$ such that

$$|y - x| < \delta \implies |f(y) - f(x)| < \varepsilon.$$

  Therefore, for sufficiently large n, we imply for any $x \in I_{k,n}$, $|I_{k,n}| < \delta$. As a result,

$$\left| \min_{y \in I_{k,n}} f(y) - f(x) \right| < \varepsilon.$$

Therefore,

$$\begin{aligned}
\int_{[a,b]} f \, \mathrm{d}m &= \lim_{n\to\infty} \int \phi_n \, \mathrm{d}m \\
&= \lim_{n\to\infty} \left[\text{Riemann lower integral of } \int_a^b f(x) \, \mathrm{d}x \right] \\
&= \int_a^b f(x) \, \mathrm{d}x
\end{aligned}$$

■

■ **Example 10.3** The Lebesgue integral gives us an alternative way to compute improper integrals. Suppose that we want to compute the integral

$$\int_0^1 (1-x)^{-1/2}\,\mathrm{d}x.$$

1. The old method is that we know the integral

$$\int_0^{1-1/n} (1-x)^{-1/2}\,\mathrm{d}x \text{ exists for any } n.$$

   Then we extend the definition of Riemann integration by taking limit of n:

$$\int_0^1 f(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_0^{1-1/n} (1-x)^{-1/2} = \lim_{n\to\infty} 2-2\sqrt{\frac{1}{n}} = 2$$

2. The Lebesgue integration does not require us to extend the definition. Consider

$$f_n(x) = (1-x)^{-1/2}\mathcal{X}_{[0,1-1/n]}$$

   Then

   - $f_n(x) \to f(x)$ on $[0,1)$
   - $f_n(x)$ is monotone increasing

   Therefore, by applying MCT I,

$$\int_{[0,1)} (1-x)^{-1/2}\,\mathrm{d}x = \lim_{n\to\infty}\int f_n\,\mathrm{d}m = \lim_{n\to\infty}\int_{[0,1-1/n]} (1-x)^{-1/2}\,\mathrm{d}x.$$

   By Proposition (10.14), $\int_{[0,1-1/n]}(1-x)^{-1/2}\,\mathrm{d}x = \int_0^{1-1/n}(1-x)^{-1/2}\,\mathrm{d}x$ for all n. Therefore, we conclude that the Lebesgue integral is equal to the (improper) Riemann integral in this case.

■

10.6. Wednesday for MAT4002

10.6.1. Reviewing On Groups

▪ **Example 10.4** Let D_{2n} be the regular polygon P with $2n$ sides in $\mathbb{R}^2$, centered at the origin. It's clear that D_{2n} is **invariant** with $2n$ rotations, or with 2 reflections. Let a denote the rotation of D_{2n} clockwise by degree π/n, and b denote the reflection over lines through the origin.

As a result, $\{e, a, a^2, \dots, a^{n-1}\}$ forms a group; and $\{e, b\}$ forms a group.

Therefore, all elements of D_g can be obtained by $a^i b^j, 0 \le i \le 3, 0 \le j \le 1$.

Any finite operations of rotation (the rotation degree is a multiple of π/n) and reflection can be represented as $a^i b^j$.

Geometrically, we can check that $ba = a^{n-1}b$. ▪

Definition 10.7 [Product Group] Let G, H be two groups. The **product group** $(G \times H, *)$ is defined as

$$\begin{aligned} G \times H &= \{(g,h) \mid g \in G, h \in H\} \\ \text{with} \quad (g_1, h_1) * (g_2, h_2) &= (g_1 g_2, h_1 h_2) \end{aligned}$$

▪

For example, $(\mathbb{R} \times \mathbb{R}, +) = \{(x,y) \mid x, y \in \mathbb{R}\}$ coincides with the usual $\mathbb{R}^2$, where

$$(x,y) * (x', y') = (x + x', y + y')$$

Definition 10.8 A map between two groups $\phi : G \to H$ is a **homomorphism** if

$$\phi(g_1 * g_2) = \phi(g_1) * \phi(g_2)$$

In other words, a homomorphism is a map preserving multiplications of groups. ▪

(R) Follow the similar idea as in MAT3040 knowledge, if $\phi : G \to H$ is a homomorphism, then $\phi(e_G) = e_H$.

■ **Example 10.5** Let $G = (\mathbb{R}, +, 0)$, and $H = \{H_2, *, I_2\}$, with H_2 of the form

$$H_2 = \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \middle| x \in \mathbb{R} \right\}$$

Define a mapping

$$\begin{aligned} \phi: \quad & G \to H \\ \text{with} \quad & x \mapsto \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Then ϕ is a homorphism:

$$\begin{aligned} \phi(x *_{\mathbb{R}} y) &= \phi(x+y) \\ &= \begin{pmatrix} 1 & x+y \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} \\ &= \phi(x) *_{H_2} \phi(y) \end{aligned}$$

■

Definition 10.9 [Isomorphism] A homomorphism $\phi : G \to H$ is an isomorphism if ϕ is bijective. The isomorphism between G and H is denoted as $G \cong H$. ■

Actually, a group can be represented as a Cayley Table:

$$G = \begin{array}{c|ccccc} \circ & g_1 & g_2 & \cdots & g_n \\ \hline g_1 & g_1 \circ g_1 & g_1 \circ g_2 & \cdots & g_1 \circ g_n \\ g_2 & g_2 \circ g_1 & g_2 \circ g_2 & \cdots & g_2 \circ g_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_n & g_n \circ g_1 & g_n \circ g_2 & \cdots & g_n \circ g_n \end{array}, \qquad H = \begin{array}{c|ccccc} \circ & h_1 & h_2 & \cdots & h_n \\ \hline h_1 & h_1 \circ h_1 & h_1 \circ h_2 & \cdots & h_1 \circ h_n \\ h_2 & h_2 \circ h_1 & h_2 \circ h_2 & \cdots & h_2 \circ h_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_n & h_n \circ h_1 & h_n \circ h_2 & \cdots & h_n \circ h_n \end{array}$$

The groups $G \cong H$ if and only if we can find a bijective $\phi : G \to H$ such that, the Cayley

Table of $(H, \circ)$ can be generated from the Cayley Table of $(G, \circ)$ by replacing each entry of G with its image under ϕ.

10.6.2. Free Groups

Definition 10.10

- Let S be a (finite) set, which is considered as an "alphabet".
- Define another set $S^{-1} := \{x^{-1} \in x \in S\}$. We insist that $S \cap S^{-1} = \emptyset$.
- A **word** in S is a finite sequence $w = w_1 \cdots w_m$, where $m \in \mathbb{N}^+ \cup \{0\}$, and each $w_i = \in \cup S^{-1}$. In particular, when $m = 0$, we view w as the empty sequence, denoted as $\emptyset$.
- The **concatenation** of two words $x_1 \cdots x_m$ and $y_1 \cdots y_n$ is the word $x_1 \cdots x_m y_1 \cdots y_n$
- Two words w, w' are **equivalent**, denoted as $w \sim w'$, if there are words $w_1, \ldots, w_n$ and $w = w_1, w' = w_n$ such that

$$w_i = \cdots y_1 x x^{-1} y_2 \cdots, \qquad w_{i+1} = \cdots y_1 y_2 \cdots$$

or

$$w_i = \cdots y_1 y_2 \cdots, \qquad w_{i+1} = \cdots y_1 x x^{-1} y_2 \cdots$$

for some $x \in S \cup S^{-1}$.

■

■ **Example 10.6** For example, $S = \{a, b\}$ and $S^{-1} = \{a^{-1} b^{-1}\}$ and

$$w = aabab^{-1}b^{-1}a^{-1}abaabb^{-1}a$$
$$w' = aabab^{-1}b^{-1}a^{-1}abaaa$$

Here w and w' differs by bb^{-1}. Therefore, $w \sim w'$, and w is said to be a elementary expansion of w'. ■

R We insist that $(s^{-1})^{-1} = s, \forall s^{-1} \in S^{-1}$, since otherwise for $x = s^{-1} \in S^{-1}$, we cannot define $(s^{-1})^{-1}$.

Moreover, for

$$w = aabab^{-1}b^{-1}a^{-1}abaabb^{-1}a$$
$$w'' = aabab^{-1}b^{-1}baabb^{-1}a,$$

w and w'' differs by $a^{-1}a$, i.e., $a^{-1}(a^{-1})^{-1}$, and therefore $w \sim w''$.

Definition 10.11 [Free Group] The **free group** $F(S)$ is defined to be the equivalence class of words, i.e.,

$$[w] := \{w' \text{ is a word in } S \mid w \sim w'\} \in F(S)$$

■

R $F(S)$ is indeed a group:

- $[w] * [w'] = [ww']$ (concatenation) check $w_1 \sim w_2, u_1 \sim u_2$ implies $w_1u_1 \sim w_2u_2$
- Identity element: $e = [\emptyset]$
- Inverse element: $[x_1 \cdots x_n]^{-1} = [x_n^{-1} \cdots x_1^{-1}]$

■ **Example 10.7** Let $S = \{a\}$ and $S^{-1} = \{a^{-1}\}$. Any word w has the form

$$w = a \cdots aa^{-1} \cdots a^{-1}a \cdots aa^{-1} \cdots a^{-1} \cdots$$

In shorthand, we denote w as $w = \cdots a^p(a^{-1})^q a^r (a^{-1})^s \cdots$, and

$$\begin{aligned}[w] = [\cdots a^p(a^{-1})^q a^r(a^{-1})^s \cdots] &= [\cdots a^{p-1}(a^{-1})^{q-1}a^r(a^{-1})^s \cdots] \\ &= [\cdots a^{p-1}(a^{-1})^{q-2}a^{r-1}(a^{-1})^s \cdots],\end{aligned}$$

e.g., we can always eliminate the adjacent terms a and a^{-1} up to equivalence class. Therefore, $F(S) = \{\cdots, [a^{-2}], [a^{-1}], [\emptyset], [a], [a^2], \cdots\}$.

It's clear that $F(S) \cong \mathbb{Z}$, where the isomorphism $\phi : \mathbb{Z} \to F(S)$ is $\phi(n) = [a^n]$. ■

■ **Example 10.8** Let $S = \{a, b\}$ and $S^{-1} = \{a^{-1}, b^{-1}\}$. In this case, $[ab] \neq [ba]$, and $[ab^{-1}a^2b^2a^{-2}b]$ cannot be reduced further.

Since S is not an abelian group in such case, we imply $F(S) \not\cong \mathbb{Z} \times \mathbb{Z}$. ■

10.6.3. Relations on Free Groups

Definition 10.12 [Group With Relations] Let S be a set. A **group with relations** is written as

$$G = \langle S \mid R(S) \rangle$$

where

- $R(S)$ consists of elements in $F(S)$
- Every element in G can be written as the form $[w] \in F(S)$, and we insist that $[w] = [w']$ in G if
  - w and w' differ by some $xx^{-1}, x \in S \cup S^{-1}$, or
  - w and w' differ by some element $z \in R(S)$, or its inverse.

■

■ **Example 10.9** Let $G = \langle a, b \mid a^2, b^2, abab^{-1}a^{-1}b^{-1} \rangle$, we want to enumerate all possible elements in G. Obseve that

$$[b^{-1}] = [b^{-1}b^2] = [b], \quad \text{similarly } [a^{-1}] = [a]$$
$$[bab] = [abab^{-1}a^{-1}b^{-1}bab] = [abab^{-1}b] = [aba]$$

As a result,

- $[a^{-n}] = [a^n]$ and $[b^{-n}] = [b^n]$
- $[a^{2n+1}] = [a], [b^{2n+1}] = [b], [a^{2n}] = [\emptyset], [b^{2n}] = [\emptyset]$
- For another type of element of G, it must be of the form $[\cdot abababab\cdots]$.

  Each aba can be changed into bab, and finally it will be reduced into the form $[ab]$.

Therefore, the elements in G are

$$[\emptyset], [a], [b], [ab], [ba], [aba]$$

In fact, $G \cong S_3$. ■

Chapter 11

Week11

11.1. Monday for MAT3040

Reviewing. Adjoint Operator: $\langle T'(\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle$.

11.1.1. Self-Adjoint Operator

Definition 11.1 [Self-Adjoint] Let V be an inner product space and $T : V \to V$ be a linear operator. Then T is **self-adjoint** if $T' = T$.

■ **Example 11.1** Let $V = \mathbb{C}^n$, and $\mathcal{B} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ be a orthonormal basis. Let $T : V \to V$ be given by

$$T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}, \quad \text{where } A \in M_{n\times n}(\mathbb{C}).$$

Or equivalently, there exists basis $\mathcal{B}$ such that $(T)_{\mathcal{B},\mathcal{B}} = \boldsymbol{A}$.

In such case, T is self-adjoint if and only if $(T')_{\mathcal{B},\mathcal{B}} = (T)_{\mathcal{B},\mathcal{B}}$, i.e., $\overline{(T)^{\mathrm{T}}_{\mathcal{B},\mathcal{B}}} = (T)_{\mathcal{B},\mathcal{B}}$, i.e., $\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}$.

Therefore, $T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$ is self-adjoint if and only if $\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}$.

Moreover, if $\mathbb{C}$ is replaced by $\mathbb{R}$, then T is seld-adjoint if and only if $\boldsymbol{A}$ is symmetric. ■

R The notion of self-adjoint for linear operator is essentially the generalized notion of Hermitian for matrix that we have stuided in MAT2040.

We also have some nice properties for self-adjoint, and the proof for which are essentially the same for the proof in the case of Hermitian matrices.

Proposition 11.1 If λ is an eigenvalue of a self-adjoint operator T, then $\lambda \in \mathbb{R}$.

Proof. Suppose there is an eigen-pair $(\lambda, \boldsymbol{w})$ for $\boldsymbol{w} \neq \mathbf{0}$, then

$$\begin{aligned}
\lambda\langle \boldsymbol{w}, \boldsymbol{w}\rangle &= \langle \boldsymbol{w}, \lambda\boldsymbol{w}\rangle \\
&= \langle \boldsymbol{w}, T(\boldsymbol{w})\rangle = \langle T'(\boldsymbol{w}), \boldsymbol{w}\rangle \\
&= \langle T(\boldsymbol{w}), \boldsymbol{w}\rangle = \langle \lambda\boldsymbol{w}, \boldsymbol{w}\rangle \\
&= \bar{\lambda}\langle \boldsymbol{w}, \boldsymbol{w}\rangle
\end{aligned}$$

Since $\langle \boldsymbol{w}, \boldsymbol{w}\rangle \neq 0$ by non-degeneracy property, we have $\lambda = \bar{\lambda}$, i.e., $\lambda \in \mathbb{R}$. ■

Proposition 11.2 If $U \le V$ is T-invariant over the self-adjoint operator T, then so is $U^{\perp}$.

Proof. It suffices to show $T(\boldsymbol{v}) \in U^{\perp}, \forall \boldsymbol{v} \in U^{\perp}$, i.e., for any $\boldsymbol{u} \in U$, check that

$$\langle \boldsymbol{u}, T(\boldsymbol{v})\rangle = \langle T'(\boldsymbol{u}), \boldsymbol{v}\rangle = \langle T(\boldsymbol{u}), \boldsymbol{v}\rangle = 0,$$

where the last equality is because that $T(\boldsymbol{u}) \in U$ and $\boldsymbol{v} \in U^{\perp}$. Therefore, $T(\boldsymbol{v}) \in U^{\perp}$. ■

Theorem 11.1 If $T : V \to V$ is self-adjoint, and $\dim(V) < \infty$, then there exists an orthonormal basis of eigenvectors of T, i.e., an orthonormal basis of V such that any element from this basis is an eigenvector of T.

Proof. We use the induction on $\dim(V)$:

- The result is trival for $\dim(V) = 1$.
- Suppose that this theorem holds for all vector spaces V with $\dim(V) \le k$, then we want to show the theorem holds when $\dim(V) = k + 1$:

  Suppose that $T : V \to V$ is self-adjoint with $\dim(V) = k + 1$, then consider

  $$\mathcal{X}_T(x) = x^{k+1} + \cdots + a_1 x + a_0, \quad a_i \in \mathbb{K}, \text{ where } \mathbb{K} \text{ denotes } \mathbb{R} \text{ or } \mathbb{C}.$$

- If $\mathbb{K} = \mathbb{C}$, then $\mathcal{X}_T(x)$ can be decomposed as

$$\mathcal{X}_T(x) = (x - \lambda_1) \cdots (x - \lambda_{k+1})$$

  In paricular, we obtain the eigen-pair $(\lambda_1, \boldsymbol{v})$

- If $\mathbb{K} = \mathbb{R}$, i.e., we treat real number as scalars, then

$$\mathcal{X}_T(x) = (x - \lambda_1) \cdots (x - \lambda_{k+1}), \text{ where } \lambda_i \in \mathbb{C}.$$

  By proposition (11.1), we imply all λ_i's are in $\mathbb{R}$. Moreover, we also obtain the eigen-pair $(\lambda_1, \boldsymbol{v})$

Consider $U = \text{span}\{\boldsymbol{v}\}$, then

- U is T-invariant
- $V = U \oplus U^\perp$, since V is finite dimensional
- $U^\perp$ is T-invariant.

Consider $T\mid_{U^\perp}$, which is a self-adjoint operator on $U^\perp$, with $\dim(U^\perp) = k$.

By induction, there exists an orthonormal basis $\{\boldsymbol{e}_2, \dots, \boldsymbol{e}_{k+1}\}$ of eigenvectors of $T\mid_{U^\perp}$.

Consider the basis $\mathcal{B} = \{\boldsymbol{v}' = \boldsymbol{v}/\|\boldsymbol{v}\|, \boldsymbol{e}_2, \dots, \boldsymbol{e}_{k+1}\}$. As a result,

1. $\mathcal{B}$ forms a basis of V
2. All $\boldsymbol{v}', \boldsymbol{e}_i$ are of norm 1 eigenvectors of T.
3. $\mathcal{B}$ is an orthonormal set, e.g., $\langle \boldsymbol{v}', \boldsymbol{e}_i \rangle = 0$, where $\boldsymbol{v}' \in U$ and $\boldsymbol{e}_i \in U^\perp$.

Therefore, $\mathcal{B}$ is a basis of orthonormal eigenvectors of V.

■

Corollary 11.1 If $\dim(V) < \infty$, and $T : V \to V$ is self-adjoint, then there exists orthonormal basis $\mathcal{B}$ such that

$$(T)_{\mathcal{B},\mathcal{B}} = \text{diag}(\lambda_1, \dots, \lambda_n)$$

In paticular, for all real symmtric matrix $\boldsymbol{A} \in \mathbb{S}^n$, there exists orthogonal matrix P ($P^{\mathrm{T}}P = \boldsymbol{I}_n$) such that

$$P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

Proof. 1. By applying theorem (11.1), there exists orthonormal basis of V, say $\mathcal{B} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ such that $T(\boldsymbol{v}_i) = \lambda_i \boldsymbol{v}_i$. Directly writing the basis representation gives

$$(T)_{\mathcal{B},\mathcal{B}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

2. For the second part, consider $T : \mathbb{R}^n \to \mathbb{R}^n$ by $T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$. Since $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$, we imply T is self-adjoint. There exists orthonormal basis $\mathcal{B} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ such that

$$(T)_{\mathcal{B},\mathcal{B}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

In particular, if $\mathcal{A} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$, then $(T)_{\mathcal{A},\mathcal{A}} = \boldsymbol{A}$. We construct $P := C_{\mathcal{A},\mathcal{B}}$, which is the change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$, then

$$P = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$$

and

$$P^{-1}(T)_{\mathcal{A},\mathcal{A}}P = (T)_{\mathcal{B},\mathcal{B}}$$

Or equivalently, $P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, with

$$P^{\mathrm{T}}P = \begin{pmatrix} \boldsymbol{v}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{v}_n^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix} = \boldsymbol{I}$$

■

11.1.2. Orthononal/Unitary Operators

Definition 11.2 A linear operator $T : V \to V$ over $\mathbb{K}$ with $\langle T(\boldsymbol{w}), T(\boldsymbol{v})\rangle = \langle \boldsymbol{w}, \boldsymbol{v}\rangle, \forall \boldsymbol{v}, \boldsymbol{w} \in V$, is called

1. **Orthogonal** if $\mathbb{K} = \mathbb{R}$
2. **Unitary** if $\mathbb{K} = \mathbb{C}$

Proposition 11.3 T is orthogonal / unitary if and only if $T' \circ T = I$

Proof. The reverse direction is by directly checking that

$$\langle T(\boldsymbol{w}), T(\boldsymbol{v})\rangle = \langle T' \circ T(\boldsymbol{w}), \boldsymbol{v}\rangle = \langle \boldsymbol{w}, \boldsymbol{v}\rangle$$

The forward direction is by checking $T' \circ T(\boldsymbol{w}) = \boldsymbol{w}, \forall \boldsymbol{w} \in V$:

$$\langle T' \circ T(\boldsymbol{w}), \boldsymbol{v}\rangle = \langle T(\boldsymbol{w}), T(\boldsymbol{v})\rangle = \langle \boldsymbol{w}, \boldsymbol{v}\rangle \implies \langle T' \circ T(\boldsymbol{w}) - \boldsymbol{w}, \boldsymbol{v}\rangle = 0, \forall \boldsymbol{v} \in V$$

By non-degeneracy, $T' \circ T(\boldsymbol{w}) - \boldsymbol{w} = 0$, i.e., $T' \circ T(\boldsymbol{w}) = \boldsymbol{w}, \forall \boldsymbol{w} \in V$. ∎

Example 11.2 Let $T : \mathbb{K}^n \to \mathbb{K}^n$ be given by $T(\boldsymbol{v}) = A\boldsymbol{v}$. Then T is orthogonal implies $(T')_{\mathcal{B},\mathcal{B}}(T)_{\mathcal{B},\mathcal{B}} = I$.

(Orthogonal) When $\mathbb{K} = \mathbb{R}$, then $A^{\mathrm{T}}A = I$

(Unitary) When $\mathbb{K} = \mathbb{C}$, then $A^{\mathrm{H}}A = I$.

Definition 11.3 [Orthogonal/Unitary Group]

$$\text{Orthognoal Group} : O(n, \mathbb{R}) = \{A \in M_{n\times n}(\mathbb{R}) \mid A^{\mathrm{T}}A = I\}$$

$$\text{Unitary Group} : U(n, \mathbb{C}) = \{A \in M_{n\times n}(\mathbb{C}) \mid A^{\mathrm{H}}A = I\}$$

11.2. Monday for MAT3006

Reviewing. Compute the integration

$$\int_{[0,1)} (1-x)^{-1/2}\,\mathrm{d}x$$

Solution. 1. Construct $g_n(x) = (1-x)^{-1/2}\mathcal{X}_{[0,1-1/n]}$, then g_n is monotone increasing and $g_n(x) \to (1-x)^{-1/2}\mathcal{X}_{[0,1)}$ pointwisely.

2. By applying MCT I and proposition (10.14),

$$\int_{[0,1)} (1-x)^{-1/2}\,\mathrm{d}x = \lim_{n\to\infty}\int g_n\,\mathrm{d}m = 2.$$

∎

Question: How to understand $\int_{[0,1]}(1-x)^{-1/2}\,\mathrm{d}x$?

Answer:

$$(1-x)^{-1/2}\mathcal{X}_{[0,1]} = (1-x)^{-1/2}\mathcal{X}_{[0,1)} + \infty\cdot\mathcal{X}_{\{1\}}$$

which follows that

$$\int (1-x)^{-1/2}\mathcal{X}_{[0,1]}\,\mathrm{d}m = \int (1-x)^{-1/2}\mathcal{X}_{[0,1)}\,\mathrm{d}m + \int \infty\cdot\mathcal{X}_{\{1\}}\,\mathrm{d}m \tag{11.1a}$$

$$= \int_{[0,1)} (1-x)^{-1/2}\,\mathrm{d}x + 0 \tag{11.1b}$$

where (11.1b) is because that $\infty\cdot\mathcal{X}_{\{1\}} = \infty\cdot 0 = 0$.

11.2.1. Consequences of MCT I

Proposition 11.4 If f,g are measurable non-negative functions, and $f=g$ a.e., then

$$\int f\,\mathrm{d}m = \int g\,\mathrm{d}m$$

Proof. Let $U = \{x \in \mathbb{R} \mid f(x) = g(x)\}$, then

$$f = f \cdot \mathcal{X}_U + f \cdot \mathcal{X}_{U^c}$$

where U^c is null. As a result,

$$\begin{aligned}\int f \, \mathrm{d}m &= \int f\mathcal{X}_U + \int f\mathcal{X}_{U^c} && (11.2a)\\ &= \int g\mathcal{X}_U + 0 && (11.2b)\\ &= \int g\mathcal{X}_U + \int g\mathcal{X}_{U^c} && (11.2c)\\ &= \int g \, \mathrm{d}m && (11.2d)\end{aligned}$$

where (11.2b) is because that $f \cdot \mathcal{X}_{U^c} = 0$ a.e., and $f \cdot \mathcal{X}_U = g \cdot \mathcal{X}_U$; (11.2c) is becasue that $g \cdot \mathcal{X}_{U^c} = 0$ a.e. ■

Proposition 11.5 — **Slight Generalization of MCT I.** Suppose that $f_n(x)$ are nonnegative measurable functions such that

1. f_n is monotone increasing a.e.
2. $f_n(x) \to f(x)$ a.e.

then

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}m = \int f \, \mathrm{d}m$$

Proof. Construct the set $V_n = \{x \mid f_n(x) \leq f_{n+1}(x)\}$ and $V = \bigcap_{n=1}^{\infty} V_n$. Since $f_n(x)$ is monotone increasing a.e., we imply $m(V_n^c) = 0$, and $m(V^c) \leq \sum_{n=1}^{\infty} m(V_n^c) = 0$.

1. Construct $\tilde{f}_n(x)$ as follows:

$$\tilde{f}_n(x) = \begin{cases} f_n(x), \text{ if } x \in V \\ 0, \text{ if } x \in V^c \end{cases}$$

   As a result,

   - $\tilde{f}_n$ is monotone increasing

- Define a function $g : \mathbb{R} \to [0, \infty]$ such that $\lim_{n\to\infty} \tilde{f}_n(x) = g_n(x)$.

Apply the MCT I gives

$$\lim_{n\to\infty} \int \tilde{f}_n \,\mathrm{d}m = \int g \,\mathrm{d}m \tag{11.3a}$$

2. Note that $\{x \mid \tilde{f}_n(x) \neq f_n(x)\} \subseteq V^c$, where V^c is null. Therefore, $f_n = f$ a.e., which implies

$$\int \tilde{f}_n \,\mathrm{d}m = \int f_n \,\mathrm{d}m \tag{11.3b}$$

3. Consider $V' = \{x \mid \lim_{n\to\infty} f_n(x) = f(x)\}$, and $(V')^c$ is null by hyphothesis. For any $x \in V \cap V'$, we imply

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \tilde{f}_n(x).$$

Since $(V \cap V')^c$ is null, we imply $\tilde{f}_n(x) \to f$ a.e. Note that $\tilde{f}_n(x) \to g$, we imply $g = f$ a.e., which follows that

$$\int g \,\mathrm{d}m = \int f \,\mathrm{d}m \tag{11.3c}$$

Combining (11.3a) to (11.3c), we conclude that

$$\lim_{n\to\infty} \int f_n \,\mathrm{d}m = \lim_{n\to\infty} \int \tilde{f}_n \,\mathrm{d}m = \int g \,\mathrm{d}m = \int f \,\mathrm{d}m$$

∎

Proposition 11.6 Let $\{f_k\}$ be non-negative measurable and

$$f := \sum_{k=1}^{\infty} f_k,$$

then

$$\int f \,\mathrm{d}m = \sum_{k=1}^{\infty} \int f_k \,\mathrm{d}m$$

Proof. Firstly, $\int f \,\mathrm{d}m$ is well-defined since $f = \lim_{n\to\infty} \sum_{k=1}^{n} f_k$ is measurable.

Secondly, take $g_n = \sum_{k=1}^{n} f_k$, which implies g_n is monotone increasing and $g_n \to f$.

Apply MCT I gives the desired result. ■

■ **Example 11.3** Consider

$$(1-x)^{-1/2} = \sum_{n=0}^{\infty} \frac{(2n)!}{4^n (n!)^2} x^n, \quad x \in [0,1)$$

Take $f_k = \frac{(2k)!}{4^k (k!)^2} x^k$. Applying proposition (11.6) gives

$$\int_{[0,1)} (1-x)^{-1/2} \, \mathrm{d}x = \sum_{n=0}^{\infty} \int_0^1 \frac{(2n)!}{4^n \cdot (n!)^2} x^n \, \mathrm{d}x$$

Or equivalently,

$$2 = \sum_{n=0}^{\infty} \frac{(2n)!}{4^n (n!)(n+1)!}$$

■

11.2.2. MCT II

We now extend our study to all measurable functions instead of non-negativity.

Definition 11.4 [Lebesgue integrable] Let f be a measurable function, then let

$$f^+(x) = \begin{cases} f(x), & \text{if } f(x) > 0 \\ 0, & \text{if } f(x) \le 0 \end{cases} = f(x) \mathcal{X}_{f^{-1}((0,\infty])}$$

and

$$f^-(x) = \begin{cases} -f(x), & \text{if } f(x) \le 0 \\ 0, & \text{if } f(x) > 0 \end{cases} = -f(x) \mathcal{X}_{f^{-1}([-\infty,0])}$$

As a result, f^+ and f^- are both measurable.

Note that

- $f(x) = f^+(x) - f^-(x)$

- $|f|(x) = f^+(x) + f^-(x)$

Now we define the Lebesgue integral of f as

$$\int f \, \mathrm{d}m = \int f^+ \, \mathrm{d}m - \int f^- \, \mathrm{d}m$$

We say f is **Lebesgue integrable** if both f^+ and f^- are integrable, i.e., $\int f^\pm \, \mathrm{d}m < \infty$

Proposition 11.7

1. If f is measurable, then f is integrable if and only if $|f|$ is integrable
2. If f is measurable, and $|f| \le g$ with g integrable, then f is also integrable

Proof.

1. If f is integrable, then $\int f^+ \, \mathrm{d}m, \int f^- \, \mathrm{d}m < \infty$. As a result,

$$\int |f| \mathrm{d}m = \int (f^+ + f^-) \mathrm{d}m = \int f^+ \, \mathrm{d}m + \int f^- \, \mathrm{d}m < \infty.$$

   For the reverse direction, if $|f|$ is integrable, then

$$\int |f| = \int f^+ + \int f^-$$

   therefore $\int f^\pm < \infty$, and hence f is interable.

2. Since $0 \le |f| \le g$, by proposition (9.9), $\int |f| \mathrm{d}m \le \int g \, \mathrm{d}m < \infty$.
   Therefore, $\int |f| \mathrm{d}m < \infty$, and hence $|f|$ is integrable, which implies f is integrable.

■

R If $|f| \le g$, and $\int |f| \mathrm{d}m = \infty$, then by proposition (9.9), we imply $\int g \, \mathrm{d}m = \infty$.

11.3. Monday for MAT4002

Reviewing. Consider the group with presentation $\langle S \mid R(S)\rangle$.

1. The elements in S are generators that have studied in abstract algebra
2. The "relations" of this group are given by the equalities on hte right-hand side, e.g., the dihedral group is defined as

$$\langle a,b \mid a^n = e, b^2 = e, bab = a^{-1}\rangle$$

Sometimes we also simplify the equality $\times = e$ as $\times$, e.g., the dihedral group can be re-written as

$$\langle a,b \mid a^n, b^2, bab = a^{-1}\rangle$$

■ **Example 11.4** Consider

$$G = < a,b \mid a^2, b^2, abab^{-1}a^{-1}b^{-1} > := < a,b \mid a^2, b^2, aba = bab > = \{e,a,b,ab,ba,aba\}$$

It's isomorphic to S^3, and the shape of S^3 is illustrated in Fig.(11.1)

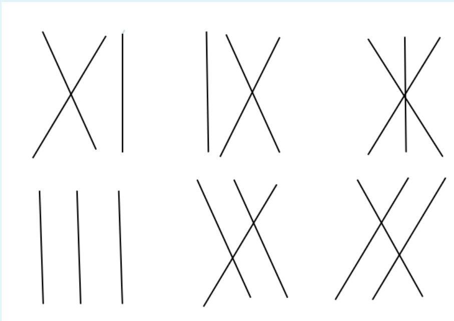


Figure 11.1: Illustration of group S^3

More precisely, the isomorphism is given by:

$$\begin{aligned} \phi: \quad & S_3 \to G \\ \text{with} \quad & X| \mapsto a, \quad |X \mapsto b \end{aligned}$$

■

■ **Example 11.5** Consider $G_2 =< a, b \mid ab = ba >$ and any word, which can be expressed as $\cdots a^s b^t a^u b^v \cdots$

- If $s \in \mathbb{N}$, we write $a^s := \underbrace{a \cdots a}_{s \text{ times}}$
- If $s \in -\mathbb{N}$, we write $a^s := \underbrace{(a^{-1}) \cdots (a^{-1})}_{-s \text{ times}}$
- For the word with the form $a \cdots b \cdots ba \cdots a$, we can always push a into the leftmost using the relation $ab = ba$
- For the word with the form $a \cdots ab \cdots ba^{-1}$, we can always push a^{-1} into the leftmost using the relation $ba^{-1} = a^{-1}b$.

