

Topology I: Elementary Topology

Yi Li

SCHOOL OF MATHEMATICS AND SHING-TUNG YAU CENTER, SOUTHEAST
UNIVERSITY, NANJING, CHINA 210096

E-mail address: yilicms@seu.edu.cn; yilicms@gmail.com; yilicms@163.com

Contents

Main References	1
Chapter 1. Basic topology	3
1.1. Metric spaces	3
1.2. Topological spaces	13
1.3. Constructions	38
1.4. Topological groups and matrix Lie groups	52
1.5. Connectedness and compactness	69
Chapter 2. Analysis on topological groups	83
2.1. Haar measures	83
2.2. Iwasawa's decomposition	89
2.3. Harish, Mellin and spherical transforms	94
Chapter 3. Fundamental groups	111
3.1. Homotopy	111
3.2. H -spaces	122
3.3. $\pi_1(\mathbf{S}^1)$	137

Main References

- Kelley, John L. *General topology*, Reprint of the 1955 edition [Van Nostrand, Toronto, Ont.], Graduate Texts in Mathematics, No. 27, Springer-Verlag, New York-Berlin, 1975. xiv+298 pp. MR0370454 (51# 6681)
- Bredon, Glen E. *Topology and geometry*, Graduate Texts in Mathematics, No. 139, Springer-Verlag, New York, 1993. xiv+557 pp.
- Lee, John M. *Introduction to topological manifolds*, Second edition, Graduate Texts in Mathematics, No. 202, Springer, New York, 2011. xviii+433 pp. ISBN: 978-1-4419-7939-1
- Munkres, James R. *Topology: a first course*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1975. xvi+413 pp.
- Burago, Dmitri; Burago, Yuri; Ivanov, Sergei. *A course in metric geometry*, Graduate Studies in Mathematics, 33, American Mathematical Society, Providence, RI, 2001. xiv+415 pp. ISBN: 0-8218-2129-6
- Spanier, Edwin H. *Algebraic topology*, Corrected reprint of the 1966 original, Springer-Verlag, New York. xvi+528 pp. ISBN: 0-387-94426-5

CHAPTER 1

Basic topology

1.1. Metric spaces

Over the n -dimensional Euclidean space \mathbf{R}^n , we have a natural distance function

$$d(x, y) := \sqrt{\sum_{1 \leq i \leq n} (x^i - y^i)^2}, \quad x = (x^1, \dots, x^n), \quad y = (y^1, \dots, y^n),$$

which satisfies the following properties: d is always nonnegative (with zero if and only if $x = y$), symmetric, and $d(x, z) \leq d(x, y) + d(y, z)$.

1.1.1. Metric spaces. A **metric** on a set X is a map $d : X \times X \rightarrow \bar{\mathbf{R}} := \mathbf{R} \cup \{+\infty\}$ satisfying, for any $x, y, z \in X$,

- (a) **(positiveness)** $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$;
- (b) **(symmetry)** $d(x, y) = d(y, x)$;
- (c) **(triangle inequality)** $d(x, z) \leq d(x, y) + d(y, z)$.

A **metric space** is a pair (X, d) where d is a metric on X .

Given a metric space (X, d) we define a relation \sim on X by

$$x \sim y \iff d(x, y) < \infty.$$

Then \sim is an equivalence relation on X , and the equivalence class $[x]$ of x can be endowed with a natural metric (still denoted by d).

Exercise 1.1.1. For any $x \in X$, show that $([x], d)$ is a finite metric space.

A map $f : (X, d_X) \rightarrow (Y, d_Y)$ between metric spaces is called **distance-preserving** if

$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2)$$

for all $x_1, x_2 \in X$. A distance-preserving bijection is called an **isometry**.

A **semi-metric** on a set X is a map $d : X \times X \rightarrow \bar{\mathbf{R}} := \mathbf{R} \cup \{\infty\}$ satisfying, for any $x, y, z \in X$,

- (a) **(nonnegativity)** $d(x, y) \geq 0$;
- (b) **(symmetry)** $d(x, y) = d(y, x)$;
- (c) **(triangle inequality)** $d(x, z) \leq d(x, y) + d(y, z)$.

A **semi-metric space** is a pair (X, d) where d is a semi-metric on X .

Given a semi-metric space (X, d) we define a relation \sim on X by

$$x \sim y \iff d(x, y) = 0.$$

Then \sim is an equivalence relation on X . Define

$$\widehat{X} := X / \sim = \{[x] : x \in X\}, \quad \widehat{d}([x], [y]) := d(x, y).$$

Then $(X/d, d) := (\widehat{X}, \widehat{d})$ is a metric space.

Exercise 1.1.2. Show that \widehat{d} is well-defined and $(X/d, d)$ is a metric space.

Exercise 1.1.3. Let $X = \mathbf{R}^2$ and define

$$d((x, y), (x', y')) := |(x - x') + (y - y')|.$$

Show that d is a semi-metric on X . Define $f : \mathbf{R}^2/d \rightarrow \mathbf{R}$ by $f([(x, y)]) := x + y$. Show that f is an isometry.

Exercise 1.1.4. Consider the set X of all continuous real-valued functions on $[0, 1]$. Show that

$$d(f, g) := \int_0^1 |f(x) - g(x)| dx$$

defines a metric on X . Is this still the case if continuity is weakened to integrability?

Exercise 1.1.5. Let (X, d_X) and (Y, d_Y) be two metric spaces. We define

$$d_{X \times Y}((x_1, y_1), (x_2, y_2)) := \sqrt{d_X(x_1, x_2) + d_Y(y_1, y_2)}.$$

Show that $(X \times Y, d_{X \times Y})$ is a metric space.

Exercise 1.1.6. As a subspace of \mathbf{R}^2 , the unit circle \mathbf{S}^1 carries the restricted Euclidean metric from \mathbf{R}^2 . We can define another (intrinsic) metric d_{int} by

$$d_{\text{int}}(x, y) := \text{the length of the shorter arc between them.}$$

Note that $d_{\text{int}}(x, y) \in [0, \pi]$ for any points $x, y \in \mathbf{S}^1$.

- (a) Show that any circle arc of length less than or equal to π , equipped with the intrinsic metric, is isometric to a straight line segment.
- (b) The whole circle with the intrinsic metric is not isometric to any subset of \mathbf{R}^2 .

Let (X, d) be a metric space. Set

$$B(x, r) := \{y \in X : d(x, y) < r\}, \quad \overline{B(x, r)} := \{y \in X : d(x, y) \leq r\}.$$

We define the **metric topology** on X by saying that $U \subseteq X$ is open if and only for any $x \in U$ there exists $\epsilon > 0$ such that $B(x, \epsilon) \subseteq U$. Let

$$(1.1.1.1) \quad \mathcal{T} := \{\text{all open subsets in } X\} \cup \{\emptyset, X\}.$$

The set \mathcal{T} satisfies the following

- (1) $U \cap V \in \mathcal{T}$ for any $U, V \in \mathcal{T}$;
- (2) $\cup_{i \in I} U_i \in \mathcal{T}$ for any collection $\{U_i\}_{i \in I} \subset \mathcal{T}$;
- (3) $\emptyset, X \in \mathcal{T}$.

In general, we can introduce a topology on any set X . A **topology** on X is a subset \mathcal{T} of 2^X satisfying

- (1) $U \cap V \in \mathcal{T}$ for any $U, V \in \mathcal{T}$;

- (2) $\cup_{i \in I} U_i \in \mathcal{T}$ for any collection $\{U_i\}_{i \in I} \subset \mathcal{T}$;
- (3) $\emptyset, X \in \mathcal{T}$.

The pair (X, \mathcal{T}) is called the **topological space**, and any element in \mathcal{T} is said to be **open**; a **closed set** is a subset of X whose complement in X is open. In particular, we see that any metric space must be a topological space.

Given a metric space (X, d) .

- (X, d) is **complete** if any Cauchy sequence¹ in X has a limit in X .
- A **diameter** of $S \subset X$ is defined by

$$(1.1.1.2) \quad \text{diam}(S) := \sup_{x, y \in S} d(x, y).$$

- (X, d) is **compact** if any sequence in X has a converging subsequence.
- $S \subset X$ is an **ϵ -net** if $S_\epsilon := \{x \in X : d(x, S) = \inf_{y \in S} d(x, y) \leq \epsilon\} = X$.
- X is **totally bounded** if for any $\epsilon > 0$ there exists a finite ϵ -net in X .

The following theorem is a basic result.

Theorem 1.1.7. *Let (X, d) be a metric space. Then X is compact if and only if X is complete and totally bounded.*

1.1.2. Length spaces. Let (X, \mathcal{T}) be a topological space. We say X is **Hausdorff** if for any different points $x, y \in X$, there exist open sets $U_x, U_y \subseteq X$ such that $x \in U_x$, $y \in U_y$, and $U_x \cap U_y = \emptyset$. Clearly that any metric space must be Hausdorff.

A **path** γ in X is a continuous² map $\gamma : I \rightarrow X$, where $I \subseteq \mathbf{R}$ is an interval (i.e., connected subset) of \mathbf{R} .

A **length structure** on X is a pair (\mathcal{A}, L) , where \mathcal{A} is a collection of paths (called **admissible paths**) and $L : \mathcal{A} \rightarrow \overline{\mathbf{R}}_+$ is a map on \mathcal{A} (called the **length of paths**), satisfying

- (a) (**\mathcal{A} is closed under restrictions**) If $\gamma : [a, b] \rightarrow X$ in \mathcal{A} , then $\gamma|_{[c, d]} \in \mathcal{A}$ for any subinterval $[c, d] \subseteq [a, b]$;
- (b) (**\mathcal{A} is closed under concatenations of paths**) If $\gamma_1 : [a, c] \rightarrow X$ and $\gamma_2 : [c, d] \rightarrow X$ are admissible with $\gamma_1(c) = \gamma_2(c)$, then the obvious product $\gamma := \gamma_1 \cdot \gamma_2 : [a, d] \rightarrow X$ is also admissible;
- (c) (**\mathcal{A} is closed under reparameterizations**) If $\gamma : [a, b] \rightarrow X$ is admissible and $\varphi : [c, d] \rightarrow [a, b]$ is a homeomorphism, then $\gamma \circ \varphi : [c, d] \rightarrow X$ is also admissible;
- (d) (**Length of paths is additive**) $L(\gamma|_{[a, b]}) = L(\gamma|_{[a, c]}) + L(\gamma|_{[c, b]})$ for any $c \in [a, b]$ and $\gamma : [a, b] \rightarrow X$ in \mathcal{A} ;
- (e) (**The length of a piece of a path continuously depends on the piece**) If $\gamma : [a, b] \rightarrow X$ is an admissible path with $L(\gamma) < +\infty$, then $L(\gamma, a, t) := L(\gamma|_{[a, t]})$ is continuous in t ;
- (f) (**$L(\cdot)$ is invariant under reparameterizations**) $L(\gamma \circ \varphi) = L(\gamma)$ for any admissible path γ and any homeomorphism φ as in (c);

¹A sequence $\{x_n\}_{n \in \mathbf{N}}$ in X is said to be **Cauchy** if for any $\epsilon > 0$ there exists an integer $n_0 \in \mathbf{N}$ such that $d(x_n, x_m) < \epsilon$ whenever $n, m \geq n_0$.

²That is, $\gamma^{-1}(U)$ is open in I for any open set U in X .

- (g) **(Length structures agree with the topology of X)** For any $x \in X$ and open set $U_x \ni x$, one has

$$\inf_{\gamma \in \mathcal{A}} \{L(\gamma) : \gamma(a) = x, \gamma(b) \in X \setminus U_x\} > 0.$$

Let (X, \mathcal{T}) is a Hausdorff space with a length structure (\mathcal{A}, L) . Define

$$(1.1.2.1) \quad d_L(x, y) := \inf_{\gamma \in \mathcal{A}} \{L(\gamma) \mid \gamma : [a, b] \rightarrow X, \gamma(a) = x, \gamma(b) = y\}.$$

If there is no path between x and y , we set $d_L(x, y) = +\infty$. We say a length structure (\mathcal{A}, L) is **complete** if for any $x, y \in X$ with $d_L(x, y) < +\infty$, there exists a path $\gamma \in \mathcal{A}$ joining x and y such that $d_L(x, y) = L(\gamma)$.

Exercise 1.1.8. Show that (X, d_L) is a metric space, where d_L is defined in (1.1.2.1).

Note that d_L is not necessarily a finite metric.

Definition 1.1.9. A **length space** is a metric space (X, d) such that $d = d_L$ for some length structure (\mathcal{A}, L) . In this case, we call d as an **intrinsic metric** on X .

In the following we shall construct a length structure on a given metric space (X, d) . If $\gamma : [a, b] \rightarrow X$ is a path in X , for any partition $\{y_i\}_{0 \leq i \leq N}$ of $[a, b]$ with $a = y_0 \leq y_1 \leq y_2 \leq \dots \leq y_N = b$, we define

$$\Sigma(\gamma, \{y_i\}_{0 \leq i \leq N}) := \sum_{1 \leq i \leq N} d(\gamma(y_{i-1}), \gamma(y_i)).$$

The **length** of γ is given by

$$(1.1.2.2) \quad L_d(\gamma) := \sup_{\{y_i\}_{0 \leq i \leq N}} \Sigma(\gamma, \{y_i\}_{0 \leq i \leq N}).$$

The path γ is said to be **rectifiable** if $L_d(\gamma) < +\infty$. The length structure (\mathcal{A}, L) now is defined by

$$(1.1.2.3) \quad \mathcal{A} := \{\text{all paths parametrized by closed intervals}\}, \quad L := L_d.$$

Theorem 1.1.10. Let (X, d) be a metric space, for the length structure (\mathcal{A}, L) defined in (1.1.2.3), we have the following properties:

- (a) $L(\gamma) \geq d(\gamma(a), \gamma(b))$ for any path $\gamma : [a, b] \rightarrow X$;
- (b) $L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]}) = L(\gamma)$ for any path $\gamma : [a, b] \rightarrow X$ and $a < c < b$;
- (c) If γ is rectifiable, then $L(\gamma, c, d) := L(\gamma|_{[c,d]})$ is continuous in c and d ;
- (d) L is a lower semi-continuous functional on $C([a, b], X)$ with respect to the point-wise convergence; that is, if $\{\gamma_i\}_{i \in \mathbf{N}}$ is a sequence of rectifiable paths in X with the same domain and $\gamma_i(t)$ converges to $\gamma(t)$ for every given $t \in [a, b]$, then $L(\gamma) \leq \liminf_{i \rightarrow \infty} L(\gamma_i)$.

PROOF. (a) For any partition $\{y_i\}_{0 \leq i \leq N}$ of $[a, b]$ we have

$$\Sigma(b\gamma, \{y_i\}_{0 \leq i \leq N}) = \sum_{1 \leq i \leq M} d(\gamma(y_{i-1}), \gamma(y_i)) \geq d(\gamma(y_0), \gamma(y_N)) = d(\gamma(a), \gamma(b))$$

by the triangle inequality, so that $L(\gamma) \geq d(\gamma(a), \gamma(b))$.

(b) Given partitions $\{y'_i\}_{0 \leq i \leq N'}$ and $\{y''_j\}_{0 \leq j \leq N''}$ of $[a, c]$ and $[c, b]$, respectively. Then $\{y'_i\}_{0 \leq i \leq N'} \cup \{y''_j\}_{0 \leq j \leq N''}$ forms a partition of $[a, b]$ and

$$\begin{aligned} L(\gamma) &\geq \Sigma\left(\gamma, \{y'_i\}_{0 \leq i \leq N'} \cup \{y''_j\}_{0 \leq j \leq N''}\right) \\ &= \sum_{1 \leq i \leq N'} d(\gamma(y'_{i-1}), \gamma(y'_i)) + \sum_{1 \leq j \leq N''} d(\gamma(y''_{j-1}), \gamma(y''_j)) \\ &= \Sigma\left(\gamma|_{[a,c]}, \{y'_i\}_{0 \leq i \leq N'}\right) + \Sigma\left(\gamma|_{[c,b]}, \{y''_j\}_{0 \leq j \leq N''}\right). \end{aligned}$$

Hence $L(\gamma) \geq L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]})$.

Conversely, for any partition $\{y_i\}_{0 \leq i \leq N}$, we see that $\{y_i\}_{0 \leq i \leq N} \cup \{c\}$ is also a partition of $[a, b]$. Consequently, we get

$$L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]}) \geq \Sigma(\gamma, \{y_i\}_{0 \leq i \leq N} \cup \{c\}) \geq \Sigma(\gamma, \{y_i\}_{0 \leq i \leq N})$$

and then $L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]}) \geq L(\gamma)$.

(c) We only prove that $L(\gamma, c, d)$ is continuous in $d \in (a, b]$. For any $\epsilon > 0$, since $L(\gamma) < +\infty$, there exists a partition $\{y_i\}_{0 \leq i \leq N}$ of $[a, b]$ such that $0 \leq L(\gamma) - \Sigma(\gamma, \{y_i\}_{0 \leq i \leq N}) < \epsilon$. We may assume that $y_{j-1} < d = y_j$ for some j , otherwise we can use the partition $\{y_i\}_{0 \leq i \leq N} \cup \{d\}$. Then

$$L(\gamma|_{[y_{j-1}, d]}) - d(\gamma(y_{j-1}), \gamma(d)) \leq L(\gamma) - \Sigma(\gamma, \{y_i\}_{0 \leq i \leq N}) < \epsilon$$

by (a) and (b). For any $d' \in [y_{j-1}, d]$, one has

$$\begin{aligned} L(\gamma|_{[y_{j-1}, d]}) - L(\gamma|_{[y_{j-1}, d']}) &\leq \epsilon + d(\gamma(y_{j-1}), \gamma(d)) - L(\gamma|_{[y_{j-1}, d']}) \\ &\leq \epsilon + d(\gamma(y_{j-1}), \gamma(d)) - d(\gamma(y_{j-1}), \gamma(d')) \\ &\leq \epsilon + d(\gamma(d'), \gamma(d)). \end{aligned}$$

Since γ is continuous, it follows that $L(\gamma, c, d') - L(\gamma, c, d) = L(\gamma, d', d) \leq 2\epsilon$ when d' is very close to d . Thus $L(\gamma, c, d)$ is continuous in d .

(d) We first assume that $L(\gamma) < +\infty$. For any $\epsilon > 0$ there exists a partition $\{y_i\}_{0 \leq i \leq N}$ of $[a, b]$ such that $L(\gamma) - \Sigma(\gamma, \{y_i\}_{0 \leq i \leq N}) < \epsilon$. Choose $j \gg 1$ so that $d(\gamma_j(y_i), \gamma(y_i)) < \epsilon/N$ for each $0 \leq i \leq N$. Compute

$$\begin{aligned} L(\gamma) &\leq \Sigma(\gamma, \{y_i\}_{0 \leq i \leq N}) + \epsilon \\ &\leq \Sigma(\gamma_j, \{y_i\}_{0 \leq i \leq N}) + \epsilon + \frac{\epsilon}{N} \cdot 2 \cdot N \\ &\leq L(\gamma_j) + 3\epsilon \end{aligned}$$

where we used

$$\begin{aligned} d(\gamma(y_{i-1}), \gamma(y_i)) &\leq d(\gamma(y_{i-1}), \gamma_j(y_{i-1})) + d(\gamma_j(y_{i-1}), \gamma_j(y_i)) \\ &\quad + d(\gamma(y_i), \gamma_j(y_i)) \\ &\leq \frac{\epsilon}{N} \cdot 2 + d(\gamma_j(y_{i-1}), \gamma_j(y_i)). \end{aligned}$$

Therefore $L(\gamma) \leq \liminf_{j \rightarrow \infty} L(\gamma_j)$.

When $L(\gamma) = +\infty$, we choose a partition $\{y_i\}_{0 \leq i \leq N}$ such that $\Sigma(\gamma, \{y_i\}_{0 \leq i \leq N}) > 1/\epsilon$. Then

$$\frac{1}{\epsilon} < \Sigma(\gamma, \{y_i\}_{0 \leq i \leq N}) \leq \Sigma(\gamma_j, \{y_i\}_{0 \leq i \leq N}) + \frac{\epsilon}{N} \cdot 2 \cdot N \leq L(\gamma_j) + 2\epsilon$$

for $j \gg 1$. Letting $j \rightarrow \infty$ yields $L(\gamma_j) \rightarrow \infty$. \square

For a metric space (X, d) , we have a length structure (\mathcal{A}, L) , where $L := L_d$, and then the **induced intrinsic metric**

$$(1.1.2.4) \quad d^\wedge := d_L.$$

However, d^\wedge may be equal to d . As a direct consequence of [Theorem 1.1.10](#), $d \leq d^\wedge$ always holds.

Example 1.1.11. Consider the subset

$$X := \bigcup_{n \in \mathbb{N}} \left[(0, 1), \left(\frac{1}{n}, 0 \right) \right] \cup [(0, 1), (0, 0)] \subset \mathbb{R}^2.$$

For the standard metric $d := d_{\mathbb{R}^2}|_X$ we see that $d((1/n, 0), (0, 0)) = 1/n \rightarrow 0$ as $n \rightarrow \infty$. However the limit $\lim_{n \rightarrow \infty} d^\wedge((1/n, 0), (0, 0))$ does not exist; in fact, for any path γ connecting $(1/n, 0)$ with $(1/m, 0)$, we have

$$\begin{aligned} L(\gamma) &= L_d(\gamma) \geq d\left(\left(\frac{1}{n}, 0\right), (0, 0)\right) + d\left(\left(\frac{1}{m}, 0\right), (0, 0)\right) \\ &= \sqrt{1 + \frac{1}{n^2}} + \sqrt{1 + \frac{1}{m^2}} > 2. \end{aligned}$$

Given a metric space (X, d) , we have the following diagram:

$$\begin{array}{ccc} (X, d) & \longrightarrow & (\mathcal{A}, L = L_d) \\ & & \downarrow \\ & & (X, d^\wedge = d_L) \\ & & \downarrow \\ (X, (d^\wedge)^\wedge = d_{\widehat{L}}) & \longleftarrow & (\mathcal{A}, \widehat{L} = L_{d^\wedge}) \end{array}$$

by [Theorem 1.1.10](#). A natural question is whether $d = (d^\wedge)^\wedge$.

Theorem 1.1.12. Let (X, d) be a metric space and \widehat{d} the intrinsic metric induced by d .

- (a) If $L_d(\gamma) < +\infty$, then $L_d(\gamma) = L_{d^\wedge}(\gamma)$.
- (b) The intrinsic metric induced by d^\wedge coincides with d^\wedge , i.e., $(d^\wedge)^\wedge = d^\wedge$.

PROOF. (a) For any two points $x, y \in X$, choose a path $\gamma : [a, b] \rightarrow X$ from x to y . Then

$$L_d(\gamma) \geq d(\gamma(a), \gamma(b)) = d(x, y)$$

which implies $d(x, y) \leq d^\wedge(x, y)$ and hence $L_d(\gamma) \leq L_{d^\wedge}(\gamma)$.

Conversely, for any path $\gamma : [a, b] \rightarrow X$ and any partition $\{y_i\}_{0 \leq i \leq N}$ of the interval $[a, b]$, we have

$$\sum_{\widehat{d}}(\gamma, \{y_i\}_{0 \leq i \leq N}) \leq \sum_{1 \leq i \leq N} d^\wedge(\gamma(y_{i-1}), \gamma(y_i)) \leq \sum_{1 \leq i \leq N} L_d(\gamma|_{[y_{i-1}, y_i]}) = L_d(\gamma)$$

according to [Theorem 1.1.10 \(b\)](#).

Part (b) follows from an obvious way. \square

Example 1.1.13. Let (\mathcal{M}, g) be a connected Riemannian manifold. Define

$$(1.1.2.5) \quad L(x, y) = L_d(\gamma) := \int_a^b |\dot{\gamma}(t)|_g dt,$$

where

$$(1.1.2.6) \quad d(x, y) = d_g(x, y) := \inf \left\{ L_g(\gamma) \mid \gamma \text{ is a piecewise } C^1\text{-path joining } x \text{ to } y \right\}.$$

Then (\mathcal{M}, d) is a metric space.

1.1.3. Gromov-Hausdorff distance. In this subsection, all proofs can be found in BBI's book "A course in metric geometry".

Let (Z, d_Z) be a metric space. Given an $\epsilon > 0$, the ϵ -neighborhood of a subset $S \subset Z$ is given by

$$(1.1.3.1) \quad S_\epsilon := \left\{ z \in Z : d_Z(z, S) = \inf_{x \in S} d_Z(z, x) \leq \epsilon \right\}.$$

We say a subset $S \subset X$ is an ϵ -net if $S_\epsilon = X$. The **Hausdorff distance** of $A, B \subset Z$ is defined to be the quantity

$$(1.1.3.2) \quad d_{\mathbf{H}}^Z(A, B) := \inf \{ \epsilon > 0 : A \subset B_\epsilon \text{ and } B \subset A_\epsilon \}.$$

If no such ϵ , we put $d_{\mathbf{H}}^Z(A, B) = +\infty$. A basic fact is that the set of all compact subsets of Z , equipped with $d_{\mathbf{H}}^Z$, forms a metric space.

For any metric spaces (X, d_X) and (Y, d_Y) , we define **Gromov-Hausdorff distance** between them by

$$(1.1.3.3) \quad d_{\mathbf{GH}}((X, d_X), (Y, d_Y)) := \inf_{(Z, d_Z), f, g} \left\{ d_{\mathbf{H}}^Z(f(X), g(Y)) : \begin{array}{l} f : (X, d_X) \rightarrow (Z, d_Z) \\ g : (Y, d_Y) \rightarrow (Z, d_Z) \\ \text{are isometric embeddings} \\ \text{into a metric space } (Z, d_Z) \end{array} \right\}.$$

One can show that if $(X, d_X), (Y, d_Y)$ are compact metric spaces with zero Gromov-Hausdorff distance, then (X, d_X) and (Y, d_Y) are isometric. Therefore, the set of all compact metric space, modulo the isometric equivalence, is a metric space with respect to $d_{\mathbf{GH}}$.

A **distortion** of a map $f : (X, d_X) \rightarrow (Y, d_Y)$ between metric spaces is defined to be

$$(1.1.3.4) \quad \text{dis}(f) := \sup_{x_1, x_2 \in X} |d_X(x_1, x_2) - d_Y(f(x_1), f(x_2))|.$$

Note that $\text{dis}(f) = 0$ if and only if f is distance-preserving. Such a f is called an ϵ -isometry if

$$\text{dis}(f) \leq \epsilon \text{ and } f(X) \text{ is an } \epsilon\text{-net in } Y.$$

The following result characterizes the Gromov-Hausdorff distances with ϵ -isometries.

Lemma 1.1.14. *Let (X, d_X) and (Y, d_Y) be metric spaces and $\epsilon > 0$.*

- (a) *If $d_{\text{GH}}((X, d_X), (Y, d_Y)) < \epsilon$, then there exist ϵ -isometries $f : (X, d_X) \rightarrow (Y, d_Y)$ and $g : (Y, d_Y) \rightarrow (X, d_X)$.*
- (b) *If $f : (X, d_X) \rightarrow (Y, d_Y)$ is an ϵ -isometry, then $d_{\text{GH}}((X, d_X), (Y, d_Y)) < 2\epsilon$.*

A **pointed metric space** is a triple (X, d_X, x_0) , where (X, d_X) is a metric space and x_0 is a fixed base-point. A **pointed map** $f : (X, d_X, x_0) \rightarrow (Y, d_Y, y_0)$ is a map $f : (X, d_X) \rightarrow (Y, d_Y)$ which preserves the base-points (that is, $f(x_0) = y_0$). An **ϵ -pointed-isometry** $f : (X, d_X, x_0) \rightarrow (Y, d_Y, y_0)$ is a pointed map f satisfying

$$|d_X(x_1, x_2) - d_Y(f(x_1), f(x_2))| < \epsilon$$

for any $x_1, x_2 \in B_X(x_0, \epsilon^{-1})$, and $B_Y(y_0, \epsilon^{-1}) \subset (f(B_X(x_0, \epsilon^{-1})))_\epsilon$. For two pointed metric spaces (X, d_X, x_0) and (Y, d_Y, y_0) define the **pointed-Gromov-Hausdorff distance** as

$$(1.1.3.5) \quad d_{\text{GH}}^{\text{pt}}((X, d_X, x_0), (Y, d_Y, y_0)) := \inf \left\{ \epsilon > 0 : \begin{array}{l} \text{there exist } \epsilon\text{-pointed-isometries} \\ f : (X, d_X, x_0) \rightarrow (Y, d_Y, y_0) \text{ and} \\ g : (Y, d_Y, y_0) \rightarrow (X, d_X, x_0) \end{array} \right\}.$$

Lemma 1.1.15. *Given $\epsilon > 0$. If $d_{\text{GH}}^{\text{pt}}((X, d_X, x_0), (Y, d_Y, y_0)) < \epsilon$, then there exist 2ϵ -isometries $f : (B_X(x_0, \epsilon^{-1}), d_X) \rightarrow (B_Y(y_0, \epsilon^{-1} + \epsilon), d_Y)$ with $f(x_0) = y_0$ and $g : (B_Y(y_0, \epsilon^{-1}), d_Y) \rightarrow (B_X(x_0, \epsilon^{-1} + \epsilon), d_X)$ with $g(y_0) = x_0$.*

PROOF. By definition, there exist 2ϵ -pointed-isometries $f : (X, d_X, x_0) \rightarrow (Y, d_Y, y_0)$ and $g : (Y, d_Y, y_0) \rightarrow (X, d_X, x_0)$. For any $x \in B_X(x_0, \epsilon^{-1})$ we have $|d_X(x, x_0) - d_Y(f(x), y_0)| < \epsilon$ and then $f(B_X(x_0, \epsilon^{-1})) \subseteq B_Y(y_0, \epsilon^{-1} + \epsilon)$. \square

We say that a sequence of pointed metric spaces (X_i, d_{X_i}, x_i) **converges** to a pointed metric space (Y, d_Y, y_0) **in the sense of pointed Gromov-Hausdorff distance**, if

$$\lim_{i \rightarrow \infty} d_{\text{GH}}^{\text{pt}}((X_i, d_{X_i}, x_i), (Y, d_Y, y_0)) = 0.$$

A subset S in a topological space X is **precompact** if any sequence in S has a subsequence that converges to a point in X .

Theorem 1.1.16. (Gromov's precompactness theorem) *Let*

$$\begin{aligned} \mathfrak{M} &:= \text{a collection of compact metric spaces,} \\ \mathfrak{M}^{\text{pt}} &:= \text{a collection of pointed metric spaces.} \end{aligned}$$

(1) *If \mathfrak{M} is uniformly totally bounded, that is,*

- (i) *there exists $D < \infty$ such that $\text{diam}(X) \leq D$ for any $X \in \mathfrak{M}$, and*
- (ii) *for any $\epsilon > 0$ there exists $N(\epsilon) \in \mathbf{N}$ such that any $X \in \mathfrak{M}$ contains an ϵ -net $S \subseteq X$ consisting of at most $N(\epsilon)$ points,*

then \mathfrak{M} is precompact in the collection of metric spaces with respect to the Gromov-Hausdorff distance.

(2) *If \mathfrak{M}^{pt} has the property that*

for any $\epsilon, \rho > 0$, there exists $N(\epsilon, \rho) \in \mathbf{N}$ such that any $(X, x_0) \in \mathfrak{M}^{\text{pt}}$ contains an ϵ -net $S \subseteq B(x_0, \rho) \subseteq X$ consisting of at most $N(\epsilon, \rho)$ points,

then \mathfrak{M}^{pt} is precompact in the collection of pointed metric spaces with respect to the pointed Gromov-Hausdorff distance.

As a corollary of volume comparison theorem, we have

Theorem 1.1.17. (Gromov) *Given $K \in \mathbf{R}$, $m \geq 2$, let*

$$\mathfrak{M}_{m,K}^{\text{pt}} := \left\{ \begin{array}{l} \text{pointed complete Riemannian } m\text{-manifolds } (\mathcal{M}, \mathbf{g}, O) \\ \text{with } \mathbf{Ric}_{\mathbf{g}} \geq (m-1)K\mathbf{g} \text{ and } O \in \mathcal{M} \end{array} \right\}.$$

Then $\mathfrak{M}_{m,K}^{\text{pt}}$ is precompact in the collection of pointed metric spaces with respect to pointed Gromov-Hausdorff distance.

The **tangent cone** of a metric space (X, d) at $x \in X$ is given by

$$(1.1.3.6) \quad (T_x X, d_x, 0_x) := \lim_{\alpha_k \rightarrow \infty} (X, \alpha_k d, x)$$

provided this limit (in the sense of pointed Gromov-Hausdorff distance) exists for any sequence $\alpha_k \rightarrow \infty$ and is independent of $\{\alpha_k\}_{k \in \mathbf{N}}$ (up to isometry).

- (1) The tangent cone is a metric space.
- (2) The tangent cone of a Riemannian manifold $(\mathcal{M}, \mathbf{g})$ at $p \in \mathcal{M}$ is isometric to $(T_x \mathcal{M}, \mathbf{g}(x), 0_x)$.
- (3) For any $c > 0$, we have an isometry $(T_x X, c d_x, 0_x) \cong (T_x X, d_x, 0_x)$.

The **Gromov-Hausdorff asymptotic cone** of a metric space (X, d) at $x \in X$ is given by

$$(1.1.3.7) \quad (\mathbf{A}X, d_{\mathbf{A}X}, 0) := \lim_{\omega_i \rightarrow 0} (X, \omega_i d, x)$$

provided this limit exists and is independent (up to isometry) of the sequence $\{\omega_i\}_{i \in \mathbf{N}}$.

- (1) For any $c > 0$, we have an isometry $(\mathbf{A}X, c d_{\mathbf{A}X}, 0) \cong (\mathbf{A}X, d_{\mathbf{A}X}, 0)$.
- (2) If $(\mathcal{M}, \mathbf{g})$ is a complete noncompact Riemannian manifold with nonnegative sectional curvature, then the asymptotic cone exists and is isometric to a Euclidean metric cone. (For a proof of this result, read Theorem I.26 in *The Ricci Flow: Techniques and Applications. Part III: Geometric-Analytic Aspects.*)

The **topological cone** of a topological space X is defined by

$$(1.1.3.8) \quad \mathbf{Cone}(X) := (X \times [0, \infty)) / (X \times \{0\}) = \{[(x, r)] : x \in X, r \in [0, \infty)\}.$$

We call the point $X \times \{0\}$ the **vertex** of $\mathbf{Cone}(X)$. It is clear that $\mathbf{Cone}(\mathbf{S}^{m-1}) \cong \mathbf{R}^m$.

We define the **Euclidean metric cone** of a metric space (X, d) with $\text{diam}(X) \leq \pi$ as follows. It is a topological cone $\mathbf{Cone}(X)$ equipped with the metric $d_{\mathbf{Cone}(X)}$ defined by

$$(1.1.3.9) \quad d_{\mathbf{Cone}(X)}([(x_1, r_1)], [(x_2, r_2)]) := \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(d(x_1, x_2))}.$$

One can show that $(\mathbf{Cone}(X), d_{\mathbf{Cone}(X)})$ is a metric space, and for any $c > 0$

$$d_{\mathbf{Cone}(X)}([(x_1, cr_1)], [(x_2, cr_2)]) = c d_{\mathbf{Cone}(X)}([(x_1, r_1)], [(x_2, r_2)]).$$

For Riemannian manifolds, we have the following equivalent description of Euclidean metric cones. Given a Riemannian manifold (\mathcal{M}, g) with diameter $\leq \pi$, define

$$(1.1.3.10) \quad g_{\text{Cone}} := r^2 g + dr \otimes dr$$

on $\mathcal{M} \times (0, \infty)$. We can show that

$$(1.1.3.11) \quad d_{g_{\text{Cone}}} = d_{\text{Cone}(\mathcal{M})}$$

on $\mathcal{M} \times (0, \infty) \subset \text{Cone}(\mathcal{M})$. Indeed, for any $x_1, x_2 \in \mathcal{M}$ and $r_1, r_2 \in (0, \infty)$, one has

$$(1.1.3.12) \quad d_{g_{\text{Cone}}}((x_1, r_1), (x_2, r_2)) = d_{\text{Cone}(\mathcal{M})}([(x_1, r_1)], [(x_2, r_2)]).$$

In fact, given two points $x_1, x_2 \in \mathcal{M}$, let $\gamma : [0, d(x_1, x_2)] \rightarrow \mathcal{M}$ be a unit speed minimal geodesic joining x_1 to x_2 , where $d := d_g$. Given $r_1, r_2 \in (0, \infty)$, join the two points $(x_1, r_1), (x_2, r_2) \in \mathcal{M} \times (0, \infty)$ by paths

$$\bar{\gamma} : [0, d(x_1, x_2)] \rightarrow \mathcal{M} \times (0, \infty), \quad u \mapsto (\gamma(u), r(u))$$

where $r : [0, d(x_1, x_2)] \rightarrow (0, \infty)$ satisfies $r(0) = r_1$ and $r(d(x_1, x_2)) = r_2$. The length of $\bar{\gamma}$ with respect to g_{Cone} is given by

$$L_{g_{\text{Cone}}}(\bar{\gamma}) = \int_0^{d(x_1, x_2)} \sqrt{r(u)^2 + \dot{r}(u)^2} du.$$

For any variation $\partial r = s$ with $s(0) = s(d(x_1, x_2)) = 0$, we have

$$\begin{aligned} \partial L_{g_{\text{Cone}}}(\bar{\gamma}) &= \int_0^{d(x_1, x_2)} [r(u)^2 + \dot{r}(u)^2]^{-1/2} [r(u)s(u) + \dot{r}(u)\dot{s}(u)] du \\ &= \int_0^{d(x_1, x_2)} [r(u)^2 + \dot{r}(u)^2]^{-1/2} r(u)s(u) \\ &\quad - \int_0^{d(x_1, x_2)} \left\{ - [r(u)^2 + \dot{r}(u)^2]^{-3/2} [r(u)\dot{r}(u) + \dot{r}(u)\ddot{r}(u)] \dot{r}(u) \right. \\ &\quad \left. + [r(u)^2 + \dot{r}(u)^2]^{-1/2} \ddot{r}(u) \right\} s(u) du \\ &= \int_0^{d(x_1, x_2)} [r(u)^2 + \dot{r}(u)^2]^{-3/2} [-r(u)\ddot{r}(u) + 2\dot{r}(u)^2 + r(u)^2] r(u)s(u) du. \end{aligned}$$

The Euler-Lagrange equation is now of the form

$$r(u) = \frac{1}{a \cos u + b \sin u}$$

for some $a, b \in \mathbf{R}$. Let $\alpha := d(x_1, x_2) \in [0, \pi]$ and consider the Euclidean planar triangle with side-angle-side equal to $r_1 - \alpha - r_2$ and vertices with polar coordinates $(r_1, 0)$, $(0, 0)$, and (r_2, α) . By the law of sines, we obtain

$$a = \frac{1}{r_1}, \quad b = \frac{r_2^{-1} - r_1^{-1} \cos \alpha}{\sin \alpha} = \frac{\frac{r_1}{r_2} - \cos \alpha}{r_1 \sin \alpha}$$

and

$$\begin{aligned} d_{g_{\text{Cone}}}((x_1, r_1), (x_2, r_2)) &= \int_0^\alpha \sqrt{r(u)^2 + \dot{r}(u)^2} du \\ &= \sqrt{a^2 + b^2} \int_0^\alpha \frac{du}{(a \cos u + b \sin u)^2} = \sqrt{a^2 + b^2} \frac{-\cos u}{b(b \sin u + a \cos u)} \Big|_0^\alpha \end{aligned}$$

$$\begin{aligned}
&= \sqrt{a^2 + b^2} \frac{ab^2 \sin \alpha}{b \sin \alpha + a \cos \alpha} = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \alpha} \\
&= d_{\text{Cone}(\mathcal{M}^m)}([(x_1, r_1)], [(x_2, r_2)]).
\end{aligned}$$

1.2. Topological spaces

A **topological space** is a pair (X, \mathcal{T}) , where X is a set and \mathcal{T} is a set of 2^X , satisfying

- (1) $U \cap V \in \mathcal{T}$ for any $U, V \in \mathcal{T}$;
- (2) $\cup_{i \in I} U_i \in \mathcal{T}$ for any collection $\{U_i\}_{i \in I} \subseteq \mathcal{T}$;
- (3) $\emptyset, X \in \mathcal{T}$.

An element of \mathcal{T} is called an **open subset** of X (or **open set** if both X and \mathcal{T} are understood), while a subset in X is called **closed** if its complement in X is open.

Suppose that (X, \mathcal{T}) is a topological space and $A \subseteq X$. Define

$$\mathcal{T}_A := \{A \cap U \mid U \in \mathcal{T}\}.$$

Then clearly (A, \mathcal{T}_A) is also a topological space.

If (X, \mathcal{T}) is a topological space, then, by definition, any intersection of finitely many open subsets of X is still an open subsets of X .

We have seen in [Section 1.1](#) that any metric space is always a topological space. Hence the Euclidean space \mathbf{R}^m is a topological space, whose topology is called the **Euclidean topology**. The following are some standard subsets of \mathbf{R}^m :

- The **unit interval**:

$$I = [0, 1] = \{x \in \mathbf{R} \mid 0 \leq x \leq 1\}.$$

- The **(open) unit ball** in \mathbf{R}^m :

$$\mathbf{B}^m := \left\{ x = (x^1, \dots, x^m) \in \mathbf{R}^m \mid |x| = \left(\sum_{1 \leq i \leq m} (x^i)^2 \right)^{1/2} < 1 \right\}.$$

When $m = 2$, we call $\mathbf{D} := \mathbf{B}^2$ the **(open) unit disk**.

- The **(closed) unit ball** in \mathbf{R}^m :

$$\mathbf{D}^m \equiv \overline{\mathbf{B}^m} \equiv \overline{\mathbf{B}^m} := \{x \in \mathbf{R}^m \mid |x| \leq 1\}.$$

When $m = 2$, we call $\mathbf{D} := \mathbf{D}^2$ the **closed unit disk**.

- The **(unit) circle**:

$$\mathbf{S}^1 := \{x \in \mathbf{R}^2 \mid |x| = 1\} = \{z \in \mathbf{C} \mid |z| = 1\}.$$

- The **(unit) m -sphere**:

$$\mathbf{S}^m := \{x \in \mathbf{R}^{m+1} \mid |x| = 1\}.$$

Let (X, \mathcal{T}) be a topological space. Then

- X and \emptyset are closed subsets of X .
- Any union of finitely many closed subsets of X is a closed subset of X .

- Any intersection of arbitrarily many closed subsets of X is a closed subset of X .

For example, let $X = \{1, 2, 3\}$ and

$$\mathcal{T} := \{\emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\}\}.$$

Then (X, \mathcal{T}) is a topological space. Clearly \emptyset , $\{3\}$, $\{2, 3\}$, and $\{1, 2, 3\}$ are all closed subsets of X . However, the subset $\{2\}$ is neither open nor closed of X . To describe $\{2\}$, introduce the following concepts.

Suppose that (X, \mathcal{T}) is a topological space and A is any subset of X .

- The **closure of A in X** is

$$\mathbf{Cle}(A) \equiv \bar{A} := \bigcap \{B \subset X \mid B \supset A \text{ and } B \text{ is closed in } X\}.$$

Then \bar{A} is closed in X .

- The **interior of A** is

$$\mathbf{Int}(A) \equiv \overset{\circ}{A} := \bigcup \{C \subset X \mid C \subset A \text{ and } C \text{ is open in } X\}.$$

Then $\overset{\circ}{A}$ is open in X .

- The **exterior of A** is

$$\mathbf{Ext}(A) := X \setminus \bar{A}.$$

- The **boundary of A** is

$$\mathbf{Bdy}(A) \equiv \partial A := X \setminus (\mathbf{Int}(A) \cup \mathbf{Ext}(A)).$$

By definition, we obtain

$$\overset{\circ}{A} \subseteq A \subseteq \bar{A}, \quad X = \mathbf{Int}(A) \bigsqcup \mathbf{Ext}(A) \bigsqcup \mathbf{Bdy}(A).$$

1.2.1. Continuous maps and bases. Let (X, \mathcal{T}) be a topological space. A **neighborhood** of $x \in X$ is a set $N \subseteq X$ containing an open subset $U \in \mathcal{T}$ with the property $x \in U \subset N$.

- (1) An neighborhood may not be open.
- (2) An neighborhood can be the entire space X .
- (3) The intersection of two neighborhoods of x is also a neighborhood of x .

A **neighborhood basis** at $x \in X$ is a set $\mathcal{B}_x \subset 2^X$ of neighborhoods of x with the property that every neighborhood N of x in X contains some $B \in \mathcal{B}_x$.

Exercise 1.2.1. Let (X, \mathcal{T}) be a topological space and $A \subseteq X$ be any subset. It is easy to prove the following statements:

- (1) A point is in $\mathbf{Int}(A)$ if and only if it has a neighborhood contained in A .
- (2) A point is in $\mathbf{Ext}(A)$ if and only if it has a neighborhood contained in $X \setminus A$.
- (3) A point is in $\mathbf{Bdy}(A)$ if and only if every neighborhood of it contains both a point of A and a point of $X \setminus A$.
- (4) A point is in \bar{A} if and only if every neighborhood of it contain a point of A .
- (5) $\bar{A} = A \cup \partial A = \overset{\circ}{A} \cup \partial A$.
- (6) $\mathbf{Int}(A)$ and $\mathbf{Ext}(A)$ are open in X , while \bar{A} and ∂A are closed in X .
- (7) The following are equivalent:
 - A is open in X .

- $A = \mathbf{Int}(A)$.
 - A contains none of its boundary points.
 - Every point of A has a neighborhood contained in A .
- (8) The following are equivalent:
- A is closed in X .
 - $A = \bar{A}$.
 - A contains all of its boundary points.
 - Every point of $X \setminus A$ has a neighborhood contained in $X \setminus A$.
- (9) $\overline{X \setminus A} = X \setminus \overset{\circ}{A}$ and $(X \setminus A)^\circ = X \setminus \bar{A}$.
- (10) Let \mathcal{A} be a collection of subsets of X . Show that

$$\overline{\bigcap_{A \in \mathcal{A}} A} \subseteq \bigcap_{A \in \mathcal{A}} \bar{A}, \quad \overline{\bigcup_{A \in \mathcal{A}} A} \supseteq \bigcup_{A \in \mathcal{A}} \bar{A}$$

and

$$\left(\bigcap_{A \in \mathcal{A}} A \right)^\circ \subseteq \bigcap_{A \in \mathcal{A}} \overset{\circ}{A}, \quad \left(\bigcup_{A \in \mathcal{A}} A \right)^\circ \supseteq \bigcup_{A \in \mathcal{A}} \overset{\circ}{A}.$$

Suppose that (X, \mathcal{T}) is a topological space and $A \subseteq X$ is a subset.

- We say that $x \in X$ is a **limit point (or accumulation point, cluster point) of A** if every neighborhood of x contains a point of A other than x (note that x itself may not be in A).
- We say that $x \in X$ is an **isolated point of A** if x has a neighborhood U in X such that $U \cap A = \{x\}$.
- A is said to be **dense in X** if $\bar{A} = X$.

Clearly every point of A is either a limit point or an isolated point, but not both.

Exercise 1.2.2. Let (X, \mathcal{T}) be a topological space and $A \subseteq X$. Show that

- A is closed if and only if it contains all of its limit points.
- A is dense if and only if every nonempty open subset of X contains a point of A .

Theorem 1.2.3. Let (X, d) be a metric space. If $A \subseteq X$, then

$$(1.2.1.1) \quad \bar{A} = \{x \in X : \lim_{n \rightarrow \infty} x_n = x \text{ for some sequence } \{x_n\}_{n \in \mathbb{N}} \subset A\}.$$

PROOF. If $x \in X$ with $\lim_{n \rightarrow \infty} x_n = x$ for some sequence $\{x_n\}_{n \in \mathbb{N}} \subset A$, any open set around x contains at least one point in A . Hence $x \in \bar{A}$. Conversely, for a point $x \in \bar{A}$, consider the ball $B(x, n^{-1})$. Then $x \in B(x, n^{-1}) \cap A$ for some x_n . Thus $d(x, x_n) < n^{-1}$. \square

A **continuous map** $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ between topological spaces is a map $f : X \rightarrow Y$ such that $f^{-1}(U) \in \mathcal{T}_X$ whenever $U \in \mathcal{T}_Y$.

We say that a map $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is **continuous at x** if for any neighborhood N of $f(x)$ in Y there exists a neighborhood M of x in X with $f(M) \subset N$. From the definition of neighborhoods, f is continuous at x if and only if for any neighborhood N of $f(x)$ in Y the inverse image $f^{-1}(N)$ is a neighborhood of x in X .

Theorem 1.2.4. A map $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is continuous if and only if f is continuous at each point in X .

PROOF. Suppose first that f is continuous. Let N be a neighborhood of $f(x)$ in Y and choose an open set U in Y with $f(x) \in U \subset N$. Then $x \in f^{-1}(U) \subset f^{-1}(N)$ and $f^{-1}(N)$ is a neighborhood of x in X . Hence f is continuous at x .

Conversely, given an open set $U \subset Y$. For any $x \in f^{-1}(U)$, we see that $f^{-1}(U)$ is a neighborhood of x so that $x \in V_x \subset f^{-1}(U)$ for some open set V_x in X . Then

$$f^{-1}(U) = \bigcup_{x \in f^{-1}(U)} V_x \subset \mathcal{T}_X$$

and f is continuous. \square

Example 1.2.5. (1) Every constant map $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is continuous.

(2) The identity map $\mathbf{1}_X : (X, \mathcal{T}_X) \rightarrow (X, \mathcal{T}_X)$ is continuous.

(3) If $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is continuous, so is the restriction of f to any open subset $U \subseteq X$.

(4) For continuous maps $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ and $g : (Y, \mathcal{T}_Y) \rightarrow (Z, \mathcal{T}_Z)$, the composite map $g \circ f : (X, \mathcal{T}_X) \rightarrow (Z, \mathcal{T}_Z)$ is also continuous. We also observe that for any continuous map $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ we have $f \circ \mathbf{1}_X = f$ and $\mathbf{1}_Y \circ f = f$.

We say a map $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is a **homeomorphism** if f is bijective and f, f^{-1} are continuous. In this case, we say that (X, \mathcal{T}_X) and (Y, \mathcal{T}_Y) are **homeomorphic** or **topologically equivalent**, and denote by $X \approx Y$.

Exercise 1.2.6. (1) Let $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ be a bijection between topological spaces. Show that f is homeomorphic if and only if $f(\mathcal{T}_X) = \mathcal{T}_Y$, that is, $U \in \mathcal{T}_X$ if and only if $f(U) \in \mathcal{T}_Y$.

(2) Let $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ be a homeomorphism and $U \in \mathcal{T}_X$. Show that $f(U) \in \mathcal{T}_Y$ and $f|_U$ is a homeomorphism from $(U, \mathcal{T}_U) \rightarrow (f(U), \mathcal{T}_{f(U)})$.

(3) Any **translation**

$$\mathbf{R}^m \longrightarrow \mathbf{R}^m, \quad x \longmapsto x + x_0$$

and **dilation**

$$\mathbf{R}^m \longrightarrow \mathbf{R}^m, \quad x \longmapsto \lambda x$$

are homeomorphic.

(4) The map $F : \mathbf{B}^m \rightarrow \mathbf{R}^m$ given by

$$F(x) := \frac{x}{1 - |x|}.$$

(5) Let $C := \{(x, y, z) \in \mathbf{R}^3 \mid \max\{|x|, |y|, |z|\} = 1\}$ be the cubical surface of side 2 centered at $(0, 0, 0)$. Define

$$F : C \longrightarrow \mathbf{S}^2, \quad (x, y, z) \longmapsto \frac{(x, y, z)}{\sqrt{x^2 + y^2 + z^2}}.$$

Show that F is homeomorphic.

(6) Let $X := [0, 1) \subseteq \mathbf{R}$ and $Y := \mathbf{S}^1 \subseteq \mathbf{C}$. Define

$$e : X \longrightarrow Y, \quad e(s) := e^{2\pi\sqrt{-1}s} = \cos(2\pi s) + \sqrt{-1}\sin(2\pi s).$$

Show that e is continuous and bijective, but is not homeomorphic.

A map $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is **open** if $f(U) \subseteq \mathcal{T}_Y$ for all $U \in \mathcal{T}_X$; f is **closed** if $f(C)$ is closed in Y for all closed set $C \subseteq X$.

Exercise 1.2.7. Let $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ be a bijective continuous map. Show that the following are equivalent:

- f is a homeomorphism.
- f is open.
- f is closed.

Proposition 1.2.8. Let $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ be a map between topological spaces. Show that

- (1) f is continuous if and only if $f(\overline{A}) \subseteq \overline{f(A)}$ for all $A \subseteq X$.
- (2) f is closed if and only if $f(\overline{A}) \supseteq \overline{f(A)}$ for all $A \subseteq X$.
- (3) f is continuous if and only if $f^{-1}(\overset{\circ}{B}) \subseteq (f^{-1}(B))^\circ$ for all $B \subseteq Y$.
- (4) f is open if and only if $f^{-1}(\overset{\circ}{B}) \supseteq (f^{-1}(B))^\circ$ for all $B \subseteq Y$.

PROOF. We only give proves of (1) and (2).

(1) Let f be a continuous map and A be a given subset of X . For any point $x \in \overline{A}$, we shall prove $f(x) \in \overline{f(A)}$. According to [Exercise 1.2.1](#) (4), it suffices to verify that every neighborhood N of $f(x)$ contains a point in $f(A)$. By definition, we have an open subset V of $f(x)$ such that $V \subseteq N$. Then, by continuity, $U := f^{-1}(V) \in \mathcal{T}_X$ and $x \in U$. Hence by [Exercise 1.2.1](#) (4) again, we can find a point $x' \in A \cap U$, so that $y' := f(x') \in f(A) \cap V$ and $f(x) \in \overline{f(A)}$.

Conversely, assume $f(\overline{A}) \subseteq \overline{f(A)}$ holds for all $A \subseteq X$. We first show that for any closed subset B of Y , $f^{-1}(B)$ is closed in X . Let $A := f^{-1}(B) \subseteq X$. To prove $f^{-1}(B)$ is closed in X , we verify that $\overline{A} = A$ by [Exercise 1.2.1](#) (8). If $x \in \overline{A}$, then

$$f(x) \in f(\overline{A}) \subseteq \overline{f(A)} = \overline{f(f^{-1}(B))} \subseteq \overline{B} = B$$

so that $x \in A$. Hence $\overline{A} = A$. Next we show the continuity of f . Let $V \in \mathcal{T}_Y$ and set $B := Y \setminus V$. Then

$$f^{-1}(B) = f^{-1}(Y) \setminus f^{-1}(V) = X \setminus f^{-1}(V).$$

Since $f^{-1}(B)$ is closed, it follows that $f^{-1}(V)$ is open.

(2) Let f be a closed map. Then for any $A \subseteq X$, we have that $f(\overline{A})$ is closed and then

$$f(A) \subseteq f(\overline{A}), \quad \overline{f(A)} \subseteq \overline{f(\overline{A})} = f(\overline{A}).$$

Conversely, assume that $f(\overline{A}) \supseteq \overline{f(A)}$ holds for all $A \subseteq X$. For a closed subset C of X , we get

$$\overline{f(C)} \subseteq f(\overline{C}) = f(C)$$

so that $f(C)$ is closed in Y . □

In [Proposition 1.2.8](#), together with [Theorem 1.2.4](#), we actually proved that for a map $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ between topological spaces, the following are equivalent:

- f is continuous,
- for any $x \in X$, f is continuous at x ,
- for any closed subset B of Y , $f^{-1}(B)$ is closed in X ,
- for any $A \subseteq X$, $f(\overline{A}) \subseteq \overline{f(A)}$.

A map $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ between topological spaces is said to be **locally homeomorphic** if every point $x \in X$ has a neighborhood $U \subseteq X$ such that $f(U) \in \mathcal{T}_Y$ and $f|_U : U \rightarrow f(U)$ is a homeomorphism.

Exercise 1.2.9. Show that

- Every homeomorphism is a local homeomorphism.
- Every local homeomorphism is continuous and open.
- Every bijective local homeomorphism is a homeomorphism.

Given a topological space (X, \mathcal{T}) . A **basis for \mathcal{T}** is a subset $\mathcal{B} \subseteq \mathcal{T}$ such that any open set U of X can be written as an union of some members of \mathcal{B} .

- (1) Let $\mathcal{B} \subseteq \mathcal{T}$ be a basis for \mathcal{T} . Then clearly
 - $X = \cup_{B \in \mathcal{B}} B$, and
 - if $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, there exists an element $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.
- (2) Consider a subset $\mathcal{C} \subseteq \mathcal{T}$ satisfying
 - for any $U \in \mathcal{T}$ and $x \in U$ there exists $C_x \in \mathcal{C}$ such that $x \in C_x \subseteq U$.

Then \mathcal{C} is a basis for \mathcal{T} . In fact, one has

$$U = \bigcup_{x \in U} C_x, \quad C_x \in \mathcal{C}.$$

Proposition 1.2.10. Let X be a set and suppose that \mathcal{B} is a collection of subsets of X . Then \mathcal{B} is a basis for some topology on X if and only if it satisfies the following two conditions:

- (1) $X = \cup_{B \in \mathcal{B}} B$.
- (2) If $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, there exists an element $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

If so, there is a unique topology on X for which \mathcal{B} is a basis, called the **topology generated by \mathcal{B}** .

PROOF. One direction is obvious. Suppose now that \mathcal{B} satisfies (1) and (2). Let $\mathcal{T}_{\mathcal{B}}$ be the collection of all unions of elements of \mathcal{B} .

We shall verify that $\mathcal{T}_{\mathcal{B}}$ is a topology on X .

- $X, \emptyset \in \mathcal{T}_{\mathcal{B}}$. By condition (1), $X \in \mathcal{T}_{\mathcal{B}}$, and $\emptyset \in \mathcal{T}_{\mathcal{B}}$ as the union of the empty collection of elements of \mathcal{B} .
- $\mathcal{T}_{\mathcal{B}}$ is closed under arbitrary unions. For any collection $\{U_i\}_{i \in I} \subseteq \mathcal{T}_{\mathcal{B}}$, write each U_i as

$$U_i = \bigcup_{j \in J_i} B_{ij}.$$

Then

$$\bigcup_{i \in I} U_i = \bigcup_{i \in I} \bigcup_{j \in J_i} B_{ij} \in \mathcal{T}_{\mathcal{B}}.$$

- $\mathcal{T}_{\mathcal{B}}$ is closed under finite intersections. It is enough to prove that whenever $U, V \in \mathcal{T}_{\mathcal{B}}$ one has $U \cap V \in \mathcal{T}_{\mathcal{B}}$. Given $x \in U \cap V$. Then there exist $B_1, B_2 \in \mathcal{B}$ such that $x \in B_1 \subseteq U$ and $x \in B_2 \subseteq V$. By condition (2), there exists $B_x \in \mathcal{B}$ such that $x \in B_x \subseteq B_1 \cap B_2 \subseteq U \cap V$. Consequently,

$$U \cap V = \bigcup_{x \in U \cap V} B_x$$

so that $U \cap V \in \mathcal{T}_{\mathcal{B}}$.

Next we show that \mathcal{B} is a basis for $\mathcal{T}_{\mathcal{B}}$. By definition of $\mathcal{T}_{\mathcal{B}}$, it is clear.

If \mathcal{T}' is another topology on X for which \mathcal{B} is a basis, then $\mathcal{B} \subseteq \mathcal{T}'$ and, for every $U \in \mathcal{T}'$, we have $U = \bigcup_{i \in I} B_i$, for a collection $\{B_i\}_{i \in I}$ of \mathcal{B} . Hence $U \in \mathcal{T}_{\mathcal{B}}$ and then $\mathcal{T}' \subseteq \mathcal{T}_{\mathcal{B}}$. Since $\mathcal{B} \subseteq \mathcal{T}'$, it follows that $\mathcal{T}_{\mathcal{B}} \subseteq \mathcal{T}'$. Therefore $\mathcal{T}' = \mathcal{T}_{\mathcal{B}}$. \square

Given a topological space (X, \mathcal{T}) . A **subbasis for \mathcal{T}** is a set $\mathcal{S} \subset 2^X$ such that \mathcal{B} is a basis for \mathcal{T} , where \mathcal{B} is the collection of all finite intersections of members of \mathcal{S} (so $\emptyset, X \in \mathcal{B}$).

Given any collection \mathcal{S} of subsets of any set X , we can construct a topology \mathcal{T} on X such that \mathcal{S} is a subbasis for \mathcal{T} . Indeed, we define \mathcal{T} to be the collection of arbitrary unions of the finite intersections of members of \mathcal{S} , assuming the convenience that \emptyset is the union of an empty collection of sets and X is the intersection of an empty collection of sets. This topology is called the **topology generated by \mathcal{S}** .

A topology space (X, \mathcal{T}) is said to be **first countable** if for any $x \in X$ there exists a countable **neighborhood basis at x** (that is, a collection \mathcal{B}_x of neighborhoods of x such that every neighborhood of contains some $B \in \mathcal{B}_x$). (X, \mathcal{T}) is said to be **second countable** if \mathcal{T} has a countable basis \mathcal{B} .

- (1) **Any metric space is first countable.** If (X, d) is a metric space, then d induces the metric topology \mathcal{T} . For each $x \in X$, the collection $\{B(x, r)\}_{r \in \mathbf{Q}^+}$ is a neighborhood basis at x .
- (2) **Euclidean spaces are second countable.** Consider the countable base $\mathcal{B} = \{B^m(x, r) \mid x \in \mathbf{Q}^m, r \in \mathbf{Q}^+\}$ of \mathbf{R}^m .
- (3) **Not every metric space is second countable**, for example $X = [0, 1]$ with the trivial metric.
- (4) **Any second countable space is first countable.**

Theorem 1.2.11. *Let (X, d) be a metric space. X is second countable if and only if X contains a countable dense subset.*

PROOF. If X is second countable, then X has a countable basis $\mathcal{B} = \{B_i\}_{i \in \mathbf{N}}$. Choose $x_i \in B_i$ for each $i \in \mathbf{N}$ and let $A := \{x_i\}_{i \in \mathbf{N}}$. Then A is a countable dense set.

Conversely, let $A = \{x_i\}_{i \in \mathbf{N}}$ be a countable dense set in X . If we take $\mathcal{B} = (B(x_i, r))_{i \in \mathbf{N}, r \in \mathbf{Q}_{>0}}$, then \mathcal{B} is a countable basis. \square

A topological space is said to be **separable** if it contains a countable dense subset.

Let (X, \mathcal{T}) be a topological space and (Y, d) is a metric space. We say a sequence of maps $\{f_n\}_{n \in \mathbf{N}}$ between X and Y **converges uniformly** to a map $f : X \rightarrow Y$ if

$$\lim_{n \rightarrow \infty} \left(\sup_{x \in X} d(f_n(x), f(x)) \right) = 0.$$

In this case, we write it as $f_n \rightrightarrows f$.

Theorem 1.2.12. *Let $\{f_n\}_{n \in \mathbf{N}}, f$ be maps between a topological space (X, \mathcal{T}) and a metric space (Y, d) . If f_n are continuous and $f_n \rightrightarrows f$, then f is continuous.*

PROOF. Given $\epsilon > 0$ there is an integer $n_0 \in \mathbf{N}$ such that $d(f(x), f(x_0)) < \frac{\epsilon}{3}$ for any $x \in X$ and $n \geq n_0$. Given $x_0 \in X$ there exists a neighborhood N of x_0 in X such that

$$d(f_{n_0}(x), f_{n_0}(x_0)) < \frac{\epsilon}{3}, \quad x \in N.$$

Hence

$$\begin{aligned} d(f(x), f(x_0)) &\leq d(f(x), f_{n_0}(x)) + d(f_{n_0}(x), f_{n_0}(x_0)) + d(f_{n_0}(x_0), f(x_0)) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

Thus f is continuous at x_0 . By [Theorem 1.2.3](#), we see that f is continuous. \square

Let X be a set and \mathbb{P} some condition given on subsets of X . Consider a topology \mathcal{T} on X whose open sets satisfy \mathbb{P} .