Therefore, all elements in G_2 are of the form $a^p b^q, p, q \in \mathbb{Z}$, and we have the relation

$$(a^{p_1} b^{q_1})(a^{p_2} b^{q_2}) = a^{p_1+p_2} b^{q_1+q_2}.$$

Therefore, $G_2 \cong \mathbb{Z} \times \mathbb{Z}$, where the isomorphism is given by:

$$\begin{aligned} \phi : \quad & \mathbb{Z} \times \mathbb{Z} \to G_2 \\ \text{with} \quad & (p, q) \mapsto a^p b^q \end{aligned}$$

■

■ **Example 11.6**

$$G_3 = \langle a \mid a^5 \rangle = \{1, a, a^2, \ldots, a^4\}$$

It's clear that $G_3 \cong \mathbb{Z}/5\mathbb{Z}$, where the isomorphism is given by:

$$\begin{aligned} \phi : \quad & \mathbb{Z}/5\mathbb{Z} \to G_3 \\ \text{with} \quad & m + 5\mathbb{Z} \mapsto a^m \end{aligned}$$

■

11.3.1. Cayley Graph for finitely presented groups

Graphs have strong connection with groups. Here we introduce a way of building graphs using groups, and the graphs are known as Cayley graphs. They describe many properties of the group in a topological way.

Definition 11.5 [Oriented Graph] An oriented graph T is specified by

1. A countable or finite set V, known as vertices
2. A countable or finite set E, known as edges
3. A function $\delta : E \longrightarrow V \times V$ given by

$$\delta(e) = (\ell(e), \tau(e))$$

where $\ell(e)$ denotes the initial vertex and $\tau(e)$ denotes the terminal vertex.

For example, let

- $V = \{a, b, c\}$
- $E = \{e_1, e_2, e_3, e_4\}$
- $\delta(e_1) = (a, a), \delta(e_2) = (b, c), \delta(e_3) = (a, c), \delta(e_4) = (b, c)$

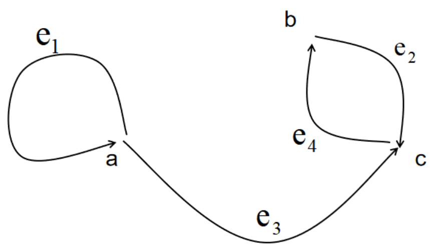


Figure 11.2: Illustration of example oriented graph

The resulted graph is plotted in Fig.(11.2)

Definition 11.6 [Cayley graph] Let $G = \langle S \mid R(S) \rangle$ with $|S| < \infty$. The **Cayley graph** associated to G is an oriented graph with

1. The vertex set G
2. The edge set $E := G \times S$
3. The function $\ell : E \longrightarrow V \times V$ is given by:

$$\begin{aligned} \ell : \quad & G \times S \to G \times G \\ \text{with} \quad & (g,s) \mapsto (g, g \cdot s) \end{aligned}$$

In particular, we link two elements in G if they differ by a generator rightside. ▪

▪ **Example 11.7** 1. The Cayley graph for $G = \langle a \rangle$ ($\cong \mathbb{Z}$) is shown in Fig.(11.3):

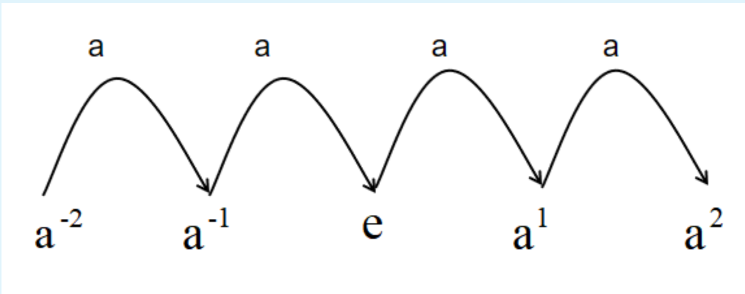


Figure 11.3: Illustration of Cayley Graph $\langle a \rangle$

2. The Cayley graph for $G = \langle a \mid a^3 \rangle$ is shown in Fig.(11.4):

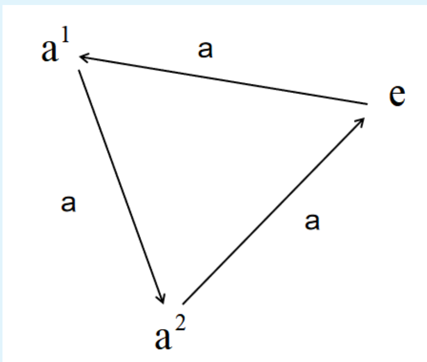


Figure 11.4: Illustration of Cayley Graph $\langle a \mid a^3 \rangle$

3. The Cayley graph for $G = \langle a, b \mid a^2, b^2, aba = bab \rangle$ is shown in Fig.(11.12):

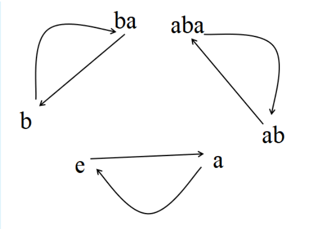


Figure 11.5: Illustration of Cayley Graph $\langle a, b \mid a^2, b^2, aba = bab \rangle$

4. The Cayley graph for $G = \langle a, b \mid ab = ba \rangle$ is shown in Fig.(11.6):

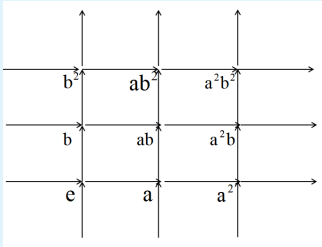


Figure 11.6: Illustration of Cayley Graph $\langle a, b \mid ab = ba \rangle$

5. The Cayley graph for $G = \langle a, b \rangle$ is shown in Fig.(11.7):

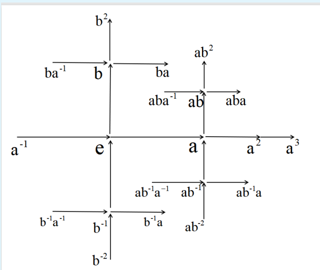


Figure 11.7: Illustration of Cayley Graph $\langle a, b \mid ab = ba \rangle$

R There could be different presentations $\langle S_1 \mid R(S_1) \rangle \cong \langle S_2 \mid R(S_2) \rangle$ of the same group.

11.3.2. Fundamental Group

Motivation. The fundamental group connects topology and algebra together, by labelling a group to each topological space, which is known as fundamental group.

Why do we need algebra in topology. Consider the S^2 (2-shpere) and $S^1 \times S^1$ (torus):

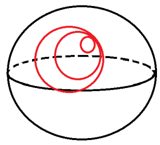


Figure 11.8: Any loop in the sphere can be contracted into a point

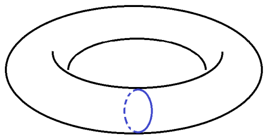


Figure 11.9: Some loops in the torus cannot be contracted into a point

As can be seen from Fig.(11.8) and Fig.(11.9), any "loop" on a sphere can be contracted to a point, while some "loop" on a torus cannot. We need the algebra to describe this phenomena formally.

Definition 11.7 [loop] Let X be a topological space. A **loop** on X is a constant map $\ell : [0,1] \to X$ such that $\ell(0) = \ell(1)$.

We say ℓ is based at $b \in X$ if $\ell(0) = \ell(1) = b$.

Definition 11.8 [composite loop] Suppose that $\boldsymbol{u}, \boldsymbol{v}$ are loops on X based at $b \in X$. The **composite loop** $u \cdot v$ is given by

$$u \cdot v = \begin{cases} u(2t), & \text{if } 0 \le t \le 1/2 \\ v(2t-1), & \text{if } 1/2 \le t \le 1 \end{cases}$$

Definition 11.9 [fundamental group] The **homotopy class of loops relative to** $\{0,1\}$ **based at** $b \in X$ forms a group. It is called the **fundamental group** of X based at b, denoted as $\pi_1(X,b)$.

More precisely, let

$$[\ell] = \{m \mid m \text{ is a loop based at } b \text{ that is homotopic to } \ell, \text{ relative to } \{0,1\}\},$$

and $\pi_1(X,b) = \{[\ell] \mid \ell \text{ are loops based at } b\}$. The operation in $\pi_1(X,b)$ is defined as:

$$[\ell] * [\ell'] := [\ell \cdot \ell'], \quad \forall [\ell],[\ell'] \in \pi_1(X,b).$$

■

R Two paths $\ell_1, \ell_2 : [0,1] \to X$ are homotopic relative to $\{0,1\}$ if we can find $H : [0,1] \times [0,1] \to X$ such that

$$H(t,0) = \ell_1(t), \quad H(t,1) = \ell_2(t)$$

and

$$H(0,s) = \ell_1(0) = \ell_2(0),\ \forall 0 \le s \le 1, \quad H(1,s) = \ell_1(1) = \ell_2(1),\ \forall 0 \le s \le 1$$

Counter example for homotopy but not relative to $\{0,1\}$:

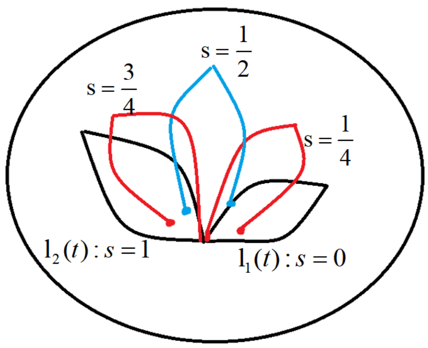


Figure 11.10: homotopy not relative to $\{0,1\}$

11.4. Wednesday for MAT3040

Reviewing. Unitary Operators

$$\langle T\boldsymbol{v}, T\boldsymbol{w}\rangle = \langle \boldsymbol{v}, \boldsymbol{w}\rangle,\ \forall \boldsymbol{v}, \boldsymbol{w} \in V.$$

11.4.1. Unitary Operator

▪ **Example 11.8** Let $V = \mathbb{R}^n$ with usual inner product. For the linear operator $T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$, T is orthogonal if and only if $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$.

Let $V = \mathbb{C}^n$ with usual inner product. For the linear operator $T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$, T is unitary if and only if $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} = \boldsymbol{I}$. ▪

Proposition 11.8 Let $T : V \to V$ be a linear operator on a vector space over $\mathbb{K}$ satisfying $T'T = I$. Then for all eigenvalues λ of T, we have $|\lambda| = 1$.

Proof. Suppose we have the eigen-pair $(\lambda, \boldsymbol{v})$, then

$$\begin{aligned}
&\langle T\boldsymbol{v}, T\boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{v}\rangle \\
\iff &\langle \lambda\boldsymbol{v}, \lambda\boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{v}\rangle \\
\iff &\bar{\lambda}\lambda\langle \boldsymbol{v}, \boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{v}\rangle
\end{aligned}$$

Since $\langle \boldsymbol{v}, \boldsymbol{v}\rangle \neq 0$ $(\boldsymbol{v} \neq \boldsymbol{0})$, we imply $|\lambda|^2 = 1$, i.e., $|\lambda| = 1$. ■

Proposition 11.9 Let $T : V \to V$ be an operator on a finite dimension V over $\mathbb{K}$ satisfying $T'T = I$. If $U \le V$ is T-invariant, then U is also T^{-1}-invariant.

Proof. Since $T'T = I$, i.e., T is invertible, we imply 0 is not a root of $\mathcal{X}_T(x)$, i.e., 0 is not a root of $m_T(x)$. Since $m_T(0) \neq 0$, $m_T(x)$ has the form

$$m_T(x) = x^m + \cdots + a_1 x + a_0,\ a_0 \neq 0,$$

which follows that

$$m_T(T) = T^m + \cdots + a_0 I = 0 \implies T(T^{m-1} + \cdots + a_1 I) = -a_0 I$$

Or equivalently,

$$T\left(-\frac{1}{a_0}(T^{m-1} + \cdots + a_1 I)\right) = I$$

Therefore,

$$T^{-1} = -\frac{1}{a_0}T^{m-1} - \cdots - \frac{a_2}{a_0}T - \frac{a_1}{a_0}I,$$

i.e., the inverse T^{-1} can be expressed as a polynomial involving T only.

Sicne U is T-invariant, we imply U is T^m-invariant for $m \in \mathbb{N}$, and therefore U is T^{-1}-invariant since T^{-1} is a polynomial of T. ■

Proposition 11.10 Let $T : V \to V$ satisfies $T'T = I$ ($\dim(V) < \infty$), then $U \le V$ is T-invariant implies $U^\perp$ is T-invariant.

Proof. Let $v \in U^\perp$, it suffices to show $T(v) \in U^\perp$.

For all $u \in U$, we have

$$\langle u, T(v)\rangle = \langle T'(u), v\rangle = \langle T^{-1}(u), v\rangle$$

Since U is T^{-1}-invaraint, we imply $T^{-1}(u) \in U$, and therefore

$$\langle u, T(v)\rangle = \langle T^{-1}(u), v\rangle = 0 \implies T(v) \in U^\perp.$$

■

Theorem 11.2 Let $T : V \to V$ be a unitary operator on finite dimension V (over $\mathbb{C}$), then there exists an orthonormal basis $\mathcal{A}$ such that

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n),\ |\lambda_i| = 1,\ \forall i.$$

Proof Outline. Note that $\mathcal{X}_T(x)$ always admits a root in $\mathbb{C}$, so we can always find an eigenvector $\boldsymbol{v} \in V$ of T.

Then the theorem follows by the same argument before on seld-adjoint operators.

- Consider $U = \text{span}\{\boldsymbol{v}\}$
- $V = U \oplus U^{\perp}$ and $U^{\perp}$ is T-invariant
- Use induction on the unitary operator $T\mid_{U^{\perp}}: U^{\perp} \to U^{\perp}$

■

R

- The argument fails for orthogonal operators

$$\begin{aligned} T \quad & : \mathbb{R} \to \mathbb{R}^2, \\ \text{with} \quad & T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v} \\ \text{where} \quad & \boldsymbol{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{aligned}$$

The matrix $\boldsymbol{A}$ is not diagonalizable over $\mathbb{R}$. It has no real eigenvalues. However, if we treat $\boldsymbol{A}$ as $T : \mathbb{C}^2 \to \mathbb{C}^2$ with $T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$, then $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} = \boldsymbol{I}$, and therefore T is unitary. Then $\boldsymbol{A}$ is diagonalizable over $\mathbb{C}$ with eigenvalues $e^{i\theta}, e^{-i\theta}$

- As a corollary of the theorem, for all $\boldsymbol{A} \in M_{n\times n}(\mathbb{C})$ satisfying $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} = \boldsymbol{I}$, there exists $P \in M_{n\times n}(\mathbb{C})$ such that

$$P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n), \quad |\lambda_i| = 1,$$

where $P = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, with $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\}$ forming orthonormal basis of $\mathbb{C}^n$.

In fact,

$$P^{\mathrm{H}}P = \begin{pmatrix} \boldsymbol{u}_1^{\mathrm{H}} \\ \vdots \\ \boldsymbol{u}_n^{\mathrm{H}} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{pmatrix} = \begin{pmatrix} \langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle & \cdots & \langle \boldsymbol{u}_1, \boldsymbol{u}_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \boldsymbol{u}_n, \boldsymbol{u}_1 \rangle & \cdots & \langle \boldsymbol{u}_n, \boldsymbol{u}_n \rangle \end{pmatrix}$$

Conclusion: all matrices $\boldsymbol{A} \in M_{n\times n}(\mathbb{C})$ with $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} = \boldsymbol{I}$ can be written as

$$\boldsymbol{A} = \boldsymbol{P}^{-1}\operatorname{diag}(\lambda_1,\dots,\lambda_n)\boldsymbol{P},$$

with some $\boldsymbol{P}$ satisfying $\boldsymbol{P}^{\mathrm{H}}\boldsymbol{P} = \boldsymbol{I}$.

Notation. Let $U(n) = \{\boldsymbol{A} \in M_{n\times n}(\mathbb{C}) \mid \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} = \boldsymbol{I})\}$ be the unitary group, then all $\boldsymbol{A} \in U(n)$ can be diagonalized by

$$A = P^{-1}\operatorname{diag}(\lambda_1,\dots,\lambda_n)P, \quad P \in U(n).$$

11.4.2. Normal Operators

Definition 11.10 [Normal] Let $T : V \to V$ be a linear operator over a $\mathbb{C}$ inner product vector space V. We say T is **normal**, if

$$T'T = TT'$$

■

■ **Example 11.9**

- All self-adjoint operators are normal:

$$T = T' \implies TT' = T'T = T^2$$

- All (finite-dimensional) unitary operators are normal:

$$T'T = TT' = I$$

■

Proposition 11.11 Let T be a normal operator on V. Then

1. $\|T(\boldsymbol{v})\| = \|T'(\boldsymbol{v})\|, \forall \boldsymbol{v} \in V$.

   In particular, $T(\boldsymbol{v}) = 0$ if and only if $T'(\boldsymbol{v}) = 0$

2. $(T - \lambda I)$ is also a normal operator, for any $\lambda \in \mathbb{C}$
3. $T(\boldsymbol{v}) = \lambda \boldsymbol{v}$ if and only if $T'(\boldsymbol{v}) = \bar{\lambda}\boldsymbol{v}$.

Proof. 1.

$$\begin{aligned}
\langle T\boldsymbol{v}, T\boldsymbol{v}\rangle &= \langle T'T\boldsymbol{v}, \boldsymbol{v}\rangle \\
&= \langle TT'\boldsymbol{v}, \boldsymbol{v}\rangle \\
&= \overline{\langle \boldsymbol{v}, TT'\boldsymbol{v}\rangle} \\
&= \overline{\langle T'\boldsymbol{v}, T'\boldsymbol{v}\rangle} \\
&= \langle T'\boldsymbol{v}, T'\boldsymbol{v}\rangle
\end{aligned}$$

Therefore, $\|T(\boldsymbol{v})\|^2 = \|T'(\boldsymbol{v})\|^2$, i.e., $\|T(\boldsymbol{v})\| = \|T'(\boldsymbol{v})\|$.

2. By hw4, $(T - \lambda I)' = T' - \overline{\lambda} I$. It suffices to check

$$(T - \lambda I)'(T - \lambda I) = (T - \lambda I)(T - \lambda I)',$$

Expanding both sides out gives the desired result, i.e.,

$$(T - \lambda I)'(T - \lambda I) = (T' - \bar{\lambda} I)(T - \lambda I) = T'T - \bar{\lambda} T - \lambda T' + \lambda\bar{\lambda} I$$

and

$$(T - \lambda I)(T - \lambda I)' = (T - \lambda I)(T' - \bar{\lambda} I) = TT' - \bar{\lambda} T - \lambda T' + \lambda\bar{\lambda} I$$

3. The proof for (3) will be discussed in the next lecture.

■

11.5. Wednesday for MAT3006

Proposition 11.12 — **Linearity.** If f, g are both integrable, then $f + g$ and αf are integrable with

$$\int (f+g)\,\mathrm{d}m = \int f\,\mathrm{d}m + \int g\,\mathrm{d}m$$
$$\int \alpha f\,\mathrm{d}m = \alpha \int f\,\mathrm{d}m \quad \alpha \in \mathbb{R}$$

Proof. 1. Construct

$$(f+g)^+ - (f+g)^- = f+g = (f^+ - f^-) + (g^+ - g^-) \implies (f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g$$

Since both sides for the equality above is non-negative, we do the Lebesgue integral both sides:

$$\int ((f+g)^+ + f^- + g^-)\,\mathrm{d}m = \int ((f+g)^- + f^+ + g^+)\,\mathrm{d}m.$$

Due to the linearity of Lebesgue integral for non-negative functions,

$$\int (f+g)^+\,\mathrm{d}m + \int f^-\,\mathrm{d}m + \int g^-\,\mathrm{d}m = \int (f+g)^-\,\mathrm{d}m + \int f^+\,\mathrm{d}m + \int g^+\,\mathrm{d}m$$

i.e.,

$$\int (f+g)\,\mathrm{d}m = \int f\,\mathrm{d}m + \int g\,\mathrm{d}m$$

2. Assume $\alpha < 0$. Then

$$\begin{aligned}
\int (\alpha f)\,\mathrm{d}m &:= \int (\alpha f)^+\,\mathrm{d}m - \int (\alpha f)^-\,\mathrm{d}m \\
&= \int (-\alpha) f^-\,\mathrm{d}m - \int (-\alpha) f^+\,\mathrm{d}m \\
&= (-\alpha)\int f^-\,\mathrm{d}m - (-\alpha)\int f^+\,\mathrm{d}m \\
&= \alpha\left(\int f^+\,\mathrm{d}m - \int f^-\,\mathrm{d}m\right) \\
&= \alpha\int f\,\mathrm{d}m
\end{aligned}$$

The proof for the case $\alpha \geq 0$ follows similarly.

■

11.5.1. Properties of Lebesgue Integrable Functions

Corollary 11.2 Suppose that f, g are integrable, then

1. If $f \leq g$, then $\int f\,\mathrm{d}m \leq \int g\,\mathrm{d}m$
2. If $f = g$ a.e., then $\int f\,\mathrm{d}m = \int g\,\mathrm{d}m$

Proof. 1. Since $g - f \geq 0$, $\int (g-f)\,\mathrm{d}m \geq \int 0\,\mathrm{d}m = 0$. By linearity, $\int g\,\mathrm{d}m - \int f\,\mathrm{d}m \geq 0$, i.e.,

$$\int g\,\mathrm{d}m \geq \int f\,\mathrm{d}m.$$

2. The proof follows similarly as in proposition (11.4). In detail, let $U = \{x \mid f(x) = g(x)\}$, then $m(U) = 0$. It follows that

$$\int f\mathcal{X}_{U^C}\,dm = \int f^+\mathcal{X}_{U^c}\,dm + \int f^-\mathcal{X}_{U^c}\,dm = 0$$

Similarly, $\int g\mathcal{X}_{U^c}dm = 0$. Therefore,

$$\begin{aligned}\int f\,\mathrm{d}m &= \int f\mathcal{X}_U\,\mathrm{d}m + \int f\mathcal{X}_{U^c}\,\mathrm{d}m \\ &= \int g\mathcal{X}_U\,\mathrm{d}m \\ &= \int g\mathcal{X}_U\,\mathrm{d}m + \int g\mathcal{X}_{U^c}\,\mathrm{d}m \\ &= \int g\,\mathrm{d}m\end{aligned}$$

■

R

1. Consider the set of integrable functions, say $\mathcal{T} = \{f : \mathbb{R} \to [-\infty, \infty], \text{ integrable}\}$, which is a vector space if we define $0_{\mathcal{T}} :=$ zero function.
   We can define a "norm" on $f \in \mathcal{T}$ by
   $$\|f\| = \int |f|\,\mathrm{d}m$$
   then $\|\alpha f\| = |\alpha|\|f\|$ and $\|f+g\| \le \|f\| + \|g\|$.
   Unfortunately, we should keep in mind that $\mathcal{T}$ is not a normed space, since there exists $f \ne 0_{\mathcal{T}}$ such that $\|f\| = 0$, e.g., $f = \mathcal{X}_{\mathbb{Q}}$.
2. To remedy this, define the equivalence relation on $\mathcal{T}$: $f \sim g$ if $f = g$ a.e.
   The equivalence classes of $\mathcal{T}$ under $\sim$ are of the form $[f] := \{g : g \sim f\}$.
   Denote the collection of equivalence classes as $L^1(\mathbb{R}) := \mathcal{T}/\sim$.
   (a) It's clear that $L^1(\mathbb{R})$ has a vector space structure
   $$[f] + [g] = [f+g]$$
   $$\alpha[f] = [\alpha f]$$
   (b) The space $L^1(\mathbb{R})$ can be viewed as a quotient space defined in linear

algebra. Consider a vector subspace $\mathcal{N}$ of $\mathcal{T}$ defined by

$$\mathcal{N} := \{g \in \mathcal{T} \mid g = 0 \text{ a.e.}\}$$

then $\mathcal{T}/\sim = \mathcal{T}/\mathcal{N}$.

(c) We define a norm on $L^1(\mathbb{R})$ by $\|[f]\| = \int |f| \, \mathrm{d}m$, which is truly a norm:

$$\begin{aligned}
\|\alpha[f]\| &= |\alpha| \|[f]\| \\
\|[f] + [g]\| &\le \|[f]\| + \|[g]\| \\
\|[f]\| = 0 &\Longleftrightarrow \int |f| \, \mathrm{d}m = 0 \Longleftrightarrow f = 0 \text{ a.e.} \Longleftrightarrow [f] = 0_{L^1(\mathbb{R})}
\end{aligned}$$

Similarly, we can study $L^2(\mathbb{R}), \dots, L^p(\mathbb{R})$, e.g., for $L^2(\mathbb{R}) = \{f : \mathbb{R} \to [-\infty, \infty] \mid \int |f|^2 \, \mathrm{d}m < \infty\}/\mathcal{N}$, define the norm

$$\|f\|_2 = (\int |f|^2 \, \mathrm{d}m)^{1/2}$$

■ **Example 11.10** There exist some improper Riemann integrable functions that are not Lebesgue integrable: Consider $f = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \mathcal{X}_{[k,k+1)}$, then the improper Riemann integral gives

$$\int_0^{\infty} f(x) \, \mathrm{d}x = \log(2) = 1 - \frac{1}{2} + \frac{1}{3} - \cdots$$

However, f is not Lebesgue integrable. Suppose on the contrary that it is , then $|f|$ is integrable:

$$|f| = \sum_{k=0}^{\infty} \frac{1}{k+1} \mathcal{X}_{[k,k+1)}$$

However,

$$\int |f|\,\mathrm{d}m = \lim_{n\to\infty}\sum_{k=0}^{n}\int\left(\frac{1}{k+1}\chi_{[k,k+1)}\right)\mathrm{d}m$$
$$= \sum_{k=0}^{\infty}\frac{1}{k+1} = \infty$$

We will also show that all the proper Riemann integrable functions are Lebesgue integrable (and the integrals have the same value)

Theorem 11.3 — **MCT II.** Let $\{f_n\}$ be a sequence of integrable functions such that

1. $f_n \le f_{n+1}$ a.e.
2. $\sup_n \int f_n\,\mathrm{d}m < \infty$

Then f_n converges to an integrable function f a.e., and

$$\int f\,\mathrm{d}m = \lim_{n\to\infty}\int f_n\,\mathrm{d}m$$

Proof. Re-define f_n by changing its values on a null set such that

1. $f_n(x) \in \mathbb{R}$, for any $x \in \mathbb{R}$
2. $f_n(x) \le f_{n+1}(x)$ for any $n \in \mathbb{R}, x \in \mathbb{R}$

Let $f(x) = \lim_{n\to\infty} f(x)$. Consider the sequence of functions $\{f_n - f_1\}_{n\in\mathbb{N}}$, then

1. $f_n - f_1 \ge 0$
2. $f_n - f_1$ is monotone increasing, integrable
3. $f_n - f_1 \to f - f_1$

Applying MCT I gives

$$\int (f - f_1)\,\mathrm{d}m = \lim_{n\to\infty}\int (f_n - f_1)\,\mathrm{d}m$$

Adding $\int f_1\,\mathrm{d}m$ and applying the linearity of integrals, we obtain

$$\int (f - f_1)\,\mathrm{d}m + \int f_1\,\mathrm{d}m = \lim_{n\to\infty}\int (f_n - f_1)\,\mathrm{d}m + \int f_1\,\mathrm{d}m = \lim_{n\to\infty}\int f_n\,\mathrm{d}m$$

Here $\lim_{n\to\infty}\int f_n \,\mathrm{d}m$ exists as $\lim_{n\to\infty}\int f_n \,\mathrm{d}m = \sup_n \int f_n \,\mathrm{d}m < \infty$; and $\int(f - f_1)\,\mathrm{d}m + \int f_1 \,\mathrm{d}m$ is integrable since it equals $\lim_{n\to\infty}\int f_n \,\mathrm{d}m < \infty$.

Therefore,

$$\text{LHS} = \int f \,\mathrm{d}m = \text{RHS} = \lim_{n\to\infty}\int f_n \,\mathrm{d}m.$$

The proof is complete. ■

11.6. Wednesday for MAT4002

11.6.1. The fundamental group

Revewing. One example for Homotopy relative to $\{0,1\}$ is illustrated in Fig.(11.4)

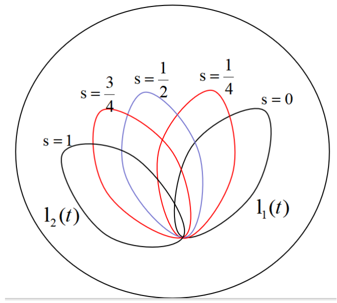


Figure 11.11: Example of homotopy relative to $\{0,1\}$

It's **essential** to study homotopy relative to $\{0,1\}$. For example, given a torus with a loop $\ell_1(t)$ and a base point b. We want to distinguish $\ell_1(t)$ and $\ell_2(t)$ as shown in Fig.(11.12):

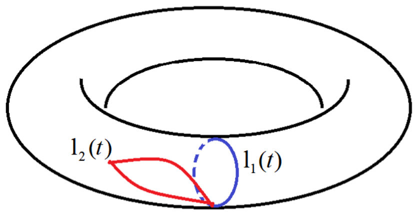


Figure 11.12: Two loops on a torus

Obviously there should be something different between $\ell_1(t)$ and $\ell_2(t)$. "Relative to$\{0,1\}$ is essential", sicne if we get rid of this condition, all loops are homotopic to the constant map $c_b(t) = b$. See the graphic illustration in Fig.(**??**):

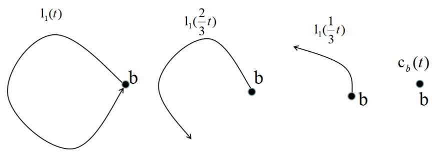


Figure 11.13: homotopy between any loop and constant map

In this case, $\ell \simeq c_b$ for any loop ℓ, there is only one trivial element $\{[c_b]\}$ in $\pi_1(X,b)$.

That's the reason why we define $\pi_1(X,b)$ as the collection of homotopy classes **relative to** $\{0,1\}$ based at b in X.

Proposition 11.13 Let $[\cdot]$ denote the homotopy class of loops relative to $\{0,1\}$ based at b, and define the operation

$$[\ell] * [\ell'] = [\ell \cdot \ell']$$

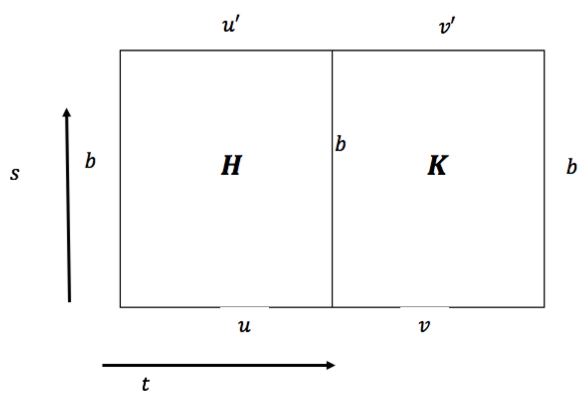
Then $(\pi_1(X,b),*)$ forms a group, where

$$\pi_1(X,b) := \{[\ell] \mid \ell : [0,1] \to X \text{ denotes loops based at } b\}$$

Proof. 1. Well-definedness: Suppose that $u \sim u'$ and $v \sim v'$, it suffices to show $u \cdot v \simeq u' \cdot v'$. Consider the given homotopies $H : u \simeq u'$, $K : v \simeq v'$. Construct a new homotopy $L : I \times I \to X$ by

$$L(t,s) = \begin{cases} H(2t,s), & 0 \le t \le 1/2 \\ K(2t-1,s), & 1/2 \le t \le 1 \end{cases}$$

The diagram below explains the ideas for constructing L. The plane denote the set $I \times I$, and the labels characterize the images of each point of $I \times I$ under L.

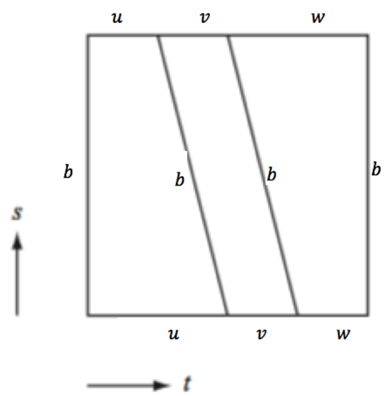


Therefore, $u \cdot v \simeq u' \cdot v'$.

2. Associate: $(u \cdot v) \cdot w \simeq u \cdot (v \cdot w)$

   Note that $(u \cdot v) \cdot w$ and $u \cdot (v \cdot w)$ are essentially different loops. Although they go with the same path, they are with different speeds. Generally speaking, the loop $(u \cdot v) \cdot w$ travels u, v using $1/4$ seconds, and w in $1/2$ seconds; but the loop $u \cdot (v \cdot w)$ travels u in $1/2$ seconds, and then v, w in $1/4$ seconds.