- (1) \mathcal{T} is the **smallest topology** or **coarsest topology** satisfying \mathbb{P} , if for any topology \mathcal{T}' on X whose open sets satisfying \mathbb{P} , then \mathcal{T} -open sets are \mathcal{T}' -open. In this case we use the notion $\mathcal{T} \prec \mathcal{T}'$.
- (2) \mathcal{T} is the **largest topology** or **finest topology** satisfying \mathbb{P} , if for any topology \mathcal{T}' on X whose open sets satisfying \mathbb{P} , then \mathcal{T}' -open sets are \mathcal{T} -open. In this case, we use the notion $\mathcal{T}' \prec \mathcal{T}$.

Example 1.2.13. (1) **(Trivial topology)** Any set X can be equipped with a topology $\mathcal{T}_t = \{\emptyset, X\}$. This is the smallest topology on X .

(2) **(Discrete topology)** Any set X can also be equipped with another topology $\mathcal{T}_d = 2^X$. This is the largest topology on X .

(3) **(Finite complement topology)** Given an infinite set X . The collection of subsets whose complement are finite, together with the empty set, forms a topology \mathcal{T}_{fc} .

(4) **(Countable complement topology)** Given an infinite set X . The collection of subsets whose complement are countable, together with the empty set, forms a topology \mathcal{T}_{cc} .

(5) **(Particular point topology)** Given an infinite set X and $x \in X$. Show that

$$\mathcal{T}_{pp} := \{U \subseteq X \mid U = \emptyset \text{ or } x \in U\}$$

is a topology on X .

(6) **(Excluded point topology)** Given an infinite set X and $x \in X$. Show that

$$\mathcal{T}_{\text{ep}} := \{U \subseteq X \mid U = X \text{ or } x \notin U\}$$

is a topology on X .

(7) Show that $(\mathbf{R}, \mathcal{T}_{\text{pp}})$ is first countable and separable, but not second countable.

(8) Show that $(\mathbf{R}, \mathcal{T}_{\text{ep}})$ is first countable, but not second countable or separable.

(9) Show that $(\mathbf{R}, \mathcal{T}_{\text{fc}})$ is separable, but not first or second countable.

1.2.2. Categories: an introduction. A category \mathcal{C} is a triple

$$(\mathbf{Ob}(\mathcal{C}), \mathbf{Hom}_{\mathcal{C}}(\cdot, \cdot), \circ),$$

where $\mathbf{Ob}(\mathcal{C})$ is a class of **objects**, $\mathbf{Hom}_{\mathcal{C}}(X, Y)$ is a set for every pair (X, Y) of objects (its element, called **morphism**, is usually written as $f : X \rightarrow Y$; we call X is the domain of f and Y the range of f), and \circ is the composition on sets (that is, given $X, Y, Z \in \mathbf{Ob}(\mathcal{C})$, we define $\circ : \mathbf{Hom}_{\mathcal{C}}(X, Y) \times \mathbf{Hom}_{\mathcal{C}}(Y, Z) \rightarrow \mathbf{Hom}_{\mathcal{C}}(X, Z)$ by $\circ(f, g) := g \circ f$), satisfying the following axioms

(1) **(Associativity)** $h \circ (g \circ f) = (h \circ g) \circ f$ for any morphisms

$$X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} W$$

(2) **(Identity)** For any $X \in \mathbf{Ob}(\mathcal{C})$ there exists a morphism $\mathbf{1}_X \in \mathbf{Hom}_{\mathcal{C}}(X, X)$, called the **identity morphism** of X , such that

$$\mathbf{1}_X \circ f = f, \quad g \circ \mathbf{1}_X = g$$

for any $f \in \mathbf{Hom}_{\mathcal{C}}(Y, X)$ and $g \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$. Note that $\mathbf{1}_X$ is unique.

\mathcal{C} is called **small** if $\mathbf{Ob}(\mathcal{C})$ is a set.

If morphisms $f \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$ and $g \in \mathbf{Hom}_{\mathcal{C}}(Y, X)$ satisfying $g \circ f = \mathbf{1}_X$, we say that g is a **left inverse** of f and f is a **right inverse** of g respectively. A **two-sided inverse** or an **inverse** of a morphism f is a morphism which is both a left inverse of f and a right inverse of f . A morphism $f : X \rightarrow Y$ is said to be an **equivalence** if there exists a morphism $g : Y \rightarrow X$ which is a two-sided inverse of f .

- (1) If $f : X \rightarrow Y$ has a left inverse and a right inverse, then they are equal and f is an equivalence. Indeed, if $g' : Y \rightarrow X$ is a left inverse and $g'' : Y \rightarrow X$ is a right inverse, then $g' \circ f = \mathbf{1}_X$ and $f \circ g'' = \mathbf{1}_Y$ so that $g' = g' \circ \mathbf{1}_Y = g' \circ (f \circ g'') = \mathbf{1}_X \circ g'' = g''$.
- (2) If $f : X \rightarrow Y$ is an equivalence, then by (1) it has a unique inverse $f^{-1} : Y \rightarrow X$ and f^{-1} is also an equivalence.
- (3) We say two objects X and Y are **equivalent**, written as $X \approx Y$, if there exists an equivalence $f : X \rightarrow Y$.
- (4) It is clear that \approx is an equivalence relation in any set of objects of \mathcal{C} .

An equivalence $f : X \rightarrow Y$ is also called an **isomorphism**. An **endomorphism** is a morphism $f : X \rightarrow X$ with same domain and range. An **automorphism** is an endomorphism which is an isomorphism. Two morphisms f and g are **parallel** if they have same domain and same range, i.e., $f, g : X \rightrightarrows Y$, or

$$X \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} Y.$$

A morphism $f : X \rightarrow Y$ is a **monomorphism** if for any pair of parallel morphisms $g_1, g_2 : Z \rightrightarrows X$, $f \circ g_1 = f \circ g_2$ implies $g_1 = g_2$.

$$Z \begin{array}{c} \xrightarrow{g_1} \\ \xrightarrow{g_2} \end{array} X \xrightarrow{f} Y$$

A morphism $f : X \rightarrow Y$ is an **epimorphism** if for any pair of parallel morphisms $g_1, g_2 : Y \rightrightarrows Z$, $g_1 \circ f = g_2 \circ f$ implies $g_1 = g_2$.

$$X \xrightarrow{f} Y \begin{array}{c} \xrightarrow{g_1} \\ \xrightarrow{g_2} \end{array} Z.$$

Note that $f : X \rightarrow Y$ is a **monomorphism** if and only if the map $f \circ : \mathbf{Hom}_{\mathcal{C}}(Z, X) \rightarrow \mathbf{Hom}_{\mathcal{C}}(Z, Y)$ is injective for any object $Z \in \mathbf{Ob}(\mathcal{C})$, and $f : X \rightarrow Y$ is an **epimorphism** if and only if the map $\circ f : \mathbf{Hom}_{\mathcal{C}}(Y, Z) \rightarrow \mathbf{Hom}_{\mathcal{C}}(X, Z)$ is injective for any object $Z \in \mathbf{Ob}(\mathcal{C})$.

If

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

are morphisms and if f and g are monomorphisms (resp. epimorphisms, resp. isomorphisms), then $g \circ f$ is a monomorphism (resp. epimorphism, resp. isomorphism).

We sometimes write $f : X \hookrightarrow Y$ or else $f : X \hookrightarrow Y$ to denote a monomorphism and $f : X \twoheadrightarrow Y$ to denote an epimorphism.

Example 1.2.14. (1) **Set** is the category of sets and set maps.

(2) **Group** is the category of groups and group homomorphisms.

(3) **AGroup** is the category of Abelian groups and group homomorphisms.

(4) **Ring** is the category of rings and ring homomorphisms.

(5) **CRing** is the category of commutative rings and ring homomorphisms.

(6) **Mod_R** is the category of R -modules and R -module homomorphisms.

(7) **Vect(**R**)** is the category of real vector spaces and **R**-linear maps.

(8) **Vect(**C**)** is the category of complex vector spaces and **C**-linear maps.

(9) **Top** is the the category of topological spaces and continuous maps.

(10) **Top_{*}** is the category of pointed topological space and pointed continuous maps.

A pointed topological space is a pair (X, x_0) where X is a topological space and $x_0 \in X$; a pointed continuous map $f : (X, x_0) \rightarrow (Y, y_0)$ is a continuous map $f : X \rightarrow Y$ such that $f(x_0) = y_0$.

The fundamental groups π_1 roughly speaking is a map from **Top_{*}** to **Group**; namely associated to a pointed topological space (X, x_0) a group $\pi_1(X, x_0)$. We also can define higher homotopy groups $\pi_i(X, x_0)$ which turns out to be Abelian for any $i \geq 2$.

Let \mathcal{C} be a category. A **subcategory** \mathcal{C}' of \mathcal{C} is itself a category and satisfies

- any object in \mathcal{C}' is also object in \mathcal{C} ,
- For any $X, Y \in \mathbf{Ob}(\mathcal{C}')$, we have $\mathbf{Hom}_{\mathcal{C}'}(X, Y) \subseteq \mathbf{Hom}_{\mathcal{C}}(X, Y)$, and
- $1_X \in \mathbf{Hom}_{\mathcal{C}'}(X, X)$ for any $X \in \mathbf{Ob}(\mathcal{C}')$.

A subcategory \mathcal{C}' is called a **full subcategory** if moreover

$$\mathbf{Hom}_{\mathcal{C}'}(X, Y) = \mathbf{Hom}_{\mathcal{C}}(X, Y)$$

for any $X, Y \in \mathbf{Ob}(\mathcal{C}')$.

Top is a subcategory, but not a full subcategory, of **Set**.

The **opposite category** \mathcal{C}° of a category \mathcal{C} is defined by

$$\mathbf{Ob}(\mathcal{C}^\circ) := \mathbf{Ob}(\mathcal{C}), \quad \mathbf{Hom}_{\mathcal{C}^\circ}(X, Y) := \mathbf{Hom}_{\mathcal{C}}(Y, X).$$

Exercise 1.2.15. Let \mathcal{C} be a category. An object $P \in \mathbf{Ob}(\mathcal{C})$ is **initial** if for any $Y \in \mathbf{Ob}(\mathcal{C})$, $\mathbf{Hom}_{\mathcal{C}}(P, Y)$ has exactly one element. An object $Q \in \mathbf{Ob}(\mathcal{C})$ is **final** if for any $X \in \mathbf{Ob}(\mathcal{C})$, $\mathbf{Hom}_{\mathcal{C}}(X, Q)$ has exactly one element. Show that two initial (resp. final) objects are isomorphic.

A **(covariant) functor** $F : \mathcal{C} \rightarrow \mathcal{C}'$ between two categories consists of

- (1) a map $F : \mathbf{Ob}(\mathcal{C}) \rightarrow \mathbf{Ob}(\mathcal{C}')$, and
- (2) for any two objects $X, Y \in \mathbf{Ob}(\mathcal{C})$, a map

$$F : \mathbf{Hom}_{\mathcal{C}}(X, Y) \rightarrow \mathbf{Hom}_{\mathcal{C}'}(F(X), F(Y)).$$

These data satisfy

$$(1.2.2.1) \quad F(\mathbf{1}_X) = \mathbf{1}_{F(X)}, \quad F(f \circ g) = F(f) \circ F(g).$$

A **contravariant functor** $G : \mathcal{C} \rightarrow \mathcal{C}'$ is a covariant functor from \mathcal{C}° to \mathcal{C}' .

Example 1.2.16. (1) Let \mathcal{C} be a category and take an object $X \in \mathbf{Ob}(\mathcal{C})$. Define

$$(1.2.2.2) \quad \mathbf{Hom}_{\mathcal{C}}(X, \cdot) : \mathcal{C} \longrightarrow \mathbf{Set}, \quad Z \longmapsto \mathbf{Hom}_{\mathcal{C}}(X, Z),$$

$$(1.2.2.3) \quad \mathbf{Hom}_{\mathcal{C}}(\cdot, X) : \mathcal{C} \longrightarrow \mathbf{Set}, \quad Z \longmapsto \mathbf{Hom}_{\mathcal{C}}(Z, X).$$

Then $\mathbf{Hom}_{\mathcal{C}}(X, \cdot)$ is a functor, while $\mathbf{Hom}_{\mathcal{C}}(\cdot, X)$ is a contravariant functor.

(2) The **forgetful functor** $F : \mathbf{Top} \rightarrow \mathbf{Set}$.

(3) The **fundamental group functor** $\pi_1 : \mathbf{Top}_* \rightarrow \mathbf{Group}$, given by $(X, x) \mapsto \pi_1(X, x)$.

Let $F_1, F_2 : \mathcal{C} \rightarrow \mathcal{C}'$ be two functors. A **morphism or natural transformation** Θ from F_1 to F_2 consists of

$$\Theta(X) \in \mathbf{Hom}_{\mathcal{C}'}(F_1(X), F_2(X)) \quad \text{whenever } X \in \mathbf{Ob}(\mathcal{C}).$$

These data satisfy the following commutative diagram:

$$(1.2.2.4) \quad \begin{array}{ccc} F_1(X) & \xrightarrow{\Theta(X)} & F_2(X) \\ F_1(f) \downarrow & & \downarrow F_2(f) \\ F_1(Y) & \xrightarrow{\Theta(Y)} & F_2(Y) \end{array} \quad F_2(f) \circ \Theta(X) = \Theta(Y) \circ F_1(f)$$

for any $X, Y \in \mathbf{Ob}(\mathcal{C})$ and any $f \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$.

Definition 1.2.17. For two categories \mathcal{C} and \mathcal{C}' , define a new category $\mathbf{Func}(\mathcal{C}, \mathcal{C}')$ as follows:

$$\mathbf{Ob}(\mathbf{Func}(\mathcal{C}, \mathcal{C}')) := \{\text{functors from } \mathcal{C} \text{ to } \mathcal{C}'\},$$

$$\mathbf{Hom}_{\mathbf{Func}(\mathcal{C}, \mathcal{C}')} (F_1, F_2) := \{\text{morphisms from } F_1 \text{ to } F_2\}.$$

Let \mathcal{C} be a category. A functor $F : \mathcal{C} \rightarrow \mathbf{Set}$ is said to be **representable** if there exists an object $X \in \mathbf{Ob}(\mathcal{C})$ such that F is isomorphic to $\mathbf{Hom}_{\mathcal{C}}(X, \cdot)$ in the category $\mathbf{Func}(\mathcal{C}, \mathbf{Set})$.

Exercise 1.2.18. (1) Show that $\mathbf{Func}(\mathcal{C}, \mathcal{C}')$ is a category.

(2) If $F : \mathcal{C} \rightarrow \mathbf{Set}$ is representable for some $X \in \mathbf{Ob}(\mathcal{C})$, then X is unique to isomorphism and is called a **representative** of F .

Let $F : \mathcal{C} \rightarrow \mathcal{C}'$ be a functor.

(1) F is **fully faithful** if for any objects $X, Y \in \mathbf{Ob}(\mathcal{C})$, the map

$$\mathbf{Hom}_{\mathcal{C}}(X, Y) \longrightarrow \mathbf{Hom}_{\mathcal{C}'}(F(X), F(Y))$$

is bijective.

(2) F is an **equivalence of categories** if F is fully faithful and for any object $X' \in \mathbf{Ob}(\mathcal{C}')$ there exists $X \in \mathbf{Ob}(\mathcal{C})$ and an isomorphism $X' \rightarrow F(X)$.

Exercise 1.2.19. $F \in \mathbf{Ob}(\mathbf{Func}(\mathcal{C}, \mathcal{C}'))$ is an equivalence of categories if and only if there exists $F' \in \mathbf{Ob}(\mathbf{Func}(\mathcal{C}', \mathcal{C}))$ and isomorphisms

$$F \circ F' \rightarrow \mathbf{1}_{\mathcal{C}'} \text{ in } \mathbf{Func}(\mathcal{C}', \mathcal{C}') \text{ and } F' \circ F \rightarrow \mathbf{1}_{\mathcal{C}} \text{ in } \mathbf{Func}(\mathcal{C}, \mathcal{C}).$$

Let \mathcal{C} be a category. Define

$$(1.2.2.5) \quad \mathcal{C}^{\vee} := \mathbf{Func}(\mathcal{C}^{\circ}, \mathbf{Set})$$

and

$$(1.2.2.6) \quad \mathbf{Hom}_{\mathcal{C}} : \mathcal{C} \longrightarrow \mathcal{C}^{\vee}, \quad X \longmapsto \mathbf{Hom}_{\mathcal{C}}(\cdot, X).$$

Theorem 1.2.20. (Yoneda's lemma) (1) For any $X \in \mathbf{Ob}(\mathcal{C})$ and any $F \in \mathbf{Ob}(\mathcal{C}^{\vee})$, one has

$$(1.2.2.7) \quad \mathbf{Hom}_{\mathcal{C}^{\vee}}(\mathbf{Hom}_{\mathcal{C}}(X, \cdot), F) \text{ is isomorphic to } F(X) \text{ in } \mathbf{Set}.$$

(2) $\mathbf{Hom}_{\mathcal{C}}$ is a fully faithful functor.

PROOF. (1) To $f \in \mathbf{Hom}_{\mathcal{C}^{\vee}}(\mathbf{Hom}_{\mathcal{C}}(\cdot, X), F)$, we associate $\phi(f) \in F(X)$ as follows:

$$f(X) : \mathbf{Hom}_{\mathcal{C}}(X, X) \longrightarrow F(X), \quad \mathbf{1}_X \longmapsto \phi(f) := f(X)(\mathbf{1}_X).$$

Conversely, to $s \in F(X)$, we can associate $\psi(s) \in \mathbf{Hom}_{\mathcal{C}^{\vee}}(\mathbf{Hom}_{\mathcal{C}}(\cdot, X), F)$ as follows: For $Y \in \mathbf{Ob}(\mathcal{C})$,

$$\mathbf{Hom}_{\mathcal{C}}(Y, X) \xrightarrow{F} \mathbf{Hom}_{\mathbf{Set}}(F(X), F(Y)) \xrightarrow{s} F(Y)$$

we define $\psi(s)(Y) := s \circ F$, where

$$s(f) := f(s), \quad f \in \mathbf{Hom}_{\mathbf{Set}}(F(X), F(Y)).$$

Then ϕ and ψ are inverse to each other.

(2) For any $X, Y \in \mathbf{Ob}(\mathcal{C})$,

$$\mathbf{Hom}_{\mathcal{C}^{\vee}}(\mathbf{Hom}_{\mathcal{C}}(\cdot, X), \mathbf{Hom}_{\mathcal{C}}(\cdot, Y)) \rightarrow \mathbf{Hom}_{\mathcal{C}}(\cdot, Y)(X) = \mathbf{Hom}_{\mathcal{C}}(X, Y)$$

is bijective by (1). □

1.2.3. Subspaces. Let (X, \mathcal{T}) be a topological space and $S \subseteq X$ is a subset. $S \subseteq X$ is a **subspace** if (S, \mathcal{T}_S) is a topological space, where $\mathcal{T}_S := \{U \cap S \mid U \in \mathcal{T}\}$ is the **relative topology** or **subspace topology**. The inclusion map

$$(1.2.3.1) \quad \iota_S : (S, \mathcal{T}_S) \longrightarrow (X, \mathcal{T})$$

is continuous.

Example 1.2.21. Consider the subspaces

$$S_1 := [0, 1] \cup (2, 3), \quad S_2 := \{1/n\}_{n \geq 1}$$

of \mathbf{R} . Note that $[0, 1]$ is not an open interval of \mathbf{R} , but is an open subset of S_1 , since $[0, 1] = S_1 \cap (-1, 2)$. In S_2 , the one-point sets $\{1/n\}$ are all open so the subspace topology on S_2 is discrete.

Let S be a subspace of (X, \mathcal{T}) . We say a subset $U \subseteq S$ is **relatively open** or **relatively closed in S** if U is open or closed in the subspace topology \mathcal{T}_S .

Proposition 1.2.22. Suppose that S is a subspace of a topological space (X, \mathcal{T}) .

- (1) If $U \subseteq S \subseteq X$, U is open in S and S is open in X , then U is open in X . The same is true with “closed” in place of “open”.
- (2) If U is a subset of S that is either open or closed in X , then it is also open or closed in S , respectively.
- (3) If $U \subseteq S \subseteq X$, then the closure of U in S is equal to $\bar{U} \cap S$.
- (4) If $U \subseteq S \subseteq X$, then the interior of U in S contains $\overset{\circ}{U} \cap S$.

PROOF. (1) U is open in S implies $U = S \cap V$ for some $V \in \mathcal{T}$. Because $S \in \mathcal{T}$, we must have $U \in \mathcal{T}$.

(2) Assume that U is open in X . Then $U = S \cap U$ is open in S .

(3) Write \bar{U}^S as the closure of U in S . We shall prove that $\bar{U}^S = \bar{U} \cap S$. Suppose $x \in \bar{U}^S$ and N is any neighborhood of x in X . Then $N' := N \cap S$ is a neighborhood of x in S , so $N' \cap U \neq \emptyset$. Because $U \subseteq S$, we get $N \cap U \neq \emptyset$ and $x \in \bar{U} \cap S$. Conversely, let $x \in \bar{U} \cap S$ and N' is any neighborhood of x in S . By definition, N' contains an open subset $V' \in \mathcal{T}_S$ satisfying $V' \ni x$. Write $V' = N \cap S$ for some $N \in \mathcal{T}$. Then N itself is a neighborhood of x in X , so $N \cap U \neq \emptyset$. Consequently, $V' \cap U \neq \emptyset$ and $N' \cap U \neq \emptyset$, so $x \in \bar{U}^S$.

(4) Write $\text{Int}_S(U)$ as the interior of U in S . Let $x \in \overset{\circ}{U} \cap S$. Then we can find a neighborhood $N \subseteq X$ containing x . Set $N' := N \cap S$ a neighborhood of x in S . Hence $x \in \text{Int}_S(U)$. \square

In **Proposition 1.2.22** (4), we can find an example so that $\overset{\circ}{U} \cap S \subsetneq \text{Int}_S(U)$. Indeed, consider $X = \mathbf{R}$, $S = [0, 1] \cup (2, 3)$ and $U = [0, 1]$. Then $\text{Int}_S(U) = [0, 1]$ but $\overset{\circ}{U} \cap S = (0, 1)$.

Theorem 1.2.23. (1) (**Characteristic property of the subspace topology**) Suppose that (X, \mathcal{T}) is a topological space and $S \subseteq X$ is a subspace. For any topological space (Y, \mathcal{T}_Y) , a map $f : (Y, \mathcal{T}_Y) \rightarrow (S, \mathcal{T}_S)$ is continuous if and only if the composite map $\iota_S \circ f : (Y, \mathcal{T}_Y) \rightarrow (X, \mathcal{T})$ is continuous:

$$\begin{array}{ccc} & & X \\ & \nearrow \iota_S \circ f & \uparrow \iota_S \\ Y & \xrightarrow{f} & S \end{array}$$

(2) **(Uniqueness of the subspace topology)** Suppose that S is a subset of a topological space (X, \mathcal{T}) . The subspace topology on S is the unique topology for which the characteristic property (1) holds.

PROOF. (1) Suppose that $\iota_S \circ f : (Y, \mathcal{T}_Y) \rightarrow (X, \mathcal{T})$ is continuous. If $U \in \mathcal{T}_S$, then $U = V \cap S = \iota_S^{-1}(V)$ for some $V \in \mathcal{T}$. Thus

$$f^{-1}(U) = f^{-1}(\iota_S^{-1}(V)) = (\iota_S \circ f)^{-1}(V) \in \mathcal{T}_Y.$$

Conversely, suppose that f is continuous. For any $V \in \mathcal{T}$, we have

$$(\iota_S \circ f)^{-1}(V) = f^{-1}(\iota_S^{-1}(V)) = f^{-1}(S \cap V) \in \mathcal{T}_S.$$

Hence $\iota_S \circ f$ is continuous.

(2) Suppose S is equipped with some topology \mathcal{T}'_S that satisfies the characteristic property (1). For convenience, let S_S denote the topological space (S, \mathcal{T}_S) and $S_{S'}$ denote the topological space (S, \mathcal{T}'_S) . To prove $\mathcal{T}_S = \mathcal{T}'_S$ we should verify that (by [Exercise 1.2.6](#)) the identity map $\mathbf{1}_S : S \rightarrow S$ is a homeomorphism between S_S and $S_{S'}$.

Clearly that the inclusion map $\iota_S : S_S \rightarrow (X, \mathcal{T})$ is continuous. Since $S_{S'}$ satisfies the characteristic property, it follows that the inclusion map $\iota_{S'} : S_{S'} \rightarrow (X, \mathcal{T})$ is also continuous. Consider the following two diagrams:

$$\begin{array}{ccc} & X & \\ \iota_{S'} \circ \mathbf{1}_{S_{S'}} \nearrow & \uparrow \iota_{S'} & \\ S_S & \xrightarrow{\mathbf{1}_{S_{S'}}} & S_{S'} \end{array} \quad \begin{array}{ccc} & X & \\ \iota_S \circ \mathbf{1}_{S'_S} \nearrow & \uparrow \iota_S & \\ S_{S'} & \xrightarrow{\mathbf{1}_{S'_S}} & S_S \end{array}$$

Here

$$\mathbf{1}_{S_{S'}} : S_S \longrightarrow S_{S'}, \quad \mathbf{1}_{S'_S} : S_{S'} \longrightarrow S_S.$$

Because $\iota_{S'} \circ \mathbf{1}_{S_{S'}} = \iota_S$ and $\iota_S \circ \mathbf{1}_{S'_S} = \iota_{S'}$, we, using again the characteristic property, obtain the continuities of both $\mathbf{1}_{S_{S'}}$ and $\mathbf{1}_{S'_S}$. Thus $\mathbf{1}_S : S \rightarrow S$ is a homeomorphism between S_S and $S_{S'}$. \square

Corollary 1.2.24. Let $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ be a continuous map between topological spaces.

- (1) **(Restricting the domain)** The restriction of f to any subspace $S \subseteq X$ is continuous.
- (2) **(Restricting the codomain)** If T is a subspace of Y that contains $f(X)$, then $f : X \rightarrow T$ is continuous.
- (3) **(Expanding the codomain)** If Y is a subspace of Z , then $f : X \rightarrow Z$ is continuous.

PROOF. (1) Because $f|_S = f \circ \iota_S$.

(2) By [Theorem 1.2.23](#).

(3) Because the map $X \rightarrow Z$ is the composite of $f : X \rightarrow Y$ with the inclusion $Y \hookrightarrow Z$. \square

Proposition 1.2.25. Suppose S is a subspace of a topological space (X, \mathcal{T}) .

- (1) If $R \subseteq S$ is a subspace of S , then R is a subspace of X .
- (2) If \mathcal{B} is a basis for \mathcal{T} , then $\mathcal{B}_S := \{B \cap S \mid B \in \mathcal{B}\}$ is a basis for \mathcal{T}_S .

- (3) *Every subspace of a first countable space is first countable.*
 (4) *Every subspace of a second countable space is second countable.*

PROOF. (1) It suffices to prove $(\mathcal{T}_S)_R = \mathcal{T}_R$. Let $U \in (\mathcal{T}_S)_R$. Then $U = R \cap V$ for some $V \in \mathcal{T}_S$. But V can be written as $V = S \cap W$ for some $W \in \mathcal{T}$, we obtain $U = R \cap (S \cap W) = R \cap W \in \mathcal{T}_R$.

Conversely, Let $U \in \mathcal{T}_R$. Then $U = R \cap W$ for some $W \in \mathcal{T}$ and

$$U = R \cap V, \quad V := S \cap W \in \mathcal{T}_S.$$

Hence $U \in (\mathcal{T}_S)_R$.

(2) By definition of subspace topology, $\mathcal{B}_S \subseteq \mathcal{T}_S$. For any $U \in \mathcal{T}_S$, we have $U = S \cap W$ for some $W \in \mathcal{T}$. Since \mathcal{B} is a basis for \mathcal{T} , it follows that

$$W = \bigcup_{i \in I} B_i, \quad B_i \in \mathcal{B}.$$

Therefore $U = \bigcup_{i \in I} B_i \cap S$ with $B_i \cap S \in \mathcal{B}_S$. Thus \mathcal{B}_S is a basis for \mathcal{T}_S .

(3) Assume that S is a subspace of a first countable topological space (X, \mathcal{T}) . For any $x \in S$, we can find a countable neighborhood basis $\mathcal{B}_x = \{B_i\}_{i \in \mathbb{N}}$ at x in X . Then $\mathcal{B}_{S,x} := \{B_i \cap S\}_{i \in \mathbb{N}}$ is a countable neighborhood basis at x in S , so (S, \mathcal{T}_S) is first countable.

(4) The proof is similar to (3). \square

Let (X, \mathcal{T}) be a topological space and S be closed in X . Clearly that if A is closed in S , then it is also closed in X .

Theorem 1.2.26. (Gluing lemma) *Let (X, \mathcal{T}) be a topological space and $X = A \cup B$, where A, B are closed in X .*

- (1) *If $f : A \rightarrow Y$ and $g : B \rightarrow Y$ are continuous and $f|_{A \cap B} = g|_{A \cap B}$, then f, g combine to give a continuous function $h : X \rightarrow Y$ given by*

$$h(x) = \begin{cases} f(x), & x \in A, \\ g(x), & x \in B. \end{cases}$$

- (2) *If $f : X \rightarrow Y$ is a map, then f is continuous provided $f|_A, f|_B$ are continuous.*

PROOF. Evidently (2) directly follows from (1). To prove (1), we use criterion of continuity below **Proposition 1.2.8**, that is, $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is continuous if and only if $f^{-1}(B)$ is closed in X for any closed subset B of Y .

For any closed subset C of Y , one has

$$h^{-1}(C) = f^{-1}(C) \cup g^{-1}(C).$$

Since f and g are continuous, we have that $f^{-1}(C)$ and $g^{-1}(C)$ are closed in A and B respectively. Hence $h^{-1}(C)$ is closed in X . \square

1.2.4. Separation axioms. Let (X, \mathcal{T}) be a topological space.

- (1) **T_0 -space:** for any $x \neq y \in X$, there exists $U \in \mathcal{T}$ containing x but not y , or $V \in \mathcal{T}$ containing y but not x .
 (2) **T_1 -space:** for any $x \neq y \in X$, there exist $U, V \in \mathcal{T}$ such that U contains x but not y and V contains y but not x .
 (3) **T_2 -space or Hausdorff space:** for any $x \neq y \in X$, there exist $U, V \in \mathcal{T}$ such that $x \in U, y \in V$, and $U \cap V = \emptyset$.

- (4) **T_3 -space or regular space:** it is T_1 -space, and for any $x \in X$, any closed set $F \subseteq X$ with $x \notin F$, there exist $U, V \in \mathcal{T}$ such that $x \in U, F \subseteq V$, and $U \cap V = \emptyset$.
- (5) **T_4 -space or normal space:** it is T_1 -space, and for any closed sets $F, G \subseteq X$ with $F \cap G = \emptyset$, there exist $U, V \in \mathcal{T}$ such that $F \subseteq U, G \subseteq V$, and $U \cap V = \emptyset$.

By definition, it is clear that

$$\{\text{normal spaces}\} \subsetneq \{\text{regular spaces}\} \subsetneq \{\text{Hausdorff spaces}\}.$$

Indeed, we shall prove that any Hausdorff space must be T_1 . By [Theorem 1.2.27](#) below, we shall prove that any one-point set $\{x\}$ in a Hausdorff space (X, \mathcal{T}) is closed. For any $y \in X$ and $y \neq x$, we can find disjoint open subsets $U, V \in \mathcal{T}$ such that $y \in U$ and $x \in V$ respectively. But $x \notin U$, we have $y \notin \overline{\{x\}}$. Thus $\overline{\{x\}} = \{x\}$.

We now give examples on strict inclusions.

- (1) **\mathbf{R}_K is Hausdorff but not regular.** Recall that \mathbf{R}_K denotes the real line with the topology generated by the basis consisting of all open intervals (a, b) and all subsets of the form $(a, b) \setminus K$, where $K = \{1/n\}_{n \in \mathbb{N}}$. Observe that K is closed in \mathbf{R}_K and does not contain 0. Suppose that there exist disjoint open sets U and V containing 0 and K , respectively. We shall find a contradiction. Choose a basis element, which must be of the form $(a, b) \setminus K$, containing 0 and lying in U . Take n sufficiently large so that $1/n \in (a, b)$ and then choose a basis element, which must be of the form (c, d) , about $1/n$ contained in V . Choose z satisfying $\max(c, 1/(n+1)) < z < 1/n$; then $z \in U \cap V$! Therefore \mathbf{R}_K is not regular.
- (2) **\mathbf{R}_ℓ is regular.** Recall that \mathbf{R}_ℓ denotes the real line with the topology generated by the basis consisting of all half-open intervals of the form $[a, b)$ with $a < b$. Clearly one-point sets are closed in \mathbf{R}_ℓ so that, according to [Theorem 1.2.27](#) below, \mathbf{R}_ℓ is T_1 . To prove normality, we suppose that A and B are disjoint closed subsets in \mathbf{R}_ℓ . For $a \in A$ choose a basis element $[a, x_a)$ not intersecting B and similarly for any $b \in B$ choose a basis element $[b, x_b)$ not intersecting A . Then open subsets

$$U := \bigcup_{a \in A} [a, x_a), \quad V := \bigcup_{b \in B} [b, x_b)$$

are disjoint open subsets about A and B respectively.

- (3) **The Sorgenfrey plane $\mathbf{R}_\ell^2 := \mathbf{R}_\ell \times \mathbf{R}_\ell$ is regular but not normal.**

Theorem 1.2.27. (1) *A topological space is T_1 if and only if one-point sets are closed sets.*

(2) *Any subspaces of a T_i -space is T_i , where $i = 2, 3$.*

(3) *A T_2 -space is T_3 if and only if for any $x \in X$ and any open subset U of x , there exists an open subset V of x such that $\overline{V} \subseteq U$.*

(4) *A T_2 -space is T_4 if and only if for any open set U and closed set A with $A \subseteq U$, there exists an open set V such that $A \subseteq V \subseteq \overline{V} \subseteq U$.*

(5) *Any metric space is normal.*

PROOF. (1) Let (X, \mathcal{T}) be a T_1 -space. Given any point $x \in X$. For any $y \in X \setminus \{x\}$ we have an open set $U_y \in \mathcal{T}$ such that $y \in U_y$ but $x \notin U_y$. Then $X \setminus \{x\} = \bigcup_{y \in X \setminus \{x\}} U_y$ is open and $\{x\}$ is closed in X .

Conversely, assume that a topological space (X, \mathcal{T}) satisfies the properties that any one-point set is closed. Consider any two distinct points $x, y \in X$. We have that $U_y := X \setminus \{x\}$, $V_x := X \setminus \{y\}$ are open. Therefore $U_x \in \mathcal{T}$ is an open set such that $y \in U_x$ but $x \notin U_x$. The similar result holds for V_y . Hence (X, \mathcal{T}) is T_1 .

(2) Assume first that (X, \mathcal{T}) is T_2 and S is a subspace of (X, \mathcal{T}) . For any $x \neq y$ in S , we can find two disjoint open subsets $U, V \in \mathcal{T}$ of x, y respectively, satisfying $U \cap V = \emptyset$. Then $(U \cap S) \cap (V \cap S) = \emptyset$ so (S, \mathcal{T}_S) is Hausdorff.

Next assume that (X, \mathcal{T}) is T_3 and S is subspace of (X, \mathcal{T}) . Clearly (S, \mathcal{T}_S) is T_1 space. Let $x \in S$ and B be a closed subset of S disjoint from x . Then

$$\overline{B} \cap S = \overline{B^S} = B$$

by **Proposition 1.2.22** (3), so $x \notin \overline{B}$. Using the regularity of X , there exist disjoint open subsets $U \ni x$ and $V \supseteq \overline{B}$. Then $U \cap S$ and $V \cap S$ are disjoint open subsets in Y containing x and B respectively. Therefore (S, \mathcal{T}_S) is regular.

(3) Let (X, \mathcal{T}) be a Hausdorff space. Suppose first that (X, \mathcal{T}) is regular. Take any $x \in X$ and any open subset U of x , and let $B := X \setminus U$. By regularity, there exist disjoint open subsets $V \ni x$ and $W \supset B$. We claim that $\overline{V} \cap B = \emptyset$. Otherwise, we can find $y \in B \cap \overline{V}$ so that W is a neighborhood of y but disjoint from V , contradicting $y \in \overline{V}$. Hence $\overline{V} \subseteq U$.

Conversely, suppose that $x \in X$ and a closed subset B not containing x . Let $U := X \setminus B$. By hypothesis, there exists an open subset V of x such that $\overline{V} \subseteq U$. The open subsets V and $X \setminus \overline{V}$ are disjoint open subsets containing x and B respectively.

(4) The proof is essentially similar to that of (3).

(5) Let (X, d) be a metric space with the induced metric topology \mathcal{T} , and A, B be disjoint closed subsets of X . For any $a \in A$, there exists $\epsilon_a > 0$ so that $B(a, \epsilon_a) \cap B = \emptyset$. Similarly, for any $b \in B$ there exists $\epsilon_b > 0$ so that $B(b, \epsilon_b) \cap A = \emptyset$. Define

$$U := \bigcup_{a \in A} B(a, \epsilon_a/2), \quad V := \bigcup_{b \in B} B(b, \epsilon_b/2).$$

Then U and V both are open and contain A and B respectively. We claim that $U \cap V = \emptyset$. Otherwise, for $z \in U \cap V$, we have $z \in B(a, \epsilon_a/2) \cap B(b, \epsilon_b/2)$ for some $a \in A$ and $b \in B$. Then

$$d(a, b) \leq d(a, z) + d(z, b) < \frac{\epsilon_a + \epsilon_b}{2}.$$

If $\epsilon_a \leq \epsilon_b$, then $d(a, b) < \epsilon_b$ so that $a \in B(b, \epsilon_b)$. If $\epsilon_b \leq \epsilon_a$, then $b \in B(a, \epsilon_a)$. Both give a contradiction. \square

Exercise 1.2.28. Prove (4) in **Theorem 1.2.27**.

Remark that a subspace of a normal space need not be normal. For example, consider the product space (for its definition see Section 1.3) \mathbf{R}^J , where J is uncountable. It can be proved that \mathbf{R}^J is regular but not normal. Moreover \mathbf{R}^J is homeomorphic to the subspace $(0, 1)^J$ of $[0, 1]^J$, which is normal.

Theorem 1.2.29. Any regular space with a countable basis is normal.

PROOF. Let (X, \mathcal{T}) be a regular space with a countable basis \mathcal{B} . Let A and B be disjoint closed subsets of X . For any $x \in A$ there exists an open subset U_x not intersecting B . By [Theorem 1.2.27](#) (3), we can find an open subset V_x of x such that $\overline{V_x} \subseteq U_x$. Choose an element of \mathcal{B} containing x and contained in V_x ; therefore we obtain a countable open subsets $\{U_n\}_{n \in \mathbf{N}} \subseteq \mathcal{B}$ such that

$$A \subseteq \bigcup_{n \in \mathbf{N}} U_n =: U, \quad \overline{U_n} \cap B = \emptyset.$$

Similarly, we can find another countable open subsets $\{V_n\}_{n \in \mathbf{N}}$ such that

$$B \subseteq \bigcup_{n \in \mathbf{N}} V_n =: V, \quad \overline{V_n} \cap A = \emptyset.$$

Then U and V are open subsets containing A and B respectively, but they may have nonempty intersection. Given $n \in \mathbf{N}$, set

$$U'_n := U_n \setminus \bigcup_{1 \leq i \leq n} \overline{V_i}, \quad V'_n := V_n \setminus \bigcup_{1 \leq i \leq n} \overline{U_i}.$$

Then U'_n and V'_n are open, and

$$A \subseteq \bigcup_{n \in \mathbf{N}} U'_n, \quad B \subseteq \bigcup_{n \in \mathbf{N}} V'_n.$$

Finally, define

$$U' := \bigcup_{n \in \mathbf{N}} U'_n, \quad V' := \bigcup_{n \in \mathbf{N}} V'_n.$$

Then $U' \cap V' = \emptyset$. Indeed, if $x \in U' \cap V'$, then $x \in U'_j \cap V'_k$ for some $j, k \in \mathbf{N}$. Assume $j \leq k$. Then $x \in U_j$ but $x \notin \bigcup_{1 \leq i \leq j} \overline{V_i}$; since $j \leq k$, $x \notin \bigcup_{1 \leq i \leq k} \overline{U_i}$ implies $x \notin \overline{U_j}$. Hence $U' \cap V' = \emptyset$. \square

Consider a metric space (X, d) , and A, B are disjoint closed subsets of X . Define a map

$$f : X \longrightarrow [0, 1], \quad x \longmapsto \frac{d(x, A)}{d(x, A) + d(x, B)}$$

where $d(x, A) := \inf\{d(x, a) \mid a \in A\}$. Clearly that

$$f|_A \equiv 0, \quad f|_B \equiv 1.$$

We claim that f is continuous. Consider a given point $x \in X$ and $\{x_n\}_{n \in \mathbf{N}} \subset X$ with $d(x_n, x) \rightarrow 0$ as $n \rightarrow +\infty$. We may assume that $x \notin A \cup B$, since A and B are closed. In this case, $d(x_n, A) > 0$ and $d(x_n, B) > 0$ for sufficiently large n . Compute

$$\begin{aligned} f(x_n) - f(x) &= \frac{d(x_n, A)}{d(x_n, A) + d(x_n, B)} - \frac{d(x, A)}{d(x, A) + d(x, B)} \\ &= \frac{d(x_n, A)d(x, B) - d(x, A)d(x_n, B)}{[d(x_n, A) + d(x_n, B)][d(x, A) + d(x, B)]} \\ &= \frac{[d(x_n, A) - d(x, A)]d(x, B) + d(x, A)[d(x, B) - d(x_n, B)]}{[d(x_n, A) + d(x_n, B)][d(x, A) + d(x, B)]}. \end{aligned}$$

From the triangle inequalities

$$|d(x_n, A) - d(x, A)| \leq d(x_n, x), \quad |d(x_n, B) - d(x, B)| \leq d(x_n, x),$$

we conclude that

$$|f(x_n) - f(x)| \leq \frac{d(x_n, x)}{d(x_n, A) + d(x_n, B)}$$

which tends to 0 as $n \rightarrow \infty$. Thus f is continuous.

According to **Theorem 1.2.27** (5), any metric spaces are normal. We ask whether the above construction holds for normal spaces.

Theorem 1.2.30. (Urysohn's lemma) *Let (X, \mathcal{T}) be a normal space and A, B be disjoint closed subsets of X . For any closed interval $[a, b] \subseteq \mathbf{R}$, there exists a continuous map $f : X \rightarrow [a, b]$ such that*

$$f|_A \equiv a \quad \text{and} \quad f|_B \equiv b.$$

PROOF. Since the map $[0, 1] \rightarrow [a, b], t \mapsto (1-t)a + tb$, is continuous, we may assume that $[a, b] = [0, 1]$.

Step 1. Let $P := [0, 1] \cap \mathbf{Q}$. We shall define for each $p \in P$ an open set $U_p \subseteq X$, in such a way that $\overline{U_p} \subseteq U_q$ whenever $p < q$. Moreover

$$A \subseteq U_0 \subseteq \overline{U_0} \subseteq U_1, \quad U_1 := X \setminus B.$$

Step 2. We define

$$U_p := \begin{cases} \emptyset, & \text{if } p < 0 \text{ and } p \in \mathbf{Q}, \\ X, & \text{if } p > 1 \text{ and } p \in \mathbf{Q}. \end{cases}$$

Clearly for any $p, q \in \mathbf{Q}$ with $p < q$, we still have $\overline{U_p} \subseteq U_q$.

Step 3. Given $x \in X$, define

$$\mathbf{Q}(x) := \{p \in \mathbf{Q} \mid U_p \ni x\} \subseteq \mathbf{Q}.$$

Then $\mathbf{Q} \cap (1, +\infty) \subset \mathbf{Q}(x)$. Thus $\mathbf{Q}(x)$ is bounded below and its greatest lower bound lies in $[0, 1]$. Define

$$f(x) := \inf \mathbf{Q}(x) = \inf \{p \in \mathbf{Q} \mid x \in U_p\}, \quad x \in X,$$

which is well-defined.

Step 4. If $x \in A$, then $x \in U_p$ for all $p \in \mathbf{Q} \cap [0, +\infty)$, so that $f(x) = 0$. If $x \in B$, then $x \in U_p$ for all $p \in \mathbf{Q} \cap (1, +\infty)$, so that $f(x) = 1$.

Step 5. We claim

$$x \in \overline{U_r} \implies f(x) \leq r \quad \text{and} \quad x \notin U_r \implies f(x) \geq r.$$

Indeed, if $x \in \overline{U_r}$, then $x \in U_s$ for all rational numbers $s > r$; therefore $f(x) \leq r$. If $x \notin U_r$, then $x \notin U_s$ for any rational numbers $s < r$; so $f(x) \geq r$.

Step 6. f is continuous. For any $x_0 \in X$ and any open interval $(c, d) \subseteq \mathbf{R}$ containing $f(x_0)$, we want to find an open subset $U \ni x_0$ such that $f(U) \subset (c, d)$. Choose $p, q \in \mathbf{Q}$ such that

$$c < p < f(x_0) < q < d.$$

By **Step 5**, $f(x_0) < q$ implies $x_0 \in U_q$, and $f(x_0) > p$ implies $x_0 \notin \overline{U_p}$. Hence $x_0 \in U := U_q \setminus \overline{U_p}$ which is open in X . To verify $f(U) \subset (c, d)$, take $x \in U$. Then $x \in U_q \subset \overline{U_q}$ so that $f(x) \leq q$ by **Step 5**. From $x \notin \overline{U_p}$ we get $f(x) \geq p$ again by **Step 5**. Thus $p \leq f(x) \leq q$ and $f(U) \subset (c, d)$.

Return back to **Step 1**. Because P is countable, we can arrange the elements of P in an infinite sequence, denote by \mathcal{P} , so that 1 and 0 are the first two elements of the sequence. Define

$$U_1 := X \setminus B \supseteq A.$$

By normality (see **Theorem 1.2.27** (4)), there exists an open subset U_0 such that $A \subset U_0 \subseteq \overline{U_0} \subseteq U_1$.

Let P_n ($n \geq 2$) denote the set of the first n rational numbers in \mathcal{P} . Assume that U_p is defined for all $p \in P_n$ satisfying the condition

$$(1.2.4.1) \quad p < q \implies \overline{U_p} \subset U_q.$$

Let r denote the next rational number in \mathcal{P} . Consider

$$P_{n+1} := P_n \cup \{r\}.$$

It is a finite subset of $[0, 1]$, and we can find $p, q \in P_{n+1}$ such that $p < r < q$ in the usual order relation and in such a way that there are no other elements of P_{n+1} in $[p, r]$ and $[r, q]$. Such p and q are called immediate predecessor and immediate successor respectively. Since p, q actually lie in P_n , we have $\overline{U_p} \subset U_q$. Applying the normality of X to the pair $(\overline{U_p}, X \setminus U_q)$ of disjoint closed subsets, there exists an open subset U_r such that $\overline{U_p} \subseteq U_r \subseteq \overline{U_r} \subseteq U_q$. We claim that (1.2.4.1) holds for every pair of elements of P_{n+1} . If both of them lie in P_n , (1.2.4.1) holds trivially. If one is r and another is $s \in P_n$, then either $s \leq p$ so that $\overline{U_s} \subset \overline{U_p} \subseteq U_r$, or $s \geq q$ so that $\overline{U_r} \subset \overline{U_q} \subseteq U_s$. Hence (1.2.4.1) also holds in P_{n+1} .

By induction on n , we have defined the open subset U_p for each $p \in P$. \square

If \mathcal{P} in **Step 1** is the following sequence

$$\mathcal{P} = \left\{ 1, 0, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \dots \right\},$$

then

$$P_3 = P_2 \cup \{1/2\} \implies 0 < \frac{1}{2} < 1 \implies U_0 \subseteq U_{\frac{1}{2}} \subseteq U_1,$$

$$P_4 = P_3 \cup \{1/3\} \implies 0 < \frac{1}{3} < \frac{1}{2} \implies U_0 \subseteq U_{\frac{1}{3}} \subseteq U_{\frac{1}{2}},$$

$$P_5 = P_4 \cup \{2/3\} \implies \frac{1}{2} < \frac{2}{3} < 1 \implies U_{\frac{1}{2}} \subseteq U_{\frac{2}{3}} \subseteq U_1,$$

$$P_6 = P_5 \cup \{1/4\} \implies 0 < \frac{1}{4} < \frac{1}{3} \implies U_0 \subseteq U_{\frac{1}{4}} \subseteq U_{\frac{1}{3}},$$

$$P_7 = P_6 \cup \{3/4\} \implies \frac{2}{3} < \frac{3}{4} < 1 \implies U_{\frac{2}{3}} \subseteq U_{\frac{3}{4}} \subseteq U_1,$$

$$P_8 = P_7 \cup \{1/5\} \implies 0 < \frac{1}{5} < \frac{1}{4} \implies U_0 \subseteq U_{\frac{1}{5}} \subseteq U_{\frac{1}{4}}.$$

Definition 1.2.31. Let A, B be two subsets of a topological space (X, \mathcal{T}) . We say that A and B **can be separated by a continuous function**, if there exists a continuous function $f : X \rightarrow [0, 1]$ such that $f|_A \equiv 0$ and $f|_B \equiv 1$.

Observe that the following are equivalent for a topological space (X, \mathcal{T}) :

- (1) (X, \mathcal{T}) is normal.
- (2) Every pair of disjoint closed subsets in X can be separated by disjoint open subsets.
- (3) Every pair of disjoint closed subsets can be separated by a continuous function.

Let (X, \mathcal{T}) be a topological space and $f : X \rightarrow \mathbf{R}$ be a continuous function. The **support of f** is

$$(1.2.4.2) \quad \mathbf{supp}(f) := \overline{\{x \in X \mid f(x) \neq 0\}}.$$

If A is a closed subset of X and U is an open subset containing A , a continuous function $f : X \rightarrow [0, 1]$ such that $f|_A \equiv 1$ and $\mathbf{supp}(f) \subseteq U$ is called a **bump function for A supported in U** .

Corollary 1.2.32. (Existence of bump functions) *Let (X, \mathcal{T}) be a normal space. If A is a closed subset of X and U is an open subset containing A , there exists a bump function for A supported in U .*

PROOF. By normality, we can find an open subset V such that $A \subseteq V \subseteq \overline{V} \subseteq U$. Apply **Theorem 1.2.30** to the pair $(A, B := X \setminus V)$. \square

Theorem 1.2.33. (Tietze's extension theorem) *Let (X, \mathcal{T}) be a normal space and $A \subseteq X$ be closed subset.*

- (1) *Any continuous function of A into $[a, b] \subset \mathbf{R}$ may be extended to a continuous map of X into $[a, b]$.*
- (2) *Any continuous function of A into \mathbf{R} may be extended to a continuous map of X into \mathbf{R} .*

PROOF. Given a continuous function $f : A \rightarrow [-r, r]$, we construct a continuous function $g : X \rightarrow \mathbf{R}$ such that

$$|g(x)| \leq \frac{1}{3}r, \quad x \in X \quad \text{and} \quad |g(a) - f(a)| \leq \frac{2}{3}r, \quad a \in A.$$

Define

$$I_1 := [-r, -r/3], \quad I_2 := [-r/3, r/3], \quad I_3 := [r/3, r], \quad B := f^{-1}(I_1), \quad C := f^{-1}(I_3).$$

Since f is continuous, it follows that B, C are closed disjoint subsets in A and then B, C are closed disjoint subset of X , by **Proposition 1.2.22** (1). According to **Theorem 1.2.30**, there exists a continuous function $g : X \rightarrow [-r/3, r/3]$ such that

$$g|_B \equiv -\frac{1}{3}r, \quad g|_C \equiv \frac{1}{3}r.$$

Then $|g(x)| \leq r/3$ for all $x \in X$. Given $a \in A$. If $a \in B$, then $|g(a) - f(a)| = |-r/3 - f(a)| \leq 2r/3$; if $a \in C$, then $|g(a) - f(a)| = |r/3 - f(a)| \leq 2r/3$; if $a \notin B \cup C$, then $f(a), g(a) \in I_2$ and $|g(a) - f(a)| \leq 2r/3$.

(1) Since $[a, b]$ is homeomorphic to $[-1, 1]$, we may assume $[a, b] = [-1, 1]$. Let $f : A \rightarrow [-1, 1]$ be a continuous map. Then there exists a continuous function $g_1 : X \rightarrow \mathbf{R}$ such that

$$|g_1(x)| \leq \frac{1}{3}, \quad x \in X \quad \text{and} \quad |g_1(a) - f(a)| \leq \frac{2}{3}, \quad a \in A.$$

The function $f - g_1$ maps A into $[-2/3, 2/3]$. We can find a continuous function $g_2 : X \rightarrow \mathbf{R}$ such that

$$|g_2(x)| \leq \frac{1}{3} \left(\frac{2}{3} \right), \quad x \in X \quad \text{and} \quad |f(a) - g_1(a) - g_2(a)| \leq \left(\frac{2}{3} \right)^2, \quad a \in A.$$

In general, for any $n \geq 2$, we can construct a continuous function g_n defined on X such that

$$|g_n(x)| \leq \frac{1}{3} \left(\frac{2}{3} \right)^{n-1}, \quad x \in X \quad \text{and} \quad \left| f(a) - \sum_{1 \leq i \leq n} g_i(a) \right| \leq \left(\frac{2}{3} \right)^n, \quad a \in A.$$

Define

$$g(x) := \sum_{n \geq 1} g_n(x), \quad x \in X.$$

Since

$$\sum_{1 \leq n \leq N} |g_n(x)| \leq \frac{1}{3} \sum_{1 \leq n \leq N} \left(\frac{2}{3} \right)^{n-1} = 1,$$

it follows that the above series is absolutely and uniformly convergent and then g is continuous on X . Clearly that $g(a) = f(a)$ for all $a \in A$, and g maps X into $[-1, 1]$.

(2) We may, without loss of generality, assume that $f : A \rightarrow (-1, 1) \cong \mathbf{R}$ is a continuous function. By (1), f can be extended to a continuous map $g : X \rightarrow [-1, 1]$. Define

$$D := g^{-1}(\{-1\}) \cup g^{-1}(\{1\})$$

which is closed in X . From $g|_A = f|_A \subseteq (-1, 1)$, we see that $A \cap D = \emptyset$. By [Theorem 1.2.30](#), there exists a continuous function $\phi : X \rightarrow [0, 1]$ such that

$$\phi|_D = 0, \quad \phi|_A = 1.$$

Define

$$h(x) := \phi(x)g(x).$$

Then g is continuous on X and is an extension of f :

$$h(a) = \phi(a)g(a) = g(a) = f(a), \quad a \in A.$$

For any $x \in X$, if $x \in D$ then $h(x) = 0 \cdot g(x) = 0$; if $x \notin D$ then $h(x) = \phi(x)g(x) \in (-1, 1)$. Thus h is a continuous from X into $(-1, 1) \cong \mathbf{R}$. \square

1.2.5. Topological manifolds. Let (X, \mathcal{T}) be a topological space.

- (1) A **cover** of X is a collection \mathcal{U} of subsets of X such that any point $x \in X$ is contained in some $U \in \mathcal{U}$. If each of the sets in \mathcal{U} is open (resp. closed), then we say \mathcal{U} is an **open cover** (resp. a **closed cover**).
- (2) Given a cover \mathcal{U} , a **subcover** of \mathcal{U} is a subcollection $\mathcal{U}' \subseteq \mathcal{U}$ that is still a cover of X .

Theorem 1.2.34. *Any open cover of a second countable topological space has a countable subcover.*

PROOF. Let \mathcal{B} be a countable basis for X and let \mathcal{U} be any open cover of X . Define

$$\mathcal{B}' := \{B \in \mathcal{B} : B \subset U \text{ for some } U \in \mathcal{U}\} \neq \emptyset.$$

It is clear that \mathcal{B}' is countable. For any $B \in \mathcal{B}'$ let $U_B \in \mathcal{U}$ be such that $U_B \supset B$ as above. Let

$$\mathcal{U}' := \{U_B \in \mathcal{U} : B \in \mathcal{B}'\}.$$

This is a countable subcover of X . □

A topological space (X, \mathcal{T}) is said to be a **Lindelöf space** if every open cover of X has a countable subcover.

Exercise 1.2.35. Recall Example 1.2.13. Show that $(\mathbf{R}, \mathcal{T}_{pp})$ is not Lindelöf, $(\mathbf{R}, \mathcal{T}_{ep})$ is Lindelöf, and $(\mathbf{R}, \mathcal{T}_{fc})$ is Lindelöf.

A topological space $\mathcal{M} = (\mathcal{M}, \mathcal{T})$ is **local Euclidean of dimension m** if for any $p \in \mathcal{M}$ there is an open set $\mathcal{U} \subset \mathcal{M}$ containing p , called **coordinate domain**, that is homeomorphic to an open subset U of \mathbf{R}^m . The homeomorphism is denoted by $\varphi : \mathcal{U} \rightarrow \varphi(\mathcal{U}) = U$, and is called a **coordinate map** on \mathcal{U} . The pair (\mathcal{U}, φ) is called a **coordinate chart** for \mathcal{M} . Write

$$(1.2.5.1) \quad \varphi(q) := (x^1(q), \dots, x^m(q)), \quad q \in \mathcal{U}.$$

We see that x^i can be viewed as a function on \mathcal{U} and $\varphi = (x^1, \dots, x^m)$.

Equivalently, \mathcal{M} is local Euclidean of dimension m if and only if any point $p \in \mathcal{M}$ has an open set $\mathcal{U} \subset \mathcal{M}$ of p that is homeomorphic to an open ball in \mathbf{R}^m . In this case we usually say \mathcal{U} is a **coordinate ball** in \mathcal{M} . We may also assume that $\varphi(p) = 0 \in \mathbf{R}^m$.

An **m -dimensional topological manifold** is a second countable Hausdorff space \mathcal{M} that is local Euclidean of dimension m . Write $\dim(\mathcal{M}) = m$ and call the **dimension** of \mathcal{M} .

A typical example of m -dimensional topological manifold is \mathbf{R}^m .

Proposition 1.2.36. *Every open subset of an m -dimensional topological manifold is still an m -dimensional topological manifold.*

PROOF. Let \mathcal{M} be an m -dimensional topological manifold and \mathcal{V} be an open subset of \mathcal{M} .

(1) \mathcal{V} is locally Euclidean of dimension m . For any $p \in \mathcal{V} \subseteq \mathcal{M}$, we have an open subset U in \mathcal{M} containing p , that is homeomorphic to an open subset U of \mathbf{R}^m . Then $U \cap \mathcal{V}$ still contains p , is open, and is homeomorphic to an open subset of \mathbf{R}^m .

(2) \mathcal{V} is second countable. By [Proposition 1.2.25](#) (4).

(3) \mathcal{V} is Hausdorff. By [Theorem 1.2.27](#) (2). \square

The **(closed) m -dimensional upper half-space** $\mathbf{H}^m \subseteq \mathbf{R}^m$ is defined by

$$(1.2.5.2) \quad \mathbf{H}^m := \{(x^1, \dots, x^m) \in \mathbf{R}^m \mid x^m \geq 0\}.$$

The boundary and interior of \mathbf{H}^m are

$$\partial\mathbf{H}^m = \{(x^1, \dots, x^m) \in \mathbf{H}^m \mid x^m = 0\}, \quad \text{Int}(\mathbf{H}^m) = \{(x^1, \dots, x^m) \in \mathbf{H}^m \mid x^m > 0\}.$$

When $m = 0$, we have $\mathbf{H}^0 = \mathbf{R}^0 = \{0\}$, so $\text{Int}(\mathbf{H}^0) = \mathbf{H}^0$ and $\partial\mathbf{H}^0 = \emptyset$.

When $m = 2$, we obtain the **upper half-plane**

$$(1.2.5.3) \quad \mathbf{H} \equiv \mathbf{H}^2 := \{\tau \in \mathbf{C} \mid \text{Im}(\tau) > 0\}.$$

Define the **modular group**

$$(1.2.5.4) \quad \mathbf{SL}_2(\mathbf{Z}) := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a, b, c, d \in \mathbf{Z}, ad - bc = 1 \right\}$$

and the associated map

$$(1.2.5.5) \quad \mathbf{SL}_2(\mathbf{Z}) \times \overline{\mathbf{C}} \longrightarrow \overline{\mathbf{C}}, \quad \left(\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tau \right) \longmapsto \gamma(\tau) := \frac{a\tau + b}{c\tau + d}.$$

Here $\overline{\mathbf{C}} := \mathbf{C} \cup \{\infty\}$ can be considered as the Riemann sphere \mathbf{S}^2 .

Exercise 1.2.37. (1) Let Γ be the subgroup of $\mathbf{SL}_2(\mathbf{Z})$ generated by

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Prove $\Gamma = \mathbf{SL}_2(\mathbf{Z})$.

(2) For any $\gamma, \gamma' \in \mathbf{SL}_2(\mathbf{Z})$ and any $\tau \in \mathbf{H}$, one has

$$(\gamma \cdot \gamma')(z) = \gamma(\gamma'(z)).$$

(3) Prove

$$\text{Im}(\gamma(\tau)) = \frac{\text{Im}(\tau)}{|c\tau + d|^2}, \quad \frac{d}{d\tau} \gamma(\tau) = \frac{1}{(c\tau + d)^2}, \quad \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{SL}_2(\mathbf{Z}).$$

Since $\mathbf{H}^m \subseteq \mathbf{R}^m$ we can give it the subspace topology of \mathbf{R}^m . An **m -dimensional topological manifold with boundary** is a second countable Hausdorff space \mathcal{M} in which every point $p \in \mathcal{M}$ has an open subset \mathcal{U} containing p , which is homeomorphic, under the map φ , either to an open subset of \mathbf{R}^m or to an open subset

of \mathbf{H}^m . The set \mathcal{U} is called a **coordinate domain** of p , φ is called a **coordinate map**, and the pair (\mathcal{U}, φ) is called a **coordinate chart** for \mathcal{M} .

We say (\mathcal{U}, φ) is an **interior chart** if $\varphi(\mathcal{U})$ is an open subset of \mathbf{R}^m , and a **boundary chart** if $\varphi(\mathcal{U})$ is an open subset of \mathbf{H}^m with $\varphi(\mathcal{U}) \cap \partial\mathbf{H}^m \neq \emptyset$.

Let \mathcal{M} be an m -dimensional topological manifold with boundary. A point $p \in \mathcal{M}$ is called an **interior point** of \mathcal{M} if it is in the domain of an interior chart, and it is called a **boundary point** of \mathcal{M} if it is in the domain of a boundary chart that takes p to $\partial\mathbf{H}^m$. The **boundary** of \mathcal{M} , denoted by $\partial\mathcal{M}$, is the set of all its boundary points, and its **interior**, denoted by $\mathbf{Int}(\mathcal{M})$ or $\overset{\circ}{\mathcal{M}}$, is the set of all its interior points.

Proposition 1.2.38. *If \mathcal{M} is an m -dimensional topological manifold with boundary, then $\mathbf{Int}(\mathcal{M})$ is an open subset of \mathcal{M} , which is itself an m -dimensional topological manifold.*

Exercise 1.2.39. Prove Proposition 1.2.38.

When \mathcal{M} is \mathbf{H}^m , an m -dimensional topological manifold with boundary, we have $\mathbf{H}^m = \mathbf{Int}(\mathbf{H}^m) \coprod \partial\mathbf{H}^m$, a disjoint union. Actually this fact holds for any m -dimensional topological manifold with boundary \mathcal{M} , i.e., $\mathcal{M} = \mathbf{Int}(\mathcal{M}) \coprod \partial\mathcal{M}$.

Proposition 1.2.40. *If \mathcal{M} is a nonempty m -dimensional topological manifold with boundary, then $\partial\mathcal{M}$ is closed in \mathcal{M} , and \mathcal{M} is an m -dimensional topological manifold if and only if $\partial\mathcal{M} = \emptyset$.*

PROOF. According to Proposition 1.2.38 and the above mentioned fact, $\partial\mathcal{M} = \mathcal{M} \setminus \mathbf{Int}(\mathcal{M})$ is closed. If \mathcal{M} is an m -dimensional topological manifold, then $\mathcal{M} = \mathbf{Int}(\mathcal{M})$ so that $\partial\mathcal{M} = \emptyset$. If $\partial\mathcal{M} = \emptyset$, then $\mathcal{M} = \mathbf{Int}(\mathcal{M})$ which is an m -dimensional topological manifold by Proposition 1.2.38. \square

Let $\mathcal{U} := \{U_i\}_{1 \leq i \leq n}$ be an open cover of a topological space (X, \mathcal{T}) . A **partition of unity subordinate to \mathcal{U}** is a finite family of continuous functions $\phi_i : X \rightarrow \mathbf{R}$, with the following properties:

- $0 \leq \phi_i(x) \leq 1$ for all i and all $x \in X$;
- $\text{Supp}(\phi_i) \subseteq U_i$;
- $\sum_{1 \leq i \leq n} \phi_i(x) = x$ for all $x \in X$.