   We want to construct a homotopy that describes the loop changes from $u \cdot (v \cdot w)$ to $(u \cdot v) \cdot w$. A graphic illustration is given below:



   An explicit homotopy $H : I \times I \longrightarrow X$ is given below:

$$H(t,s) = \begin{cases} u(4t/(2-s)), & 0 \le t \le 1/2 - 1/4s \\ v(4t-2+s), & 1/2 - 1/4s \le t \le 3/4 - 1/4s \\ w(4t-3+s/(1+s)), & 3/4 - 1/4s \le t \le 1 \end{cases}$$

   Therefore,

$$[u] * ([v] * [w]) = ([u] * [v]) * [w]$$

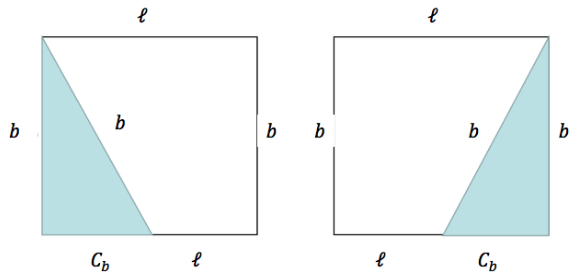
3. Intuitively, the identity should be the constant map, i.e., let $c_b : I \longrightarrow X$ by $c_b(t) = b, \forall t$, and let $\ell = [c_b]$, it suffices to show

$$[c_b] * [\ell] = [\ell] * [c_b] = [\ell] \Longleftrightarrow [c_b \cdot \ell] = [\ell \cdot c_b] = [\ell]$$

Or equivalently,

$$c_b \cdot \ell \simeq \ell, \quad \ell \cdot c_b \simeq \ell$$

The graphic homotopy is shown below. (You should have been understood this diagram)



4. Inverse: the inverse of $[u]$, where u is a loop, should be $[u']$, where u' is the reverse of the traveling of u. Therefore, for all $u : I \longrightarrow X$ (loop based at b), define $u^{-1} : I \longrightarrow X$ by $u^{-1}(t) = u(1-t)$. Note that

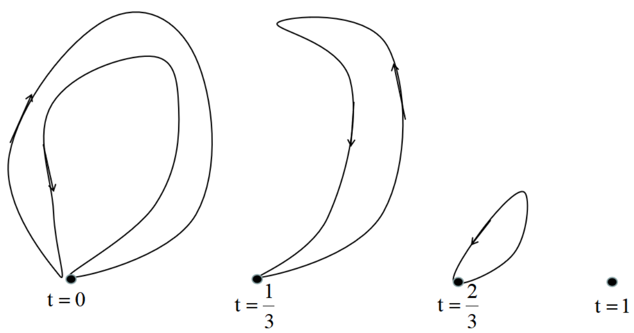
$$[u] * [u^{-1}] = [u \cdot u^{-1}], \quad e = [c_b]$$

It suffices to show $u \cdot u^{-1} \simeq c_b$ and $u^{-1} \cdot u \simeq c_b$:

The homotopy below gives $u \cdot u^{-1} \simeq c_b$, and the $u^{-1} \cdot u \simeq c_b$ follows similarly.

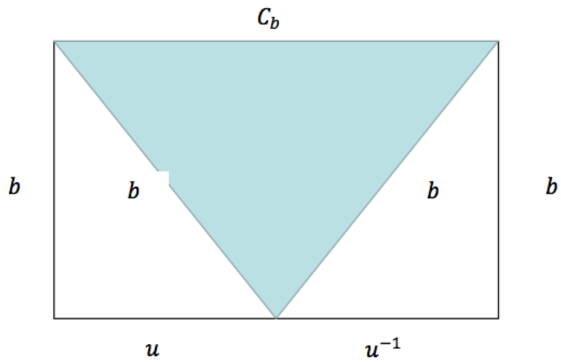
$$H(t,s) = \begin{cases} u(2t(1-s)), & 0 \le t \le 1/2 \\ u((2-2t)(1-s)), & 1/2 \le t \le 1 \end{cases}$$

The graphic illustration is given below:



■

R Note that the figure below does not define a homotopy from $u \cdot u^{-1}$ to c_b!



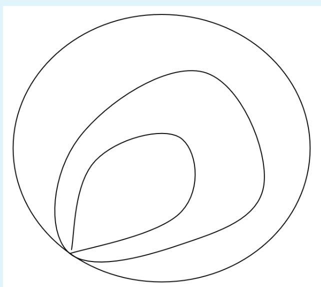
The reason is that for the upper part, as $s \longrightarrow 1$, the time for traveling u and u^{-1} becomes very small, i.e., a particle has to pass u and u^{-1} in infinitely small time, which is not well-defined.

■ **Example 11.11** The reason why $\pi_1(\mathbb{R}^2, b) = \{e\}$ is trivial:

- For any $u : I \longrightarrow \mathbb{R}^2$ with $u(0) = u(1) = b$, consider the homotopy

$$H(t,s) = (1-s)u(t) + sb.$$

  Therefore, $u \simeq c_b$ for any loop u based at b. Check the diagram below for graphic illustration of this homotopy.



More generally, if $X \simeq \{x\}$ is contractible, then $\pi_1(X,b) = \{e\}$. The same argument cannot work for $(\mathbb{R}^2\{0\}, \boldsymbol{b})$, since the mapping $H : \mathbb{R}^2 \setminus \{0\} \times I \to \mathbb{R}^2 \setminus \{0\}$ with $H(\boldsymbol{t}, s) = (1-s)u(\boldsymbol{t}) + s\boldsymbol{b}$ is not well-defined. In particular, the value $H(s,t)$ may hit the origin $\mathbf{0}$.

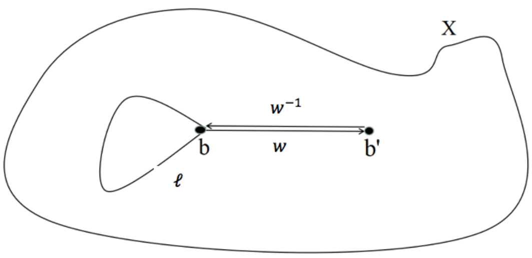
However, $\pi_1(S^1, 1)$ is non-trivial. We cannot deform the loop in S^1 into a constant loop. We will see that $\pi_1(S^1, 1) \cong \mathbb{Z}$. ▪

Proposition 11.14 If b, b' are path-connected in X, then $\pi_1(X,b) \cong \pi_1(X,b')$.

Proof. Let w be a path from b to b', and define

$$\begin{aligned} w_\# : \quad & \pi_1(X,b) \to \pi_1(X,b') \\ \text{with} \quad & [\ell] \mapsto [w^{-1}\ell w] \end{aligned}$$

1. Well-definedness: Check that $\ell \simeq \ell'$ implies $w^{-1}\ell w \simeq w^{-1}\ell' w$. See the figure below for graphic illustration.



2. $w_\#$ is a homomorphism:

$$w_\#([\ell_1]) \cdot w_\#([\ell_2]) = [w^{-1} \cdot \ell_1 w] \cdot [w^{-1} \cdot \ell_2 w] \tag{11.4a}$$
$$= [w^{-1} \cdot \ell_1 \ell_2 w] \tag{11.4b}$$
$$= w_\#([\ell_1 \ell_2]) \tag{11.4c}$$

where (11.4b) is because that $w \cdot w^{-1} = c_b$.

3. And $w_\#$ is also injective. If loops ℓ_1, ℓ_2 are such that $w_\#(\ell_1) = w_\#(\ell_2)$, then

$$[w^{-1}\ell_1 w] = [w^{-1}\ell_2 w],$$

which follows that

$$[\ell_1] = [w][w^{-1}\ell_1 w][w^{-1}] = [w][w^{-1}\ell_2 w][w^{-1}] = [\ell_2] \tag{11.5}$$

4. Finally, $w_\#$ is surjective, because for any $u \in \pi_1(X, b')$, let $v = wuw^{-1}$, then v is based at b, so $[v] \in \pi_1(X, b)$, and $w_\#(v) = [u]$. Therefore $w_\#$ is surjective.

In conclusion, $w_\#$ is a group isomorphism between $\pi_1(X, b)$ and $\pi_1(X, b')$. ■

R In (11.5) we extended the meaning of $[\ell]$ to allow ℓ to be a path, and the equivalence class is defined by the relation "$\sim$": $\ell_1 \sim \ell_2$ iff they are homotopic relative to $\{0, 1\}$. The multiplication rules are defined similarly.

Chapter 12

Week12

12.1. Monday for MAT3040

12.1.1. Remarks on Normal Operator

Proposition 12.1 If T is normal, then

1. $\|T(\boldsymbol{v})\| = \|T'(\boldsymbol{v})\|$ for any $\boldsymbol{v} \in V$
2. $(T - \lambda I)$ is normal for any $\lambda \in \mathbb{C}$
3. $T(\boldsymbol{v}) = \lambda\boldsymbol{v}$ if and only if $T'(\boldsymbol{v}) = \bar{\lambda}\boldsymbol{v}$
4. If $T(\boldsymbol{v}) = \lambda\boldsymbol{v}$ and $T(\boldsymbol{w}) = \mu\boldsymbol{w}$ with $\lambda \neq \mu$, then $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0$.

Proof. (3) • For the forward direction, if $(T - \lambda I)\boldsymbol{v} = 0$, then by part (2), $(T - \lambda I)$ is normal, which follows that

$$\|(T - \lambda I)'(\boldsymbol{v})\| = 0 \implies (T - \lambda I)'(\boldsymbol{v}) = 0 \implies T'\boldsymbol{v} = \bar{\lambda}\boldsymbol{v}.$$

• For the reverse direction, suppose that $(T' - \bar{\lambda}I)\boldsymbol{v} = 0$. Since T is normal, we imply T' is normal. Then by part (2), $(T' - \bar{\lambda}I)$ is normal. By applying the same trick,

$$(T' - \bar{\lambda}I)'\boldsymbol{v} = 0 \implies ((T')' - \bar{\bar{\lambda}}I)\boldsymbol{v} = 0.$$

By hw4, $(T')' = T$. Therefore, $(T - \lambda I)\boldsymbol{v} = 0$.

(4) Observe that

$$\lambda\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \langle \bar{\lambda}\boldsymbol{v}, \boldsymbol{w}\rangle \xRightarrow{\text{by (3)}} \lambda\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \langle T'(\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle = \langle \boldsymbol{v}, \mu\boldsymbol{w}\rangle = \mu\langle \boldsymbol{v}, \boldsymbol{w}\rangle$$

Since $\lambda \neq \mu$, we imply $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$. The proof is complete.

■

Theorem 12.1 Let T be an operator on a finite dimensional ($\dim(V) = n$) $\mathbb{C}$-inner product vector space V satisfying $T'T = TT'$. Then there is an orthonormal basis of eigenvectors of V, i.e., an orthonormal basis of V such that any element from this basis is an eigenvector of T.

Proof. Since $\mathcal{X}_T(x)$ must have a root in $\mathbb{C}$, there must exist an eigen-pair $(\boldsymbol{v}, \lambda)$ of T.

- Construct $U = \text{span}\{\boldsymbol{v}\}$, and it follows that

$$T\boldsymbol{v} = \lambda\boldsymbol{v} \implies U \text{ is } T\text{-invariant.}$$

$$T'\boldsymbol{v} = \bar{\lambda}\boldsymbol{v} \implies U \text{ is } T'\text{-invariant.}$$

- Moreover, we claim that $U^\perp$ is T and T' invariant: let $\boldsymbol{w} \in U^\perp$, and for all $\boldsymbol{u} \in U$, we have

$$\langle \boldsymbol{u}, T(\boldsymbol{w}) \rangle = \langle T'(\boldsymbol{u}), \boldsymbol{w} \rangle = \langle \bar{\lambda}\boldsymbol{u}, \boldsymbol{w} \rangle = \lambda \langle \boldsymbol{u}, \boldsymbol{w} \rangle = 0,$$

  i.e., $U^\perp$ is T invariant.

$$\langle \boldsymbol{u}, T'(\boldsymbol{w}) \rangle = \langle T(\boldsymbol{u}), \boldsymbol{w} \rangle = \langle \lambda\boldsymbol{u}, \boldsymbol{w} \rangle = \bar{\lambda} \langle \boldsymbol{u}, \boldsymbol{w} \rangle = 0,$$

  which implies $U^\perp$ is T' invariant.

- Therefore, we construct the operator $T\mid_{U^\perp} : U^\perp \to U^\perp$, and

$$TT' = T'T \implies (T\mid_{U^\perp})(T'\mid_{U^\perp}) = (T'\mid_{U^\perp})(T\mid_{U^\perp}),$$

  i.e., $(T\mid_{U^\perp})$ is normal on $U^\perp$. Moreover, $\dim(U^\perp) = n - 1$.

- Applying the same trick as in Theorem (11.1), we imply there exists an orthonormal basis $\{\boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ of eigenvectors of $(T\mid_{U^\perp})$. Then we can argue that

$$\mathcal{B} = \{\boldsymbol{v}' = \boldsymbol{v}/\|\boldsymbol{v}\|, \boldsymbol{e}_2, \dots, \boldsymbol{e}_{k+1}\}$$

is a basis of orthonormal eigenvectors of V.

∎

Corollary 12.1 [Spectral Theorem for Normal Operator] Let $T : V \to V$ be a normal operator on a $\mathbb{C}$-inner product space with $\dim(V) < \infty$. Then there exists self-adjoint operators $P_1, \dots, P_k$ such that

$$P_i^2 = P_i, \quad P_iP_j = 0, i \neq j, \quad \sum_{i=1}^k P_i = I,$$

and $T = \sum_{i=1}^k \lambda_i P_i$, where λ_i's are the eigenvalues of T.

(R) These P_i's are the **orthogonal projections** from V to the λ_i-eigenspace $\ker(T - \lambda_i I)$ of T, i.e.,we have

$$\begin{aligned} & v = P_i(v) + (v - P_i(v)), \\ \text{where} \quad & P_i(v) \in \ker(T - \lambda_i I), \text{ and } v - P_i(v) \in (\ker(T - \lambda_i I))^\perp. \end{aligned}$$

You should know how to compute P_i's when $T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$ in the course MAT2040.

Proof. Since T has a basis of eigenvectors, by definition, T is diagonalizable. By proposition (8.2),

$$m_T(x) = (x - \lambda_1) \cdots (x - \lambda_k),$$

where λ_i's are distinct. By spectral decomposition corollary (9.2), it suffices to show P_i's are self-disjoint.

- Recall that $P_i = a_i(T)q_i(T) := b_m T^m + \cdots + b_1 T + b_0 T$, i.e., a polynomial of T, and therefore

$$P_i' = \bar{b}_m(T')^m + \cdots + \bar{b}_1(T') + \bar{b}_0 I.$$

We claim that P_i is normal: Since $T'T = TT'$, we imply

$$(T')^p T^q = T^q (T')^p, \forall p, q \in \mathbb{N}$$

which follows that

$$\begin{aligned}
P_i P_i' &= (b_m T^m + \cdots + b_0 I)(\bar{b}_m(T')^m + \cdots + \bar{b}_1(T') + \bar{b}_0 I) \\
&= \sum_{1 \le x, y \le m} b_x \bar{b}_y (T)^x (T')^y \\
&= \sum_{1 \le x, y \le m} \bar{b}_y b_x (T')^y (T)^x \\
&= (\bar{b}_m(T')^m + \cdots + \bar{b}_1(T') + \bar{b}_0 I)(b_m T^m + \cdots + b_0 I) \\
&= P_i' P_i
\end{aligned}$$

- In general, S is self-adjoint, which implies S is normal, but not vice versa. However, the converse holds if further all eigenvalues of S are real numbers:

  By Theorem (12.1), we imply S is orthonormally diagonalizable, and its diagonal representation is of the form

  $$(S)_{\mathcal{B},\mathcal{B}} = \operatorname{diag}(\lambda_1, \dots, \lambda_k).$$

  Note that $\mathcal{B}$ is also a basis for S' and elements of $\mathcal{B}$ are eigenvalues of S', by part (3) in proposition (12.1). Therefore,

  $$(S')_{\mathcal{B},\mathcal{B}} = \operatorname{diag}(\lambda_1, \dots, \lambda_k).$$

  Therefore, $S = S'$.

In particular, for $S = P_i$, we can easily show all eigenvalues of P_i are 0 or 1, which are real. Therefore, P_i's are self-adjoint. ∎

Corollary 12.2 Let $T: V \to V$ be a linear operator on $\mathbb{C}$-inner product space with $\dim(V) < \infty$. Then T is normal if and only if $T' = f(T)$ for some polynomial $f(x) \in \mathbb{C}[x]$.

Proof.

- For the reverse direction, if $T' = f(T)$, then $T'T = f(T)T = Tf(T) = TT'$.
- For the forward direction, suppose that T is normal, then by corollary (12.1),

$$T = \sum_{i=1}^{k} \lambda_i P_i, \ P_i = f_i(T), \text{ where } P_i\text{'s are self-adjoint,}$$

which follows that

$$T' = \left(\sum_{i=1}^{k} \lambda_i P_i\right)' = \sum_{i=1}^{k} \bar{\lambda}_i P_i' = \sum_{i=1}^{k} \bar{\lambda}_i P_i = \sum_{i=1}^{k} \bar{\lambda}_i f_i(T)$$

∎

R The normal operator is a generalization of Hermitian matrices, and it inherits many nice properties of Hermitian.

12.1.2. Tensor Product

Motivation. Let U, V, W be vector spaces. We want to study bilinear maps $f: U \times W \to U$, i.e.,

$$f(av_1 + bv_2, w) = af(v_1, w) + bf(v_2, w)$$

$$f(v, cw_1 + dw_2) = cf(v, w_1) + df(v, w_2)$$

Unfortunately, bilinear form usually is not a linear transformation!

Example 12.1

- Let $f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be with $(u, v) \mapsto \langle u, v \rangle$.
- Let $f: M_{n\times n}(\mathbb{F}) \times M_{n\times n}(\mathbb{F}) \to M_{n\times n}(\mathbb{F})$ be with $f(A, B) = AB$.
- Let $f: \mathbb{F}[x] \times \mathbb{F}[x] \to \mathbb{F}$ be with $f(p(x), q(x)) = p(1)q(2)$

- Let $f : \mathbb{F}[x] \times \mathbb{F}[x] \longrightarrow \mathbb{F}[x]$ be with $f(p(x), q(x)) = p(x)q(x)$.

12.2. Monday for MAT3006

12.2.1. Remarks on MCT

■ **Example 12.2** The MCT can help us to compute the integral

$$\lim_{n\to\infty}\int_{[0,n\pi]}\cos\left(\frac{x}{2n}\right)xe^{-x^2}\,\mathrm{d}x$$

Construct $f_n(x) = \cos\left(\frac{x}{2n}\right)xe^{-x^2}\mathcal{X}_{[0,n\pi]}$.

- Since $\cos(x/2n) < \cos(x/2(n+1))$ for any $x \in [0, n\pi]$, we imply f_n is monotone increasing with n
- $f_n(x)$ is integrable for all n.
- f_n converges pointwise to $xe^{-x^2}\mathcal{X}_{[0,\infty)}$

Therefore, MCT I applies and

$$\lim_{n\to\infty}\int_{[0,n\pi]}\cos\left(\frac{x}{2n}\right)xe^{-x^2}\,\mathrm{d}x = \int\left(\lim_{n\to\infty}f_n\right)\mathrm{d}m$$

with

$$\lim_{n\to\infty}f_n = xe^{-x^2}\mathcal{X}_{[0,\infty)}.$$

Moreover,

$$\int\left(\lim_{n\to\infty}f_n\right)\mathrm{d}m = \lim_{m\to\infty}\int_{[0,m]}xe^{-x^2}\,\mathrm{d}x \tag{12.1a}$$

$$= \int_0^\infty xe^{-x^2}\,\mathrm{d}x \tag{12.1b}$$

$$= \frac{1}{2} \tag{12.1c}$$

where (12.1a) is by applying MCT I with $g_m(x) = xe^{-x^2}\mathcal{X}_{[0,m]}$ and proposition (10.14) to compute a Lebesgue integral by evaluating a proper Riemann integral. ■

Then we discuss the Lebesgue integral for series:

Corollary 12.3 [Lebesgue Series Theorem] Let $\{f_n\}$ be a series of measurable functions such that

$$\sum_{n=1}^{\infty}\int |f_n|\,\mathrm{d}m < \infty,$$

then $\sum_{n=1}^{k} f_n$ converges to an integrable function $f = \sum_{n=1}^{\infty} f_n$ a.e., with

$$\int f\,\mathrm{d}m = \sum_{n=1}^{\infty}\int f_n\,\mathrm{d}m$$

Proof. • For each f_n, consider

$$f_n = f_n^+ - f_n^-, \text{ where } f_n^+, f_n^- \text{ are nonnegative.}$$

By proposition (11.6),

$$\int \sum_{n=1}^{\infty} f_n^+\,\mathrm{d}m = \sum_{n=1}^{\infty}\int f_n^+\,\mathrm{d}m \le \sum_{n=1}^{\infty}\int |f_n|\,\mathrm{d}m < \infty.$$

Therefore, $f^+ := \sum_{n=1}^{\infty} f_n^+ = \lim_{k\to\infty}\sum_{n=1}^{k} f_n^+$ is integrable. The same follows by replacing f^+ with f^-. By corollary (9.6), $f^+(x), f^-(x) < \infty, \forall x \in U$, where U^c is null.

• Therefore, construct

$$f(x) = \begin{cases} f^+(x) - f^-(x), & x \in U \\ 0, & x \in U^c \end{cases}$$

Moreover, for $x \in U$,

$$\begin{aligned}
f(x) &= \left(\lim_{k\to\infty}\sum_{n=1}^{k} f_n^+(x)\right) - \left(\lim_{k\to\infty}\sum_{n=1}^{k} f_n^-(x)\right) \\
&= \lim_{k\to\infty}\left(\sum_{n=1}^{k} f_n^+(x) - \sum_{n=1}^{k} f_n^-(x)\right) \\
&= \lim_{k\to\infty}\left[\sum_{n=1}^{k}(f_n^+(x) - f_n^-(x))\right] \\
&= \sum_{n=1}^{\infty} f_n(x)
\end{aligned}$$

where the first equality is because that both terms are finite.

- It follows that

$$\int f\,\mathrm{d}m = \int f^+\,\mathrm{d}m - \int f^-\,\mathrm{d}m \tag{12.2a}$$

$$= \int \sum_{n=1}^{\infty} f_n^+\,\mathrm{d}m - \int \sum_{n=1}^{\infty} f_n^-\,\mathrm{d}m \tag{12.2b}$$

$$= \left(\sum_{n=1}^{\infty}\int f_n^+\,\mathrm{d}m\right) - \left(\sum_{n=1}^{\infty}\int f_n^-\,\mathrm{d}m\right) \tag{12.2c}$$

$$= \sum_{n=1}^{\infty}\left(\int f_n^+\,\mathrm{d}m - \int f_n^-\,\mathrm{d}m\right) \tag{12.2d}$$

$$= \sum_{n=1}^{\infty}\int f_n\,\mathrm{d}m \tag{12.2e}$$

where (12.2a),(12.2d) is because that summation/subtraction between series holds when these series are finite; (12.2c) is by proposition (11.6); (12.2e) is by definition of f_n.

■

■ **Example 12.3** Compute the integral

$$\int_{(0,1]} e^{-x}x^{\alpha-1}\,\mathrm{d}x,\ \alpha > 0.$$

- Construct $f_n(x) = (-1)^n \frac{x^{\alpha+n-1}}{n!} \mathcal{X}_{(0,1]}, n \ge 0,$ and

$$\sum_{n=0}^{N} f_n(x) \to e^{-x} x^{\alpha-1}, \text{ pointwisely}, x \in (0,1].$$

By applying MCT I,

$$\int |f_n| \, \mathrm{d}m = \frac{1}{(\alpha+n)n!}$$

Therefore,

$$\sum_{n=0}^{\infty} \int |f_n| \, \mathrm{d}m = \sum_{n=0}^{\infty} \frac{1}{(\alpha+n)n!} < \infty$$

- Applying the Lebesgue Series Theorem,

$$\int_{(0,1]} e^{-x} x^{\alpha-1} \, \mathrm{d}x = \int_{(0,1]} (\sum_{n=0}^{\infty} f_n) \, \mathrm{d}m = \sum_{n=0}^{\infty} \int f_n \, \mathrm{d}m = \sum_{n=0}^{\infty} \frac{(-1)^n}{(\alpha+n)n!}$$

■

R It's essential to have $\sum \int |f| \, \mathrm{d}m < \infty$ rather than $\sum \int f_n \, \mathrm{d}m < \infty$ in the Lebesgue Series Theorem. For example, let

$$f_n = \frac{(-1)^{n+1}}{(n+1)} \mathcal{X}_{[n,n+1)} \implies \sum_{n=1}^{\infty} \int f_n \, \mathrm{d}m = \log(2) < \infty$$

However, $f := \sum f_n$ is not integrable.

12.2.2. Dominated Convergence Theorem

Theorem 12.2 Let $\{f_n\}$ be a sequence of measruable functions such that $|f_n| \le g$ a.e., and g is integrable. Suppose that $\lim_{n\to\infty} f_n(x) = f(x)$ a.e., then

1. f is integrable,
2. $$\int f \, \mathrm{d}m = \lim_{n\to\infty} \int f_n \, \mathrm{d}m$$

Proof. • Observe that

$$|f_n| \le g \implies \lim_{n\to\infty} |f_n| \le g \implies |f| \le g$$

By comparison test, g is integrable implies $|f|$ is integrable, and further f is integrable.

• Consider the sequence of non-negative functions $\{g - f_n\}_{n\in\mathbb{N}}$ and $\{g + f_n\}_{n\in\mathbb{N}}$.

By Fatou's Lemma,

$$\begin{aligned}
\liminf_{n\to\infty} \int (g - f_n)\,\mathrm{d}m &\ge \int \liminf_{n\to\infty} (g - f_n)\,\mathrm{d}m \\
&= \int (g - f)\,\mathrm{d}m \\
&= \int g\,\mathrm{d}m - \int f\,\mathrm{d}m
\end{aligned}$$

which follows that

$$\int g\,\mathrm{d}m - \limsup_{n\to\infty} \int f_n\,\mathrm{d}m \ge \int g\,\mathrm{d}m - \int f\,\mathrm{d}m$$

i.e.,

$$\int f\,\mathrm{d}m \ge \limsup_{n\to\infty} \int f_n\,\mathrm{d}m$$

• Similarly,

$$\liminf_{n\to\infty} (g + f_n)\,\mathrm{d}m \ge \int \liminf_{n\to\infty} (g + f_n)\,\mathrm{d}m = \int g\,\mathrm{d}m + \int f\,\mathrm{d}m$$

which implies

$$\liminf_{n\to\infty} \int f_n\,\mathrm{d}m \ge \int f\,\mathrm{d}m$$

As a result,

$$\limsup_{n\to\infty} \int f_n\,\mathrm{d}m \le \int f\,\mathrm{d}m \le \liminf_{n\to\infty} \int f_n\,\mathrm{d}m,$$

which implies

$$\int f \, dm = \lim_n \int f_n \, dm$$

■

Corollary 12.4 [Bounded Convergence Theorem] Suppose that $E \in \mathcal{M}$ be such that $m(E) < \infty$. If

- $|f_n(x)| \le K < \infty$ for any $x \in E, n \in \mathbb{N}$
- $f_n \to f$ a.e. in E,

then f is integrable in E with

$$\int_E f \, dm = \lim_{n\to\infty} \int f_n \, dm$$

Proof. Take $g = K\mathcal{X}_E$ in DCT. ■

Proposition 12.2 Every Riemann integrable function f on $[a, b]$ is Lebesgue integrable, without the condition that f is continuous a.e.

Proof. Since f is Riemann integrable, we imply f is bounded. We construct the Riemann lower abd upper functions with 2^n equal intervals, denoted as $\{\phi_n\}$ and $\{\psi_n\}$, which follows that

- ϕ_n is monotone increasing; ψ_n is monotone decreasing;
- $\phi_n \le f \le \psi_n$, and

$$\lim_{n\to\infty} \int_{[a,b]} \phi_n = \int_a^b f(x) \, dx = \lim_{n\to\infty} \int_{[a,b]} \psi_n.$$

Construct $g = \sup_n \phi_n$ and $h = \inf_n \psi_n$. Now we can apply the bounded convergence theorem:

- ϕ_n is bounded on $[a, b]$
- $\phi_n \to g$ on $[a, b]$

which implies g is Lebesgue integrable on $[a,b]$, with

$$\int_{[a,b]} g \, \mathrm{d}m = \lim_{n\to\infty} \int_{[a,b]} \phi_n \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x.$$

Similarly, h is Lebesgue integrable, with

$$\int_{[a,b]} h \, \mathrm{d}m = \lim_{n\to\infty} \int_{[a,b]} \psi_n \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x.$$

Moreover, $g \le f \le h$, and

$$\int_{[a,b]} (h-g) \, \mathrm{d}m = \int_{[a,b]} h \, \mathrm{d}m - \int_{[a,b]} g \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x - \int_a^b f(x) \, \mathrm{d}x = 0,$$

which implies $h = g$ a.e., and further $f = g$ a.e., which implies

$$\int_{[a,b]} f \, \mathrm{d}m = \int_{[a,b]} g \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x.$$

∎

R However, an improper Riemann integral does not necessarily has the corresponding Lebesgue integral:

$$f(x) = \sum_{n=1}^{\infty} (-1)^n n \cdot \mathcal{X}_{(1/(n+1),1/n]}, \ x \in [0,1]$$

In this case, f is Riemann integrable but not Lebesgue integrable.

12.3. Monday for MAT4002

Proposition 12.3 If b, b' are path connected in X, then

$$\pi_1(X,b) \cong \pi_1(X,b')$$

R Last lecture we have given the isomorphism

$$\begin{aligned} W_\#: \quad & \pi_1(X,b) \to \pi_1(X,b') \\ \text{with} \quad & [\ell] \mapsto [w^{-1} \cdot \ell \cdot w] \end{aligned}$$

where w denotes a path from b to b'. The inverse of $W_\#$ is given by:

$$\begin{aligned} W_\#^{-1}: \quad & \pi_1(X,b') \to \pi_1(X,b) \\ \text{with} \quad & [m] \mapsto [w \cdot m \cdot w^{-1}] \end{aligned}$$

Notation. For path connected space X, we will just write $\pi_1(X)$ instead of $\pi_1(X,x)$.

Proposition 12.4 Let (X,x) and (Y,y) be spaces with basepoints x and y, and $f: X \to Y$ be a continuous map with $f(x) = y$. Then every loop $\ell: I \to X$ based at x gives a loop $f \circ \ell: I \to Y$ based at y, i.e., the continous map f induces a homomorphism of groups

$$\begin{aligned} f_*: \quad & \pi_1(\pi,x) \to \pi_1(Y,y) \\ & [\ell] \mapsto [f \circ \ell] := f_*([\ell]) \end{aligned}$$

Moreover,

1. $(\mathrm{id}_{X\to X})_* = \mathrm{id}_{\pi_1(X,x)\to\pi_1(X,x)}$
2. $(g \circ f)_* = g_* \circ f_*$
3. If $f \simeq f'$ relative to $x \in X$, then $f_* = (f')_*$

Proof. • Well-definedness: Suppose that $\ell \simeq \ell'$, then $f \circ \ell \simeq f \circ \ell'$ by propositon (9.4). Therefore, $[f \circ \ell] = [f \circ \ell']$.

- Homomorphism: It's clear that

$$f \circ (\ell \circ \ell') = (f \circ \ell) \circ (f \circ \ell')$$

  Therefore, $f_*[\ell \ell'] = (f_*[\ell]) * (f_*[\ell'])$

The other three statements are obvious.

■

Proposition 12.5 Let X, Y be path-connected such that $X \simeq Y$ (i.e., there exists $f : X \longrightarrow Y$ and $g : Y \longrightarrow X$ such that $g \circ f \simeq \mathrm{id}_X, f \circ g \simeq \mathrm{id}_Y$). Then $\pi_1(X) \cong \pi_1(Y)$.

In particular, if X, Y are path-connected with $X \cong Y$, then $\pi_1(X) \cong \pi_1(Y)$

Proof. Consider the mapping

$$\pi_1(X, x_0) \xrightarrow{f_*} \pi_1(Y, y_0) \xrightarrow{g_*} \pi_1(X, x_1)$$

It suffices to show that f_* and g_* are bijective. (The homomorphism follows from proposition (12.4))

- Wrong proof: $g \circ f \simeq \mathrm{id}_X$ implies $(g \circ f)_* = (\mathrm{id}_X)_*$ implies $g_* \circ f_* = \mathrm{id}_{\pi_1(X, x_0)}$.

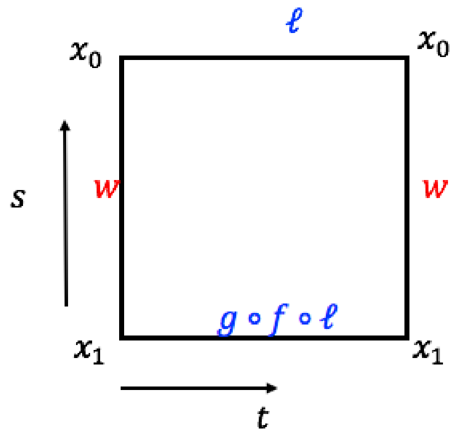
  Reason: note that $(g \circ f) \simeq \mathrm{id}_X$ is **not** relative to x_0.

Consider the homotopy $H : g \circ f \simeq \mathrm{id}_X$, where $H(x_0, s)$ is not necessarily a constant for $s \in I$. It follows that $H(x_0, 0) = x_1$ and $H(x_0, 1) = x_0$, i.e., $w(s) := H(x_0, s)$ defines a path from x_1 to x_0.

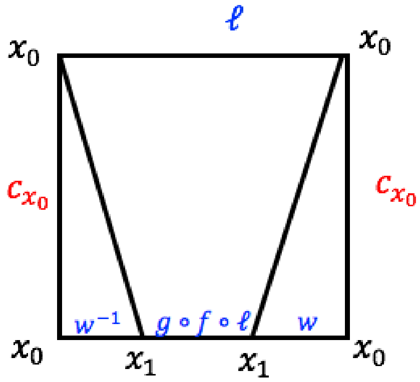
For any loop $\ell : I \longrightarrow X$ based at x_0, consider the homotopy

$$\begin{aligned} K = H \circ (\ell \times \mathrm{id}_I) : \quad & I \times I \longrightarrow X \\ \text{where} \quad & K(t, s) = H((\ell(t), s)) \\ & K(t, 0) = H(\ell(t), 0) = g \circ f(\ell(t)) \\ & K(t, 1) = H(\ell(t), 1) = \ell(t) \\ & K(0, s) = w(s) = K(1, s) \end{aligned}$$

The graphic plot of K is given in the figure below:



The homotopy between ℓ and $g \circ f \circ \ell$ motivates us to construct a homotopy between ℓ and $w^{-1} \circ g \circ f \circ \ell \circ w$ relative to $\{0,1\}$:



Therefore,

$$[\ell] = [w^{-1} g f \ell w] = W_\#([g f \ell]) = (W_\# \circ g_* \circ f_*)[\ell]$$

which follows that $W_\# \circ g_* \circ f_* = \mathrm{id}_{\pi_1(X,x_0)}$. Therfore, f_* is injective, g_* is surjective.