Theorem 1.2.41. (Existence of finite partitions of unity) *If $\mathcal{U} = \{U_i\}_{1 \leq i \leq n}$ is a finite open cover of a normal space (X, \mathcal{T}) , then there exists a partition of unity subordinate to \mathcal{U} .*

PROOF. Basic idea is to construct two finite open covers $\mathcal{W} = \{W_i\}_{1 \leq i \leq n}$ and $\mathcal{V} = \{V_i\}_{1 \leq i \leq n}$ of X so that

$$\overline{W_i} \subseteq V_i \subseteq \overline{V_i} \subseteq U_i.$$

Then, by Theorem 1.2.30 we can find $\{\phi_i\}_{1 \leq i \leq n}$.

Step 1: We can shrink $\mathcal{U} = \{U_i\}_{1 \leq i \leq n}$ to an open cover $\mathcal{V} = \{V_i\}_{1 \leq i \leq n}$ of X such that $\overline{V_i} \subseteq U_i$ for each i .

Let $A := X \setminus \bigcup_{2 \leq i \leq n} U_i$. Then A is closed in X and is contained in U_1 . By normality, [Theorem 1.2.27](#) (4), there exists an open subset V_1 of X such that

$$A \subseteq V_1 \subseteq \overline{V_1} \subseteq U_1.$$

So $\{V_1, U_2, \dots, U_n\}$ is still an open cover of X . In general, given open subsets V_1, \dots, V_{k-1} such that $\{V_1, \dots, V_{k-1}, U_k, U_{k+1}, \dots, U_n\}$ is an open cover of X , let

$$A := X \setminus \left(\left(\bigcup_{1 \leq i \leq k-1} V_i \right) \cup \left(\bigcup_{k+1 \leq j \leq n} U_j \right) \right).$$

Then A is closed in X and is contained in U_k . By normality again, there exists an open subset V_k such that $A \subseteq V_k \subseteq \overline{V_k} \subseteq U_k$. Then $\{V_1, \dots, V_{k-1}, V_k, U_{k+1}, \dots, U_n\}$ is an open cover of X . After finite steps, we prove **Step 1**.

Step 2: Complete the proof. Using **Step 1** again, we can choose an open cover $\mathcal{W} = \{W_i\}_{1 \leq i \leq n}$ of X such that

$$\overline{W_i} \subseteq V_i \subseteq \overline{V_i} \subseteq U_i.$$

By [Theorem 1.2.30](#), for each i , there exists a continuous function $\psi_i : X \rightarrow [0, 1]$ such that $\psi_i(\overline{W_i}) = 1$ and $\psi_i(X \setminus V_i) = 0$. Because $\psi_i^{-1}(\mathbf{R} \setminus \{0\}) \subseteq V_i$, we have $\text{Supp}(\psi_i) \subseteq \overline{V_i} \subseteq U_i$. Since \mathcal{W} covers X , the finite sum $\sum_{1 \leq i \leq n} \psi_i(x) > 0$ for all $x \in X$. Define

$$\phi_i(x) := \frac{\psi_i(x)}{\sum_{1 \leq j \leq n} \psi_j(x)}, \quad x \in X.$$

Then $\{\phi_i\}_{1 \leq i \leq n}$ is a partition of unity subordinate to \mathcal{W} . □

1.3. Constructions

We in this section discuss two important constructions of new topological spaces from old one. At first we introduce topological embeddings.

An injective continuous map that is a homeomorphism onto its image (in the subspace topology) is called a **topological embedding**.

Proposition 1.3.1. (1) *A continuous injective map that is either open or closed is a topological embeddings.*

(2) *A surjective topological embedding is a homeomorphism.*

PROOF. (1) Let $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ be a continuous injective map. Then f defines a bijective map $f : X \rightarrow f(X)$ that is continuous. To prove $f : (X, \mathcal{T}_X) \rightarrow (f(X), \mathcal{T}_{f(X)})$ is homeomorphic, we need to check that it is open or closed by [Exercise 1.2.7](#).

If $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is open (resp. closed) and $A \in \mathcal{T}_X$, then $f(A)$ is open (resp. closed) in Y . By [Proposition 1.2.22](#) (2), $f(A)$ is also open (resp. closed) in $f(X)$. Thus $f : (X, \mathcal{T}_X) \rightarrow (f(X), \mathcal{T}_{f(X)})$ is open (resp. closed).

(2) Clearly. □

If S is a subspace of a topological space (X, \mathcal{T}) , then the inclusion map $\iota_S : (S, \mathcal{T}_S) \rightarrow (X, \mathcal{T})$ is a topological embedding.

Example 1.3.2. (Examples of topological embeddings) (1) Let $F : \mathbf{R} \rightarrow \mathbf{R}^2$ be the injective continuous map

$$F(s) := (s, s^2).$$

Its image is the parabola $P := \{(x, y) \in \mathbf{R}^2 \mid y = x^2\}$. Hence F is an injective continuous map. Denote $\pi : \mathbf{R}^2 \rightarrow \mathbf{R}$ by $\pi(x, y) := x$. Then $\pi|_P$ is continuous and gives the inverse of $F : \mathbf{R} \rightarrow P$. Therefore F is a topological embedding.

(2) Let $f : [0, 1) \rightarrow \mathbf{R}^2$ be the map $f(s) := (\cos(2\pi s), \sin(2\pi s))$. It is clear that f is not a homeomorphism onto its image in the subspace topology. However the restriction of f to any interval $[0, b)$ for $b \in (0, 1)$ is an topological embedding.

(3) Let $U \subseteq \mathbf{R}^n$ be an open subset and $\mathbf{U} \rightarrow \mathbf{R}^k$ be any continuous map. The **graph** of f is the subset $\Gamma_f \subseteq \mathbf{R}^{n+k}$ defined by

$$\Gamma_f := \{(x, y) \in U \times \mathbf{R}^k \mid y = f(x)\},$$

with the subspace topology inherited from \mathbf{R}^{n+k} . Define two maps

$$\Phi_f : U \longrightarrow \mathbf{R}^{n+k}, \quad x \longmapsto (x, f(x))$$

and

$$\pi : \mathbf{R}^{n+k} \longrightarrow \mathbf{R}^n, \quad (x^1, \dots, x^n, x^{n+1}, \dots, x^{n+k}) \longmapsto (x^1, \dots, x^n).$$

Then Φ_f is a continuous bijection from U onto Γ_f and $\pi|_{\Gamma_f}$ is a continuous inverse for Φ_f . Hence Φ_f is a topological embedding and Γ_f is homeomorphic to U , so that Γ_f is a topological manifold.

(4) Consider the **n -dimensional unit sphere or unit n -sphere**

$$\mathbf{S}^n = \left\{ x = (x^1, \dots, x^{n+1}) \in \mathbf{R}^{n+1} \mid \sum_{1 \leq i \leq n+1} (x^i)^2 = 1 \right\}.$$

For example, $\mathbf{S}^0 = \{-1, 1\}$, \mathbf{S}^1 is the unit circle in \mathbf{R}^2 , and \mathbf{S}^2 is the spherical surface of radius 1 in \mathbf{R}^3 . For each $i \in \{1, \dots, n+1\}$, define

$$U_i^\pm := \{x \in \mathbf{S}^n \mid \pm x^i > 0\}.$$

Then

$$\mathbf{S}^n = \bigcup_{1 \leq i \leq n+1} U_i^+ \cup \bigcup_{1 \leq i \leq n+1} U_i^-.$$

On each U_i^\pm , we see that $x \in \mathbf{S}^n \cap U_i^\pm$ if and only if

$$x^i = \pm \sqrt{1 - \sum_{1 \leq j \neq i \leq n+1} (x^j)^2}.$$

Hence $\mathbf{S}^n \cap U_i^\pm$ is the graph of a continuous function, is therefore locally Euclidean of dimension n . Thus \mathbf{S}^n is an n -dimensional topological manifold.

Using the stereographic projection we can give an explicit coordinate map. Let $N := (0, \dots, 0, 1) \in \mathbf{S}^n$ and consider the map

$$\varphi : \mathbf{S}^n \setminus \{N\} \longrightarrow \mathbf{R}^n, \quad (x^1, \dots, x^n, x^{n+1}) \longmapsto \left(\frac{x^1}{1 - x^{n+1}}, \dots, \frac{x^n}{1 - x^{n+1}} \right).$$

Evidently,

$$\varphi^{-1}(u^1, \dots, u^n) = \left(\frac{2u^1}{1 + |u|^2}, \dots, \frac{2u^n}{1 + |u|^2}, \frac{|u|^2 - 1}{1 + |u|^2} \right), \quad u = (u^1, \dots, u^n).$$

Similarly, we let $S := \{0, \dots, 0, -1\}$ and consider the map

$$\psi : \mathbf{S}^n \setminus \{S\} \longrightarrow \mathbf{R}^n, \quad (x^1, \dots, x^n, x^{n+1}) \longmapsto \left(\frac{x^1}{1+x^{n+1}}, \dots, \frac{x^n}{1+x^{n+1}} \right).$$

Evidently,

$$\psi^{-1}(u^1, \dots, u^n) = \left(\frac{2u^1}{1+|u|^2}, \dots, \frac{2u^n}{1+|u|^2}, \frac{1-|u|^2}{1+|u|^2} \right), \quad u = (u^1, \dots, u^n).$$

Consequently, $\{(\mathbf{S}^n \setminus N, \varphi), (\mathbf{S}^n \setminus S, \psi)\}$ covers \mathbf{S}^n .

(5) Suppose that \mathcal{M} is a topological manifold and $S \subset \mathcal{M}$. Under the relative topology, S is itself a topological space and furthermore a topological manifold. When \mathcal{U} is open, we say that \mathcal{U} is an **open submanifold** of \mathcal{M} .

Let us denote $\mathbf{GL}(n, \mathbf{R})$ the set of all $n \times n$ singular real matrices. Since any $n \times n$ matrix can be viewed as a point in \mathbf{R}^{n^2} , the “determinant function” given by

$$\det : \mathbf{R}^{n^2} \supset \mathbf{GL}(n, \mathbf{R}) \longrightarrow \mathbf{R}^{n^2}, \quad A \longmapsto \det(A)$$

shows that $\mathbf{GL}(n, \mathbf{R}) = \mathbf{R}^{n^2} \setminus \det^{-1}(0)$ is a topological manifold (actually it is an open submanifold).

If $f_i : X_i \rightarrow Y_i$ are maps for $i = 1, \dots, n$, their **product map** is

$$f_1 \times \dots \times f_n : X_1 \times \dots \times X_n \longrightarrow Y_1 \times \dots \times Y_n$$

given by

$$f_1 \times \dots \times f_n(x_1, \dots, x_n) := (f_1(x_1), \dots, f_n(x_n)).$$

1.3.1. Product spaces. Let (X_i, \mathcal{T}_i) , $1 \leq i \leq n$ be topological spaces. Define

$$\mathcal{B} := \{U_1 \times \dots \times U_n \mid U_i \in \mathcal{T}_i, 1 \leq i \leq n\}.$$

According to **Proposition 1.2.10**, \mathcal{B} is a basis for some topology. We call such a topology the **product topology** $\mathcal{T}_{X_1 \times \dots \times X_n}$ or $\mathcal{T}_1 \otimes \dots \otimes \mathcal{T}_n$ on $X_1 \times \dots \times X_n$. The topological space $(X_1 \times \dots \times X_n, \mathcal{T}_1 \otimes \dots \otimes \mathcal{T}_n)$ is called the **product space**. The basis subsets of the form $U_1 \times \dots \times U_n$ are called **product open subsets**.

For example, on $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$, the product topology is generated by sets of the form $I \times J$, where I and J are open subsets of \mathbf{R} . Hence the product topology on \mathbf{R}^2 is generated by open rectangle, so that it is the same as the metric topology induced by the Euclidean distance function.

Consider the above product space $(X_1 \times \dots \times X_n, \mathcal{T}_1 \otimes \dots \otimes \mathcal{T}_n)$. The canonical projection

$$(1.3.1.1) \quad \pi_i : X_1 \times \dots \times X_n \longrightarrow X_i$$

is actually continuous. Indeed, for any $U_i \in \mathcal{T}_i$ we have

$$\pi_i^{-1}(U_i) = X_1 \times \dots \times X_{i-1} \times U_i \times X_{i+1} \times \dots \times X_n \in \mathcal{T}_{X_1 \times \dots \times X_n}$$

so that π_i is continuous.

Theorem 1.3.3. (1) (**Characteristic property of product topologies**) Let $X_1 \times \cdots \times X_n$ be a product space. For any topological space B , a map $f : B \rightarrow X_1 \times \cdots \times X_n$ is continuous if and only if each of its component functions $f_i := \pi_i \circ f$ is continuous:

$$\begin{array}{ccc} & X_1 \times \cdots \times X_n & \\ & \nearrow f & \downarrow \pi_i \\ B & \xrightarrow{f_i} & X_i \end{array}$$

(2) (**Uniqueness of the product topology**) Let X_1, \dots, X_n be topological spaces. The product topology on $X_1 \times \cdots \times X_n$ is the unique topology that satisfies the characteristic property (1).

(3) Let X_1, \dots, X_n be topological spaces.

- (a) The projection maps $\pi_i : X_1 \times \cdots \times X_n \rightarrow X_i$ are all continuous and open.
- (b) The product topology is “associative” in the sense that the three product topologies $X_1 \times X_2 \times X_3$, $(X_1 \times X_2) \times X_3$, and $X_1 \times (X_2 \times X_3)$ on the set $X_1 \times X_2 \times X_3$ are all equal.
- (c) For any i and points $x_j \in X_j$, $j \neq i$, the map $f_i : X_i \rightarrow X_1 \times \cdots \times X_n$ given by

$$f_i(x) := (x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$$

is a topological embedding of X_i into the product space.

- (d) If for each i , \mathcal{B}_i is a basis for the topology of X_i , then the set $\{B_1 \times \cdots \times B_n : B_i \in \mathcal{B}_i\}$ is a basis for the product topology on $X_1 \times \cdots \times X_n$.
- (e) If A_i is the subspace of X_i for $i = 1, \dots, n$, the product topology and the subspace topology on $A_1 \times \cdots \times A_n \subseteq X_1 \times \cdots \times X_n$ are equal.
- (f) If each X_i is Hausdorff so is $X_1 \times \cdots \times X_n$.
- (g) If each X_i is second countable so is $X_1 \times \cdots \times X_n$.

(4) A product of continuous maps is continuous, and a product of homeomorphisms is a homeomorphism.

(5) If $\mathcal{M}_1, \dots, \mathcal{M}_n$ are topological manifolds of dimensions m_1, \dots, m_n , respectively, the product space $\mathcal{M}_1 \times \cdots \times \mathcal{M}_n$ is a topological manifold of dimension $m_1 + \cdots + m_n$.

PROOF. (1) Suppose first that all f_i are continuous. To prove f is continuous, it suffices to prove that $f^{-1}(U_1 \times \cdots \times U_n)$ is open in B . According to

$$f^{-1}(U_1 \times \cdots \times U_n) = f_1^{-1}(U_1) \cap \cdots \cap f_n^{-1}(U_n),$$

we see that f is continuous.

Conversely, if f is continuous, then for any $U_i \in \mathcal{T}_i$ we have

$$f_i^{-1}(U_i) = f^{-1}(X_1 \times \cdots \times X_{i-1} \times U_i \times X_{i+1} \times \cdots \times X_n)$$

that is open in B . Hence f_i is continuous.

(2) – (5): as an exercise. □

Exercise 1.3.4. Prove (2) – (5) in Theorem 1.3.3.

From Theorem 1.3.3, we have the m -torus

$$(1.3.1.2) \quad \mathbf{T}^m := \underbrace{\mathbf{S}^1 \times \cdots \times \mathbf{S}^1}_m$$

which is an m -dimensional topological manifold. When $m = 2$, we call \mathbf{T}^2 the **torus**.

For any indexed family $(X_\alpha)_{\alpha \in J}$ of sets, define its Cartesian product by

$$\prod_{\alpha \in J} X_\alpha := \left\{ \mathbf{x} : J \rightarrow \bigcup_{\alpha \in J} X_\alpha \mid x_\alpha := \mathbf{x}(\alpha) \in X_\alpha \right\}.$$

When $X_\alpha = X$ for all $\alpha \in J$, we write $X^J := \prod_{\alpha \in J} X_\alpha$.

Given a collection $(X_\alpha, \mathcal{T}_\alpha)_{\alpha \in J}$ of topological spaces, define the topology \mathcal{T} to be the topology generated by the basis

$$\left\{ \prod_{\alpha \in J} U_\alpha \mid U_\alpha \in \mathcal{T}_\alpha \text{ and } U_\alpha = X_\alpha \text{ for all but a finite number of } \alpha \right\}.$$

We call \mathcal{T} the **Tychonoff topology** or **product topology** on $\prod_{\alpha \in J} X_\alpha$. The **box topology** is generalized by the basis

$$\left\{ \prod_{\alpha \in J} U_\alpha \mid U_\alpha \in \mathcal{T}_\alpha \text{ for all } \alpha \in J \right\}.$$

Most theorems about finite products will also hold for arbitrary products if we use the product topology. Whenever we consider the product $\prod_{\alpha \in J} X_\alpha$ we shall assume it is given by the product topology.

Theorem 1.3.5. (1) *Suppose that the topology \mathcal{T}_α on each X_α is generated by a basis \mathcal{B}_α , $\alpha \in J$. Then the basis for the box topology on $\prod_{\alpha \in J} X_\alpha$ is*

$$\left\{ \prod_{\alpha \in J} B_\alpha \mid B_\alpha \in \mathcal{B}_\alpha \right\},$$

and the basis for the product topology on $\prod_{\alpha \in J} X_\alpha$ is

$$\left\{ \prod_{\alpha \in J} B_\alpha \mid \begin{array}{l} B_\alpha \in \mathcal{B}_\alpha \text{ for finitely many indices } \alpha \text{ and} \\ B_\alpha = X_\alpha \text{ for all the remaining indices} \end{array} \right\}.$$

(2) Let $\pi_\beta : \prod_{\alpha \in J} X_\alpha \rightarrow X_\beta$ be the natural projection, that is, $\pi_\beta(\mathbf{x}) = x_\beta$. Define

$$\mathcal{S} := \bigcup_{\beta \in J} \mathcal{S}_\beta, \quad \mathcal{S}_\beta := \left\{ \pi_\beta^{-1}(U_\beta) \mid U_\beta \in \mathcal{T}_\beta \right\}.$$

Then \mathcal{S} is a subbasis for the product topology.

(3) Let A_α , $\alpha \in J$, be subspace of X_α .

(3.1) $\prod_{\alpha \in J} A_\alpha$ is a subspace of $\prod_{\alpha \in J} X_\alpha$ if both products are given by the box product, or if both products are given by the product product.

(3.2) If $\prod_{\alpha \in J} X_\alpha$ is given either the box or the product topology, then

$$\prod_{\alpha \in J} \overline{A_\alpha} = \overline{\prod_{\alpha \in J} A_\alpha}.$$

(4) If each X_α , $\alpha \in J$, is Hausdorff, then $\prod_{\alpha \in J} X_\alpha$ is Hausdorff in both the box and product topologies.

(5) **(Characteristic property of infinite product spaces)** Let $(X_\alpha, \mathcal{T}_\alpha)_{\alpha \in J}$ be an indexed family of topological spaces. For any topological space (B, \mathcal{T}_B) , a map $f : B \rightarrow$

$\prod_{\alpha \in J} X_\alpha$ (with the product topology) is continuous if and only if its component functions $f_\alpha = \pi_\alpha \circ f$ is continuous. The product topology is the unique topology on $\prod_{\alpha \in J} X_\alpha$ that satisfies this property.

PROOF. (1) is clear.

(2) Let $\mathcal{B}_{\mathcal{S}}$ be the basis on $\prod_{\alpha \in A} X_\alpha$ consisting of finite intersections of elements of \mathcal{S} . Then for any $B \in \mathcal{B}_{\mathcal{S}}$ we have

$$B = \bigcap_{1 \leq i \leq n} \pi_{\beta_i}^{-1}(U_{\beta_i})$$

where β_1, \dots, β_n are distinct and $U_{\beta_i} \in \mathcal{T}_{\beta_i}$. Defining $U_\alpha = X_\alpha$ if $\alpha \neq \beta_1, \dots, \beta_n$, we obtain $B = \prod_{\alpha \in J} U_\alpha$ which is a basis element for the product topology.

(3) – (4): as an exercise.

(5) Consider the diagram

$$\begin{array}{ccc} B & \xrightarrow{f} & \prod_{\alpha \in J} X_\alpha \\ & \searrow f_\beta & \downarrow \pi_\beta \\ & & X_\beta \end{array}$$

Assume first that f is continuous. For any $U_\beta \in \mathcal{T}_\beta$, $\pi_\beta^{-1}(U_\beta)$ is a subbasis element for the product topology on $\prod_{\alpha \in J} X_\alpha$, we see that π_β is continuous and then $f_\beta = \pi_\beta \circ f$ is also continuous.

Conversely, assume that each f_α is continuous. For any subbasis element $\pi_\beta^{-1}(U_\beta)$ for the product topology on $\prod_{\alpha \in J} X_\alpha$, where $U_\beta \in \mathcal{T}_\beta$, one has

$$f^{-1}(\pi_\beta^{-1}(U_\beta)) = f_\beta^{-1}(U_\beta) \in \mathcal{T}_B.$$

Hence f is continuous. The uniqueness is similar to that of finite product topology. \square

Exercise 1.3.6. Prove **Theorem 1.3.5** (3) – (4).

Example 1.3.7. This example shows that **Theorem 1.3.5** (5) may not hold if the product topology is replaced by the box topology. Consider $\mathbf{R}^\omega := \prod_{n \geq 1} X_n$ with $X_n \equiv \mathbf{R}$, and define

$$\begin{array}{ccc} \mathbf{R} & \xrightarrow{f} & \mathbf{R}^\omega \\ & \searrow f_n & \downarrow \pi_n \\ & & \mathbf{R} = X_n \end{array}$$

where $f(t) := (t, t, t, \dots)$ and $f_n(t) := t$. If \mathbf{R}^ω is given the box topology, then f is not continuous. For example, let

$$B := \prod_{n \geq 1} \left(-\frac{1}{n}, \frac{1}{n} \right)$$

be an open subset of \mathbf{R}^ω relative to the box topology. We prove that $f^{-1}(B)$ is not open in \mathbf{R} . Otherwise, $f^{-1}(B)$ contains some interval $(-\delta, \delta)$ about $0 \in \mathbf{R}$. So $f((-\delta, \delta)) \subseteq B$ and

$$f_n((-\delta, \delta)) = (-\delta, \delta) \subseteq \left(-\frac{1}{n}, \frac{1}{n}\right)$$

for all $n \geq 1$, a contradiction!

Theorem 1.3.8. A topological space (X, \mathcal{T}) is Hausdorff if and only if $\Delta_X := \{(x, y) \in X \times X : x = y\} \subset X \times X$ is closed in $(X \times X, \mathcal{T}_{X \times X})$.

PROOF. Suppose first that Δ_X is closed in $X \times X$ with respect to the product topology. Then $X \times X \setminus \Delta_X$ is open. If $x \neq y$ in X , then $(x, y) \in U \times V \subset \mathcal{T}_{X \times X} \cap (X \times X \setminus \Delta_X)$. Hence X is Hausdorff.

For any $(x, y) \in X \times X \setminus \Delta_X$, we have $x \in U \in \mathcal{T}_X$ and $y \in V \in \mathcal{T}_Y$ for some open sets U, V with $U \cap V = \emptyset$. Hence $(x, y) \in U \times V$. If $U \times V \subset \Delta_X$, then $z \in U \cap V$, a contradiction. Therefore $U \times V \subset X \times X \setminus \Delta_X$ and Δ_X is closed. \square

1.3.2. Quotient spaces. Let (X, \mathcal{T}_X) be a topological space, Y be a set, and $q : X \rightarrow Y$ a surjective map.

(1) Define a topology \mathcal{T}_Y on Y by

$$(1.3.2.1) \quad \mathcal{T}_Y := \{V \subseteq Y : q^{-1}(V) \in \mathcal{T}_X\}.$$

This is called the **quotient topology** induced by q . In this topology, $q : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is continuous (called a **quotient map**).

(2) (Y, \mathcal{T}_Y) is a topological space and \mathcal{T}_Y is the largest topology on Y which makes q continuous.

Let (X, \mathcal{T}_X) be a topological space and \sim be an equivalence relation on X .

(1) Let $Y := X / \sim := \{[x] : x \in X\}$ be the set of equivalence classes, and define

$$\pi : X \longrightarrow Y, \quad x \longmapsto [x],$$

be the natural projection.

(2) X / \sim together with the topology determined by π is called the **quotient space** of X by the given equivalence relation. The quotient map π is called the **projection**.

More generally, a surjective map $\pi : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ between topological spaces is called a **quotient map** if the following condition holds:

$$V \in \mathcal{T}_Y \iff \pi^{-1}(V) \in \mathcal{T}_X.$$

It is clear that a quotient map is continuous.

Example 1.3.9. (1) **(Quotient spaces may not be Hausdorff)** Let $X = \mathbf{R}$ and $x \sim y$ if and only if $x - y \in \mathbf{Q}$. Then X / \sim is noncountable. If $\mathcal{T}_X := \mathcal{T}_{\mathbf{R}}$, then $\mathcal{T}_{X/\sim}$ is the trivial topology on X / \sim .

(2) Let $\mathbf{P}^n := (\mathbf{R}^{n+1} \setminus \{0\}) / \sim$ be the quotient space of $\mathbf{R}^{n+1} \setminus \{0\}$ where the equivalence relation is defined by

$$x \sim y \iff y = \lambda x \text{ for some } \lambda \neq 0.$$

Equivalently, \mathbf{P}^n is the set of 1-dimensional linear subspaces in \mathbf{R}^{n+1} . Denote the quotient map by

$$(1.3.2.2) \quad \pi : \mathbf{R}^{n+1} \setminus \{0\} \longrightarrow \mathbf{P}^n, \quad x \longmapsto [x] = [x^1, \dots, x^{n+1}].$$

If \mathcal{U}_i is given by

$$\mathcal{U}_i := \left\{ [x] \in \mathbf{P}^n : x^i \neq 0 \right\}, \quad 1 \leq i \leq n+1,$$

then $(\mathcal{U}_i)_{1 \leq i \leq n+1}$ is an open cover of \mathbf{P}^n with respect to the quotient topology induced by π . Define

$$\varphi_i([x]) := \left(\frac{x^1}{x^i}, \dots, \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, \dots, \frac{x^{n+1}}{x^i} \right);$$

then φ_i is homeomorphic and hence \mathbf{P}^n is locally Euclidean. **Lemma 1.3.11** below implies that \mathbf{P}^n is second countable. The Hausdorff property of \mathbf{P}^n will be proved in Section 1.5.

(3) Let (X, \mathcal{T}) be any topological space and A be any subset of X . Define a relation \sim on X by

$$x_1 \sim x_2 \iff (x_1, x_2) \in A \times A \text{ or } (X \setminus A) \times (X \setminus A).$$

Clearly that \sim is an equivalence relation on X . The quotient space $X \setminus A := X / \sim$ is said to be obtained by **collapsing A to a point**.

For example, $\overline{\mathbf{B}^n} \setminus \mathbf{S}^{n-1}$ is homeomorphic to \mathbf{S}^{n-1} .

(4) If (X, \mathcal{T}) is any topological space, the quotient

$$(1.3.2.3) \quad \mathbf{C}(X) := (X \times [0, 1]) / (X \times \{0\})$$

obtained from $X \times [0, 1]$ by collapsing one end to a point is called the **cone on X** . For example, $\mathbf{C}(\mathbf{S}^n)$ is homeomorphic to \mathbf{B}^n .

(5) Let $(X_i, \mathcal{T}_i)_{1 \leq i \leq n}$ be a finite family of nonempty topological spaces. For each $i \in \{1, \dots, n\}$, let $x_i \in X_i$ be a fixed point. The **wedge product** $\bigvee_{1 \leq i \leq n} X_i$ of X_1, \dots, X_n is the quotient space obtained from the disjoint union $\bigsqcup_{1 \leq i \leq n} X_i$ by collapsing the set $\{x_1, \dots, x_n\}$ to a point.

For example, the wedge sum $\mathbf{R} \vee \mathbf{R}$ is homeomorphic to the union of the x -axis and the y -axis in the plane, and the wedge sum $\mathbf{S}^1 \vee \mathbf{S}^1$ is homeomorphic to the figure-eight space consisting of the union of the two circles of radius 1 centered at $(0, 1)$ and $(0, -1)$ in the plane.

(6) Quotient maps may not be open or closed. For example, consider

$$\pi_1 : \mathbf{R} \times \mathbf{R} \longrightarrow \mathbf{R}, \quad (x, y) \longmapsto x$$

and $A := \{(x, y) \in \mathbf{R} \times \mathbf{R} \mid x \geq 0 \text{ or } y = 0\}$. Let $q : A \rightarrow \mathbf{R}$ be obtained by restricting π_1 on A . Then q is a quotient map but is neither open or closed.

(7) Consider $X := [0, 1] \cup [2, 3] \subseteq \mathbf{R}$ with the subspace topology, $Y := [0, 2] \subseteq \mathbf{R}$ with the subspace topology. Define a map

$$p : X \longrightarrow Y, \quad x \longmapsto \begin{cases} x, & x \in [0, 1], \\ x - 1, & x \in [2, 3]. \end{cases}$$

Then p is surjective, continuous, closed, quotient, but is not open. For $A := [0, 1] \cup [2, 3] \subseteq X$, define $q : A \rightarrow Y$ to be the restriction of p on A ; then q is continuous, surjective, but is not quotient.

(8) Consider the projection $\pi_1 : \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}, (x, y) \mapsto x$. Then π_1 is continuous, surjective, open, but is not closed (consider $C := \{(x, y) \in \mathbf{R} \times \mathbf{R} \mid xy = 1\}$). Let $A := C \cup \{(0, 0)\}$ and $q := \pi_1|_A$; then q is continuous, surjective, but is not a quotient map.

(9) Define

$$q : \mathbf{R} \longrightarrow A := \{a, b, c\}, \quad x \longmapsto \begin{cases} a, & x > 0, \\ b, & x < 0, \\ c, & x = 0. \end{cases}$$

Then the quotient topology on A induced by q is $\mathcal{T}_A = \{\{a\}, \{b\}, \{a, b\}, \{a, b, c\}, \emptyset\}$.

If $q : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is a quotient map, a subset $U \subseteq X$ is said to be **saturated** (with respect to q) if $U = q^{-1}(V)$ for some subset $V \subseteq Y$.

- (1) U is saturated if and only if $U = q^{-1}(q(U))$.
- (2) A subset $q^{-1}(y) \subseteq X$ for $y \in Y$ is called a **fiber** of q ; a saturated set is one that is a union of fibers.

Theorem 1.3.10. (1) **(Characteristic property of quotient topologies)** Let $q : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ be a quotient map. For any topological space (B, \mathcal{T}_B) , a map $f : Y \rightarrow B$ is continuous if and only if the composite map $f \circ q$ is continuous:

$$\begin{array}{ccc} X & & \\ q \downarrow & \searrow f \circ q & \\ Y & \xrightarrow{f} & B \end{array}$$

(2) **(Characterization of quotient maps)** Let (X, \mathcal{T}_X) and (Y, \mathcal{T}_Y) be topological spaces, and let $q : X \rightarrow Y$ be any surjective map. Then q is a quotient map if and only if the characteristic property holds.

(3) **(Passing to the quotient)** Suppose that $q : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is a quotient map. (B, \mathcal{T}_B) is a topological space, and $f : (X, \mathcal{T}_X) \rightarrow (B, \mathcal{T}_B)$ is any continuous map that is constant on the fibers of q (i.e., if $q(x) = q(x')$, then $f(x) = f(x')$). Then there exists a unique continuous map $\tilde{f} : (Y, \mathcal{T}_Y) \rightarrow (B, \mathcal{T}_B)$ such that $f = \tilde{f} \circ q$:

$$\begin{array}{ccc} X & & \\ q \downarrow & \searrow f & \\ Y & \xrightarrow{\tilde{f}} & B \end{array}$$

In this situation, we say that the map f **passes to the quotient** or **descends to the quotients**.

(4) **(Uniqueness of quotient spaces)** Suppose $q_1 : (X, \mathcal{T}_X) \rightarrow (Y_1, \mathcal{T}_{Y_1})$ and $q_2 : (X, \mathcal{T}_X) \rightarrow (Y_2, \mathcal{T}_{Y_2})$ are quotient maps that make the same identifications (i.e., $q_1(x) = q_1(x')$ if and only if $q_2(x) = q_2(x')$). Then there is a unique homeomorphism $\varphi : (Y_1, \mathcal{T}_{Y_1}) \rightarrow (Y_2, \mathcal{T}_{Y_2})$ such that $\varphi \circ q_1 = q_2$.

(5) **(Composition property of quotient maps)** Suppose $q_1 : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ and $q_2 : (Y, \mathcal{T}_Y) \rightarrow (Z, \mathcal{T}_Z)$ are quotient maps. Then their composition $q_2 \circ q_1 : (X, \mathcal{T}_X) \rightarrow (Z, \mathcal{T}_Z)$ is also a quotient map.

(6) A continuous surjective map $q : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is a quotient map if and only if it takes saturated open sets to open sets, or saturated closed sets to closed set.

(7) Suppose that $q : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is a quotient map. The restriction of q to any saturated open or closed set is a quotient map.

(8) If $q : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is a surjective continuous map that is also an open or closed map, then it is a quotient map.

PROOF. (1) For any $U \in \mathcal{T}_B$, $f^{-1}(U)$ is open in Y if and only if $q^{-1}(f^{-1}(U)) = (f \circ q)^{-1}(U)$ is open in X .

(2): as an exercise.

(3) For any $y \in Y$, we choose some $x \in X$ with $q(x) = y$ (because q is surjective). Define

$$\tilde{f}(y) := f(x).$$

Since f is constant on the fibers of q , it follows that \tilde{f} is well-defined. The continuity of \tilde{f} follows from (1).

If there exists another continuous map $\tilde{f}' : (Y, \mathcal{T}_Y) \rightarrow (B, \mathcal{T}_B)$ such that $f = \tilde{f}' \circ q$. Then

$$\tilde{f}'(y) = \tilde{f}'(q(x)) = f(x) = (\tilde{f} \circ q)(x) = \tilde{f}(y).$$

(4) Consider three diagrams

$$\begin{array}{ccc} \begin{array}{ccc} X & & \\ q_1 \downarrow & \searrow q_2 & \\ Y_1 & \cdots \xrightarrow{\exists \tilde{q}_2} & Y_2 \end{array} & \begin{array}{ccc} X & & \\ q_2 \downarrow & \searrow q_1 & \\ Y_2 & \cdots \xrightarrow{\exists \tilde{q}_1} & Y_1 \end{array} & \begin{array}{ccc} X & & \\ q_1 \downarrow & \searrow q_1 & \\ Y_1 & \xrightarrow[\tilde{q}_1 \circ \tilde{q}_2]{\mathbf{1}_{Y_1}} & Y_1 \end{array} \end{array}$$

By (3), there exist \tilde{q}_2 and \tilde{q}_1 . Since

$$\tilde{q}_1 \circ \tilde{q}_2 \circ q_1 = \tilde{q}_1 \circ q_2 = q_1,$$

we obtain from (2) that $\tilde{q}_1 \circ \tilde{q}_2 = \mathbf{1}_{Y_1}$. Similarly we have $\tilde{q}_2 \circ \tilde{q}_1 = \mathbf{1}_{Y_2}$.

(5) – (6): as an exercise.

(7) Let A be a saturated open set and let $\pi = q|_A : A \rightarrow q(A)$. Given $V \subseteq q(A)$ and assume $\pi^{-1}(V)$ is open in A , we shall show that V is open in $q(A)$. We first claim

$$\pi^{-1}(V) = q^{-1}(V).$$

Indeed, if $V \subseteq q(A)$ and A is saturated, then $q^{-1}(V) \subseteq A$ (why?) and hence $\pi^{-1}(V) = q^{-1}(V)$ (why?)

Now $\pi^{-1}(V)$ is open in A and A is open in X , so that $\pi^{-1}(V)$ is open in X . From $q^{-1}(V) = \pi^{-1}(V)$, we see that $q^{-1}(V)$ is open in X . Since q is a quotient map, it follows that V is open in Y and in particular open in $q(A)$.

(8) Assume that q is open. Given $V \subseteq Y$ and assume that $q^{-1}(V)$ is open in X , we shall show that V is open in Y . Because q is surjective, we get $q(q^{-1}(V)) = V$. Hence $V = q(q^{-1}(V))$ is open in Y (because q is open). \square

Exercise 1.3.11. Prove Theorem 1.3.10 (2), (5), and (6).

Note that the product of two quotient maps need not be a quotient map. If $q : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is a quotient map and \mathcal{T}_X is Hausdorff, then \mathcal{T}_Y need not be Hausdorff.

Lemma 1.3.12. Suppose $\pi : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is a quotient map and X is second countable. If Y is locally Euclidean, it is second countable.

PROOF. Since Y is locally Euclidean, it follows that Y has an open cover $\mathcal{U} = (\mathcal{U}_\alpha)_{\alpha \in A}$ with \mathcal{U}_α is homeomorphic to some open ball. Then $\pi^{-1}(\mathcal{U}) := \{\pi^{-1}(\mathcal{U}_\alpha) : \mathcal{U}_\alpha \in \mathcal{U}\}$ is an open cover of X . The second countability of X shows that (by **Theorem 1.2.34**) there exists a countable subcover of $\pi^{-1}(\mathcal{U})$. Let $\mathcal{U}' \subseteq \mathcal{U}$ denote a subcover such that $\pi^{-1}(\mathcal{U}')$ is a countable subcover of X and \mathcal{U}' is a countable cover of Y by Euclidean balls. Each such ball has a countable basis, and the union of all these bases is a countable basis for Y . \square

1.3.3. The de Rham cohomology groups on \mathbf{R}^n . Let U be an open set in \mathbf{R}^n and $\Omega^p(U)$ denote the space of smooth p -forms on U generated by dx^I , where

$$(1.3.3.1) \quad dx^I := dx^{i_1} \wedge \cdots \wedge dx^{i_p}, \quad I = (1 \leq i_1 < \cdots < i_p \leq n).$$

over $C^\infty(U)$. Any element of $\Omega^p(U)$ can be written as

$$\omega = \sum_{|I|=p} \omega_I dx^I.$$

Then

$$(1.3.3.2) \quad \Omega^\bullet(U) := \bigoplus_{0 \leq p \leq n} \Omega^p(U)$$

is a graded algebra. Define a differential operator d (**exterior differentiation**) by

$$(1.3.3.3) \quad d : \Omega^p(U) \longrightarrow \Omega^{p+1}(U)$$

as follows:

(i) if $f \in \Omega^0(U)$, then

$$df = \sum_{1 \leq i \leq n} \frac{\partial f}{\partial x^i} dx^i = df;$$

(ii) if $\omega = \sum_{|I|=p} \omega_I dx^I \in \Omega^p(U)$, then

$$d\omega := \sum_{|I|=p} d\omega_I \wedge dx^I = \sum_{|I|=p} d\omega_I \wedge dx^I.$$

It is easy to show that $d^2 = d \circ d = 0$ and

$$(1.3.3.4) \quad d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^{\deg(\omega)} \omega \wedge d\eta,$$

where $\deg(\omega) = p$ if $\omega \in \Omega^p(U)$.

The pair $(\Omega^\bullet(U), d)$ is called the **de Rham cohomological complex** on U . We denote by $Z^p(U)$ and $B^p(U)$ the set of **closed p -forms** and **exact p -forms** respectively. Note that exact forms are always closed, however the closed forms may not be exact³. The **p -th de Rham cohomology group** of U is

$$(1.3.3.5) \quad H_{\text{DR}}^p(U) := \frac{Z^p(U)}{B^p(U)}, \quad 0 \leq p \leq n.$$

Set

$$(1.3.3.6) \quad H_{\text{DR}}^\bullet(U) := \bigoplus_{0 \leq p \leq n} H_{\text{DR}}^p(U).$$

³Consider $U = \mathbf{R}^2 \setminus \{(0,0)\}$ and $\omega = (xdy - ydx)/(x^2 + y^2)$

For any $[\omega], [\eta] \in H_{\mathbf{DR}}^{\bullet}(U)$ we can define

$$(1.3.3.7) \quad [\omega] \cdot [\eta] := [\omega \wedge \eta]$$

which is independent of the choice of representatives.

If $F : U \rightarrow V$ is a smooth map from an open subset $U \subset \mathbf{R}^n$ to another open subset $V \subset \mathbf{R}^m$, and if $\omega \in \Omega^p(V)$, the **pull-back** $F^*\omega$ is the smooth p -form on U obtained by making the substitution $\mathbf{y} = F(\mathbf{x}) = (F^1(\mathbf{x}), \dots, F^m(\mathbf{x}))$ in ω

$$(1.3.3.8) \quad (F^*\omega)(\mathbf{x}) = \sum_{|I|=p} \omega_I(F(\mathbf{x})) dF^{i_1} \wedge \dots \wedge dF^{i_p}.$$

If we have a second smooth map $G : V \rightarrow W$ and if η is a smooth form on W , then $F^*(G^*\eta)$ is obtained by substituting $\mathbf{z} = G(\mathbf{y})$ and $\mathbf{y} = F(\mathbf{x})$, thus

$$(1.3.3.9) \quad F^*(G^*\eta) = (G \circ F)^*\eta.$$

Moreover, we have

$$(1.3.3.10) \quad d(F^*\omega) = F^*(d\omega)$$

for any smooth form ω on V . Consequently, the smooth map $F : U \rightarrow V$ induces a morphism on cohomology groups:

$$(1.3.3.11) \quad F^* : H_{\mathbf{DR}}^p(V) \longrightarrow H_{\mathbf{DR}}^p(U), \quad [\omega] \longmapsto [F^*\omega].$$

Let ω be a smooth form on $[0, 1] \times U$. For $(t, \mathbf{x}) \in [0, 1] \times U$ we write

$$(1.3.3.12) \quad \omega(t, \mathbf{x}) = \sum_{|I|=p} \omega_I(t, \mathbf{x}) dx^I + \sum_{|J|=p-1} \tilde{\omega}_J(t, \mathbf{x}) dt \wedge dx^J.$$

Define

$$(1.3.3.13) \quad \mathbf{K} : \Omega^p([0, 1] \times U) \longrightarrow \Omega^{p-1}(U), \quad \omega \longmapsto \sum_{|J|=p-1} \left(\int_0^1 \tilde{\omega}_J(t, \mathbf{x}) dt \right) dx^J.$$

We then have the following diagram

$$\begin{array}{ccc} \Omega^p([0, 1] \times U) & \xrightarrow{d} & \Omega^{p+1}([0, 1] \times U) \\ \downarrow \mathbf{K} & & \downarrow \mathbf{K} \\ \Omega^{p-1}(U) & \xrightarrow{d} & \Omega^p(U) \end{array}$$

Observe that

$$(1.3.3.14) \quad (\mathbf{K}d + d\mathbf{K})\omega = \sum_{|I|=p} \left(\int_0^1 \frac{\partial \omega_I}{\partial t}(t, \mathbf{x}) dt \right) dx^I = \sum_{|I|=p} [\omega_I(1, \mathbf{x}) - \omega_I(0, \mathbf{x})] dx^I$$

where ω is given in (1.3.3.12). Thus

$$(1.3.3.15) \quad (\mathbf{K}d + d\mathbf{K})\omega = \iota_1^*\omega - \iota_0^*\omega$$

where $\iota_t : U \rightarrow [0, 1] \times U$ is the injection $\mathbf{x} \mapsto (t, \mathbf{x})$.

Theorem 1.3.13. (Poincaré's lemma) *Let $U \subseteq \mathbf{R}^n$ be a star-shaped open set with respect to $\mathbf{0}$ (that is, $\mathbf{0} \in U$ and any line segment $\overline{\mathbf{0}\mathbf{x}}$ is entirely contained in U). If $\omega \in Z^p(U)$ and $p \geq 1$, then $\omega = d\eta$ for some $\eta \in \Omega^{p-1}(U)$. Consequently, $H_{\mathbf{DR}}^p(U) = \{0\}$ for any star-shaped open set (with respect to some point in U) and $p \geq 1$, and $H_{\mathbf{DR}}^0(U) = \mathbf{R}$ for any star-shaped open set.*

PROOF. Let $H : [0, 1] \times U \rightarrow U$ be defined by $H(t, x) := tx$ and set

$$F(x) := x, \quad G(x) := \mathbf{0}, \quad x \in \Omega.$$

Then

$$F = H \circ \iota_1, \quad G = H \circ \iota_0,$$

so that

$$d(\mathbf{K}(H^*\omega)) = \iota_{*1}H^*\omega - \iota_{*0}H^*\omega = F^*\omega - G^*\omega = \begin{cases} \omega - \omega(\mathbf{0}), & p = 0, \\ \omega, & p \geq 1. \end{cases}$$

When $p \geq 1$, we get $\eta = \mathbf{K}(H^*\omega)$. \square

A direct sum of vector spaces

$$C^\bullet := \bigoplus_{p \in \mathbf{Z}} C^p$$

is called a **differential complex** if there are homomorphisms

$$\dots \longrightarrow C^{p-1} \xrightarrow{d^{p-1}} C^p \xrightarrow{d^p} C^{p+1} \xrightarrow{d^{p+1}} \dots$$

such that $d^{p+1} \circ d^p = 0$ for any $p \in \mathbf{Z}$. d is the **differential operator** of the complex C^\bullet . The **cohomology** of C^\bullet is the direct sum of vector spaces

$$(1.3.3.16) \quad H^\bullet(C^\bullet) := \bigoplus_{p \in \mathbf{Z}} H^p(C^\bullet)$$

where

$$H^p(C^\bullet) := \frac{\text{Ker}(d^p) \cap C^p}{\text{Im}(d^{p-1}) \cap C^p}.$$

A map $f^\bullet : A^\bullet \rightarrow B^\bullet$ between two differential complexes is a **chain map** if it commutes with the differential operators of A^\bullet and B^\bullet : $f^\bullet \circ d_{A^\bullet}^p = d_{B^\bullet}^p \circ f^\bullet$. That is,

$$\begin{array}{ccccccc} \dots & \longrightarrow & A^{p-1} & \xrightarrow{d_{A^\bullet}^{p-1}} & A^p & \xrightarrow{d_{A^\bullet}^p} & A^{p+1} & \xrightarrow{d_{A^\bullet}^{p+1}} & \dots \\ & & \downarrow f^{p-1} & & \downarrow f^p & & \downarrow f^{p+1} & & \\ \dots & \longrightarrow & B^{p-1} & \xrightarrow{d_{B^\bullet}^{p-1}} & B^p & \xrightarrow{d_{B^\bullet}^p} & B^{p+1} & \xrightarrow{d_{B^\bullet}^{p+1}} & \dots \end{array}$$

for any $p \in \mathbf{Z}$,

$$f^p \circ d_{A^\bullet}^{p-1} = d_{B^\bullet}^{p-1} \circ f^{p-1}.$$

A sequence of vector spaces

$$\dots \longrightarrow V^{i-1} \xrightarrow{f^{i-1}} V^i \xrightarrow{f^i} V^{i+1} \longrightarrow \dots$$

is said to be **exact** if for all $i \in \mathbf{Z}$, $\text{Ker}(f^i) = \text{Im}(f^{i-1})$. An exact sequence of vector spaces $0 \rightarrow U \rightarrow V \rightarrow W \rightarrow 0$ is called a **short exact sequence**.

Given a short exact sequence of differential complexes

$$0 \longrightarrow A^\bullet \xrightarrow{f^\bullet} B^\bullet \xrightarrow{g^\bullet} C^\bullet \longrightarrow 0$$

in which the maps f^\bullet and g^\bullet are chain maps, there is a long exact sequence of cohomology groups:

$$\dots \longrightarrow H^p(A^\bullet) \xrightarrow{H^p(f^\bullet)} H^p(B^\bullet) \xrightarrow{H^p(g^\bullet)} H^p(C^\bullet) \xrightarrow{\delta^p} H^{p+1}(A^\bullet) \longrightarrow$$

Indeed, the map $H^p(f^\bullet) : H^p(A^\bullet) \rightarrow H^p(B^\bullet)$ is defined by

$$H^p(f^\bullet)([a]) := [f^p(a)].$$

If $a \in \text{Ker}(d_{A^\bullet}^p)$, then

$$0 = f^{p+1} \circ d_{A^\bullet}^p(a) = d_{B^\bullet}^p \circ f^p(a)$$

which implies $f^p(a) \in \text{Ker}(d_{B^\bullet}^p)$. If $[a_1] = [a_2] \in H^p(A^\bullet)$, then

$$a_2 = a_1 + d_{A^\bullet}^{p-1}(a)$$

for some $a \in A^{p-1}$. Then

$$f^p(a_2) = f^p(a_1) + f^p \circ d_{A^\bullet}^{p-1}(a) = f^p(a_1) + d_{B^\bullet}^{p-1} \circ f^{p-1}(a).$$

Thus $[f^p(a_1)] = [f^p(a_2)]$.

Similarly, we can define $H^p(g^\bullet) : H^p(B^\bullet) \rightarrow H^p(C^\bullet)$.

Next we define $\delta^p : H^p(C^\bullet) \rightarrow H^{p+1}(A^\bullet)$ as follows.

$$\begin{array}{ccccccc} & & \vdots & & \vdots & & \vdots \\ & & \uparrow d_{A^\bullet}^{p+1} & & \uparrow d_{B^\bullet}^{p+1} & & \uparrow d_{C^\bullet}^{p+1} \\ 0 & \longrightarrow & A^{p+1} \ni a & \xrightarrow{f^{p+1}} & B^{p+1} \ni d_{B^\bullet}^p(b) & \xrightarrow{g^{p+1}} & C^{p+1} \longrightarrow 0 \\ & & \uparrow d_{A^\bullet}^p & & \uparrow d_{B^\bullet}^p & & \uparrow d_{C^\bullet}^p \\ 0 & \longrightarrow & A^p & \xrightarrow{f^p} & B^p \ni b & \xrightarrow{g^p} & C^p \ni c \longrightarrow 0 \\ & & \uparrow d_{A^\bullet}^{p-1} & & \uparrow d_{B^\bullet}^{p-1} & & \uparrow d_{C^\bullet}^{p-1} \\ & & \vdots & & \vdots & & \vdots \end{array}$$

For any $c \in \text{Ker}(d_{C^\bullet}^p)$, since g^p is surjective, it follows that $g^p(b) = c$ for some $b \in B^p$. Because

$$0 = d_{C^\bullet}^p \circ g^p(b) = g^{p+1} \circ d_{B^\bullet}^p(b),$$

we get $d_{B^\bullet}^p(b) \in \text{Ker}(g^{p+1}) = \text{Im}(f^{p+1})$. Thus $f^{p+1}(a) = d_{B^\bullet}^p(b)$ for some $a \in A^{p+1}$. From

$$0 = d_{B^\bullet}^{p+1} \circ d_{B^\bullet}^p(b) = d_{B^\bullet}^{p+1} \circ f^{p+1}(a) = f^{p+2} \circ d_{A^\bullet}^{p+1}(a)$$

and the injectivity of f^{p+2} , we see that $d_{A^\bullet}^{p+1}(a) = 0$ and then a is closed. Define

$$\delta^p([c]) := [a].$$

It is clear that the map δ^p is independent of the choice of a, b, c .

Proposition 1.3.14. *The map δ^p is independent of the choice of a, b, c .*

PROOF. Assume $[c] = [c']$. Then $c' = c + d_{C^\bullet}^{p-1}(\tilde{c})$ for some $\tilde{c} \in C^{p-1}$. For c' , we can also find $b' \in B^p$ and $a' \in A^{p+1}$ such that

$$g^p(b') = c', \quad d_{A^\bullet}^{p+1}(a') = 0, \quad f^{p+1}(a') = d_{B^\bullet}^p(b').$$

Since $g^p(b' - b) = c' - c = d_{C^\bullet}^{p-1}(\tilde{c})$, it follows that

$$0 = d_{C^\bullet}^p \circ g^p(b' - b) = g^{p+1} \circ d_{B^\bullet}^p(b' - b)$$

and then $d_{B^\bullet}^p(b' - b) \in \mathbf{Ker}(g^{p+1}) = \mathbf{Im}(f^{p+1})$. Hence $f^{p+1}(\tilde{a}) = d_{B^\bullet}^p(b' - b)$ for some $\tilde{a} \in A^{p+1}$. Moreover

$$0 = d_{B^\bullet}^{p+1} \circ f^{p+1}(\tilde{a}) = f^{p+2} \circ d_{A^\bullet}^{p+1}(\tilde{a}) \implies d_{A^\bullet}^{p+1}(\tilde{a}) = 0.$$

From $f^{p+1}(\tilde{a}) = f^{p+1}(a' - a)$ and the injectivity of f^{p+1} , we obtain $a' = a + \tilde{a}$ for some closed vector \tilde{a} . Therefore $[a'] = [a]$. \square

1.4. Topological groups and matrix Lie groups

A **topological group** is a group G endowed with a topology such that

$$\mu : G \times G \longrightarrow G, \quad (g_1, g_2) \longmapsto g_1 g_2$$

and

$$\iota : G \longrightarrow G, \quad g \longmapsto g^{-1}$$

are both continuous. If the topology is discrete, G is called a **discrete group**.

For example, \mathbf{R} , together with additive group structure and Euclidean topology, is a topological group; $\mathbf{R}^* := \mathbf{R} \setminus \{0\}$, together with multiplication and the relative topology, is a topological group; the **general linear group** $\mathbf{GL}(n, \mathbf{R}) := \{n \times n \text{ real nonsingular matrices}\}$, together with matrix multiplication and the relative topology inherited from \mathbf{R}^{n^2} , is a topological group.

1.4.1. Subgroups and left/right translations. Let G be a topological group. A **subgroup** H of G is a subset which is also a subgroup in the algebraic sense. Under the relative topology, H is also a topological group.

If A, B are subsets of G we let

$$AB := \{ab : a \in A, b \in B\}, \quad A^{-1} := \{a^{-1} : a \in A\}.$$

We say a subset A of G is **symmetric** if $A = A^{-1}$.

Proposition 1.4.1. (1) *Any finite product of topological groups is a topological group.*

(2) *Let G be a topological group with unity element e .*

- (a) *The symmetric neighborhoods of e form a neighborhood basis at e .*
- (b) *If $g \in G$ and U is any neighborhood of g , then there is a symmetric neighborhood V of e such that $VgV^{-1} \subseteq U$.*
- (c) *If U is any neighborhood of e and n is any positive integer, then there exists a symmetric neighborhood V of e such that $V^n \subseteq U$.*

(3) *If H is any subgroup of G , then \overline{H} is also a subgroup of G . If H is a normal subgroup then so is \overline{H} .*

PROOF. (1): as an exercise.

(2) If U is a neighborhood of e then so is U^{-1} , and hence $U \cap U^{-1}$ is a symmetric neighborhood of e . This implies (a).

(3) **to be added!**

\square

By **Proposition 1.4.1**, $\mathbf{R}^n, \mathbf{S}^1, \mathbf{T}^n$ are topological groups. Since the **orthogonal group**

$$\mathbf{O}(n) = \{A \in \mathbf{GL}(n, \mathbf{R}) : AA^T = I\}$$

is a subgroup of $\mathbf{GL}(m, \mathbf{R})$, $\mathbf{O}(m)$ is a topological group.

Let G be a topological group and $g \in G$.

(1) The **left translation** is defined by

$$(1.4.1.1) \quad L_g : G \longrightarrow G, \quad g' \longmapsto gg'.$$

Then $L_g^{-1} = L_{g^{-1}}$ and $L_{g_1} \circ L_{g_2} = L_{g_1g_2}$.

(2) The **right translation** is defined by

$$(1.4.1.2) \quad R_g : G \longrightarrow G, \quad g' \longmapsto g'g^{-1}.$$

Then $R_g^{-1} = R_{g^{-1}}$ and $R_{g_1} \circ R_{g_2} = R_{g_1g_2}$.

(3) The **adjoint homomorphism** is defined by

$$(1.4.1.3) \quad \mathbf{ad}_g : G \longrightarrow G, \quad g' \longmapsto gg'g^{-1}.$$

Then $\mathbf{ad}_g = L_g \circ R_g = R_g \circ L_g$.

1.4.2. Group actions. Let X be a topological space and G a topological group.

(1) The **left action of G on X** is a continuous map

$$(1.4.2.1) \quad G \times X \longrightarrow X, \quad (g, x) \longmapsto g \cdot x$$

with the properties that

$$g_1 \cdot (g_2 \cdot x) = (g_1g_2) \cdot x, \quad 1 \cdot x = x$$

for any $g_1, g_2 \in G$ and $x \in X$.

(2) The **right action of G on X** is a continuous map

$$(1.4.2.2) \quad X \times G \longrightarrow X, \quad (x, g) \longmapsto x \cdot g$$

with the properties that

$$(x \cdot g_1) \cdot g_2 = x \cdot (g_1g_2), \quad x \cdot 1 = x$$

for any $g_1, g_2 \in G$ and $x \in X$.

Observe that L_g induces a left action of G on itself by setting

$$g \cdot x := L_g x.$$

Similarly, R_g induces a right action of G on itself by setting

$$x \cdot g := R_{g^{-1}} x.$$

We now always assume that $G \curvearrowright X$ is a left action. For any $x \in X$, we define the **orbit of x** by

$$(1.4.2.3) \quad G \cdot x := \{g \cdot x : g \in G\} \subseteq X.$$

If $G \cdot x \cap G \cdot y \neq \emptyset$, then $g \cdot x = y$ for some $g \in G$ and $G \cdot x = G \cdot y$. We can define an equivalence relation \sim on X by

$$(1.4.2.4) \quad x \sim y \iff g \cdot x = y \text{ for some } g \in G.$$

Then the equivalence class $[x]$ of x is precisely the orbit $G \cdot x$ of x . Define the **orbit space** by

$$(1.4.2.5) \quad G \backslash X := X / \sim = \{[x] : x \in X\} = \bigcup_{x \in X} G \cdot x.$$

Example 1.4.2. (1) The left action $\mathbf{GL}(n, \mathbf{R}) \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ given by $(A, x) \mapsto Ax$ has orbits that are $\{0\}$ and $\mathbf{R}^n \setminus \{0\}$.

(2) The left action $\mathbf{O}(n) \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ given by $(A, x) \mapsto Ax$ has orbits that are $\{0\}$ and spheres at 0 .

(3) The left action $\mathbf{R}^* \times (\mathbf{R}^n \setminus \{0\}) \rightarrow \mathbf{R}^n \setminus \{0\}$ given by $(r, x) \mapsto rx$ has orbits that are lines through 0 .

Theorem 1.4.3. Let G be a topological group acted from left on a topological space X . Then

- (i) $\pi : X \rightarrow G \backslash X$ is open, and
- (ii) $G \backslash X$ is Hausdorff if and only if $\{(x, y) \in X \times X : y = g \cdot x \text{ for some } g \in G\}$ is closed in $X \times X$.

PROOF. (i) Let $U \subset X$ be open and consider $\pi^{-1}(\pi(U))$. To prove π is open, we shall verify $\pi^{-1}(\pi(U))$ is open in X . For any $x \in \pi^{-1}(\pi(U))$, set $U_x := \{g \cdot x : g \in G\}$. The continuity implies that U_x is a neighborhood of x and $U_x \subset \pi^{-1}(\pi(U))$.

(ii) $\Delta_{G \backslash X} = \{([x], [x]) : [x] \in G \backslash X\}$ implies that

$$(\pi \times \pi)^{-1} \Delta_{G \backslash X} = \{(x, y) \in X \times X : g \cdot x = y \text{ for some } g \in G\}.$$

Since $\pi \times \pi$ is open, surjective, and continuous, it follows that $\Delta_{G \backslash X}$ is closed if and only if $(\pi \times \pi)^{-1}(\Delta_{G \backslash X})$ is closed. \square

In general, we can show that if $f : X \rightarrow Y$ is open, surjective, and continuous, then a subset $S \subseteq X$ is closed if and only if $f^{-1}(S) \subseteq X$ is closed. Indeed, if S is closed, then

$$f^{-1}(Y \setminus S) = f^{-1}(Y) \setminus f^{-1}(S) = X \setminus f^{-1}(S)$$

is open in X . If $f^{-1}(S)$ is closed, then $f(X \setminus f^{-1}(S)) = Y \setminus S$ is open in Y .

1.4.3. Matrix Lie groups. Let $\mathbf{K} = \mathbf{R}$ or \mathbf{C} . The **general linear map over \mathbf{K}** is the group

$$(1.4.3.1) \quad \mathbf{GL}(n, \mathbf{K}) := \{n \times n \text{ invertible matrices with } \mathbf{K}\text{-entries}\} \subseteq \mathbf{K}^{n^2}.$$

Then $\mathbf{GL}(n, \mathbf{K})$ is a topological group. Define

$$(1.4.3.2) \quad \mathbf{Max}(n, \mathbf{K}) := \{n \times n \text{ matrices with } \mathbf{K}\text{-entries}\} \cong \mathbf{K}^{n^2}.$$

We say a sequence $\{A_m\}_{m \geq 1}$ of complex matrices in $\mathbf{Max}(n, \mathbf{C})$ **converges** to a matrix $A \in \mathbf{Max}(n, \mathbf{C})$ if $(A_m)_{ij} \rightarrow A_{ij}$ as $m \rightarrow \infty$ for all $1 \leq i, j \leq n$.

Definition 1.4.4. A **matrix Lie group** is any subgroup G of $\mathbf{GL}(n, \mathbf{C})$ with the following property: if $\{A_m\}_{m \geq 1}$ is any sequence of matrices in G and A_m converges to some matrix A , then either $A \in G$ or A is not invertible.

Equivalently $G \subseteq \mathbf{GL}(n, \mathbf{C})$ is a matrix Lie group if and only if it is a closed subgroup of $\mathbf{GL}(n, \mathbf{C})$.

Example 1.4.5. (Counterexamples of matrix Lie groups) (1) *The set of all $n \times n$ invertible matrices all of whose entries are rational.*

(2) *Let a be irrational and consider*

$$G := \left\{ \begin{bmatrix} e^{\sqrt{-1}t} & 0 \\ 0 & e^{\sqrt{-1}ta} \end{bmatrix} \mid t \in \mathbf{R} \right\}.$$

Clearly $-I \notin G$. However we can find a sequence of matrices in G which converges to $-I$, so G is not a matrix Lie group. Actually

$$\bar{G} = \left\{ \begin{bmatrix} e^{\sqrt{-1}t} & 0 \\ 0 & e^{\sqrt{-1}s} \end{bmatrix} \mid s, t \in \mathbf{R} \right\} \cong \mathbf{S}^1 \times \mathbf{S}^1 = \mathbf{T}^2.$$

We introduce some examples of matrix Lie groups.

(1) **The general linear groups** $\mathbf{GL}(n, \mathbf{R})$ and $\mathbf{GL}(n, \mathbf{C})$. Moreover $\mathbf{GL}(n, \mathbf{R})$ is a subgroup of $\mathbf{GL}(n, \mathbf{C})$.

(2) **The special linear groups** $\mathbf{SL}(n, \mathbf{K})$:

$$(1.4.3.3) \quad \mathbf{SL}(n, \mathbf{C}) := \{A \in \mathbf{GL}(n, \mathbf{K}) \mid \det A = 1\}.$$

(3) **The orthogonal and special orthogonal groups.** We say $A \in \mathbf{Max}(n, \mathbf{R})$ is **orthogonal** if

$$\sum_{1 \leq k \leq n} A_{ki} A_{kj} = \delta_{ij}, \quad 1 \leq i, j \leq n.$$

The usual inner product on \mathbf{R}^n is

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{1 \leq k \leq n} x^k y^k, \quad \mathbf{x} = (x^1, \dots, x^n), \quad \mathbf{y} = (y^1, \dots, y^n) \in \mathbf{R}^n.$$

Then A is orthogonal is equivalent to $\langle \mathbf{x}, \mathbf{y} \rangle = \langle A\mathbf{x}, A\mathbf{y} \rangle$ for any $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$. Hence A is orthogonal if and only if $A^\top A = I$, where A^\top denotes the transpose of A , i.e., $(A^\top)_{ij} = A_{ji}$.

If A is orthogonal, then $1 = \det(A^\top A) = (\det A)^2$ and $\det A = \pm 1$.

Let

$$(1.4.3.4) \quad \mathbf{O}(n) := \{A \in \mathbf{GL}(n, \mathbf{R}) \mid A \text{ orthogonal}\}$$

and

$$(1.4.3.5) \quad \mathbf{SO}(n) := \{A \in \mathbf{O}(n) \mid \det A = 1\}.$$

For $n = 2$, $\mathbf{SO}(2)$ is the group of rotations on the plane, i.e.,

$$(1.4.3.6) \quad \mathbf{SO}(2) = \left\{ \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \mid \theta \in \mathbf{R} \right\}.$$

(4) **The unitary and special unitary groups.** We say $A \in \mathbf{Max}(n, \mathbf{C})$ is **unitary** if

$$\sum_{1 \leq k \leq n} \overline{A_{ki}} A_{kj} = \delta_{ij}, \quad 1 \leq i, j \leq n.$$

The usual inner product on \mathbf{C}^n is

$$\langle \mathbf{z}, \mathbf{w} \rangle := \sum_{1 \leq k \leq n} \overline{z^k} w^k, \quad \mathbf{z} = (z^1, \dots, z^n), \quad \mathbf{w} = (w^1, \dots, w^n) \in \mathbf{C}^n.$$

Then A is unitary is equivalent to $\langle z, w \rangle = \langle Az, Aw \rangle$ for any $z, w \in \mathbf{C}^n$. Hence A is unitary if and only if $A^*A = I$, where A^* denotes the adjoint of A , i.e., $(A^*)_{ij} = \overline{A_{ji}}$.

If A is unitary, then $1 = \det(A^*A) = |\det A|^2$ and $\det A \in \mathbf{S}^1$. Let

$$(1.4.3.7) \quad \mathbf{U}(n) := \{A \in \mathbf{L}(n, \mathbf{C}) \mid A \text{ unitary}\}$$

and

$$(1.4.3.8) \quad \mathbf{SU}(n) := \{A \in \mathbf{U}(n) \mid \det A = 1\}.$$

- (5) **The complex orthogonal groups.** Instead the inner product on \mathbf{C}^n , we consider the bilinear form (\cdot, \cdot) on \mathbf{C}^n ,

$$(z, w) := \sum_{1 \leq k \leq n} z^k w^k, \quad z = (z^1, \dots, z^n), \quad w = (w^1, \dots, w^n) \in \mathbf{C}^n.$$

We say $A \in \mathbf{Max}(n, \mathbf{C})$ is **(complex) orthogonal** if $(z, w) = (Az, Aw)$ for all $z, w \in \mathbf{C}^n$. If A is orthogonal, then $1 = \det A \cdot \det A$ and $\det A = \pm 1$. Let

$$(1.4.3.9) \quad \mathbf{O}(n, \mathbf{C}) := \{A \in \mathbf{Max}(n, \mathbf{C}) \mid A \text{ orthogonal}\}$$

and

$$(1.4.3.10) \quad \mathbf{SO}(n, \mathbf{C}) := \{A \in \mathbf{O}(n, \mathbf{C}) \mid \det A = 1\}.$$

- (6) **The generalized orthogonal and Lorentz groups.** For any positive integers n and k , define a symmetric bilinear form $[\cdot, \cdot]_{n,k}$ on \mathbf{R}^{n+k} :

$$(1.4.3.11) \quad [x, y]_{n,k} := \sum_{1 \leq i \leq n} x^i y^i - \sum_{n+1 \leq j \leq n+k} x^j y^j.$$

Let

$$(1.4.3.12) \quad \mathbf{O}(n, k) := \{A \in \mathbf{Max}(n+k, \mathbf{R}) \mid [Ax, Ay]_{n,k} = [x, y]_{n,k}\}.$$

For $A \in \mathbf{Max}(n+k, \mathbf{R})$, let

$$A^{(i)} := \begin{bmatrix} A_{1,i} \\ \vdots \\ A_{n+k,i} \end{bmatrix}, \quad 1 \leq i \leq n+k.$$

Then

$$A \in \mathbf{O}(n, k) \iff \begin{cases} [A^{(i)}, A^{(j)}]_{n,k} = 0, & i \neq j, \\ [A^{(i)}, A^{(i)}]_{n,k} = 1, & 1 \leq i \leq n \\ [A^{(i)}, A^{(i)}]_{n,k} = -1, & n+1 \leq i \leq n+k. \end{cases}$$

Let

$$g := \begin{bmatrix} I_n & 0 \\ 0 & -I_k \end{bmatrix}$$

Then

$$(1.4.3.13) \quad \mathbf{AO}(n, k) \iff A^T g A = g.$$

Hence $\det A = \pm 1$.

When $(n, k) = (3, 1)$, we obtain the Lorentz group $\mathbf{O}(3, 1)$.

- (7) **The symplectic groups.** Define the following skew-symmetric bilinear form B on \mathbf{R}^{2n} :

$$(1.4.3.14) \quad B[x, y] := \sum_{1 \leq k \leq n} (x^k y^{n+k} - x^{n+k} y^k).$$

Let us define the **real symplectic group** by

$$(1.4.3.15) \quad \mathbf{Sp}(n, \mathbf{R}) := \{A \in \mathbf{Max}(2n, \mathbf{R}) \mid B[Ax, Ay] = B[x, y], \forall x, y \in \mathbf{R}^{2n}\}.$$

Clearly $\mathbf{Sp}(1, \mathbf{R}) = \mathbf{SL}(2, \mathbf{R})$. If define

$$J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

then $B[x, y] = \langle x, Jy \rangle$ and

$$(1.4.3.16) \quad A \in \mathbf{Sp}(n, \mathbf{R}) \implies A^T J A = J.$$

Hence $\det A = \pm 1$ for all $A \in \mathbf{Sp}(n, \mathbf{R})$. Actually, we can prove that $\det A = 1$ for all $A \in \mathbf{Sp}(n, \mathbf{R})$.

Similarly we can define the **complex symplectic group**

$$(1.4.3.17) \quad \mathbf{Sp}(n, \mathbf{C}) := \{A \in \mathbf{Max}(2n, \mathbf{C}) \mid B[Ax, Ay] = B[x, y], \forall x, y \in \mathbf{C}^{2n}\}.$$

Then

$$(1.4.3.18) \quad A \in \mathbf{Sp}(n, \mathbf{C}) \iff A^T J A = J.$$

Hence $\det A = \pm 1$ for all $A \in \mathbf{Sp}(n, \mathbf{C})$. Actually, we can prove that $\det A = 1$ for all $A \in \mathbf{Sp}(n, \mathbf{C})$. Clearly $\mathbf{Sp}(1, \mathbf{C}) = \mathbf{SL}(2, \mathbf{C})$.

The **compact symplectic group**⁴ is defined to be

$$(1.4.3.19) \quad \mathbf{Sp}(n) := \mathbf{Sp}(n, \mathbf{C}) \cap \mathbf{U}(2n).$$

Clearly $\mathbf{Sp}(1) = \mathbf{SU}(2)$.

The special and general linear groups, the orthogonal and unitary groups, and the symplectic groups make up the **classical groups**.

- (8) The **Heisenberg group** H is

$$(1.4.3.20) \quad H := \left\{ A = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \in \mathbf{Max}(3, \mathbf{R}) \mid a, b, c \in \mathbf{R} \right\} \subseteq \mathbf{SL}(3, \mathbf{R}).$$

- (9) $\mathbf{R}^* = \{a \in \mathbf{R} \mid a \neq 0\} \cong \mathbf{GL}(1, \mathbf{R})$, $\mathbf{C}^* = \{a \in \mathbf{C} \mid a \neq 0\} \cong \mathbf{GL}(1, \mathbf{C})$, and $\mathbf{S}^1 = \{z \in \mathbf{C}^* \mid |z| = 1\} \cong \mathbf{U}(1)$. The additive group \mathbf{R} is isomorphic to $\mathbf{GL}^+(1, \mathbf{R})$ the group of 1×1 matrices with positive determinant, via the map $x \mapsto [e^x]$. Moreover, the additive group \mathbf{R}^n is isomorphic to the group of diagonal real matrices with positive diagonal entries, i.e.,

$$\mathbf{R}^n \cong \mathbf{GL}^+(1, \mathbf{R}) \times \cdots \times \mathbf{GL}^+(1, \mathbf{R}) \quad (n \text{ times}).$$

- (10) **The Euclidean and Poincaré groups.** Let

$$(1.4.3.21) \quad \mathbf{E}(n) := \{\text{distance-preserving bijections on } \mathbf{R}^n\}$$

that is, $f \in \mathbf{E}(n)$ if and only if $f : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is bijective and $|x - y| = |f(x) - f(y)|$ for all $x, y \in \mathbf{R}^n$. Note that f need not to be linear.

⁴Some books used $\mathbf{Sp}(n)$ for our $\mathbf{Sp}(n, \mathbf{C})$ and $\mathbf{USp}(n)$ for our $\mathbf{Sp}(n)$.

Then $\mathbf{O}(n)$ is a subgroup of $\mathbf{E}(n)$ and is the group of all linear distance-preserving maps on \mathbf{R}^n .

For $x \in \mathbf{R}^n$ define the **translation by x** by

$$T_x : \mathbf{R}^n \longrightarrow \mathbf{R}^n, \quad y \longmapsto x + y.$$

A basic fact about $\mathbf{R}(n)$ is any $T \in \mathbf{E}(n)$ can be written uniquely as an orthogonal linear transformation followed by a translation, i.e., $T = T_x \circ R$ for some $x \in \mathbf{R}^n$ and $R \in \mathbf{O}(n)$. Write

$$\{x, R\} := T_x \circ R.$$

Then for $y \in \mathbf{R}^n$, one has

$$\{x, R\}y = Ry + x$$

and

$$\{x_1, R_1\}\{x_2, R_2\}y = R_1(R_2y + x_2) + x_1 = R_1R_2y + (x_1 + R_1x_2).$$

The product on $\mathbf{E}(n)$ is

$$(1.4.3.22) \quad \{x_1, R_1\}\{x_2, R_2\} = \{x_1 + R_1x_2, R_1 \circ R_2\}$$

and the inverse of $\{x, R\}$ is

$$\{x, R\}^{-1} = \{-R^{-1}x, R^{-1}\}.$$

Because translations are not linear, we conclude that $\mathbf{E}(n)$ is not a subgroup of $\mathbf{GL}(n, \mathbf{R})$. However, $\mathbf{E}(n)$ is isomorphic to a subgroup of $\mathbf{GL}(n+1, \mathbf{R})$ via the following map

$$\{x, R\} \longmapsto \begin{bmatrix} R & x^T \\ \mathbf{0} & 1 \end{bmatrix} \in \mathbf{GL}(n+1, \mathbf{R}).$$

The Poincaré group is defined to be

$$(1.4.3.23) \quad \mathbf{P}(n, 1) := \{T = T_x \circ A \mid x \in \mathbf{R}^{n+1}, A \in \mathbf{O}(n, 1)\}.$$

This is the group of affine transformations of \mathbf{R}^{n+1} which preserves the Lorentz distance

$$d_L(x, y) := \sum_{1 \leq i \leq n} (x^i - y^i)^2 - (x^{n+1} - y^{n+1})^2, \quad x, y \in \mathbf{R}^{n+1}.$$

Here an affine transformation is one of the form $x \mapsto Ax + b$, where A is a linear transformation and b is a constant. Similarly, we can prove that $\mathbf{P}(n, 1)$ is isomorphic to the group of $(n+2) \times (n+2)$ matrices of the form

$$\begin{bmatrix} A & x^T \\ \mathbf{0} & 1 \end{bmatrix} \quad A \in \mathbf{O}(n, 1), \quad x \in \mathbf{R}^{n+1}$$

Exercise 1.4.6. (1) Show that $\mathbf{O}(n)$ is a group, a subgroup of $\mathbf{GL}(n, \mathbf{C})$, and is a matrix Lie group.

(2) Show that $\mathbf{SO}(n)$ is a group, a subgroup of $\mathbf{O}(n)$, and is a matrix Lie group.

(3) Deduce (1.4.3.6) from (1.4.3.5).

(4) Show that $\mathbf{U}(n)$ is a group, a subgroup of $\mathbf{GL}(n, \mathbf{C})$, and is a matrix Lie group.

(5) Show that $\mathbf{SU}(n)$ is a group, a subgroup of $\mathbf{U}(n)$, and is a matrix Lie group.

- (6) Show that $\mathbf{O}(n, k)$ is a group, a subgroup of $\mathbf{GL}(n + k, \mathbf{R})$, and a matrix Lie group.
- (7) Determine elements in $\mathbf{O}(1, 1)$.
- (8) Verify (1.4.3.13).
- (9) Show that $\mathbf{Sp}(n, \mathbf{R})$ is a group, a subgroup of $\mathbf{GL}(2n, \mathbf{R})$, and a matrix Lie group.
- (10) Show that the Heisenberg H is a group, a subgroup of $\mathbf{GL}(3, \mathbf{R})$, and a matrix Lie group.

Actually, every $A \in \mathbf{SU}(2)$ can be expressed in the form

$$A = \begin{bmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{bmatrix}, \quad \alpha, \beta \in \mathbf{C}, \quad |\alpha|^2 + |\beta|^2 = 1.$$

Hence we can view $\mathbf{SU}(2)$ as the three-dimensional sphere \mathbf{S}^3 inside $\mathbf{C}^2 = \mathbf{R}^4$.

Definition 1.4.7. A matrix Lie group G is said to be **compact** if the following conditions holds:

- (1) if A_m is any sequence of matrices in G and A_m converges to a matrix A , then $A \in G$, and
- (2) there exists a constant C such that for all $A \in G$, $\|A\|_\infty = \max_{1 \leq i, j \leq n} |A_{ij}| \leq C$.

If we view $\mathbf{Max}(n, \mathbf{C})$ as \mathbf{C}^{n^2} , then G is compact in the sense of **Definition 1.4.7** if and only if G is a closed and bounded subset of \mathbf{C}^{n^2} .

Example 1.4.8. (1) *Compact matrix Lie groups:* $\mathbf{O}(n)$, $\mathbf{SO}(n)$, $\mathbf{U}(n)$, $\mathbf{SU}(n)$, $\mathbf{Sp}(n)$.
 (2) *Noncompact matrix Lie groups:* $\mathbf{GL}(n, \mathbf{R})$, $\mathbf{GL}(n, \mathbf{C})$, $\mathbf{SL}(n, \mathbf{R})$, $\mathbf{SL}(n, \mathbf{C})$, $\mathbf{O}(n, \mathbf{C})$, $\mathbf{SO}(n, \mathbf{C})$, $\mathbf{O}(n, k)$, $\mathbf{SO}(n, k)$, the Heisenberg group H , $\mathbf{Sp}(n, \mathbf{R})$, $\mathbf{Sp}(n, \mathbf{C})$, $\mathbf{E}(n)$, $\mathbf{P}(n, 1)$, \mathbf{R} , \mathbf{R}^* , \mathbf{C} , \mathbf{C}^* .

Exercise 1.4.9. Verify **Example 1.4.8**.

Definition 1.4.10. A matrix Lie group G is **path-connected** if given any two matrices A and B in G , there exists a continuous path $A(t)$ in G , $0 \leq t \leq 1$, such that $A(0) = A$ and $A(1) = B$.

Since any matrix Lie group G can be viewed as a subset of \mathbf{C}^{n^2} that is simply-connected, it follows that any matrix Lie group is path-connected if and only if it is path-connected.

A matrix Lie group which is not path-connected can be decomposed uniquely as a union of **path-connected components**, such that two elements in the same path-connected component can be joined by a continuous path, but two elements in the different path-connected components cannot.

Proposition 1.4.11. (1) *If G is a matrix Lie group, then the path-connected component of G containing the identity is a subgroup of G .*

- (2) *For any $n \geq 1$, the group $\mathbf{GL}(n, \mathbf{C})$ is path-connected.*
- (3) *For any $n \geq 1$, the group $\mathbf{SL}(n, \mathbf{C})$ is path-connected.*
- (4) *For any $n \geq 1$, the groups $\mathbf{U}(n)$ and $\mathbf{SU}(n)$ are path-connected.*

PROOF. (1) Let H be the path-connected component of G containing the identity e . For $A, B \in H \subseteq G$, there exists a continuous paths $A(t), B(t)$, such that

$$A(0) = B(0) = e, \quad A(1) = A, \quad B(1) = B.$$

Then $(A(t)B(t))$ is a continuous path connecting e and AB . Hence the product AB is still in H . Similarly the inverse of A is again in H . Hence H is a subgroup.

(2) When $n = 1$, $\mathbf{GL}(1, \mathbf{C}) \cong \mathbf{C}^*$ that is obviously path-connected.

When $n \geq 2$, we claim that any element $A \in \mathbf{GL}(n, \mathbf{C})$ can be connected to the identity I by a continuous path lying in $\mathbf{GL}(n, \mathbf{C})$. From linear algebra, we know that any matrix is similar to an upper triangular matrix. That is, for any $A \in \mathbf{Max}(n, \mathbf{C})$ there exists $C \in \mathbf{GL}(n, \mathbf{C})$ such that

$$A = CBC^{-1}, \quad B = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \ddots & * & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

If $A \in \mathbf{GL}(n, \mathbf{C})$, then all λ_i are nonzero. Set

$$B(t) = \begin{bmatrix} \lambda_1 & (1-t)* & \cdots & (1-t)* \\ 0 & \ddots & (1-t)* & \vdots \\ \vdots & \ddots & \ddots & (1-t)* \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \quad (0 \leq t \leq 1), \quad D = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Then $A(t) := CB(t)C^{-1}$ is a continuous path in $\mathbf{GL}(n, \mathbf{C})$ from A to CDC^{-1} and $A(t) \in \mathbf{GL}(n, \mathbf{C})$. Next, for each $i \in \{1, \dots, n\}$, choose a continuous path $\lambda_i(t) \in \mathbf{C}^*$, $1 \leq t \leq 2$, such that $\lambda_i(1) = \lambda_i$ and $\lambda_i(2) = 1$. Then $\tilde{A}(t)$ is a continuous path in $\mathbf{GL}(n, \mathbf{C})$ from CDC^{-1} to I_n , where

$$\tilde{A}(t) = C \cdot \text{diag}(\lambda_1(t), \dots, \lambda_n(t)) \cdot C^{-1}, \quad 1 \leq t \leq 2.$$

Hence we find a continuous path in $\mathbf{GL}(n, \mathbf{C})$ from A to I_n .

(3) In the proof of (2), define $\lambda_i(t)$, $1 \leq i \leq n-1$ as before and

$$\lambda_n(t) = \frac{1}{\lambda_1(t) \cdots \lambda_{n-1}(t)}.$$

(4) From linear algebra, any unitary matrix $U \in \mathbf{U}(n)$ can be written as

$$U = U_1 \cdot \text{diag}(e^{\sqrt{-1}\theta_1}, \dots, e^{\sqrt{-1}\theta_n}) \cdot U_1^{-1}, \quad U_1 \in \mathbf{U}(n), \quad \theta_1, \dots, \theta_n \in \mathbf{R}.$$

If we define

$$U(t) = U_1 \cdot \text{diag}(e^{\sqrt{-1}(1-t)\theta_1}, \dots, e^{\sqrt{-1}(1-t)\theta_n}) \cdot U_1^{-1}, \quad 0 \leq t \leq 1,$$

then $U(t) \in \mathbf{U}(n)$ and connects U to I_n . □

Group	Connected?	Components
$\mathbf{GL}(n, \mathbf{C})$	Y	1
$\mathbf{SL}(n, \mathbf{C})$	Y	1
$\mathbf{GL}(n, \mathbf{R})$	N	2
$\mathbf{SL}(n, \mathbf{R})$	Y	1
$\mathbf{O}(n)$	N	2
$\mathbf{SO}(n)$	Y	1
$\mathbf{U}(n)$	Y	1
$\mathbf{SU}(n)$	Y	1
$\mathbf{O}(n, 1)$	N	4
$\mathbf{SO}(n, 1)$	N	2
Heisenberg	Y	1
$\mathbf{E}(n)$	N	2
$P(n, 1)$	N	4

Definition 1.4.12. A matrix Lie group G is **simply-connected** if it is path-connected and any loop in G can be shrunk continuously to a point in G . Precisely, assume that G is path-connected. Then G is simply-connected if for any given continuous path $A(t)$, $0 \leq t \leq 1$, lying in G with $A(0) = A(1)$, there exists a continuous function $A(s, t)$, $0 \leq s, t \leq 1$, such that

$$A(s, 0) = A(s, 1) \quad (0 \leq s \leq 1), \quad A(0, t) = A(t), \quad A(1, t) = A(1, 0) \quad (0 \leq t \leq 1).$$

Define $A_s(t) := A(s, t)$. Then $A(s, 0) = A(s, 1)$ means that A_s is a loop; $A(0, t) = A(t)$ means $A_0 = A$; $A(1, t) = A(1, 0)$ means A_1 is the point $A(0) = A(1)$.

Proposition 1.4.13. *The group $\mathbf{SU}(2)$ is simply-connected.*

PROOF. Any $A \in \mathbf{SU}(2)$ can be written as

$$A = \begin{bmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{bmatrix}, \quad \alpha, \beta \in \mathbf{C} \text{ with } |\alpha|^2 + |\beta|^2 = 1.$$

Define

$$\mathbf{SU}(2) \longrightarrow \mathbf{S}^3 \subseteq \mathbf{C}^2 \cong \mathbf{R}^4, \quad \begin{bmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{bmatrix} \longmapsto (\alpha, \beta).$$

The above map is homeomorphic so that $\mathbf{SU}(2)$ is simply-connected. \square

Group	Simply-connected?	Fundamental group
$\mathbf{SO}(2)$	N	\mathbf{Z}
$\mathbf{SO}(n)$ ($n \geq 3$)	N	$\mathbf{Z}/2\mathbf{Z}$
$\mathbf{U}(n)$	X	\mathbf{Z}
$\mathbf{SU}(n)$	Y	$\{1\}$
$\mathbf{Sp}(n)$	Y	$\{1\}$
$\mathbf{GL}^+(n, \mathbf{R})$	N	\mathbf{Z} ($n = 2$), $\mathbf{Z}/2\mathbf{Z}$ ($n \geq 3$)
$\mathbf{GL}(n, \mathbf{C})$	N	\mathbf{Z}
$\mathbf{SL}(n, \mathbf{R})$ ($n \geq 2$)	N	\mathbf{Z} ($n = 2$), $\mathbf{Z}/2\mathbf{Z}$ ($n \geq 3$)
$\mathbf{SL}(n, \mathbf{C})$	Y	$\{1\}$
$\mathbf{SO}(n, \mathbf{C})$	N	\mathbf{Z} ($n = 2$), $\mathbf{Z}/2\mathbf{Z}$ ($n \geq 3$)
$\mathbf{Sp}(n, \mathbf{R})$	N	\mathbf{Z}
$\mathbf{Sp}(n, \mathbf{C})$	Y	$\{1\}$

Let $X \in \mathbf{Max}(n, \mathbf{K})$. Define formally

$$(1.4.3.24) \quad e^X \equiv \exp X := \sum_{m \geq 0} \frac{X^m}{m!}.$$

Regarding $\mathbf{Max}(n, \mathbf{C})$ as \mathbf{C}^{n^2} , we can define the **Hilbert-Schmidt norm** by

$$(1.4.3.25) \quad \|X\| := \left(\sum_{1 \leq i, j \leq n} |X_{ij}|^2 \right)^{1/2}.$$

Then we can pullback all analysis on \mathbf{C}^{n^2} to $\mathbf{Max}(n, \mathbf{C})$. A sequence $\{X_m\}_{m \geq 1}$ in $\mathbf{Max}(n, \mathbf{C})$ **converges** to X , if $\|X_m - X\| \rightarrow 0$ as $m \rightarrow \infty$. A sequence $\{X_m\}_{m \geq 1}$ in $\mathbf{Max}(n, \mathbf{C})$ is a **Cauchy sequence** if $\|X_m - X_\ell\| \rightarrow 0$ as $m, \ell \rightarrow \infty$. Therefore, any Cauchy sequence in $\mathbf{Max}(n, \mathbf{C})$ converges.

Proposition 1.4.14. *For any $X \in \mathbf{Max}(n, \mathbf{K})$, the series (1.4.3.24) converges. Moreover e^X is continuous in X .*

PROOF. Since

$$\sum_{0 \leq m \leq M} \left\| \frac{X^m}{m!} \right\| \leq \sum_{0 \leq m \leq M} \frac{\|X\|^m}{m!} \leq e^{\|X\|},$$

it follows that the series $\sum_{m \geq 0} \|X^m/m!\|$ is convergent. Hence $\{\sum_{0 \leq i \leq m} X^i/i!\}_{m \geq 0}$ is a Cauchy sequence and then is a convergent sequence. Thus e^X is well-defined. Because the series (1.4.3.24) uniformly converges on $\{X \in \mathbf{Max}(n, \mathbf{C}) \mid \|X\| \leq R\}$ and each $X^m/m!$ is continuous, we can conclude that e^X is continuous. \square

Consider the following 2×2 matrices

$$X = \begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} b & c \\ 0 & b \end{bmatrix}$$

Clearly

$$X^2 = \begin{bmatrix} -a^2 & 0 \\ 0 & -a^2 \end{bmatrix}, \quad X^3 = \begin{bmatrix} 0 & a^3 \\ -a^3 & 0 \end{bmatrix}, \quad X^4 = \begin{bmatrix} a^4 & 0 \\ 0 & a^4 \end{bmatrix}, \quad \dots$$

Then

$$e^X = \begin{bmatrix} 1 - \frac{a^2}{2!} + \frac{a^4}{4!} + \cdots & - \left(a - \frac{a^3}{3!} + \cdots \right) \\ a - \frac{a^3}{3!} + \cdots & 1 - \frac{a^2}{2!} + \frac{a^4}{4!} + \cdots \end{bmatrix} = \begin{bmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{bmatrix}$$

On the other hand

$$XY \neq YX, \quad e^Y = \begin{bmatrix} e^b & e^b c \\ 0 & e^b \end{bmatrix}, \quad e^X e^Y \neq e^Y e^X.$$

To compute e^X for more general matrices X , we first prove the following

Proposition 1.4.15. *Let $X, Y \in \mathbf{Max}(n, \mathbf{C})$. Then*

- (1) $e^0 = I_n$.
- (2) $(e^X)^* = e^{X^*}$.
- (3) e^X is invertible and $(e^X)^{-1} = e^{-X}$.
- (4) $e^{(\alpha+\beta)X} = e^{\alpha X} e^{\beta X}$ for all $\alpha, \beta \in \mathbf{C}$.
- (5) If $XY = YX$, then $e^{X+Y} = e^X e^Y = e^Y e^X$.
- (6) If C is invertible, then $e^{CXC^{-1}} = C e^X C^{-1}$.
- (7) $\|e^X\| \leq e^{\|X\|}$.
- (8) e^{tX} is a smooth curve in $\mathbf{Max}(n, \mathbf{C})$ and

$$\frac{d}{dt} e^{tX} = X e^{tX} = e^{tX} X, \quad \left. \frac{d}{dt} \right|_{t=0} e^{tX} = X.$$

PROOF. (1) – (2) and (6) – (8) are clearly. Also (3) – (4) follow from (5). Since $XY = YX$, it follows that for any $m \geq 1$ we have

$$(X + Y)^m = \sum_{0 \leq k \leq m} \binom{m}{k} X^k Y^{m-k}.$$

Hence

$$\begin{aligned} e^X e^Y &= \left(\sum_{m \geq 0} \frac{X^m}{m!} \right) \left(\sum_{m \geq 0} \frac{Y^m}{m!} \right) = \sum_{m \geq 0} \left(\sum_{0 \leq k \leq m} \frac{X^k}{k!} \frac{Y^{m-k}}{(m-k)!} \right) \\ &= \sum_{m \geq 0} \frac{1}{m!} \left(\sum_{0 \leq k \leq m} \binom{m}{k} X^k Y^{m-k} \right) = \sum_{m \geq 0} \frac{1}{m!} (X + Y)^m = e^{X+Y}. \end{aligned}$$

Now we have $e^X e^Y = e^{X+Y} = e^{Y+X} = e^Y e^X$. \square

Exercise 1.4.16. Prove (1) – (2) and (6) – (8) in Proposition 1.4.15.

Return back to the above matrix

$$X = \begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}$$

Because the eigenvectors of X are

$$\begin{bmatrix} 1 \\ \sqrt{-1} \end{bmatrix}, \quad \begin{bmatrix} \sqrt{-1} \\ 1 \end{bmatrix}$$

with eigenvalues $-\sqrt{-1}a$ and $\sqrt{-1}a$, respectively, we get

$$X = CDC^{-1}, \quad C := \begin{bmatrix} 1 & \sqrt{-1} \\ \sqrt{-1} & 1 \end{bmatrix}, \quad D = \begin{bmatrix} -\sqrt{-1}a & 0 \\ 0 & \sqrt{-1}a \end{bmatrix}$$

From [Proposition 1.4.15](#) (6), we have

$$\begin{aligned} e^X &= Ce^DC^{-1} = \begin{bmatrix} 1 & \sqrt{-1} \\ \sqrt{-1} & 1 \end{bmatrix} \begin{bmatrix} e^{-\sqrt{-1}a} & 0 \\ 0 & e^{\sqrt{-1}a} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{-1}}{2} \\ -\frac{\sqrt{-1}}{2} & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{bmatrix} \end{aligned}$$

For any $X, Y \in \mathbf{Max}(n, \mathbf{C})$, we know from [Proposition 1.4.15](#) (8) that e^{tX} is smooth and then

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} \left(e^{tX} Y e^{-tX} \right) &= \left. \frac{d}{dt} \right|_{t=0} \left(\left(I_n + tX + t^2 \frac{X^2}{2} + \dots \right) Y \left(I_n - tX + t^2 \frac{X^2}{2} + \dots \right) \right) \\ &= \left. \frac{d}{dt} \right|_{t=0} \left(Y + t(XY - YX) + t^2(\dots) \right) = XY - YX. \end{aligned}$$

Definition 1.4.17. For any $X \in \mathbf{Max}(n, \mathbf{K})$ define

$$(1.4.3.26) \quad \mathbf{ad}_X : \mathbf{Max}(n, \mathbf{K}) \longrightarrow \mathbf{Max}(n, \mathbf{K}), \quad Y \longmapsto [X, Y] := XY - YX.$$

We also call $[X, Y]$ the **Lie bracket** of X and Y .

If $g(z) = \sum_{m \geq 0} a_m (z - a)^m$ is a formal series, then define

$$g(\mathbf{ad}_X) := \sum_{m \geq 0} a_m \left(\mathbf{ad}_X - a \mathbf{1}_{\mathbf{Max}(n, \mathbf{K})} \right)^m : \mathbf{Max}(n, \mathbf{K}) \longrightarrow \mathbf{Max}(n, \mathbf{K}).$$

In particular, we obtain the operator $e^{\mathbf{ad}_X}$ for each $X \in \mathbf{Max}(n, \mathbf{K})$. Moreover we can prove

$$(1.4.3.27) \quad e^X Y e^{-X} = e^{\mathbf{ad}_X}(Y).$$

Indeed,

$$e^{\mathbf{ad}_X}(Y) = \left(\sum_{m \geq 0} \frac{(\mathbf{ad}_X)^m}{m!} \right) Y, \quad e^X Y e^{-X} = \sum_{m \geq 0} \sum_{0 \leq k \leq m} \frac{X^k Y (-X)^{m-k}}{k! (m-k)!}.$$

To prove (1.4.3.27) we need to verify

$$(\mathbf{ad}_X)^m Y = \sum_{0 \leq k \leq m} \binom{m}{k} X^k Y (-X)^{m-k}.$$

When $m = 1$, it is trivial. Assume that $(\mathbf{ad}_X)^m Y = \sum_{0 \leq k \leq m} \binom{m}{k} X^k Y (-X)^{m-k}$. Then

$$\begin{aligned} (\mathbf{ad}_X)^{m+1} Y &= \mathbf{ad}_X \left((\mathbf{ad}_X)^m Y \right) = \left[X, \sum_{0 \leq k \leq m} \binom{m}{k} X^k Y (-X)^{m-k} \right] \\ &= \sum_{0 \leq k \leq m} \binom{m}{k} \left(X^{k+1} Y (-X)^{m-k} - X^k Y (-X)^{m+1-k} \right) \\ &= \sum_{1 \leq k \leq m+1} \binom{m}{k-1} X^k Y (-X)^{m+1-k} + \sum_{0 \leq k \leq m} \binom{m}{k} X^k Y (-X)^{m+1-k} \\ &= \sum_{0 \leq k \leq m+1} \binom{m+1}{k} X^k Y (-X)^{m+1-k}. \end{aligned}$$

A **finite-dimensional \mathbf{K} -Lie algebra** is a finite-dimensional \mathbf{K} -vector space \mathfrak{g} , together with a map $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g}$ into \mathfrak{g} , with the following properties

- (1) $[\cdot, \cdot]$ is bilinear,
- (2) $[X, Y] = -[Y, X]$ for all $X, Y \in \mathfrak{g}$,
- (3) **(Jacobi identity)** $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in \mathfrak{g}$.

It is clear that $\mathbf{Max}(n, \mathbf{K})$ together with $[\cdot, \cdot]$ defined in (1.4.3.26) is a finite-dimensional \mathbf{K} -Lie algebra.

Theorem 1.4.18. *Every invertible $n \times n$ matrix can be expressed as e^X for some $X \in \mathbf{Max}(n, \mathbf{C})$.*

Recall that the function

$$\ln z = \sum_{m \geq 1} (-1)^{m+1} \frac{(z-1)^m}{m}$$

is defined and holomorphic in $\{z \in \mathbf{C} \mid |z-1| < 1\}$. In particular, when $|u| < \ln 2$, $\ln(e^u) = u$, when $|z-1| < 1$, $e^{\ln z} = z$.

Definition 1.4.19. For any $A \in \mathbf{Max}(n, \mathbf{C})$, define

$$(1.4.3.28) \quad \ln A := \sum_{m \geq 1} (-1)^{m+1} \frac{(A - I_n)^m}{m}.$$

As in the proof of **Proposition 1.4.14**, we can prove that the series (1.4.3.28) is defined and continuous in $\{A \in \mathbf{Max}(n, \mathbf{C}) \mid \|A - I_n\| < 1\}$. Moreover, when $\|X\| < \ln 2$, $\ln e^X = X$, when $\|A - I_n\| < 1$, $e^{\ln A} = A$.

To extend **Proposition 1.4.15** (5) to the case that $[X, Y] \neq 0$, we need the following definition:

$$(1.4.3.29) \quad g(z) = \frac{\ln z}{1 - \frac{1}{z}} = \frac{z \ln z}{z-1} = 1 + \sum_{m \geq 1} \frac{(-1)^{m+1}}{m(m+1)} (z-1)^m, \quad |z-1| < 1.$$

Exercise 1.4.20. Verify (1.4.3.29).

For any $X, Y \in \mathbf{Max}(n, \mathbf{C})$ with sufficiently small $\|X\|$ and $\|Y\|$, we have

$$\begin{aligned} e^X e^Y &= I_n + (X + Y) + \left(\frac{1}{2} X^2 + XY + \frac{1}{2} Y^2 \right) \\ &\quad + \left(\frac{1}{6} X^3 + \frac{1}{2} X^2 Y + \frac{1}{2} X Y^2 + \frac{1}{6} Y^3 \right) + \dots \end{aligned}$$

Then

$$e^X e^Y - I_n = (X + Y) + \left(\frac{1}{2} X^2 + XY + \frac{1}{2} Y^2 \right) + \dots$$

and

$$\begin{aligned} (e^X e^Y - I_n)^2 &= (X + Y)^2 + \left(X^3 + Y^3 + XYX + YXY \right. \\ &\quad \left. + \frac{1}{2} Y^2 X + \frac{1}{2} Y X^2 + \frac{3}{2} X^2 Y + \frac{2}{3} X Y^2 \right) + \dots \end{aligned}$$

Therefore

$$\ln(e^X e^Y) = \sum_{m \geq 1} \frac{(-1)^{m+1}}{m} (e^X e^Y - I_n)^m = (e^X e^Y - I_n) - \frac{1}{2} (e^X e^Y - I_n)^2 + \dots$$

$$\begin{aligned}
&= X + Y + \frac{1}{2}X^2 + XY + \frac{1}{2}Y^2 - \frac{1}{2}(X^2 + XY + YX + Y^2) + \cdots \\
&= X + Y + \frac{XY - YX}{2} + \cdots = X + Y + \frac{1}{2}[X, Y] + \cdots
\end{aligned}$$

The higher order terms are given by the famous Baker-Campbell-Hausdorff formula.

Theorem 1.4.21. (Baker-Campbell-Hausdorff) For all $X, Y \in \mathbf{Max}(n, \mathbf{C})$ with $\|X\|$ and $\|Y\|$ sufficiently small, we have

$$(1.4.3.30) \quad \ln(e^X e^Y) = X + \int_0^1 g(e^{\text{ad}_X} \circ e^{t\text{ad}_Y})(Y) dt.$$

The above formula gives the next terms in the series of $\ln(e^X e^Y)$.

Exercise 1.4.22. Show by (1.4.3.30) that

$$(1.4.3.31) \quad \ln(e^X e^Y) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + \cdots$$

for sufficiently small X and Y .

We will give a proof of **Theorem 1.4.21** in a special case that X and Y commute with $[X, Y]$ (we now do not assume further that $\|X\|$ and $\|Y\|$ are sufficiently small).

Theorem 1.4.23. If $X, Y \in \mathbf{Max}(n, \mathbf{C})$ satisfies

$$(1.4.3.32) \quad [X, [X, Y]] = [Y, [X, Y]] = 0,$$

then

$$(1.4.3.33) \quad e^X e^Y = e^{X+Y+\frac{1}{2}[X, Y]}.$$

PROOF. For any $t \in \mathbf{R}$ we shall prove

$$e^{tX} e^{tY} = e^{tX+tY+\frac{t^2}{2}[X, Y]}$$

which is equivalent, under the assumption (1.4.3.32),

$$e^{t(X+Y)} = e^{tX} e^{tY} e^{-\frac{t^2}{2}[X, Y]}.$$

Set

$$A(t) := e^{t(X+Y)}, \quad B(t) := e^{tX} e^{tY} e^{-\frac{t^2}{2}[X, Y]}.$$

Clearly that

$$\frac{d}{dt} A(t) = A(t)(X + Y)$$

by **Proposition 1.4.15** (8). For $B(t)$ one has

$$\begin{aligned}
\frac{d}{dt} B(t) &= \frac{d}{dt} \left(e^{tX} e^{tY} e^{-\frac{t^2}{2}[X, Y]} \right) \\
&= \left(\frac{d}{dt} e^{tX} \right) e^{tY} e^{-\frac{t^2}{2}[X, Y]} + e^{tX} \left(\frac{d}{dt} e^{tY} \right) e^{-\frac{t^2}{2}[X, Y]} + e^{tX} e^{tY} \left(\frac{d}{dt} e^{-\frac{t^2}{2}[X, Y]} \right) \\
&= e^{tX} X e^{tY} e^{-\frac{t^2}{2}[X, Y]} + e^{tX} e^{tY} Y e^{-\frac{t^2}{2}[X, Y]} + e^{tX} e^{tY} e^{-\frac{t^2}{2}[X, Y]} (-t[X, Y]).
\end{aligned}$$

Since Y commutes with $[X, Y]$, it follows that

$$e^{tX}e^{tY}Ye^{-\frac{t^2}{2}[X,Y]} = e^{tX}e^{tY}e^{-\frac{t^2}{2}[X,Y]}Y.$$

Using (1.4.3.27) and $[Y, [X, Y]] = 0$ yields

$$Xe^{tY} = e^{tY}e^{-tY}Xe^{tY} = e^{tY}e^{-\text{ad}_Y}(X) = X - t[Y, X] = X + t[X, Y].$$

Consequently,

$$\begin{aligned} \frac{d}{dt}B(t) &= e^{tX}(X + t[X, Y])e^{-\frac{t^2}{2}[X,Y]} + e^{tX}e^{tY}e^{-\frac{t^2}{2}[X,Y]}Y + e^{tX}e^{tY}e^{-\frac{t^2}{2}[X,Y]}(-t[X, Y]) \\ &= e^{tX}e^{tY}e^{-\frac{t^2}{2}[X,Y]}(X + t[X, Y] + Y - t[X, Y]) = B(t)(X + Y). \end{aligned}$$

Because both $A(t)$ and $B(t)$ satisfy the same differential equation, and $A(0) = B(0) = I_n$, we can conclude that $A(t) \equiv B(t)$. \square

Exercise 1.4.24. (1) Show that for any two elements X, Y of the Heisenberg group H , we always have (1.4.3.32).

(2) Show by direct computation that for any $X, Y \in \mathfrak{h}$, we have (1.4.3.33), where

$$(1.4.3.34) \quad \mathfrak{h} := \left\{ \begin{bmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{bmatrix} \mid \alpha, \beta, \gamma \in \mathbf{R} \right\}.$$

Actually, \mathfrak{h} defined in (1.4.3.34) is the Lie algebra of the Heisenberg group H in the sense of **Definition 1.4.25**.

Definition 1.4.25. Let G be a matrix Lie group. The **Lie algebra of G** , denoted \mathfrak{g} , is the set of all matrices X such that e^{tX} is in G for all $t \in \mathbf{R}$.

If we write $A(t) := e^{tX}$, then $A(0) = I_n$, $A(t)$ is continuous in t , and $A(t+s) = A(t)A(s)$ for all $t, s \in \mathbf{R}$. Such a function $A : \mathbf{R} \rightarrow \mathbf{GKL}(n, \mathbf{C})$ is called a **one-parameter subgroup** of $\mathbf{GL}(n, \mathbf{C})$. Conversely, we can prove that for a given one-parameter subgroup $A(t)$ of $\mathbf{GL}(n, \mathbf{C})$ there exists a unique $X \in \mathbf{Max}(n, \mathbf{C})$ such that $A(t) = e^{tX}$.

Theorem 1.4.26. For any $X \in \mathbf{Max}(n, \mathbf{C})$ we have

$$(1.4.3.35) \quad \det(e^X) = e^{\text{tr}(X)}.$$

PROOF. From linear algebra, we know that X can be uniquely written as

$$X = S + N,$$

where $S, N \in \mathbf{Mat}(n, \mathbf{C})$, $SN = NS$, S is diagonalizable and N is nilpotent. From $SN = NS$, we obtain

$$e^X = e^{S+N} = e^S e^N, \quad \det(e^X) = \det(e^S) \det(e^N).$$

Write $S = C \text{diag}(\lambda_1, \dots, \lambda_n) C^{-1}$ for some $C \in \mathbf{GL}(n, \mathbf{C})$. Then

$$e^S = C \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) C^{-1}$$

and

$$\det(e^S) = \prod_{1 \leq i \leq n} e^{\lambda_i} = e^{\sum_{1 \leq i \leq n} \lambda_i} = e^{\text{tr}(S)}.$$

Since N is nilpotent, from linear algebra, we can find some $D \in \mathbf{GL}(n, \mathbf{C})$ such that

$$N = D \begin{bmatrix} 0 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} D^{-1}$$

Hence

$$e^N = D \begin{bmatrix} 1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} D^{-1}$$

So $\det(e^N) = 1$ and $\operatorname{tr}(N) = 0$. Consequently $\det(e^X) = e^{\operatorname{tr}(S)} = e^{\operatorname{tr}(X)}$. \square

We now can determine Lie algebras of classical groups.

- (1) $G = \mathbf{GL}(n, \mathbf{C})$: If $X \in \mathbf{Max}(n, \mathbf{C})$, we have $e^{tX} \in \mathbf{GL}(n, \mathbf{C})$. Hence the Lie algebra $\mathfrak{gl}(n, \mathbf{C})$ of $\mathbf{GL}(n, \mathbf{C})$ is $\mathbf{Max}(n, \mathbf{C})$.
- (2) $G = \mathbf{GL}(n, \mathbf{R})$: The Lie algebra $\mathfrak{gl}(n, \mathbf{R})$ of $\mathbf{GL}(n, \mathbf{R})$ is $\mathbf{Max}(n, \mathbf{R})$.
- (3) $G = \mathbf{SL}(n, \mathbf{C})$: If $e^{tX} \in \mathbf{SL}(n, \mathbf{C})$, then from (1.4.3.35) we have

$$1 = \det(e^{tX}) = e^{\operatorname{tr}(tX)} = e^{t \operatorname{tr}(X)} \iff \operatorname{tr}(X) = 0.$$

Hence the Lie algebra $\mathfrak{sl}(n, \mathbf{C})$ of $\mathbf{SL}(n, \mathbf{C})$ is given by

$$(1.4.3.36) \quad \mathfrak{sl}(n, \mathbf{C}) = \{X \in \mathbf{Max}(n, \mathbf{C}) \mid \operatorname{tr}(X) = 0\}.$$

- (4) $G = \mathbf{SL}(n, \mathbf{R})$: The Lie algebra $\mathfrak{sl}(n, \mathbf{R})$ of $\mathbf{SL}(n, \mathbf{R})$ is given by

$$(1.4.3.37) \quad \mathfrak{sl}(n, \mathbf{R}) = \{X \in \mathbf{Max}(n, \mathbf{R}) \mid \operatorname{tr}(X) = 0\}.$$

- (5) $G = \mathbf{U}(n)$: $e^{tX} \in \mathbf{U}(n)$ if and only if $(e^{tX})^* = (e^{tX})^{-1} = e^{-tX}$, or $e^{tX^*} = e^{-tX}$. Thus the Lie algebra $\mathfrak{u}(n)$ of $\mathbf{U}(n)$ is

$$(1.4.3.38) \quad \mathfrak{u}(n) = \{X \in \mathbf{Max}(n, \mathbf{C}) \mid X^* = -X\}.$$

- (6) $G = \mathbf{SU}(n)$: The Lie algebra $\mathfrak{su}(n)$ of $\mathbf{SU}(n)$ is

$$(1.4.3.39) \quad \mathfrak{su}(n) = \{X \in \mathbf{Max}(n, \mathbf{C}) \mid X^* = -X, \operatorname{tr}(X) = 0\}.$$

- (7) $G = \mathbf{O}(n)$: $e^{tX} \in \mathbf{O}(n)$ if and only if $(e^{tX})^\top = (e^{tX})^{-1}$ or $e^{tX^\top} = e^{-tX}$. Thus the Lie algebra $\mathfrak{o}(n)$ of $\mathbf{O}(n)$ is

$$(1.4.3.40) \quad \mathfrak{o}(n) = \{X \in \mathbf{Max}(n, \mathbf{R}) \mid X^\top = -X\}.$$

If $X \in \mathfrak{o}(n)$, then $\operatorname{tr}(X) = 0$ and $\det(e^X) = 0$.

- (8) $G = \mathbf{SO}(n)$: The Lie algebra $\mathfrak{so}(n)$ of $\mathbf{SO}(n)$ is the same as $\mathfrak{o}(n)$.

- (9) $G = \mathbf{O}(n, k)$: The Lie algebra of $\mathbf{O}(n, k)$ is given by

$$(1.4.3.41) \quad \mathfrak{o}(n, k) = \left\{ X \in \mathbf{Max}(n+k, \mathbf{R}) \mid gX^\top g = -X, g = \begin{bmatrix} I_n & 0 \\ 0 & -I_k \end{bmatrix} \right\}.$$

- (10) $G = \mathbf{SO}(n, k)$: the Lie algebra $\mathfrak{so}(n, k)$ is the same as $\mathfrak{o}(n, k)$.

(11) $G = \mathbf{Sp}(n, \mathbf{C})$: The Lie algebra of $\mathbf{Sp}(n, \mathbf{C})$ is given by

$$(1.4.3.42) \quad \mathfrak{sp}(n, \mathbf{C}) = \left\{ \begin{bmatrix} A & B \\ C & -A^T \end{bmatrix} \mid A \in \mathbf{Max}(n, \mathbf{C}) \text{ and } B, C \text{ symmetric} \right\}.$$

(12) $G = \mathbf{Sp}(n, \mathbf{R})$: The Lie algebra of $\mathbf{Sp}(n, \mathbf{R})$ is given by

$$(1.4.3.43) \quad \mathfrak{sp}(n, \mathbf{R}) = \left\{ X \in \mathbf{Max}(2n, \mathbf{R}) \mid JX^T J = X, J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \right\}.$$

(13) $G = \mathbf{Sp}(n)$: The Lie algebra of $\mathbf{Sp}(n)$ is given by $\mathfrak{sp}(n) = \mathfrak{sp}(n, \mathbf{C}) \cap \mathfrak{u}(2n)$.

(14) $G = \text{Heisenberg group } H$: The Lie algebra of H is given by (1.4.3.34).

(15) $G = \mathbf{E}(n)$: The Lie algebra of $\mathbf{E}(n)$ is given by

$$(1.4.3.44) \quad \mathfrak{e}(n) = \left\{ \begin{bmatrix} Y & y^T \\ \mathbf{0} & 1 \end{bmatrix} \mid Y \in \mathbf{Max}(n, \mathbf{R}), Y^T = -Y \right\}.$$

(16) $G = \mathbf{P}(n, 1)$: The Lie algebra of $\mathbf{P}(n, 1)$ is given by

$$(1.4.3.45) \quad \mathfrak{p}(n, 1) = \left\{ \begin{bmatrix} Y & y^T \\ \mathbf{0} & 0 \end{bmatrix} \mid Y \in \mathfrak{so}(n, 1) \right\}.$$

1.5. Connectedness and compactness

In this section we give general definitions of connectedness and compactness.

1.5.1. Connectedness. Let (X, \mathcal{T}) be a topological space.

(1) A **separation** of X is a pair (U, V) such that

$$\emptyset \neq U, V \subseteq \mathcal{T}, \quad U \cap V = \emptyset, \quad X = U \cup V.$$

(2) X is called **disconnected** if there exists a separation of X . If there does not, X is called **connected**. Thus, a topological space X is called connected if it is not the disjoint union of two nonempty open subsets.

(3) A subset A of X is called **clopen** if it is both open and closed in X . It is clear that X is connected if and only if its only clopen subsets are X and \emptyset . If A is a nonempty proper subset of X that is clopen, then the pair $(U = A, V = X \setminus A)$ is a separation of X . Conversely, if X is connected, then for any subset $A \subseteq X$ that is clopen in X , we conclude that $(U = A, V = X \setminus A)$ is not a separation.

(4) A **discrete valued map** is a continuous map from X to a discrete space D .

(5) X is connected if and only if every discrete valued map on X is constant. If X is connected and $f : X \rightarrow D$ is a discrete valued map and if $y \in D$ is in the range of f , then $f^{-1}(y)$ is clopen and nonempty and so must equal X , and so f is constant with only value y . Conversely, if X is not connected then $X = U \cup V$ for some disjoint clopen sets U and V . Then the map $f : X \rightarrow \{0, 1\}$ which is 0 on U and is 1 on V is a nonconstant discrete valued map.

(6) A subset A of X is called **connected**, if A is connected with respect to the subspace topology.

(7) Define a relation \sim on X by

$$x \sim y \iff x \text{ and } y \text{ belong to a connected subset of } X.$$

Then \sim is an equivalence relation on X , and the equivalence class $[x]$ is called a **connected component** or **component** of X .

(8) Define a relation \sim on X by

$$x \sim y \iff f(x) = f(y) \text{ for every discrete valued map } f \text{ on } X.$$

Then \sim is an equivalence relation on X , and the equivalence class $[x]$ is called a **quasi-component** of X .

Proposition 1.5.1. (1) *If $f : (X, \mathcal{T}_X) \rightarrow (Y, \mathcal{T}_Y)$ is continuous and X is connected, then $f(X)$ is connected.*

(2) *If $(Y_i)_{i \in I}$ is a collection of connected subsets in a topological space X and if no two of the Y_i are disjoint, then $\cup_{i \in I} Y_i$ is connected.*

(3) *Components of a topological space X are connected and closed. Each connected subset is contained in a component. Thus the components are "maximal connected subsets". Components are either equal or disjoint, and fill out X .*

(4) *Quasi-components of a topological space X are closed. Each connected subset is contained in a quasi-component. In particular, each component is contained in a quasi-component. Quasi-components are either equal or disjoint, and fill out X .*

(5) **(Intermediate value theorem)** *If $f : X \rightarrow \mathbf{R}$ is continuous and $x, y \in X$, then for any M between $f(x)$ and $f(y)$ we have $f(z) = M$ for some $z \in X$.*

(6) *If X_1, \dots, X_k are connected, then $X_1 \times \dots \times X_k$ is connected.*

(7) *If \sim is an equivalence relation on a topological space X , then X / \sim is connected.*

(8) *Suppose that (X, \mathcal{T}) is a topological space and U, V are disjoint open subsets of X . If S is a connected subset of X contained in $U \cup V$, then either $S \subseteq U$ or $S \subseteq V$.*

(9) *If (X, \mathcal{T}) is a topological space that contains a dense connected subset S , then X is connected.*

(10) *Let (X, \mathcal{T}) be a topological space. The union of a collection of connected subspaces of X that have a point in common is connected.*

(11) *Suppose that (X, \mathcal{T}) is a topological space and $A \subseteq X$ is connected. If a subset B of X satisfies $A \subseteq B \subseteq \bar{A}$, then B is connected. In particular, \bar{A} is connected.*

PROOF. (1) Let $d : f(X) \rightarrow D$ be a discrete valued map. Then $d \circ f : X \rightarrow D$ is a discrete valued map on X and hence must be constant. This implies that d is constant on $f(X)$ and then $f(X)$ is connected.

(2) Let $d : \cup_{i \in I} Y_i \rightarrow D$ be a discrete valued map. Let p, q be any two points in $\cup_{i \in I} Y_i$. Suppose $x \in Y_i, y \in Y_j$, and $z \in Y_i \cap Y_j$. Because d is constant on each Y_i , we conclude that $d(x) = d(z) = d(y)$. Hence d is constant.

(3) By definition, the component of X containing x is the union of all connected subsets containing x , hence it is connected by (2). Since the closure of a connected set is connected⁵, it follows that components of X is closed.

(4) If $x \in X$ then the quasi-component containing x is the set

$$\{y \in X : d(y) = d(x) \text{ for all discrete valued maps } d\} = \bigcap_{d \text{ discrete value maps}} d^{-1}(d(x))$$

⁵Suppose that A is a connected subset. If U and V are disjoint open subsets of X that separates \bar{A} , then A is connected in one of them, say $A \subseteq U$. Since $U \cap V = \emptyset$, it follows that $\bar{A} \subseteq U$, which means that $\bar{A} \cap V = \emptyset$, a contradiction.

which is an intersection of closed subsets and hence is closed.

(5) This follows from the fact that $f(X)$ is an interval (a nonempty subset of \mathbf{R} is connected if and only if it is an interval).

(6) By induction, we need only to consider a product of two topological spaces. Let X, Y be connected and $d : X \times Y \rightarrow D$ a discrete valued map. Given a point $(x_0, y_0) \in X \times Y$ and define maps

$$\iota_{x_0} : Y \longrightarrow X \times Y, \quad y \longmapsto (x_0, y) \quad \iota_{y_0} : X \longrightarrow X \times Y, \quad x \longmapsto (x, y_0).$$

Then we get maps $d \circ \iota_{x_0} : Y \rightarrow D$ and $d \circ \iota_{y_0} : X \rightarrow D$. Since X and Y are connected, it follows that $d \circ \iota_{x_0} = c_{x_0}$ and $d \circ \iota_{y_0} = c_{y_0}$ for some constants c_{x_0}, c_{y_0} (may depending on (x_0, y_0)). In particular, $c_{x_0} = c_{y_0} = d(x_0, y_0) =: c_{x_0, y_0}$. Thus $d(x_0, y) = d(x, y_0) = c_{x_0, y_0}$ for any $x \in X, y \in Y$. We now show that the constant c_{x_0, y_0} is independent of the choice of (x_0, y_0) . Indeed, for another pair (x_1, y_1) , we also have $d(x_1, y) = d(x, y_1) = c_{x_1, y_1}$ for any $x \in X, y \in Y$. Then

$$c_{x_1, y_1} = d(x_0, y_1) = c_{x_0, y_0}.$$

Thus $d(x_0, y_0) = c$ for some constant c that is independent of the choice of (x_0, y_0) . Since (x_0, y_0) was arbitrary, we get d is constant and then $X \times Y$ is connected.

(7) This follows from the fact that $X \rightarrow X/\sim$ is surjective and (1).

(8) Since U, V are open in X , it follows that $A \cap U$ and $A \cap V$ are open in A . The connectedness of A implies that $A \cap U = \emptyset$ or $A \cap V = \emptyset$. Thus $A \subseteq V$ or $A \subseteq U$.

(9) Let $A \subseteq X$ be a dense connected subset of X . Assume that (U, V) is a separation of X . By (8), we have $A \subseteq U$ or $A \subseteq V$. If $A \subseteq U$, then $X = \overline{A} \subseteq \overline{U} = U$ (since U is both open and closed). But this implies $V = \emptyset$. Hence X is connected.

(10) Let $\{A_\alpha\}_{\alpha \in I}$ be a collection of connected subspaces of X and $p \in \bigcap_{\alpha \in I} A_\alpha$. Set $Y := \bigcup_{\alpha \in I} A_\alpha$ and suppose (C, D) is a separation of Y . Then $p \in C$ or $p \in D$. Assume that $p \in C$. Since A_α is connected, from (8), we must have $A_\alpha \subseteq C$ for all $\alpha \in I$. Thus $Y = \bigcup_{\alpha \in I} A_\alpha \subseteq C$. This contradiction shows that Y is connected.

(11) Consider a separation (C, D) of B . According to (8), $A \subseteq C$ or $A \subseteq D$. Suppose that $A \subseteq C$. Then $\overline{A} \subseteq \overline{C}$. Since $\overline{C} \cap D = C \cap D = \emptyset$, it follows that $B \cap D = \emptyset$. This contradicts the fact that $\emptyset \neq D \subseteq B$. \square

Example 1.5.2. \mathbf{R}^ω is not connected in the box topology, but it is connected in the product topology. Indeed, we have

$$\mathbf{R}^\omega = A \cup B$$

where $A \cap B = \emptyset$, and

$$A := \{\text{bounded sequences of } \mathbf{R}\} \neq \emptyset, \quad B := \{\text{unbounded sequences of } \mathbf{R}\} \neq \emptyset.$$

We claim that both A and B are clopen in the box topology. For all $\mathbf{a} = (a^i)_{i \geq 1} \in \mathbf{R}^\omega$ consider the open set

$$U_{\mathbf{a}} := \prod_{i \geq 1} (a^i - 1, a^i + 1).$$

Then $U_{\mathbf{a}} \subseteq A$ if $\mathbf{a} \in A$ and $U_{\mathbf{a}} \in B$ if $\mathbf{a} \in B$.

To prove that \mathbf{R}^ω is connected in the product topology we shall use the fact that \mathbf{R} is connected (see **Theorem 1.5.3**). Define

$$\begin{aligned}\tilde{\mathbf{R}}^n &:= \left\{ x = (x^i)_{i \geq 1} \in \mathbf{R}^\omega \mid x^i = 0 \text{ for } i > n \right\} \cong \mathbf{R}^n, \\ \mathbf{R}^\infty &:= \bigcup_{n \geq 1} \tilde{\mathbf{R}}^n\end{aligned}$$

From **Proposition 1.5.1** (6), \mathbf{R}^n is connected. Since $\tilde{\mathbf{R}}^n$ have the common point $\mathbf{0}$, it follows from **Proposition 1.5.1** (10) that \mathbf{R}^∞ is also connected. We now claim that

$$\overline{\mathbf{R}^\infty} = \mathbf{R}^\omega$$

which, by **Proposition 1.5.1** (11), implies the connectedness of \mathbf{R}^ω in the product topology. For $a = (a^i)_{i \geq 1} \in \mathbf{R}^\omega$, let $U := \prod_{i \geq 1} U_i$ be a basis element for the product topology that contains a . Then there exists an integer $N \in \mathbf{N}$ such that $U_i = \mathbf{R}$ for $i > N$. Hence $x = (x^1, \dots, x^N, 0, \dots, 0) \in \mathbf{R}^\infty \cap U$. Thus $U \cap \mathbf{R}^\infty \neq \emptyset$ and $a \in \overline{\mathbf{R}^\infty}$.

A subset $I \subseteq \mathbf{R}$ is said to be an **interval** if it contains more than one element, and whenever $a, b \in I$, every c such that $a < c < b$ is also in I . An interval must be one of the nine types of sets

$$[a, b], (a, b), [a, b), (a, b], (-\infty, b], (-\infty, b), [a, +\infty), (a, +\infty), (-\infty, +\infty).$$

Theorem 1.5.3. \mathbf{R} is connected. Moreover a nonempty subset of \mathbf{R} is connected if and only if it is a singleton or an interval.

PROOF. We assume that $I \subseteq \mathbf{R}$ contains at least two points.

Assume first that I is not an interval. By definition, there exists $a, b, c \in \mathbf{R}$ so that $a < c < b$, $a, b \in I$ but $c \notin I$. Then the sets $(-\infty, c) \cap I$ and $(c, +\infty) \cap I$ form a separation for I . Thus I is not connected.

Now we assume that I is an interval. It is not connected, there are open subsets $U, V \subseteq \mathbf{R}$ such that $(U \cap I, V \cap I)$ forms a separation for I . Choose $a \in U \cap I$ and $b \in V \cap I$. We may without loss of generality that $a < b$. Since I is an interval, it follows that $[a, b] \subseteq I$. Since U and V are open, there exists $\epsilon > 0$ such that $[a, a + \epsilon) \subseteq U \cap I$ and $(b - \epsilon, b] \subseteq V \cap I$. Let $c := \sup(U \cap [a, b])$. Then $a + \epsilon \leq c < b - \epsilon$ and $c \in (a, b) \subseteq I \subseteq U \cup V$. If $c \in U$, then $(c - \epsilon', c + \epsilon') \subseteq U \cap (a, b)$ for some $\epsilon' > 0$; if $c \in V$, then $(c - \epsilon'', c + \epsilon'') \subseteq V$ for some $\epsilon'' > 0$. Both cases contradict with the definition of c . Therefore I is connected. \square

But in general the components of X need not to be open in X . For example, consider $\mathbf{Q} \subseteq \mathbf{R}$. Then each component of \mathbf{Q} consists of a single point so that this component is not open.

1.5.2. Path-connectedness. Let (X, \mathcal{T}) be a topological space and $x, y \in X$. A **path in X from x to y** is a continuous map $f : [0, 1] \rightarrow X$ such that $f(0) = x$ and $f(1) = y$.

- (1) X is **path-connected** if for any two points x, y in X there is a path in X from x to y .
- (2) Define a relation \sim on X by

$$x \sim y \iff \text{there exists a path in } X \text{ from } x \text{ to } y.$$

Then \sim is an equivalence relation on X , and the equivalence class $[x]$ is called a **path-component** of X .

- (3) X is called **locally connected** if for each $x \in X$ and each neighborhood N of x , there is a connected neighborhood V of x with $V \subseteq N$. Equivalently, X is locally connected if it admits a basis of connected open subsets.
- (4) X is called **locally path-connected** if it admits a basis of path-connected open subsets. Equivalently, X is locally path-connected if for each $x \in X$ and each neighborhood N of x , there is a path-connected neighborhood V of x with $V \subseteq N$.

Remark 1.5.4. (1) *The following sets*

$$\mathbf{R}^n, \mathbf{R}^n \setminus \{0\} \ (n \geq 2), \mathbf{B}^n \ (n \geq 1), \mathbf{S}^n \ (n \geq 1), \mathbf{T}^n, \text{ convex sets}$$

are all path-connected.