The similar argument gives

$$W_\# \circ f_* \circ g_* = \mathrm{id}_{\pi_1(Y,y_0)}$$

Therefore, f_* is surjective, g_* is injective. The bijectivity is shown.

■

Definition 12.1 [Simply-Connected] A space X is **simply-connected** if X is path connected, and X has trivial fundamental group, i.e., $\pi_1(X) = \{e\}$ for some point $e \in X$.

Example 12.4 If X is contractible, then X is path-connected. By proposition (12.5), since $X \simeq \{e\}$, we imply

$$\pi_1(X) \cong \pi_1(\{e\}) = \{e\}.$$

Therefore, all contractible spaces (e.g., $\mathbb{R}^n$) are simply-connected.

However, not all simply-connected spaces are contractible, e.g., $\pi_1(S^2) \cong \{e\}$, but S^2 is not homotopy equivalent to a point.

12.3.1. Some basic results on $\pi_1(X,b)$

We will study $\pi_1(X,b)$ for some simplicial complexes.

Definition 12.2 [Edge Loop] Let $K = (V, \Sigma)$ be a simplicial complex.

1. An edge path $(v_0, \dots, v_n)$ is such that
   (a) $a_i \in V(K)$
   (b) For each i, $\{a_{i-1}, a_i\}$ spans a simplex of K
2. An edge loop is an edge path with $a_n = a_0$.
3. Let $\alpha = (v_0, \dots, v_n), \beta = (w_0, \dots, w_m)$ be two edge paths with $v_n = w_0$, then we define

$$\alpha \circ \beta = (v_0, \dots, v_n, w_1, \dots, w_m)$$

Definition 12.3 [Elementary Contraction/Expansion] Let α, β be two edge paths.

1. An elementary contraction of α is a new edge path obtained by performing one of the followings on α:

- Replacing $\cdots a_{i-1}a_i\cdots$ by $\cdots a_{i-1}\cdots$ provided that $a_{i-1}=a_i$
- Replacing $\cdots a_{i-1}a_ia_{i+1}\cdots$ by $\cdots a_{i-1}\cdots$ provided that $a_{i-1}=a_{i+1}$
- Replacing $\cdots a_{i-1}a_ia_{i+1}\cdots$ by $\cdots a_{i-1}a_{i+1}\cdots$ provided that $\{a_{i-1},a_i,a_{i+1}\}$ spans a 2-simplex of K.

2. An elementary expansion is the reverse of the elementary contraction.
3. Two edge paths α,β are equivalent if α and β differs by a finite sequence of elementary contractions or expansions.

12.4. Wednesday for MAT3040

12.4.1. Introduction to Tensor Product

Reviewing. Bilinear map: $f : V \times W \to U$, e.g.,

$$\begin{aligned} f: &\quad \mathbb{R}^3 \times \mathbb{R}^3 \\ \text{with} &\quad f(u,v) = u \times v \end{aligned}$$

Note that f is usually not a linear transformation, e.g.,

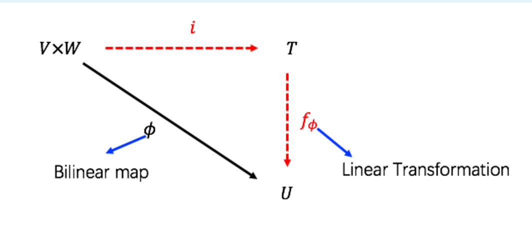
$$f(3(\boldsymbol{v},\boldsymbol{w})) = f(3\boldsymbol{v},3\boldsymbol{w}) = (3\boldsymbol{v}) \times (3\boldsymbol{w}) = 9\boldsymbol{v} \times \boldsymbol{w} \neq 3f(\boldsymbol{v},\boldsymbol{w}).$$

The vector space structure of $V \times W$ is not suited to study bilinear map, and the proper way is to study its induced linear transformation.

Definition 12.4 [Universal Property of Tensor Product] Let V, W be vector spaces. Consider the set

$$\text{Obj} := \{\phi : V \times W \to U \mid \phi \text{ is a bilinear map}\}$$

We say T, or $(i : V \times W \to T) \in \text{Obj}$ satisfies the **universal property** if for any $(\phi : V \times W \to T) \in \text{Obj}$, there exists an unique linear transformation $f_\phi : T \to U$ such that the diagram below commutes:



i.e., $\phi = f_\phi \circ i$. ■

Therefore, rather than studying bilinear map ϕ, it is better to study the linear transformation f_ϕ instead.

Question: does T exist?

Definition 12.5 [Spanning Set] Let V, W be vector spaces. Let $S = \{(\boldsymbol{v}, \boldsymbol{w}) \mid \boldsymbol{v} \in V, \boldsymbol{w} \in W\}$, then we define

$$\mathfrak{X} = \operatorname{span}(S).$$

R

1. The spanning set $\mathfrak{X}$ is not addictive, e.g., $\mathfrak{x}_1 = 3(0, \boldsymbol{w}) \in \mathfrak{X}$ and $\mathfrak{x}_2 = 1(0, \boldsymbol{w}) + 1(0, 2\boldsymbol{w}) \in \mathfrak{X}$, but $\mathfrak{x}_1 \neq \mathfrak{x}_2$.
2. Note that we assume no relations on the elements $(\boldsymbol{v}, \boldsymbol{w}) \in \mathcal{S}$. More precisely, the set $S = \{(\boldsymbol{v}, \boldsymbol{w}) \mid \boldsymbol{v} \in V, \boldsymbol{w} \in W\}$ is linearly independent in $\mathfrak{X}$. For example, $(0, \boldsymbol{w}) \perp (0, 2\boldsymbol{w})$.
3. The only legitimate relationship is

$$2(\boldsymbol{v}_1, \boldsymbol{w}_1) + 3(\boldsymbol{v}_1, \boldsymbol{w}_1) = 5(\boldsymbol{v}, \boldsymbol{w}),$$

   which is not equal to $(5\boldsymbol{v}, 5\boldsymbol{w})$
4. $\mathcal{S}$ is a basis of $\mathfrak{X}$, and therefore $\mathcal{X}$ is of uncountable dimension.

Definition 12.6 [Special subspace of $\mathfrak{X}$] Let $y \leq \mathfrak{X}$ be a vector subspace spanned by vectors of the form

$$\{1(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}) - 1(\boldsymbol{v}_1, \boldsymbol{w}) - 1(\boldsymbol{v}_2, \boldsymbol{w})\}, \quad \text{and} \quad \{1(\boldsymbol{v}, \boldsymbol{w}_1 + \boldsymbol{w}_2) - 1(\boldsymbol{v}, \boldsymbol{w}_1) - 1(\boldsymbol{v}, \boldsymbol{w}_2)\}$$

and

$$\{1(k\boldsymbol{v}, \boldsymbol{w}) - k(\boldsymbol{v}, \boldsymbol{w}) \mid k \in \mathbb{F}\}$$

and

$$\{1(\boldsymbol{v}, k\boldsymbol{w}) - k(\boldsymbol{v}, \boldsymbol{w}) \mid k \in \mathbb{F}\}$$

Definition 12.7 [Tensor Product] We define the **tensor product** $V \otimes W$ by

$$V \otimes W = \mathcal{X}/y.$$

Therefore, $\boldsymbol{v} \otimes \boldsymbol{w} = (\boldsymbol{v}, \boldsymbol{w}) + y \in \mathcal{X}/y$

R

1. As a result, the tensor product is finitely addictive:

$$\begin{aligned}(\boldsymbol{v}_1 + \boldsymbol{v}_2) \otimes \boldsymbol{w} &= (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) + y \\ &= (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) - [(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) - (\boldsymbol{v}_1, \boldsymbol{w}) - (\boldsymbol{v}_2, \boldsymbol{w})] + y \\ &= 0(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) + (\boldsymbol{v}_1, \boldsymbol{w}) + (\boldsymbol{v}_2, \boldsymbol{w}) + y \\ &= [(\boldsymbol{v}_1, \boldsymbol{w}) + y] + [(\boldsymbol{v}_2, \boldsymbol{w}) + y] \\ &= \boldsymbol{v}_1 \otimes \boldsymbol{w} + \boldsymbol{v}_2 \otimes \boldsymbol{w}\end{aligned}$$

   Similarly,

$$\begin{aligned}\boldsymbol{v} \otimes (\boldsymbol{w}_1 + \boldsymbol{w}_2) &= (\boldsymbol{v} \otimes \boldsymbol{w}_1) + (\boldsymbol{v} \otimes \boldsymbol{w}_2) \\ (k\boldsymbol{v}) \otimes \boldsymbol{w} &= k(\boldsymbol{v} \otimes \boldsymbol{w}) \\ \boldsymbol{v} \otimes (k\boldsymbol{w}) &= k(\boldsymbol{v} \otimes \boldsymbol{w})\end{aligned}$$

2. The product space $V \times W$ is different from the tensor product space $V \otimes W$:

   (a) $(\boldsymbol{v}, \mathbf{0}) \neq \mathbf{0}_{V \times W}$ in $V \times W$; but $\boldsymbol{v} \otimes 0 \in 0_{V \otimes W}$:

$$\begin{aligned}V \otimes 0 &= V \otimes (0\boldsymbol{w}) \\ &= 0(V \otimes w) \\ &= 0_{V \otimes W}\end{aligned}$$

   Moreover, f is bilinear implies $f(\boldsymbol{v}, 0) = \mathbf{0}$.

(b) $(\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2, \boldsymbol{w}_2) = (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}_1 + \boldsymbol{w}_2)$; but $\boldsymbol{v}_1 \otimes \boldsymbol{w}_1 + \boldsymbol{v}_2 \otimes \boldsymbol{w}_2$ cannot be simplified further, unless $\boldsymbol{v}_1 = \boldsymbol{v}_2$:

$$\boldsymbol{v} \otimes \boldsymbol{w}_1 + \boldsymbol{v} \otimes \boldsymbol{w}_2 = \boldsymbol{v} \otimes (\boldsymbol{w}_1 + \boldsymbol{w}_2)$$

Theorem 12.3 The bilinear map

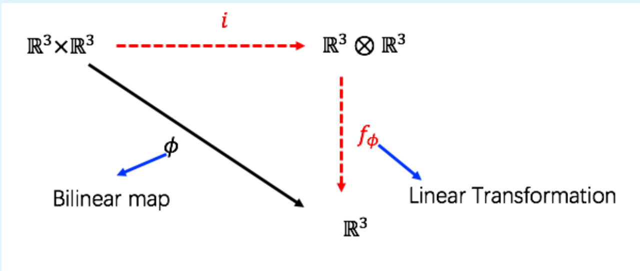
$$\begin{aligned} i: \quad & V \times W \to V \otimes W \quad (i \in \mathrm{Obj}) \\ \text{with} \quad & (\boldsymbol{v}, \boldsymbol{w}) \mapsto \boldsymbol{v} \otimes \boldsymbol{w} \end{aligned}$$

satisfies the universal property of tensor products.

■ **Example 12.5** Consider a common bilinear map

$$\begin{aligned} \phi: \quad & \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \\ \text{with} \quad & (\boldsymbol{v}, \boldsymbol{w}) \mapsto \boldsymbol{v} \times \boldsymbol{w} \end{aligned}$$

By the universal property, there exists the linear transformation $f_\phi : \mathbb{R}^3 \otimes \mathbb{R}^3 \to \mathbb{R}^3$ such that the diagram below commutes:



■

12.5. Wednesday for MAT3006

12.5.1. Riemann Integration & Lebesgue Integration

■ **Example 12.6** Compute the integral

$$L = \lim_{n\to\infty} \int_0^1 \frac{nx\log(x)}{1+n^2x^2}\,\mathrm{d}x.$$

Let $f_n(x) = \frac{nx\log(x)}{1+n^2x^2}\mathcal{X}_{(0,1]}$, which is continuous on $[0,1]$, i.e., integrable on $[0,1]$. The goal is to show $L = 0$.

- Note that $f_n(x) \to 0, \forall x \in [0,1]$ pointwisely, as $n \to \infty$.
- Note that $t/(1+t^2) \le \frac{1}{2}, \forall t \ge 0$. Take $t = nx$, we imply

$$|f_n(x)| \le \frac{1}{2}|\log(x)|\mathcal{X}_{(0,1]}$$

  We claim that $\frac{1}{2}|\log(x)|\mathcal{X}_{(0,1]} := -\frac{1}{2}\log(x)\mathcal{X}_{(0,1]}$ is integrable: by MCT I,

$$\int -\frac{1}{2}\log(x)\mathcal{X}_{(0,1]}\,\mathrm{d}m = \lim_{n\to\infty}\int_{1/n}^1 -\frac{1}{2}\log(x)\,\mathrm{d}x = \frac{1}{2} < \infty.$$

Therefore, the DCT applies, and

$$\lim_{n\to\infty}\int_{(0,1]} \frac{nx\log(x)}{1+n^2x^2}\,\mathrm{d}x = \int_{(0,1]} \lim_{n\to\infty} \frac{nx\log(x)}{1+n^2x^2}\,\mathrm{d}x = \int_{(0,1]} 0\,\mathrm{d}x = 0$$

However, $f_n(x)$ does not converge to $f(x) \equiv 0$ uniformly on $[0,1]$:

$$\sup_{0\le x\le 1} |f_n(x) - 0| \ge |f_n(1/n) - 0| = \frac{1}{2}\log(n) \to \infty, \text{as } n \to \infty$$

Therefore, we cannot switch integral symbol and limit by using the tools in MAT2006. ■

Proposition 12.6 Suppose that $f(x)$ is a proper Riemann integrable function on $[a,b]$.

Then $f(x)$ is Lebesgue integrable on $[a, b]$ with

$$\int_{[a,b]} f \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x.$$

Proof. Since f is properly Riemann inregrable, we imply $f(x)$ is bounded on $[a, b]$, i.e., $|f(x)| \le K, \forall x \in [a, b]$. Construct the Riemann lower and upper functions with 2^n equal subintervals, denoted as ϕ_n, ψ_n, which follows that

- $\phi_n(x) \le f(x) \le \psi_n(x), \forall n$
- $\phi_n(x)$ is monotone increasing
- $\psi_n(x)$ is monotone decreasing

Now apply bounded convergence theorem on $\psi_n - \phi_n$:

- $|\psi_n(x) - \phi_n(x)| \le 2K$ on $[a, b]$
- $\psi_n - \phi_n \to \psi - \phi$

which implies

$$\begin{aligned}
\int |\psi - \phi| \, \mathrm{d}m &= \int \psi - \phi \, \mathrm{d}m \\
&= \lim_{n\to\infty} \int \psi_n - \phi_n \, \mathrm{d}m = \lim_{n\to\infty} \int \psi_n \, \mathrm{d}m - \lim_{n\to\infty} \int \phi_n \, \mathrm{d}m \\
&= \text{Riemann Upper Sum} - \text{Riemann Lower Sum} \\
&= 0
\end{aligned}$$

Therefore, $\int |\psi - \phi| \, \mathrm{d}m = 0$ implies $\psi(x) = \phi(x)$ a.e. By sandwich theorem,

$$\psi(x) = f(x) = \phi(x) \text{ a.e.}$$

Therefore,

$$\int f \, \mathrm{d}m = \int \phi \, \mathrm{d}m = \lim_{n\to\infty} \int \phi_n \, \mathrm{d}m = \int_a^b f(x) \, \mathrm{d}x$$

where the second equality is by MCT II.

∎

R The improper Riemann integrable functions $f(x)$ is not necessarily Lebesgue integrable. However, if we assume $f(x) \geq 0$, then $f(x)$ is improper Riemann integrable implies $f(x)$ is Lebesgue integrable, with the same integral value.

Proof Outline. Suppose $f(x)$ is improper Riemann integrable on $[a,b]$, where $a,b \in \mathbb{R} \cup \{\pm\infty\}$.

- Construct $f_n = f\mathcal{X}_{[a_n,b_n]}$, with $[a_n,b_n] \subseteq [a_{n+1},b_{n+1}] \subseteq \cdots \subseteq [a,b]$.
- By previous proposition, f_n is proper Riemann integrable implies f_n is Lebesgue integrable.
- Then we apply the MCT I to $\{f_n\}$.

$\blacksquare$

12.5.2. Continuous Parameter DCT

Theorem 12.4 — **Continous parameter DCT.** Let $I, J \subseteq \mathbb{R}$ be intervals, and $f : I \times J \to \mathbb{R}$ be such that

1. for fixed $y \in J$, the function $f(x) := f(x,y)$ is an integrable function over I.
2. for fixed $y \in J$,
$$\lim_{y' \to y} f(x,y') = f(x,y)$$
for almost all $x \in I$
3. There exists integrable $g(x)$ (do not depend on y) such that for all $y \in J$,
$$|f(x,y)| \leq g(x)$$
for almost all $x \in I$.

As a result,
$$F(y) = \int_I f(x,y)\,\mathrm{d}x$$
is a continuous function on J.

R Note that the integrability of $f(x)$ in hypothesis (1) can be weaken into the measurability of $f(x)$: The measurability of $f(x)$ together with hypothesis (3), and DCT implies the integrability of $f(x)$.

Proof. Let $\{y_n\}$ be a sequence on J such that $y_n \to y$. It suffices to show $F(y_n) \to F(y)$.

Construct $f_n(x) = f(x, y_n)$, which follows that

- $f_n(x)$ is integrable for all n (by hypothesis (1)) (why check integrable)
- $|f_n(x)| \le g(x)$ a.e. for all n, and $g(x)$ is integrable (by hypothesis (3))
- By hypothesis (2),

$$\lim_{n\to\infty} f_n(x) = f(x, y)$$

Therefore, the DCT applies, and

$$\lim_{n\to\infty} \int_I f_n(x, y_n)\,\mathrm{d}m = \int \lim_{n\to\infty} f_n(x, y_n)\,\mathrm{d}m = \int_I f(x, y)\,\mathrm{d}m$$

Or equivalently,

$$\lim_{n\to\infty} F(y_n) = F(y)$$

∎

■ **Example 12.7** Consider $f(x,y) = e^{-x}x^{y-1}$ with $I \times J = (0,\infty) \times [m, M]$, where $0 < m < M < \infty$. We will study the integral

$$\Gamma(y) = \int_0^\infty e^{-x}x^{y-1}\,\mathrm{d}x$$

We check the hypothesis in the Theorem (12.4):

1. For fixed $y \in [m, M]$, $f(x) := f(x,y)$ is indeed measurable on $(0,\infty)$, since $f(x)$ is continous on $(0,\infty)$.
2. The hypothesis (2) follows directly from the contiuity of $f(x,y)$

3.

$$\begin{aligned}|f(x,y)| &\le e^{-x}x^{m-1}\mathcal{X}_{[0,1]} + e^{-x}x^{M-1}\mathcal{X}_{(1,\infty)}\\ &\le x^{m-1}\mathcal{X}_{[0,1]} + e^{-x}x^{M-1}\mathcal{X}_{(1,\infty)}\end{aligned}$$

Here $x^{m-1}\mathcal{X}_{[0,1]}$ is integrable. Following the similar argument in (1), we imply $e^{-x}x^{M-1}\mathcal{X}_{(1,\infty)}$ is integrable as well.

Therefore, $\Gamma(y)$ is continuous for any $m \le y \le M$. Since the choice of $0 < m < M < \infty$ is arbitrary, we imply $T(y)$ is continous on $(0,\infty)$.

In the next lecture we wish to show that

$$F'(y) = \int_I \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$$

■

12.6. Wednesday for MAT4002

Reviewing.

- Edge loop based at $b \in V$:

$$\alpha = (b, v_1, \cdots, v_n, b)$$

- Equivalence class of edge loops:

$$[\alpha] = \{\alpha' \mid \alpha' \sim \alpha, \alpha' \text{ is the edge loop based at } b\}$$

  Note that $\alpha' \sim \alpha$ if they differ from finitely many elementary contractions or expansions.

  For instance, let K in the figure below denote a triangle:

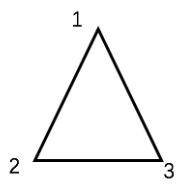


Figure 12.1: Triangle K

  Then the canonical form of any equivalence form $[\alpha]$ can be expressed as:

$$[\alpha] = [bcabc \cdots ab],$$

  where $a, b, c \in \{1,2,3\}$ are distinct.

12.6.1. Groups & Simplicial Complices

Proposition 12.7 The $E(K,b) = \{[\alpha] \mid \alpha \text{ is edge loop based at } b\}$ is a group, with the operation

$$[\alpha] * [\beta] = [\alpha \cdot \beta]$$

Proof. 1. Well-definedness of $*$:

$$\alpha \sim \alpha', \beta \sim \beta' \implies \alpha \cdot \beta \sim \alpha' \cdot \beta'$$

2. Associativity is clear.
3. The identity is $e := [b]$: for any edge loop $[\alpha] = [bv_1 \cdots b]$,

$$\begin{aligned}[\alpha] * e &= [bv_1 \cdots v_n b] * [b] \\ &= [bv_1 \cdots v_n bb] \\ &= [bv_1 \cdots v_n b] = [\alpha].\end{aligned}$$

   Also, $e * [\alpha] = [\alpha]$.
4. The inverse of any edge loop $[bv_1 \cdots v_n b]$ is $[bv_n \cdots v_1 b]$:

$$\begin{aligned}[bv_1 \cdots v_n b]^{-1} * [bv_1 \cdots v_n b] &= [bv_n \cdots v_1 bbv_1 \cdots v_n b] \\ &= [bv_n \cdots v_1 bv_1 \cdots v_n b] \\ &= [bv_n \cdots v_2 v_1 v_2 \cdots v_n b] \\ &= \cdots \\ &= [b]\end{aligned}$$

   Similarly, $[bv_1 \cdots v_n b] * [bv_1 \cdots v_n b]^{-1} = [b]$.

■

We will see that for K defined in Fig.(12.1), $E(K,1) \cong \mathbb{Z}$, in the next class.

Theorem 12.5 $E(K,b) \cong \pi_1(|K|,b)$.

This is the most difficult theorem that we have faced so far. Let's recall the simplicial approximation proposition first:

R

Proposition 12.8 — Simplicial Approximation Proposition. Suppose that $f : |K| \longrightarrow |L|$ be such that for all $v \in V(K)$, there exists $g(v) \in V(L)$ satisfying

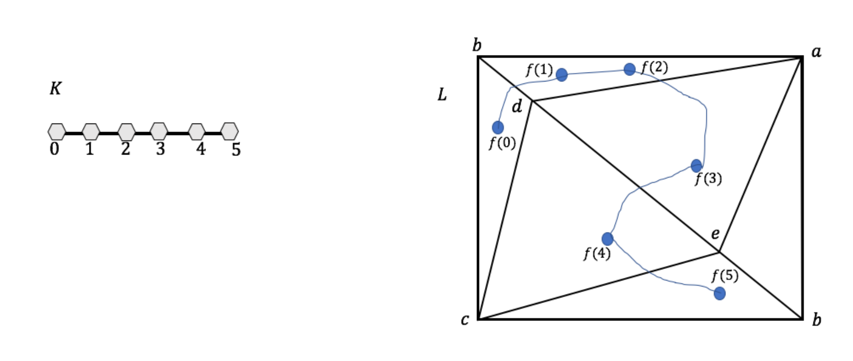
$$f(\mathrm{st}_K(v)) \subseteq \mathrm{st}_L(g(v)).$$

As a result,

1. the mapping
$$\begin{aligned} g : \quad & K \longrightarrow L \\ \text{with} \quad & v \mapsto g(v) \end{aligned}$$
is a simplicial map, i.e., for all $\sigma_K \in \Sigma_K, g(\sigma_K) \in \Sigma_L$
2. Moreover, $|g| \simeq f$.

Furthermore, if $A \subseteq K$ is a simplicial subcomplex such that $f(|A|) \subseteq |B|$, where $B \subseteq L$ is a simplicial subcomplex, then we can choose g such that $g\mid_A : A \longrightarrow B$ and the homotopy $|g| \simeq f$ sends $|A|$ to $|B|$.

■ **Example 12.8** Consider the simplicial complex K and L shown in the figure below:



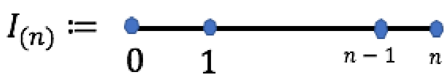
Let A_1 denote the subcomplex with $V(A_1) = \{0\}, \Sigma_{A_1} = \{\{0\}\}$, and A_2 denote the subcomplex wit $V(A_2) = \{1,2\}$ and $\Sigma_{A_2} = \{\{1,2\},\{1\},\{2\}\}$. Therefore,

$$f(|A_1|) \subseteq |\Delta_{\{b,c,d\}}|, \quad f(|A_2|) \subseteq |\Delta_{\{a,b,d\}}|,$$

There exists simplicial mapping g with

$$g(0) = b, \quad g(1) = b, \quad g(2) = d, \quad g(3) = e, \quad g(4) = c, \quad g(5) = c$$

Proof. 1. For each edge loop $\alpha = (v_0, \ldots, v_n)$ based at b, consider the simplicial complex



Together with the simplicial map

$$\begin{aligned} g_\alpha : \quad & I_{(n)} \to K \\ \text{with} \quad & g_\alpha(i) = v_i \end{aligned}$$

Note that it is well-defined since $\{i, i+1\} \in \Sigma_{I_{(n)}}$, and $\{v_i, v_{i+1}\} \in \Sigma_K$.

Now construct the mapping

$$\begin{aligned} \theta : \quad & \{\text{edge loop based at } b\} \to \pi_1(K, b) \\ \text{with} \quad & \alpha \mapsto [|g_\alpha|] \\ \text{where} \quad & |g_\alpha| : |I_{(n)}| (\cong [0,1]) \to |K| \\ & |g_\alpha|(i/n) = v_i \end{aligned}$$

For example,

$$\alpha = (bdeabcb), \implies |g_\alpha|(0) = b, |g_\alpha|(1/6) = d, |g_\alpha|(2/6) = e, \cdots, |g_\alpha|(1) = b,$$

i.e., $|g_\alpha|$ is a loop based at b.

Therefore, $[|g_\alpha|] \in \pi_1(|K|, b)$.

2. Now, suppose $\alpha \sim \alpha'$ be two edge loops differ by an elemenary contraction, e.g.,

$$\alpha' = (bdebcb) \sim \alpha = (bdeabcb).$$

As a result, $|g_{\alpha'}| \simeq |g_\alpha|$ relative to $\{0,1\}$, i.e., $[|g_\alpha|] = [|g_{\alpha'}|]$.

Therefore, we have a well-defined map:

$$\begin{aligned} \tilde{\theta} : &\quad \{\text{edge loops based at } b\}/\sim \longrightarrow \pi_1(|K|, b) \\ \text{with} &\quad [\alpha] \mapsto [|g_\alpha|] \end{aligned}$$

Therefore, $\tilde{\theta} : E(K,b) \to \pi_1(|K|,b)$ is the desired map.

3. $\tilde{\theta}$ is a homomorphism: it suffices to show that

$$\tilde{\theta}([\alpha] * [\beta]) = \tilde{\theta}([\alpha])\tilde{\theta}([\beta]),$$

which suffices to show $[|g_{\alpha\cdot\beta}|] = [|g_\alpha||g_\beta|]$, i.e., $|g_{\alpha\cdot\beta}| \simeq |g_\alpha||g_\beta|$. Note that $|g_{\alpha\cdot\beta}|$ and $|g_\alpha||g_\beta|$ are the same path with different "speed", i.e., homotopy.

4. The mapping $\tilde{\theta}$ is surjective: Let $\ell : [0,1] \to |K|$ be a loop based at b. It suffices to find an edge loop α such that $[|g_\alpha|] = [\ell]$, i.e., $|g_\alpha| \simeq \ell$.

   (a) Applying the simplicial approximation theorem, there exist large n and $g : I_{(n)} \to K$ such that $|g| \simeq \ell$. Here we can choose $g : I_{(n)} \to K$ to be such that $g(\{0\}) = \{b\}$, $g(\{n\}) = \{b\}$, and $|g| \simeq \ell$ relative to $\{0,1\}$.

   (b) Take $\alpha = (g(0), g(1), \dots, g(n))$ so that $g(0) = b = g(n)$, with $g_\alpha = g$. Therefore, $[|g_\alpha|] = [\ell]$, and hence $\tilde{\theta}$ is surjective.

■

Chapter 13

Week13

13.1. Monday for MAT3040

Reviewing.

1. Define $S = \{(\boldsymbol{v},\boldsymbol{w}) \mid \boldsymbol{v} \in V, \boldsymbol{w} \in W\}$ and $\mathfrak{X} = \text{span}(S)$. In $\mathfrak{X}$, there are no relations between distinct elements of S, e.g.,

$$2(\boldsymbol{v},0) + 3(0,\boldsymbol{w}) \neq 1(2\boldsymbol{v},3\boldsymbol{w})$$

   General element in $\mathfrak{X}$:

$$a_1(\boldsymbol{v}_1,\boldsymbol{w}_1) + \cdots + a_n(\boldsymbol{v}_n,\boldsymbol{w}_n),$$

   where $(\boldsymbol{v}_i,\boldsymbol{w}_i)$ are distinct.
2. Define the space $V \otimes W = \mathfrak{X}/y$, with

$$\boldsymbol{v} \otimes \boldsymbol{w} = 1(\boldsymbol{v},\boldsymbol{w}) + y \in V \otimes W.$$

   General element in $\mathfrak{X}/y := V \otimes W$:

$$\begin{aligned}
a_1(\boldsymbol{v}_1,\boldsymbol{w}_1) + \cdots + a_n(\boldsymbol{v}_n,\boldsymbol{w}_n) + y &= a_1((\boldsymbol{v}_1,\boldsymbol{w}_1) + y) + \cdots + a_n((\boldsymbol{v}_n,\boldsymbol{w}_n) + y) \\
&= a_1(\boldsymbol{v}_1 \otimes \boldsymbol{w}_1) + \cdots + a_n(\boldsymbol{v}_n \otimes \boldsymbol{w}_n) \\
&= (a_1\boldsymbol{v}_1) \otimes \boldsymbol{w}_1 + \cdots + (a_n\boldsymbol{v}_n) \otimes \boldsymbol{w}_n
\end{aligned}$$

Therefore, a general element in $V \otimes W$ is of the form

$$\boldsymbol{v}_1' \otimes \boldsymbol{w}_1 + \cdots + \boldsymbol{v}_n' \otimes \boldsymbol{w}_n, \ \boldsymbol{v}_i' \in V, \boldsymbol{w}_i \in W. \tag{13.1}$$

Note that $V \otimes W$ is different from $V \times W$, where all elements in $V \times W$ can be expressed as $(\boldsymbol{v}, \boldsymbol{w})$.

3. The tensor product mapping

$$\begin{aligned} i: \quad & V \times W \to V \otimes W \\ \text{with} \quad & (\boldsymbol{v}, \boldsymbol{w}) \mapsto \boldsymbol{v} \otimes \boldsymbol{w} \end{aligned}$$

satisfies the universal property.

Here we present an example for computing tensor product by making use of the rules below:

$$\begin{aligned} (\boldsymbol{v}_1 + \boldsymbol{v}_2) \otimes \boldsymbol{w} &= \boldsymbol{v}_1 \otimes \boldsymbol{w} + \boldsymbol{v}_2 \otimes \boldsymbol{w} \\ \boldsymbol{v} \otimes (\boldsymbol{w}_1 + \boldsymbol{w}_2) &= (\boldsymbol{v} \otimes \boldsymbol{w}_1) + (\boldsymbol{v} \otimes \boldsymbol{w}_2) \\ (k\boldsymbol{v}) \otimes \boldsymbol{w} &= k(\boldsymbol{v} \otimes \boldsymbol{w}) \\ \boldsymbol{v} \otimes (k\boldsymbol{w}) &= k(\boldsymbol{v} \otimes \boldsymbol{w}) \end{aligned}$$

▪ **Example 13.1** Let $V = W = \mathbb{R}^2$, with

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Here we have

$$\begin{aligned}
\begin{pmatrix}3\\1\end{pmatrix}\otimes\begin{pmatrix}-4\\2\end{pmatrix} &= (3\boldsymbol{e}_1+2\boldsymbol{e}_2)\otimes(-4\boldsymbol{e}_1+2\boldsymbol{e}_2)\\
&= (3\boldsymbol{e}_1)\otimes(-4\boldsymbol{e}_1+2\boldsymbol{e}_2)+(\boldsymbol{e}_2)\otimes(-4\boldsymbol{e}_1+2\boldsymbol{e}_2)\\
&= (3\boldsymbol{e}_1)\otimes(-4\boldsymbol{e}_1)+(3\boldsymbol{e}_1)\otimes(2\boldsymbol{e}_2)+(\boldsymbol{e}_2)\otimes(-4\boldsymbol{e}_1)+\boldsymbol{e}_2\otimes(2\boldsymbol{e}_2)\\
&= -12(\boldsymbol{e}_1\otimes\boldsymbol{e}_1)+6(\boldsymbol{e}_1\otimes\boldsymbol{e}_2)-4(\boldsymbol{e}_2\otimes\boldsymbol{e}_1)+2(\boldsymbol{e}_2\otimes\boldsymbol{e}_2)
\end{aligned}$$

Exercise: Check that $\boldsymbol{e}_1\otimes\boldsymbol{e}_2+\boldsymbol{e}_2\otimes\boldsymbol{e}_1$ cannot be re-written as

$$(a\boldsymbol{e}_1+b\boldsymbol{e}_2)\otimes(c\boldsymbol{e}_1+d\boldsymbol{e}_2),\quad a,b,c,d\in\mathbb{R}.$$

13.1.1. Basis of $V\otimes W$

Motivation. Given that $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\}$ is a basis of V, and $\{\boldsymbol{w}_1,\dots,\boldsymbol{w}_m\}$ a basis of W, we aim to find a basis of $V\otimes W$ using $\boldsymbol{v}_i$'s and $\boldsymbol{w}_i$'s.