(2) *The topologist's sin curve*

$$X = \left\{ (x, y) \in \mathbf{R}^2 \mid x = 0, -1 \leq y \leq 1 \right\} \cup \left\{ (x, y) \in \mathbf{R}^2 \mid y = \sin \frac{1}{x}, 0 < x \leq 1 \right\}$$

is connected, but not path-connected.

PROOF. Indeed, Write $S := \{(x, \sin(1/x)) \mid x \in (0, 1]\}$. Then S is connected since $S = (1 \times f)((0, 1])$ for $f(x) := \sin(1/x)$ (by **Proposition 1.5.1** (1)), \bar{S} is also connected (by **Proposition 1.5.1** (11)), and $\bar{S} = X$.

We now show that \bar{S} is not path-connected. Suppose there exists a path $f : [a, c] \rightarrow \bar{S}$ beginning at $(0, 0)$ and ending at a point of S . Set

$$I := \{t \in [a, c] \mid f(t) \in \{0\} \times [-1, 1]\}.$$

Then I is closed in $[a, c]$ and $b := \sup I$ exists. Moreover $f : [b, c] \rightarrow \bar{S}$ is a path that maps b into $\{0\} \times [-1, 1]$ and maps $(b, c]$ to S . Write $f(t) = (x(t), y(t))$, $t \in [b, c]$. Then $x(b) = 0$, $x(t) > 0$ for all $t \in (b, c]$, and $y(t) = \sin(1/x(t))$ for $t \in (b, c]$. Given $n \in \mathbf{N}$, there exists $u \in (0, x(b + \frac{1}{n}))$ such that $\sin \frac{1}{u} = (-1)^n$. Then there exists $t_n \in (b, b + \frac{1}{n})$ such that $x(t_n) = u$. Hence $y(t_n) = \sin(1/u) = (-1)^n$. But $t_n \rightarrow b$, on the other hand we have $y(t_n) \rightarrow y(b) \in [-1, 1]$. This contradiction indicates X is not path-connected. \square

(3) *If $U \subseteq \mathbf{R}^2$ is an open and connected subspace, then U is path-connected.*

PROOF. Given $x_0 \in U$, define

$$\tilde{U} := \{x \in U \mid \exists \text{ path in } U \text{ joining } x_0 \text{ to } x\} \subseteq U.$$

Since $x_0 \in \tilde{U}$, we claim that \tilde{U} is both open and closed in U so that $\tilde{U} = U$. For $x \in \tilde{U}$, there exists a path $f \in [0, 1] \rightarrow U$ such that $f(0) = x_0$ and $f(1) = x$. Since U is open, it follows that there exists an open ball $\mathbf{B}^2(x, \epsilon) \subseteq U$ and then $\mathbf{B}^2(x, \epsilon) \subseteq \tilde{U}$. To prove the closedness of \tilde{U} , we take a sequence $\{x_n\}_{n \geq 1}$ in \tilde{U} such that $x_n \rightarrow x \in U$. For any $\epsilon > 0$, there exists $N \in \mathbf{N}$ such that $x_n \in \mathbf{B}^2(x, \epsilon)$ for all $n \geq N$. Since $x_n \in \tilde{U}$, we can find a path connecting x_0 to x_n . Therefore we can find a path joining x_0 to x . Thus $x \in \tilde{U}$. \square

(4) We have the following relations:

$$\begin{array}{ccccc}
 \mathbf{PC} & \xleftarrow{+\mathbf{LPC}} & \mathbf{C} & \xleftarrow{\quad} & \mathbf{PC} \\
 & & & & \uparrow +\mathbf{C} \\
 & & \mathbf{LC} & \xleftarrow{\quad} & \mathbf{LPC}
 \end{array}$$

where \mathbf{C} = connected, \mathbf{PC} = path-connected, \mathbf{LC} = locally connected, and \mathbf{LPC} = locally path-connected.

Proposition 1.5.5. Let (X, \mathcal{T}) be a topological space.

- (1) If X is path-connected, then X is connected.
- (2) Each path-component is contained in a single component, and each component is a disjoint union of path components.
- (3) If $A \subset X$ is path-connected, then A is contained in a single path-component.
- (4) If X is locally connected, then any open subset of X is locally connected and each component of X is open in X .
- (5) If X is locally path-connected, then X is locally connected, any open subset of X is locally path-connected, each path-component of X is open, and its path-components are the same as its components.
- (6) If X is locally path-connected, then X is connected if and only if it is path-connected.
- (7) If $f : X \rightarrow Y$ is continuous and X is path-connected, then $f(X)$ is path-connected.
- (8) Any finite product of path-connected spaces is path-connected.
- (9) If \sim is an equivalence relation on a path-connected space X , then X/\sim is path-connected.

PROOF. (1) Suppose that X is path-connected but not connected. Choose a separation (U, V) of X , choose $p \in U$ and $q \in V$, and choose a path γ from p to q in X . Then $f^{-1}(U)$ and $f^{-1}(V)$ are disjoint open subsets of $[0, 1]$ that cover $[0, 1]$; moreover $0 \in f^{-1}(U)$ and $1 \in f^{-1}(V)$, so that $(f^{-1}(U), f^{-1}(V))$ is a separation of $[0, 1]$, which is a contradiction.

(2) and (3) can be proved as **Proposition 1.4.1** (3).

(4) The first statement is obvious. Let A be a component of X . If $x \in A$, then x has a connected neighborhood U by local connectedness. Then $U \subset A$. Thus A is open.

(5) The first two statements are obvious. Let B be a path-component and A the single component containing A . By (2), A can be written as a disjoint union of path-components, each of which is open in X and thus in A . If $B \neq A$, then $(B, A \setminus B)$ is a separation of A , which is a contradiction since A is connected.

(6) If X is connected, then X has only one component and hence one path-component; thus X is path-connected. Conversely, if X is path-connected, by (5), X has only one component so that X is connected. \square

Exercise 1.5.6. Prove **Proposition 1.5.5** (7)-(9).

Consider the topologist's sin curve $X = \bar{S}$ in **Remark 1.5.4**. It has one component, two path components S and $V := \{0\} \times [-1, 1]$, S is open in \bar{S} but not closed,

and V is closed but not open in \bar{S} . Let

$$\hat{S} := \bar{S} \setminus (\{0\} \times (\mathbf{Q} \cap [-1, 1])).$$

Then \hat{S} has still one component but uncountably many path components.

Example 1.5.7. (1) Each interval in \mathbf{R} is both connected and locally connected.

(2) $[-1, 0) \cup (0, 1] \subseteq \mathbf{R}$ is not connected, but is locally connected.

(3) The topologist's sine curve X is connected but not locally connected. Otherwise, X must be path-connected!

(4) \mathbf{Q} is neither connected nor locally connected.

There are no relations between connectedness and local connecteness.

Theorem 1.5.8. Any topological manifold is locally path-connected. Consequently, by **Proposition 1.5.5**, for any topological manifold \mathcal{M} , \mathcal{M} is connected if and only if it is path-connected.

PROOF. By definition, for any m -dimensional topological manifold \mathcal{M} we have $\mathcal{M} = \cup_{i \in \mathbf{N}} \mathcal{U}_i$ where $\varphi_i : \mathcal{U}_i \rightarrow \mathbf{B}_i \subset \mathbf{R}^m$ is homeomorphic for each i . Hence \mathcal{M} is locally path-connected. \square

1.5.3. Compactness. Let (X, \mathcal{T}) be a topological space.

(1) X is **compact** if any open cover of X has a finite subcover.

(2) A collection \mathcal{U} of sets has the **finite intersection property** if the intersection of any finite subcollection is nonempty.

A topological space (X, \mathcal{T}) is compact if and only if for every collection of closed subsets of X which has the finite intersection property, the intersection of the entire collection is nonempty.

PROOF. Given a collection \mathcal{U} of subsets of X , let $\mathcal{C} := \{X \setminus U \mid U \in \mathcal{U}\}$ be the collection of their complements.

(a) \mathcal{U} is a collection of open subsets if and only if \mathcal{C} is a collection of closed subsets.

(b) $\{\mathcal{U}_i\}_{1 \leq i \leq n} \subseteq \mathcal{U}$ covers X if and only if $\cap_{1 \leq i \leq n} C_i = \emptyset$ where $C_i := X \setminus \mathcal{U}_i$.

(c) \mathcal{U} covers X if and only if $\cap_{C \in \mathcal{C}} C = \emptyset$.

Then

$$X \text{ is compact} \iff \left\{ \text{given any collection } \mathcal{U} \text{ of open subsets of } X \right.$$

$$\left. \text{that covers } X, \exists \text{ finite subcollection of } \mathcal{U} \text{ covers } X \right\} \iff$$

$$\left\{ \text{given any collection } \mathcal{U} \text{ of open subsets of } X, \text{ if no finite subcollection of} \right.$$

$$\left. \mathcal{U} \text{ covers } X, \text{ then } \mathcal{U} \text{ does not cover } X \right\} \iff \text{the desired statement}$$

by the above statements (a) – (c). \square

(3) X is **locally compact** if for any $x \in X$, there is a compact subset of X containing a neighborhood of x .

(4) A subset $S \subseteq X$ is **precompact** or **relatively compact** if $\bar{S} \subseteq X$ is compact.

- (5) A continuous map $f : X \rightarrow Y$ between topological spaces is **proper** if $f^{-1}(K) \subseteq X$ is compact whenever $K \subseteq Y$ is compact.

Recall that a subset $S \subseteq \mathbf{R}^n$ is compact if it is closed and bounded.

Example 1.5.9. (1) \mathbf{R} is not compact, since the open cover $\mathcal{U} = \{(n, n+2)\}_{n \in \mathbf{Z}}$ of \mathbf{R} contains no finite subcollection that cover \mathbf{R} .

(2) $X = \{0\} \cup \{1/n\}_{n \geq 1}$ is compact. Given an open cover \mathcal{U} of X , there exists an open subset $U \in \mathcal{U}$ containing 0. Take $N \in \mathbf{N}$ such that $1/n \in U$ for all $n > N$. For each $i \in \{1, \dots, N\}$, we choose $U_i \in \mathcal{U}$ containing $1/i$. Then U, U_1, \dots, U_N form a finite subcollection of \mathcal{U} that contains X .

(3) If X is finite, then X is compact.

(4) $(0, 1]$ is not compact, because the open cover $\mathcal{U} = \{(1/n, 1]\}_{n \geq 1}$ contains no finite subcollection covering $(0, 1]$.

Let Y be a subset of a topological space (X, \mathcal{T}) . Then Y is compact (in the subspace topology) if and only if every cover of Y by open subsets of X contains a finite subcollection covering Y . Let $\mathcal{U} = \{U_\alpha\}_{\alpha \in I}$ be a cover of Y by subsets open in X . Then $\mathcal{U}_Y := \{U_\alpha \cap Y \mid \alpha \in I\}$ is a cover of Y by subsets open in Y . Since Y is compact, there exists a finite subcollection $\{U_{\alpha_i} \cap Y\}_{1 \leq i \leq n}$ covers Y . Thus $\{U_{\alpha_i}\}_{1 \leq i \leq n}$ is a subcollection of \mathcal{U} that covers Y . Conversely, let $\mathcal{U}' = \{U'_\alpha \mid \alpha \in I'\}$ be a cover of Y by subsets open in Y . For each $\alpha \in I'$, $U'_\alpha = U_\alpha \cap Y$ for some open subset U_α in X . Then $\mathcal{U} = \{U_\alpha\}_{\alpha \in I'}$ is a cover of Y by subsets open in X . By hypothesis, there is a finite subcollection $\{U_{\alpha_i}\}_{1 \leq i \leq n}$ covers Y . Then $\{U'_{\alpha_i}\}_{1 \leq i \leq n}$ is a subcollection of \mathcal{U}' that covers Y .

Proposition 1.5.10. (1) If $f : X \rightarrow Y$ is a continuous map between topological spaces and X is compact, then $f(X)$ is compact.

(2) (**Extreme value theorem**) If $f : X \rightarrow \mathbf{R}$ is a continuous map and X is compact, then f is bounded and f attains its minimum and maximum in X .

(3) If X is compact and $U \subset X$ is closed, then U is compact.

(4) If X is Hausdorff and $U \subset X$ is compact, then U is closed.

(5) If X is Hausdorff, $K, L \subset X$ are compact, and $K \cap L = \emptyset$, then $K \subset U$ and $L \subset V$ for some disjoint open subsets $U, V \subset X$.

(6) The finite product of compact spaces is compact.

(7) If X is compact and \sim is any equivalence relation, then X/\sim is compact.

(8) (**Closed map lemma**) Let $f : X \rightarrow Y$ be a continuous map from a compact space X to a Hausdorff space Y . Then f is closed. If f is moreover surjective (injective, bijective), then f is a quotient map (topological embedding, homeomorphism).

(9) If X is Hausdorff, then the following are equivalent:

- X is locally compact,
- any $x \in X$ has a precompact neighborhood,
- X has a basis of precompact open sets.

(10) Any open or closed subset of a locally compact Hausdorff space are locally compact Hausdorff.

(11) If $f : X \rightarrow Y$ is a proper map between locally compact Hausdorff spaces, then f is a closed map.

PROOF. (1) Let \mathcal{U} be an open cover of $f(X)$. For each $U \in \mathcal{U}$, $f^{-1}(U)$ is an open subset of X . Since \mathcal{U} covers $f(X)$, it follows that $f^{-1}(\mathcal{U}) := \{f^{-1}(U) : U \in \mathcal{U}\}$ is a cover of X . Since X is compact, there is a finite subcollection $\{f^{-1}(U_i)\}_{1 \leq i \leq n}$ covers X . Then $\{U_i\}_{1 \leq i \leq n}$ is a finite subcollection of \mathcal{U} covers $f(X)$.

\mathcal{U} is an open cover of X . By compactness of X , there exists $N \in \mathbf{N}$ such that $(f^{-1}(U_i))_{1 \leq i \leq N}$ cover X . Then $(U_i)_{1 \leq i \leq N}$ covers $f(X)$.

(2) By (1), $f(X)$ is a compact subset of \mathbf{R} so that $f(X)$ is closed and bounded. In particular, it contains its maximum and minimum in X .

(3) Let \mathcal{U} be a cover of U by open subsets of X . Then $\mathcal{U} \cup \{X \setminus U\}$ is an open cover of X , which has a finite subcover $(U_1, \dots, U_N, X \setminus U)$. Therefore U can be covered by $(U_i)_{1 \leq i \leq N}$ and U is compact.

(5) If the statement holds for L being a one-point set, then the same result holds for general L . For each $y \in L$ there exist disjoint open subsets $U_y, V_y \subset X$ such that $K \subset U_y$ and $y \in V_y$. By compactness of L , finitely many of these, say $(V_{y_i})_{1 \leq i \leq N}$, cover L . Then setting $U := \bigcap_{1 \leq i \leq N} U_{y_i}$ and $V := \bigcup_{1 \leq i \leq N} V_{y_i}$ proves the result.

We now prove the result for the case $L = \{y\}$. For each $x \in K$, there exists disjoint open sets U_x containing x and V_x containing y by the Hausdorff property. The collection $\{U_x : x \in K\}$ is an open cover of K , so it has a finite subcover, say $\{U_{x_1}, \dots, U_{x_N}\}$. Let $U := \bigcup_{1 \leq i \leq N} U_{x_i}$ and $V := \bigcap_{1 \leq i \leq N} V_{x_i}$. Then U and V are disjoint open sets with $K \subset U$ and $\{y\} \subset V$.

(4) For any $y \in X \setminus U$, by (5), we have $U \subset U$ and $y \in V$ for some disjoint open subsets U and V . Thus $X \setminus U$ is open.

(6) Without loss of generality, we may consider the product $X \times Y$ of compact spaces. Let \mathcal{U} be an open cover of $X \times Y$. Choose $x \in X$; the slice $\{x\} \times Y$ is homeomorphic to Y , so finitely many of the sets of \mathcal{U} cover it, say $(U_i)_{1 \leq i \leq N}$. Since product open sets are a basis for the product topology, for each $y \in Y$ there is a product open set $V \times W \subset X \times Y$ such that $(x, y) \in V \times W \subset \bigcup_{1 \leq i \leq N} U_i$. Finitely many of these product sets cover $\{x\} \times Y$, say $(V_i \times W_i)_{1 \leq i \leq M}$. If we set $Z_x := \bigcap_{1 \leq i \leq M} V_i$, then the whole strip $Z_x \times Y$ is contained in $\bigcup_{1 \leq i \leq N} U_i$. Because $(Z_x)_{x \in X}$ is an open cover of X , by compactness of X , we have a finite subcover, say $(Z_{x_i})_{1 \leq i \leq K}$, of X . Since finitely many sets of \mathcal{U} cover each strip $Z_{x_i} \times Y$, and finitely many such strips cover $X \times Y$.

(7) This follows from the fact that a quotient of a compact space is the image of a compact space by a continuous map.

(8) HW.

(9) $(c) \Rightarrow (b) \Rightarrow (a)$ are obvious. We now prove $(a) \Rightarrow (c)$. It suffices to show that every $x \in X$ has a neighborhood of precompact open sets. Being locally compact, X has a compact subset K containing a neighborhood U of x . Set

$$\mathcal{U}_x := \{\text{neighborhoods of } x \text{ contained in } U\}$$

so that \mathcal{U}_x is a neighborhood of x . Since X is Hausdorff and K is compact, by (4), K must be closed. If $V \in \mathcal{U}_x$, then $V \subset U \subset K$ implies $\overline{V} \subset \overline{K} = K$. Thus \overline{V} is compact by (3), and V is precompact.

(10) We need only to verify the local compactness. If U is open in X , then, **Theorem 1.5.12** yields that for any $x \in Y$ there exists a precompact neighborhood V_x of x such that $\overline{V_x} \subset U$. Thus U is locally compact.

Suppose that K is closed in X . By (9), any $x \in K$ has a precompact neighborhood U_x in X . Since $\overline{U_x \cap Z} = \overline{U_x} \cap Z$ is closed in $\overline{U_x}$ that is compact in X , it follows that $\overline{U_x \cap Z}$ is compact in X by (3). Hence $U_x \cap Z$ is a precompact neighborhood of x in Z , and Z is locally compact.

(11) For any closed set $K \subset X$, we shall show that $f(K)$ is closed in Y . Let $y \in \partial(f(K)) \subset Y$. There exists a precompact neighborhood U of y so that $y \in \partial(f(K) \cap U)$

\bar{U}). The properness of f implies that $f^{-1}(\bar{U})$ is compact in X and then $K \cap f^{-1}(\bar{U})$ is closed in $f^{-1}(\bar{U})$. Hence $K \cap f^{-1}(\bar{U})$ is compact in X . The continuity of f implies that $f(K) \cap \bar{U} = f(K \cap f^{-1}(\bar{U}))$ is compact in Y , and then $f(K \cap f^{-1}(\bar{U}))$ is closed in Y . Therefore $y \in f(K) \cap \bar{U} \subset f(K)$. Thus $f(K)$ is closed in Y . \square

Let (X, d) be a metric space and $S \subseteq X$. Define

$$(1.5.3.1) \quad \text{diam}(S) := \sup_{x, y \in S} d(x, y)$$

the **diameter** of S . If \mathcal{U} is an open cover of (X, d) , a positive number $\delta > 0$ is called a **Lebesgue number** for \mathcal{U} if any set whose diameter less than δ is contained in one of $U \in \mathcal{U}$.

Theorem 1.5.11. (Lebesgue number lemma) *Any open cover of a compact metric space has a Lebesgue number.*

PROOF. Let \mathcal{U} be an open cover of a compact metric space (X, d) . For any $x \in X$, there is an open subset $U \in \mathcal{U}$ with $x \in U$. The openness of U implies that $B_{2r(x)}(x) \subseteq U$ for some $r(x) > 0$. Since $\{B_{r(x)}(x)\}_{x \in X}$ forms an open cover of X , it follows that $B_{r(x_1)}(x_1), \dots, B_{r(x_n)}(x_n)$ cover X for some $n \in \mathbf{N}$. Define

$$\delta := \min_{1 \leq i \leq n} r(x_i) > 0.$$

If $S \subseteq X$ with $\text{diam}(S) < \delta$, we claim that $S \subseteq B_{2r(x_i)}(x_i)$ for some $i \in \{1, \dots, n\}$. Let $y \in S$ and choose $y \in B_{r(x_i)}(x_i)$ for some i . Therefore for any $z \in S$,

$$d(z, x_i) \leq d(z, y) + d(y, x_i) < \delta + d(y, x_i) < \delta + r(x_i) < 2r(x_i).$$

Thus δ is a Lebesgue number for \mathcal{U} . \square

Theorem 1.5.12. (Shrinking lemma) *Let X be a locally compact Hausdorff space and $x \in X$. If U is a neighborhood of x , then there exists a precompact neighborhood V of x such that $\bar{V} \subset U$.*

PROOF. By **Proposition 1.5.19** (9), there is a precompact neighborhood W of x so that $\bar{W} \setminus U$ is closed in \bar{W} . Then $\bar{W} \setminus U$ is compact by **Proposition 1.5.19** (3). Part (5) in **Proposition 1.5.19** implies that there exist open sets $Y, Y' \subseteq X$ such that $x \in Y$ and $\bar{W} \setminus U \subseteq Y'$. Let $V := Y \cap W$ be a neighborhood of x . Since $\bar{V} \subseteq \bar{W}$, it follows that \bar{V} is compact by (3). Because $\bar{V} \subseteq \bar{Y}$, we get $\bar{V} \subseteq \bar{W} \setminus Y' \subseteq U$. \square

Example 1.5.13. (1) \mathbf{R}^n is locally compact Hausdorff so that any open or closed subset of \mathbf{R}^n is locally compact Hausdorff.

(2) Any topological manifold is locally compact Hausdorff.

(3) \mathbf{B}^n and \mathbf{S}^{n-1} are compact.

(4) The set $A = \{(x, 1/x) \mid 0 < x \leq 1\}$ is closed in \mathbf{R}^2 , but is not compact.

(5) The set $A = \{(x, \sin(1/x)) \mid 0 < x \leq 1\}$ is bounded in \mathbf{R}^2 , but is not compact.

A point x of a topological space (X, \mathcal{T}) is said to be an **isolated point** of X if $\{x\}$ is open in X .

Theorem 1.5.14. *If (X, \mathcal{T}) is a nonempty compact Hausdorff space without any isolated points, then X is uncountable.*

PROOF. (1) Given any nonempty open subset $U \subseteq X$ and any $x \in X$, there exists a nonempty open subset $V \subseteq U$ such that $x \notin \overline{V}$. Choose $y \in U$ different from x (when $x \notin U$, we choose any point in U because $U \neq \emptyset$; when $x \in U$, we can also choose such a y because x is not an isolated point of X). Since X is Hausdorff, there exist disjoint open subsets W_1 and W_2 about x and y respectively. Then we can set $V := W_2 \cap U$.

(2) Given a map $f : \mathbf{Z}_+ \rightarrow X$. We show that f is not surjective (and hence X is uncountable). Let $x_n := f(n)$. From (1), we can find a nonempty open subset $V_1 \subseteq X$ such that $x_1 \notin \overline{V_1}$. In general, given $V_{n-1} \neq \emptyset$, there exists a nonempty open subset $V_n \subseteq V_{n-1}$ such that $x_n \notin \overline{V_n}$. Now we have a family $\{\overline{V_n}\}_{n \geq 1}$ of nonempty closed subsets of X with the property that $\overline{V_n} \supseteq \overline{V_{n+1}}$. Because X is compact, we can find $x \in \bigcap_{n \geq 1} \overline{V_n}$. Consequently, for any $n \geq 1$, $x \neq x_n$, so that f is not surjective. \square

Corollary 1.5.15. *Any nonempty closed interval in \mathbf{R} is uncountable.*

Exercise 1.5.16. Verify Corollary 1.5.15.

1.5.4. Paracompact spaces. Let (X, \mathcal{T}) be a topological space.

- (1) A collection $\mathcal{U} \subseteq 2^X$ is **locally finite** if any point $x \in X$ has a neighborhood that intersects at most finitely many of the sets in \mathcal{U} .
- (2) Given a cover \mathcal{U} of X , a cover \mathcal{V} is called a **refinement** of \mathcal{U} (written as $\mathcal{V} \prec \mathcal{U}$) if any $V \in \mathcal{V}$ is contained in some $U \in \mathcal{U}$. It is an **open refinement** if each $V \in \mathcal{V}$ is open in X .
- (3) X is **paracompact** if any open cover of X has a locally finite open refinement.

Example 1.5.17. (1) $\mathcal{U} = \{(n, n+2)\}_{n \in \mathbf{Z}}$ is locally finite in \mathbf{R} .

(2) $\mathcal{U} = \{(0, 1/n)\}_{n \geq 1}$ or $\{(1/(n+1), 1/n)\}_{n \geq 1}$ is locally finite in $(0, 1)$ but not in \mathbf{R} .

Proposition 1.5.18. (1) Let X be a topological space and $\mathcal{U} \subset 2^X$.

- (a) \mathcal{U} is locally finite if and only if $\overline{\mathcal{U}} := \{\overline{U} : U \in \mathcal{U}\}$ is locally finite.
- (b) If \mathcal{U} is locally finite, then

$$\overline{\bigcup_{U \in \mathcal{U}} U} = \bigcup_{U \in \mathcal{U}} \overline{U}.$$

(2) Any second countable, locally compact Hausdorff space X is paracompact. In fact, each open cover of X has a countable, locally finite refinement consisting of open subsets with compact closure (i.e., there exists a countable, locally finite, precompact refinement).

(3) Any paracompact Hausdorff space is normal.

(4) Any closed subspace of a paracompact space is paracompact.

PROOF. (1) (a) Suppose that \mathcal{U} is locally finite. For any $x \in X$, there exists a neighborhood W of x such that W intersects $U_1, \dots, U_m \in \mathcal{U}$ for some $m \in \mathbf{N}$. If $W \cap \overline{U} \neq \emptyset$, then $W \cap U \neq \emptyset$ so that $U \in \{U_1, \dots, U_m\}$. Thus W intersects \overline{U} for finitely many $\overline{U} \in \overline{\mathcal{U}}$. The converse is obvious.

(b) Observe that $\cup_{U \in \mathcal{U}} \overline{U} \subseteq \overline{\cup_{U \in \mathcal{U}} U}$. Let $x \notin \cup_{U \in \mathcal{U}} \overline{U}$. There exists some $U \in \mathcal{U}$ such that $x \in U$ and intersects $\overline{U_1}, \dots, \overline{U_k} \in \mathcal{U}$ by (1). Then $U \setminus \cup_{1 \leq i \leq k} \overline{U_k}$ contains x and intersects none of the sets in \mathcal{U} ; hence $x \notin \overline{\cup_{U \in \mathcal{U}} U}$.

(2) Let X be a second countable locally compact Hausdorff space. We claim that there exists a sequence $(G_i)_{i \in \mathbb{N}}$ of open sets such that $\overline{G_i}$ is compact, $\overline{G_i} \subseteq G_{i+1}$, and $X = \cup_{i \in \mathbb{N}} G_i$.

In fact, local compactness implies that X has a basis of precompact open sets by **Proposition 1.5.19**. The second countable assumption implies that there exists a countable basis $(U_i)_{i \in \mathbb{N}}$ of precompact open sets by **Theorem 1.2.34**. Let $G_1 := U_1$. Assume $G_k := U_1 \cup \dots \cup U_{j_k}$. Let j_{k+1} be the smallest positive integer greater than j_k such that $\overline{G_k} \subseteq \cup_{1 \leq i \leq j_{k+1}} U_i$. Define $G_{k+1} := \cup_{1 \leq i \leq j_{k+1}} U_i$. Then $(G_i)_{i \in \mathbb{N}}$ is countable, $\overline{G_i}$ is compact, $\overline{G_i} \subseteq G_{i+1}$, and

$$X = \bigcup_{i \in \mathbb{N}} U_i \subseteq \bigcup_{i \in \mathbb{N}} G_i \subseteq \bigcup_{i \in \mathbb{N}} \overline{G_i} \subseteq X.$$

Let $\mathcal{U} = (U_\alpha)_{\alpha \in I}$ be an arbitrary open cover of X . Since $\overline{G_i} \setminus G_{i-1} \subseteq G_{i+1} \setminus \overline{G_{i-2}}$ ($i \geq 3$) is compact, it follows that we can choose a finite subcover of the open cover $(U_\alpha \cap (G_{i+1} \setminus \overline{G_{i-2}}))_{\alpha \in I}$ of $\overline{G_i} \setminus G_{i-1}$ and choose a finite subcover of the open cover $(U_\alpha \cap G_3)_{\alpha \in I}$ of G_2 . Indeed, let $W_i := \overline{G_i} \setminus G_{i-1}$; the local compactness of X implies that there exists a finite subcover $(U_{ij})_{1 \leq j \leq m_i}$ of the open cover $(U_\alpha \cap (G_{i+1} \setminus \overline{G_{i-2}}))_{\alpha \in I}$ of W_i . Set $V_{ij} := U_{ij} \cap (G_{i+1} \setminus \overline{G_{i-2}})$ for $1 \leq j \leq m_i$. Similarly, we can define V_{ij} for $i = 1, 2$. Hence $(V_{ij})_{1 \leq j \leq m_i}$ is an open subcover of W_i and

$$\bigcup_{i \in \mathbb{N}} \bigcup_{1 \leq j \leq m_i} V_{ij} = \bigcup_{i \in \mathbb{N}} W_i = X.$$

If $\mathcal{V} := (V_{ij})_{i \in \mathbb{N}, 1 \leq j \leq m_i}$, then $\mathcal{V} \prec \mathcal{U}$ and $W_i \cap \cup_{1 \leq j \leq m_i} V_{kj} = \emptyset$ for $k \neq i-2, i-1, i, i+1, i+1$. Thus \mathcal{V} is locally finite.

(3) We first prove that X is regular. Let $a \in X$ and B be closed subset of X disjoint from a . The Hausdorff property implies that for any $b \in B$ there is an open subset $U_b \ni b$ such that $\overline{U_b} \cap \{a\} = \emptyset$. Cover X by $\{U_b\}_{b \in B}$ and $X \setminus B$, and take a locally finite open refinement \mathcal{C} that covers X . Set

$$\mathcal{D} := \{C \in \mathcal{C} \mid C \cap B \neq \emptyset\}.$$

Then \mathcal{D} covers B . For any $D \in \mathcal{D}$, $\overline{D} \cap \{a\} = \emptyset$. Let

$$V := \bigcup_{D \in \mathcal{D}} D$$

that is open in X and contains B . Since \mathcal{D} is locally finite, it follows that $\overline{V} = \cup_{D \in \mathcal{D}} \overline{D}$. Hence $\overline{V} \cap \{a\} = \emptyset$ so that X is regular (because V and $X \setminus \overline{V}$ are disjoint open subsets of B and a respectively).

Let A and B be two disjoint closed subsets of X . For any $a \in A$, there exist open subsets U_a of a and V_a containing B , respectively, such that $\overline{U_a} \cap V_a = \emptyset$. As above, we can find two disjoint open subsets V and $X \setminus V$ of B and A respectively.

(4) Let Y be a closed subspace of the paracompact space X and let \mathcal{U} be an open cover of Y . For any $U \in \mathcal{U}$, there exists an open subset U' of X such that $U' \cap Y = U$. Cover X by the open subsets U' , together with $X \setminus Y$. We then can take a locally finite open refinement \mathcal{V} of \mathcal{U} that still covers X . The collection $\mathcal{C} := \{V \cap Y \mid V \in \mathcal{V}\}$ is a locally finite open refinement of \mathcal{U} . \square

Remark 1.5.19. (1) A paracompact subspace of a Hausdorff space X need not be closed in X . For example, $(0, 1)$ is paracompact, but it is not closed in \mathbf{R} .

(2) **(A. H. Stone)** Every metric space is paracompact.

(3) Every regular Lindelöf space is paracompact.

(4) The product of two paracompact spaces need not be paracompact. For example, \mathbf{R}_ℓ is regular and Lindelöf, it follows from (3) that \mathbf{R}_ℓ is paracompact, but \mathbf{R}_ℓ^2 , that is Hausdorff and not normal, is not paracompact.

(5) \mathbf{R}^ω is paracompact in the product topology.

(6) \mathbf{R}^J is not paracompact if J is uncountable.

Let $\mathcal{U} = \{U_\alpha\}_{\alpha \in J}$ be an index open cover of a topological space X . An indexed family of continuous functions $\phi_\alpha : X \rightarrow [0, 1]$ is a **partition of unity on X dominated by \mathcal{U}** , if

- (1) $\text{supp}(\phi_\alpha) \subseteq U_\alpha$ for all $\alpha \in J$,
- (2) $\{\text{supp}(\phi_\alpha)\}_{\alpha \in J}$ is locally finite, and
- (3) $\sum_{\alpha \in J} \phi_\alpha(x) \equiv 1$ for all $x \in X$.

Lemma 1.5.20. (Shrinking lemma) Suppose that (X, \mathcal{T}) is a paracompact Hausdorff space. If $\mathcal{U} = \{U_\alpha\}_{\alpha \in J}$ is an indexed family of open cover of X , then \mathcal{U} admits a locally finite open refinement $\mathcal{V} = \{V_\alpha\}_{\alpha \in J}$ indexed by the same index set J , such that $\overline{V_\alpha} \subseteq U_\alpha$ for each $\alpha \in J$.

PROOF. Let $\mathcal{A} := \{A \in \mathcal{T} \mid \overline{A} \subseteq U_\alpha \text{ for some } \alpha \in J\}$. According to **Proposition 1.5.18**, X is normal so that for any $x \in X$ we can find a open subset A_x such that $\overline{A_x} \subseteq U_\alpha$ for some $\alpha \in J$. Thus \mathcal{A} covers X . The paracompactness implies that \mathcal{A} has a locally finite open refinement $\mathcal{B} = \{B_\beta\}_{\beta \in K}$. We then find a map $f : K \rightarrow J$ such that for any $\beta \in K$ one has $\overline{B_\beta} \subseteq \overline{A} \subseteq U_{f(\beta)}$. Define, for each $\alpha \in J$,

$$V_\alpha := \begin{cases} \bigcup_{B_\beta \in \mathcal{B}_\alpha} B_\beta, & \mathcal{B}_\alpha := \{B_\beta \in \mathcal{B} \mid f(\beta) = \alpha\} \neq \emptyset, \\ \emptyset, & \mathcal{B}_\alpha = \emptyset. \end{cases}$$

For $B_\beta \in \mathcal{B}_\alpha$, we have $\overline{B_\beta} \subseteq U_{f(\beta)} = U_\alpha$. Since \mathcal{B}_α is locally finite, by **Proposition 1.5.18**,

$$\overline{V_\alpha} = \bigcup_{B_\beta \in \mathcal{B}_\alpha} \overline{B_\beta} \subseteq U_\alpha.$$

Finally we check the local finiteness of $\mathcal{V} = \{V_\alpha\}_{\alpha \in J}$. Given $x \in X$. There is a neighborhood W of x that intersects B_β for only finitely many values of $\beta = \beta_1, \dots, \beta_N$. Then W can intersect V_α only if $\alpha \in \{f(\beta_1), \dots, f(\beta_N)\}$. \square

Theorem 1.5.21. Let (X, \mathcal{T}) be a paracompact Hausdorff space. If $\mathcal{U} = \{U_\alpha\}_{\alpha \in J}$ is an indexed open cover of X , then there is a partition of unity on X dominated by \mathcal{U} .

PROOF. By **Lemma 1.5.20**, we can find two locally finite indexed families of open sets $\mathcal{W} = \{W_\alpha\}_{\alpha \in J}$ and $\mathcal{V} = \{V_\alpha\}_{\alpha \in J}$ covering X such that

$$\overline{W_\alpha} \subseteq V_\alpha, \quad \overline{V_\alpha} \subseteq U_\alpha.$$

The normality, coming from **Proposition 1.5.18**, implies that there is a continuous function $\psi_\alpha : X \rightarrow [0, 1]$ such that

$$\psi_\alpha(\overline{W_\alpha}) = 1, \quad \psi_\alpha(X \setminus V_\alpha) = 0,$$

by **Theorem 1.2.30**. Hence $\text{supp}(\psi_\alpha) \subseteq \overline{V_\alpha} \subseteq U_\alpha$. Because \mathcal{W} covers X , for any $x \in X$ we can find some ψ_α such that $\psi_\alpha(x) > 0$. Define

$$\Psi(x) := \sum_{\alpha \in J} \psi_\alpha(x) \geq 0$$

which is strictly positive for each $x \in X$. Since for each $x \in X$ there is a neighborhood W_x of x that intersects $\text{supp}(\psi_\alpha)$ for only finitely many values of α , it follows that the above infinite sum at x is actually finite. Hence Ψ is well-defined and continuous on W_x and then on X . Finally set $\phi_\alpha(x) := \psi_\alpha(x)/\Psi(x)$. \square

Analysis on topological groups

2.1. Haar measures

From now on let \mathcal{G} be a **local compact group**, which in this note means a locally compact and Hausdorff topological group.

A collection \mathcal{S} of subsets of \mathcal{G} is called a **σ -algebra** if

- (1) $\emptyset, X \in \mathcal{S}$,
- (2) $A \in \mathcal{S}$ implies $X \setminus A \in \mathcal{S}$, and
- (3) $\{A_n\}_{n \geq 1} \subseteq \mathcal{S}$ implies $\bigcup_{n \geq 1} A_n \in \mathcal{S}$.

Clearly that the topology \mathcal{T} for X is a σ -algebra. Define

$$\mathcal{B}_{\mathcal{G}} := \bigcap \{ \mathcal{S} \mid \mathcal{S} \text{ is a } \sigma\text{-algebra of } X \text{ and contains } \mathcal{T} \}.$$

Exercise 2.1.1. Show that $\mathcal{B}_{\mathcal{G}}$ is also a σ -algebra and is the smallest σ -algebra containing the open subsets of \mathcal{G} .

Elements in $\mathcal{B}_{\mathcal{G}}$ are called **Borel subsets in \mathcal{G}** .

For any locally compact group \mathcal{G} , Haar's theorem gives, up to a positive multiplicative constant, a unique countably additive, nontrivial measure μ on the Borel sets of \mathcal{G} satisfying the following properties:

- (1) **the measure μ is left-invariant:** $\mu(g\mathcal{S}) = \mu(\mathcal{S})$ for every $g \in \mathcal{G}$ and all Borel sets $\mathcal{S} \subseteq \mathcal{G}$;
- (2) **the measure μ is finite on every compact set:** $\mu(\mathcal{K}) < \infty$ for all compact $\mathcal{K} \subseteq \mathcal{G}$;
- (3) **the measure μ is outer regular on Borel sets $\mathcal{S} \subseteq \mathcal{G}$:** $\mu(\mathcal{S}) = \inf\{\mu(\mathcal{U}) : \mathcal{S} \subseteq \mathcal{U}, \mathcal{U} \text{ open}\}$;
- (4) **the measure μ is inner regular on open sets $\mathcal{U} \subseteq \mathcal{G}$:** $\mu(\mathcal{U}) = \sup\{\mu(\mathcal{K}) : \mathcal{K} \subseteq \mathcal{U}, \mathcal{K} \text{ compact}\}$.

Such a measure on \mathcal{G} is called a **(left) Haar measure**. If μ is a Haar measure and $g_1 \in \mathcal{G}$, then $\mathcal{B}_{\mathcal{G}} \ni \mathcal{S} \mapsto \mu(\mathcal{S}g_1)$ is left-invariant. Thus, by uniqueness of the Haar measure, there exists a function Δ from the group \mathcal{G} to \mathbb{R}^+ , called the **modular function**, such that

$$\mu(\mathcal{S}g_1) = \Delta(g_1)\mu(\mathcal{S})$$

for every Borel set $\mathcal{S} \subseteq \mathcal{G}$.

We say \mathcal{G} is **unimodular** if the modular function Δ is identically 1, or equivalently, if the Haar measure is both left and right-invariant. Examples of unimodular groups are Abelian groups, compact groups, discrete groups, semisimple Lie groups and connected nilpotent Lie groups.

An example of a non-unimodular group is the group of affine transformations

$$\{x \mapsto ax + b : a \in \mathbf{R}^*, b \in \mathbf{R}\} = \left\{ \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \mid a \in \mathbf{R}^*, b \in \mathbf{R} \right\}$$

on the real line. This example shows that a solvable Lie group need not be unimodular. In this group a left Haar measure is given by $da \wedge db/a^2$, and a right Haar measure by $da \wedge db/|a|$.

Let \mathcal{G} be a locally compact group with Haar measure μ . For $0 < p < +\infty$, we define $\widehat{L}^p(\mathcal{G}, \mu)$ the set of all complex-valued μ -measurable functions on \mathcal{G} whose modulus to the p -th power is integrable. $\widehat{L}^\infty(\mathcal{G}, \mu)$ is the set of all complex-valued μ -measurable functions f on \mathcal{G} such that for some $M > 0$ the set $\{x \in \mathcal{G} \mid |f(x)| > M\}$ has μ -measure zero. We say two functions in $L^p(\mathcal{G}, \mu)$ are considered equal (or equivalent) if they are equal μ -almost everywhere. The equivalence class of $\widehat{L}^p(\mathcal{G}, \mu)$ is denoted $L^p(\mathcal{G}, \mu)$. Similarly we can define $L^\infty(\mathcal{G}, \mu)$.

We simply write $L^p(\mathcal{G})$ for $L^p(\mathcal{G}, \mu)$, when $p > 0$ or $p = \infty$. For $0 < p < +\infty$ define

$$\|f\|_{L^p(\mathcal{G})} := \left(\int_{\mathcal{G}} |f(x)|^p d\mu(x) \right)^{1/p}$$

and for $p = \infty$ define

$$\|f\|_{L^\infty(\mathcal{G})} := \inf \{M > 0 \mid \mu(\{x \in \mathcal{G} \mid |f(x)| > M\}) = 0\}.$$

The left invariance of μ is equivalent to the fact that for all $t \in \mathcal{G}$ and all nonnegative measurable functions f on \mathcal{G} we have

$$\int_{\mathcal{G}} f(tx) d\mu(x) = \int_{\mathcal{G}} f(x) d\mu(x).$$

2.1.1. Examples of Haar measures. Let $\mathcal{G} := \mathbf{R}^* = \mathbf{R} \setminus \{0\}$ with group law the usual multiplication. Then \mathcal{G} is locally compact and the measure

$$\mu := \frac{dx}{|x|}$$

is invariant under multiplicative translation, because of the following identity

$$\int_{-\infty}^{+\infty} f(tx) \frac{dx}{|x|} = \int_{-\infty}^{+\infty} f(x) \frac{dx}{|x|}$$

for all $f \in L^1(\mathbf{R}^*, \mu)$ and all $t \in \mathbf{R}^*$. Then

$$\mu(A) = \int_A \frac{dx}{x} = \int_{-\infty}^{+\infty} \chi_A(x) \frac{dx}{x}.$$

Example 2.1.2. (1) A Haar measure on the multiplicative group \mathbf{R}^+ is dx/x . That is

$$\mu(A) = \int_{-\infty}^{+\infty} \chi_A(x) \frac{dx}{x}.$$

(2) The Heisenberg group \mathfrak{H}^n is the set $\mathbf{C}^n \times \mathbf{R}$ with the group operation

$$(z, t)(w, s) := \left(z + w, t + s + 2\operatorname{Im} \sum_{1 \leq i \leq n} z^i \overline{w^i} \right).$$

The identity in \mathfrak{H}^n is $(\mathbf{0}, 0) \in \mathbf{C}^n \times \mathbf{R}$ and $(z, t)^{-1} = (-z, -t)$. \mathfrak{H}^n is topologically identified with $\mathbf{C}^n \times \mathbf{R}$ and hence is a locally compact group. Both the left and right Haar measure on \mathfrak{H}^n is Lebesgue measure $dz \wedge dt$.

(3) Let $\mathcal{G} := \mathbf{R}^2 \setminus \{(0, y) | y \in \mathbf{R}\}$ with group operation

$$(x, y)(z, w) := (xz, xw + y) \iff \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z & w \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} xz & xw + y \\ 0 & 1 \end{bmatrix}$$

For $A \subseteq \mathcal{B}_{\mathcal{G}}$ define

$$\mu(A) := \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi_A(x, y) \frac{dx dy}{x^2}.$$

When $\mathbf{t} = (t_1, t_2) \in \mathcal{G}$ and $(x, y) \in A$ we have

$$\mathbf{t}(x, y) = (t_1 x, t_1 y + t_2)$$

and then

$$\begin{aligned} \mu(\mathbf{t}A) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi_{\mathbf{t}A}(x, y) \frac{dx dy}{x^2} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi_A(x', y') \frac{d(x'/t) d((y' - t_2)/t_1)}{(x'/t_1)^2} \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi_A(x', y') \frac{dx' dy'}{x'^2} = \mu(A). \end{aligned}$$

Thus μ is a left Haar measure on \mathcal{G} . Similarly we can prove that

$$\nu(A) := \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi_A(x, y) \frac{dx dy}{|x|}$$

is a right Haar measure on \mathcal{G} .

2.1.2. Convolution. Let $f, g \in L^1(\mathcal{G})$. Define the **convolution** $f * g$ by

$$(2.1.2.1) \quad (f * g)(x) := \int_{\mathcal{G}} f(y) g(y^{-1}x) d\mu(y).$$

Since

$$\begin{aligned} \left| \int_{\mathcal{G}} \int_{\mathcal{G}} f(y) g(y^{-1}x) d\mu(y) d\mu(x) \right| &\leq \int_{\mathcal{G}} \int_{\mathcal{G}} |f(y)| |g(y^{-1}x)| d\mu(x) d\mu(y) \\ &= \int_{\mathcal{G}} |f(y)| \int_{\mathcal{G}} |g(y^{-1}x)| d\mu(x) d\mu(y) = \int_{\mathcal{G}} |f(y)| \int_{\mathcal{G}} |g(x)| d\mu(x) d\mu(y) \\ &= \|f\|_{L^1(\mathcal{G})} \|g\|_{L^1(\mathcal{G})} < +\infty, \end{aligned}$$

it follows that $f * g$ is well-defined and

$$(2.1.2.2) \quad \|f * g\|_{L^1(\mathcal{G})} \leq \|f\|_{L^1(\mathcal{G})} \|g\|_{L^1(\mathcal{G})}.$$

Letting $z = x^{-1}y$ in (2.1.2.1) yields

$$(2.1.2.3) \quad (f * g)(x) = \int_{\mathcal{G}} f(xz) g(z^{-1}) d\mu(z).$$

Exercise 2.1.3. (1) For all $f, g, h \in L^1(\mathcal{G})$, we have

$$f * (g * h) = (f * g) * h$$

and

$$f * (g + h) = f * g + f * h, \quad (f + g) * h = f * h + g * h.$$

(2) Verify (2.1.2.4).

For $f, g \in L^1(\mathcal{G})$ and $t \in \mathcal{G}$, define

$${}^t f(x) := f(tx), \quad f^t(x) := f(xt).$$

Then we have

$$(2.1.2.4) \quad {}^t f * g = {}^t(f * g), \quad f * g^t = (f * g)^t.$$

Theorem 2.1.4. (Minkowski's inequality) *Let $1 \leq p \leq +\infty$. For $f \in L^p(\mathcal{G})$ and $g \in L^1(\mathcal{G})$ we have $g * f$ exists μ -a.e. and satisfies*

$$(2.1.2.5) \quad \|g * f\|_{L^p(\mathcal{G})} \leq \|g\|_{L^1(\mathcal{G})} \|f\|_{L^p(\mathcal{G})}.$$

PROOF. The case $p = 1$ was proved in (2.1.2.2), and the case $p = \infty$ is obvious. Now we assume $1 < p < +\infty$. Write $p' = \frac{p}{p-1}$, and compute

$$\begin{aligned} (|g| * |f|)(x) &= \int_{\mathcal{G}} |f(y^{-1}x)| |g(y)| d\mu(x) = \int_{\mathcal{G}} |f(y^{-1}x)| |g(y)|^{\frac{1}{p}} |g(y)|^{\frac{1}{p'}} d\mu(x) \\ &\leq \left(\int_{\mathcal{G}} |f(y^{-1}x)|^p |g(y)| d\mu(y) \right)^{1/p} \left(\int_{\mathcal{G}} |g(y)| d\mu(y) \right)^{1/p'} \end{aligned}$$

so that

$$\begin{aligned} \| |g| * |f| \|_{L^p(\mathcal{G})} &\leq \left(\|g\|_{L^1(\mathcal{G})}^{p-1} \int_{\mathcal{G}} \int_{\mathcal{G}} |f(y^{-1}x)|^p |g(y)| d\mu(y) d\mu(x) \right)^{1/p} \\ &= \left(\|g\|_{L^1(\mathcal{G})}^{p-1} \int_{\mathcal{G}} |f(x)|^p d\mu(x) \int_{\mathcal{G}} |g(y)| d\mu(y) \right)^{1/p} = \|f\|_{L^p(\mathcal{G})} \|g\|_{L^1(\mathcal{G})}. \end{aligned}$$

Now the inequality (2.1.2.5) follows from $|g * f| \leq |g| * |f|$. \square

If f is a function on \mathcal{G} , we define

$$(2.1.2.6) \quad \tilde{f}(x) := f(x^{-1}), \quad x \in \mathcal{G}.$$

Then (2.1.2.3) can be written as

$$(f * g)(x) = \int_{\mathcal{G}} f(xz) \tilde{g}(z) d\mu(z).$$

Theorem 2.1.5. (Young's inequality) *Let $1 \leq p, q, r \leq +\infty$ satisfy*

$$\frac{1}{q} + 1 = \frac{1}{p} + \frac{1}{r}.$$

*Then for all $f \in L^p(\mathcal{G})$ and all $g \in L^r(\mathcal{G})$ satisfying $\|g\|_{L^r(\mathcal{G})} = \|\tilde{g}\|_{L^r(\mathcal{G})}$ we have $f * g$ exists μ -a.e. and satisfies*

$$(2.1.2.7) \quad \|f * g\|_{L^q(\mathcal{G})} \leq \|g\|_{L^r(\mathcal{G})} \|f\|_{L^p(\mathcal{G})}.$$

PROOF. If $r = +\infty$, then $p = 1$ and $q = +\infty$ in which case the inequality (2.1.2.7) is trivial. Now we assume that $r < +\infty$. From $\frac{1}{q} + 1 = \frac{1}{p} + \frac{1}{r}$ we have

$$\frac{1}{r'} + \frac{1}{p'} + \frac{1}{q} = 1, \quad \frac{1}{r} + \frac{1}{r'} = 1 = \frac{1}{p} + \frac{1}{p'}.$$

By Hölder's inequality we get

$$|(f * g)(x)| \leq \int_{\mathcal{G}} |f(y)| |g(y^{-1}x)| d\mu(x)$$

$$\begin{aligned}
&= \int_{\mathcal{G}} |f(y)|^{p/r'} \left(|f(y)|^{p/q} |g(y^{-1}x)|^{r/q} \right) |g(y^{-1}x)|^{r/p'} d\mu(y) \\
&\leq \|f\|_{L^p(\mathcal{G})}^{p/r'} \left(\int_{\mathcal{G}} |f(y)|^p |g(y^{-1}x)|^r d\mu(y) \right)^{1/q} \left(\int_{\mathcal{G}} |g(y^{-1}x)|^r d\mu(y) \right)^{1/p'} \\
&= \|f\|_{L^p(\mathcal{G})}^{p/r'} \left(\int_{\mathcal{G}} |f(y)|^p |g(y^{-1}x)|^r d\mu(y) \right)^{1/q} \left(\int_{\mathcal{G}} |\tilde{g}(x^{-1}y)|^r d\mu(y) \right)^{1/p'} \\
&= \left(\int_{\mathcal{G}} |f(y)|^p |g(y^{-1}x)|^r d\mu(y) \right)^{1/q} \|f\|_{L^p(\mathcal{G})}^{p/r'} \|\tilde{g}\|_{L^r(\mathcal{G})}^{r/p'}
\end{aligned}$$

so that

$$\begin{aligned}
\| |f| * |g| \|_{L^q(\mathcal{G})} &\leq \|f\|_{L^p(\mathcal{G})}^{p/r'} \|\tilde{g}\|_{L^r(\mathcal{G})}^{r/p'} \left(\int_{\mathcal{G}} \int_{\mathcal{G}} |f(y)|^p |g(y^{-1}x)|^r d\mu(x) d\mu(y) \right)^{1/q} \\
&= \|f\|_{L^p(\mathcal{G})}^{p/r'} \|\tilde{g}\|_{L^r(\mathcal{G})}^{r/p'} \|f\|_{L^p(\mathcal{G})}^{p/q} \|g\|_{L^r(\mathcal{G})}^{r/q} = \|g\|_{L^r(\mathcal{G})} \|f\|_{L^p(\mathcal{G})}
\end{aligned}$$

using again the assumption $\|g\|_{L^r(\mathcal{G})} = \|\tilde{g}\|_{L^r(\mathcal{G})}$. \square

Exercise 2.1.6. (Hardy) Use **Theorem 2.1.4** to prove the following inequalities.

(1) For $1 < p < +\infty$,

$$\left(\int_0^{+\infty} \left(\frac{1}{x} \int_0^x |f(t)| dt \right)^p dx \right)^{1/p} \leq \frac{p}{p-1} \|f\|_{L^p((0,+\infty))}$$

and

$$\left(\int_0^{+\infty} \left(\int_x^{+\infty} |f(t)| dt \right)^p dx \right)^{1/p} \leq p \left(\int_0^{+\infty} |f(t)|^p t^p dt \right)^{1/p}.$$

(2) For $0 < b < +\infty$ and $1 \leq p < +\infty$,

$$\left(\int_0^{+\infty} \left(\int_0^x |f(t)| dt \right)^p x^{-b-1} dx \right)^{1/p} \leq \frac{p}{b} \left(\int_0^{+\infty} |f(t)|^p t^{p-b-1} dt \right)^{1/p}$$

and

$$\left(\int_0^{+\infty} \left(\int_x^{+\infty} |f(t)| dt \right)^p x^{b-1} dx \right)^{1/p} \leq \frac{p}{b} \left(\int_0^{+\infty} |f(t)|^p t^{p+b-1} dt \right)^{1/p}.$$

2.1.3. Modular functions on \mathcal{G} . Let \mathcal{G} be a locally compact group with Haar measure $\mu := dg$. Recall that the modular function $\Delta \equiv \Delta_{\mathcal{G}}$ on \mathcal{G} is the function (actually continuous homomorphism into \mathbb{R}^+) such that for all $f \in C_c(\mathcal{G})$,

$$(2.1.3.1) \quad \int_{\mathcal{G}} f(gg_1) dg = \Delta(g_1) \int_{\mathcal{G}} f(g) dg.$$

Proposition 2.1.7. Let \mathcal{P} be a locally compact group with two closed subgroups \mathcal{A} , \mathcal{U} such that \mathcal{A} normalizes \mathcal{U} , and such that the product

$$(2.1.3.2) \quad \mathcal{A} \times \mathcal{U} \longrightarrow \mathcal{A}\mathcal{U} = \mathcal{P}$$

is a topological isomorphism. Then for $f \in C_c(\mathcal{P})$ the functional

$$(2.1.3.3) \quad f \longmapsto \int_{\mathcal{U}} \int_{\mathcal{A}} f(au) da du = \int_{\mathcal{A}} \int_{\mathcal{U}} f(au) du da$$

is a Haar (left-invariant) functional on \mathcal{P} .

PROOF. Left invariant by \mathcal{A} is obvious. Let $u_1 \in \mathcal{U}$. Then

$$\begin{aligned} \int_{\mathcal{A}} \int_{\mathcal{U}} f(u_1 a u) d u d a &= \int_{\mathcal{A}} \int_{\mathcal{U}} f(a a^{-1} u_1 a u) d u d a \\ &= \int_{\mathcal{A}} \int_{\mathcal{U}} (f \circ a)(u_1^a u) d u d a \quad \text{where } u_1^a := a^{-1} u_1 a \in \mathcal{U} \\ &= \int_{\mathcal{A}} \int_{\mathcal{U}} f(a u) d u d a \end{aligned}$$

by the Haar measure property. \square

Proposition 2.1.8. *Notation is as in Proposition 2.1.7. There exists a unique continuous homomorphism $\delta : \mathcal{A} \rightarrow \mathbb{R}^+$ such that for $f \in C_c(\mathcal{U})$,*

$$(2.1.3.4) \quad \int_{\mathcal{U}} f(a^{-1} u a) d u = \delta(a) \int_{\mathcal{U}} f(u) d u,$$

or in other words, for $f \in C_c(\mathcal{P})$,

$$(2.1.3.5) \quad \int_{\mathcal{U}} f(u a) d u = \delta(a) \int_{\mathcal{U}} f(a u) d u.$$

If \mathcal{U} is unimodular, then δ is the modular function on \mathcal{P} , that is,

$$(2.1.3.6) \quad \Delta_{\mathcal{P}}(p) = \Delta_{\mathcal{P}}(a u) = \delta(a).$$

PROOF. For the first statement, we note that the map $u \mapsto u^a = a^{-1} u a$ is a topological group automorphism of \mathcal{U} , which preserves Haar measure up to a constant factor. For the second statement, we first note that the functional

$$f \mapsto \int_{\mathcal{U}} \int_{\mathcal{A}} f(a u) d a d u$$

is right-invariant under \mathcal{U} . Let $a_1 \in \mathcal{A}$. Then

$$\begin{aligned} \int_{\mathcal{U}} \int_{\mathcal{A}} f(a u a_1) d a d u &= \int_{\mathcal{U}} \int_{\mathcal{A}} f(a a_1 a^{-1} u a_1) d a d u = \int_{\mathcal{U}} \int_{\mathcal{A}} f(a a_1 u) d a d u \\ &= \int_{\mathcal{U}} \int_{\mathcal{A}} f(a a_1 a^{-1} a u) d a d u = \delta(a_1) \int_{\mathcal{U}} \int_{\mathcal{A}} f(a u) d a d u, \end{aligned}$$

which proves the formula (2.1.3.6). \square

Proposition 2.1.9. *Let \mathcal{G} be a locally compact group with two closed subgroups \mathcal{P}, \mathcal{K} such that*

$$(2.1.3.7) \quad \mathcal{P} \times \mathcal{K} \longrightarrow \mathcal{P}\mathcal{K} = \mathcal{G}$$

is a topological isomorphism (not group isomorphism). Assume that \mathcal{G}, \mathcal{K} are unimodular. Let $d g, d p, d k$ be given Haar measures on $\mathcal{G}, \mathcal{P}, \mathcal{K}$ respectively. Then there is a constant c such that for all $f \in C_c(\mathcal{G})$,

$$(2.1.3.8) \quad \int_{\mathcal{G}} f(g) d g = c \int_{\mathcal{P}} \int_{\mathcal{K}} f(p k) d p d k.$$

If in addition $\mathcal{P} = \mathcal{A}\mathcal{U}$ as in Proposition 2.1.7, with \mathcal{U} unimodular, so we have the product decomposition

$$(2.1.3.9) \quad \mathcal{U} \times \mathcal{A} \times \mathcal{K} \longrightarrow \mathcal{G},$$

then

$$(2.1.3.10) \quad \int_{\mathcal{G}} f(g) dg = \int_{\mathcal{U}} \int_{\mathcal{A}} \int_{\mathcal{K}} f(uak) \delta(a)^{-1} dudak.$$

PROOF. The first assertion comes from a standard fact of homogeneous spaces, that is, there exists a left-invariant Haar measure on $\mathcal{G}/\mathcal{K} = \mathcal{P}$. The second integral formula comes from **Proposition 2.1.8**. \square

2.2. Iwasawa's decomposition

The **Iwasawa decomposition** for $\mathcal{G} := \mathbf{SL}_2(\mathbf{C}) = \mathbf{SL}(2, \mathbf{C})$ is $\mathcal{G} = \mathcal{U}\mathcal{A}\mathcal{K}$, where

- \mathcal{U} is the subgroup of upper triangular unipotent matrices;
- \mathcal{A} is the subgroup of diagonal matrices with positive diagonal components;
- \mathcal{K} is the unitary subgroup.

2.2.1. Iwasawa's decomposition. Thus \mathcal{U} is the group of 2×2 matrices of the form

$$(2.2.1.1) \quad u(x) := \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \quad \text{with } x \in \mathbf{C}.$$

The elements of \mathcal{A} are of the form

$$(2.2.1.2) \quad a = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \quad \text{with } a_1, a_2 > 0 \text{ and } a_1 a_2 = 1.$$

We also define the diagonal group \mathbf{D} consisting of all elements

$$(2.2.1.3) \quad z = \begin{bmatrix} z_1 & 0 \\ 0 & z_1^{-1} \end{bmatrix} \quad \text{with } 0 \neq z_1 \in \mathbf{C}.$$

The elements k of \mathcal{K} are the matrices satisfying (besides determinant 1)

$$(2.2.1.4) \quad k^{-1} = k^*, \quad \text{where } M^* := {}^t \overline{M} \text{ for all matrix } M.$$

Theorem 2.2.1. (Iwasawa's decomposition) *Every element $g \in \mathcal{G}$ has a unique expression as a product*

$$(2.2.1.5) \quad g = uak \quad \text{with } u \in \mathcal{U}, a \in \mathcal{A}, k \in \mathcal{K}.$$

Furthermore, let $\text{Pos}_2 := \text{Pos}_2(\mathbf{C})$ be the space of positive definite 2×2 Hermitian matrices. Then the map $g \mapsto gg^$ is a bijection $\mathcal{G}/\mathcal{K} \approx \text{SPos}_2$, so of $\mathcal{U}\mathcal{A} \approx \text{SPos}_2$.*

PROOF. Let $g \in \mathcal{G}$ and e_1, e_2 be the standard (vertical) unit vectors of \mathbf{C}^2 . Let $g^{(i)} := ge_i$. We orthonormalize $g^{(1)}, g^{(2)}$ by the standard Gram-Schmidt process, which means we can find an upper triangular matrix $B = (b_{ij})$ ($b_{21} = 0$) such that if we let

$$e'_1 := b_{11}g^{(1)}, \quad e'_2 := b_{12}g^{(1)} + b_{22}g^{(2)},$$

then e'_1, e'_2 are orthonormal unit vectors, and we can choose both b_{11} and b_{22} greater than 0 (This corresponds to dividing by the norm). Thus $B \in \mathcal{AU} = \mathcal{UA}$.

Let k be the matrix such that $ke_i = e'_i$ for $i = 1, 2$. Then k is unitary and $gB = k$. Since $\det g = 1$ and the diagonal components of B are positive, it follows that

$\det k = 1$, so $\mathcal{G} = \mathcal{K}\mathcal{A}\mathcal{U} = \mathcal{U}\mathcal{A}\mathcal{K}$. To show uniqueness, we use the map $g \mapsto gg^*$. Note that $gg^* = \mathbf{1}$ (identity) if and only if $g \in \mathcal{K}$. Suppose that

$$u_1 a k_1 (u_1 a k_1)^* = u_1 a^2 u_1^* = u_2 b^2 u_2^* = u_2 b k_2 (u_2 b k_2)^*$$

with $u_1, u_2 \in \mathcal{U}$, $a, b \in \mathcal{A}$, and $k_1, k_2 \in \mathcal{K}$. Putting $u := u_2^{-1}u_1$, we obtain $ua^2 = b^2(u^*)^{-1}$. Since u, u^* are triangular in opposite directions, they must be diagonal, and finally $a^2 = b^2$, so $a = b$ because the diagonal elements are positive. \square

Corollary 2.2.2. *Every element $g \in \mathcal{G}$ has a decomposition, called **polar**,*

$$(2.2.1.6) \quad g = k_1 b k_2 \quad \text{with } k_1, k_2 \in \mathcal{K} \text{ and } b \in \mathcal{A}.$$

The element b is uniquely determined up to permutation of diagonal components. Let \mathcal{A}^+ be the set of $a = \text{diag}(a_1, a_2) \in \mathcal{A}$ such that $a_1 \geq 1$. Then $\mathcal{G} = \mathcal{K}\mathcal{A}^+\mathcal{K}$, and the decomposition (2.2.1.6) determines b uniquely in \mathcal{A}^+ .

PROOF. Since gg^* can be written as $gg^* = k_1 b^2 k_1^{-1}$ with $b \in \mathcal{A}$ and $k_1 \in \mathcal{K}$, where b^2 is the matrix of eigenvalues of gg^* . By the bijection in [Theorem 2.2.1](#), there exists $k_2 \in \mathcal{K}$ such that $g = k_1 b k_2$. \square

Given $g = uak \in \mathcal{G}$ with its Iwasawa decomposition, we let

$$(2.2.1.7) \quad g_{\mathcal{A}} := a \quad \text{or also} \quad a_g := a$$

denotes the \mathcal{A} -projection, and similarly for $g_{\mathcal{U}}$ and $g_{\mathcal{K}}$.

2.2.2. Characters. A **character** of \mathcal{A} is a continuous homomorphism $\chi : \mathcal{A} \rightarrow \mathbb{C}^\times$ into the multiplication group of nonzero complex numbers. Additively, let $\mathfrak{a} := \text{Lie}(\mathcal{A})$ be the \mathbb{R} -vector space of diagonal matrices $H = \text{diag}(h_1, h_2)$ with trace zero, so $\mathcal{A} = \exp \mathfrak{a}$. Let \mathfrak{a}^\vee be the dual space. Let α be the (additive) character on \mathfrak{a} defined by $\alpha(H) := h_1 - h_2$. If $a = \exp H = \text{diag}(a_1, a_2) \in \mathcal{A} = \exp \mathfrak{a}$, then we define the **Iwasawa coordinate** $\mathbf{y} : \mathcal{A} \rightarrow \mathbb{R}^+$ by

$$(2.2.2.1) \quad \mathbf{y}(a) \equiv \mathbf{y}_a \equiv \chi_\alpha(a) \equiv a^\alpha := \frac{a_1}{a_2} = a_1^2.$$

We note that α is a basis of \mathfrak{a}^\vee . Furthermore, \mathbf{y} gives a group isomorphism of \mathcal{A} with the positive multiplicative group. For complex s , we get

$$(2.2.2.2) \quad \mathbf{y}_a^s = a^{s\alpha} = \chi_{s\alpha}(a),$$

with a complex (multiplicative) character $\chi_{s\alpha}$. We often write y instead of \mathbf{y}_a and

$$(2.2.2.3) \quad a = a_y \quad \text{or} \quad a(y) = \begin{bmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{bmatrix}.$$

Let \mathbf{D} = diagonal subgroup of \mathcal{G} with elements

$$(2.2.2.4) \quad z = \begin{bmatrix} z_1 & 0 \\ 0 & z_1^{-1} \end{bmatrix}$$

and \mathbf{T} = torus subgroup = subgroup of \mathbf{D} consisting of elements

$$(2.2.2.5) \quad \epsilon = \begin{bmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_1^{-1} \end{bmatrix} \quad \text{with } |\epsilon_1| = 1.$$

Then we have a direct product decomposition

$$(2.2.2.6) \quad \mathbf{D} = \mathcal{A}\mathbf{T} \quad \text{with } z = a_z \epsilon \quad \text{and} \quad \mathbf{T} = \mathbf{D} \cap \mathcal{K}.$$

Indeed, writing $z_1 = |z_1|\epsilon_1$,

$$z = \begin{bmatrix} z_1 & 0 \\ 0 & z_1^{-1} \end{bmatrix} = \begin{bmatrix} |z_1| & 0 \\ 0 & |z_1|^{-1} \end{bmatrix} \begin{bmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_1^{-1} \end{bmatrix} =: a_z \epsilon.$$

The character α is actually the restriction of a character on \mathbf{D} , namely

$$(2.2.2.7) \quad z^\alpha = \frac{z_1}{z_2} = z_1^2.$$

The kernel of α in \mathbf{D} is the group $\{\pm \mathbf{1}\}$. We define $z \in \mathbf{D}$ to be **regular** if its diagonal components are distinct, in other words if $z^\alpha \neq 1$. More generally, an arbitrary square matrix is called **regular** if it can be conjugated to a regular diagonal matrix.

2.2.3. \mathcal{K} -bi-invariant functions. Let f be a function on \mathcal{G} , which we write $f \in \mathbf{Fu}(\mathcal{G})$. We say that f is **\mathcal{K} -bi-invariant** if

$$(2.2.3.1) \quad f(k_1 g k_2) = f(g), \quad \text{for all } k_1, k_2 \in \mathcal{K} \text{ and } g \in \mathcal{G}.$$

By the polar decomposition, if b is the \mathcal{A} -polar component of g , then $f(g) = f(b)$ and f is determined by its values on \mathcal{A} . Let

$$(2.2.3.2) \quad w := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

be called the **Weyl element**, and let the group of order 2 that it generates mod $\pm \mathbf{1}$ be called the **Weyl group** $\mathcal{W} \subseteq \mathcal{K}$. Then w acts on \mathbf{D} and \mathcal{A} by conjugation, and for $z \in \mathbf{D}$,

$$(2.2.3.3) \quad z^w := w z w^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} z_1 & 0 \\ 0 & z_1^{-1} \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} z_1^{-1} & 0 \\ 0 & z_1 \end{bmatrix} = z^{-1}.$$

Proposition 2.2.3. *A \mathcal{K} -bi-invariant function is necessarily even.*

PROOF. By the polar decomposition (2.2.1.6), $g = k_1 b k_2$, where $b \in \mathcal{A}$, so if φ is \mathcal{K} -bi-invariant, then

$$\varphi(g) = \varphi(b) = \varphi(w b w^{-1}) = \varphi(b^w) = \varphi(b^{-1})$$

because of $b = w b^{-1}$ and (2.2.3.3). Thus the restriction map from \mathcal{G} to \mathcal{A} induces bijections

$$(2.2.3.4) \quad \mathbf{Fu}(\mathcal{K} \backslash \mathcal{G} / \mathcal{K}) \rightleftharpoons \mathbf{Fu}_{\text{even}}(\mathcal{A}) \quad \text{and} \quad C(\mathcal{K} \backslash \mathcal{G} / \mathcal{K}) \rightleftharpoons C_{\text{even}}(\mathcal{A}).$$

Since the Iwasawa coordinate y gives an isomorphism of \mathcal{A} with \mathbb{R}^+ , the function $v := \ln y$ gives an isomorphism of \mathcal{A} with \mathbb{R} . We obtain a linear isomorphism

$$(2.2.3.5) \quad \mathbf{Fu}(\mathcal{K} \backslash \mathcal{G} / \mathcal{K}) \rightleftharpoons \mathbf{Fu}_{\text{even}}(\mathbb{R}).$$

Given a \mathcal{K} -bi-invariant function $f \in \mathbf{Fu}(\mathcal{K} \backslash \mathcal{G} / \mathcal{K})$, the corresponding even function on \mathbb{R} will be denoted by f_+ , so that we have $f(g) = f(b) = f_+(v)$. \square

We denote the **conjugation action** of \mathcal{G} on itself by \mathbf{c} , so

$$(2.2.3.6) \quad \mathbf{c}(g)g' := g g' g^{-1}.$$

This action induces what we also call the conjugation action on various functors.

Let \mathfrak{u} be the Lie algebra of \mathcal{U} , by definition the algebra of strictly upper triangular matrices, so of the form

$$(2.2.3.7) \quad \begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix}, \quad x \in \mathbb{C}.$$

Then \mathfrak{u} is 1-dimensional as a complex vector space, with \mathbb{C} -basis

$$(2.2.3.8) \quad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

For \mathfrak{u} as a vector space over \mathbb{R} , a natural basis consists of E_{12} and iE_{12} . We have

$$(2.2.3.9) \quad \mathcal{U} = \mathbf{1} + \mathfrak{n} = \exp \mathfrak{n},$$

where \exp is the exponential given by the usual power series.

Let $a = \text{diag}(a_1, a_1^{-1}) \in \mathcal{A}$ as above. Then E_{12} and iE_{12} are eigenvectors for the conjugation action by elements of \mathcal{A} . In fact¹,

$$(2.2.3.10) \quad a(E_{12})a^{-1} = a^\alpha(E_{12}), \quad a(iE_{12})a^{-1} = a^\alpha(iE_{12}).$$

Thus $\mathfrak{u} = \mathfrak{u}_\alpha$ is an eigenspace, with eigencharacter α , which has multiplicity 2 viewing \mathfrak{u} as a vector space of dimension 2 over \mathbb{R} . Note that ${}^t\mathfrak{u}$ is a $(-\alpha)$ -eigenspace. We call $\alpha, -\alpha$ the **regular characters**. The trace of the representation on \mathfrak{u} is 2α , and the half-trace is $\rho := \alpha = \rho_{\mathcal{G}}$ (this is notation that becomes significant in the higher-dimensional case). More generally, for $z \in \mathbf{D}$,

$$(2.2.3.11) \quad zu(x)z^{-1} = \begin{bmatrix} 1 & z_1^2 x \\ 0 & 1 \end{bmatrix} = u(z^\alpha x).$$

We let $\delta = \delta_{\mathcal{G}}$ be the **Iwasawa character** on \mathcal{A} , namely

$$(2.2.3.12) \quad \delta_{\mathcal{G}}(a) := a^{2\alpha} = y_a^2 = a_1^4.$$

The exponent is the trace of the representation, viewed as being on a 2-dimensional \mathbb{R} -space. The absolute Jacobian of conjugation on \mathfrak{u} viewed as an \mathbb{R} -space is

$$(2.2.3.13) \quad |\delta(z)| = a_z^{2\alpha}.$$

2.2.4. Group Fubini theorem. In the concrete application to $\mathcal{G} = \mathbf{SL}_2(\mathbb{C})$ and its Iwasawa decomposition $\mathcal{G} = \mathcal{U}\mathcal{A}\mathcal{K} = \mathcal{A}\mathcal{U}\mathcal{K}$, we have

$$(2.2.4.1) \quad \mathcal{P} = \mathcal{G}/\mathcal{K} = \mathcal{U}\mathcal{A} = \mathcal{A}\mathcal{U},$$

and \mathcal{U} is normal in \mathcal{P} . All four of $\mathcal{G}, \mathcal{K}, \mathcal{U}, \mathcal{A}$ are unimodular, so all the hypotheses in the previous propositions are satisfied. The character $a \mapsto \delta(a)$ in **Proposition 2.1.8** is the Iwasawa character, $\delta(a) = a^{2\alpha}$. We normalize the Haar measure by taking $c = 1$. Thus we define the **Iwasawa measure** to be the Haar measure such that for $f \in C_c(\mathcal{G})$ we have

$$(2.2.4.2) \quad \int_{\mathcal{G}} f(g) dg := \int_{\mathcal{K}} \int_{\mathcal{U}} \int_{\mathcal{A}} f(auk) da dudk = \int_{\mathcal{K}} \int_{\mathcal{A}} \int_{\mathcal{U}} f(uak) a^{-2\alpha} dudadk,$$

¹For example,

$$a(E_{12})a^{-1} = \begin{bmatrix} a_1 & 0 \\ 0 & a_1^{-1} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1^{-1} & 0 \\ 0 & a_1 \end{bmatrix} = \begin{bmatrix} 0 & a_1^2 \\ 0 & 0 \end{bmatrix} = a_1^2 E_{12} = a^\alpha E_{12}.$$

where

$$da := \frac{dy}{y} \text{ with } y = a^\alpha,$$

$$du := \text{Euclidean measure } dx = dx^1 dx^2 \text{ if } x = x^1 + \mathbf{i}x^2 \text{ with } x^1, x^2 \in \mathbb{R},$$

$$dk := \text{Haar measure on } \mathcal{K} \text{ giving } \mathcal{K} \text{ total measure 1.}$$

Unless otherwise specified, all integral formulas are with respect to the Iwasawa measure.

The Iwasawa measure is canonically determined by the Iwasawa decomposition. We may use the notation

$$(2.2.4.3) \quad dg \equiv d_{\mathbf{IW}}g \equiv d\mu_{\mathbf{IW}}(g)$$

for the Iwasawa measure, when comparing it with other measures. Thus the basic Iwasawa measure integration formula can be written

$$(2.2.4.4) \quad \int_{\mathcal{G}} f(g) d\mu_{\mathbf{IW}}(g) = \int_{\mathcal{K}} \int_{\mathcal{A}} \int_{\mathcal{U}} f(uak) a^{-2\alpha} du da dk, \quad \mathbf{INT1.}$$

The above normalization of the Haar measure fits perfectly the normalization that comes from the quaternionic model for $\mathcal{G}/\mathcal{K} \cong \mathbb{H}^3$, i.e., the set of quaternions $z = x^1 + \mathbf{i}x^2 + \mathbf{j}y$, and the standard normalization of the measure on this homogeneous space for the action of $\mathbf{SL}_2(\mathbb{C})$ given by $(az + b)/(cz + d)$ with complex a, b, c, d . This measure is $dx^1 dx^2 dy/y^3$.

For z diagonal regular in \mathcal{G} , and $u \in \mathcal{U}$, the **commutator** is $[z, u] = zu(uz)^{-1} = zuz^{-1}u^{-1}$. Define

$$(2.2.4.5) \quad J_{\mathcal{G}, \mathbf{com}}(z) \equiv J_{\mathbf{com}}(z) := |1 - z^\alpha|^{m(\alpha)} \text{ with } m(\alpha) = 2.$$

The **exponent** is the multiplicity $m(\alpha)$ of α . Then

$$(2.2.4.6) \quad \int_{\mathcal{U}} f(z^{-1}uzu^{-1}) du = \frac{1}{J_{\mathbf{com}}(z^{-1})} \int_{\mathcal{U}} f(u) du, \quad \mathbf{INT2.}$$

Indeed, matrix multiplication shows that

$$[z^{-1}, u] = z^{-1}uzu^{-1} = \begin{bmatrix} z_1^{-1} & 0 \\ 0 & z_1 \end{bmatrix} \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 & 0 \\ 0 & z_1^{-1} \end{bmatrix} \begin{bmatrix} 1 & -x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & (z_1^{-2} - 1)x \\ 0 & 1 \end{bmatrix}$$

whence

$$\begin{aligned} \int_{\mathcal{U}} f(z^{-1}uzu^{-1}) du &= \int_{\mathbb{C}} f[u((z^{-\alpha} - 1)x)] \frac{\mathbf{i}}{2} dx \wedge d\bar{x} \\ &= \left[(z^{-\alpha} - 1) \overline{(z^{-\alpha} - 1)} \right]^{-1} \int_{\mathbb{C}} f(u(x)) \frac{\mathbf{i}}{2} dx \wedge d\bar{x} = \frac{1}{J_{\mathbf{com}}(z^{-1})} \int_{\mathcal{U}} f(u) du. \end{aligned}$$

From the product decomposition $\mathbf{D} = \mathcal{A}\mathbf{T}$, we give the compact group \mathbf{T} the total Haar measure 1. Then we give \mathbf{D} the product measure of $da = dy/y$ and this normalized measure on \mathbf{T} , which we write dw . Then $\mathbf{D} \setminus \mathcal{G}$ has the **homogeneous space measure**, which makes the **group Fubini theorem** valid, that is, $\mu_{\mathbf{D} \setminus \mathcal{G}}$ such that

$$(2.2.4.7) \quad \int_{\mathcal{G}} f(g) d\mu_{\mathbf{IW}}(g) = \int_{\mathbf{D} \setminus \mathcal{G}} \int_{\mathbf{D}} f(w\dot{g}) dw d\mu_{\mathbf{D} \setminus \mathcal{G}}(\dot{g}).$$

Under the correspondence $\mathbf{D} \setminus \mathcal{G} \leftrightarrow \mathcal{U}(\mathbf{T} \setminus \mathcal{K})$, for right \mathcal{K} -invariant functions, we have with respect to a decomposition $g = wuk$ ($w \in \mathbf{D}$, $u \in \mathcal{U}$, $k \in \mathcal{K}$),

$$(2.2.4.8) \quad d\mu_{\mathbf{D} \setminus \mathcal{G}}(g) = du.$$

Theorem 2.2.4. *Let $\varphi \in C_c(\mathcal{K} \setminus \mathcal{G}/\mathcal{K})$ and let z be a regular diagonal element. Then*

$$(2.2.4.9) \quad g \mapsto \varphi(g^{-1}zg)$$

has compact support on $\mathbf{D} \setminus \mathcal{G}$, and we have

$$(2.2.4.10) \quad \int_{\mathbf{D} \setminus \mathcal{G}} \varphi(g^{-1}zg) d\mu_{\mathbf{D} \setminus \mathcal{G}}(g) = \frac{1}{J_{\text{com}}(z^{-1})} \int_{\mathcal{U}} \varphi(a_z u) du.$$

PROOF. For (2.2.4.10), using $g^{-1}zg = k^{-1}(u^{-1}w^{-1}z w u)k$, one has

$$\begin{aligned} \int_{\mathbf{D} \setminus \mathcal{G}} \varphi(g^{-1}zg) dg &= \int_{\mathcal{U}} \varphi(u^{-1}zu) du = \int_{\mathcal{U}} \varphi(z z^{-1} u^{-1} z u) du \\ &= \int_{\mathcal{U}} (\varphi \circ z)(z^{-1} u^{-1} z u) du = \int_{\mathcal{U}} (\varphi \circ z)(z^{-1} u z u^{-1}) du = \frac{1}{J_{\text{com}}(z^{-1})} \int_{\mathcal{U}} \varphi(zu) du \end{aligned}$$

by (2.2.4.6). Since, writing $z_1 = |z_1| \epsilon$,

$$\begin{bmatrix} \epsilon^{-1} & 0 \\ 0 & \epsilon \end{bmatrix} \begin{bmatrix} z_1 & z_1 x \\ 0 & z_1^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} |z_1| & |z_1| x \\ 0 & |z_1|^{-1} \end{bmatrix} = a_z u,$$

we see that $\varphi(zu) = \varphi(a_z u)$. \square

2.3. Harish, Mellin and spherical transforms

The Iwasawa decomposition leads into the Harish transform and Mellin transform, whose composite is the spherical transform, of Harish-Chandra, relating analytic (Fourier-type) expansion on \mathcal{G}/\mathcal{K} to those on \mathcal{A} .

2.3.1. The Harish transform and the orbital integral. We factorize $J_{\text{com}}(z)$ for diagonal z as follows. For z as given by (2.2.2.4), we have $z^{\alpha/2} = z_1$ and $z^{-\alpha/2} = z_1^{-1}$, so that $|z^{\alpha/2} - z^{-\alpha/2}| = |z_1 - z_1^{-1}| = |z_1^{-1} - z_1| = |(z^{-1})^{\alpha/2} - (z^{-1})^{-\alpha/2}|$. Put

$$(2.3.1.1) \quad |\mathbf{D}|(z^{-1}) = |\mathbf{D}|(z) := \left| z^{\frac{\alpha}{2}} - z^{-\frac{\alpha}{2}} \right|^{m(\alpha)} = \left| z^{\frac{\alpha}{2}} - z^{-\frac{\alpha}{2}} \right|^2 = \left| z_1 - z_1^{-1} \right|^2.$$

Then

$$(2.3.1.2) \quad J_{\text{com}}(z^{-1}) = |\mathbf{D}|(z) \delta(a_z)^{-1/2} = |\mathbf{D}|(z) z^{-\alpha}.$$

Note that we can take $z^{\alpha/2} = z_1$ and $z^{-\alpha/2} = z_1^{-1}$. We define the **orbital integral** by

$$(2.3.1.3) \quad \mathbf{Orb}_{\mathbf{D} \setminus \mathcal{G}}(\varphi, z) := \int_{\mathbf{D} \setminus \mathcal{G}} \varphi(g^{-1}zg) d\mu_{\mathbf{D} \setminus \mathcal{G}}(g)$$

with $\mu_{\mathbf{D} \setminus \mathcal{G}}$ as in (2.2.4.8). We define the **Harish transform** of $\varphi \in C_c(\mathcal{K} \setminus \mathcal{G}/\mathcal{K})$ by

$$(2.3.1.4) \quad (\mathbf{H}\varphi)(z) := \delta(a_z)^{1/2} \int_{\mathcal{U}} \varphi(a_z u) du = |\mathbf{D}|(z) \mathbf{Orb}_{\mathbf{D} \setminus \mathcal{G}}(\varphi, z).$$

The equality on the right comes from **Theorem 2.2.4** and (2.3.1.2). The first formula is applicable to all diagonal z , but the second, with the orbital integral, is applicable only to regular z . As usual in measure-integration theory, the equality extends to functions for which the integrals are absolutely convergent, by continuity. This

will be important for us, and certain convergence properties will be dealt with at the appropriate time and place.

Theorem 2.3.1. *For $\varphi \in C_c(\mathcal{K} \backslash \mathcal{G} / \mathcal{K})$, the Harish transform $\mathbf{H}\varphi$ is invariant under the Weyl group, i.e.,*

$$(2.3.1.5) \quad (\mathbf{H}\varphi)(a) = (\mathbf{H}\varphi)(a^{-1}).$$

PROOF. $\mathbf{H}\varphi$ is invariant if and only if $(\mathbf{H}\varphi)(wzw^{-1}) = (\mathbf{H}\varphi)(z)$, or equivalently, if and only if $(\mathbf{H}\varphi)(z^{-1}) = (\mathbf{H}\varphi)(z)$ by (2.2.3.3). Using (2.2.2.6) and $\mathbf{T} \subseteq \mathcal{K}$ and the fact that φ is \mathcal{K} -bi-invariant, it suffices to verify that $(\mathbf{H}\varphi)(a) = (\mathbf{H}\varphi)(a^{-1})$.

By continuity, it suffices to prove the assertion when a is regular, so $|\mathbf{D}|(a) = |\mathbf{D}|(a^{-1})$. From the fact that \mathcal{K} -bi-invariant implies that φ is even, we get

$$\begin{aligned} (\mathbf{H}\varphi)(a^{-1}) &= |\mathbf{D}|(a^{-1}) \mathbf{Orb}_{\mathbf{D} \backslash \mathcal{G}}(\varphi, a^{-1}) = |\mathbf{D}|(a) \int_{\mathbf{D} \backslash \mathcal{G}} \varphi(g^{-1}a^{-1}g) d\mu_{\mathbf{D} \backslash \mathcal{G}}(g) \\ &= |\mathbf{D}|(a) \int_{\mathbf{D} \backslash \mathcal{G}} \varphi((g^{-1}a^{-1}g)^{-1}) d\mu_{\mathbf{D} \backslash \mathcal{G}}(g) = (\mathbf{H}\varphi)(a). \end{aligned}$$

One can also argue directly on the other integral defining \mathbf{H} , using the Jacobian for $u \mapsto a^{-1}ua$. \square

The Harish transform is therefore a linear map

$$(2.3.1.6) \quad \mathbf{H} : C_c(\mathcal{K} \backslash \mathcal{G} / \mathcal{K}) \longrightarrow C_c(\mathcal{A})^{\mathcal{W}},$$

where the upper index \mathcal{W} means the space of functions invariant under \mathcal{W} .

We consider next the behavior under products, notably the **convolution product** on \mathcal{G} , defined for a specified Haar measure by

$$(2.3.1.7) \quad (h * f)(z) := \int_{\mathcal{G}} h(zg^{-1})f(g) d\mu_{\mathbf{I}\mathbf{W}}g$$

on a class of pairs of functions for which the integral is absolutely convergent. For non-commutative \mathcal{G} , the \mathcal{K} -bi-invariance (and the evenness) allows a choice of writing variables on one side or another, with inverses or not, namely for ψ, f, \mathcal{K} -bi-invariant such that for all z the function $z \mapsto \psi(zg)f(g)$ is in L^1 , the convolution $(\psi * f)(z)$ is defined by

$$\begin{aligned} (\psi * f)(z) &:= \int_{\mathcal{G}} \psi(zg)f(g) d\mu_{\mathbf{I}\mathbf{W}}(g) = \int_{\mathcal{G}} \psi(zg^{-1})f(g^{-1}) d\mu_{\mathbf{I}\mathbf{W}}(g) \\ (2.3.1.8) \quad &= \int_{\mathcal{G}} \psi(zg^{-1})f(g) d\mu_{\mathbf{I}\mathbf{W}}(g) = \int_{\mathcal{G}} \psi(g^{-1})f(gz) d\mu_{\mathbf{I}\mathbf{W}}(g) \\ &= \int_{\mathcal{G}} \psi(g)f(gz) d\mu_{\mathbf{I}\mathbf{W}}(g). \end{aligned}$$

Theorem 2.3.2. (Gelfand) *Let \mathcal{G} be a locally compact unimodular group, and let \mathcal{K} be a compact subgroup. Let τ be an anti-automorphism of \mathcal{G} of order 2 such that given $g \in \mathcal{G}$, there exist $k_1, k_2 \in \mathcal{K}$ satisfying ${}^\tau g = k_1 g k_2$. Then the convolution is commutative (whenever the convolution integral is absolutely convergent).*

PROOF. Since ψ, f are \mathcal{K} -bi-invariant and \mathcal{G} is unimodular, it follows that ψ, f are τ -invariant and so is the Haar measure. Therefore

$$\begin{aligned}
(\psi * f)(z) &= \int_{\mathcal{G}} \psi(zg^{-1})f(g)dg = \int_{\mathcal{G}} \psi(\tau z \tau g^{-1})f(\tau g)dg \\
&= \int_{\mathcal{G}} \psi(\tau g^{-1})f(\tau z \tau g)dg \quad (\text{by } g \mapsto gz) = \int_{\mathcal{G}} \psi(g^{-1})f(k_1 z k_2 \tau g)dg \\
&= \int_{\mathcal{G}} \psi(\tau g^{-1})f(k_1 z k_2 g)dg = \int_{\mathcal{G}} \psi(g^{-1}k_2)f(k_1 z g)dg \\
&= \int_{\mathcal{G}} f(zgk_2^{-1})\psi(k_2g^{-1})dg = \int_{\mathcal{G}} f(z(k_2g^{-1})^{-1})\psi(k_2g^{-1})dg.
\end{aligned}$$

Thus $\psi * f = f * \psi$. \square

Theorem 2.3.3. For $\varphi, \psi \in C_c(\mathcal{K} \backslash \mathcal{G} / \mathcal{K})$, we have

$$(2.3.1.9) \quad \mathbf{H}(\varphi * \psi) = \mathbf{H}\varphi * \mathbf{H}\psi,$$

i.e., on $C_c(\mathcal{K} \backslash \mathcal{G} / \mathcal{K})$ the Harish transform is an algebra homomorphism for the convolution product.

PROOF. By continuity, we may assume that $z = a \in \mathcal{A}$. Then, in this case, $a^\alpha = \delta(a)^{1/2}$ and $a_z = a_a = a$. Hence

$$\begin{aligned}
(\mathbf{H}(\varphi * \psi))(a) &= a^\alpha \int_{\mathcal{U}} (\varphi * \psi)(au)du = a^\alpha \int_{\mathcal{U}} \int_{\mathcal{G}} \varphi(aug)\psi(g^{-1})dgdu \\
&= a^\alpha \int_{\mathcal{U}} \int_{\mathcal{G}} \varphi(ag)\psi(g^{-1}u)dgdu.
\end{aligned}$$

Let $g = bvk$ be the Iwasawa decomposition with $b \in \mathcal{A}, v \in \mathcal{U}$, and $k \in \mathcal{K}$. Then, using (2.2.4.2), we get

$$\begin{aligned}
(\mathbf{H}(\varphi * \psi))(a) &= a^\alpha \int_{\mathcal{U}} \int_{\mathcal{K}} \int_{\mathcal{U}} \int_{\mathcal{A}} \varphi(abvk)\psi(k^{-1}v^{-1}b^{-1}u)dbdvdkdu \\
&= a^\alpha \int_{\mathcal{U}} \int_{\mathcal{K}} \int_{\mathcal{U}} \int_{\mathcal{A}} \varphi(abv)\psi(v^{-1}b^{-1}u)dbdvdkdu \\
&= a^\alpha \int_{\mathcal{U}} \int_{\mathcal{U}} \int_{\mathcal{A}} \varphi(abv)\psi(v^{-1}b^{-1}u)dbdvdu \\
&= a^\alpha \int_{\mathcal{U}} \int_{\mathcal{U}} \int_{\mathcal{A}} \varphi(ab^{-1}v)\psi(v^{-1}bu)dbdudv \\
&= a^\alpha \int_{\mathcal{U}} \int_{\mathcal{U}} \int_{\mathcal{A}} \varphi(ab^{-1}v)\psi(v^{-1}ub)\delta(b)^{-1}dbdvdu \\
&= a^\alpha \int_{\mathcal{U}} \int_{\mathcal{U}} \int_{\mathcal{A}} \varphi(ab^{-1}v)\psi(ub)\delta(b)^{-1}dbdvdu \\
&= a^\alpha \int_{\mathcal{U}} \int_{\mathcal{U}} \int_{\mathcal{A}} \varphi(ab^{-1}v)\psi(bu)dbdvdu \\
&= (ab^{-1})^\alpha b^\alpha \int_{\mathcal{U}} \int_{\mathcal{U}} \int_{\mathcal{A}} \varphi(ab^{-1}v)\psi(bu)dbdvdu
\end{aligned}$$

which is exactly the convolution $(\mathbf{H}\varphi * \mathbf{H}\psi)(a)$. \square

2.3.2. The Mellin and spherical transforms. Let χ be a character on \mathcal{A} (continuous homomorphism into \mathbb{C}^\times). Let $F \in C_c(\mathcal{A})$. We define the **Mellin transform** of F to be the function on the group $\mathbf{ch}(\mathcal{A})$ of characters given by

$$(2.3.2.1) \quad (\mathbf{M}F)(\chi) := \int_{\mathcal{A}} F(a)\chi(a) da.$$

We use the same Haar measure da on \mathcal{A} as before, with the coordinate character \mathbf{y} . Viewing \mathbf{y} as a basis for the characters, we can write a character in the form

$$(2.3.2.2) \quad \chi(a) := \mathbf{y}_a^s = a^{s\alpha}$$

with a complex variable s . Then the Mellin transform is written with $F_\alpha(\mathbf{y}) = F(a)$:

$$(2.3.2.3) \quad (\mathbf{M}_\alpha F)(s) \equiv (\mathbf{M}F_\alpha)(s) = \int_0^\infty F_\alpha(\mathbf{y})\mathbf{y}^s \frac{d\mathbf{y}}{\mathbf{y}}.$$

Here $\mathbf{y} := \mathbf{y}_a \in [0, \infty)$. This is the ordinary Mellin transform of elementary analysis. If F does not have compact support, the integral may converge only in a restricted domain of s , usually some half-plane where the transform is analytic.

If F is even ($F(a) = F(a^{-1})$), then $\mathbf{M}F$ is also even. In other words,

$$(2.3.2.4) \quad (\mathbf{M}F)(\chi) = (\mathbf{M}F)(\chi^{-1}).$$

For $g \in \mathcal{G}$ let $g_{\mathcal{A}}$ be its Iwasawa projection on \mathcal{A} . Then the function $g \mapsto \chi(g_{\mathcal{A}})$, also written $\chi(g)$, is well-defined, and is called the **Iwasawa lifting** of $\chi \in \mathbf{ch}(\mathcal{A})$ to \mathcal{G} . As before, let $\delta = \delta_{\mathcal{G}}$ be the Iwasawa character, $\delta(a) = a^{2\alpha}$. We define the **spherical kernel** on the product of \mathcal{G} with the character group $\mathbf{ch}(\mathcal{A})$ to be the function given by

$$(2.3.2.5) \quad \Phi_\chi(g) \equiv \Phi(\chi, g) := \int_{\mathcal{K}} (\chi\delta^{1/2})(kg) dk = \int_{\mathcal{K}} (\chi\delta^{1/2})((kg)_{\mathcal{A}}) dk.$$

Let $\mathfrak{a} := \text{Lie}(\mathcal{A})$ be the real vector space of diagonal 2×2 matrices of trace 0,

$$(2.3.2.6) \quad H = \begin{bmatrix} h_1 & 0 \\ 0 & h_2 \end{bmatrix} \quad \text{with } h_2 = -h_1.$$

Let ζ denote a complex (additive) character of \mathcal{A} . Then χ may be written as χ_ζ , with ζ such that if $a = \exp H$, then

$$(2.3.2.7) \quad \chi_\zeta(a) := e^{\zeta(H)} = e^{\zeta(\ln a)}.$$

Thus the Mellin transform can also be viewed as defined on the complex dual $(\mathfrak{a}_{\mathbb{C}})^\vee$, and the spherical kernel may be written in terms of the additive variable ζ , in the form

$$(2.3.2.8) \quad \Phi_\zeta(g) \equiv \Phi(\zeta, g) = \Phi(\chi_\zeta, g), \quad \zeta \in (\mathfrak{a}_{\mathbb{C}})^\vee.$$

We let $\mathbf{S} \equiv \mathbf{S}_{\text{IW}}$ be the **integral (spherical) transform** defined by this kernel,

$$(2.3.2.9) \quad (\mathbf{S}f)(\chi) := \int_{\mathcal{G}} \Phi(\chi, g)f(g) dg.$$

Theorem 2.3.4. *Let f be measurable and \mathcal{K} -bi-invariant on \mathcal{G} . Also suppose that the repeated integral*

$$\int_{\mathcal{G}} \int_{\mathcal{K}} (|\chi|\delta^{1/2})((kg)_{\mathcal{A}}) |f(g)| dk dg$$

is absolutely convergent. Then

$$(2.3.2.10) \quad \mathbf{S}f = \mathbf{M}(\mathbf{H}f),$$

that is, $\mathbf{S} = \mathbf{M} \circ \mathbf{H}$ on $C_c(\mathcal{K} \backslash \mathcal{G}/\mathcal{K})$.

PROOF. From (2.3.1.6), consider the following diagram

$$\begin{array}{ccc} C_c(\mathcal{K} \backslash \mathcal{G}/\mathcal{K}) & \xrightarrow{\mathbf{H}} & C_c(\mathcal{A})^{\mathcal{W}} \subseteq C_c(\mathcal{A}) \\ & \searrow \mathbf{S} & \downarrow \mathbf{M} \\ & & \mathbf{Fu}(\mathbf{ch}(\mathcal{A})) \end{array}$$

By Fubini's theorem and \mathcal{K} -invariance of f and dg , we have

$$\begin{aligned} (\mathbf{S}f)(\chi) &= \int_{\mathcal{G}} \Phi(\chi, g) f(g) dg = \int_{\mathcal{K}} \int_{\mathcal{G}} (\chi \delta^{1/2})((kg)_{\mathcal{A}}) f(g) dg dk \\ &= \int_{\mathcal{K}} \int_{\mathcal{G}} (\chi \delta^{1/2})(g_{\mathcal{A}}) f(g) dg dk = \int_{\mathcal{A}} \int_{\mathcal{U}} \chi(a) \delta^{1/2}(a) f(au) da du \\ &= \int_{\mathcal{A}} \chi(a) (\mathbf{H}f)(a) da = \mathbf{M}(\mathbf{H}f)(\chi). \end{aligned}$$

Thus $\mathbf{S}f = \mathbf{M}(\mathbf{H}f)$. □

Corollary 2.3.5. *The spherical kernel is invariant under \mathcal{W} in the variable χ , that is,*

$$(2.3.2.11) \quad \Phi_{\chi} = \Phi_{\chi^{-1}}.$$

PROOF. From (2.3.2.10), we have, for any $f \in C_c(\mathcal{K} \backslash \mathcal{G}/\mathcal{K})$,

$$\begin{aligned} (\mathbf{S}f)(\chi^{-1}) &= \int_{\mathcal{A}} \chi^{-1}(a) (\mathbf{H}f)(a) da = \int_{\mathcal{A}} \chi(a^{-1}) (\mathbf{H}f)(a) da \\ &= \int_{\mathcal{A}} \chi(a) (\mathbf{H}f)(a^{-1}) da = \int_{\mathcal{A}} \chi(a) (\mathbf{H}f)(a) da = (\mathbf{S}f)(\chi). \end{aligned}$$

Therefore $\Phi_{\chi} = \Phi_{\chi^{-1}}$. □

Theorem 2.3.6. *The Mellin transform and the spherical transform are multiplicative homomorphisms, that is, for $\varphi, \psi \in C_c(\mathcal{K} \backslash \mathcal{G}/\mathcal{K})$, and $f, h \in C_c(\mathcal{A})$, we have*

$$(2.3.2.12) \quad \mathbf{S}(\varphi * \psi) = \mathbf{S}(\varphi)\mathbf{S}(\psi), \quad \mathbf{M}(f * g) = \mathbf{M}(f)\mathbf{M}(h).$$

PROOF. By definitions,

$$\begin{aligned} \mathbf{M}(f * h)(\chi) &= \int_{\mathcal{A}} (g * h)(a) \chi(a) da = \int_{\mathcal{A}} \int_{\mathcal{A}} f(ab^{-1}) h(b) \chi(a) db da \\ &= \int_{\mathcal{A}} \int_{\mathcal{A}} f(a) h(b) \chi(a) \chi(b) db da = (\mathbf{M}f)(\chi) (\mathbf{M}h)(\chi). \end{aligned}$$

Combining this with (2.3.1.9) and (2.3.2.10) gets the result for \mathbf{S} . □

If $\chi = \chi_{s\alpha}$ (that is, $\chi(a) = \chi_{s\alpha}(a) = a_1^{2s}$), we may write

$$(2.3.2.13) \quad (\mathbf{S}_{\alpha}f)(s) := (\mathbf{S}f)(\chi_{s\alpha}).$$

If $\chi = \chi_{\zeta}$ with $\zeta \in (\mathfrak{a}_{\mathbb{C}})^{\vee}$, we may write

$$(2.3.2.14) \quad (\mathbf{S}f)(\zeta) := (\mathbf{S}f)(\chi_{\zeta}).$$

2.3.3. Computation of the orbital integral. Let $z \in \mathcal{G}$ be regular, that is, z is conjugate to a diagonal matrix with distinct diagonal elements (the eigenvalues)

$$(2.3.3.1) \quad z \sim \tilde{\eta} = \begin{bmatrix} \eta & 0 \\ 0 & \eta^{-1} \end{bmatrix} \quad \text{with } 0 \neq \eta \in \mathbb{C}.$$

Without loss of generality, we may suppose that $|\eta| \geq 1$. If $|\eta| > 1$, then η is uniquely determined among the two eigenvalues η, η^{-1} of z . From the conjugation action on matrices, one sees that the centralizer in \mathcal{G} of a regular diagonal element is the set of diagonal matrices. In particular, the isotropy group of $\tilde{\eta}$ under the conjugation action is the diagonal group \mathbf{D} . Hence the isotropy group \mathcal{G}_z is conjugate to \mathbf{D} by the same conjugation that diagonalizes z to $\tilde{\eta}$. We write $\eta \equiv \eta(z)$ and $\tilde{\eta} \equiv \tilde{\eta}(z)$.

Since conjugation preserves Haar measure, we have

$$(2.3.3.2) \quad \int_{\mathcal{G}_z \backslash \mathcal{G}} \varphi(g^{-1}zg) dg = \int_{\mathbf{D} \backslash \mathcal{G}} \varphi(g^{-1}\tilde{\eta}g) dg.$$

Thus the Harish transform can be computed for z in diagonal form, and we use indiscriminately identity

$$(2.3.3.3) \quad (\mathbf{H}\varphi)(z) = |\mathbf{D}|(\tilde{\eta}(z)) \mathbf{Orb}_{\mathcal{G}_z \backslash \mathcal{G}}(\varphi, z) = |\mathbf{D}|(\eta) \mathbf{Orb}_{\mathbf{D} \backslash \mathcal{G}}(\varphi, \tilde{\eta}).$$

We now go on with a theorem that allows for a more explicit determination of the orbital integral, and therefore of the Harish transform. As in **Section 2.3.2**, we normalize Haar measure so that \mathbf{T} as well as \mathcal{K} has measure 1. On $\mathbf{D} = \mathcal{A}\mathbf{T}$ we put the product measure of $da = dy/y$ on \mathcal{A} and this normalized measure on \mathbf{T} , which we write dw (w variable in \mathbf{D}). Then for $f \in L^1(\mathcal{G}/\mathcal{K})$, we have

$$(2.3.3.4) \quad \int_{\mathcal{G}} f(g) dg = \int_{\mathcal{U}} \int_{\mathbf{D}} f(wu) dw du.$$

The dg on the left is of course the Iwasawa measure. From the modified Iwasawa decomposition $\mathcal{G} = \mathbf{D}\mathcal{U}\mathcal{K}$, we get

$$(2.3.3.5) \quad \begin{aligned} \mathbf{Orb}_{\mathbf{D} \backslash \mathcal{G}}(\varphi, \tilde{\eta}) &= \int_{\mathbf{D} \backslash \mathcal{G}} \varphi(g^{-1}\tilde{\eta}g) d\mu_{\mathbf{D} \backslash \mathcal{G}}(g) = \int_{\mathcal{U}\mathcal{K}} \varphi(k^{-1}u^{-1}z^{-1}\tilde{\eta}zuk) duk \\ &= \int_{\mathcal{U}} \varphi(u^{-1}z^{-1}\tilde{\eta}zu) du = \int_{\mathcal{U}} \varphi(u^{-1}\tilde{\eta}u) du \\ &= \int_{\mathbb{C}} \varphi(u(-x)\tilde{\eta}u(x)) dx = \int_{\mathbb{C}} \varphi(M_{\eta}(x)) dx, \end{aligned}$$

where

$$(2.3.3.6) \quad M_{\eta}(x) := u(x)^{-1}\tilde{\eta}u(x) = \begin{bmatrix} \eta & x(\eta - \eta^{-1}) \\ 0 & \eta^{-1} \end{bmatrix}$$

The function φ is given in terms of polar coordinates, so we have to express the polar \mathcal{A} -coordinate of $M_{\eta}(x)$ in terms of x . We shall deal with the (x, y) coordinates, polar \mathcal{A} -coordinates, and ordinary polar coordinates in \mathbb{C} .

Theorem 2.3.7. *Let the Haar measure on \mathcal{G} be the Iwasawa measure as in **Subsection 2.2.4**. Let φ be measurable on \mathcal{G} , \mathcal{K} -bi-invariant (so even), and such that the orbital integral $\mathbf{Orb}_{\mathbf{D} \backslash \mathcal{G}}(\varphi, \tilde{\eta})$ is absolutely convergent for regular $\tilde{\eta}$. For $b \in \mathcal{A}$, $y = b^{\alpha} \geq 1$, write*

$$(2.3.3.7) \quad \varphi_{\alpha}(y) \equiv \varphi_{\alpha}(y(b)) := \varphi(b).$$

Then for $|\eta| \geq 1$ and $\tilde{\eta}$ regular, we have

$$(2.3.3.8) \quad \mathbf{Ord}_{\mathbf{D} \setminus \mathcal{G}}(\varphi, \tilde{\eta}) = \frac{2\pi}{|\eta - \eta^{-1}|^2} \int_{|\eta|^2}^{\infty} \varphi_{\alpha}(y) \frac{y - y^{-1}}{2} \frac{dy}{y}.$$

Or in terms of the additive variable $v := \ln y$, using the notation

$$(2.3.3.9) \quad \varphi_{\alpha,+}(v) \equiv \varphi_{\alpha,+}(\ln y) := \varphi(b),$$

the formula becomes

$$(2.3.3.10) \quad \begin{aligned} \mathbf{Ord}_{\mathbf{D} \setminus \mathcal{G}}(\varphi, \tilde{\eta}) &= \frac{2\pi}{|\eta - \eta^{-1}|^2} \int_{|\eta|^2}^{\infty} \varphi_{\alpha,+}(\ln y) \frac{y - y^{-1}}{2} \frac{dy}{y} \\ &= \frac{2\pi}{|\eta - \eta^{-1}|^2} \int_{2 \ln |\eta|}^{\infty} \varphi_{\alpha,+}(v) \sinh(v) dv. \end{aligned}$$

PROOF. To go back and forth between the variable x in $M_{\eta}(x)$ and the \mathcal{A} -variable, we use the quadratic map, namely,

$$(2.3.3.11) \quad \text{If } M_{\eta}(x) = k_1 b k_2 \text{ (} k_1, k_2 \in \mathcal{K}, b \in \mathcal{A} \text{), then } M_{\eta}(x) M_{\eta}(x)^* = k_1 b^2 k_1^{-1}.$$

For (2.3.3.6), we have

$$(2.3.3.12) \quad M_{\eta}(x) M_{\eta}(x)^* = \begin{bmatrix} |\eta|^2 + |x|^2 (\eta - \eta^{-1})^2 & x \tilde{\eta}^{-1} (\eta - \eta^{-1}) \\ \bar{x} \eta^{-1} (\tilde{\eta} - \tilde{\eta}^{-1}) & |\eta|^{-2} \end{bmatrix}$$

So b^2 can be computed from the eigenvalues of $M_{\eta}(x) M_{\eta}(x)^*$, namely, the roots of the characteristic polynomial

$$T^2 - \tau(r)T + 1 = 0,$$

where $r = |x|$, and $\tau(r)$ is the trace of the matrix $M_{\eta}(x) M_{\eta}(x)^*$:

$$(2.3.3.13) \quad \tau(r) := |\eta|^2 + |\eta|^{-2} + r^2 |\eta - \eta^{-1}|^2.$$

We have $r^2 = x \bar{x}$ and $dx = dx^1 dx^2 = r dr d\theta$. With $b = \text{diag}(b_1, b_2)$, $b_1 > 1$, we can solve for b as a function of r (and so as a function of x). Introduce the discriminant of the characteristic polynomial

$$\Delta := \tau(r)^2 - 4.$$

Then the roots of the characteristic polynomial are

$$(2.3.3.14) \quad y \equiv \mathbf{y}_b = \mathbf{y}(b) = b_1^2 = \frac{\tau(r) + \Delta^{1/2}}{2} \quad \text{and} \quad b_2^2 = \frac{\tau(r) - \Delta^{1/2}}{2},$$

or also $\tau + \Delta^{1/2} = 2y$. Since $b_2^2 = b_1^{-2}$ we get from (2.3.3.14) the value of Δ in terms of y ,

$$(2.3.3.15) \quad \Delta^{1/2} = y - y^{-1}.$$

Note that y is a variable on the positive multiplicative group. In terms of the multiplicative coordinate $y = \mathbf{y}(b) = b^{\alpha}$, we write

$$\varphi(b) = \varphi_{\alpha}(y).$$

We change variables in the integral form (2.3.3.5), $dx^1 dx^2 = r dr d\theta$, to get

$$(2.3.3.16) \quad \mathbf{Orb}_{\mathbf{D} \setminus \mathcal{G}}(\varphi, \tilde{\eta}) = \int_{\mathcal{C}} \varphi(M_{\eta}(x)) dx = 2\pi \int_0^{\infty} \varphi(b(r)) r dr.$$

We have to write down the Jacobian between the variables y and r . By (2.3.3.13), we have

$$d\tau = |\eta - \eta^{-1}|^2 2r dr.$$

By (2.3.3.14),

$$dy = \frac{1}{2} \left(d\tau + \frac{2\tau d\tau}{2\Delta^{1/2}} \right) = \frac{\Delta^{1/2} + \tau}{2\Delta^{1/2}} d\tau = \frac{y}{\Delta^{1/2}} d\tau.$$

Hence, using (2.3.3.15),

$$(2.3.3.17) \quad \frac{dy}{y} = \frac{|\eta - \eta^{-1}|^2}{\Delta^{1/2}} 2r dr \quad \text{or} \quad r dr = \frac{y - y^{-1}}{2} \frac{1}{|\eta - \eta^{-1}|^2} \frac{dy}{y}.$$

Putting (2.3.3.16) and (2.3.3.17) together yields (2.3.3.8), except determining the limits of integration.

The polar r ranges over $(0, +\infty)$. The expression (2.3.3.13) for $\tau(r)$ shows that if we put $\ell := |\eta|^2 + |\eta|^{-2}$, then

$$(2.3.3.18) \quad \tau \in (\ell, +\infty), \quad \Delta \in (\ell^2 - 4, +\infty), \quad \ell^2 - 4 = (|\eta|^2 - |\eta|^{-2})^2, \quad \ell + \sqrt{\ell^2 - 4} = 2|\eta|^2.$$

From (2.3.3.14) we get

$$(2.3.3.19) \quad y = \frac{\tau + \Delta^{1/2}}{2} \in \left(\frac{\ell + \sqrt{\ell^2 - 4}}{2}, \infty \right) = (|\eta|^2, \infty),$$

determining the limits of integration. The last formula (2.3.3.10) is obvious. \square

Corollary 2.3.8. *For $a \in \mathcal{A}^+$ (i.e., $a^\alpha > 1$, so a regular), we have*

$$(2.3.3.20) \quad (\mathbf{H}\varphi)(a) = 2\pi \int_{a^\alpha}^{\infty} \varphi_\alpha(y) \frac{y - y^{-1}}{2} \frac{dy}{y}.$$

PROOF. It follows from (2.3.1.1), (2.3.3.4), and (2.3.3.9). \square

2.3.4. Gaussians on \mathcal{G} and their spherical transform. By a **basic \mathcal{G} -Gaussian** (**basic Gaussian** for short) we mean a function on \mathcal{G} that is \mathcal{K} -bi-invariant, and with some constant $c > 0$, for all $b \in \mathcal{A}$ has the formula

$$(2.3.4.1) \quad \begin{aligned} \varphi_c(b) &= e^{-(\alpha(\ln b))^2/2c} \frac{\alpha(\ln b)}{\sinh(\alpha(\ln b))} \\ &= e^{-(\ln y)^2/2c} \frac{\ln y}{(y - y^{-1})/2} \quad \text{with } y = y(b), \\ &= e^{-v^2/2c} \frac{v}{\sin hv} \quad \text{with } v = \ln y. \end{aligned}$$

We can then deal with the vector space $\mathbf{Gauss}(\mathcal{G})$ generated by the basic \mathcal{G} -Gaussian functions as a natural space of test functions.

Theorem 2.3.9. *For the orbital integral, with the basic \mathcal{G} -Gaussian $\varphi_c \in \mathbf{Gauss}(\mathcal{G})$, and regular diagonal z , we have the value*

$$(2.3.4.2) \quad \mathbf{Orb}_{\mathbf{D} \setminus \mathcal{G}}(\varphi_c, z) = \frac{2\pi c}{|z^{\alpha/2} - z^{-\alpha/2}|^2} e^{-(\ln |z^\alpha|)^2/2c}.$$

For $a \in \mathcal{A}$ regular, $y = \mathbf{y}(a)$, the Harish transform is given by

$$(2.3.4.3) \quad (\mathbf{H}\varphi_c)(a) = 2\pi c e^{-(\alpha(\ln a))^2/2c} = 2\pi c e^{-(\ln y)^2/2c}.$$

Here we use the y -variable, $\alpha(\ln a) = \ln y$.

PROOF. Letting $\tilde{\eta} = z$ (so that $\eta = z^{\alpha/2}$) and $\varphi = \varphi_c$ in (2.3.3.10), we have

$$\begin{aligned} \mathbf{Orb}_{\mathbf{D} \setminus \mathcal{G}}(\varphi_c, z) &= \frac{2\pi c}{|\eta - \eta^{-1}|^2} \int_{2\ln \eta}^{\infty} e^{-v^2/2c} d\frac{v^2}{2c} \\ &= \frac{2\pi c}{|\eta - \eta^{-1}|^2} e^{-(2\ln \eta)^2/2c} = \frac{2\pi c}{|z^{\alpha/2} - z^{-\alpha/2}|^2} e^{-(\ln |z^{\alpha}|)^2/2c}. \end{aligned}$$

The last formula follows from (2.2.2.2), (2.3.1.4), and (2.3.4.2). \square

The factor 2π came originally from having taken the ordinary Euclidean measure on $\mathbb{C} \cong \mathcal{U}$, so polar coordinates for which the circle has measure 2π , which is standard, but is given priority over the counterconvention to given compact groups measure 1.

Theorem 2.3.10. *Write a character on $\mathfrak{a} = \text{Lie}(\mathcal{A})$ in the form $\chi = s\alpha$ with complex s . Then the spherical transform is*

$$(2.3.4.4) \quad (\mathbf{S}\varphi_c)(s\alpha) = (\mathbf{M}(\mathbf{H}\varphi_c))(s\alpha) = (2\pi c)^{3/2} e^{s^2 c/2}.$$

On the imaginary axis $s = ir$, this gives

$$(2.3.4.5) \quad (\mathbf{S}\varphi_c)(ir\alpha) = (\mathbf{M}(\mathbf{H}\varphi_c))(ir\alpha) = (2\pi c)^{3/2} e^{-r^2 c/2}.$$

PROOF. By **Theorem 2.3.9** and $\mathbf{S} = \mathbf{M} \circ \mathbf{H}$ (cf., **Theorem 2.3.4**) we have (cf., equations (2.3.2.3) and (2.3.2.13))

$$\frac{1}{2\pi c} (\mathbf{S}\alpha\varphi_c)(s) = \frac{1}{2\pi c} (\mathbf{S}\varphi_c)(s\alpha) = \int_0^{\infty} e^{-(\ln y)^2/2c} y^s \frac{dy}{y} = \int_{-\infty}^{\infty} e^{-v^2/2c} e^{vs} dv.$$

Because of the rapid decay of Gaussians, the Mellin transform is entire in s , so it suffices to determine the integral when $s = ir$ (r real) is pure imaginary. One then uses the self-duality of the function $e^{-v^2/2}$ with respect to $dv/\sqrt{2\pi}$, and a change of variables to get rid of c . \square

We let the **Gauss space on \mathbb{R}** , $\mathbf{Gauss}(\mathbb{R})$, be the vector space generated by the Gaussian functions $r \mapsto e^{-r^2 c/2}$, with all constants $c > 0$. Sometimes, we deal with the imaginary axis $i\mathbb{R}$, and $\mathbf{Gauss}(i\mathbb{R})$ is the space generated by the functions having the above value at ir .

Corollary 2.3.11. *The spherical transform gives a linear isomorphism*

$$(2.3.4.6) \quad \mathbf{Gauss}(\mathcal{G}) \longrightarrow \mathbf{Gauss}(\mathbb{R}), \quad \varphi_c \longmapsto (\mathbf{S}\varphi_c)(\cdot\alpha).$$

Corollary 2.3.12. *With the Iwasawa measure $d\mu_{\mathbf{IW}}$, the total integral of φ_c is*

$$(2.3.4.7) \quad \int_{\mathcal{G}} \varphi_c(g) d\mu_{\mathbf{IW}}(g) = (2\pi c)^{3/2} e^{c/2}.$$

PROOF. Taking $s = 1$ in (2.3.4.4) yields

$$(2\pi c)^{3/2} e^{c/2} = (\mathbf{S}\varphi_c)(-\alpha) = \int_{\mathcal{G}} \Phi(-\alpha, g) \varphi_c(g) dg$$

by (2.3.2.9). The spherical kernel is given by the integral

$$\Phi(-\alpha, g) = \int_{\mathcal{K}} (-\alpha \delta^{1/2})((kg)_{\mathcal{A}}) dk = \int_{\mathcal{K}} ((kg)_{\mathcal{A}})^{-\alpha+\alpha} dk = \int_{\mathcal{K}} dk = 1.$$

Thus, we get (2.3.4.7). \square

Let $g \in \mathcal{G}$ with polar decomposition

$$(2.3.4.8) \quad g = k_1 b k_2, \quad k_1, k_2 \in \mathcal{K} \text{ and } b \in \mathcal{A}.$$

The element $b \in \mathcal{A}$ is determined only up to a permutation of its diagonal components. We call $y = \mathbf{y}(b) = b^\alpha$ or y^{-1} a **polar coordinate** (or polar \mathcal{A} -coordinate) of g . Note that $(\ln y)^2$ is uniquely determined by g , and is real analytic on \mathcal{G} . Define the **polar height** σ to be the \mathcal{K} -bi-invariant function such that $\sigma \geq 0$ and

$$(2.3.4.9) \quad \sigma^2(g) = \sigma^2(b) = (\ln y)^2 = \nu^2.$$

Then the basic \mathcal{G} -Gaussian φ_c , given in (2.3.4.1), can be written in the form

$$(2.3.4.10) \quad \varphi_c = e^{-\sigma^2/2c} \frac{\sigma}{\sinh \sigma} = e^{-(\ln y)^2/2c} \frac{\ln y}{\sinh(\ln y)}.$$

Note that $\nu / \sinh \nu$ is a power series in $\nu^2 = (\ln y)^2$, so real analytic on \mathcal{G} .

For $H \in \mathfrak{a}$ given in (2.3.2.6), we define the symmetric bilinear form $\mathcal{B} : \mathfrak{a} \times \mathfrak{a} \rightarrow \mathbb{R}$ by

$$(2.3.4.11) \quad \mathcal{B}(H, H) := \operatorname{tr}(HH) = |H|_{\mathcal{B}}^2.$$

For $b = \operatorname{diag}(b_1, b_2) \in \mathcal{A}$, we have $\ln b = \operatorname{diag}(\ln b_1, \ln b_2) \in \mathfrak{a}$, and

$$|\ln b|_{\mathcal{B}}^2 = (\ln b_1)^2 + (\ln b_2)^2 = 2(\ln b_1)^2 = \frac{(\ln y)^2}{2},$$

so

$$(2.3.4.12) \quad \sigma^2(b) = 2|\ln b|_{\mathcal{B}}^2 = (\ln y)^2.$$

The trace form on \mathfrak{a} can be extended in two natural ways to the complexification, namely, a *symmetric way* and a *Hermitian way*. The symmetric way will be relevant in the next subsection. The trace form induces a corresponding positive definite scalar product on the dual space \mathfrak{a}^\vee , because it induces a natural isomorphism of \mathfrak{a} and \mathfrak{a}^\vee . Specifically, if $\lambda \in \mathfrak{a}^\vee$, then there is an element $H_\lambda \in \mathfrak{a}$ such that for all $H \in \mathfrak{a}$,

$$(2.3.4.13) \quad \mathcal{B}(H_\lambda, H) = \lambda(H).$$

Then by definition, for $\lambda, \xi \in \mathfrak{a}^\vee$, we have

$$(2.3.4.14) \quad \mathcal{B}^\vee(\xi, \lambda) := \mathcal{B}(H_\xi, H_\lambda),$$

so for example², $\mathcal{B}^\vee(\alpha, \alpha) = 2$.

²Let $H_\alpha \in \mathfrak{a}$ be the correspondence of $\alpha \in \mathfrak{a}^\vee$. According to $\operatorname{tr}(H_\alpha H) = \alpha(H)$, we have $H_\alpha = \operatorname{diag}(1, -1)$ and then $\mathcal{B}^\vee(\alpha, \alpha) = \operatorname{tr}(\operatorname{diag}(1, 1)) = 2$.

We can then extend the form from \mathfrak{a}^\vee to a \mathbf{C} -bilinear form $\mathcal{B}_\mathbf{C}^\vee := (\mathcal{B}^\vee)_\mathbf{C} = \mathcal{B}^\vee \otimes_{\mathbf{R}} \mathbf{C}$ on its complexification, i.e., to characters $\zeta \in \mathfrak{a}_\mathbf{C}^\vee = (\mathfrak{a}^\vee)_\mathbf{C} = \mathfrak{a}^\vee \otimes_{\mathbf{R}} \mathbf{C}$ that can be written in the form $\zeta = \xi + i\lambda$, with $\xi, \lambda \in \mathfrak{a}^\vee$. We call ξ and λ the **real** and **imaginary parts** of ζ respectively. Then by definition

$$(2.3.4.15) \quad \mathcal{B}_\mathbf{C}^\vee(\zeta, \zeta) := \mathcal{B}^\vee(\xi, \xi) + 2i\mathcal{B}^\vee(\xi, \lambda) - \mathcal{B}^\vee(\lambda, \lambda).$$

In particular, for $\zeta = s\alpha$ with $s \in \mathbf{C}$, we have

$$(2.3.4.16) \quad \mathcal{B}_\mathbf{C}^\vee(i\lambda, i\lambda) = -\mathcal{B}^\vee(\lambda, \lambda), \quad \mathcal{B}_\mathbf{C}^\vee(s\alpha, s\alpha) = 2s^2.$$

Thus the form $\mathcal{B}_\mathbf{C}^\vee$ is \mathbf{C} -bilinear, not Hermitian. Its restriction to the imaginary axis $\mathfrak{I}\mathfrak{a}^\vee$ is negative definite, corresponding to minus the original form on \mathfrak{a} .

2.3.5. The polar Haar measure and inversion. A \mathcal{K} -bi-invariant function is determined by its values on the polar \mathcal{A} -component b in the polar decomposition $\mathcal{G} = \mathcal{K}\mathcal{A}\mathcal{K}$, i.e., $g = k_1bk_2$. We define the **polar Jacobian**

$$(2.3.5.1) \quad J_{\mathbf{P}}(b) := \left(\frac{b^\alpha - b^{-\alpha}}{2} \right)^{m(\alpha)} = \left(\frac{y(b) - y(b)^{-1}}{2} \right)^2$$

with $y = y(b) = b^\alpha = b_1^2$ and $m(\alpha) = 2$. The associated **polar measure** $d\mu_{\mathbf{P}}$ is given by

$$(2.3.5.2) \quad d\mu_{\mathbf{P}}(b) := J_{\mathbf{P}}(b)db = \left(\frac{e^v - e^{-v}}{2} \right)^2 = (\sinh v)^2 dv$$

in terms of the v -variable, $v = \ln y$. It is a Jacobian computation due to Harish-Chandra in general that the functional

$$(2.3.5.3) \quad \varphi \mapsto \int_{\mathcal{K}} \int_{\mathcal{K}} \int_{\mathcal{A}} \varphi(k_1bk_2) J_{\mathbf{P}}(b) db dk_1 dk_2$$

is a Haar functional on \mathcal{G} . For \mathcal{K} -bi-invariant functions, one does not need the two \mathcal{K} integrations in the above formula. The total dk -measure of \mathcal{K} is assumed to be normalized to 1. The question arises how the polar Haar measure is related to the Iwasawa measure.

Theorem 2.3.13. *For $\varphi \in \mathbf{Gauss}(\mathcal{G})$, we have*

$$(2.3.5.4) \quad \frac{1}{2\pi} \int_{\mathcal{G}} \varphi(g) d\mu_{\mathbf{IW}}(g) = \int_{\mathcal{A}} \varphi(b) J_{\mathbf{P}}(b) db = \int_{\mathcal{A}} \varphi d\mu_{\mathbf{P}}.$$

PROOF. For $\varphi_c \in \mathbf{Gauss}(\mathcal{G})$, one has

$$\frac{1}{2\pi} \int_{\mathcal{G}} \varphi_c(g) d\mu_{\mathbf{IW}}(g) = (2\pi)^{1/2} c^{3/2} e^{c/2}.$$

On the other hand,

$$\int_{\mathcal{A}} \varphi_c(b) d\mu_{\mathbf{P}}(b) = \int_{-\infty}^{\infty} e^{-v^2/2c} v \left(\frac{e^v - e^{-v}}{2} \right) dv = \int_{-\infty}^{\infty} e^{-v^2/2c} e^v v dv.$$

Using the identity

$$(2.3.5.5) \quad \frac{1}{s} \int_{-\infty}^{\infty} e^{sx} e^{-x^2/2x} x dx = (2\pi)^{1/2} c^{3/2} e^{s^2c/2}, \quad s \in \mathbf{C},$$

we get (2.3.5.4). \square

Corollary 2.3.14. *For all $f \in C_c(\mathcal{K} \backslash \mathcal{G}/\mathcal{K})$ one has*

$$(2.3.5.6) \quad \frac{1}{2\pi} \int_{\mathcal{G}} f(g) d\mu_{\text{IW}}(g) = \int_{\mathcal{A}} f(b) d\mu_{\mathbf{P}}(b).$$

PROOF. It follows from **Corollary 2.3.11** and the fact the **Gauss**(\mathbb{R}) is dense in $C_c(\mathbb{R})$. \square

Having two natural measures, the Iwasawa measure and the polar measure, we get two normalizations of the spherical transform, which we denote by

$$(2.3.5.7) \quad \mathbf{S}_{\text{IW}} \text{ and } \mathbf{S}_{\mathbf{P}} \text{ so that } \mathbf{S}_{\mathbf{P}} = \frac{1}{2\pi} \mathbf{S}_{\text{IW}}.$$

The \mathbf{S} of proceeding sections is \mathbf{S}_{IW} . We may call $\mathbf{S}_{\mathbf{P}}$ the **polar normalized spherical transform**, or simply the **polar spherical transform**.