Proposition 13.1 The set $\{\boldsymbol{v}_i\otimes\boldsymbol{w}_j \mid 1\le i\le n, 1\le j\le m\}$ spans the tensor product space $V\otimes W$.

Proof. Consider any $\boldsymbol{v}\in V$ and $\boldsymbol{w}\in W$, and we want to express $\boldsymbol{v}\otimes\boldsymbol{w}$ in terms of $\boldsymbol{v}_i,\boldsymbol{w}_j$. Suppose that $\boldsymbol{v}=\alpha_1\boldsymbol{v}_1+\cdots+\alpha_n\boldsymbol{v}_n$ and $\boldsymbol{w}=\beta_1\boldsymbol{w}_1+\cdots+\beta_m\boldsymbol{w}_m$.

Substituting $\boldsymbol{v}=\alpha_1\boldsymbol{v}_1+\cdots+\alpha_n\boldsymbol{v}_n$ into the expression $\boldsymbol{v}\otimes\boldsymbol{w}$, we imply

$$\begin{aligned}
\boldsymbol{v}\otimes\boldsymbol{w} &= (\alpha_1\boldsymbol{v}_1+\cdots+\alpha_n\boldsymbol{v}_n)\otimes\boldsymbol{w}\\
&= (\alpha_1\boldsymbol{v}_1)\otimes\boldsymbol{w}_1+\cdots+(\alpha_n\boldsymbol{v}_n)\otimes\boldsymbol{w}_n\\
&= \alpha_1(\boldsymbol{v}_1\otimes\boldsymbol{w})+\cdots+\alpha_n(\boldsymbol{v}_n\otimes\boldsymbol{w})
\end{aligned}$$

For each $\boldsymbol{v}_i\otimes\boldsymbol{w}$, $i=1,\dots,n$, similarly,

$$\boldsymbol{v}_i\otimes\boldsymbol{w}=\beta_1(\boldsymbol{v}_i\otimes\boldsymbol{w}_1)+\cdots+\beta_m(\boldsymbol{v}_i\otimes\boldsymbol{w}_m).$$

Therefore,

$$\boldsymbol{v} \otimes \boldsymbol{w} = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j (\boldsymbol{v}_i \otimes \boldsymbol{w}_j) \tag{13.2}$$

By (13.4), any vector in $V \otimes W$ is of the form

$$\boldsymbol{v}^{(1)} \otimes \boldsymbol{w}^{(1)} + \cdots + \boldsymbol{v}^{(\ell)} \otimes \boldsymbol{w}^{(\ell)}$$

By (13.5), each $\boldsymbol{v}^{(k)} \otimes \boldsymbol{w}^{(k)}, k = 1, \ldots, \ell$, can be expressed as

$$\boldsymbol{v}^{(k)} \otimes \boldsymbol{w}^{(k)} = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i^{(k)} \beta_j^{(k)} (\boldsymbol{v}_i \otimes \boldsymbol{w}_j)$$

Therefore,

$$\boldsymbol{v}^{(1)} \otimes \boldsymbol{w}^{(1)} + \cdots + \boldsymbol{v}^{(\ell)} \otimes \boldsymbol{w}^{(\ell)} = \sum_{k=1}^{\ell} \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i^{(k)} \beta_j^{(k)} (\boldsymbol{v}_i \otimes \boldsymbol{w}_j)$$

In other words, $\{\boldsymbol{v}_i \otimes \boldsymbol{w}_j \mid 1 \le i \le n, 1 \le j \le m\}$ spans $V \otimes W$. ■

Theorem 13.1 A basis of $V \otimes W$ is $\{\boldsymbol{v}_i \otimes \boldsymbol{w}_j \mid 1 \le i \le n, 1 \le j \le m\}$

Proof. By proposition (13.1), it suffices to show that the set $\{\boldsymbol{v}_i \otimes \boldsymbol{w}_j \mid 1 \le i \le n, 1 \le j \le m\}$ is linear independent. Suppose that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} (\boldsymbol{v}_i \otimes \boldsymbol{w}_j) = \boldsymbol{0} \tag{13.3}$$

Suppose that $\{\phi_1, \ldots, \phi_n\}$ is a dual basis of V^*, and $\{\psi_1, \ldots, \psi_m\}$ is a dual basis of W^*. Construct the mapping

$$\begin{aligned} \pi_{p,q} : \quad & V \times W \to \mathbb{F} \\ \text{with} \quad & \pi_{p,q} = \phi_p(\boldsymbol{v}) \psi_q(\boldsymbol{w}) \end{aligned}$$

- The mapping $\pi_{p,q}$ is actually bilinear: for instance,

$$\begin{aligned}\pi_{p,q}(a\boldsymbol{v}_1+b\boldsymbol{v}_2,\boldsymbol{w})&=\phi_p(a\boldsymbol{v}_1+b\boldsymbol{v}_2)\psi_q(\boldsymbol{w})\\&=(a\phi_p(\boldsymbol{v}_1)+b\phi_p(\boldsymbol{v}_2))\psi_q(\boldsymbol{w})\\&=a\phi_p(\boldsymbol{v}_1)\psi_q(\boldsymbol{w})+b\phi_p(\boldsymbol{v}_2)\psi_q(\boldsymbol{w})\\&=a\pi_{p,q}(\boldsymbol{v}_1,\boldsymbol{w})+b\pi_{p,q}(\boldsymbol{v}_2,\boldsymbol{w}).\end{aligned}$$

  Following the similar ideas, we can check that $\pi_{p,q}(\boldsymbol{v},a\boldsymbol{w}_1+b\boldsymbol{w}_2)=a\pi_{p,q}(\boldsymbol{v},\boldsymbol{w}_1)+b\pi_{p,q}(\boldsymbol{v},\boldsymbol{w}_2)$.

- Therefore, $\pi_{p,q}\in\mathrm{Obj}$. By the universal property of the tensor product, $\pi_{p,q}$ induces the unique linear transformation

$$\begin{aligned}\textstyle\prod_{p,q}:&\quad V\otimes W\to\mathbb{F}\\\text{with}&\quad\textstyle\prod_{p,q}(\boldsymbol{v}\otimes\boldsymbol{w})=\pi_{p,q}(\boldsymbol{v},\boldsymbol{w})\end{aligned}$$

  In other words, $\prod_{p,q}(\boldsymbol{v}\otimes\boldsymbol{w})=\phi_p(\boldsymbol{v})\psi_q(\boldsymbol{w})$.

- Applying the mapping $\Pi_{p,q}$ on both sides of (13.3), we imply

$$\Pi_{p,q}\left(\sum_{i=1}^n\sum_{j=1}^n\alpha_{ij}(\boldsymbol{v}_i\otimes\boldsymbol{w}_j)\right)=\Pi_{p,q}(\mathbf{0})$$

  Or equivalently,

$$\sum_{i=1}^n\sum_{j=1}^n\alpha_{ij}\Pi_{p,q}(\boldsymbol{v}_i\otimes\boldsymbol{w}_j)=0,$$

  i.e.,

$$\sum_{i=1}^n\sum_{j=1}^n\alpha_{ij}\phi_p(\boldsymbol{v}_i)\psi_q(\boldsymbol{w}_j)=\alpha_{p,q}=0$$

  Following this procedure, we can argue that $\alpha_{ij}=0,\forall i,\forall j$.

$\blacksquare$

Corollary 13.1 If $\dim(V), \dim(W) < \infty$, then $\dim(V \otimes W) = \dim(V)\dim(W)$

Proof. Check dimension of the basis of $V \otimes W$. ■

(R) The universal property can be very helpful. In particular, given a bilinear mapping, say $\phi : V \times W \to U$, we imply $\phi \in \text{Obj}$. By theorem (12.3), since i satisfies the universal property of tensor product, we can induce an unique linear transformation $\psi : V \otimes W \to U$.

Let's try another example for making use of the universal property:

Theorem 13.2 For finite dimension U and V,

$$V \otimes U \cong U \otimes V$$

Proof. Construct the mapping

$$\begin{aligned} \phi : \quad & V \times U \to U \otimes V \\ \text{with} \quad & \phi(\boldsymbol{v}, \boldsymbol{u}) = \boldsymbol{u} \otimes \boldsymbol{v} \end{aligned}$$

Indeed, ϕ is bilinear: for instance,

$$\begin{aligned} \phi(a\boldsymbol{v}_1 + b\boldsymbol{v}_2, \boldsymbol{u}) &= u \otimes (a\boldsymbol{v}_1 + b\boldsymbol{v}_2) \\ &= a(\boldsymbol{u} \otimes \boldsymbol{v}_1) + b(u \otimes \boldsymbol{v}_2) \\ &= a\phi(\boldsymbol{v}_1, \boldsymbol{u}) + b\phi(\boldsymbol{v}_2, \boldsymbol{u}) \end{aligned}$$

Therefore, $\phi \in \text{Obj}$. By the universal property of tensor product, we induce an unique linear transformation

$$\begin{aligned} \Phi : \quad & V \otimes U \to U \otimes V \\ \text{with} \quad & \Phi(\boldsymbol{v} \otimes \boldsymbol{u}) = \boldsymbol{u} \otimes \boldsymbol{v} \end{aligned}$$

Similarly, we may induce the linear transformation

$$\begin{aligned} \Psi : \quad & U \otimes V \to V \otimes U \\ \text{with} \quad & \Psi(\boldsymbol{u} \otimes \boldsymbol{v}) = \boldsymbol{v} \otimes \boldsymbol{u} \end{aligned}$$

Given any $\sum_i \boldsymbol{u}_i \otimes \boldsymbol{v}_i \in U \otimes V$, observe that

$$\begin{aligned} (\Phi \circ \Psi)\left(\sum_i \boldsymbol{u}_i \otimes \boldsymbol{v}_i\right) &= \Phi\left(\sum_i \Psi(\boldsymbol{u}_i \otimes \boldsymbol{v}_i)\right) \\ &= \Phi\left(\sum_i \boldsymbol{v}_i \otimes \boldsymbol{u}_i\right) \\ &= \sum_i \Phi(\boldsymbol{v}_i \otimes \boldsymbol{u}_i) \\ &= \sum_i \boldsymbol{u}_i \otimes \boldsymbol{v}_i \end{aligned}$$

Therefore, $\Phi \circ \Psi = \mathrm{id}_{U \otimes V}$. Similarly, $\Psi \circ \Phi = \mathrm{id}_{V \otimes U}$. Therefore,

$$U \otimes V \cong V \otimes U.$$

■

13.1.2. Tensor Product of Linear Transformation

Motivation. Given two linear transformations $T : V \to V'$ and $S : W \to W'$, we want to construct the tensor product

$$T \otimes S : V \otimes W \to V' \otimes W'$$

Question: is $T \otimes S$ a linear transformation?

Answer: Yes. Universal property plays a role!

13.2. Monday for MAT3006

Notations. In this lecture, we let $\int_I f(x,y)\,\mathrm{d}x$ denote the Lebesgue integral.

Theorem 13.3 Let I,J be intervals in $\mathbb{R}$, and $f : I\times J\to\mathbb{R}$ be a function such that

1. For fixed $y\in J$, the function $f(x):=f(x,y)$ is integrable on I
2. $\frac{\partial f}{\partial y}$ exists for any $(x,y)\in I\times J$
3. $\left|\frac{\partial f}{\partial y}(x,y)\right|\le g(x)$ for some integrable function $g(x)$ on I.

Then $F(y):=\int_I f(x,y)\,\mathrm{d}x$ is differentiable on J, with

$$F'(y)=\int_I \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$$

Proof. Fix $y\in J$, and consider any sequence $\{y_n\}$ (with $y_n\ne y$) in J converging to y.

Construct the function

$$g_n(x):=\frac{f(x,y_n)-f(x,y)}{y_n-y}$$

which follows that

1. The function g_n is integrable by hypothesis (1)
2. The function $g_n(x)$ converges to $\frac{\partial f}{\partial y}(x,y)$ as $n\to\infty$
3. By MVT, $|g_n(x)|=|\frac{\partial f}{\partial y}(x,\xi)|$, which is bounded by $g(x)$ by hypothesis (3).

Therefore, the DCT applies, and

$$\int_I g_n(x)\,\mathrm{d}x=\frac{1}{y_n-y}\left[\int f(x,y_n)\,\mathrm{d}x-\int f(x,y)\,\mathrm{d}x\right]\to\int_I\frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$$

In other words, for all sequences $\{y_n\}\to y$ with $y_n\ne y$,

$$\lim_{n\to\infty}\frac{F(y_n)-F(y)}{y_n-y}=\int_I\frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x$$

From the elementary analysis knowledge, in particular, $\lim_{y'\to y}H(y')$ exists (equal to

L) if and only if $\lim_{n\to\infty} H(y_n) = L$ for all sequences $\{y_n\} \to y$ with $y_n \neq y$. Therefore,

$$F'(y) := \lim_{y'\to y} \frac{F(y') - F(y)}{y' - y} = \int_I \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x.$$

■

13.2.1. Double Integral

Definition 13.1 [Measure in $\mathbb{R}^2$] In $\mathbb{R}^2$, we can define the **measure** of the rectangle $A \times B \subseteq \mathbb{R}^2$ with $A, B \in \mathcal{M}$ by

$$m^*(A \times B) = m(A)m(B)$$

In particular, we define

$$x \cdot \infty = \infty \cdot x = (-x) \cdot (-\infty) = \begin{cases} \infty, & \text{if } x > 0 \\ -\infty, & \text{if } x < 0 \\ 0, & \text{if } x = 0 \end{cases}$$

Definition 13.2 [Outer Measure in $\mathbb{R}^2$] Then the outer measure of any $E \subseteq \mathbb{R}^2$ is defined as

$$m^*(E) := \inf\left\{ \sum_{i=1}^{\infty} m(R_i) \,\middle|\, E \subseteq \bigcup_{i=1}^{\infty} R_i, R_i = A_i \times B_i, A_i, B_i \in \mathcal{M} \right\}$$

Definition 13.3 [Lebesgue Measurable in $\mathbb{R}^2$] A subset $E \subseteq \mathbb{R}^2$ is Lebesgue measurable if E satisfies the Carathedory Property:

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E),$$

for any subset $A \subseteq \mathbb{R}^2$.

Product Space of $\mathbb{R}^2$. Given two measurable spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$, in particular, we are interested in

$$(X, \mathcal{A}, \mu) = (Y, \mathcal{B}, \lambda) = (\mathbb{R}, \mathcal{M}, m).$$

Now we want to construct another measurable space in $X \times Y := \mathbb{R}^2$.

1. Start from the "measurable rectangles"
$$\mathcal{A} \times \mathcal{B} = \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$$
2. Define the function $\pi : \mathcal{A} \times \mathcal{B} \to [0, \infty]$ by
$$\pi(A \times B) = \mu(A)\lambda(B).$$
3. Let $\mathcal{A} \otimes \mathcal{B}$ be the smallest σ-algebra containing $\mathcal{A} \times \mathcal{B}$. Then by Caratheodory extension theorem, we can extend $\pi : \mathcal{A} \times \mathcal{B} \to [0, \infty]$ to $\tilde{\pi} : \mathcal{A} \otimes \mathcal{B} \to [0, \infty]$ such that
   (a) $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \tilde{\pi})$ is a measurable space
   (b) $\tilde{\pi} \mid_{\mathcal{A} \times \mathcal{B}} = \pi$.

R

- If further we have $\mathcal{A}$ and $\mathcal{B}$ are σ-finite, i.e., there exists $E_i \in \mathcal{A}$ such that $X = \cup_{i=1}^{\infty} E_i$, $\mu(E_i) < \infty, \forall i$, then we can imply the extension $\tilde{\pi}$ is unique. (For instance, $\mathbb{R} = \cup_{n \in \mathbb{Z}} [n, n+1]$ and $m([n, n+1]) = 1 < \infty$, i.e., $(\mathbb{R}, \mu, m)$ is σ-finite.)
- Question: we have constructed two measurable space $(\mathbb{R} \times \mathbb{R}, \mathcal{M} \otimes \mathcal{M}, \tilde{\pi})$ and $(\mathbb{R}^2, \mathcal{M}_{\mathbb{R}^2}, m)$. Are they the same?

  Answer : no, but the latter can be obtained from the former by completion process. In particular,
$$m \mid_{\mathcal{M} \otimes \mathcal{M}} = \tilde{\pi}.$$

Let's study the measurable space $(\mathbb{R}\times\mathbb{R}, \mathcal{M}\otimes\mathcal{M}, \pi)$ first, where $f:\mathbb{R}^2\to[-\infty,\infty]$ is a measurable function, i.e., $f^{-1}((a,\infty])\in\mathcal{A}\otimes\mathcal{B}$. In particular, we say $E\subseteq\mathbb{R}\times\mathbb{R}$ is measurable if $E\in\mathcal{M}\otimes\mathcal{M}$ for the moment being (but we will generalize the notion of measurable into $\mathcal{M}_{\mathbb{R}^2}$ in the future).

Definition 13.4 [x-section and y-section] Let $E\subseteq X\times Y$, with $(x,y)\in E$. Define

- the x-section $E_x=\{y\in Y \mid (x,y)\in E\}$, for fixed $x\in X$
- the y-section $E_y=\{x\in X \mid (x,y)\in E\}$, for fixed $y\in Y$.

Proposition 13.2 Suppose that $E\subseteq X\times Y$ is measurable (i.e., $E\in\mathcal{A}\otimes\mathcal{B}$), then $E_x\in\mathcal{B}$ and $E_y\in\mathcal{A}$.

Proof. Construct the set $\mathfrak{A}=\{E\in\mathcal{A}\otimes\mathcal{B}\mid E_x\in\mathcal{B}\}$. It suffices to show $\mathfrak{A}=\mathcal{A}\otimes\mathcal{B}$. We claim that

1. $\mathfrak{A}$ is a σ-algebra
2. $\mathfrak{A}$ contains all $A\times B\in\mathcal{A}\times\mathcal{B}$

If the claim (1) and (2) hold, and since $\mathcal{A}\otimes\mathcal{B}$ is the smallest-σ-algebra containing $\mathcal{A}\times\mathcal{B}$, we imply $\mathcal{A}\otimes\mathcal{B}\subseteq\mathfrak{A}\subseteq\mathcal{A}\otimes\mathcal{B}$, i.e., the proof is complete.

1. (a) Note that $\emptyset\in\mathfrak{A}$, and $X\times Y\in\mathfrak{A}$ since $(X\times Y)_x=Y\in\mathcal{B}$.
   (b) Suppose that $E_i\in\mathfrak{A}$, $i\geq 1$, i.e., $(E_i)_x\in\mathcal{B}$. Observe that
   $$\left(\bigcup_{i=1}^{\infty}E_i\right)_x=\bigcup_{i=1}^{\infty}(E_i)_x\in\mathcal{B},$$
   since $\mathcal{B}$ is a σ-algebra. Therefore, $\cup_{i=1}^{\infty}E_i\in\mathfrak{A}$.
   (c) Suppose that $E\in\mathfrak{A}$, i.e., $(E)_x\in\mathcal{B}$, then
   $$\begin{aligned}(E^c)_x&=\{y\mid(x,y)\in E^c\}\\&=\{y\mid(x,y)\notin E\}\\&=(E_x)^c\in\mathcal{B}\end{aligned}$$

which implies $E^c \in \mathfrak{A}$.

2. For any $A \times B \in \mathcal{A} \times \mathcal{B}$, since $(A \times B)_x = B \in \mathcal{B}$, we imply $(A \times B) \in \mathfrak{A}$.

In conclusion, $\mathfrak{A} = \mathcal{A} \otimes \mathcal{B}$. For all $E \in \mathcal{A} \otimes \mathcal{B}$, we imply $E \in \mathfrak{A}$, i.e., $E_x \in \mathcal{B}$. ∎

Proposition 13.3 Sippose that $f : X \times Y \to [-\infty, \infty]$ is measurable. (i.e., $f^{-1}((a, \infty]) \in \mathcal{A} \otimes \mathcal{B}$), then the maps

$$\begin{cases} f_x : & Y \to [-\infty, \infty] \\ \text{with} & f_x(y) := f(x, y) \end{cases}, \qquad \begin{cases} f_y : & X \to [-\infty, \infty] \\ \text{with} & f_y(x) := f(x, y) \end{cases}$$

are measurable. More precisely, $f_x^{-1}((a, \infty]) \in \mathcal{B}$ and $f_y^{-1}((a, \infty]) \in \mathcal{A}$.

Proof.

$$\begin{aligned} f_x^{-1}((a, \infty]) &= \{y \in Y \mid f_x(y) \in (a, \infty]\} \\ &= \{y \in Y \mid f(x, y) > a\} \\ &= \{(u, y) \in X \times Y \mid f(u, y) > a\}_x \\ &= (f^{-1}((a, \infty]))_x \in \mathcal{B} \end{aligned}$$

∎

13.3. Monday for MAT4002

13.3.1. Isomorphsim between Edge Loop Group and the Fundamental Group

Recall that

$$\pi_1(X,b) := \{[\ell] \mid \ell : [0,1] \to X \text{ denotes the loops based at } b\}$$

and

$$E(K,b) = \{[\alpha] \mid \alpha \text{ is an edge loop in } K \text{ based at } b\}$$

Now we show that the mapping defined below is injective:

$$\begin{aligned} \theta : \quad & E(K,b) \to \pi_1(|K|,b) \\ \text{with} \quad & [\alpha] \mapsto [|g_\alpha|] \end{aligned}$$

- Let $\alpha = (v_0,\dots,v_n)$ be an edge loop based at b such that $\theta([\alpha]) = e$, i.e., $|g_\alpha| \simeq c_b$. It suffices to show that $[\alpha]$ is the identity element of $E(K,b)$.
- Choose a homotopy $H : |g_\alpha| \simeq c_b$ such that $H : I \times I \to |K|$. The graphic illustration for H is shown in Fig. (13.8).

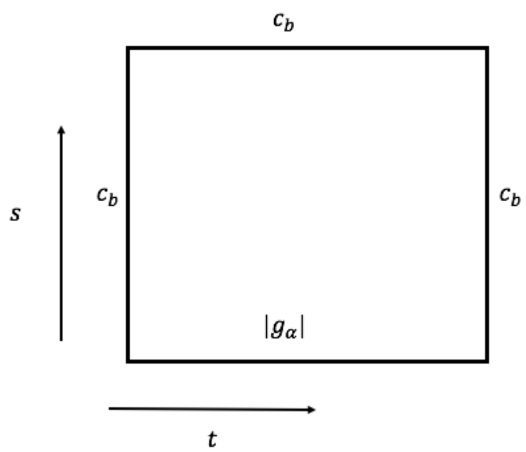


Figure 13.1: Graphic illustration for $H : I \times I \to |K|$

Now apply the simplicial approximation theorem, there exists a subdivision of $I \times I$, denoted as $(I \times I)_{(r)}$ (for sufficiently large r), shown in the Fig. (13.9)

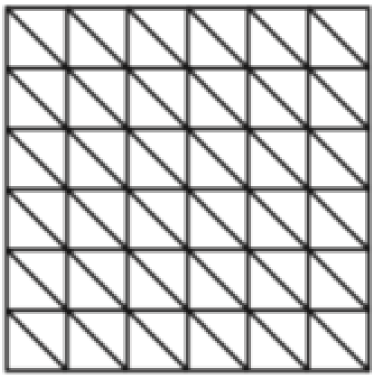


Figure 13.2: Graphic illustration for $(I \times I)_{(r)}$. In particular, divide $I \times I$ into r^2 congruent squares, and then further divide each of these squares along the diagonal to form $(I \times I)_{(r)}$.

such that $|(I \times I)_{(r)}| = I \times I$, and there exists the simplicial map

$$\begin{aligned} G: \quad & (I \times I)_{(r)} \to K \\ \text{such that} \quad & |G| \simeq H. \end{aligned}$$

Without loss of generality, assume r is a sufficiently large multiple of n.

The graphic illustration of $|G|$ is shown in Fig. (13.3):

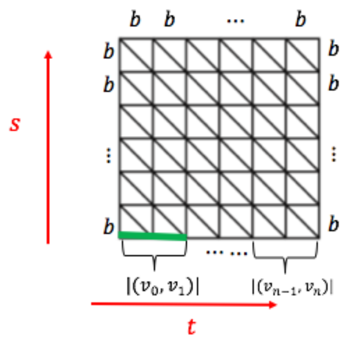


Figure 13.3: Graphic illustration for the mapping $|G|$.

In particular, $|G|$ maps $\{0,1\} \times I$ into $\{b\}$; $I \times \{1\}$ into $\{b\}$; $(i/n, 0)$ into $\{v_i\}, i =$

$0, \ldots, n$, and $[i/n, (i+1)/n]$ into $|(v_i, v_{i+1})|, i = 0, \ldots, n-1$.

- Consider the simplicial subcomplex of $(I \times I)_{(r)}$ shown in Fig. (13.4)

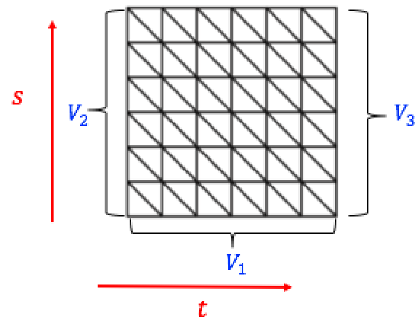


Figure 13.4: Graphic illustration for the simplicial subcomplex V_1, V_2, V_3.

For instance, V_1 has $(r+1)$ 0-simplicies and r 1-simplies. It follows that

$$H(|V_1|) = H(|V_2|) = H(|V_3|) = \{b\}.$$

By proposition (10.6), we can pick G be such that

$$G(V_1) = G(V_2) = G(V_3) = \{b\}.$$

Consider W_1 as the simplicial subcomplex of $(I \times I)_{(r)}$ given by the green line shown in Fig. (13.3), which follows that

$$H(|W_1|) = \{v_0, v_1\} \implies G(W_1) = \{v_0, v_1\}$$

Similarly,

$$H(|W_i|) = \{v_{i-1}, v_i\} \implies G(W_i) = \{v_{i-1}, v_i\}, \forall 1 \leq i \leq n.$$

As a result, $|G|(|V_1|) = \beta := (bv_0 \cdots v_0 v_1 \cdots v_1 \cdots v_n \cdots v_n b)$, and clearly,

$$\begin{aligned} \beta &\sim (bv_0 v_1 v_2 \cdots v_{n-1} v_n b) \\ &\sim (bv_1 v_2 \cdots v_{n-1} b) = \alpha \end{aligned}$$

- Now it suffices to show $\beta \simeq e$. This is true by the sequence of elementary contractions and expansions as shown in the Fig. (13.5).

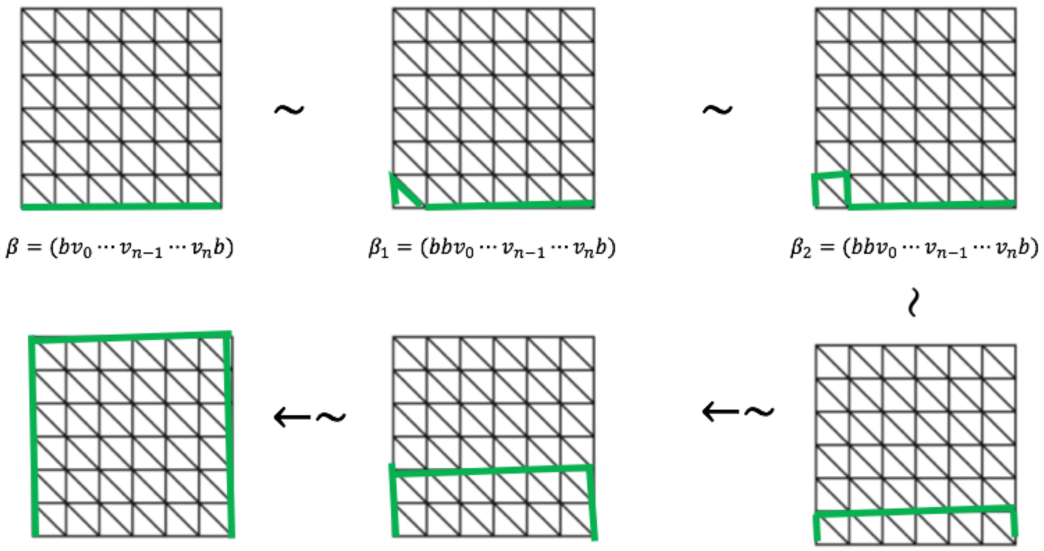


Figure 13.5: A sequence of elementary contractions and expansions to show that $\beta \sim (b \cdots b) = (b)$.

R The definition of $E(K,b)$ only involves n-simplicials for $n \leq 2$.

Proposition 13.4 For any simplicial complex K, consider the simplicial subcomplex $\text{Skel}^n(K) = (V_k, \Sigma_K^n)$, where Σ_K^n consists of $\sigma \in \Sigma_K$ with $|\sigma| \leq n+1$ (this is the n-skeleton of K). Then

$$\pi_1(|K|, b) \cong \pi_1(|\text{Skel}^2(K)|, b)$$

Proof. Since $E(K,b)$ only involves n-simplicials for $n \leq 2$, we imply $E(K,b) \cong E(\text{Skel}^2(K), b)$.

Moreover, $\pi_1(|K|,b) \cong E(K,b)$ and $\pi_1(|\text{Skel}^2(K)|, b) \cong E(\text{Skel}^2(K), b)$.

The proof is complete. ∎

Corollary 13.2 For $n \geq 2$, $\pi_1(S^n)$ is a trivial fundamental group.

Proof. Consider the simplicial complex K with

$$V = \{1,2,\dots,n+2\}, \quad \Sigma = \{\text{all proper subsets of } V\}$$

It's clear that $|K| \cong S^n$, and $\text{Skel}^2(K)$ has

- $V : \{1,\dots,n+2\}$
- Σ^2 : all subsets of V with less or equal to 3 elements.

For any edge loop a in $\pi_1(|\text{skel}^2(K)|)$, we have

$$\begin{aligned} a &= (bv_0v_1v_2\cdots v_n) \\ &\sim (bv_1v_2\cdots v_{n-2}v_{n-1}b) \\ &\sim \cdots \\ &\sim (b) \end{aligned}$$

Therefore, all edge loops α in $\pi_1(|\text{skel}^2(K)|)$ satisfies $[\alpha] = [(b)] = e.$, i.e.,

$$\pi_1(|\text{skel}^2(K)|) \cong \{e\},$$

which implies $\pi_1(|K|) \cong \pi_1(|\text{skel}^2(K)|) \cong \{e\}$. Since $|K| \cong S^n$, we imply

$$\pi_1(S^n) \cong \pi_1(|K|) \cong \{e\}.$$

∎

R The Corollary (13.2) does not hold for S^1 since the constructed Σ^2 for S^1 does not contain $\{1,2,3\}$.

Theorem 13.4 $\pi_1(S^1) \cong \mathbb{Z}$.

Proof. Construct the triangle K shown in Fig. (13.6), and it's clear that $|K| \cong S^1$.

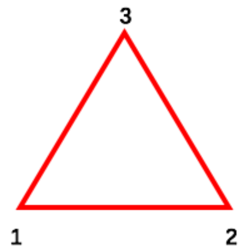


Figure 13.6: Triangle K such that $|K| \cong S^1$

It suffices to show $E(K,1) \cong \mathbb{Z}$. Define the orientation of $|K|$ as shown in Fig. (13.7).

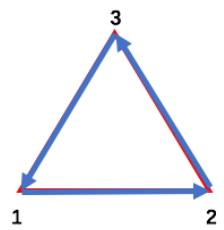


Figure 13.7: Orientation of $|K|$

Any edge loop α based at 1 is equivalent to the canonical form

$$\alpha \sim (1bc1bc\cdots 1bc1), \quad \text{where } bc = 32 \text{ or } 23.$$

We construct the isomorphism between $E(K,b)$ and $\mathbb{Z}$ directly:

$$\begin{aligned} \phi: \quad & E(K,b) \to \mathbb{Z} \\ \text{with} \quad & [\alpha] \mapsto \text{winding number of } \alpha \end{aligned}$$

where the winding number of α is the number of times it traverses $(2,3)$ in the forwards direction minus the number of times it traverses $(3,2)$ in the backwards direction.