We let

$$(2.3.5.8) \quad \Phi_{\mathbf{P}}(\zeta, a) := \frac{\mathcal{B}^{\vee}(\alpha, \alpha) a^{\zeta} - a^{-\zeta}}{\mathcal{B}_{\mathbb{C}}^{\vee}(\alpha, \zeta) a^{\alpha} - a^{-\alpha}} = \frac{2}{\mathcal{B}_{\mathbb{C}}^{\vee}(\alpha, \zeta)} \frac{a^{\zeta} - a^{-\zeta}}{a^{\alpha} - a^{-\alpha}}.$$

Writing $\zeta = s\alpha$ with $s \in \mathbb{C}$, we can write this formula in the form

$$(2.3.5.9) \quad \Phi_{\mathbf{P}, \alpha}(s, a) := \Phi_{\mathbf{P}}(s\alpha, a) = \frac{1}{s} \frac{y^s - y^{-s}}{y - y^{-1}} = \frac{1}{s} \frac{e^{sv} - e^{-sv}}{e^v - e^{-v}} = \frac{1}{s} \frac{\sinh(sv)}{\sinh v}.$$

Theorem 2.3.15. *For all basic \mathcal{G} -Gaussians, taking the polar measure convolution $*_{\mathcal{A}, \mathbf{P}}$ on \mathcal{A} , we have*

$$(2.3.5.10) \quad (\Phi_{\mathbf{P}} *_{\mathcal{A}, \mathbf{P}} \varphi_c)(s\alpha) = (\mathbf{S}_{\mathbf{P}} \varphi_c)(s\alpha)$$

so $\Phi_{\mathbf{P}}$ is an integral kernel for $\mathbf{S}_{\mathbf{P}}$ (with polar measure).

PROOF. By definition, (2.3.4.1) and, (2.3.5.2),

$$\begin{aligned} (\Phi_{\mathbf{P}} *_{\mathcal{A}, \mathbf{P}} \varphi_c)(s\alpha) &= \int_{\mathcal{A}} \Phi_{\mathbf{P}}(s\alpha, a) \varphi_c(a) d\mu_{\mathbf{P}}(a) \\ &= \frac{1}{s} \int_{-\infty}^{\infty} \frac{e^{sv} - s^{-sv}}{e^v - e^{-v}} e^{-v^2/2c} \frac{2v}{e^v - e^{-v}} \left(\frac{e^v - e^{-v}}{2} \right)^2 dv \\ &= \frac{1}{s} \int_{-\infty}^{\infty} \left(\frac{e^{sv} - e^{-sv}}{2} \right) e^{-v^2/2c} v dv = \frac{1}{s} \int_{-\infty}^{\infty} e^{sv} e^{-v^2/2c} v dv = (\mathbf{S}_{\mathbf{P}} \varphi_c)(s\alpha) \end{aligned}$$

where we used (2.3.5.5), (2.3.4.4), and (2.3.5.7). \square

The spherical kernel (2.3.5.8) in polar coordinates on $\mathbf{SL}_2(\mathbb{C})$ is originally due to **Gelfand** and **Naimark** (1950).

We observe that the polar spherical kernel has a product structure. The first factor in (2.3.5.8) has a name, the **Harish-Chandra c-function**, defined for our purpose by

$$(2.3.5.11) \quad \mathbf{c}(\zeta) := \frac{\mathcal{B}^{\vee}(\alpha, \alpha)}{\mathcal{B}_{\mathbb{C}}^{\vee}(\alpha, \zeta)} = \frac{1}{s} \text{ if } \zeta = s\alpha.$$

The second factor in (2.3.5.8) comes from a skew-symmetrization procedure, namely,

$$\frac{a^\zeta - a^{-\zeta}}{a^\alpha - a^{-\alpha}}.$$

We know from [Corollary 2.3.11](#) that the spherical transform is a linear isomorphism between the Gauss spaces. We want to exhibit the measure on \mathfrak{a}^\vee with respect to which the transpose of the spherical kernel induces the inverse integral transform. As to the Haar measure on \mathfrak{a}^\vee , we have the Haar measure $da = dy/y = dv$ on $\mathfrak{a}^\vee \cong \mathbb{R}$. We let $d\mu_{\text{FOU}}(\lambda)$ be the Haar measure on \mathfrak{a}^\vee that makes Fourier inversion come out without any extra constant factor. The **Fourier transform F** is defined by

$$(2.3.5.12) \quad (\mathbf{F}f)(\lambda) := \int_{\mathfrak{a}} f(H) e^{-i\lambda(H)} dH, \quad \lambda \in \mathfrak{a}^\vee.$$

The isomorphism of \mathfrak{a} with \mathbb{R} is taken via the coordinate v . That is,

$$\begin{array}{ccc} \mathcal{A} & \xrightarrow{\exp^{-1}} & \mathfrak{a} \\ & & \downarrow y \\ & & \mathbb{R}^+ \\ & & \downarrow \ln \\ & & \mathbb{R} \end{array} \quad \begin{array}{ccc} & & \searrow v \\ & & \mathbb{R} \end{array} \quad \begin{array}{ccc} a = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} & \longrightarrow & H = \begin{bmatrix} h_1 & 0 \\ 0 & h_2 \end{bmatrix} \\ & & \downarrow \\ & & a_1/a_2 \end{array} \quad \begin{array}{ccc} & & \searrow v \\ & & h_1 - h_2 \\ & \longrightarrow & \end{array}$$

Thus if $\lambda = r\alpha$ on \mathfrak{a} with $\lambda \in \mathbb{R}$, then

$$(2.3.5.13) \quad (\mathbf{F}_{\alpha,+}f)(r) \equiv (\mathbf{F}f)(\lambda) = \int_{\mathfrak{a}} f(H) e^{-ir\alpha(H)} dH = \int_{\mathbb{R}} f_{\alpha,+}(v) e^{-irv} dv.$$

The **Fourier normalized measure** is then defined to be

$$(2.3.5.14) \quad d\mu_{\text{FOU}}(\lambda) := \frac{dr}{2\pi}.$$

We define the **Harish-Chandra measure** on \mathfrak{a}^\vee by

$$(2.3.5.15) \quad d\mu_{\text{HC}}(\lambda) := \frac{1}{|\mathbf{c}(\mathbf{i}\lambda)|^2} d\mu_{\text{FOU}}(\lambda) = \frac{r^2 dr}{2\pi}, \quad \lambda = s\alpha.$$

We let $\Phi_{\mathbf{P},\text{HC}}$ be the integral operator having the kernel function $\Phi_{\mathbf{P}}(\mathbf{i}\lambda, a)$ with respect to the Harish-Chandra measure $d\mu_{\text{HC}}(\lambda)$, so for an even function h such that the integral is absolutely convergent, by definition

$$(2.3.5.16) \quad \begin{aligned} (\Phi_{\mathbf{P},\text{HC}}h)(a) &:= \int_{\mathfrak{a}^\vee} \Phi_{\mathbf{P}}(\mathbf{i}\lambda, a) h(\mathbf{i}\lambda) d\mu_{\text{HC}}(\lambda) \\ &= \int_{\mathbb{R}} h_{\alpha}(\mathbf{i}r) (-\mathbf{i}r) \frac{e^{irv} - e^{-irv}}{e^v - e^{-v}} \frac{dr}{2\pi} \quad \text{with } v = \alpha(\ln a). \end{aligned}$$

Theorem 2.3.16. *The integral operator $\Phi_{\mathbf{P},\text{HC}}$ is the inverse of $\mathbf{S}_{\mathbf{P}}$ on the Gauss spaces*

$$\mathbf{S}_{\mathbf{P}} : \mathbf{Gauss}(\mathcal{G}) \longrightarrow \mathbf{Gauss}(\mathbf{i}\mathbb{R}).$$

The map $\mathbf{S}_{\mathbf{P}}$ is an L^2 -isometry with respect to the polar measure on \mathcal{G} and the Harish-Chandra measure on \mathfrak{a}^\vee , the integral scalar product taken on \mathfrak{a}^\vee .

PROOF. For given $\varphi_c \in \mathbf{Gauss}(\mathcal{G})$, let $h := \mathbf{S_P}\varphi_c$. By (2.3.4.5),

$$h_{\mathbf{a}}(\mathbf{ir}) = (2\pi)^{1/2}c^{3/2}e^{-r^2c/2}.$$

Then

$$\begin{aligned} (\Phi_{\mathbf{P,HC}}h)(a) &= \int_{\mathbb{R}} (2\pi)^{1/2}c^{3/2}e^{-r^2c/2}(-\mathbf{ir}) \frac{e^{\mathbf{ir}v} - e^{-\mathbf{ir}v}}{e^v - e^{-v}} \frac{dr}{2\pi} \\ &= \int_{\mathbb{R}} \frac{(2\pi)^{-1/2}c^{3/2}}{\sinh v} (-\mathbf{ir}) e^{-r^2c/2 + \mathbf{ir}v} dr \\ &= e^{-v^2/2c} \frac{v}{\sinh v} = \varphi_c(a). \end{aligned}$$

Thus $(\Phi_{\mathbf{P,HC}} \circ \mathbf{S_P})\varphi_c = \varphi_c$.

To prove the second statement, we first make explicit the L^2 -scalar product on \mathcal{G} , for the polar measure. For two basic Gaussians $\varphi_c, \varphi_{c'}$, we get, integrating over $\mathbf{a} \cong \mathbb{R}$:

$$\begin{aligned} \langle \varphi_c, \varphi_{c'} \rangle_{\mathbf{P}} &= \int_{\mathbb{R}} e^{-v^2/2c} \frac{2v}{e^v - e^{-v}} e^{-v^2/2c'} \frac{2v}{e^v - e^{-v}} \left(\frac{e^v - e^{-v}}{2} \right)^2 dv \\ &= \int_{\mathbb{R}} e^{-v^2(c+c')/2} v^2 dv = (cc')^{3/2} \int_{\mathbb{R}} e^{-v^2(c+c')/2} v^2 dv. \end{aligned}$$

On the other hand,

$$\begin{aligned} \langle \mathbf{S_P}\varphi_c, \mathbf{S_P}\varphi_{c'} \rangle_{\mathbf{HC}} &= \int_{\mathbb{R}} (\mathbf{S_P}\varphi_c)(\mathbf{ir}\mathbf{a})(\mathbf{S_P}\varphi_{c'})(\mathbf{ir}\mathbf{a}) \frac{r^2 dr}{2\pi} \\ &= \int_{\mathbb{R}} (2\pi)^{1/2}c^{3/2}e^{-r^2c/2} (2\pi)^{1/2}c'^{3/2}e^{-r^2c'/2} \frac{r^2 dr}{2\pi} \\ &= (cc')^{3/2} \int_{\mathbb{R}} e^{-r^2(c+c')/2} r^2 dr = \left(\frac{2cc'}{c+c'} \right)^{3/2} \Gamma(3/2). \end{aligned}$$

Thus $\langle \varphi_c, \varphi_{c'} \rangle_{\mathbf{P}} = \langle \mathbf{S_P}\varphi_c, \mathbf{S_P}\varphi_{c'} \rangle_{\mathbf{HC}}$. \square

Jorgenson and **Lang** (2005) showed that the above isometry can be extended by continuity to $L^2(\mathcal{K} \setminus \mathcal{G}/\mathcal{K})$ because of the denseness of the Gauss space.

2.3.6. Point-pair invariants, the polar height, and the polar distance. Let $\varphi \in \mathbf{Fu}(\mathcal{K} \setminus \mathcal{G}/\mathcal{K})$ be a \mathcal{K} -bi-invariant function on \mathcal{G} . Following Selberg, we define its associated **point-pair invariant** \mathbf{K}_φ to be the function defined by

$$(2.3.6.1) \quad \mathbf{K}_\varphi(z, w) := \varphi(z^{-1}w), \quad z, w \in \mathcal{G}.$$

The \mathcal{K} -bi-invariance of φ implies that \mathbf{K}_φ is defined on $\mathcal{G}/\mathcal{K} \times \mathcal{G}/\mathcal{K}$.

- (1) \mathbf{K}_φ is symmetric, i.e., $\mathbf{K}_\varphi(z, w) = \mathbf{K}_\varphi(w, z)$, because φ is even.

In higher dimensions, one has to assume the evenness condition in addition to the \mathcal{K} -bi-invariance.

- (2) \mathbf{K}_φ is \mathcal{G} -invariant, i.e., $\mathbf{K}_\varphi(gz, gw) = \mathbf{K}_\varphi(z, w)$.

A \mathcal{K} -bi-invariant function on \mathcal{G} is determined by its values on \mathcal{A} by the polar decomposition $\mathcal{G} = \mathcal{K}\mathcal{A}\mathcal{K}$. If $g = k_1ak_2$, then b is uniquely determined up to a permutation of its diagonal elements. Hence a function on \mathcal{A} that is invariant under permutations extends to a \mathcal{K} -bi-invariant function on \mathcal{G} .

Recall that the polar height σ is the \mathcal{K} -bi-invariant function such that $\sigma \geq 0$ and

$$\sigma^2(g) = \sigma^2(b) = (\ln \mathbf{y}(b))^2 = (\ln b^\alpha)^2,$$

where $g = k_1 b k_2$ is the polar decomposition of $g \in \mathcal{G}$.

Proposition 2.3.17. *Let $g = uak = k_1 b k_2$ be the Iwasawa, respectively polar, decomposition. Then*

$$(2.3.6.2) \quad \sigma(a) \leq \sigma(b) \quad \text{or} \quad |\ln a| \leq |\ln b|.$$

In general reductive Lie group theory, this result is due to **Harish-Chandra**. The inequality (2.3.6.2) can be viewed from a differential geometric point of view. The space \mathcal{G}/\mathcal{K} is a Cartan-Hadamard manifold, whose exponential map is metric semi-increasing. The \mathcal{U} -components are perpendicular to \mathcal{A} . The metric increasing property shows that the law of cosines with the inequality in the right direction is valid on \mathcal{G}/\mathcal{K} , and the desired inequality (2.3.6.2) then follows at once.

PROOF. Let e_1, e_2 be the unit column vectors of \mathbb{C}^2 . Let $a = \text{diag}(a_1, a_2)$ and $b = \text{diag}(b_1, b_2)$. We may without loss of generality that $a_1, b_1 \geq 1$. It suffices to prove that $a_1 \leq b_1$. From $\mathcal{U}\mathcal{A} = \mathcal{A}\mathcal{U}$, we have $ua = av$ for some $v \in \mathcal{U}$. Writing $av = ua = k_1 b k_2 k^{-1} = k_1 b k'$ with $k' := k_2 k^{-1} \in \mathcal{K}$, we have

$$a_1^2 = |ave_1|^2 = |k_1 b k' e_1|^2,$$

where $|\cdot|$ stands for the Euclidean metric on \mathbb{C}^2 . Since k' is unitary, there are numbers $c_1, c_2 \in \mathbb{C}$ such that $k' e_1 = c_1 e_1 + c_2 e_2$ and $|c_1|^2 + |c_2|^2 = 1$. Then

$$|b k' e_1|^2 = |b_1 c_1|^2 + |b_2 c_2|^2 = b_1^2 |c_1|^2 + b_2^2 |c_2|^2.$$

It suffices to show that for $y \geq 1$ ($y := b_1^2$) we have

$$y|c_1|^2 + y^{-1}|c_2|^2 \leq y.$$

Using $|c_2|^2 = 1 - |c_1|^2$ yields

$$y|c_1|^2 + y^{-1}|c_2|^2 = \frac{|c_1|^2(y^2 - 1) + 1}{y} \leq \frac{(y^2 - 1) + 1}{y} = y$$

which proves the inequality (2.3.6.2). \square

The following deep property is due to **Cartan** and **Harish-Chandra**.

Theorem 2.3.18. (Triangle inequality) *For $g, g' \in \mathcal{G}$, we have*

$$(2.3.6.3) \quad \sigma(gg') \leq \sigma(g) + \sigma(g').$$

Lang (1999) provided an elementary self-contained treatment of **Theorem 2.3.18** on Bruhat-Tits spaces.

Definition 2.3.19. We then define the **polar distance**

$$(2.3.6.4) \quad d_{\mathbf{p}}(z, w) := \sigma(z^{-1}w), \quad z, w \in \mathcal{G},$$

which is the associated point-pair invariant. According to **Theorem 2.3.18**, $d_{\mathbf{p}}$ is a distance defined on \mathcal{G}/\mathcal{K} .

Recall from [Theorem 2.2.1](#) that $\mathcal{G}/\mathcal{K} \approx \mathbf{Pos}_2$ by $g \mapsto gg^*$ (with $g^* = \bar{g}$). Let

$$\mathbf{y}_{\mathbf{P}}(g) = b^\alpha = \frac{b_1}{b_2} = b_1^2$$

be the polar y -coordinate of $g = k_1bk_2 \in \mathcal{G}$. Then

$$(2.3.6.5) \quad \sigma(g) = |\ln \mathbf{y}_{\mathbf{P}}(g)| = d_{\mathbf{P}}(g\mathcal{K}, e_{\mathcal{G}/\mathcal{K}}),$$

where $e_{\mathcal{G}/\mathcal{K}}$ is the unit coset of \mathcal{G}/\mathcal{K} .

Theorem 2.3.20. *The function $d_{\mathbf{P}}$ is a distance function on \mathcal{G}/\mathcal{K} .*

PROOF. [Lang](#) (1999) defined a distance on $\mathbf{Pos}_2(\mathbb{R})$. The analogous treatment for $\mathbf{Pos}_n(\mathbb{C})$ and $\mathbf{SPos}_n(\mathbb{C})$ runs exactly the same way. Thus we get a natural \mathcal{G} -invariant distance on \mathcal{G}/\mathcal{K} , which is a Riemannian manifold, with \mathcal{G} acting by translation as a group of isometries. \square

Fundamental groups

3.1. Homotopy

Recall the category **Top** of topological spaces and continuous maps. For any $f, g \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$, a **homotopy from f to g** is a continuous map $H : X \times [0, 1] \rightarrow Y$, $(x, t) \mapsto H(x, t) \equiv H_t(x)$, such that

$$H_0 = f, \quad H_1 = g.$$

In this case, we also say that **f and g are homotopic** and write $f \simeq g$ or $H : f \simeq g$. If f is homotopic to a constant map, we say f is **null homotopic**.

3.1.1. Homotopy relation. It is clear that $f \simeq f$. Actually, we can prove that homotopy is an equivalence relation.

Proposition 3.1.1. (1) *Homotopy is an equivalence relation on $\mathbf{Hom}_{\mathbf{Top}}(X, Y)$.*
 (2) *Homotopy relation is preserved by composition. That is, if*

$$f_0, f_1 \in \mathbf{Hom}_{\mathbf{Top}}(X, Y), \quad g_0, g_1 \in \mathbf{Hom}_{\mathbf{Top}}(Y, Z),$$

and $f_0 \simeq f_1$ and $g_0 \simeq g_1$, then $g_0 \circ f_0 \simeq g_1 \circ f_1$.

PROOF. (1) If $H : f \simeq g$, then $H^{-1} : g \simeq f$, where

$$\bar{H}(x, t) := H(x, 1 - t), \quad (x, t) \in X \times [0, 1].$$

If $F : f \simeq g$ and $G : g \simeq h$, then $H : f \simeq h$, where

$$H(x, t) := \begin{cases} F(x, 2t), & 0 \leq t \leq \frac{1}{2}, \\ G(x, 2t - 1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Since $F(x, 1) = g(x) = G(x, 0)$, the map H is well-defined and hence, by **Theorem 1.2.26**, H is continuous.

(2) Let $F : f_0 \simeq f_1$ and $G : g_0 \simeq g_1$. Then we define $H_t := G_t \circ F_t$ which implies that $H : g_0 \circ f_0 \simeq g_1 \circ f_1$. \square

Because \simeq is an equivalence relation by **Proposition 3.1.1**, we denote by $[X, Y]$ the set of equivalence classes of continuous maps from X to Y , that is,

$$(3.1.1.1) \quad [X, Y] := \mathbf{Hom}_{\mathbf{Top}}(X, Y) / \simeq.$$

3.1.2. The fundamental group. Recall the category \mathbf{Top}_* consisting of pointed topological space (X, x) and point-preserving continuous maps. For any $(X, x) \in \mathbf{Ob}(\mathbf{Top}_*)$, denote by $\Omega(X, x)$ the set of all loops in X based at x , that is,

$$(3.1.2.1) \quad \Omega(X, x) := \{f \in [0, 1] \rightarrow X \mid f \text{ is continuous and } f(0) = f(1) = x\}.$$

The set $\Omega(X, x)$ consists at least one element c_x given by

$$(3.1.2.2) \quad c_x(s) := x, \quad s \in [0, 1].$$

For $f \in \Omega(X, x)$, define its **reverse path** \bar{f} by

$$\bar{f}(s) := f(1 - s), \quad s \in [0, 1].$$

Then $\bar{f} \in \Omega(X, x)$. We say $f, g \in \Omega(X, x)$ is **path homotopic** and write $f \sim g$, if there exists a continuous map $G : [0, 1] \times [0, 1] \rightarrow X$, $(s, t) \mapsto H(s, t) = H_t(s)$, such that

$$H_0 = f, \quad H_1 = g, \quad H_t(0) = H_t(1) = x \quad (t \in [0, 1]).$$

We also say that H is a **path homotopy from f to g** and write $H : f \sim g$.

Proposition 3.1.2. \sim is an equivalence relation on $\Omega(X, x)$.

PROOF. The proof is almost the same as **Proposition 3.1.1**. □

We use the above proposition to define the **fundamental group of X based at x** as

$$(3.1.2.3) \quad \pi_1(X, x) := \Omega(X, x) / \sim = \{[f] \mid f \in \Omega(X, x)\},$$

the set of path classes of loops based at x .

For $f, g \in \Omega(X, x)$, we can define their **product** $f * g : [0, 1] \rightarrow X$ by

$$(3.1.2.4) \quad f * g(s) := \begin{cases} f(2s), & 0 \leq s \leq \frac{1}{2}, \\ g(2s - 1), & \frac{1}{2} \leq s \leq 1. \end{cases}$$

For $[f], [g] \in \pi_1(X, x)$, we can also define

$$(3.1.2.5) \quad [f] * [g] := [f * g].$$

Firstly, we shall verify that the above definition (3.1.2.5) is well-defined, that is, if $F : f \sim f'$ and $G : g \sim g'$, then $H : f * g \sim f' * g'$. Indeed, define

$$H(s, t) = \begin{cases} F(2s, t), & 0 \leq s \leq \frac{1}{2}, 0 \leq t \leq 1, \\ G(2s - 1, t), & \frac{1}{2} \leq s \leq 1, 0 \leq t \leq 1. \end{cases}$$

Theorem 3.1.3. For any $(X, x) \in \mathbf{Ob}(\mathbf{Top}_*)$, $(\pi_1(X, x), *)$ is a group.

PROOF. We shall verify

(1) $[c_x]$ is the identity element in $\pi_1(X, x)$:

$$[c_x] * [f] = [f] * [c_x] = [f].$$

(2) The inverse of $[f]$ is $[\bar{f}]$:

$$[f] * [\bar{f}] = [\bar{f}] * [f] = [c_x].$$

(3) $*$ is associative:

$$[f] * ([g] * [h]) = ([f] * [g]) * [h].$$

For (1): we have $H : f \sim c_x * f$, where

$$H(s, t) = \begin{cases} x, & t \geq 2s, \\ f\left(\frac{2s-t}{2-t}\right), & t \leq 2s. \end{cases}$$

Also, we have $G : f \sim f * c_x$, where

$$G(s, t) = \begin{cases} f\left(\frac{2s}{2-t}\right), & t \leq 2 - 2s, \\ x, & t \geq 2 - 2s. \end{cases}$$

For (2): we have $H : c_x \sim f * \bar{f}$, where

$$H(s, t) = \begin{cases} f(2s), & 0 \leq s \leq \frac{t}{2}, \\ f(t), & \frac{t}{2} \leq s \leq 1 - \frac{t}{2}, \\ f(2 - 2s), & 1 - \frac{t}{2} \leq s \leq 1. \end{cases}$$

Also, we have $G : c_x \sim \bar{f} * f$, where

$$H(s, t) = \begin{cases} f(1 - 2s), & 0 \leq s \leq \frac{t}{2}, \\ f(1 - t), & \frac{t}{2} \leq s \leq 1 - \frac{t}{2}, \\ f(2s - 1), & 1 - \frac{t}{2} \leq s \leq 1. \end{cases}$$

For (3): we have $H : (f * g) * h \sim f * (g * h)$. \square

Exercise 3.1.4. Prove $(f * g) * h \sim f * (g * h)$ in Theorem 3.1.3.

Suppose now that X is path-connected. For any $x, y \in X$ we can define a natural map

$$(3.1.2.6) \quad \Phi_g : \pi_1(X, x) \longrightarrow \pi_1(X, y), \quad [f] \longmapsto [\bar{g} * f * g] = [\bar{g}] * [f] * [g],$$

where g is a path from x to y .

Proposition 3.1.5. *If X is path-connected, then the map (3.1.2.6) is a group isomorphism with inverse $\Phi_{\bar{g}}$.*

PROOF. For any $[f_1], [f_2] \in \pi_1(X, x)$ we have

$$\begin{aligned} \Phi_g([f_1]) * \Phi_g([f_2]) &= ([\bar{g}] * [f_1] * [g]) * ([\bar{g}] * [f_2] * [g]) \\ &= [\bar{g}] * [f_1] * ([g] * [\bar{g}]) * [f_2] * [g] = [\bar{g}] * [f_1] * [f_2] * [g] = \Phi_g([f_1] * [f_2]). \end{aligned}$$

Hence Φ_g is a group homomorphism. To prove that Φ_g is an isomorphism we notice that

$$\Phi_g \circ \Phi_{\bar{g}}([f]) = \Phi_g([\bar{g}] * [f] * [g]) = [\bar{g}] * ([g] * [f] * [\bar{g}]) * [g] = [f].$$

Similarly one has $\Phi_{\bar{g}} \circ \Phi_g([f]) = [f]$. \square

As a consequence of **Theorem 3.1.5**, for any path-connected topological space X and any two points $x, y \in X$, we have that $\pi_1(X, x)$ is isomorphic to $\pi_1(X, y)$. Hence we can use the notation “ $\pi_1(X)$ ” to denote “the” group $\pi_1(X, x)$, relative to any point $x \in X$.

We say a path-connected topological space is **simply-connected** if $\pi_1(X, x) = \{[c_x]\}$ (that is, $\pi_1(X)$ is trivial).

Example 3.1.6. (1) $\pi_1(\mathbf{R}^n)$ is trivial. For any fixed point $x \in \mathbf{R}^n$ and any $f \in \Omega(\mathbf{R}^n, x)$, we have

$$f \sim c_x,$$

since

$$H(s, t) = (1 - t)f(s) + tc_x(s), \quad (s, t) \times [0, 1] \times [0, 1],$$

is a homotopic from f to c_x . Hence $[f] = [c_x]$ for any $f \in \Omega(\mathbf{R}^n, x)$ and then $c_1(\mathbf{R}^n) = \{[c_x]\}$.

(2) For any convex domain $X \subseteq \mathbf{R}^n$, the fundamental group $c_1(X, x)$ is trivial. Indeed, for any $f \in \Omega(X, x)$ define H as in (1), so that $f \sim c_x$.

(3) $\pi_1(\mathbf{S}^n)$ is trivial, for any $n \geq 2$.

Consider a morphism $\varphi \in \mathbf{Hom}_{\mathbf{Top}_*}((X, x), (Y, y))$. Then $\varphi : X \rightarrow Y$ is a continuous map with $\varphi(x) = y$. Define

$$(3.1.2.7) \quad \varphi_* : \pi_1(X, x) \longrightarrow \pi_1(Y, y), \quad [f] \longmapsto [\varphi \circ f].$$

The same argument as for **Proposition 3.1.5** shows that φ_* is a group homomorphism. We call φ_* the **homomorphism induced by φ** .

The above discussion lets us to define a functor

$$(3.1.2.8) \quad \pi_1 : \mathbf{Top}_* \longrightarrow \mathbf{Group}, \quad (X, x) \longmapsto \pi_1(X, x),$$

with $\pi_1(\varphi) = \varphi_*$.

Theorem 3.1.7. π_1 defined in (3.1.2.8) is a functor from \mathbf{Top}_* to \mathbf{Group} .

PROOF. We shall verify the following two things:

(1) If $\varphi \in \mathbf{Hom}_{\mathbf{Top}_*}((X, x), (Y, y))$ and $\psi \in \mathbf{Hom}_{\mathbf{Top}_*}((Y, y), (Z, z))$, then

$$(\psi \circ \varphi)_* = \psi_* \circ \varphi_*.$$

(2) For any $x \in X$, we have

$$(\mathbf{1}_X)_* = \mathbf{1}_{\pi_1(X, x)}.$$

For (1),

$$(\psi \circ \varphi)_*([f]) = [\psi \circ \varphi \circ f] = \psi_*([\varphi \circ f]) = \psi_*(\varphi_*([f])).$$

For (2),

$$(\mathbf{1}_X)_*([f]) = [\mathbf{1}_X \circ f] = [f] = \mathbf{1}_{\pi_1(X, x)}([f]).$$

Consequently, π_1 is a functor. \square

Corollary 3.1.8. If $\varphi : (X, x) \rightarrow (Y, y)$ is a homeomorphism in \mathbf{Top}_* , then the induced homomorphism $\varphi_* : \pi_1(X, x) \rightarrow \pi_1(Y, y)$ is a group isomorphism.

3.1.3. Fundamental groups of manifolds. Recall the Lebesgue number lemma, [Theorem 1.5.11](#).

Theorem 3.1.9. *The fundamental group of a topological manifold is countable.*

PROOF. Let \mathcal{M} be a topological m -manifold and take any fixed point $x \in \mathcal{M}$. To prove $\pi_1(\mathcal{M}, x)$ is countable, we shall analyze any loop $f : [0, 1] \rightarrow \mathcal{M}$ at x .

(1) According to [Proposition 1.5.18](#), we can choose a countable open cover \mathcal{U} of \mathcal{M} in such a way that each element $\mathcal{U} \in \mathcal{U}$ is isomorphic to some open ball in \mathbf{R}^m (so that \mathcal{U} is path-connected). For any given $\mathcal{U}, \mathcal{U}' \in \mathcal{U}$, the intersection $\mathcal{U} \cap \mathcal{U}'$ has at most countably many connected components $(\mathcal{U} \cap \mathcal{U}')_{[i]}$, $i \in I(\mathcal{U}, \mathcal{U}') \subseteq \mathbf{N}$; otherwise \mathcal{U} is uncountable. For each $i \in I(\mathcal{U}, \mathcal{U}')$, choose a point $x_i \in (\mathcal{U} \cap \mathcal{U}')_{[i]}$. Consider the set

$$\Delta := \left\{ x_i \in (\mathcal{U} \cap \mathcal{U}')_{[i]} \mid i \in I(\mathcal{U}, \mathcal{U}'), \mathcal{U}, \mathcal{U}' \in \mathcal{U} \right\}.$$

Because \mathcal{U} and $I(\mathcal{U}, \mathcal{U}')$ are countable, we see that the set Δ is also countable.

(2) For each $\mathcal{U} \in \mathcal{U}$ and each pair (x, x') in Δ with $x, x' \in \mathcal{U}$, we can take a path $h_{x, x'}^{\mathcal{U}}$ from x to x' in \mathcal{U} , since \mathcal{U} is path-connected. Applying [Theorem 1.5.11](#) to the open cover

$$\mathcal{U}_f := f^{-1}(f([0, 1]) \cap \mathcal{U}) = \left\{ f^{-1}(\mathcal{U} \cap f([0, 1])) \right\}$$

of the compact interval $[0, 1]$, we obtain a number $\delta > 0$ such that any subset $A \subseteq [0, 1]$ with diameter $\text{diam}(A) < \delta$ is contained in some element of \mathcal{U}_f . In particular, taking $n \in \mathbf{N}$ with $n > 1/\delta$ yields that $f([(k-1)/n, k/n]) \subseteq \mathcal{U}_k \in \mathcal{U}$. Write

$$f_k(s) := f|_{[(k-1)/n, k/n]} \left(\frac{s+k-1}{n} \right), \quad s \in [0, 1].$$

Then $[f] = [f_1] \cdots [f_n]$.

(3) For each $k = 1, \dots, n-1$, take a point $x_k \in \mathcal{U}_k \cap \mathcal{U}_{k+1}$. Since $f(k/n) \in \mathcal{U}_k \cap \mathcal{U}_{k+1}$ and $\mathcal{U}_k \cap \mathcal{U}_{k+1}$ is path-connected, it follows that we can choose a path g_k in $\mathcal{U}_k \cap \mathcal{U}_{k+1}$ from x_k to $f(k/n)$. Let

$$f'_k := g_{k-1} * f_k * \overline{g}_k.$$

For $k = 0$ or $k = n$, take $x_k = p$ and $g_k = c_x$ the constant path. Therefore

$$[f] = [f'_1] * \cdots * [f'_n].$$

Now f'_k and $h_{x_{k-1}, x_k}^{\mathcal{U}_k}$ both connect x_{k-1} to x_k in \mathcal{U}_k , using the simple-connectedness of \mathcal{U}_k , we see that f'_k is homotopic to $h_{x_{k-1}, x_k}^{\mathcal{U}_k}$. Consequently

$$[f] = \left[h_{x_0, x_1}^{\mathcal{U}_1} \right] * \cdots * \left[h_{x_{n-1}, x_n}^{\mathcal{U}_n} \right].$$

Hence $\pi_1(\mathcal{M}, x)$ is countable. \square

The above argument also shows that the fundamental group of a compact manifold is finite.

To improve the result in [Theorem 3.1.9](#) we shall assume some curvature conditions.

Theorem 3.1.10. (Bonnet-Myers) *If (\mathcal{M}, g) is a complete Riemannian m -manifold with $\text{Ric}_g \geq (m-1)Kg$, where $K > 0$, then $\text{diam}(\mathcal{M}, g) \leq \pi/\sqrt{K}$. In particular, \mathcal{M} is compact and $\pi_1(\mathcal{M}) < +\infty$.*

PROOF. We will give a proof in the course of Differential Manifold. \square

If the Ricci curvature bound is replaced by the sectional sectional curvature bound, we have the following outstanding result.

Theorem 3.1.11. (Cheeger-Gromoll soul theorem) *If (\mathcal{M}, g) is a complete Riemannian m -manifold with $\text{Sec}_g > 0$ everywhere, then \mathcal{M} is diffeomorphic to \mathbf{R}^m .*

In 1994, **Perelman** proved the so-called **soul conjecture**.

Theorem 3.1.12. (Soul conjecture, solved by Perelman) *If (\mathcal{M}, g) is a complete Riemannian m -manifold with $\text{Sec}_g \geq 0$ everywhere and > 0 at some point, then \mathcal{M} is diffeomorphic to \mathbf{R}^m .*

For nonnegative Ricci curvature we have the following well-known conjecture.

Conjecture 3.1.13. (Milnor's conjecture) *If (\mathcal{M}, g) is a complete Riemannian m -manifold with $\text{Ric}_g \geq 0$, then $\pi_1(\mathcal{M})$ is finite generated.*

When $m = 3$, the above conjecture was recently solved by **Gang Liu** under the previous work of Schoen-Yau. Under an extra condition that g has maximum volume growth, **Peter Li** proved **Conjecture 3.1.13** in 1986 (actually, he proved that $\pi_1(\mathcal{M})$ is finite). Recall that if $\text{Ric}_g \geq 0$, then the volume ratio

$$\frac{\text{Vol}_g(\mathcal{B}_g(x, r))}{\omega_m r^m}$$

is nonincreasing. Then

$$\lim_{r \rightarrow \infty} \frac{\text{Vol}_g(\mathcal{B}_g(x, r))}{\omega_m r^m} = \theta \in [0, 1].$$

We say the Riemannian metric g has **maximum volume growth**, if $\theta \in (0, 1]$. When $\theta = 1$, we can prove that in this case (\mathcal{M}, g) is isometric to \mathbf{R}^m .

On the other hand, a simply-connected manifold with suitable curvature conditions has a rigidity result.

Theorem 3.1.14. (Rauch-Klingenberg-Berger topological sphere theorem) *If (\mathcal{M}, g) is a complete, simply-connected Riemannian n -manifold with*

$$\frac{1}{4} \leq \text{Sec}_g(\Pi) \leq 1$$

for all 2-planes Π , then \mathcal{M} is homeomorphic to \mathbf{S}^n . In particular, \mathcal{M}^3 is diffeomorphic to \mathbf{S}^3 .

How about the diffeomorphisms between \mathcal{M}^m and \mathbf{S}^m when the dimension $m \geq 4$? The answer is YES! It was proved by **Brendle** and **Schoen** by using the Ricci flow.

Theorem 3.1.15. (Diffeomorphic spherical theorem) *If (\mathcal{M}, g) is a complete, simply-connected Riemannian n -manifold with*

$$\frac{1}{4} \leq \text{Sec}_g(\Pi) \leq 1$$

for all 2-planes Π , then \mathcal{M} is diffeomorphic to \mathbf{S}^m .

3.1.4. Fundamental groups of product spaces. Let $\{(X_i, \mathcal{T}_i)\}_{1 \leq i \leq n}$ be a finite family of topological spaces, and let

$$(3.1.4.1) \quad p_i : X_1 \times \cdots \times X_n \longrightarrow X_i, \quad (x_1, \cdots, x_n) \longmapsto x_i$$

denote the i -th projection. Then we get group homomorphisms

$$(3.1.4.2) \quad p_{i,*} : \pi_1(X_1 \times \cdots \times X_n, (x_1, \cdots, x_n)) \longrightarrow \pi_1(X_i, x_i)$$

by (3.1.2.7). Consequently, we can define a group homomorphism

$$(3.1.4.3) \quad \mathbf{p}_* : \pi_1(X_1 \times \cdots \times X_n, (x_1, \cdots, x_n)) \longrightarrow \pi_1(X_1, x_1) \times \cdots \times \pi_1(X_n, x_n)$$

given by

$$(3.1.4.4) \quad \mathbf{p}_*([f]) := (p_{1,*}[f], \cdots, p_{n,*}[f]).$$

Theorem 3.1.16. *The map (3.1.4.4) is an isomorphism.*

PROOF. We shall prove the injectivity and surjectivity of \mathbf{p}_* .

(1) \mathbf{p}_* is injective. Suppose $\mathbf{p}_*([f]) = ([c_{x_1}], \cdots, [c_{x_n}])$. Write

$$f = (f_1, \cdots, f_n).$$

Then $p_i \circ f = f_i$ and

$$([c_{x_1}], \cdots, [c_{x_n}]) = ([f_1], \cdots, [f_n]).$$

Thus $f_i \sim c_{x_i}$, $1 \leq i \leq n$ and hence $f \sim c_{(x_1, \cdots, x_n)}$.

(2) \mathbf{p}_* is surjective. For $([f_1], \cdots, [f_n]) \in \pi_1(X_1, x_1) \times \cdots \times \pi_1(X_n, x_n)$, define

$$f(s) := (f_1(s), \cdots, f_n(s)), \quad s \in [0, 1].$$

Then $\mathbf{p}_*([f]) = ([f_1], \cdots, [f_n])$. □

3.1.5. Retractions and deformation retractions. Let (X, \mathcal{T}) be a topological space and $A \subseteq X$ a subset. Under the subspace topology $\mathcal{T}_A := \mathcal{T} \cap A$, we see that A is itself a topological space.

- (1) A continuous map $r : X \rightarrow A$ is called a **retraction** if $r \circ \iota_A = \mathbf{1}_A$, where $\iota_A : A \rightarrow X$ is the inclusion. If there exists a retraction from X to A , we say that A is a **retract of X** .

It is clear that a retract of a connected space (resp., compact space) is connected (resp., compact). Moreover, if $A \subseteq B \subseteq X$, A is a retract of B , and B is a retract of X , then A is a retract of X .

- (2) A retraction $r : X \rightarrow A$ is said to be a **deformation retraction** if $\iota_A \circ r$ is homotopic to $\mathbf{1}_X$. If there exists a deformation retraction from X to A , then A is said to be a **deformation retract of X** . That is, $A \subseteq X$ is

a deformation retract of X if and only if there exists a homotopy $H : X \times [0, 1] \rightarrow X$ satisfying

$$\begin{aligned} H(x, 0) &\equiv H_0(x) = x, & x \in X, \\ H(x, 1) &\equiv H_1(x) \in A, & x \in X, \\ H(a, 1) &\equiv H_1(a) = a, & a \in A. \end{aligned}$$

Let $f \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$. A morphism $g \in \mathbf{Hom}_{\mathbf{Top}}(Y, X)$ is a **homotopy inverse for f** if

$$g \circ f \simeq \mathbf{1}_X, \quad f \circ g \simeq \mathbf{1}_Y.$$

If there exists a homotopy inverse for f , then f is called a **homotopy equivalence** and we say that X is **homotopic to Y** (write $X \simeq Y$).

Exercise 3.1.17. Homotopy equivalence is an equivalence relation on the class of all topological spaces $\mathbf{Ob}(\mathbf{Top})$.

If $r : X \rightarrow A$ is a deformation retraction, then, according to $r \circ \iota_A = \mathbf{1}_A$, r and ι_A are homotopy equivalences.

Proposition 3.1.18. (1) If $r : X \rightarrow A$ is a retraction, then for any $x \in A$, $\iota_{A,*} : \pi_1(X, x) \rightarrow \pi_1(A, x)$ is injective and $r_* : \pi_1(X, x) \rightarrow \pi_1(A, x)$ is surjective.

(2) A retract of a simply-connected space is also simply-connected.

PROOF. (1) From $r \circ \iota_A = \mathbf{1}_A$ we obtain $\mathbf{1}_{\pi_1(A, x)} = r_* \circ \iota_{A,*}$ which implies the desired results.

(2) Since $\pi_1(A, x)$ is trivial and $\iota_{A,*} : \pi_1(A, x) \rightarrow \pi_1(X, x)$ is injective, it follows that $\pi_1(X, x)$ is also trivial. \square

Example 3.1.19. (1) The map $r : \mathbf{R}^m \setminus \{0\} \rightarrow \mathbf{S}^{m-1}$ given by

$$r(x) := \frac{x}{|x|}$$

is a retraction. Since $\pi_1(\mathbf{S}^1) \cong \mathbf{Z}$ (we shall prove soon), it follows from Proposition 3.1.18 that $\mathbf{R}^2 \setminus \{0\}$ is not simply-connected. Consequently, $\mathbf{R}^2 \setminus \{0\}$ is not homeomorphic to \mathbf{R}^2 .

(2) Since the map $r : \mathbf{T}^2 = \mathbf{S}^1 \times \mathbf{S}^1 \rightarrow \mathbf{S}^1 \times \{1\}$ given by

$$r(x, y) := (x, 1)$$

is a retraction, it follows that \mathbf{T}^2 is not simply-connected and then \mathbf{T}^2 is not homeomorphic to \mathbf{S}^2 .

Let $X, Y \in \mathbf{Ob}(\mathbf{Top})$ and $A \subseteq X$. A homotopy H between $f, g \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$ is said to be **stationary on A** if

$$H(x, t) = f(x), \quad \forall (x, t) \in A \times [0, 1].$$

If there exists such a homotopy H , we say that f and g are **homotopic relative to A** , written as

$$H : f \simeq g \text{ (rel } A),$$

and H is called a **homotopy relative to A** . In this case, $f = g$ on A .

It is clear that homotopy relative to A is an equivalence relation in $\mathbf{Hom}_{\mathbf{Top}}(X, Y)$.

A retraction $r : X \rightarrow A$ is called a **strong deformation retraction** if $\mathbf{1}_X$ is homotopic to $\iota_A \circ r$ relative to A . That is, $A \subseteq X$ is a strong deformation retract of X if and only if there exists a homotopy $H : X \times [0, 1] \rightarrow X$ satisfying

$$\begin{aligned} H(x, 0) &\equiv H_0(x) = x, & x \in X, \\ H(x, 1) &\equiv H_1(x) \in A, & x \in X, \\ H(a, 1) &\equiv H_1(a) = a, & a \in A, \\ H(a, t) &\equiv H_t(a) = a, & a \in A, t \in [0, 1]. \end{aligned}$$

Example 3.1.20. (1) For any $m \geq 1$, \mathbf{S}^{m-1} is a strong deformation retract of $\mathbf{R}^m \setminus \{\mathbf{0}\}$ and of $\overline{\mathbf{B}^m} \setminus \{\mathbf{0}\}$. Indeed, consider the homotopy

$$H : (\mathbf{R}^n \setminus \{\mathbf{0}\}) \times [0, 1] \longrightarrow \mathbf{R}^m \setminus \{\mathbf{0}\}, \quad (x, t) \longmapsto (1-t)x + t \frac{x}{|x|}.$$

(2) For $m \geq 3$, $\mathbf{R}^m \setminus \{\mathbf{0}\}$ and $\overline{\mathbf{B}^m} \setminus \{\mathbf{0}\}$ are both simply-connected.

Next result generalizes Corollary 3.1.8.

Theorem 3.1.21. (Homotopy invariance) If $\varphi : X \rightarrow Y$ is a homotopy equivalence, then for any point $x \in X$, $\varphi_* : \pi_1(X, x) \rightarrow \pi_1(Y, \varphi(x))$ is an isomorphism.

PROOF. Let $\psi : Y \rightarrow X$ be a homotopy inverse for φ . Then, by definition, we have $\psi \circ \varphi \simeq \mathbf{1}_X$ and $\varphi \circ \psi \simeq \mathbf{1}_Y$. Moreover φ, ψ induce

$$\pi_1(X, x) \xrightarrow{\varphi_*} \pi_1(Y, \varphi(x)) \xrightarrow{\psi_*} \pi_1(X, \psi(\varphi(x))) \xrightarrow{\varphi_*} \pi_1(Y, \varphi(\psi(\varphi(x)))).$$

Lemma 3.1.22. If $\varphi, \psi \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$ and $H : \varphi \simeq \psi$ is a homotopy. For any $x \in X$, let h be the path in Y from $\varphi(x)$ to $\psi(x)$ defined by

$$h(t) := H(x, t), \quad t \in [0, 1],$$

and let $\Phi_h : \pi_1(Y, \varphi(x)) \rightarrow \pi_1(Y, \psi(x))$ be the isomorphism defined in (3.1.2.6). Then the following diagram

$$\begin{array}{ccc} \pi_1(X, x) & \xrightarrow{\varphi_*} & \pi_1(Y, \varphi(x)) \\ \parallel & & \downarrow \Phi_h \\ \pi_1(X, x) & \xrightarrow{\psi_*} & \pi_1(Y, \psi(x)) \end{array}$$

is commutative.

PROOF. For $[f] \in \pi_1(X, x)$ we need to verify

$$\psi_*([f]) = \Phi_h(\varphi_*([f]))$$

which is equivalent to $h * (\psi \circ f) \sim (\varphi \circ f) * h$ or $(\varphi \circ f) * h \sim h * (\psi \circ f)$. Consider the map

$$F : [0, 1] \times [0, 1] \longrightarrow Y, \quad (s, t) \longmapsto H(f(s), t).$$

Then

$$F(s, 0) = H(f(s), 0) = \varphi(f(s)), \quad F(1, s) = H(f(1), s) = H(x, s) = h(s)$$

and

$$F(0, s) = H(f(0), s) = H(x, s) = h(s), \quad F(s, 1) = H(f(s), 1) = \psi(f(s)).$$

Lemma 3.1.23 shows that $(\varphi \circ f) * h \sim h * (\psi \circ f)$. \square

Lemma 3.1.23. (Square lemma) Let $F : [0, 1] \times [0, 1] \rightarrow X$ be a continuous map, and let f, g, h, k be the paths in X defined by

$$f(s) := F(s, 0), \quad g(s) := F(1, s), \quad h(s) := F(0, s), \quad k(s) := F(s, 1).$$

Then $f * g \sim h * k$.

PROOF. Consider the diagram map $d(s) := F(s, s)$. It is clear that $f * g \sim d$ (**Exercise 3.1.24**). Similarly, $h * k \sim d$. Hence $f * g \sim h * k$. \square

Since $H : \mathbf{1}_X \simeq \psi \circ \varphi$, it follows that h defined by

$$h(t) := H(x, t), \quad t \in [0, 1],$$

is a path in X from x to $\psi(\varphi(x))$. By **Lemma 3.1.22**, we have the following commutative diagram:

$$\begin{array}{ccc} \pi_1(X, x) & \xrightarrow{\mathbf{1}_{X,*}} & \pi_1(X, x) \\ \parallel & & \downarrow \Phi_h \\ \pi_1(X, x) & \xrightarrow{(\psi \circ \varphi)_*} & \pi_1(Y, \psi(\varphi(x))) \end{array}$$

Thus

$$\Phi_h = \psi_* \circ \varphi_*.$$

Consequently, ψ_* is surjective and φ_* is injective.

Similarly, since $G : \mathbf{1}_Y \simeq \varphi \circ \psi$, it follows that g defined by

$$g(t) := G(\varphi(x), t), \quad t \in [0, 1],$$

is a path in Y from $\varphi(x)$ to $\varphi(\psi(\varphi(x)))$. By **Lemma 3.1.22**, we have the following commutative diagram:

$$\begin{array}{ccc} \pi_1(Y, \varphi(x)) & \xrightarrow{\mathbf{1}_{Y,*}} & \pi_1(Y, \varphi(x)) \\ \parallel & & \downarrow \Phi_g \\ \pi_1(Y, \varphi(x)) & \xrightarrow{(\varphi \circ \psi)_*} & \pi_1(Y, \varphi(\psi(\varphi(x)))) \end{array}$$

Thus

$$\Phi_g = \varphi_* \circ \psi_*.$$

Consequently, φ_* is surjective and ψ_* is injective. Therefore φ_* is isomorphic. \square

Exercise 3.1.24. Verify $f * g \sim d$ in **Lemma 3.1.23**.

3.1.6. Contractible spaces. A topological space X is said to be **contractible** if $\mathbf{1}_X$ is homotopic to some constant map. Equivalently, X is contractible, if there exists a point $x_0 \in X$ and a continuous map $H : X \times [0, 1] \rightarrow X$ such that

$$H(x, 0) = x \quad (x \in X), \quad H(x, 1) = x_0 \quad (x \in X).$$

Thus, X is contractible if, intuitively, X can be continuously shrunk to a point.

- (1) \mathbf{R}^n is contractible, because we can take $H(x, t) = (1 - t)x$.
- (2) $[0, 1]$ is contractible, because we can take $H(x, t) := (1 - t)x$.
- (3) Any convex subset X of \mathbf{R}^n is contractible, because we can take $H(x, t) = (1 - t)x + tx_0$.
- (4) Any star-shaped subset X of \mathbf{R}^n is contractible.

- Proposition 3.1.25.** (1) If $f, g \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$ and Y is contractible, then $f \simeq g$.
 (2) If X is contractible, then any two constant maps on X are homotopic and the identity map $\mathbf{1}_X$ is homotopic to any constant map on X .
 (3) A topological space X is contractible if and only if X is homotopic to a one-point space.
 (4) Two contractible spaces are homotopic and any continuous map between contractible spaces is a homotopy equivalence.

PROOF. (1) Assume $\mathbf{1}_Y \simeq c_{y_0}$, where c_{y_0} is a constant map on Y . Then $f = \mathbf{1}_Y f \simeq c_{y_0} f = c_{y_0} g \simeq \mathbf{1}_Y g = g$.

(2) Trivial.

(3) If X is homotopic to a one-point space $Y = \{y_0\}$, then we have a homotopy equivalence $f : X \rightarrow Y$ with homotopy inverse $g : Y \rightarrow X$. Then $\mathbf{1}_X \simeq gf = c_{x_0}$, where $x_0 := g(y_0)$. Thus X is contractible.

Conversely, assume that X is contractible. We can find a point $x_0 \in X$ and continuous map $H : X \times [0, 1] \rightarrow X$ such that $H(x, 0) = x$ and $H(x, 1) = x_0$ for all $x \in X$. Set $Y := \{x_0\}$ and consider two continuous maps $f : X \rightarrow Y$ given by $f(x) := x_0$ and $g : Y \rightarrow X$ given by $g(x_0) = x_0$. Then

$$f \circ g = \mathbf{1}_Y, \quad g \circ f = c_{x_0} \simeq \mathbf{1}_X.$$

Therefore f is a homotopy equivalence from X to Y with homotopy inverse g .

(4) Let X and Y be two contractible spaces. By (3), X is homotopic to some one-point space $\{x_0\}$; thus $X \simeq \{x_0\}$. Similarly, $Y \simeq \{y_0\}$ for some $y_0 \in Y$. Since $\{x_0\} \simeq \{y_0\}$, it follows that $X \simeq Y$.

Next, take a map $f \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$ with X, Y being contractible. The first part gives two maps $f_1 \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$ and $g_1 \in \mathbf{Hom}_{\mathbf{Top}}(Y, X)$ such that $f_1 \circ g_1 \simeq \mathbf{1}_Y$ and $g_1 \circ f_1 \simeq \mathbf{1}_X$. Because $f \circ g_1 \in \mathbf{Hom}_{\mathbf{Top}}(Y, X)$ and $g_1 \circ f \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$. According to (1), we get $f \circ g_1 \simeq \mathbf{1}_Y$ and $g_1 \circ f \simeq \mathbf{1}_X$. Thus f is a homotopy equivalence. \square

Example 3.1.26. Proposition 3.1.25 (1) is not true in the case of relative homotopy. That is, if $f, g \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$, $f|_A = g|_A$ for some open subset $A \subseteq X$, and Y is contractible, then we may not have $f \simeq g$ (rel A).

Consider the **comb space** Y defined by

$$Y = (\{0\} \times [0, 1]) \cup ([0, 1] \times \{0\}) \cup \left\{ (x, y) \in \mathbf{R}^2 \mid y \in [0, 1], x = \frac{1}{n}, n = 1, 2, \dots \right\}.$$

Let $H : Y \times [0, 1] \rightarrow Y$ be defined by

$$H((x, y), t) := (x, (1-t)y).$$

Then H is a homotopy from $\mathbf{1}_Y$ to the projection of Y to the x -axis, and so $\mathbf{1}_Y$ is homotopic to some constant map on Y . Thus Y is contractible. Define $c : Y \rightarrow Y$, $(x, y) \mapsto (0, 1)$. Using Proposition 3.1.25 yields that $\mathbf{1}_Y \simeq c$ and $\mathbf{1}_Y|_A = c|_A$ where $A := \{(0, 1)\}$. However we can not find a homotopy from $\mathbf{1}_Y$ to c relative to A .

We say a map $f \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$ is **null homotopic** if it is homotopic to some constant map.

Theorem 3.1.27. Let $x_0 \in \mathbf{S}^n$ and let $f \in \mathbf{Hom}_{\mathbf{Top}}(\mathbf{S}^n, Y)$. Then TFAE:

- (1) f is null homotopic.

- (2) f can be continuously extended over $\mathbf{D}^{n+1} := \overline{\mathbf{B}^{n+1}}$.
 (3) f is null homotopic relative to x_0 , that is, f is homotopic to some constant map relative to x_0 .

PROOF. (1) \Rightarrow (2): Let $H : f \simeq c_{y_0}$ for some $y_0 \in Y$. Define an extension \tilde{f} of f over \mathbf{D}^{n+1} by

$$\tilde{f}(x) := \begin{cases} y_0, & 0 \leq |x| \leq \frac{1}{2}, \\ H(x/|x|, 2(1-|x|)), & \frac{1}{2} \leq |x| \leq 1. \end{cases}$$

Then \tilde{f} is continuous.

(2) \Rightarrow (3): Assume that f has a continuous extension $\tilde{f} : \mathbf{D}^{n+1} \rightarrow Y$. Define $H : \mathbf{S}^n \times [0, 1] \rightarrow Y$ by

$$H(x, t) := \tilde{f}((1-t)x + tx_0).$$

Then $H(x, 0) = \tilde{f}(x) = f(x)$ for all $x \in \mathbf{S}^n$ and $H(x, 1) = \tilde{f}(x_0) = c_{x_0}$.

(3) \Rightarrow (1): Trivial. \square

Corollary 3.1.28. *Any continuous map from \mathbf{S}^n to a contractible space has a continuous extension over \mathbf{D}^{n+1} .*

PROOF. By **Proposition 3.1.25** (1) and **Theorem 3.1.26**. \square

3.2. H -spaces

The **homotopy category** \mathbf{HTop} is the category whose objects are topological spaces and whose morphisms are homotopy classes of continuous maps. That is

$$\mathbf{Ob}(\mathbf{HTop}) = \mathbf{Ob}(\mathbf{Top}), \quad \mathbf{Hom}_{\mathbf{HTop}}(X, Y) := [X, Y]$$

where the last one was defined in (3.1.1.1). By **Proposition 3.1.1**, we see that \mathbf{HTop} is indeed a category and the composition is defined as

$$\mathbf{Hom}_{\mathbf{HTop}}(X, Y) \times \mathbf{Hom}_{\mathbf{HTop}}(Y, Z) \longrightarrow \mathbf{Hom}_{\mathbf{HTop}}(X, Z), \quad ([f], [g]) \longmapsto [g \circ f].$$

Define a functor $[\cdot] : \mathbf{Top} \rightarrow \mathbf{HTop}$ to be

$$[\cdot] : \mathbf{Ob}(\mathbf{Top}) \longrightarrow \mathbf{Ob}(\mathbf{HTop}), \quad X \longmapsto X$$

and

$$[\cdot] : \mathbf{Hom}_{\mathbf{Top}}(X, Y) \longrightarrow \mathbf{Hom}_{\mathbf{HTop}}(X, Y), \quad f \longmapsto [f].$$

Thus the above functor $[\cdot]$ is covariant.

We extend the homotopy category of topological spaces to the homotopy category of topological pairs.

Definition 3.2.1. A **topological pair** (X, A) consists of a topological space X and a subspace $A \subseteq X$. If A is empty, we write simply X for (X, \emptyset) . A **topological subpair** $(X', A') \subseteq (X, A)$ consists of a topological pair (X', A') with $X' \subseteq X$ and $A' \subseteq A$. A **topological pair map** $f : (X, A) \rightarrow (Y, B)$ between topological pairs is a continuous map $f \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$ such that $f(A) \subseteq B$.

The above concept generalizes the pointed topological spaces and the pointed-preserving maps. Let \mathbf{Top}_{**} denote the category of topological pairs and topological pair maps. Hence \mathbf{Top} and \mathbf{Top}_* are both (fully) subcategories of \mathbf{Top}_{**} .

Given a topological pair $(X, A) \in \mathbf{Ob}(\mathbf{Top}_{**})$, we let $(X, A) \times [0, 1]$ denote the topological pair $(X \times [0, 1], A \times [0, 1]) \in \mathbf{Ob}(\mathbf{Top}_{**})$.

- (1) Let $X' \subseteq X$ and suppose $f_0, f_1 \in \mathbf{Hom}_{\mathbf{Top}_{**}}((X, A), (Y, B))$ agree on X' , that is, $f_0|_{X'} = f_1|_{X'}$. We say f_0 is **homotopic to f_1 relative to X'** , denoted by $f_0 \simeq f_1 \text{ (rel } X')$, if there exists a map $F \in \mathbf{Hom}_{\mathbf{Top}_{**}}((X, A) \times [0, 1], (Y, B))$ such that

$$F(x, 0) = f_0(x), \quad F(x, 1) = f_1(x), \quad x \in X$$

and

$$F(x, t) = f_0(x), \quad x \in X', \quad t \in [0, 1].$$

We also say F is a **homotopy relative to X' from f_0 to f_1** and denote by $F : f_0 \simeq f_1 \text{ (rel } X')$.

- (2) When $X' = \emptyset$, we return back to the previous definition of $f_0 \simeq f_1$.
(3) Note that $f_0 \simeq f_1 \text{ (rel } X')$ implies $f_0 \simeq f_1 \text{ (rel } X'')$ whenever $X'' \subseteq X'$.
(4) We also write $F_t(x) := F(x, t)$ so that $F_t \in \mathbf{Hom}_{\mathbf{Top}_{**}}((X, A), (Y, B))$ is a continuous one-parameter family of topological pair maps from (X, A) to (Y, B) .

Proposition 3.1.1 can be generalized to the following

Proposition 3.2.2. (1) *Homotopy relative to X' is an equivalence relation in*

$$\mathbf{Hom}_{\mathbf{Top}_{**}}((X, A), (Y, B)).$$

(2) *Homotopy relative to X' is preserved by composition. That is, if*

$$f_0, f_1 \in \mathbf{Hom}_{\mathbf{Top}_{**}}((X, A), (Y, B)), \quad g_0, g_1 \in \mathbf{Hom}_{\mathbf{Top}_{**}}((Y, B), (Z, C)),$$

and $f_0 \simeq f_1 \text{ (rel } X')$, $g_0 \simeq g_1 \text{ (rel } Y')$, and $f_0(X') = f_1(X') \subseteq Y'$, then $g_0 \circ f_0 \simeq g_1 \circ f_1 \text{ (rel } X')$.

Exercise 3.2.3. Prove **Proposition 3.2.2**.

The above proposition allows us to define the **homotopy class relative to X'**

$$[(X, A'), (Y, B)]_{X'} := \{[f]_{X'} : f \in \mathbf{Hom}_{\mathbf{Top}_{**}}((X, A), (Y, B))\}.$$

We also can define the **homotopy category of topological pairs \mathbf{HTop}_{**}** whose objects are topological pairs and whose morphisms are homotopy classes (relative to \emptyset). We also have a covariant functor from \mathbf{Top}_{**} to \mathbf{HTop}_{**} .

- (1) For any $(P, Q) \in \mathbf{Ob}(\mathbf{HTop}_{**})$, there is a covariant functor

$$[(P, Q), (\cdot, \cdot)] := \mathbf{Hom}_{\mathbf{HTop}_{**}}((P, Q), (\cdot, \cdot))$$

from \mathbf{HTop}_{**} to \mathbf{Set} , satisfying

$$[(P, Q), (\cdot, \cdot)]((X, A)) := \mathbf{Hom}_{\mathbf{HTop}_{**}}((P, Q), (X, A))$$

and for $[f] \in \mathbf{Hom}_{\mathbf{HTop}_{**}}((X, A), (Y, B))$,

$$f_{\#} \equiv [(P, Q), (\cdot, \cdot)]([f]) \in \mathbf{Hom}_{\mathbf{Set}}([(P, Q), (X, A)], [(P, Q), (Y, B)])$$

where

$$f_{\#}([g]) := [f \circ g], \quad [g] \in [(P, Q), (X, A)].$$

- (2) A topological pair map $f : (X, A) \rightarrow (Y, B)$ is called a **homotopy equivalence** if $[f]$ is an equivalence in \mathbf{HTop}_{**} . Equivalently, there is a topological pair map $g : (Y, B) \rightarrow (X, A)$, called a **homotopy inverse of f** , such that

$$f \circ g \simeq \mathbf{1}_Y, \quad g \circ f \simeq \mathbf{1}_X.$$

Pairs (X, A) and (Y, B) are said to have the **same homotopy type** if there are equivalent in \mathbf{HTop} .

3.2.1. Two notions on functors. Let \mathcal{C} be a category and take an object $Y \in \mathcal{C}$.

- (1) Define a covariant functor $\pi_Y : \mathcal{C} \rightarrow \mathbf{Set}$ by

$$(3.2.1.1) \quad \pi_Y := \mathbf{Hom}(Y, \cdot) \text{ on } \mathbf{Ob}(\mathcal{C})$$

and

$$(3.2.1.2) \quad \pi_Y(h) := h_{\#} : \mathbf{Hom}_{\mathcal{C}}(Y, Z) \longrightarrow \mathbf{Hom}_{\mathcal{C}}(Y', Z), \quad g \longmapsto h \circ g,$$

for any $h \in \mathbf{Hom}_{\mathcal{C}}(Z, Z')$.

- (2) Define a covariant functor $\pi^Y : \mathcal{C} \rightarrow \mathbf{Set}$ by

$$(3.2.1.3) \quad \pi^Y := \mathbf{Hom}(\cdot, Y) \text{ on } \mathbf{Ob}(\mathcal{C})$$

and

$$(3.2.1.4) \quad \pi^Y(f) := h_{\#} : \mathbf{Hom}_{\mathcal{C}}(X', Y) \longrightarrow \mathbf{Hom}_{\mathcal{C}}(X, Y), \quad g' \longmapsto g' \circ f,$$

for any $f \in \mathbf{Hom}_{\mathcal{C}}(X, X')$.

3.2.2. H -spaces. If $(X, x_0), (Y, y_0) \in \mathbf{Ob}(\mathbf{Top}_*)$, we denote by

$$(3.2.2.1) \quad [X, Y] := [(X, x_0), (Y, y_0)]$$

the set of morphisms from X to Y in \mathbf{HTop}_* .