The difficult part is to show the well-definedness of ϕ, which can be done by using canonical form of α. ■

13.4. Wednesday for MAT3040

13.4.1. Tensor Product for Linear Transformations

Proposition 13.5 Suppose that $T : V \to V'$ and $S : W \to W'$ are linear transformations, then there exists an unique linear transformation

$$\begin{aligned} T \otimes S : &\quad V \otimes W \to V' \otimes W' \\ \text{satisfying} &\quad (T \otimes S)(v \otimes w) = T(v) \otimes S(w) \end{aligned}$$

Proof. We construct the mapping

$$\begin{aligned} T \times S : &\quad V \times W \to V' \otimes W' \\ \text{with} &\quad (T \times S)(v, w) = T(v) \otimes S(w) \end{aligned}$$

This mapping is indeed bilinear: for instance, we can show that

$$(T \times S)(av_1 + bv_2, w) = a(T \times S)(v_1, w) + b(T \times S)(v_2, w)$$

Therefore, $T \times S \in \mathrm{Obj}$. Since the tensor product satisfies the universal property, we imply there exists an unique linear transformation

$$\begin{aligned} T \otimes S &\quad V \otimes W \to V' \otimes W' \\ \text{satisfying} &\quad (T \otimes S)(v \otimes w) = T(v) \otimes S(w) \end{aligned}$$

∎

Notation Warning. Does the notion $T \otimes S$ really form a tensor product, i.e., do we obtain the addictive rules for tensor product such as

$$(aT_1 + bT_2) \otimes S = a(T_1 \otimes S) + b(T_2 \otimes S)?$$

■ **Example 13.2** Let $V = V' = \mathbb{F}^2$ and $W = W' = \mathbb{F}^3$. Define the matrix-multiply mappings:

$$\begin{cases} T: & V \to V \\ \text{with} & \boldsymbol{v} \mapsto \boldsymbol{A}\boldsymbol{v} \\ & \boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{cases} \qquad \begin{cases} S: & W \to W \\ \text{with} & \boldsymbol{w} \mapsto \boldsymbol{B}\boldsymbol{w} \\ & \boldsymbol{B} = \begin{pmatrix} p & q & r \\ s & t & u \\ v & w & x \end{pmatrix} \end{cases}$$

How does $T \otimes S : V \otimes W \to V \otimes W$ look like?

- Suppose $\{e_1, e_2\}, \{f_1, f_2, f_3\}$ are usual basis of V, W, respectively. Then the basis of $V \otimes W$ is given by:

$$C = \{e_1 \otimes f_1, e_1 \otimes f_2, e_1 \otimes f_3, e_2 \otimes f_1, e_2 \otimes f_2, e_2 \otimes f_3\}.$$

- As a result, we can compute $(T \otimes S)(e_i \otimes f_j)$ for $i = 1,2$ and $j = 1,2,3$. For instance,

$$\begin{aligned} (T \otimes S)(e_1 \otimes e_1) &= T(e_1) \otimes S(e_1) \\ &= (ae_1 + ce_2) \otimes (pe_1 + se_2 + ve_3) \\ &= (ap)e_1 \otimes e_1 + (as)e_1 \otimes e_2 + (av)e_1 \otimes e_3 + (cp)e_2 \otimes e_1 + (cs)e_2 \otimes e_2 + (cv)e_2 \otimes e_3 \end{aligned}$$

- Therefore, we obtain a matrix representation for the linear transformation $(T \otimes S)$:

We want a matrix representation for $(T \otimes S)$:

$$(T \otimes S)_{C,C} = \begin{pmatrix} aB & bB \\ cB & dB \end{pmatrix},$$

which is a large matrix formed by taking all possible products between the elements of $\boldsymbol{A}$ and those of $\boldsymbol{B}$. This operation is called the **Kronecker Tensor Product**, see the command *kron* in MATLAB for detail. ■

Proposition 13.6 More generally, given the linear operator $T : V \to V$ and $S : W \to W$, let $\mathcal{A} = \{v_1, \ldots, v_n\}, \mathcal{B} = \{w_1, \ldots, w_m\}$ be a basis of V, W respectively, with

$$(T)_{\mathcal{A},\mathcal{A}} = (a_{ij}) \quad (S_{\mathcal{B},\mathcal{B}}) = (b_{ij}) := B$$

As a result, $(T \otimes S)_{C,C} = A \otimes B$, where $C = \{v_1 \otimes w_1, \ldots, v_n \otimes w_m\}$, and $A \otimes B$ denotes the Kronecker tensor product, defined as the matrix

$$\begin{pmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \cdots & a_{n,n}B \end{pmatrix}.$$

Proof. Following the similar procedure as in Example (13.2) and applying the relation

$$\begin{aligned}
(T \otimes S)(v_i \otimes w_j) &= T(v_i) \otimes S(w_j) \\
&= \left(\sum_{k=1}^{n} a_{ki} v_k\right) \otimes \left(\sum_{\ell=1}^{m} b_{\ell j} w_\ell\right) \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{m} (a_{ki} b_{\ell j}) v_k \otimes w_\ell
\end{aligned}$$

■

Proposition 13.7 The operation $T \otimes S$ satisfies all the properties of tensor product. For example,

$$(aT_1 + bT_2) \otimes S = a(T_1 \otimes S) + b(T_2 \otimes S)$$

$$T \otimes (cS_1 + dS_2) = c(T \otimes S_1) + d(T \otimes S_2)$$

Therefore, the usage of the notion "$\otimes$" is justified for the definition of $T \otimes S$.

Proof using matrix multiplication. For instance, consider the operation $(T + T') \otimes S$, with $(T)_{\mathcal{A},\mathcal{A}} = (a_{ij})$, $(T')_{\mathcal{A},\mathcal{A}} = (c_{ij}), (S)_{\mathcal{B},\mathcal{B}} = B$.

We compute its matrix representation directly:

$$
\begin{aligned}
((T+T')\otimes S)_{\mathcal{C},\mathcal{C}} &= (T+T')_{\mathcal{A},\mathcal{A}} \otimes (S)_{\mathcal{B},\mathcal{B}} \\
&= [(T)_{\mathcal{A},\mathcal{A}} + (T')_{\mathcal{A},\mathcal{A}}] \otimes (S)_{\mathcal{B},\mathcal{B}} \\
&= (T)_{\mathcal{A},\mathcal{A}} \otimes (S)_{\mathcal{B},\mathcal{B}} + (T')_{\mathcal{A},\mathcal{A}} \otimes (S)_{\mathcal{B},\mathcal{B}}
\end{aligned}
$$

where the last equality is by the addictive rule for kronecker product for matrices. Therefore,

$$
((T+T')\otimes S)_{\mathcal{C},\mathcal{C}} = (T\otimes S)_{\mathcal{C},\mathcal{C}} + (T'\otimes S)_{\mathcal{C},\mathcal{C}} \implies (T+T')\otimes S = T\otimes S + T'\otimes S
$$

■

Proof using basis of $T\otimes S$. Another way of the proof is by computing

$$
((T+T')\otimes S)(v_i\otimes w_j),
$$

where $\{v_i\otimes w_j \mid 1\le i\le n, 1\le j\le m\}$ forms a basis of $(T+T')\otimes S$:

$$
\begin{aligned}
((T+T')\otimes S)(v_i\otimes w_j) &= (T+T')(v_i)\otimes S(w_j) \\
&= (T(v_i)+T'(v_i))\otimes S(w_j) \\
&= T(v_i)\otimes S(w_j) + T'(v_i)\otimes S(w_j) \\
&= (T\otimes S)(v_i\otimes w_j) + (T'\otimes S)(v_i\otimes w_j)
\end{aligned}
$$

Since $((T+T')\otimes S)(v_i\otimes w_j)$ coincides with $(T\otimes S+T'\otimes S)(v_i\otimes w_j)$ for all basis vectors $v_i\otimes w_j\in\mathcal{C}$, we imply

$$
(T+T')\otimes S = T\otimes S + T'\otimes S
$$

■

Proposition 13.8 Let A,C be linear operators from V to V, and B,D be linear operators from W to W, then

$$
(A\otimes B)\circ(C\otimes D) = (AC)\otimes(BD)
$$

Proposition 13.9 Define linear operators $A: V \to V$ and $B: W \to W$ with $\dim(V), \dim(W) < \infty$. Then

$$\det(A \otimes B) = (\det(A))^{\dim(W)}(\det(B))^{\dim(V)}$$

Corollary 13.3 There exists a linear transformation

$$\begin{aligned} \Phi: \quad & \mathrm{Hom}(V,V) \otimes \mathrm{Hom}(W,W) \to \mathrm{Hom}(V \otimes W, V \otimes W) \\ \text{with} \quad & A \otimes B \mapsto A \otimes B \end{aligned}$$

where the input of Φ is the tensor product of linear transformations, and the output is the linear transformation.

Proof. Construct the mapping

$$\begin{aligned} \Phi \quad & : \mathrm{Hom}(V,V) \times \mathrm{Hom}(W,W) \to \mathrm{Hom}(V \otimes W, V \otimes W) \\ \text{with} \quad & \Phi(A,B) = A \otimes B \end{aligned}$$

The Φ is indeed bilinear: for instance,

$$\begin{aligned} \Phi(pA + qC, B) &= (pA + qC) \otimes B \\ &= p(A \otimes B) + q(C \otimes B) \\ &= p\Phi(A,B) + q\Phi(C,B) \end{aligned}$$

This corollary follows from the universal property of tensor product. ■

R If assuming that $\dim(V), \dim(W) < \infty$, we imply

$$\begin{aligned} \dim(\text{Input space of } \Phi) = \dim(\mathrm{Hom}(V,V))\dim(\mathrm{Hom}(W,W)) \\ &= [\dim(V)\dim(V)] \cdot [\dim(W)\dim(W)] = [\dim(V)\dim(W)]^2 \\ &= [\dim(V \otimes W)]^2 \\ &= \dim(\mathrm{Hom}(V \otimes W, V \otimes W)) \\ &= \dim(\text{Output space of } \Phi) \end{aligned}$$

Therefore, is Φ is an isomorphism? If so, then every linear operator $\alpha : V \otimes W \to V \otimes W$ can be expressed as

$$\alpha = A_1 \otimes B_1 + \cdots + A_k \otimes B_k$$

where $A_i : V \to V$ and $B_j : W \to W$.

13.5. Wednesday for MAT3006

13.5.1. Fubini's and Tonell's Theorem

Motivation. Given two measurable space $(\mathbb{R}, \mathcal{M}, \mathrm{d}x)$ and $(\mathbb{R}, \mathcal{M}, \mathrm{d}y)$, we have constructed the product measurable space $(\mathbb{R}^2, \mathcal{M} \otimes \mathcal{M}, \mathrm{d}\pi)$. Suppose $f : \mathbb{R}^2 \to [-\infty, \infty]$ is measurable on this space, now we want to show that

$$\int f(x,y)\,\mathrm{d}\pi = \int \left(\int f_y(x)\,\mathrm{d}x \right) \mathrm{d}y = \int \left(\int f_x(y)\,\mathrm{d}y \right) \mathrm{d}x$$

Easier Goal. The proof for the statement above is hard. Consider the easier case where f is a simple function first, i.e., $f(x,y) = \mathcal{X}_E(x,y), E \in \mathcal{M} \otimes \mathcal{M}$, which follows that

$$\begin{aligned}
&\int \mathcal{X}_E(x,y)\,\mathrm{d}\pi = \pi(E) \\
&\int (\mathcal{X}_E)_y(x)\,\mathrm{d}x = \int \mathcal{X}_{E_y}(x)\,\mathrm{d}x = m_X(E_y) \\
&\int (\mathcal{X}_E)_x(y)\,\mathrm{d}y = \int \mathcal{X}_{E_x}(y)\,\mathrm{d}x = m_Y(E_x)
\end{aligned}$$

Therefore, our easier goal is to show that

$$\pi(E) = \int m_X(E_y)\,\mathrm{d}y = \int m_Y(E_x)\,\mathrm{d}x, \quad \forall E \in \mathcal{M} \otimes \mathcal{M}. \tag{13.4}$$

Easiest Goal. Consider the simplest case where $E = A \times B \in \mathcal{M} \otimes \mathcal{M}$, where $A \in \mathcal{M}_X, B \in \mathcal{M}_Y$, which implies

- $\pi(A \times B) = m_X(A)m_Y(B)$
- As shown in the figure (13.8), for fixed $y \in Y$,

$$m_X((A \times B)_y) = \begin{cases} m_X(A), & \text{if } y \in B \\ m_X(\emptyset) = 0, & \text{if } y \notin B \end{cases} = m_X(A)\mathcal{X}_B(y)$$

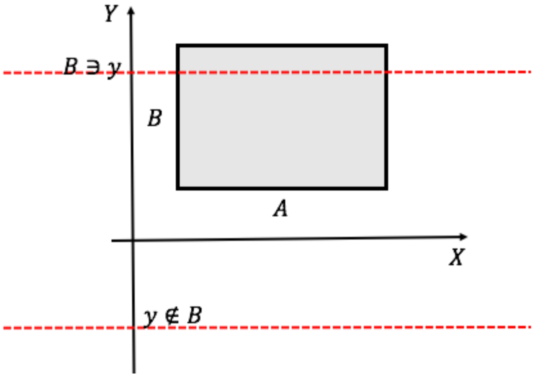


Figure 13.8: Illustration for $m_X((A\times B)_y)$

Therefore, we imply

$$\begin{aligned}\int m_X((A\times B)_y)\,\mathrm{d}y &= \int m_X(A)\mathcal{X}_B(y)\,\mathrm{d}y \\ &= m_X(A)\int \mathcal{X}_B(y)\,\mathrm{d}y \\ &= m_X(A)m_Y(B)\end{aligned}$$

Similarly,

$$\int m_Y((A\times B)_x)\,\mathrm{d}x = m_X(A)m_Y(B).$$

Therefore, the easiest goal (Eq. (13.4)) holds for $E = A\times B \in \mathcal{M}\times\mathcal{M}$.

(R) Generalization from the easier goal to the real goal is trivial, i.e., applying MCT is ok. The difficulty is that how to show the easier goal (Eq. (13.4)) holds for any $E\in\mathcal{M}\otimes\mathcal{M}$, given that the easier goal (Eq. (13.4)) holds for any $E\in\mathcal{M}\times\mathcal{M}$.

Definition 13.5 [Monotone Class] Let X be a non-empty set. A **monotone class** $\mathcal{T}$ is a collection of subsets of X closed under countable increasing unions and countable decreasing intersections, i.e.,

1. If $E_i \in \mathcal{T}(i \in \mathbb{N})$ and $E_i \subseteq E_{i+1}, \forall i$, then

$$\bigcup_{i=1}^{\infty} E_i \in \mathcal{T}$$

2. If $F_i \in \mathcal{T}(i \in \mathbb{N})$ with $F_i \supseteq F_{i+1}, \forall i$, then

$$\bigcap_{i=1}^{\infty} F_i \in \mathcal{T}$$

■

R Every σ-algebra is a monotone classs. In particular, for $X = \mathbb{R}$, the collection of subsets $\mathcal{M}$ and $\mathcal{B}$ are both monotone classes.

Definition 13.6 [Smallest Monotone Class] For any $S \subseteq \mathcal{P}(X)$, denote

$$\mathcal{M}(S) := \bigcap_{\mathcal{T} \text{ is a monotone class such that } S \subseteq \mathcal{T}} \mathcal{T},$$

which is also the monotone class. We call $\mathcal{M}(S)$ as the smallest monotone class containing S. ■

R It's clear that $\mathcal{M}(S) \subseteq \sigma(S)$, where $\sigma(S)$ is the smallest σ-algebra containing S. Question: when do we have $\mathcal{M}(S) = \sigma(S)$?

Theorem 13.5 — Monotone Class Theorem. Let X be a non-empty set. If $S \subseteq \mathcal{P}(X)$ is an **algebra** (i.e., $E_1, E_2 \in S \implies E_1 \cup E_2 \in S, E_1 \cap E_2 \in S, E_1^c \in S$), then $\mathcal{M}(S) = \sigma(S)$.

We skip the proof for the monotone class theorem, but you may refer to the proof in the blackboard.

■ **Example 13.3** 1. Let $X = \mathbb{R}$, and $S^1 = \{\text{all intervals}\}$ is **not** an algebra, e.g.,

$$[1,2] \in S^1 \implies [1,2]^c = (-\infty, 1) \cup (2, \infty) \notin S^1.$$

However, $S = \{\text{finite disjoint union of intervals}\}$ is an algebra. Therefore,

$$\mathcal{M}(S) = \sigma(S) := \mathcal{B}(\text{Borel } \sigma\text{-algebra}).$$

2. Let $X = \mathbb{R}^2$, and define

$$S = \left\{ \text{finite disjoint union of measurable rectangles } \bigcup_{i=1}^{k} (A_i \times B_i) \,\middle|\, A_i, B_i \in \mathcal{M} \right\}$$

Then S is an algebra, for instance, as shown in the Fig. (13.9), $(A \times B)^c = (A^c \times \mathbb{R}) \cup (A \times B^c)$ is a disjoint union of 2 measurable rectangles.

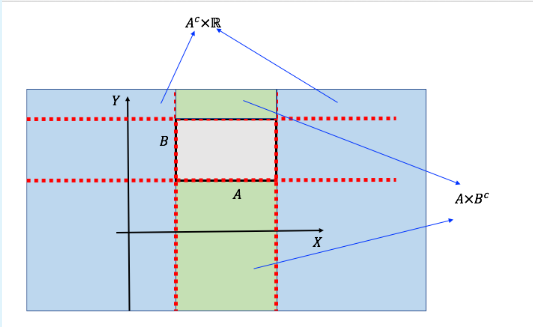


Figure 13.9: Illustration for $(A \times B)^c$

Therefore, $\mathcal{M}(S) = \sigma(S) := \mathcal{M} \otimes \mathcal{M}$

Proposition 13.10 For all $E \in \mathcal{M} \otimes \mathcal{M}$, we have

$$\pi(E) = \int m_Y(E_x)\,\mathrm{d}x = \int m_X(E_y)\,\mathrm{d}y \tag{13.5}$$

Proof. Construct

$$\mathcal{A} = \left\{ E \in \mathcal{M} \otimes \mathcal{M} \,\middle|\, \begin{array}{l} x \mapsto m_Y(E_x) \text{ is a measurable function of } x \\ y \mapsto m_X(E_y) \text{ is a measurable function of } y \\ \text{Eq. (13.5) holds} \end{array} \right\}$$

- Claim 1: $\mathcal{A}$ is a monotone class

- Claim 2: Any finite disjoint union of measurable rectangles is in $\mathcal{A}$:

$$\bigcup_{i=1}^{k}(A_i \times B_i) \in \mathcal{A}, \quad k \in \mathbb{N}$$

If claim (1),(2) holds, then $S \subseteq \mathcal{A}$, where

$$S = \{\text{finite disjoint union of measurable rectangles}\}$$

which follows that

$$\mathcal{M}(S) \subseteq \mathcal{A}.$$

By monotone class theorem, $\sigma(S) = \mathcal{M}(S) \subseteq \mathcal{A}$, i.e.,

$$\mathcal{M} \otimes \mathcal{M} = \sigma(S) = \mathcal{M}(S) \subseteq \mathcal{A} \subseteq \mathcal{M} \otimes \mathcal{M} \implies \mathcal{M} \otimes \mathcal{M} = \mathcal{A}.$$

Therefore, (13.5) holds for all $E \in \mathcal{A} = \mathcal{M} \otimes \mathcal{M}$.

We left the proof for claim (1) in next class. Now we give a proof for claim (2):

- For any $E = \cup_{i=1}^{k}(A_i \times B_i)$,

  $$m_Y(E_x) = \sum_{i=1}^{k} m_Y(B_i)\mathcal{X}_{A_i}(x)$$

  is a simple function on x, and therefore measurable.

- Similarly,

  $$m_X(E_y) = \sum_{i=1}^{k} m_X(A_i)\mathcal{X}_{B_i}(y)$$

  is also measurable.

- By the easiest goal, (13.5) also holds.

Therefore, claim (2) is true. ■

13.6. Wednesday for MAT4002

13.6.1. Applications on the isomorphism of fundamental group

Theorem 13.6

$$\pi_1(S^1) \cong (\mathbb{Z}, +)$$

Proof. Define the orientation of $|K|$ as shown in Fig. (13.10).

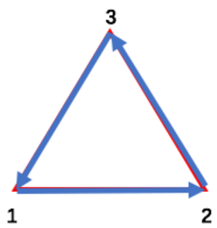


Figure 13.10: Orientation of $|K|$

Following the proof during last lecture, we construct

$$\begin{aligned} \phi : \quad & E(K,1) \to (\mathbb{Z},+) \\ \text{with} \quad & [\alpha] \mapsto \text{winding number of } \alpha \end{aligned}$$

where the winding number of α is the

$$\text{number of 23 appearing in } \alpha - \text{number of 32 appearing in } \alpha.$$

Note that

1. The winding number is invariant under elementary contraction and elementary expansion.

2. In particular,

$$\text{winding number for } (1 \underbrace{23\cdots123}_{23 \text{ shows for } m \text{ times}} 1) = m$$

$$\text{winding number for } (1 \underbrace{32\cdots132}_{32 \text{ shows for } n \text{ times}} 1) = -n$$

3. For any given α, it is equivalent to a unique $(123123\cdots1231)$ or $(132\cdots1321)$, since otherwise α will have different winding numbers.

Therefore, (1) and (3) shows the well-definedness of ϕ. In particular, (1) shows that as $\alpha \sim \alpha'$, we have $\phi([\alpha]) = \phi([\alpha'])$; (2) shows that the winding number of α is an unique integer.

- Homomorphism: For given any two edge loops α, β based at 1, suppose that $[\alpha] = [(1bc1bc\cdots1bc1)]$ and $[\beta] = [(1pq1pq\cdots1pq1)]$, then

$$\phi([\alpha]\cdot[\beta]) = \phi([\alpha\cdot\beta]) = [(1bc1bc\cdots1bc11pq1pq\cdots1pq1)]$$

  Discuss the case for the sign of $\phi([\alpha])$ and $\phi([\beta])$ separately gives the desired result.

- Surjectivity: for a given $m \in \mathbb{Z}$, construct α such that $\phi([\alpha]) = m$ is easy.

- Injectivity: suppose that $\phi([\alpha]) = 0$, then by definition of ϕ, $[\alpha] = [(1)] = e$, which is the trivial element in $E(K, 1)$.

Therefore, ϕ is an isomorphism. ■

R Actually, we can show that the loop based at 1 given by:

$$\begin{aligned} \ell \quad & I \longrightarrow S^1 \\ \text{with} \quad & t \mapsto e^{2\pi i t} \end{aligned}$$

is a generator for $\pi_1(S^1, 1)$:

- $\phi([\ell]) = 1$, where $\phi : \pi_1(S^1, 1) \cong \mathbb{Z}$.
- The loop

$$\begin{aligned} \ell^m : \quad & I \to S^1 \quad m \in \mathbb{Z} \\ \text{with} \quad & \ell^m(t) = e^{2\pi i m t} \end{aligned}$$

gives $\phi([\ell^m]) = m$

Corollary 13.4 [Fundamental Theorem of Algebra] All non-constant polynomials in $\mathbb{C}$ has at least one root in $\mathbb{C}$

Proof. • Suppose on the contrary that

$$p(x) = a_n x^n + \cdots + a_1 x + a_0 \; a_n \neq 0$$

has no roots, then p is a mapping from $\mathbb{C}$ to $\mathbb{C} \setminus \{0\}$. It's clear that $\mathbb{C} \setminus \{0\} \simeq \{z \in \mathbb{C} \mid |z| = 1\}$, and therefore

$$\pi_1(\mathbb{C} \setminus \{0\}) = \pi_1(S^1) \cong \mathbb{Z}.$$

- The induced homomorphism p^* of p is given by:

$$\begin{aligned} p_* : \quad & \pi_1(\mathbb{C}) \to \pi_1(\mathbb{C} \setminus \{0\}) \\ \text{with} \quad & \{e\} \mapsto \mathbb{Z} \end{aligned}$$

Note that $\pi_1(\mathbb{C})$ is trival as $\mathbb{C}$ is contractible. Also, $p_*(e) = 0$.

- Consider the inclusion from $C_r = \{z \in \mathbb{C} \mid |z| = r\}$ to $\mathbb{C}$:

$$\begin{aligned} i : \quad & C_r \to \mathbb{C} \\ \text{with} \quad & z \mapsto z \end{aligned}$$

It satisfies the diagram given below:

$$\begin{tikzcd} C_r \arrow[r, dashed, "i"] \arrow[dr, "p|_{C_r}"'] & \mathbb{C} \arrow[d, dashed, "p"] \\ & \mathbb{C} \setminus \{0\} \end{tikzcd}$$

As a result, the induced homomorphism i^* of i satisfies the diagram

$$\begin{tikzcd} \pi_1(C_r) \arrow[r, dashed, "i_*"] \arrow[dr, "(p|_{C_r})_*"'] & \pi_1(\mathbb{C}) \arrow[d, dashed, "p_*"] \\ & \pi_1(\mathbb{C} \setminus \{0\}) \end{tikzcd}$$

Or equivalently,

$$\begin{tikzcd} \mathbb{Z} \cong \pi_1(C_r) \arrow[r, dashed, "i_*"] \arrow[dr, "(p|_{C_r})_*"'] & \{e\} \arrow[d, dashed, "p_*"] \\ & \mathbb{Z} \cong \pi_1(\mathbb{C} \setminus \{0\}) \end{tikzcd}$$

Therefore, $p_* \circ i_*$ is a zero map since $p_*(e) = 0$, i.e., $(p\,|_{C_r})_*$ is a zero homomorphism.

- Then it's natural to study $p\,|_{C_r}\colon C_r \to \mathbb{C} \setminus \{0\}$. Construct

$$\begin{cases} q(z) = k \cdot z^n, \quad k := \dfrac{p(r)}{r^n} \text{is a constant} \\ p(z) = a_n z^n + \cdots + a_1 z + a_0 \end{cases}$$

Therefore, $p(r) = q(r)$, and $p|_{C_r}, q|_{C_r} : C_r \to \mathbb{C} \setminus \{0\}$.

– We claim that $p|_{C_r} \simeq q|_{C_r}$ for large r. First construct the mapping

$$\begin{aligned} H: \quad & C_r \times [0,1] \to \mathbb{C} \\ \text{with} \quad & H(z,t) = tp(z) + (1-t)q(z) \\ \text{and} \quad & H(z,0) = q(z),\ H(z,1) = p(z) \end{aligned}$$

If we want to show H is the homotopy between $p|_{C_r}$ and $q|_{C_r}$, it suffices to show that H is well-defined, i.e., $H : C_r \times [0,1] \to \mathbb{C} \setminus \{0\}$.

Suppose on the contrary that there exists (z,t) such that

$$(1-t)p(z) + tq(z) = 0, \quad |z| = r, t \in [0,1]$$

Or equivalently,

$$(1-t)(a_n z^n + \cdots + a_1 z + a_0) + t \cdot kz^n = 0.$$

Substituting k with $p(r)/r^n$ gives

$$a_n z^n + \cdots + a_1 z + a_0 = t\left(a_{n-1}z^{n-1} + \cdots + a_0 - a_{n-1}\frac{z^n}{r} - \cdots - a_1\frac{z^n}{r^{n-1}} - a_0\frac{z^n}{r^n}\right)$$

The LHS has leading order n, while the RHS has leading order less or equal to $n-1$. As $r = |z| \to \infty$, $t \to \infty$. Therefore, the equality does not hold in the range $t \in [0,1]$ when r is sufficiently large.

For this choice of $r = |z|$,

$$H : C_r \times [0,1] \to \mathbb{C} \setminus \{0\}$$

gives the homotopy $p|_{C_r} \simeq q|_{C_r}$.

- Therefore, we imply $(p|_{C_r})_* = (q|_{C_r})_*$. Now we check the mapping $(q|_{C_r})_* : \mathbb{Z} \to \mathbb{Z}$. In particular, we check the value of $(q|_{C_r})_*(1)$, where 1 is the generator in $\pi_1(C_r)$.

Here we construct the loop

$$\begin{aligned} \ell : \quad & I \to C_r \\ \text{with} \quad & \ell(t) = re^{2\pi i t} \end{aligned}$$

and therefore $[\ell] = 1$. It follows that

$$(q|_{C_r})_*(1) = (q|_{C_r})_*([\ell]) = [q|_{C_r}(\ell)] = q(re^{2\pi i t}) = k \cdot r^n \cdot e^{2\pi i n t} \neq 0.$$

Therefore, $(q|_{C_r})_*$ is not a zero homomorphism, i.e., $(q|_{C_r})_* : \mathbb{Z} \cong \pi_1(C_r) \to \pi_1(C\{0\}) \cong \mathbb{Z}$ is the map $1 \mapsto n$, which gives a contradiction.

■

Chapter 14

Week14

14.1. Monday for MAT3040

14.1.1. Multilinear Tensor Product

Definition 14.1 [Tensor Product among More spaces] Let $V_1,\dots,V_p$ be vector spaces over $\mathbb{F}$. Let $S = \{(v_1,\dots,v_p) \mid v_i \in V_i\}$ (We assume no relations among distinct elements in S), and define $\mathfrak{X} = \operatorname{span}(S)$.

1. Then define the tensor product space $V_1 \otimes \cdots \otimes V_p = \mathfrak{X}/y$, where y is the vector subspace of $\mathfrak{X}$ spanned by vectors of the form

$$(v_1,\dots,v_i + v_i',\dots,v_p) - (v_1,\dots,v_i,\dots,v_p) - (v_1,\dots,v_i',\dots,v_p),$$

   and

$$(v_1,\dots,\alpha v_i,\dots,v_p) - \alpha(v_1,\dots,v_i,\dots,v_p)$$

   where $i = 1,2,\dots,p$.

2. The tensor product for vectors is defined as

$$v_1 \otimes \cdots \otimes v_p := \{(v_1,\dots,v_p) + y\} \in V_1 \otimes \cdots \otimes V_p$$

■

(R) Similar as in tensor product among two space,

1. We have

$$v_1 \otimes \cdots \otimes (\alpha v_i + \beta v_i') \otimes \cdots \otimes v_p = \alpha(v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_p) + \beta(v_1 \otimes \cdots \otimes v_i' \otimes \cdots \otimes v_p)$$

2. A general vector in $V_1 \otimes \cdots \otimes V_p$ is

$$\sum_{i=1}^{n} (W_1^{(i)} \otimes \cdots \otimes W_p^{(i)}), \quad \text{where } W_j^{(i)} \in V_j, j = 1, \dots, p$$

3. Let $\mathcal{B}_i = \{v_i^{(1)}, \dots, v_i^{(\dim(V_i))}\}$ be a basis of $V_i, i = 1, \dots, p$, then

$$\mathcal{B} = \{V_1^{(\alpha_1)} \otimes \cdots \otimes V_p^{(\alpha_p)} \mid 1 \le \alpha_i \le \dim(V_i)\}$$

is a basis of $V_1 \otimes \cdots \otimes V_p$. As a result,

$$\dim(V_1 \otimes \cdots \otimes V_p) = (\dim(V_1)) \times \cdots \times (\dim(V_p))$$

Theorem 14.1 — Universal Property of multi-linear tensor. Let $\text{Obj} = \{\phi : V_1 \times \cdots \times V_p \to W \mid \phi \text{ is a } p\text{-linear map}\}$, i.e.,

$$\phi(v_1, \dots, \alpha v_i + \beta v_i', \dots, v_o) = \alpha\phi(v_1, \dots, v_i, \dots, v_p) + \beta\phi(v_1, \dots, v_i', \dots, v_p),$$
$$\forall v_i, v_i' \in V_i, i = 1, \dots, p, \forall \alpha, \beta \in \mathbb{F}.$$

For instance, the multiplication of p matrices is a p-linear map.

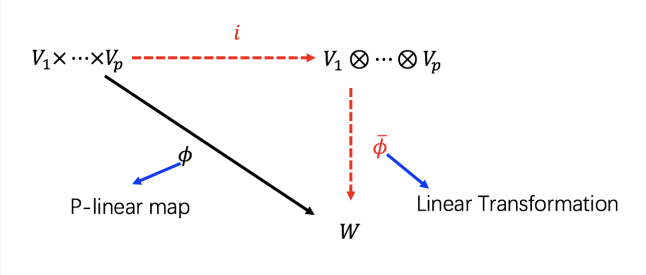
Then the mapping in the Obj,

$$\begin{aligned} i : \quad & V_1 \times V_p \to V_1 \otimes \cdots \otimes V_p \\ \text{with} \quad & (v_1, \dots, v_p) \mapsto v_1 \otimes \cdots \otimes v_p \end{aligned}$$

satisfies the universal property. In other words, for any $\phi : V_1 \times \cdots \times V_p \in \text{Obj}$, there exists the unqiue linear transformation

$$\bar{\phi} : V_1 \otimes \cdots \otimes V_p \to W$$

such that the diagram below commutes:



In other words, $\phi = \bar{\phi} \circ i$.

Corollary 14.1 Let $T_i : V_i \to V_i'$ be a linear transformation, $1 \le i \le p$. There is a unique linear transformation

$$\begin{aligned} (T_1 \otimes \cdots \otimes T_p) : \quad & V_1 \otimes \cdots \otimes V_p \to V_1' \otimes \cdots \otimes V_p' \\ \text{satisfying} \quad & (T_1 \otimes \cdots \otimes T_p)(v_1 \otimes \cdots \otimes v_p) = T_1(v_1) \otimes \cdots \otimes T_p(v_p) \end{aligned}$$

Proof. Construct the mapping

$$\begin{aligned} \phi : \quad & V_1 \times \cdots \times V_p \to V_1' \otimes \cdots \otimes V_p' \\ \text{with} \quad & (v_1, \ldots, v_p) \mapsto T_1(v_1) \otimes \cdots \otimes T_p(v_p) \end{aligned}$$

which is indeed p-linear.

By the universal property, we induce the unique linear transformation

$$\bar{\phi} : V_1 \otimes \cdots \otimes V_p \to V_1' \otimes \cdots \otimes V_p'$$

∎

Notation. To make life easier, from now on, we only consider $V_1 = \cdots = V_p = V$. Then for any linear transformation $T : V \to W$, we have

$$T^{\otimes p} : V \otimes \cdots \otimes V \to W \otimes \cdots \otimes W$$

We use the short-hand notation $V^{\otimes p}$ to denote $\underbrace{V \otimes \cdots \otimes V}_{p \text{ terms in total}}$

Final Exam Ends Here.

14.1.2. Exterior Power

Definition 14.2 A p-linear map $\phi : V \times \cdots \times V \to W$ is called **alternating** if

$$\phi(v_1, \dots, v_i, \dots, v_j, \dots, v_p) = \mathbf{0}_W, \quad \text{provided that there exists some } v_i = v_j \text{ for } i \neq j.$$

Also, we say ϕ is p-alternating

Example 14.1 1. The cross product mapping

$$\begin{aligned} \phi : \quad & \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \\ \text{with} \quad & (\boldsymbol{v}, \boldsymbol{w}) \mapsto \boldsymbol{v} \times \boldsymbol{w} \end{aligned}$$

is alternating:

- ϕ is bilinear
- $\phi(\boldsymbol{v}, \boldsymbol{v}) = \boldsymbol{v} \times \boldsymbol{v} = \mathbf{0}$.

2. The determinant mapping

$$\begin{aligned} \phi : \quad & \underbrace{\mathbb{F}^n \times \cdots \times \mathbb{F}^n}_{n \text{ terms in total}} \to \mathbb{F} \\ \text{with} \quad & (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \mapsto \det([\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n]) \end{aligned}$$

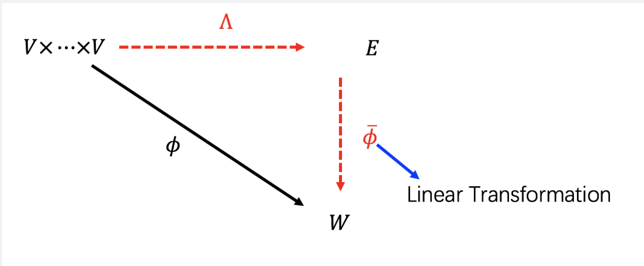
is alternating:

- ϕ is n-linear by MAT2040 knowledge
- ϕ is alternating by MAT2040 knowledge

■

Theorem 14.2 — **Universal Property for exterior power.** Let Obj $:= \{\phi : \underbrace{V \times \cdots V}_{p \text{ terms}} \to W \mid \phi$ is p-alternating map$\}$. Then there exists $\{\Lambda : V \times \cdots \times V \to E\} \in$ Obj satisfying the following:

- For all $\phi : V \times \cdots \times V \to W \in$ Obj, there exists unique linear transformation $\bar{\phi} : E \to W$ satisfying



  In other words, $\phi = \bar{\phi} \circ \Lambda$.

14.2. Monday for MAT3006

14.2.1. Tonelli's and Fubini's Theorem

Proposition 14.1 For all $E \in \mathcal{M} \otimes \mathcal{M}$,we have

$$\int m_Y(E_x)\,\mathrm{d}x = \int m_X(E_y)\,\mathrm{d}y = \pi(E), \tag{14.1}$$

where $\pi(\cdot)$ is a measure on $\mathcal{M} \otimes \mathcal{M}$.

Here note that

$$\begin{aligned} m_X(E_y) &:= \int (\mathcal{X}_E)_y(x)\,\mathrm{d}x \\ m_Y(E_x) &:= \int (\mathcal{X}_E)_x(y)\,\mathrm{d}y \end{aligned}$$

Proof. Construct

$$\mathcal{A} = \left\{ E \in \mathcal{M} \otimes \mathcal{M} \;\middle|\; \begin{array}{l} x \mapsto m_Y(E_x) \text{ measurable} \\ y \mapsto m_X(E_y) \text{ measurable} \\ \text{(14.1) holds for } E \end{array} \right\}$$

Following the proof given in the last lecture, it suffices to show $\mathcal{A}$ is a monotone class:

- Construct
$$\mathcal{A}_k = \mathcal{A} \cap \{E \in \mathcal{M} \otimes \mathcal{M} \mid E \subseteq [-k,k] \times [-k,k]\}.$$

  We first show that $\mathcal{A}_k$ is a monotone class for all $k \in \mathbb{N}$:

  1. Suppose that $E_n \subseteq E_{n+1}, \forall n$ and $E_n \in \mathcal{A}_k$, and we aim to show $E := \cup_{n=1}^{\infty} E_n \in \mathcal{A}_k$. Consider the function $f_n(x) = m_Y((E_n)_x)$, which is measurable for all n, and $f_n(x) \le f_{n+1}(x)$ for all n, since $E_n \subseteq E_{n+1}$.
  The MCT I implies that $f(x) = m_Y(E_x)$ is measurable with

$$\int m_Y(E_x)\,\mathrm{d}x = \lim_{n\to\infty} \int m_Y((E_n)_x)\,\mathrm{d}x \overset{(a)}{=} \lim_{n\to\infty} \pi(E_n) \overset{(b)}{=} \pi(E)$$

where (a) is because that $E_n \in \mathcal{A}$; and (b) is due to the exercise in Hw3. Similarly, $y \mapsto m_X(E_y)$ is measurable, with $\int m_X(E_y)\,\mathrm{d}y = \pi(E)$. Therefore, $E \in \mathcal{A}$, i.e., $E \in \mathcal{A}_k$ as well.

2. Suppose that $F_i \in \mathcal{A}_k, F_i \supseteq F_{i+1}$, and we aim to show $F := \cap_{i=1}^{\infty} F_i \in \mathcal{A}_k$. Construct the measurable function $g_n(x) = m_Y((F_n)_x)$, and $g_n(x) \ge g_{n+1}(x)$; $|g_n(x)| \le g_1(x)$, with $g_1(x)$ integrable. (You may see the bounded rectangle in $\mathcal{A}_k$ matters here)

   The DCT implies that $g(x) = m_Y(F_x)$ is measurable, with

$$\int m_Y(F_x)\,\mathrm{d}x = \lim_{n\to\infty} \int g_n\,\mathrm{d}x = \lim_{n\to\infty} \pi(F_n) = \pi(F).$$

   Similarly, $y \mapsto m_X(F_y)$ is measurable, with $\int m_X(F_y)\,\mathrm{d}y = \pi(F)$. Therefore, $F \in \mathcal{A}_k$.

Together with the results from last lecture, we conclude that claim (1) and (2) holds for $\mathcal{A}_k$. Following the similar idea of the results obtained from last lecture, we conclude that $\mathcal{A}_k = \{E \in \mathcal{M} \otimes \mathcal{M} \mid E \subseteq [-k,k] \times [-k,k]\}$.

- Then we show $\mathcal{A}$ is a monotone class, i.e., closed under countable decreasing intersections.

Suppose that $F_i \in \mathcal{A}, F_i \supseteq F_{i+1}$, we aim to show that $F := \cap F_i \in \mathcal{A}$.

Construct

$$F_i^{(k)} = F_i \cap ([-k,k] \times [-k,k]),$$

which follows that $F_i^{(k)} \supseteq F_{i+1}^{(k)}$, and $F_i^{(k)} \in \mathcal{A}_k$ since $F_i^{(k)} \in \mathcal{M} \otimes \mathcal{M}$ and $F_i^{(k)} \subseteq [-k,k] \times [-k,k]$. We denote $F^{(k)} = \cap_{i=1}^{\infty} F_i^{(k)}$. The previous result implies that $F^{(k)} \in \mathcal{A}_k$, i.e.,

$$\int m_Y((F^{(k)})_x)\,\mathrm{d}x = \pi(F^{(k)})$$

Now note that $F^{(1)} \subseteq F^{(2)} \subseteq \cdots$, and $F = \cup_{k\in\mathbb{N}} F^{(k)}$. Therefore, applying MCT gives

$$\int m_Y(F_x)\,\mathrm{d}x = \lim_{k\to\infty} \int m_Y((F^{(k)})_x)\,\mathrm{d}x = \lim_{k\to\infty} \pi(F^{(k)}) = \pi(F).$$

Therefore, F satisfies (14.1), i.e., $F \in \mathcal{A}$

■

Theorem 14.3 — **Tonelli's Theorem.** Let $F : \mathbb{R}^2 \to [0, \infty]$ be measurable under the space $(\mathbb{R}^2, \mathcal{M} \otimes \mathcal{M}, \pi)$. Then

$$\begin{cases} x \mapsto \int F(x,y)\,\mathrm{d}y \\ y \mapsto \int F(x,y)\,\mathrm{d}x \end{cases} \text{is measurable,}$$

and

$$\int F\,\mathrm{d}\pi = \int \left(\int F(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int \left(\int F(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

Proof. Let

$$\phi_n(x,y) = \sum_{k=0}^{4^n} (k \cdot 2^{-n}) \mathcal{X}_{F^{-1}([k \cdot 2^{-n}, (k+1) \cdot 2^{-n}])} + 2^n \mathcal{X}_{F^{-1}(2^n, \infty]}$$

We just re-write the terms above as $\sum_k \alpha_k \mathcal{X}_{E_k}$. Our constructed $\phi_n(x,y)$ is a monotone increasing simple function such that $\phi_n \to F$ pointwise. It follows that

$$\begin{aligned}
\int F\,\mathrm{d}\pi &= \lim_{n\to\infty} \int \phi_n\,\mathrm{d}\pi && (14.2a)\\
&= \lim_{n\to\infty} \int \left(\sum_k \alpha_k \mathcal{X}_{E_k} \right) \mathrm{d}\pi && (14.2b)\\
&= \lim_{n\to\infty} \sum_k \alpha_k \int \mathcal{X}_{E_k}\,\mathrm{d}\pi = \lim_{n\to\infty} \sum_k \alpha_k \pi(E_k) && (14.2c)\\
&= \lim_{n\to\infty} \sum_k \alpha_k \int \left(\int \mathcal{X}_{E_k}(x,y)\,\mathrm{d}x \right) \mathrm{d}y && (14.2d)\\
&= \lim_{n\to\infty} \int \int \left(\sum_k \alpha_k \mathcal{X}_{E_k}(x,y) \right) \mathrm{d}x\,\mathrm{d}y && (14.2e)\\
&= \lim_{n\to\infty} \int \left(\int \phi_n(x,y)\,\mathrm{d}x \right) \mathrm{d}y && (14.2f)\\
&= \int \lim_{n\to\infty} \left(\int \phi_n(x,y)\,\mathrm{d}x \right) \mathrm{d}y && (14.2g)\\
&= \int \int \lim_{n\to\infty} \phi_n(x,y)\,\mathrm{d}x\,\mathrm{d}y && (14.2h)\\
&= \int \int F(x,y)\,\mathrm{d}x\,\mathrm{d}y && (14.2i)
\end{aligned}$$

where (14.2a) is by the MCT I on ϕ_n; (14.2c) is by the linearity of integral; (14.2d) is by proposition (14.1) (14.2e) is by the linearity of integral; (14.2g) is by the MCT I on $f_n(y) = \int \phi_n(x,y)\,\mathrm{d}x$; (14.2h) is by the MCT I on $g_n(x) = \phi_n(x,y)$; (14.2i) is because that $\phi_n(x,y) \to F(x,y)$. ■

Theorem 14.4 — Fubini's Theorem. Suppose that $F : \mathbb{R}^2 \to [-\infty,\infty]$ is integrable, then

$$\int F\,\mathrm{d}\pi = \int \left(\int F(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int \left(\int F(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

Proof. Suppose $F = F^+ - F^-$, where $F^\pm$ are both integrable. Applying Tonell's theorem on both F^- and F^+ and the linearity of integrals gives the desired result. ■

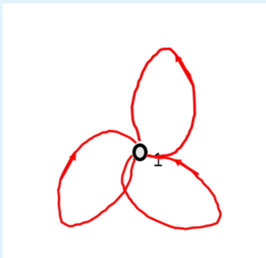
14.3. Monday for MAT4002

14.3.1. Fundamental group of a Graph

Definition 14.3 [Graph] A graph $T = (V, E)$ is defined by the following components:

- V is a finite or countable set, called vertex set;
- E is a finite or countable set, called edge set;
- A function $\delta : E \to V \times V$ with $\delta(e) = (\ell(e), \tau(e))$, where $\ell(e), \tau(e)$ is known as the endpoints of e.

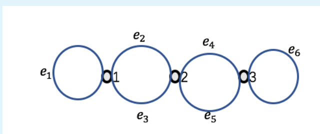
■ **Example 14.2** 1. Let $V = \{1\}, E = \{e_1, e_2, e_3\}$, and define $\delta(e_i) = (1,1), i = 1,2,3$. The graph (V, E) is represented below:



2. Let $V = \{e_1, e_2, e_3\}$ and $E = \{e_1, \dots, e_6\}$, and define

$$\delta(e_1) = (1,1), \quad \delta(e_2) = (1,2), \quad \delta(e_3) = (1,2),$$
$$\delta(e_4) = (2,3), \quad \delta(e_5) = (2,3), \quad \delta(e_6) = (3,3).$$

The graph (V, E) is represented below (We do not care the direction of edges for this graph):

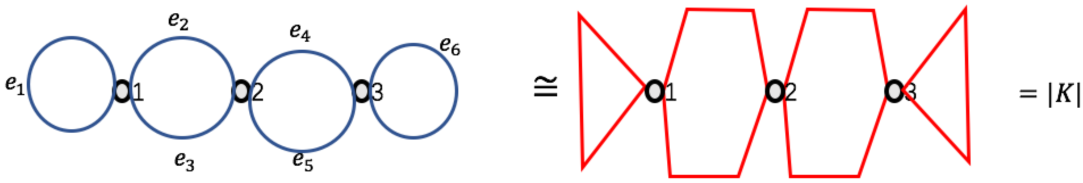


Definition 14.4 [Realizatin of a Graph] For a given graph $\Gamma = (V, E)$, construct a realization by

$$\{|V| \times \{\text{zero simplies}\} \coprod |E| \times \{\text{1-simplies}\}\}/\sim$$

where the equivalence class is induced from the function δ. We still call this realization of the graph as Γ.

In general, graphs are not simplicial complexes. But we can "sub-divide" each edge of Γ into three parts such that there exists simplicial complex K with $|K| \cong \Gamma$. For instance,



where $|K|$ is a simplicial complex.

Definition 14.5

- Subgraph $\Gamma' \subseteq \Gamma$: $\Gamma' = (V', E')$ with $V' \subseteq V$ and $E' \subseteq E$, and

$$\delta\mid_{V'} : E' \to V' \times V'$$

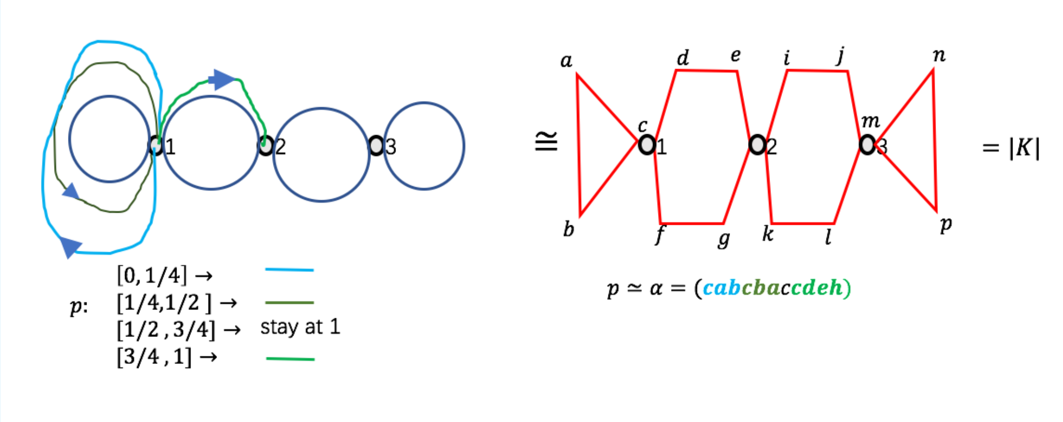
- Edge path: A continous function $p : [0,1] \to \Gamma$ such that there exists $n \in \mathbb{N}$ satisfying

$$p\mid_{[i/n, i+1/n]} : \left[\frac{i}{n}, \frac{i+1}{n}\right] \to T$$

is a path along an edge of Γ, or a constant function on a vertex of Γ, for $0 \le i \le n-1$.

Under the homeomorphism $\Gamma \cong |K|$, each edge path is homotopic to

$|g_\alpha|$ for some edge path α in the simplicial complex K. For instance,



- An Edge loop is an edge path p such that $p(0) = p(1) = b \in V$.
- Embedded Edge Loop: An injective edge loop, i.e., $p : [0,1] \longrightarrow \Gamma$ such that

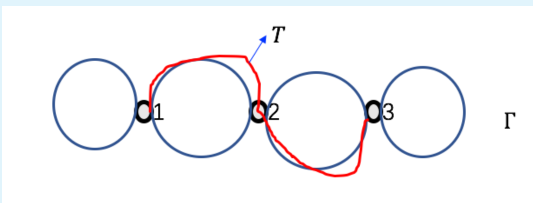
$$\text{for } x \notin V, \quad p^{-1}(x) = \emptyset \text{ or a single point.}$$

- Tree: a connected graph T that contains no embedded edge loop $p : [0,1] \longrightarrow T$.

  For instance, as shown in the figure, T_1 contains no edge loop, in particular, the edge loop (a,b,a) is not embedded; T_2 contains embedded edge loop (a,b,c,d,a).

- Maximal Tree of a connected graph Γ:
  - A subgraph T of Γ such that T is a tree.
  - By adding an edge $e \in E(\Gamma) \setminus E(T)$ into T, the new graph is no longer a tree.

  For instance, $T \subseteq \Gamma$ shown in the figure below is a maximal tree.



Theorem 14.5 Let Γ be a connected graph, and T is a subgraph of Γ such that T is a tree. Then T is a maximal tree if and only if $V(T) = V(\Gamma)$.

Moreover, there always exists a maximal tree for all Γ.

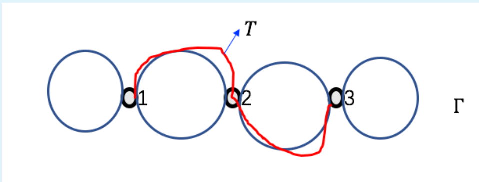
Proof Outline for second part. Construct an ordering of $\{v_1, \ldots, v_i\} \subseteq V(\Gamma)$, such that for each integer $i \geq 2$, there is an edge connecting v_{i+1} with some vertex in $\{v_1, \ldots, v_i\}$.

Then construct $T_1 \subseteq T_2 \subseteq \cdots$, where T_i is a tree containing vertices $\{v_1, \ldots, v_i\}$. As a result, $T = \cup_{i \in \mathbb{N}} T_i$ is a maximal tree. ■

Theorem 14.6 Let Γ be a connected graph. Then $\pi(\Gamma)$ is isomorphic to the free group generated by $\#\{E(\Gamma) \setminus E(T)\}$ elements, for any maximal tree of Γ.

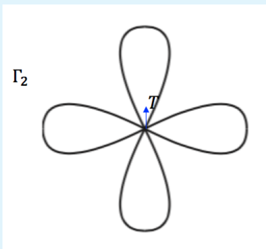
▪ **Example 14.3**

1. The graph $T \subseteq \Gamma_1$ shown in the figure below is a maximal tree.



   Therefore, $\pi_1(\Gamma_1) \cong \langle a, b, c, d \rangle$ since $\#\{E(\Gamma_1) \setminus E(T)\} = 4$.

2. The graph $T \subseteq \Gamma_2$ shown in the figure below is a maximal tree.



   Therefore, $\pi_1(\Gamma_2) \cong \langle a, b, c, d \rangle$ since $\#\{E(\Gamma_2) \setminus E(T)\} = 4$.

3. Note that $\Gamma_1 \simeq \Gamma_2$. The reason for such homotopy equivalence is in the link

   `https://www.math3ma.com/blog/clever-homotopy-equivalences`

Chapter 15

Week15

15.1. Monday for MAT3040

15.1.1. More on Exterior Power

Reviewing. Let $\text{Obj} := \{\phi : V \times \cdots \times V \to W \mid \phi \text{ is alternating}\}$, then there exists

$$\{\Lambda : V \times \cdots \times V \to E\} \in \text{Obj}$$

such that

$$\phi = \bar{\phi} \circ \Lambda, \quad \text{where } \bar{\phi} : E \to W \text{ is the unique linear transformation}$$

Here we give one way for constructing E:

$$E = V^{\otimes p}/U,$$

where U is spanned by vectors of the form

$$v_1 \otimes \cdots \otimes v_p \in V^{\otimes p}, \quad v_i = v_j \text{ where for some } i \neq j.$$

For instance, $v \otimes v \otimes \cdots \otimes v_p \in U$.

Definition 15.1 [Wedge Product] Define the wedge product space

$$\wedge^p V := V^{\otimes p}/U = E,$$

with the wedge product among vectors

$$v_1 \wedge \cdots \wedge v_p = v_1 \otimes \cdots \otimes v_p + U \in \wedge^p V$$

As a result, the mapping

$$\begin{aligned} \wedge : \quad & V \times \cdots \times V \to E := \wedge^p V \\ & (v_1, \ldots, v_p) \mapsto v_1 \wedge \cdots \wedge v_p \end{aligned}$$

will satisfy the universal property of exterior power.

Proposition 15.1

1. We have the p-linearity for $\wedge^p V$, i.e.,

$$v_1 \wedge \cdots \wedge (av_i + bv_i') \wedge \cdots \wedge v_p = a(v_1 \wedge \cdots \wedge v_i \wedge \cdots \wedge v_p) + b(v_1 \wedge \cdots \wedge v_i' \wedge \cdots \wedge v_p)$$

for $i = 1, \ldots, p$.

2. The wedge product is alternating:

$$\begin{aligned} v_1 \wedge \cdots \wedge v \wedge \cdots \wedge v \wedge \cdots \wedge v_p &:= v_1 \otimes \cdots \otimes v \otimes \cdots \otimes v \otimes \cdots \otimes v_p + U \\ &= 0 + U \\ &= 0_{\wedge^p V} \end{aligned}$$

3. The wedge product reverses sign reversal property:

$$v_1 \wedge \cdots \wedge v \wedge \cdots \wedge w \wedge \cdots \wedge v_p = -v_1 \wedge \cdots \wedge w \wedge \cdots \wedge v \wedge \cdots \wedge v_p$$

Reason: $(v + w) \wedge (v + w) = 0$, which implies $v \wedge w + w \wedge v = 0$.

Proposition 15.2 1. If $\dim(V) = n$, and $0 \le p \le n$, then

$$\dim(\wedge^p V) = \binom{n}{p}$$

2. For all linear operators $T : V \to V$, there is an unique linear operator from $\wedge^p V$ to $\wedge^p V$:

$$\begin{aligned} T^{\wedge^p} : \quad & \wedge^p V \to \wedge^p V \\ \text{with} \quad & v_1 \wedge \cdots \wedge v_p \mapsto T(v_1) \wedge \cdots \wedge T(v_p) \end{aligned}$$

Proof. 1. Let $\{v_1, \dots, v_n\}$ be basis of V, then $\{v_{i_1} \otimes \cdots \otimes v_{i_p} \mid 1 \le i_k \le n\}$ forms basis of $V^{\otimes p}$. Note that $\{v_{i_1} \wedge \cdots \wedge v_{i_p} \mid 1 \le i_k \le n\}$ spans $\wedge^p V$, since $\pi_V : V \to V/U$ is surjective. We claim that

$$\mathcal{B} = \{v_{i_1} \wedge \cdots \wedge v_{i_p} \mid 1 \le i_1 < i_2 < \cdots < i_p \le n\}$$

is a basis of $\wedge^p V$

- $\mathcal{B}$ spans $\wedge^p V$: we can use (3) in proposition (15.1) to "rearrange" the indices $j_1, \dots, j_p$ into ascending order, and $\text{span}(\mathcal{B}) = \text{span}\{v_{i_1} \wedge \cdots \wedge v_{i_p} \mid 1 \le i_k \le n\}$.
- We omit the proof that $\mathcal{B}$ is linear independent due to time limit.

The numbre of vectors in $\mathcal{B}$ is equal to $\binom{n}{p}$.

■

15.1.2. Determinant

Previous Approach for defining determinant. We define the determinant for $\boldsymbol{A} = M_{n\times n}(\mathbb{F})$ directly. From such complicated definition, we come up with $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$, which implies that the similar matrices share with the same determinant, then we define the determinant for any linear operator $T : V \to V$ as

$$\det(T) = \det((T)_{\mathcal{B},\mathcal{B}}), \quad \text{for some basis } \mathcal{B} \text{ of } T$$

New Approach. We will define $\det(T)$ for linear operators without fixing a basis, and then we will imply $\det(T \circ S) = \det(T)\det(S)$ easily. Then $\det(\boldsymbol{A})$ for $\boldsymbol{A} \in M_{n\times n}(\mathbb{F})$ belongs to our special case.

Definition 15.2 [Determinant for Linear Operators]

1. Suppose that $\dim(V) = n$, then
$$\dim(\wedge^n V) = \binom{n}{n} = 1$$
More precisely, for any basis $\{v_1, \ldots, v_n\}$ of V, we have $\wedge^n(V) = \text{span}\{v_1 \wedge \cdots \wedge v_n\}$.
2. Note tht $T^{\wedge^n} : \wedge^n V \to \wedge^n V$ is a linear operator on $\wedge^n V \cong \mathbb{F}$. Therefore, for all $\tau \in \wedge^n V$, there exists $\alpha_T \in \mathbb{F}$ such that
$$T^{\wedge^n}(\tau) = \alpha_T \tau$$
3. Now we define
$$\det(T) = \alpha_T$$
This definition of determinant does not depend on any choice of basis of V.

■

▪ **Example 15.1**

1. Suppose that $T = I : V \to V$ be identity. Take a basis $\{v_1, \ldots, v_n\}$ of V, then
$$T^{\wedge^n}(v_1 \wedge \cdots \wedge v_n) = T(v_1) \wedge \cdots \wedge T(v_n)$$
Or equivalently,
$$\det(T) \cdot (v_1 \wedge \cdots \wedge v_n) = v_1 \wedge \cdots \wedge v_n$$
Therefore, $\det(T) = 1$.
2. Suppose that $T : V \to V$ is diagonalizable with $\{w_1, \ldots, w_n\}$ forming eigen-basis of T.

As a result,

$$T^{\wedge^n}(w_1 \wedge \cdots \wedge w_n) = T(w_1) \wedge T(w_2) \cdots \wedge T(w_n),$$

which implies

$$\det(T)(w_1 \wedge \cdots \wedge w_n) = (\lambda_1 w_1) \wedge \cdots \wedge (\lambda_n w_n),$$

which implies

$$\det(T) w_1 \wedge \cdots \wedge w_n = (\lambda_1 \cdots \lambda_n) w_1 \wedge \cdots \wedge w_n,$$

i.e., $\det(T) = \lambda_1 \cdots \lambda_n$.

■

Proposition 15.3 Let $T, S : V \to V$ be linear transformations, then

$$\begin{aligned} &(T \circ S)^{\wedge^p} : \quad \wedge^p V \longrightarrow \wedge^p V \\ &\text{with} \qquad T^{\wedge^p}, S^{\wedge^p} : \wedge^p V \longrightarrow \wedge^p V \end{aligned}$$

satisfies

$$(T \circ S)^{\wedge^p} = (T^{\wedge^p}) \circ (S^{\wedge^p})$$

Proof. Pick any basis $\{v_{i_1} \wedge \cdots \wedge v_{i_p} \mid 1 \le i_1 < \cdots < i_p \le n\}$ of $\wedge^p V$. Then

$$(T \circ S)^{\wedge^p}(v_{i_1} \wedge \cdots \wedge v_{i_p}) = (T \circ S)(v_{i_1}) \wedge \cdots \wedge (T \circ S)(v_{i_p})$$

On the other hand,

$$\begin{aligned} (T^{\wedge^p}) \circ (S^{\wedge^p})(v_{i_1} \wedge \cdots \wedge v_{i_p}) &= (T^{\wedge^p})(S(v_{i_1}) \wedge \cdots \wedge S(v_{i_p})) \\ &= (T \circ S)(v_{i_1}) \wedge \cdots (T \circ S)(v_{i_p}) \end{aligned}$$

■

Corollary 15.1

$$\det(T \circ S) = \det(T)\det(S)$$

Proof. Pick any basis $\{v_1 \wedge \cdots \wedge v_n\}$ of $\wedge^n v$, then

$$\begin{aligned}
\det(T \circ S) v_1 \wedge \cdots \wedge v_n &= (T \circ S)^{\wedge^n} v_1 \wedge \cdots \wedge v_n \\
&= (T^{\wedge^n}) \circ ((S^{\wedge^n}) v_1 \wedge \cdots \wedge v_n) \\
&= (T^{\wedge^n})(\det(S) v_1 \wedge \cdots \wedge v_n) \\
&= \det(S) T^{\wedge^n}(v_1 \wedge \cdots \wedge v_n) \\
&= \det(S)\det(T) v_1 \wedge \cdots \wedge v_n
\end{aligned}$$

Therefore, $\det(T \circ S) = \det(T)\det(S)$.

∎

Theorem 15.1 Let $V = \mathbb{F}^n$, and

$$\begin{aligned}
T: &\quad V \to V \\
\text{with} &\quad T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}, \quad \boldsymbol{A} \in M_{n \times n}(\mathbb{F})
\end{aligned}$$

Then $\det(T) = \det(\boldsymbol{A})$

Proof. Take $\{e_1, \ldots, e_n\}$ as the usual basis of $V \equiv \mathbb{F}^n$, then

$$\begin{aligned}
\det(T) e_1 \wedge \cdots \wedge e_n &= T(e_1) \wedge \cdots T(e_n) \\
&= a_1 \wedge \cdots \wedge a_n
\end{aligned}$$

where a_i denotes the i-th column of $\boldsymbol{A}$.

As we have studied before [c.f. p141 in MAT2040 Notebook], the previous definition of determinant is based on three basic properties. It suffices to show these three basis properties:

1. The determinant of the n by n identity matrix is 1: See part (1) in Example (15.1)

2. The determianant changes sign when two columns (w.l.o.g., "rows" are relaced with "columns") are exchanged: due to the sign reversal property for wedge product
3. The determinant is a linear function of each column separately, i.e.,

$$a_1 \wedge \cdots \wedge (ta_i) \wedge \cdots \wedge a_n = t(a_1 \wedge \cdots \wedge a_i \wedge \cdots \wedge a_n)$$

Once we verify these three properties, we conclude that the explicit formula for $\det(\boldsymbol{A})$ is a special case for our new definition. ■

Or we can come into the previous definition for determinant directly. For instance, consider the mapping

$$\begin{aligned} T: &\quad \mathbb{R}^2 \to \mathbb{R}^2 \\ \text{with} &\quad T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

Then we imply

$$\begin{aligned} \det(T)(e_1 \wedge e_2) = \begin{pmatrix} a \\ c \end{pmatrix} \wedge \begin{pmatrix} b \\ d \end{pmatrix} \\ &= (ae_1) \wedge (be_1) + (ae_1) \wedge (de_2) + (ce_2) \wedge (de_1) + (ce_2) \wedge (de_2) \\ &= (ad)e_1 \wedge e_2 + (bc)e_2 \wedge e_1 \\ &= (ad - bc)e_1 \wedge e_2 \end{aligned}$$

Therefore, we imply $\det(T) = ad - bc$.

15.2. Monday for MAT3006

15.2.1. Applications on the Tonell's and Fubini's Theorem

Theorem 15.2 — **Tonell.** Let $f:\mathbb{R}^2\to[0,\infty]$ be a measurable function (i.e., $f^{-1}((a,\infty])\in \mathcal{M}\otimes\mathcal{M}$), then

$$\int f\,\mathrm{d}\pi=\int\left(\int f(x,y)\,\mathrm{d}x\right)\mathrm{d}y=\int\left(\int f(x,y)\,\mathrm{d}y\right)\mathrm{d}x$$

Theorem 15.3 — **Fubini.** Let $f:\mathbb{R}^2\to[-\infty,\infty]$ be integrable (i.e., $f=f^+-f^-$ with $f^\pm:\mathbb{R}^2\to[0,\infty]$ measurable and $\int f^\pm\,\mathrm{d}x<\infty$), then

$$\int f\,\mathrm{d}\pi=\int\left(\int f(x,y)\,\mathrm{d}x\right)\mathrm{d}y=\int\left(\int f(x,y)\,\mathrm{d}y\right)\mathrm{d}x$$

Corollary 15.2 Suppose that $f:\mathbb{R}^2\to[-\infty,\infty]$ is measurable, and either

$$\int\left(\int|f(x,y)|\,\mathrm{d}x\right)\mathrm{d}y \tag{15.2a}$$

or

$$\int\left(\int|f(x,y)|\,\mathrm{d}y\right)\mathrm{d}x \tag{15.2b}$$

exists, then f is integrable, and the result of Fubini follows. (i.e., one can switch the order of integration as long as the integral of $|f|$ exists)

Proof. If (15.2a) or (15.2b) exists (is finite), then Tonell's Theorem implies that $|f|$ is integrable, which implies f is integrable.

Therefore, the assumption of Fubini's theorem holds, and the proof is comptete. ∎

(R) The advantage for corollary (15.2) is that computing (15.2a) or (15.2b) is easier than showing the integrability of f in general.

■ **Example 15.2** Compute the double integral

$$\int_0^1 \int_0^x \sqrt{\frac{1-y}{x-y}}\,\mathrm{d}y\,\mathrm{d}x.$$

Construct the function $f(x,y) := \sqrt{\frac{1-y}{x-y}}\mathcal{X}_E(x,y)$, with E shown in Fig. (15.1).

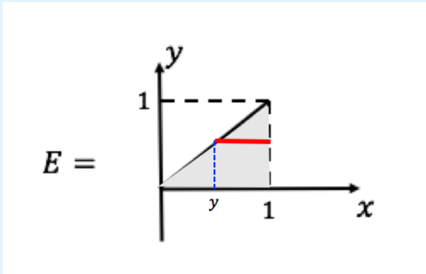


Figure 15.1: Illustration for integral domain E

We want to compute $\int f(x,y)\,\mathrm{d}\pi$ and show that

$$\int_0^1 \int_0^x \sqrt{\frac{1-y}{x-y}}\,\mathrm{d}y\,\mathrm{d}x = \int f(x,y)\,\mathrm{d}\pi.$$

- Consider the integral

$$\begin{aligned}
\int\left(\int f(x,y)\,\mathrm{d}x\right)\mathrm{d}y &= \int_0^1\left(\int_y^1 \sqrt{\frac{1-y}{x-y}}\,\mathrm{d}x\right)\mathrm{d}y && (15.3a)\\
&= \int_0^1 \sqrt{1-y}\left(\int_y^1 \frac{1}{\sqrt{x-y}}\,\mathrm{d}x\right)\mathrm{d}y && (15.3b)\\
&= \int_0^1 \sqrt{1-y}\left(\int_0^{1-y} \frac{1}{\sqrt{t}}\,\mathrm{d}t\right)\mathrm{d}y && (15.3c)\\
&= \int_0^1 \sqrt{1-y}\cdot(2\sqrt{1-y})\,\mathrm{d}y && (15.3d)\\
&= 2\int_0^1 (1-y)\,\mathrm{d}y && (15.3e)\\
&= 1 && (15.3f)
\end{aligned}$$

where the justification of (15.4a) is from Fig. (15.1).

- Therefore, $\int(\int |f(x,y)|\,\mathrm{d}x)\,\mathrm{d}y < \infty$. Mreover, f is continous on E°, i.e., measurable on E° (it's clear that a continous function is measurable). Since ∂E is null, we imply f is measurable on $E := E^\circ \cup \partial E$.
- Therefore, the assumption of Corollary (15.2) holds, and we imply that

$$\int\left(\int f(x,y)\,\mathrm{d}y\right)\mathrm{d}x = \int\left(\int f(x,y)\,\mathrm{d}x\right)\mathrm{d}y$$

It's clear that

$$\int_0^1\int_0^x \sqrt{\frac{1-y}{x-y}}\,\mathrm{d}y\,\mathrm{d}x = \int\left(\int f(x,y)\,\mathrm{d}y\right)\mathrm{d}x,$$

and therefore

$$\int_0^1\int_0^x \sqrt{\frac{1-y}{x-y}}\,\mathrm{d}y\,\mathrm{d}x = 1.$$

■

Process of Completion. We have two measures on $\mathbb{R}^2$:

- $\mathcal{M}\otimes\mathcal{M}$, and
- $\mathcal{M}_{\mathbb{R}^2}$, given by

$$\mathcal{M}_{\mathbb{R}^2} = \{E \subseteq \mathbb{R}^2 \mid m^*(A) = m^*(A\cap E) + m^*(A\cap E^c) \text{ for all subsets } A \subseteq \mathbb{R}^2\}$$

Here $\mathcal{M}_{\mathbb{R}^2}$ equals the completion of $\mathcal{M}\otimes\mathcal{M}$, i.e., all $E \subseteq \mathcal{M}_{\mathbb{R}^2}$ can be decomposed as

$$E = B \cup (E \setminus B),$$

where $B \in \mathcal{M}\otimes\mathcal{M}$ and $E\setminus B \in \mathcal{M}_{\mathbb{R}^2}$ with $\pi(E\setminus B) = 0$.

Question: does Tonell's theorem holds for (Lebesgue) measurable functions $f : \mathbb{R}^2 \to [0,\infty]$ (i.e., $f^{-1}((a,\infty]) \in \mathcal{M}_{\mathbb{R}^2}$ for any $a \in [0,\infty)$?)

Answer: Yes. To see so, we just need the following proposition

Proposition 15.4 Let $(\mathbb{R}^2, \mathcal{M}_{\mathbb{R}^2}, \pi)$ be the Lebesgue measure on $\mathbb{R}^2$, and $N \in \mathcal{M}_{\mathbb{R}^2}$ be such that $\pi(N) = 0$. Then for almost all values of $x \in \mathbb{R}$, $N_x \in \mathcal{M}$ and $m_Y(N_x) = 0$.

Proof. For $N \in \mathcal{M}_{\mathbb{R}^2}$. By hw3, there exists $B' \in \mathcal{M} \otimes \mathcal{M}$ such that $N \subseteq B'$, with

$$\pi(B') = \pi(N).$$

If N is null, then $\pi(B') = 0$. By Tonell's theorem on $M \otimes \mathcal{M}$, we imply

$$\pi(B') = \int m_Y(B'_x)\,\mathrm{d}x = \int m_X(B'_y)\,\mathrm{d}y = 0$$

Therefore, $m_Y(B'_x) = 0$ for almost all $x \in \mathbb{R}$. Since $N \subseteq B'$, we imply $N_x \subseteq B'_x$, i.e., N_x is also a null set. Therefore, $N_x \in \mathcal{M}$ and $m_Y(N_x) = 0$. ■

■ **Example 15.3** Consider the integral

$$\int_0^\infty \int_0^\infty y e^{-y^2(1+x^2)}\,\mathrm{d}y\,\mathrm{d}x$$

Define $f(x,y) = ye^{-y^2(1+x^2)}$, which is continous on $(0,\infty) \times (0,\infty)$, and therefore measurable. It follows that

$$\begin{aligned}
\int_0^\infty \int_0^\infty f(x,y)\,\mathrm{d}y\,\mathrm{d}x &= \int_0^\infty \left(\lim_{n\to\infty} \int_0^n f(x,y)\,\mathrm{d}y \right) \mathrm{d}x && (15.4a)\\
&= \int_0^\infty \left(\frac{1}{1+x^2}\frac{1}{2} \right) \mathrm{d}x && (15.4b)\\
&= \lim_{n\to\infty} \int_0^n \frac{1}{2}\frac{1}{1+x^2}\,\mathrm{d}x && (15.4c)\\
&= \frac{\pi}{4} && (15.4d)
\end{aligned}$$

where (15.4a) is by applying MCT I on the function $f(x,y)\mathcal{X}_{[0,n]}$; (15.4b) and (15.4d) is by computation; (15.4c) is by applying MCT I on the function $\frac{1}{1+x^2}\frac{1}{2}\mathcal{X}_{[0,n]}$.

By corollary (15.2),

$$\int_0^\infty \int_0^\infty y e^{-y^2(1+x^2)}\,\mathrm{d}x\,\mathrm{d}y = \frac{\pi}{4}$$

Or equivalently,

$$\int_0^\infty ye^{-y^2}\int_0^\infty e^{-x^2y^2}\,\mathrm{d}x\,\mathrm{d}y = \frac{\pi}{4}$$

By applying MCT I on $e^{-x^2y^2}\mathcal{X}_{[0,n]}$, we have

$$\int_0^\infty ye^{-y^2}\lim_{n\to\infty}\int_0^n e^{-x^2y^2}\,\mathrm{d}x\,\mathrm{d}y = \frac{\pi}{4}$$

By change of variable with $t = xy$, we imply

$$\int_0^\infty ye^{-y^2}\lim_{n\to\infty}\frac{1}{y}\int_0^{ny} e^{-t^2}\,\mathrm{d}t\,\mathrm{d}y = \frac{\pi}{4}$$

Or equivalently,

$$\int_0^\infty e^{-y^2}\int_0^\infty e^{-t^2}\,\mathrm{d}t\,\mathrm{d}y = \frac{\pi}{4}$$

Therefore, we conclude that

$$\left(\int_0^\infty e^{-y^2}\,\mathrm{d}y\right)^2 = \frac{\pi}{4} \implies \int_0^\infty e^{-x^2}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2}$$

15.2.2. Final Review

Test: Tuesday, 9:00 - 11:30 am

Venue: TD105

Except ODE, i.e., Picard-Lindorf Theorem.

Contraction theorem and Weierstrass theorem will be tested.

Lucky: choose three oringinal problems from tutorial (question: in BB?)

Theorem mentioned today, how to use, and remember how to proof.

Some problems, Fotou;s lemma, MCT, DCT, Stone-weierstrasss theorem.

How say? how to proof?

8 problems at most.

1. Part I: metric space, and set theorem.

   - a metric $d : X^2 \to \mathbb{R}^+ \cup \{0\}$. Properties:

- Non-negativity, i.e., $d \geq 0$, and $d(x,y) = 0$ iff $x = y$.
- Symmetry: $d(x,y) = d(y,x)$
- Triangle inequality: $d(x,y) \leq d(x,z) + d(z,y)$

- An open set $G \subseteq X$: for all $x \in G$, there exists $\rho > 0$, such that $B_\rho(x) = \{y \in X \mid d(y,x) < \rho\} \subseteq G$.

  An interior point $x_0 \in G$: for some $\varepsilon > 0$, $B_\varepsilon(x_0) \subseteq G$.

  Important: If all the x in G are interior points, then G is open

- A closed set G^c, when G is open

- A limit point $x_0 \in G^c$: for all $\varepsilon > 0$, there exists $y \in X$ and $y \neq x_0$ such that

$$y \in B_\varepsilon(x_0)$$

  Important: If all the points in G^c are limit points, then G^c is closed.

- A compact set $K \subseteq X$: every open cover of K has a finite subcover.

  In Euclidean space ($\mathbb{R}^n$), when K is closed and bounded

  Important: In metric space (X,d), every sequence $\{x_n\}$ in K has a subsequence $\{x_{n_i}\} \to x \in K$ as $i \to \infty$

- A complete space X: all the Cauchy sequence converges in X.

  Counter-example for in-complete space

- Compactness implies completeness, and further completeness implies closedness

- Important: Stone-Weierstrass Theorem: what is Stone-Weierstrass approximation Theorem. Not full version, but approximation theorem.

  A sub-algebra: $\mathcal{A} \subseteq C(X)$.

  $\mathcal{A}$ is dense in $C(X)$ (X is compact), i.e., $\bar{\mathcal{A}} = C(X)$.

  Equivalent to say $\mathcal{A}$ is equipped with two properties:

  - separation property
  - non-vanishing property

Weierstrass Approximation Theorem, read the proof for polynomials

- Arzela-Ascoli Theorem: for a closed subset $\mathcal{F} \subseteq C(K)$, where K is compact in $\mathbb{R}^n$, then $\mathcal{F}$ is compact, i.e., $\mathcal{F}$ is bounded and equi-continuous.

  Equi-continuity: for all $\varepsilon > 0$, there exists $\delta > 0$ such that

  $$|f(x) - f(y)| < \varepsilon, \quad \text{provided that } |x - y| < \delta, f \in \mathcal{F}$$

  how to find δ.

- Baire-Category Theorem: A subset sequence $\{E_n\}$ of complete (X, d) is nowhere dense implies that $\cup_n E_n$ has empty interior.

2. Measure Theory:

- Outer measure: $m^*(\cdot)$, other kind of measure, definition

  $$m^*(E) = \inf\left\{\sum_n m(I_n) \middle| E \subseteq \bigcup_{n=1}^{\infty} I_n \text{ open rectangles}\right\}$$

  For Euclidean space, $m(I_n) = \prod_{i=1}^n (b_i^{(n)} - a_i^{(n)})$ and $I_n =$.

- Inner measure:

  $$m_*(E) = \sup\left\{\sum_n m(I_n) \middle| \sqcup_{n=1}^{\infty} I_n \subseteq E, \quad I_n\text{'s are disjoint union}\right\}$$

  relate to Riemann integral

- Lebesgue measurable: Carthedory extension theorem

  $$m^*(S) = m^*(S \cap E) + m^*(S \cap E^c) \implies m^*(E) = m(E)$$

  (Important) Equivalent proposition: For any $A \subseteq E$ and $B \in E^c$,

  $$m^*(A \cup B) = m^*(A) + m^*(B)$$

  Examples of non-measurable: Vitali Set

- Measurable Function: $f : X \to Y$,

$$f^{-1}(E) = \{x \in X \mid f(x) \in E\} \in \mathcal{F}, \forall E \in \mathcal{G},$$

where $\mathcal{F}$ and $\mathcal{G}$ are σ-algebras w.r.t X, Y.

Lebesgue Measurable (easier): $f : X \to \mathbb{R}$.

$$\{x \in X : f(x) > a\} \text{ is measurable for } \forall a \in \mathbb{R}$$

Equivalent to $\{x \in X \mid f(x) \leq a\}$.

- Important: almost everywhere convergence: $f_n \to f$ a.e., i.e., $m\{x \in X : \lim_{n\to\infty} f_n(x) \neq f(x)\} = 0$.

Convergent in measure: $f_n \to f$ is equivalent to say

$$\lim_{n\to\infty} m\{x \in X \mid |f_n(x) - f(x)| < \varepsilon\} = 0, \quad \forall \varepsilon > 0$$

- Part III: Lebesgue Integral.

Simple function: $\phi : X \to \mathbb{R}$ by

$$\phi(x) = \sum_{i=1}^{n} \alpha_i \mathcal{X}_{A_i}(x)$$

where $a_1, \ldots, a_n \in \mathbb{R}$ and $A_1, \ldots, A_n \subseteq X$ are disjoint, with

$$\mathcal{X}_{A_i}(x) =$$

$\int_X \mathcal{X}_{A_i} \, \mathrm{d}m = m(A_i)$

- Lebesgue integral for ϕ on $E \subseteq X$:

$$\int_E \phi \, \mathrm{d}m = \sum_{i=1}^{n} a_i m(A_i \cap E)$$

- Lebesgue integral for non-negative f:

$$\int_E f\,\mathrm{d}m = \sup\{\int_E \phi\,\mathrm{d}m \mid 0 \le \phi \le f\}$$

- For general f, $f = f^+ - f^-$, with

$$f^+ = \begin{cases} f, & f > 0 \\ 0, & \text{otherwise} \end{cases}$$

- If $\int |f|\,\mathrm{d}m < \infty$, then f is Lebesgue integrable.

3. - No proof for Fatuou's lemma.

   For non-negative $\{f_n\}$ on measurable X,

$$\int_X \lim_{n\to\infty} \int f_n\,\mathrm{d}m \le \lim_{n\to\infty} \int \int_X f_n\,\mathrm{d}m$$

- Proof is important: Monotonce convergence theorem: Assumption:
  - $f_n : X \to [0, \infty]$
  - $f_n \le f_{n+1}$
  - $f_n \to f$ a.e.

  Conclusion:

$$\lim_{n\to\infty} \int_X f_n\,\mathrm{d}m = \int_X \lim_{n\to\infty} f_n\,\mathrm{d}m = \int_X f\,\mathrm{d}m$$

- Dominated Convergence Theorem:
  - $|f_n| \le g$, where g is Lebesgue measurable
  - $f_n \to f$ a.e.
- Fubini's Theorem: (X, Y) is measurable, $(X \times Y)$, f is measurable function. If

$$\int_{X\times Y} |f|\,\mathrm{d}(x, y) = \int_X \int_Y |f|\,\mathrm{d}y\,\mathrm{d}x = \int_Y \int_X |f|\,\mathrm{d}x\,\mathrm{d}y$$

  then

$$\int_{X\times Y} f\,\mathrm{d}(x, y) = \int_X \int_Y f\,\mathrm{d}y\,\mathrm{d}x = \int_Y \int_X f\,\mathrm{d}x\,\mathrm{d}y$$

how does this theorem expresses in Riemann integral.

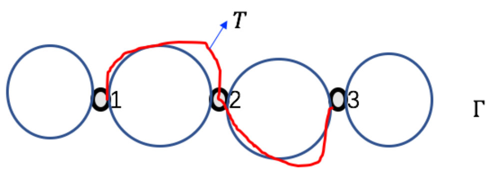
Proof in Riemann integral case.

Check counter-examples.

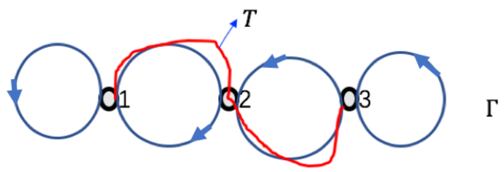
15.3. Monday for MAT4002

Theorem 15.4 Let Γ be a connected graph. Then $\pi(\Gamma)$ is isomorphic to the free group generated by $\#\{E(\Gamma)\setminus E(T)\}$ elements, for any maximal tree of Γ.

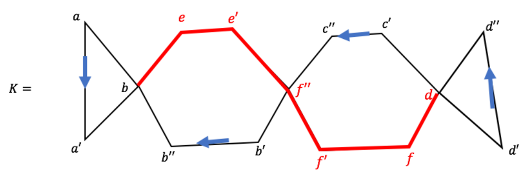
Now we give a proof for this theorem on one special case of Γ:



Proof. • Fix an orientation for each $e \in E(\Gamma)\setminus E(T)$:



- Now let K be a simplicial complex with $|K| \cong \Gamma$:



  As a result, $E(K,b) \cong \pi_1(\Gamma)$

- Now we construct the group homomorphism

$$
\begin{aligned}
\phi: \quad & \langle \alpha,\beta,\gamma,\delta\rangle \to E(K,b)\\
\text{with} \quad & \phi(\alpha) = [ba'a''b]\\
& \phi(\beta) = [bee'f''b'b''b]\\
& \phi(\gamma) = [bee'f''f'fdc'c''f''e'eb]\\
& \phi(\delta) = [bee'f''f'fdd''d'dff'f''e'eb]
\end{aligned}
$$

- We can show the group homomorphism ϕ is bijective. In particular, the inverse of ϕ is given by:

$$\Psi: \quad E(K,b) \to \langle \alpha,\beta,\gamma,\delta \rangle$$

where for any $[\ell] := [bv_1 \cdots v_n] \in E(K,b)$, the mapping $\Psi[\ell]$ is constructed by

(a) Remove all other letters appearing in ℓ except $b, a', a'', b', b'', c', c'', d', d''$

(b) Assign

$$\alpha, \ \alpha^{-1}, \ \beta, \ \beta^{-1}, \ \gamma, \ \gamma^{-1}, \ \delta, \ \delta^{-1}$$

for each appearance of

$$a'a'', \ a''a', \ b'b'', \ b''b', \ c'c'', \ c''c', \ d'd'', \ d''d',$$

respectively.

■

15.3.1. The Selfert-Van Kampen Theorem

Theorem 15.5 Let $K = K_1 \cup K_2$ be the union of two **path-connected open** sets, where $K_1 \cap K_2$ is also path-connected. Take $b \in K_1 \cap K_2$, and suppose the group presentations for $\pi_1(K_1,b), \pi_1(K_2,b)$ are

$$\pi_1(K_1,b) \cong \langle X_1 \mid R_1 \rangle, \quad \pi_1(K_2,b) \cong \langle X_2 \mid R_2 \rangle.$$

Let the inclusions be

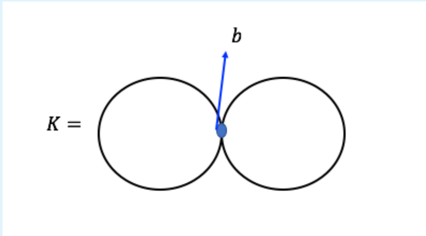
$$i_1 : K_1 \cap K_2 \hookrightarrow K_1, \quad i_2 : K_1 \cap K_2 \hookrightarrow K_2,$$

then a presentation of $\pi_1(K,b)$ is given by:

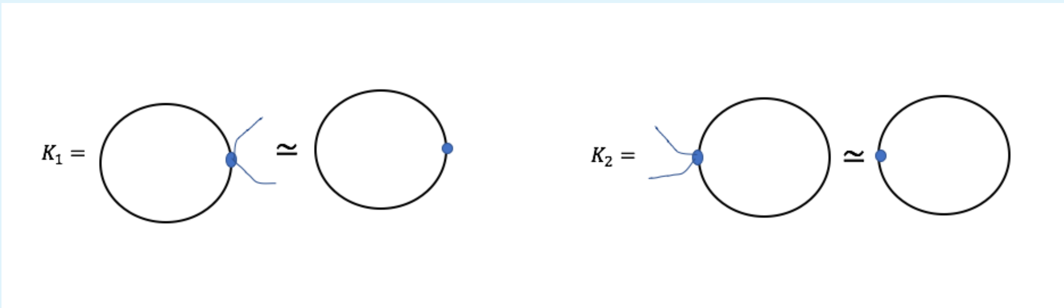
$$\pi_1(K,b) \cong \langle X_1 \cup X_2 \mid R_1 \cup R_2 \cup \{(i_1)_*(g) = (i_2)_*(g) : \forall g \in \pi_1(K_1 \cap K_2, b)\} \rangle.$$

(Here $(i_1)_*: \pi_1(K_1 \cap K_2, b) \hookrightarrow \pi_1(K_1, b)$ and $(i_2)_*: \pi_1(K_1 \cap K_2, b) \hookrightarrow \pi_1(K_2, b)$.)

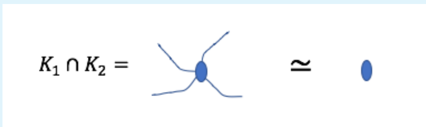
■ **Example 15.4** 1. Let $K = S^1 \wedge S^1$ given by



(a) Then construct b as the intersection between two circles, and construct K_1, K_2 as shown below:



We can see that $K_1 \cap K_2$ is contractible:



(b) As we have shown before, $\pi_1(S^1) \cong \mathbb{Z}$, which follows that

$$\pi_1(K_1, b) \cong \langle \alpha \rangle, \quad \pi_1(K_2, b) \cong \langle \beta \rangle$$

Also, $\pi_1(K_1 \cap K_2, b) \cong \pi_1(\{b\}, b) \cong \{e\}$.

(c) It's easy to compute $(i_1)_*$ and $(i_2)_*$:

$$\begin{array}{ll} (i_1)_*: & \pi_1(K_1 \cap K_2) \longrightarrow \pi_1(K_1) \\ \text{with} & e \longmapsto e \end{array}, \qquad \begin{array}{ll} (i_2)_*: & \pi_1(K_1 \cap K_2) \longrightarrow \pi_1(K_2) \\ \text{with} & e \longmapsto e \end{array},$$

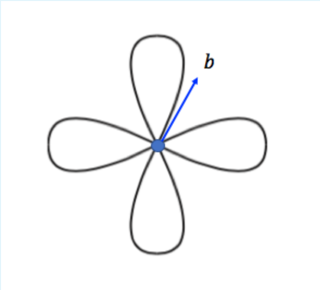
(d) Therefore, by Seifert-Van Kampen Theorem,

$$\pi_1(K,b) \cong \langle \alpha, \beta \mid e = e \rangle \cong \langle \alpha, \beta \rangle$$

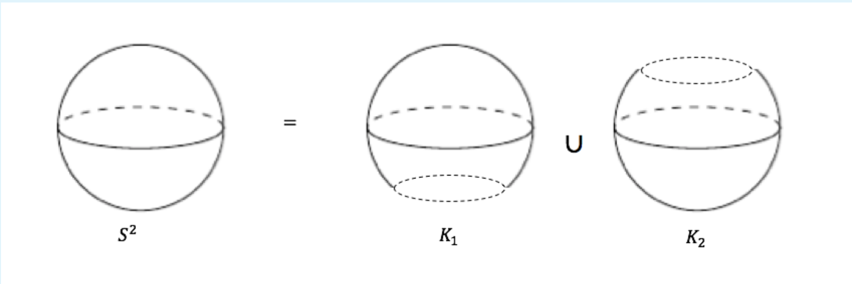
2. By induction,

$$\pi_1(\wedge^n S^1, b) \cong \langle a_1, \ldots, a_n \rangle$$

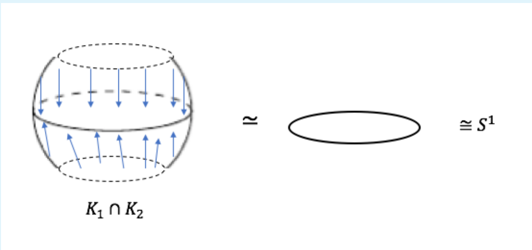
For instance, the figure illustration for $\wedge^4 S^1$ and the basepoint b is given below:



3. (a) Construct $S^2 = K_1 \cup K_2$, which is shown below:



Therefore, we see that $K_1 \cap K_2 \simeq S^1$:



(b) It's clear that K_1 and K_2 are contractible, and therefore

$$\pi_1(K_1) \cong \langle \beta \mid \beta \rangle, \quad \pi_1(K_2) \cong \langle \gamma \mid \gamma \rangle$$

and $\pi_1(K_1 \cap K_2) \cong \pi_1(S^1) \cong \langle \alpha \rangle$.

(c) Then we compute $(i_1)_*$ and $(i_2)_*$. In particular, the mapping $(i_1)_*$ is defined as

$$\begin{aligned} (i_1)_* : &\quad \pi_1(K_1 \cap K_2) \to \pi_1(K_1) \\ \text{with} &\quad [\alpha] \mapsto [i_1(\alpha)] \end{aligned}$$

where α is any loop based at b. Since K_1 is contractible, we imply α in K_1 is homotopic to c_b, i.e.,

$$(i_1)_*([\alpha]) = [i_1(\alpha)] = e, \forall \alpha \in \pi_1(K_1 \cap K_2).$$

Similarly, $(i_2)_*([\alpha]) = e$.

(d) By Seifert-Van Kampen Theorem, we conclude that

$$\pi_1(S^2) \cong \langle \beta, \gamma \mid \beta, \gamma, e = e \rangle \cong \{e\}$$

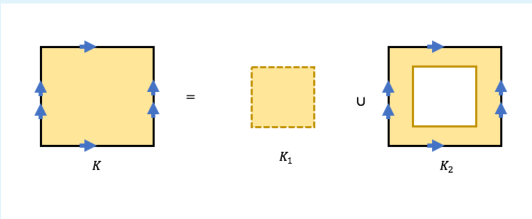
4. Homework: Use the same trick to check that $\pi_1(S^n) = \{e\}$ for all $n \geq 2$. Hint: for S^3, construct

$$K_1 = \{(x_1, \ldots, x_4) \in S^3 \mid x_4 > -1/2\}$$

and

$$K_1 = \{(x_1, \ldots, x_4) \in S^3 \mid x_4 < 1/2\}$$

5. (a) Consider the quotient space $K \cong \mathbb{T}^2$, and we construct $K = K_1 \cup K_2$ as follows:



Therefore, we can see that K_1 is contractible, and K_2 is homotopy equivalent to $S^1 \wedge S^1$:

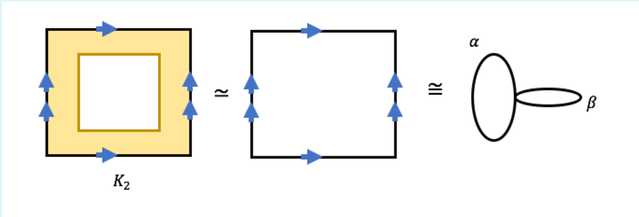
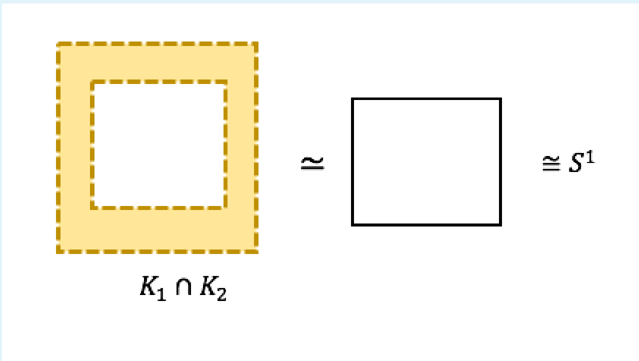


Figure 15.2: Illustration for $K_2 \simeq S^1 \wedge S^1$

and $K_1 \cap K_2$ is homotopic equivalent to the circle:



(b) It follows that

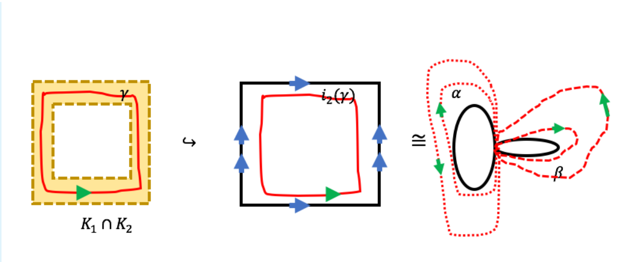
$$\pi_1(K_1) \cong \{e\}, \quad \pi_1(K_2) \cong \langle \alpha, \beta \rangle,$$

and $\pi_1(K_1 \cap K_2) \cong \langle \gamma \rangle$.

(c) Then we compute $(i_1)_*$ and $(i_2)_*$. In particular, $(i_1)_*$ is trivial:

$$\begin{aligned} (i_1)_* : \quad & \pi_1(K_1 \cap K_2) \to \pi_1(K_1) \\ \text{with} \quad & [\alpha] \mapsto e \end{aligned}$$

Then compute $(i_2)_*$. In particular, for any loop γ, we draw the graph for $i_2(\gamma)$:



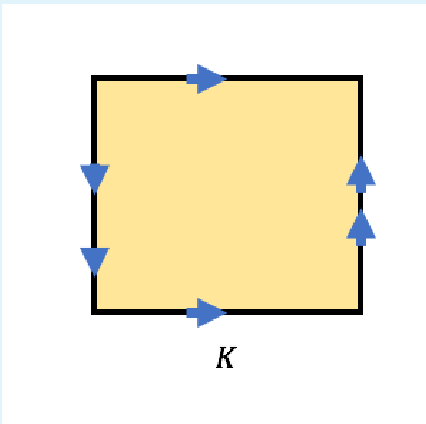
Therefore,

$$(i_2)_*[\gamma] = [i_2(\gamma)] = [\alpha\beta\alpha^{-1}\beta^{-1}]$$

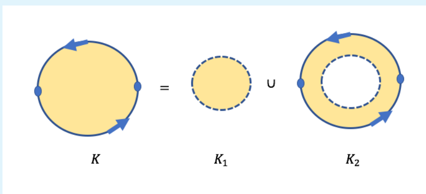
(d) By Seifert-Van Kampen Theorem, we conclude that

$$\pi_1(K) \cong \langle \alpha, \beta \mid \beta, \alpha\beta\alpha^{-1}\beta^{-1} = e \rangle \cong \langle \alpha, \beta \mid, \alpha\beta = \beta\alpha \rangle \cong \mathbb{Z} \times \mathbb{Z}$$

6. Exerise: The Klein bottle K shown in graph below satisfies $\pi_1(K) = \langle a, b \mid aba^{-1}b \rangle$.



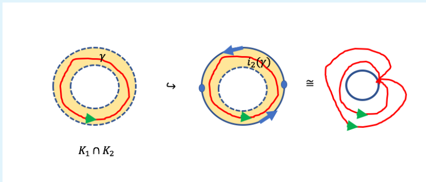
7. Consider the quotient space $K = \mathbb{R}P^2$. We construct $K = K_2 \cup K_2$, which is shown below:



(a) It's clear that K_1 is contractible. In hw3, question 1, we can see that $K_2 \simeq S^1$. Moreover, similar as in (5), $K_1 \cap K_2 \simeq S^1$.

(b) Therefore, $\pi_1(K_1) = \{e\}$ and $\pi_1(K_2) = \langle \alpha \rangle$, $\pi_1(K_1 \cap K_2) = \langle \gamma \rangle$.

(c) It's easy to see that $(i_1)_*([\gamma]) = e$ for any loop γ. For any loop γ, we draw the graph for $i_2(\gamma)$:



Therefore, $(i_2)_*([\gamma]) = [i_2(\gamma)] = [\alpha^2]$.

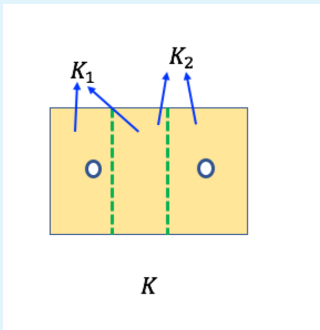
(d) By Seifert-Van Kampen Theorem, we conclude that

$$\pi_1(K) \cong \langle \alpha \mid \alpha^2 = e \rangle \cong \mathbb{Z}/2\mathbb{Z} \cong \{0,1\}_{\text{mod }(2)}$$

8. Let $K = \mathbb{R}^2 \setminus \{2 \text{ points } \alpha, \beta\}$. As have shown in hw3, $K \simeq S^1 \wedge S^1$, which implies

$$\pi_1(K) \cong \pi_1(S^1 \wedge S^1) \cong \langle \alpha, \beta \rangle.$$

We can compute the fundamental group for K directly. Construct $K = K_1 \cup K_2$ as follows:



(a) It's clear that $K_1 \cong \mathbb{R}^2 \setminus \{\text{one point}\} \simeq S^1$ and similarly $K_2 \simeq S^1$. Moreover, $K_1 \cap K_2$ is contractible

(b) Therefore,

$$\pi_1(K_1) \cong \langle \alpha \rangle, \quad \pi_1(K_2) \cong \langle \beta \rangle, \quad \pi_1(K_1 \cap K_2) \cong \{e\}$$

(c) Therefore, $(i_1)_*$ and $(i_2)_*$ is trivial since $\pi_1(K_1 \cap K_2) \cong \{e\}$.

(d) By Seifert-Van Kampen Theorem, we conclude that

$$\pi_1(K) \cong \langle \alpha, \beta \mid e = e \rangle \cong \langle \alpha, \beta \rangle$$

■