Theorem 3.2.4. *If P is a topological group with the identity element, π^P is a contravariant functor from \mathbf{HTop}_* to \mathbf{Group} .*

PROOF. (1) $[X, P]$ is a group for any given pointed topological space X . If $[g_1], [g_2] \in \mathbf{Hom}_{\mathbf{HTop}_*}(X, P)$, we define

$$[g_1][g_2] := [g_1g_2]$$

with $g_1g_2 : X \rightarrow P$ given by $(g_1g_2)(x) := g_1(x)g_2(x)$. Clearly that under this multiplication, $[X, P]$ is a group.

- (2) π^P is a contravariant functor. □

Exercise 3.2.5. Complete the proof of Theorem 3.2.4.

It is clear that $[X, P]$ is Abelian if P is Abelian. For example, $[X, \mathbf{S}^1]$ is an Abelian group.

An H -space (P, μ) consists of a pointed topological space $P = (P, p_0)$ together with a continuous multiplication

$$(3.2.2.2) \quad \mu : P \times P \longrightarrow P$$

for which the constant map $c : P \rightarrow P$, given by $c = c_{p_0}$, is a **homotopy identity**, that is, each composite, $1 := \mathbf{1}_P$

$$P \xrightarrow{(c,1)} P \times P \xrightarrow{\mu} P, \quad P \xrightarrow{(1,c)} P \times P \xrightarrow{\mu} P$$

is homotopic to 1.

- (1) The multiplication μ is said to be **homotopy associative** if the square

$$\begin{array}{ccc} P \times P \times P & \xrightarrow{\mu \times 1} & P \times P \\ 1 \times \mu \downarrow & & \downarrow \mu \\ P \times P & \xrightarrow{\mu} & P \end{array}$$

is homotopy commutative, that is,

$$\mu \circ (\mu \times 1) \simeq \mu \circ (1 \times \mu).$$

- (2) A continuous map $\varphi : P \rightarrow P$ is called a **homotopy inverse** for P and μ if each composite

$$P \xrightarrow{(\varphi,1)} P \times P \xrightarrow{\mu} P, \quad P \xrightarrow{(1,\varphi)} P \times P \xrightarrow{\mu} P$$

is homotopic to c .

- (3) An **H-group** (P, μ, φ) is a homotopy associative H -space with a homotopy inverse.
(4) Any topological group is an H -group.
(5) A multiplication μ in an H -space P is said to be **homotopy Abelian** if the triangle

$$\begin{array}{ccc} P \times P & \xrightarrow{T} & P \times P \\ \mu \downarrow & & \downarrow \mu \\ P & \xlongequal{\quad} & P \end{array} \quad T(p_1, p_2) := (p_2, p_1),$$

is homotopy commutative, that is,

$$\mu \circ T \simeq \mu.$$

- (6) An H -group with homotopy Abelian multiplication is called an **Abelian H-group**.
(7) An **H -space homomorphism** between two H -spaces (P, μ) and (P', μ') is a continuous map $\alpha : P \rightarrow P'$ such that the square

$$\begin{array}{ccc} P \times P & \xrightarrow{\mu} & P \\ \alpha \times \alpha \downarrow & & \downarrow \alpha \\ P' \times P' & \xrightarrow{\mu'} & P' \end{array}$$

is homotopy commutative, that is,

$$\alpha \circ \mu \simeq \mu' \circ (\alpha \times \alpha).$$

Theorem 3.2.6. *A pointed topological space having the same homotopy type as an H -space (or an H -group, or an Abelian H -group), is itself an H -space (or H -group, or an Abelian H -group) in such a way that the homotopy equivalence is an H -space homomorphism.*

PROOF. (1) Let (P, μ) be an H -space, and (P', p'_0) is a pointed topological space that has the same homotopy type as P . Let $f : P \rightarrow P'$ and $g : P' \rightarrow P$ be homotopy inverses. Define

$$P' \times P' \xrightarrow{g \times g} P \times P \xrightarrow{\mu} P \xrightarrow{f} P'$$

μ as

$$\mu' := f \circ \mu \circ (g \times g).$$

Then μ' is a continuous map. Next we shall verify that each composite

$$P' \xrightarrow{(c', 1')} P' \times P' \xrightarrow{\mu'} P', \quad P' \xrightarrow{(1', c')} P' \times P' \xrightarrow{\mu'} P'$$

is homotopic to $\mathbf{1}_{P'}$. Since

$$\begin{array}{ccc} P' & \xrightarrow{(1', c')} & P' \times P' \xrightarrow{\mu'} P' \\ g \downarrow & & \uparrow f \\ P & \xrightarrow{(1, c)} & P \times P \xrightarrow{\mu} P \end{array}$$

the above diagram shows $f \circ \mu \circ (1, c) \circ g = \mu' \circ (1, c')$ and $\mu \circ (1, c) \simeq \mathbf{1}_P$, it follows that

$$\mu' \circ (1', c') \simeq \mathbf{1}_{P'}.$$

Similarly, from

$$\begin{array}{ccc} P' & \xrightarrow{(c', 1')} & P' \times P' \xrightarrow{\mu'} P' \\ g \downarrow & & \uparrow f \\ P & \xrightarrow{(c, 1)} & P \times P \xrightarrow{\mu} P \end{array}$$

the above diagram shows $f \circ \mu \circ (c, 1) \circ g = \mu' \circ (c', 1')$ and $\mu \circ (c, 1) \simeq \mathbf{1}_P$, it follows that

$$\mu' \circ (c', 1') \simeq \mathbf{1}_{P'}.$$

Consequently, (P', μ') is an H -space.

(2) Consider the diagrams

$$\begin{array}{ccc} P \times P & \xrightarrow{\mu} & P \\ f \times f \downarrow & & \downarrow f \\ P' \times P' & \xrightarrow{\mu'} & P' \end{array} \quad \begin{array}{ccc} P' \times P' & \xrightarrow{\mu'} & P' \\ g \times g \downarrow & & \downarrow g \\ P \times P & \xrightarrow{\mu} & P \end{array}$$

Because

$$\begin{aligned} \mu' \circ (f \times f) &= f \circ \mu \circ (g \times g) \circ (f \times f) \\ &= f \circ \mu \circ (g \circ f) \times (g \circ f) \simeq f \circ \mu \circ (\mathbf{1}_P \times \mathbf{1}_P) \simeq f \circ \mu \end{aligned}$$

and

$$g \circ \mu' = g \circ f \circ \mu \circ (g \times g) \simeq \mu \circ (g \times g),$$

we see that g and f are both H -space homomorphisms.

- (3) If μ is homotopy associative, so is μ' .
 (4) If μ is homotopy Abelian, so is μ' .
 (5) If $\varphi : P \rightarrow P$ is a homotopy inverse for P , then $\varphi' := f \circ \varphi \circ g : P' \rightarrow P'$ is a homotopy inverse for P' . \square

Exercise 3.2.7. Verify (3), (4), and (5) in **Theorem 3.2.6**.

Given an H -space (P, μ) (with base point p_0) and a pointed topological space (X, x_0) , for any $[g_1], [g_2] \in [X, P]$ define

$$(3.2.2.3) \quad [g_1] * [g_2] := [\mu \circ (g_1, g_2)].$$

Clearly the above multiplication is well-defined. If (P, μ, φ) is an H -group, then $[X, P]$ is a group under (3.2.2.3).

Note that in (3.2.2.3), the multiplication $*$ is determined only by μ and then by P . If we consider $[g_1], [g_2] \in [Y, P]$, we still use the same notion $*$ for the multiplication of $[Y, P]$.

Exercise 3.2.8. Show that if (P, μ, φ) is an H -group, then $[X, P]$ is a group under (3.2.2.3).

We then have a contravariant functor

$$\pi^P : \mathbf{HTop}_* \longrightarrow \mathbf{Group}, \quad X \longmapsto [X, P]$$

and for any morphism $[f] \in \mathbf{Hom}_{\mathbf{HTop}_*}(X, Y)$,

$$f^\# := \pi^P([f]) \in \mathbf{Hom}_{\mathbf{Group}}([Y, P], [X, P]), \quad [Y, P] \ni g' \longmapsto g' \circ f.$$

Consequently, we obtain the following

Theorem 3.2.9. *If $P \equiv (P, \mu, \varphi)$ is an H -group, then $\pi^P : \mathbf{HTop}_* \rightarrow \mathbf{Group}$ is a contraovariant functor. If P is an Abelian H -group, then π^P is a contravariant functor from \mathbf{HTop}_* into \mathbf{AG} .*

Theorem 3.2.10. *If $P \equiv (P, p_0)$ is a pointed topological space such that π^P takes values in \mathbf{Group} , then P is an H -group (or Abelian H -group if π^P takes values in \mathbf{AG}).*

PROOF. Consider the pointed topological space $P \times P \equiv (P \times P, (p_0, p_0))$ and $\pi^P(P \times P) = [P \times P, P] \in \mathbf{Ob}(\mathbf{Group})$. Denote by $*$ the group multiplication of $[P \times P, P]$. The projections

$$\begin{array}{ccc} P \times P & \xrightarrow{\text{pr}_1} & P \\ \text{pr}_2 \downarrow & & \\ & & P \end{array}$$

gives rise to a continuous map $\mu : P \times P \rightarrow P$ such that

$$[\text{pr}_1] * [\text{pr}_2] = [\mu].$$

Write $c : P \rightarrow P$ the constant map c_{p_0} .

(1) (P, μ) is an H -space. Consider the following diagrams

$$P \xrightarrow[(1,c)]{(c,1)} P \times P \xrightarrow{\mu} P$$

and try to prove $\mu \circ (c, 1) \simeq 1 \simeq \mu \circ (1, c)$, where $1 := \mathbf{1}_P$. We first establish a general formula for $[\mu \circ (f, g)]$, where $f, g \in \mathbf{Hom}_{\mathbf{Top}_*}(X, P)$. In fact,

$$\begin{aligned} [\mu \circ (f, g)] &= (f, g)^\#[\mu] = (f, g)^\#([\mathbf{pr}_1] * [\mathbf{pr}_2]) \\ &= \left((f, g)^\#[\mathbf{pr}_1] \right) * \left((f, g)^\#[\mathbf{pr}_2] \right) = [f] * [g]. \end{aligned}$$

Consequently,

$$[\mu \circ (1, c)] = [1] * [c].$$

Consider the following two continuous maps

$$f : P \longrightarrow \{p_0\}, \quad g : \{p_0\} \longrightarrow P$$

defined, respectively, by

$$f(p) := p_0, \quad g(p_0) := p_0.$$

Since π^P takes values in **Group**, it follows that

$$\pi^P([f]) \in \mathbf{Hom}_{\mathbf{Group}}(\{\{p_0\}, P\}, [P, P]).$$

In particular,

$$\pi^P([f])(e_{\{\{p_0\}, P\}}) = e_{[P, P]},$$

where e_G stands for the identity element of the given group G . Because $\{\{p_0\}, P\}$ consists only one element $[g]$, we must have $[g] = e_{\{\{p_0\}, P\}}$. Hence

$$e_{[P, P]} = \pi^P([f])([g]) = [g \circ f] = [c].$$

Thus $[c]$ is the identity element in the group $[P, P]$. Consequently,

$$[\mu \circ (1, c)] = [1] * [c] = [1] \iff \mu \circ (1, c) \simeq 1.$$

Similarly, one has $\mu \circ (c, 1) \simeq 1$.

(2) (P, μ, φ) is an H -group. We next show that μ is homotopy associative, i.e.,

$$\begin{array}{ccc} P \times P \times P & \xrightarrow{\mu \times 1} & P \times P \\ \downarrow 1 \times \mu & & \downarrow \mu \\ P \times P & \xrightarrow{\mu} & P \end{array} \quad \mu \circ (\mu \times 1) \simeq \mu \circ (1 \times \mu).$$

Let $q_1, q_2, q_3 : P \times P \times P \rightarrow P$ be three projections. Then

$$\begin{aligned} [\mu \circ (1 \times \mu)] &= (1 \times \mu)^\#[\mu] = (1 \times \mu)^\#([\mathbf{pr}_1] * [\mathbf{pr}_2]) \\ &= (1 \times \mu)^\#[\mathbf{pr}_1] *_{P \times P \times P} (1 \times \mu)^\#[\mathbf{pr}_2] = [q_1] * [\mu \circ (q_2, q_3)] = [q_1] * ([q_2] * [q_3]). \end{aligned}$$

Similarly,

$$[\mu \circ (\mu \times 1)] = ([q_1] * [q_2]) * [q_3].$$

Since $*$ is associative, it follows that $[\mu \circ (1 \times \mu)] = [\mu \circ (\mu \times 1)]$.

Thirdly, we show that P has a homotopy inverse. Let $[\varphi] \in [P, P]$ be the inverse of $[1]$ in $[P, P]$, i.e., $[1] * [\varphi] = [c] = [\varphi] * [1]$. Then $c \simeq \mu(1, \varphi)$ and $c \simeq \mu(\varphi, 1)$.

(3) (P, μ, φ) is an Abelian H -group if π^P takes values in **AG**. In this case, $*$ is Abelian. Consider the following diagram

$$\begin{array}{ccc} P \times P & \xrightarrow{T} & P \times P \\ \mu \downarrow & & \downarrow \mu \\ P & \xlongequal{\quad} & P \end{array} \quad T(p_1, p_2) = (p_2, p_1).$$

Because $T = (\text{pr}_2, \text{pr}_1)$, we obtain

$$[\mu \circ T] = [\mu \circ (\text{pr}_2, \text{pr}_1)] = [\text{pr}_2] * [\text{pr}_1] = [\text{pr}_1] * [\text{pr}_2] = [\mu].$$

Thus P is Abelian. \square

Corollary 3.2.11. *Let $\alpha : P \rightarrow P'$ be a map between H -groups. Then $\alpha_{\#}$ is a natural transformation from π^P to $\pi^{P'}$ in **Group** if and only if α is an H -space homomorphism.*

PROOF. For $[f] \in \mathbf{Hom}_{\mathbf{HTop}_*}(X, Y)$, consider the following diagram

$$\begin{array}{ccc} \pi^P(X) & \xrightarrow{\alpha_{\#}(X)} & \pi^{P'}(X) \\ \pi^P(f) \uparrow & & \uparrow \pi^{P'}(f) \\ \pi^P(Y) & \xrightarrow{\alpha_{\#}(Y)} & \pi^{P'}(Y) \end{array}$$

For any $[g] \in [Y, P]$, we have

$$(\pi^{P'}(f) \circ \alpha_{\#}(Y)) [g] = \pi^{P'}(f)[\alpha \circ g] = [(\alpha \circ g) \circ f]$$

and

$$(\alpha_{\#}(X) \circ \pi^P(f)) [g] = \alpha_{\#}(X)[g \circ f] = [\alpha \circ (g \circ f)].$$

Hence $\pi^{P'}(f) \circ \alpha_{\#}(Y) = \alpha_{\#}(X) \circ \pi^P(f)$. We need only to verify that $\alpha_{\#}(X) : [X, P] \rightarrow [X, P']$ is a group homomorphism. For $[g_1], [g_2] \in [X, P]$ we have

$$\alpha_{\#}(X)([g_1] * [g_2]) = \alpha_{\#}(X)[\mu \circ (g_1, g_2)] = [\alpha \circ \mu \circ (g_1, g_2)]$$

and

$$\alpha_{\#}(X)[g_1] * \alpha_{\#}(X)[g_2] = [\alpha \circ g_1] * [\alpha \circ g_2] = [\mu' \circ (\alpha \circ g_1, \alpha \circ g_2)].$$

Consequently, $\alpha_{\#}(X)$ is a group homomorphism if and only if $\alpha \circ \mu \simeq \mu' \circ (\alpha \times \alpha)$, i.e., $\alpha_{\#}(X)$ is a group homomorphism for any $X \in \mathbf{Ob}(\mathbf{HTop}_*)$ if and only if α is a H -space homomorphism. \square

3.2.3. A classical example of H -groups. A classical example of H -groups is as follows. Let $Y \equiv (Y, y_0)$ be a pointed topological space. The **loop space of Y (based at y_0)**, denoted by $\Omega(Y, y_0)$, is defined to be

(3.2.3.1)

$$\Omega(Y, y_0) := \left\{ \omega : ([0, 1], \{0, 1\}) \longrightarrow (Y, y_0) \left| \begin{array}{l} \text{continuous functions} \\ \text{topologized by the} \\ \text{compact-open topology} \end{array} \right. \right\}.$$

We recall the compact-open topology. If $(X, A), (Y, B) \in \mathbf{Ob}(\mathbf{Top}_{**})$, then $\mathbf{Hom}_{\mathbf{Top}_{**}}((X, A), (Y, B))$ denotes the set of continuous functions $f : X \rightarrow Y$ such

that $f(A) \subseteq B$. Let Y^X denote the space of continuous functions from X to Y , given the **compact-open topology** (which is the topology generated by the subbasis

$$(3.2.3.2) \quad \{ \mathbf{Hom}_{\mathbf{Top}_{**}}((X, K), (Y, U)) : K \text{ compact in } X \text{ and } U \text{ open in } Y \}.$$

Thus the compact-open topology is the collection of subsets consisting of Y^X , \emptyset , and all unions of finite intersections of elements of (3.2.3.2).

(1) Let

$$(3.2.3.3) \quad \mathbf{E} : Y^X \times X \longrightarrow Y, \quad (f, x) \longmapsto f(x),$$

denote the **evaluation map**. Given a function $g : Z \rightarrow Y^X$, the composite

$$(3.2.3.4) \quad Z \times X \xrightarrow{g \times 1} Y^X \times X \xrightarrow{\mathbf{E}} Y$$

is a function from $Z \times X$ to Y .

(2) **(Theorem of exponential correspondence)** *If X is locally compact Hausdorff space and $Y, Z \in \mathbf{Ob}(\mathbf{Top})$, a map $g : Z \rightarrow Y^X$ is continuous if and only if $\mathbf{E} \circ (g \times 1) : Z \times X \rightarrow Y$ is continuous.*

Theorem 3.2.12. $\Omega(Y, y_0)$ is an H -group.

PROOF. Define $\omega_0 : [0, 1] \rightarrow Y$ to be $\omega_0(t) \equiv y_0$. Then $\omega_0 \in \Omega(Y, y_0)$ and then $\Omega(Y, y_0) \in \mathbf{Ob}(\mathbf{Top}_*)$.

(1) Define a map

$$\mu : \Omega(Y, y_0) \times \Omega(Y, y_0) \longrightarrow \Omega(Y, y_0)$$

by

$$\mu(\omega, \omega')(t) := (\omega * \omega')(t) = \begin{cases} \omega(2t), & 0 \leq t \leq \frac{1}{2}, \\ \omega'(2t - 1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

To prove μ is continuous, we use the above Theorem of exponential correspondence. Let $\mathbf{E} : \Omega(Y, y_0) \times [0, 1] \rightarrow Y$ be the evaluation map. It suffices to show that the composite

$$\Omega(Y, y_0) \times \Omega(Y, y_0) \times [0, 1] \xrightarrow{\mu \times 1} \Omega(Y, y_0) \times [0, 1] \xrightarrow{\mathbf{E}} Y$$

is continuous. By definition of μ , we see that $\mathbf{E} \circ (\mu \times 1)$ is continuous so that μ is continuous.

(2) To prove

$$\Omega(Y, y_0) \xrightarrow[(1, c)]{(c, 1)} \Omega(Y, y_0) \times \Omega(Y, y_0) \xrightarrow{\mu} \Omega(Y, y_0), \quad \mu \circ (c, 1) \simeq 1 \simeq \mu \circ (1, c)$$

where $1 := \mathbf{1}_{\Omega(Y, y_0)}$, we define a map

$$F : \Omega(Y, y_0) \times [0, 1] \longrightarrow \Omega(Y, y_0), \quad F(\omega, t)(s) = \begin{cases} \omega\left(\frac{2s}{t+1}\right), & 0 \leq s \leq \frac{t+1}{2}, \\ y_0, & \frac{t+1}{2} \leq s \leq 1. \end{cases}$$

Since $\mathbf{E} \circ (F \times 1) : \Omega(Y, y_0) \times [0, 1] \times [0, 1] \rightarrow Y$ is continuous, it follows that F is continuous and then $F \circ (\mu \circ (1, c)) \simeq 1$. Similarly, we have $\mu \circ (c, 1) \simeq 1$. Hence $(\Omega(Y, y_0), \mu)$ is an H -space.

(3) To show that μ is homotopy associative, consider the following diagram

$$\begin{array}{ccc} \Omega(Y, y_0) \times \Omega(Y, y_0) \times \Omega(Y, y_0) & \xrightarrow{\mu \times 1} & \Omega(Y, y_0) \times \Omega(Y, y_0) \\ 1 \times \mu \downarrow & & \downarrow \mu \\ \Omega(Y, y_0) \times \Omega(Y, y_0) & \xrightarrow{\mu} & \Omega(Y, y_0) \end{array}$$

Define a homotopy by

$$G : \Omega(Y, y_0) \times \Omega(Y, y_0) \times \Omega(Y, y_0) \times [0, 1] \longrightarrow \Omega(Y, y_0)$$

by

$$G(\omega, \omega', \omega'', t)(s) := \begin{cases} \omega\left(\frac{4s}{t+1}\right), & 0 \leq s \leq \frac{t+1}{4}, \\ \omega'(4s - t - 1), & \frac{t+1}{4} \leq s \leq \frac{t+2}{4}, \\ \omega''\left(\frac{4s-t-2}{2-t}\right), & \frac{t+2}{4} \leq s \leq 1. \end{cases}$$

Then G is continuous and $G : \mu \circ (\mu \times 1) \simeq \mu \circ (1 \times \mu)$.

(4) Define a homotopy inverse $\varphi : \Omega(Y, y_0) \rightarrow \Omega(Y, y_0)$ by

$$\varphi(\omega)(s) := \omega(1 - s), \quad s \in [0, 1].$$

Define a homotopy by

$$H : \Omega(Y, y_0) \times [0, 1] \longrightarrow \Omega(Y, y_0)$$

by

$$H(\omega, t)(s) = \begin{cases} y_0, & 0 \leq s \leq \frac{t}{2}, \\ \omega(2s - t), & \frac{t}{2} \leq s \leq \frac{1}{2}, \\ \omega(2 - 2s - t), & \frac{1}{2} \leq s \leq 1 - \frac{t}{2}, \\ y_0, & 1 - \frac{t}{2} \leq s \leq 1. \end{cases}$$

Then H is continuous and $H : \mu \circ (1, \varphi) \simeq c$. Similarly, we have $\mu \circ (\varphi, 1) \simeq c$.

In summary, $(\Omega(Y, y_0), \mu, \varphi)$ is an H -group. \square

Let **HGroup** denote the category of H -groups and continuous H -space homomorphisms.

Proposition 3.2.13. *The loop functor Ω is a covariant functor from \mathbf{Top}_* to **HGroup**.*

PROOF. For $(Y, y_0) \in \mathbf{Ob}(\mathbf{Top}_*)$, define $\Omega((Y, y_0)) := (\Omega(Y, y_0), \mu, \varphi) \in \mathbf{Ob}(\mathbf{HGroup})$. If $h \in \mathbf{Hom}_{\mathbf{Top}_*}(Y, Y')$ ($h(y_0) = y'_0$), we define

$$\Omega(h) : \Omega(Y, y_0) \longrightarrow \Omega(Y', y'_0)$$

by

$$\Omega(h)(\omega)(s) := h(\omega(s)), \quad s \in [0, 1].$$

To verify $\Omega(h) \in \mathbf{Hom}_{\mathbf{HGroup}}(\Omega(Y, y_0), \Omega(Y', y'_0))$ we consider the following diagram

$$\begin{array}{ccc} \Omega(Y, y_0) \times \Omega(Y, y_0) & \xrightarrow{\mu} & \Omega(Y, y_0) \\ \Omega(h) \times \Omega(h) \downarrow & & \downarrow \Omega(h) \\ \Omega(Y', y'_0) \times \Omega(Y', y'_0) & \xrightarrow{\mu'} & \Omega(Y', y'_0) \end{array}$$

For any $\omega, \omega' \in \Omega(Y, y_0)$ we have

$$(\Omega(h) \circ \mu)(\omega, \omega') = \Omega(h)(\mu(\omega, \omega')) = h \circ \mu(\omega, \omega') = h \circ (\omega * \omega')$$

$$\begin{aligned} &\simeq (h \circ \omega) * (h \circ \omega') = \mu'(h \circ \omega, h \circ \omega') = \mu'(\Omega(h)(\omega), \Omega(h)(\omega')) \\ &= (\mu' \circ (\Omega(h) \times \Omega(h))) (\omega, \omega'). \end{aligned}$$

Thus $\Omega(h) \circ \mu \simeq \mu' \circ (\Omega(h) \times \Omega(h))$. In summary, Ω is a covariant functor. \square

3.2.4. Suspension. If (X, x_0) and (Y, y_0) are pointed topological spaces, we define

$$(3.2.4.1) \quad X \vee Y := (X \times \{y_0\}) \cup \{x_0\} \times Y$$

the subspace of $X \times Y$. If $f \in \mathbf{Hom}_{\mathbf{Top}_*}(X, Z)$ and $g \in \mathbf{Hom}_{\mathbf{Top}_*}(Y, Z)$, then

$$(3.2.4.2) \quad (f, g) : X \vee Y \longrightarrow Z$$

is given by $(f, g)|_X = f$ and $(f, g)|_Y = g$. Note that $X \vee Y$ can be regarded as a pointed topological space $(X \vee Y, (x_0, y_0))$.

Definition 3.2.14. Let \mathcal{C} be a category and $X_i \in \mathbf{Ob}(\mathcal{C}), i \in I$. The **direct sum** of $X_i (i \in I)$ is an object $\bigoplus_{i \in I} X_i$ together with a family of morphisms

$$k_i : X_i \longrightarrow \bigoplus_{i \in I} X_i$$

with the universal property: For any object Y and any family of morphisms $f_i : X_i \rightarrow Y$, there exists a unique morphism $f : \bigoplus_{i \in I} X_i \rightarrow Y$ such that $f \circ k_i = f_i$ for any $i \in I$.

$$\begin{array}{ccc} \bigoplus_{i \in I} X_i & \xleftarrow{k_i} & X_i \\ \downarrow \exists! f & & \downarrow f_i \\ \mathcal{C} & \xlongequal{\quad} & \mathcal{C} \end{array}$$

If the direct sum of X_i exists, it is unique up to unique morphism.

We now show that $X \vee Y, X \equiv (X, x_0), Y \equiv (Y, y_0)$ are pointed topological spaces, is the direct sum $X \oplus Y$ of X and Y in \mathbf{Top}_* . Consider

$$k_1 : X \rightarrow X \vee Y, \quad k_2 : Y \rightarrow X \vee Y$$

given, respectively, by $k_1(x) := (x, y_0)$ and $k_2(y) := (x_0, y)$. For any $Z \in \mathbf{Ob}(\mathbf{Top}_*)$ and $f_1 \in \mathbf{Hom}_{\mathbf{Top}_*}(X, Z)$ and $f_2 \in \mathbf{Hom}_{\mathbf{Top}_*}(Y, Z)$, define

$$f : X \vee Y \longrightarrow Z, \quad f(x, y) := \begin{cases} f_1(x), & (x, y) \in X \times \{y_0\}, \\ f_2(y), & (x, y) \in \{x_0\} \times Y. \end{cases}$$

Hence $f \circ k_i = f_i, i = 1, 2$ and therefore $X \vee Y \cong X \oplus Y$ in \mathcal{C} .

An **H-cogroup** consists of a pointed topological space $Q \equiv (Q, q_0)$ together with a continuous multiplication

$$(3.2.4.3) \quad v : Q \longrightarrow Q \vee Q$$

for which the constant $c : P \rightarrow P$, given by $c = c_{q_0}$, is a **homotopy identity**, that is, each composite, $1 = 1_Q$,

$$Q \xrightarrow{v} Q \wedge Q \xrightarrow{(c, 1)} Q \quad Q \xrightarrow{v} Q \vee Q \xrightarrow{(1, c)} Q$$

is homotopic to 1.

- (1) The multiplication ν is said to be **homotopy associative** if the square

$$\begin{array}{ccc} Q & \xrightarrow{\nu} & Q \vee Q \\ \nu \downarrow & & \downarrow 1 \vee \nu \\ Q \vee Q & \xrightarrow{\nu \vee 1} & Q \vee Q \vee Q \end{array}$$

is homotopy commutative, i.e.,

$$(1 \vee \nu) \circ \nu \simeq (\nu \vee 1) \circ \nu.$$

- (2) A continuous map $\psi : Q \rightarrow Q$ is called a **homotopy inverse** for Q and ν if each composite

$$Q \xrightarrow{\nu} Q \vee Q \xrightarrow{(1, \psi)} Q, \quad Q \xrightarrow{\nu} Q \vee Q \xrightarrow{(\psi, 1)} Q$$

is homotopic to c .

- (3) An **H-cogroup** (Q, ν, ψ) is a homotopy associative H -cospace with a homotopy inverse.
 (4) A multiplication ν in an H -cospace Q is said to be **homotopy Abelian** if the triangle

$$\begin{array}{ccc} Q & \xrightarrow{\nu} & Q \vee Q \\ \parallel & & \downarrow T \\ Q & \xrightarrow{\nu} & Q \vee Q \end{array} \quad T(q_1, q_2) := (q_2, q_1),$$

is homotopy commutative, that is,

$$T \circ \nu \simeq \nu.$$

- (5) An H -cogroup with homotopy Abelian multiplication is called an **Abelian H-cogroup**.
 (6) An **H-cospace homomorphism** between two H -cospaces (Q, ν) and (Q', ν') is a continuous map $\beta : Q \rightarrow Q'$ such that the square

$$\begin{array}{ccc} Q & \xrightarrow{\nu} & Q \vee Q \\ \beta \downarrow & & \downarrow \beta \vee \beta \\ Q' & \xrightarrow{\nu'} & Q' \vee Q' \end{array}$$

is homotopy commutative, that is,

$$(\beta \vee \beta) \circ \nu \simeq \nu' \circ \beta.$$

Theorem 3.2.15. *A pointed topological space having the same homotopy type as an H -cospace (or an H -cogroup, or an Abelian H -cogroup), is itself an H -cospace (or an H -cogroup, or an Abelian H -cogroup) in such a way that the homotopy equivalence is an H -cospace homomorphism.*

PROOF. Similar to **Theorem 3.2.6**. □

Given an H -cospace (Q, ν) (with base point q_0) and a pointed topological space (X, x_0) , for any $[f_1], [f_2] \in [Q, X]$ define

$$(3.2.4.4) \quad [f_1] * [f_2] := [(f_1, f_2) \circ \nu].$$

Clearly the above multiplication is well-defined. If (P, μ, φ) is an H -group, then $[X, P]$ is a group under (3.2.4.4).

We then have a covariant functor

$$\pi_Q : \mathbf{HTop}_* \longrightarrow \mathbf{Group}, \quad X \longmapsto [Q, X]$$

and for any morphism $[f] \in \mathbf{Hom}_{\mathbf{HTop}_*}(X, Y)$,

$$f\# := \pi_Q([f]) \in \mathbf{Hom}_{\mathbf{Group}}([Q, X], [Q, Y]), \quad [Q, X] \ni g' \longmapsto f \circ g'.$$

Consequently, we obtain the following

Theorem 3.2.16. *If $Q \equiv (Q, \nu, \psi)$ is an H -cogroup, then $\pi_Q : \mathbf{HTop}_* \rightarrow \mathbf{Group}$ is a covariant functor. If P is an Abelian H -cogroup, then π_Q is a covariant functor from \mathbf{HTop}_* into \mathbf{AG} .*

PROOF. Similar to **Theorem 3.2.9**. □

Theorem 3.2.17. *If $Q \equiv (Q, q_0)$ is a pointed topological space such that π_Q takes values in \mathbf{Group} , then Q is an H -cogroup (or Abelian H -cogroup if π_Q takes values in \mathbf{AG}).*

PROOF. Similar to **Theorem 3.2.10**. □

Corollary 3.2.18. *Let $\beta : Q \rightarrow Q'$ be a map between H -cogroups. Then $\beta\#$ is a natural transformation from $\pi_{Q'}$ to π_Q in \mathbf{Group} if and only if β is an H -cospace homomorphism.*

PROOF. Similar to **Corollary 3.2.11**. □

Exercise 3.2.19. Verify **Theorem 3.2.15**, **Theorem 3.2.16**, **Theorem 3.2.17**, and **Corollary 3.2.18**.

3.2.5. A classical example of H -cogroups. A classical example of H -cogroups is as follows. Let $Z \equiv (Z, z_0)$ be a pointed topological space. The **suspension of Z (based at z_0)**, denoted by $\Xi(Z, z_0)$, is defined to be

$$(3.2.5.1) \quad \Xi(Z, z_0) := Z \times [0, 1] / ((Z \times \{0\}) \cup (\{z_0\} \times [0, 1]) \cup (Z \times \{1\})).$$

The equivalence class of $(z, t) \in Z \times [0, 1]$ is denoted by $[z, t]$ so that

$$[z, 0] = [z', 1] = [z_0, t], \quad \text{for all } z, z' \in Z \text{ and } t \in [0, 1].$$

Moreover, $\Xi(Z, z_0)$ is a pointed topological space with base point $[z_0, 0]$. Define

$$\nu : \Xi(Z, z_0) \longrightarrow \Xi(Z, z_0) \vee \Xi(Z, x_0)$$

by

$$\nu([z, t]) := \begin{cases} ([z, 2t], [z_0, 0]), & 0 \leq t \leq \frac{1}{2}, \\ ([z_0, 0], [z, 2t - 1]), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Theorem 3.2.20. $\Xi(Z, z_0)$ is an H -cogroup.

Let $\mathbf{HCGroup}$ denote the category of H -cogroups and continuous H -cospace homomorphisms.

Proposition 3.2.21. *The loop functor Ξ is a covariant functor from \mathbf{Top}_* to $\mathbf{HCGroup}$.*

Lemma 3.2.22. *For $n \geq 0$, $\Xi(\mathbf{S}^n)$ is homeomorphic to \mathbf{S}^{n+1} .*

PROOF. Let $\mathbf{p}_0 := (1, 0, \dots, 0) \in \mathbf{S}^n$ be a base point of \mathbf{S}^n . Then

$$\mathbf{S}^n = \{z \in \mathbf{R}^{n+2} \mid \mathbf{x} = (x^1, \dots, x^{n+2}), |\mathbf{x}| = 1, x^{n+2} = 0\}$$

and

$$\mathbf{D}^{n+1} = \{\mathbf{x} \in \mathbf{R}^{n+2} \mid \mathbf{x} = (x^1, \dots, x^{n+2}), |\mathbf{x}| \leq 1, x^{n+2} = 0\}.$$

It is clear that

$$\mathbf{S}^{n+1} = H_+ \cup H_-, \quad \mathbf{S}^n = H_+ \cap H_-,$$

where

$$H_{\pm} := \{\mathbf{x} \in \mathbf{S}^{n+1} \mid \pm x^{n+2} \geq 0\}.$$

The natural projection $\mathbf{R}^{n+2} \rightarrow \mathbf{R}^{n+1}$ gives rise two projections

$$p_{\pm} : H_{\pm} \rightarrow \mathbf{D}^{n+1}, \quad \mathbf{x} = (x^1, \dots, x^{n+2}) \mapsto (x^1, \dots, x^{n+1}).$$

Define $F : \Xi(\mathbf{S}^n) \rightarrow \mathbf{S}^{n+1}$ to be, $(\mathbf{x}, t) \in \mathbf{S}^n \times [0, 1]$,

$$F([\mathbf{x}, t]) := \begin{cases} p_-^{-1}(2t\mathbf{x} + (1-2t)\mathbf{p}_0), & 0 \leq t \leq \frac{1}{2}, \\ p_+^{-1}(2(1-t)\mathbf{x} + (2t-1)\mathbf{p}_0), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Since

$$F([\mathbf{x}, 0]) = p_-^{-1}(\mathbf{p}_0) = (1, 0, \dots, 0, 0) = p_+^{-1}(\mathbf{p}_0) = F([\mathbf{x}, 1]) = F([\mathbf{p}_0, t]) \in \mathbf{S}^{n+1},$$

it follows that $F \in \mathbf{Hom}_{\mathbf{Top}_*}(\Xi(\mathbf{S}^n), \mathbf{S}^{n+1})$. Solving $F \circ G = \mathbf{1}$ and $G \circ F = \mathbf{1}$, we can construct the inverse $G : \mathbf{S}^{n+1} \rightarrow \Xi(\mathbf{S}^n)$. Therefore $F : \Xi(\mathbf{S}^n) \rightarrow \mathbf{S}^{n+1}$ is homeomorphic. \square

Definition 3.2.23. For $n \geq 1$ the n -th homotopy group functor π_n is the covariant functor on \mathbf{HTop}_* defined by $\pi_n := \pi_{\mathbf{S}^n}$. For any $X \equiv (X, x_0) \in \mathbf{Ob}(\mathbf{HTop}_*)$, we call $\pi_n(X) \equiv \pi_n(X, x_0)$ the n -th homotopy group.

Note π_1 is just the fundamental group functor proved in **Subsection 3.3.2**.

Proposition 3.2.24. *A map $\alpha : \mathbf{S}^n \rightarrow X$ represents the trivial element of $\pi_n(X)$ for $n \geq 1$, where $X \in \mathbf{Ob}(\mathbf{Top}_*)$, if and only if α can be continuously extended over \mathbf{D}^{n+1} .*

PROOF. By **Theorem 3.1.27**. \square

3.2.6. Higher homotopy groups. If $P \equiv (P, \mu, \varphi) \in \mathbf{Ob}(\mathbf{HGroup})$ and $Q \equiv (Q, \nu, \psi) \in \mathbf{Ob}(\mathbf{HCGroup})$, then we have two different group multiplication on $[P, Q]$ given, respectively, by (3.2.2.3) and (3.2.4.4).

Proposition 3.2.25. *Let X and Y be two objects in a category \mathcal{C} and let $*$ and $*'$ be two laws of composition in $\mathbf{Hom}_{\mathcal{C}}(X, Y)$ such that*

(a) *$*$ and $*'$ have a common two-sided identity element, that is, there is a morphism $f_0 \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$ such that*

$$f * f_0 = f_0 * f = f = f *' f_0 = f_0 *' f,$$

for any $f \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$;

(b) $*$ and $'$ are mutually distributive, i.e., for any $f_1, f_2, g_1, g_2 \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$

$$(f_1 * f_2) *' (g_1 * g_2) = (f_1 *' g_1) * (f_2 *' g_2).$$

Then $*$ and $'$ are equal, and each is commutative and associative.

PROOF. For any $f, g \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$ one has

$$f * g = (f *' f_0) * (f_0 *' g) = (f * f_0) *' (f_0 * g) = f *' g$$

and

$$g * f = (f_0 *' g) * (f *' f_0) = (f_0 * f) *' (g * f_0) = f *' g.$$

Therefore $f * g = f *' g = g * f$. For any $f, g, h \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$ we have

$$(f * g) * h = (f * g) *' h = (f * g) *' (f_0 * h) = (f *' f_0) * (g *' h) = f * (g * h).$$

This proves the associativity of $*$ and then of $'$. \square

Corollary 3.2.26. *If $P = (P, \mu)$ is an H -space and $Q \equiv (Q, \nu, \psi)$ is an H -cogroup, then $[Q, P]$ is an Abelian group and the group structure is defined by the multiplication map in P .*

PROOF. Define two multiplications $*$ and $'$ in $[Q, P] = \mathbf{Hom}_{\mathbf{HTop}_*}(Q, P)$ by

$$[f] * [g] := [\mu \circ (f, g)], \quad [f] *' [g] := [(f, g) \circ \nu], \quad [f], [g] \in [Q, P].$$

Because $\mu \circ (c, 1) \simeq 1 \simeq \mu \circ (1, c)$, with $c := c_{p_0}$, we get

$$\mu \circ (f, c) \simeq f \simeq \mu \circ (c, f) \implies [f] * [c] = [f] = [c] * [f].$$

Similarly, using the definition of ν and noting that $f(q_0) = p_0$, we have

$$[f] *' [c] = [f] = [c] *' [f].$$

The mutual distribution of $*$ and $'$ can be proved, see [Exercise 3.2.29](#), in the same way.

According to [Proposition 3.2.25](#), we can conclude that $*$ and $'$ are equal, commutative and associative. \square

Corollary 3.2.27. *If $P = (P, \mu)$ is an H -space, then $\pi_n(P)$ is Abelian for all $n \geq 1$.*

PROOF. Recall from [Lemma 3.2.22](#) that

$$\pi_n(P) = \pi_{\mathbf{S}^n}(P) = [\mathbf{S}^n, P] = [\mathbb{E}(\mathbf{S}^{n-1}), P], \quad n \geq 1.$$

Using [Theorem 3.2.20](#) and [Corollary 3.2.26](#) yields that $\pi_n(P)$ is Abelian for any H -space and any $n \geq 1$. \square

To extend [Corollary 3.2.27](#) to general pointed topological spaces, we consider double suspensions

$$\pi_n(X) = [\mathbb{E}(\mathbf{S}^{n-1}), X] = \left[\mathbb{E} \left(\mathbb{E}(\mathbf{S}^{n-2}) \right), X \right], \quad n \geq 2.$$

Define two multiplications $*$ and $'$ on $[\mathbb{E}(\mathbb{E}(\mathbf{S}^{n-2})), P]$ as follows:

$$[f] * [g] := [f * g], \quad [f] *' [g] := [f *' g], \quad [f], [g] \in [\mathbb{E}(\mathbb{E}(\mathbf{S}^{n-2})), P],$$

where

$$(f * g)([z, t], s) := \begin{cases} f([z, 2t], s), & 0 \leq t \leq \frac{1}{2}, \\ g([z, 2t - 1], s), & \frac{1}{2} \leq t \leq 1, \end{cases}$$

and

$$(f *' g)([[z, t], s]) := \begin{cases} f([[z, t], 2s]), & 0 \leq s \leq \frac{1}{2}, \\ g([[z, t], 2s - 1]), & \frac{1}{2} \leq s \leq 1. \end{cases}$$

Let $c : \Xi(\Xi(\mathbf{S}^{n-2})) \rightarrow P$ be the constant map, i.e., $c([[z, t], s]) = p_0$.

Corollary 3.2.28. *For $n \geq 2$, $\pi_n : \mathbf{HTop}_* \rightarrow \mathbf{AG}$ is a covariant functor.*

PROOF. (1) $[f] * [c] = [f] = [f] *' [c]$. We first prove that $f * c \simeq f$. Consider the map

$$F : \Xi(\Xi(\mathbf{S}^{n-2})) \times [0, 1] \longrightarrow P, \quad ([[z, t], s], u) \longmapsto \begin{cases} f([[z, \frac{2t}{u+1}], s]), & 0 \leq t \leq \frac{u+1}{2}, \\ p_0, & \frac{u+1}{2} \leq t \leq 1. \end{cases}$$

When $t = \frac{u+1}{2}$, we see that $f([[z, \frac{2t}{u+1}], s])f([[z, 1], s]) = p_0$, since $[z, 1]$ is the base point of $\Xi(\mathbf{S}^{n-2})$. Hence F is continuous. Observe that

$$F([[z, t], s], 0) = f * c, \quad F([[z, t], s], 1) = f.$$

Hence $f * c \simeq f$. Similarly, one can prove $c * f \simeq f$ and $f *' c \simeq f \simeq c *' f$.

The mutual distribution of $*$ and $*'$ can be proved, see [Exercise 3.2.29](#), in the same way.

According to [Proposition 3.2.25](#), we can conclude that $*$ and $*'$ are equal, commutative and associative. Thus $\pi_n(X)$ is Abelian for all $n \geq 2$. \square

Exercise 3.2.29. Prove the mutual distribution in [Corollary 3.2.26](#) and [Corollary 3.2.28](#).

3.3. $\pi_1(\mathbf{S}^1)$

Recall some facts from categories. A category \mathcal{C} is said to be **small** if $\mathbf{Ob}(\mathcal{C})$ is a set. A morphism $f \in \mathbf{Hom}_{\mathcal{C}}(X, Y)$ is called an **equivalence** if there is a morphism $g \in \mathbf{Hom}_{\mathcal{C}}(Y, X)$ satisfying

$$g \circ f = \mathbf{1}_X, \quad f \circ g = \mathbf{1}_Y.$$

If $g' \in \mathbf{Hom}_{\mathcal{C}}(Y, X)$ is another morphism satisfying

$$g' \circ f = \mathbf{1}_X, \quad f \circ g' = \mathbf{1}_Y,$$

then

$$g' = g' \circ \mathbf{1}_Y = g' \circ (f \circ g) = (g' \circ f) \circ g = \mathbf{1}_X \circ g = g.$$

Thus, g is unique and denoted by f^{-1} . A **groupoid** \mathfrak{P} is a small category in which each morphism is an equivalence.

Proposition 3.3.1. *Let \mathfrak{P} be a groupoid.*

(1) *For any $A, B \in \mathbf{Ob}(\mathfrak{P})$ define a relation*

$$A \sim B \iff \mathbf{Hom}_{\mathfrak{P}}(A, B) \neq \emptyset.$$

*Then \sim is an equivalence relation on the set $\mathbf{Ob}(\mathfrak{P})$. The equivalence classes are called the **components** of \mathfrak{P} . The groupoid \mathfrak{P} is said to be **connected** if it has only one component.*

(2) *For any $A \in \mathbf{Ob}(\mathfrak{P})$, the law of composition*

$$\mathbf{Hom}_{\mathfrak{P}}(A, A) \times \mathbf{Hom}_{\mathfrak{P}}(A, A) \longrightarrow \mathbf{Hom}_{\mathfrak{P}}(A, A), \quad (g, f) \longmapsto f \circ g$$

is a group operation in $\mathbf{Hom}_{\mathfrak{P}}(A, A)$.

(3) There is a covariant functor $\mathbf{ad} : \mathfrak{P} \rightarrow \mathbf{Group}$ which assigns to an object $A \in \mathbf{Ob}(\mathfrak{P})$ the group $\mathbf{ad}(A) = \mathbf{Hom}_{\mathfrak{P}}(A, A)$ and to a morphism $f \in \mathbf{Hom}_{\mathfrak{P}}(A, B)$ the group homomorphism $\mathbf{ad}_f \in \mathbf{Hom}_{\mathbf{Group}}(\mathbf{ad}(A), \mathbf{ad}(B))$,

$$\mathbf{ad}_f : \mathbf{Hom}_{\mathfrak{P}}(A, A) \longrightarrow \mathbf{Hom}_{\mathfrak{P}}(B, B), \quad g \longmapsto \mathbf{ad}_f(g),$$

defined by $\mathbf{ad}_f(g) := f \circ g \circ f^{-1}$. Actually, \mathbf{ad}_f is a group isomorphism.

(4) If $A \sim B$ in \mathfrak{P} , the collection of isomorphisms $\{\mathbf{ad}_f : f \in \mathbf{Hom}_{\mathfrak{P}}(A, B)\}$ is a conjugacy class of isomorphisms $\mathbf{Hom}_{\mathfrak{P}}(A, A) \rightarrow \mathbf{Hom}_{\mathfrak{P}}(B, B)$.

(5) Choose another groupoid \mathfrak{P}' and let \mathbf{F} be a covariant functor from \mathfrak{P} to \mathfrak{P}' . Then \mathbf{F} maps each component of \mathfrak{P} into some component of \mathfrak{P}' .

PROOF. (1) Obvious.

(2) The identity element of $\mathbf{Hom}_{\mathfrak{P}}(A, A)$ is the identity morphism

$$\mathbf{1}_A \in \mathbf{Hom}_{\mathfrak{P}}(A, A).$$

It is clearly that the group axioms are satisfied.

(3) For any $g_1, g_2 \in \mathbf{Hom}_{\mathfrak{P}}(A, A)$, one has

$$\mathbf{ad}_f(\mathbf{1}_A) = f \circ \mathbf{1}_A \circ f^{-1} = f \circ f^{-1} = \mathbf{1}_A$$

and

$$\begin{aligned} \mathbf{ad}_f(g_1 \circ g_2) &= f \circ (g_1 \circ g_2) \circ f^{-1} \\ &= (f \circ g_1 \circ f^{-1}) \circ (f \circ g_2 \circ f^{-1}) = \mathbf{ad}_f(g_1) \circ \mathbf{ad}_f(g_2). \end{aligned}$$

Thus \mathbf{ad}_f is a group homomorphism. Hence \mathbf{ad} is a covariant functor from \mathfrak{P} to \mathbf{Group} . Since $\mathbf{ad}_f^{-1} = \mathbf{ad}_{f^{-1}}$, it follows that \mathbf{ad}_f is a group isomorphism.

(4) It follows from (3).

(5) Obvious. □

Exercise 3.3.2. Complete the proof of **Proposition 3.3.1**.

3.3.1. The fundamental groupoid. A path ω in a topological space X is a continuous map $\omega : [0, 1] \rightarrow X$; the point $\omega(0)$ is called the **origin** of ω , while the point $\omega(1)$ the **end** of ω . Two paths ω, ω' in X are said to be **homotopic** if they are homotopic relative to $\partial[0, 1] = \{0, 1\}$; in this case we denote by $\omega \simeq \omega'$ instead of $\omega \simeq \omega'(\text{rel } \{0, 1\})$. It is clear that “ \simeq ” is an equivalence relation so the resulting equivalence classes, denoted as $[\omega]$, are called **path classes**.

Recall the first fundamental group $(\pi_1(X, x), *)$ defined in (3.1.2.3). We now relax one point x to two points x_1, x_2 , and then define a “fundamental group”, called **fundamental groupoid**, which is a groupoid.

Given a topological space $X \in \mathbf{Ob}(\mathbf{Top})$, we define a category $\mathfrak{P}(X)$ whose objects are the points of X , whose morphisms from x_1 to x_0 , $x_0, x_1 \in X$, are the path class $[\omega]$ with x_0 as an origin and x_1 as end, and whose composite is the product of path classes:

$$[\omega] * [\omega'] = [\omega * \omega'], \quad [\omega] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x_0, x_1), \quad [\omega'] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x_1, x_2),$$

given by

$$(\omega * \omega')(s) := \begin{cases} \omega(2s), & 0 \leq s \leq \frac{1}{2}, \\ \omega(2s-1), & \frac{1}{2} \leq s \leq 1. \end{cases}$$

Theorem 3.3.3. $\mathfrak{P}(X)$ is a groupoid for any topological space $X \in \mathbf{Ob}(\mathbf{Top})$.

PROOF. (1) $\mathfrak{P}(X)$ is a category. To prove the associativity of the composite, consider

$$[\omega] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x_0, x_1), \quad [\omega'] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x_1, x_2), \quad [\omega''] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x_2, x_3).$$

We should verify $(\omega * \omega') * \omega'' \simeq \omega * (\omega' * \omega'')$. Define the corresponding homotopy $F : [0, 1] \times [0, 1] \rightarrow X$, $(s, t) \mapsto F(s, t)$, by

$$F(s, t) = \begin{cases} \omega\left(\frac{4s}{t+1}\right), & 0 \leq s \leq \frac{t+1}{4}, \\ \omega'(4s-t-1), & \frac{t+1}{4} \leq s \leq \frac{t+2}{4}, \\ \omega''\left(\frac{4s-t-2}{4}\right), & \frac{t+2}{4} \leq s \leq 1. \end{cases}$$

For any $x \in X$, we should construct the identity morphism $\mathbf{1}_x \in \mathbf{Hom}_{\mathfrak{P}(X)}(x, x)$. Let $c_x : [0, 1] \rightarrow X$ be the constant map given by $c_x(s) := x$ for any $s \in [0, 1]$. We claim that

$$[c_x] = \mathbf{1}_x \in \mathbf{Hom}_{\mathfrak{P}(X)}(x, x), \quad x \in \mathbf{Ob}(\mathfrak{P}(X)).$$

For any $[\omega] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x', x)$, define a homotopy $G : [0, 1] \times [0, 1] \rightarrow X$ by

$$G(s, t) = \begin{cases} \omega\left(\frac{2s}{t+1}\right), & 0 \leq s \leq \frac{t+1}{2}, \\ x, & \frac{t+1}{2} \leq s \leq 1. \end{cases}$$

Then $G : \omega * c_x \simeq \omega$. Thus $[\omega] * [c_x] = [\omega]$ for any $[\omega] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x', x)$. Similarly, we have

$$[c_x] * [\omega] = [\omega], \quad \text{for all } [\omega] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x, x').$$

(2) $\mathfrak{P}(X)$ is a groupoid. For any $[\omega] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x, x')$ we should prove that $[\omega]^{-1} = [\omega^{-1}]$, where $\omega^{-1}(s) := \omega(1-s)$. Equivalently, we should prove that $\omega * \omega^{-1} \simeq c_x$ and $\omega^{-1} * \omega \simeq c_{x'}$. We give a proof of the first claim. The corresponding homotopy $H : [0, 1] \times [0, 1] \rightarrow X$ is

$$H(s, t) = \begin{cases} x, & 0 \leq s \leq \frac{t}{2}, \\ \omega(2s-t), & \frac{t}{2} \leq s \leq \frac{1}{2}, \\ \omega(2-2s-t), & \frac{1}{2} \leq s \leq 1 - \frac{t}{2}, \\ x, & 1 - \frac{t}{2} \leq s \leq 1. \end{cases}$$

In summary, $\mathfrak{P}(X)$ is a groupoid. \square

The category $\mathfrak{P}(X)$ is called the **category of path classes** of X or the **fundamental groupoid** of X . From the definition, we see that the components of the fundamental groupoid are path components of X .

Given a continuous map $f \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$, we can construct a covariant functor $f_{\#} : \mathfrak{P}(X) \rightarrow \mathfrak{P}(Y)$ as follows. For $x \in \mathbf{Ob}(\mathfrak{P}(X))$, let $f_{\#}(x) = f(x) \in \mathbf{Ob}(\mathfrak{P}(Y))$, and for $[\omega] \in \mathbf{Hom}_{\mathfrak{P}(X)}(x, x')$, let

$$f_{\#}([\omega]) := [f \circ \omega] \in \mathbf{Hom}_{\mathfrak{P}(Y)}(f(x), f(x')).$$

Then, $f_{\#}$ maps each path component of X into some path component of Y . Hence, we obtain a covariant functor

$$(3.3.1.1) \quad \pi_0 : \mathbf{Top} \longrightarrow \mathbf{Set}$$

with

$$\pi_0(X) = \{\text{path components of } X\}$$

and for $f \in \mathbf{Hom}_{\mathbf{Top}}(X, Y)$,

$$\pi_0(f) := f_{\#} : \pi_0(X) \longrightarrow \pi_0(Y)$$

maps the path component of x in X to the path component of $f(x)$ in Y . The above covariant functor (3.3.1.1) can be regarded as a covariant functor

$$(3.3.1.2) \quad \pi_0 : \mathbf{HTop} \longrightarrow \mathbf{Set}$$

with

$$\pi_0(X) = \{\text{path components of } X\}$$

and for $[f] \in \mathbf{Hom}_{\mathbf{HTop}}(X, Y)$,

$$\pi_0([f]) := f_{\#} : \pi_0(X) \longrightarrow \pi_0(Y).$$

Indeed, if $F : f \simeq f'$, then, for any $x \in X$, the path $\omega_x(t) := F(x, t)$ connects $f(x)$ to $f'(x)$. Therefore $f(x)$ and $f'(x)$ belong to the same path component of Y and then $f_{\#} = f'_{\#}$.

Proposition 3.3.4. *If X is contractible, then π_X and π_0 are naturally equivalent functors on \mathbf{HTop} .*

PROOF. Since X and one-point space P have the same homotopy type, it suffices to prove that π_P and π_0 are naturally equivalent functors on \mathbf{HTop} . For any $X \in \mathbf{Ob}(\mathbf{HTop})$, define

$$\psi(X)([g]) := g_{\#}(P), \quad \text{for all } [g] \in \pi_P(X) = [P, X].$$

Here we observe that $\pi_0(P)$ consists of the only path component P . Consider the digram

$$\begin{array}{ccc} \pi_P(X) & \xrightarrow{\psi(X)} & \pi_0(X) \\ \pi_P([f]) \downarrow & & \downarrow \pi_0([f]) \\ \pi_P(Y) & \xrightarrow[\psi(Y)]{} & \pi_0(Y) \end{array}$$

Because, for any $[g] \in \pi_P(X)$,

$$(\pi_0([f]) \circ \psi(X))([g]) = \pi_0([f])(g_{\#}(P)) = f_{\#}(g_{\#}(P)) = (f \circ g)_{\#}(P)$$

and

$$(\psi(Y) \circ \pi_P([f]))([g]) = \psi(Y)(f_{\#}([g])) = \psi(Y)([f \circ g]) = (f \circ g)_{\#}(P).$$

Thus ψ is a natural transformation. \square

3.3.2. The equivalent of π_1 . From [Lemma 3.1.22](#) we obtain a covariant functor

$$(3.3.2.1) \quad \pi : \mathbf{HTop}_* \longrightarrow \mathbf{Group}$$

given by, for $X \equiv (X, x_0) \in \mathbf{Ob}(\mathbf{HTop}_*)$,

$$\pi(X) := \pi_1(X, x_0)$$

and, for $[f] \in \mathbf{Hom}_{\mathbf{HTop}_*}(X, Y)$,

$$\pi([f]) := f_* : \pi_1(X, x_0) \longrightarrow \pi_1(Y, y_0).$$

Recall from [Definition 3.2.23](#) that $\pi_1 := \pi_{\mathbf{S}^1}$ is a covariant functor from \mathbf{HTop}_* to \mathbf{Group} .

Theorem 3.3.5. *Two covariant functors $\pi, \pi_1 : \mathbf{HTop}_* \rightarrow \mathbf{Group}$ are naturally equivalent.*

PROOF. Define a continuous map $\lambda : [0, 1] \rightarrow \Xi(\mathbf{S}^0)$ by $\lambda(s) := [-1, s]$, where $\mathbf{S}^0 = \{-1, 1\}$ and 1 is the base point of \mathbf{S}^0 . Then λ induces, by the definition of $\Xi(\mathbf{S}^0) \cong \mathbf{S}^1$, a bijection

$$\lambda^\# : [\Xi(\mathbf{S}^0), X] = \pi_1(X) \longrightarrow \pi(X) = \pi_1(X, x_0), \quad [g] \longmapsto [g \circ \lambda].$$

Moreover we see that $\lambda^\#$ is a natural transformation between π_1 and π . \square

A topological space $X \in \mathbf{Ob}(\mathbf{Top})$ is said to be **n -connected** for $n \geq 0$ if every continuous map $f : \mathbf{S}^k \rightarrow X$, for $k \leq n$, has a continuous extension over \mathbf{D}^{k+1} .

- (1) A 1-connected space is also said to be **simply-connected**.
- (2) If $0 \leq m \leq n$, an n -connected space is m -connected.
- (3) According to [Proposition 3.2.24](#), a topological space X is n -connected for $n \geq 1$ if and only if it is path connected and $\pi_k(X, x)$ is trivial for every base point $x \in X$ and $1 \leq k \leq n$.
- (4) From [Corollary 3.1.28](#), any contractible space is n -connected for every $n \geq 0$.

3.3.3. $\pi_1(\mathbf{S}^1) = \mathbf{Z}$. We compute the fundamental group $\pi(\mathbf{S}^1) = \pi_1(\mathbf{S}^1, 1)$ of $\mathbf{S}^1 = \{e^{\sqrt{-1}\theta}\}_{\theta \in [0, 2\pi]}$ following [Tucker](#) (1945). Consider the exponential map

$$(3.3.3.1) \quad \exp : \mathbf{R} \longrightarrow \mathbf{S}^1, \quad t \longmapsto \exp(t) = e^{2\pi\sqrt{-1}t} \equiv e(t).$$

Observe that

$$e(t_1 + t_2) = e(t_1)e(t_2), \quad e(t_1) = e(t_2) \iff t_1 - t_2 \in \mathbf{Z}.$$

Moreover, $e|_{(-1/2, 1/2)}$ is a homeomorphism of $(-\frac{1}{2}, \frac{1}{2})$ onto $\mathbf{S}^1 \setminus \{e^{\pi\sqrt{-1}}\}$. We then can define the inverse of $e|_{(-1/2, 1/2)}$ to be

$$(3.3.3.2) \quad e^{-1} : \mathbf{S}^1 \setminus \{e^{\pi\sqrt{-1}}\} \longrightarrow \left(-\frac{1}{2}, \frac{1}{2}\right).$$

Recall that a subset $X \subseteq \mathbf{R}^n$ is **star-shaped with respect to $x_0 \in X$** if $\overline{x_0x} \subseteq X$ whenever $x \in X$.

Lemma 3.3.6. *Let $X \subset \mathbf{R}^n$ be a star-shaped (with respect to $x_0 \in X$), compact, and continuous set. For any given continuous map $f : X \rightarrow \mathbf{S}^1$ with $e(t_0) = f(x_0)$ for some $t_0 \in \mathbf{R}$, we can find a unique continuous map $\tilde{f} : X \rightarrow \mathbf{R}$ such that*

$$(3.3.3.3) \quad \tilde{f}(x_0) = t_0, \quad e \circ \tilde{f} = f.$$

Thus, f can be uniquely lifted to \mathbf{R} :

$$\begin{array}{ccc} & & \mathbf{R} \\ & \nearrow \exists! \tilde{f} & \downarrow e \\ X & \xrightarrow{f} & \mathbf{S}^1 \end{array}$$

PROOF. (1) **Uniqueness.** Let $\tilde{g} : X \rightarrow \mathbf{R}$ be another continuous map satisfying $\tilde{g}(x_0) = t_0$ and $e \circ \tilde{g} = f$. Then $e \circ (\tilde{f} - \tilde{g}) = 0$ and $(\tilde{f} - \tilde{g})(x_0) = 0$. Thus $\tilde{f} - \tilde{g}$ is a continuous map from X onto \mathbf{Z} . The continuity of X implies that $\tilde{f} - \tilde{g}$ is constant so that $\tilde{f} = \tilde{g}$ on X .

(2) **Existence.** Because the parallel translation is continuous, we may without loss of generality assume that $x_0 = \mathbf{0}$. Since X is compact, it follows that f is uniformly continuous. There exists $\delta > 0$ such that

$$|f(x) - f(y)| < 2 \quad \text{whenever} \quad |x - y| < \delta$$

(in particular, $f(x)$ and $f(y)$ are not antipodes in \mathbf{S}^1). The boundedness of X implies that there exists $n \in \mathbf{N}$ such that $|x|/n < \delta$ for all $x \in X$. For each $j \in \{0, 1, \dots, n-1\}$ and all $x \in X$, we have

$$\left| f\left(\frac{j+1}{n}x\right) - f\left(\frac{j}{n}x\right) \right| < 2$$

because $|(j+1)x/n - jx/n| = |x|/n < \delta$. Particularly, $f((j+1)x/n) \neq -f(jx/n)$. Then we, for each $0 \leq j \leq n-1$, can define a continuous map

$$g_j : X \longrightarrow \mathbf{S}^1 \setminus \{e^{\pi\sqrt{-1}}\}, \quad x \longmapsto \frac{f((j+1)x/n)}{f(jx/n)}.$$

Consequently,

$$\prod_{0 \leq j \leq n-1} g_j(x) = \frac{f(x)}{f(0)} \quad \text{or} \quad f(x) = f(0) \prod_{0 \leq j \leq n-1} g_j(x).$$

Define $\tilde{f} : X \rightarrow \mathbf{R}$ to be $\tilde{f}(x) := e^{-1}(f(x))$; precisely,

$$\tilde{f}(x) := t_0 + \sum_{0 \leq j \leq n-1} e^{-1}(g_j(x)).$$

Then $\tilde{f}(0) = t_0$ and $e \circ \tilde{f} = f$. □

Taking $X = [0, 1]$ in **Lemma 3.3.6** yields that for any loop $\alpha : [0, 1] \rightarrow \mathbf{S}^1$ based at $p_0 = 1$, there exists a unique continuous map $\tilde{\alpha} : [0, 1] \rightarrow \mathbf{R}$ such that

$$\tilde{\alpha}(0) = 0, \quad e \circ \tilde{\alpha} = \alpha.$$

In particular,

$$e(\tilde{\alpha}(1)) = \alpha(1) = p_0 = 1.$$

Thus $\tilde{\alpha}(1) \in \mathbf{Z}$. The **degree** of α is now given by

$$(3.3.3.4) \quad \deg(\alpha) := \tilde{\alpha}(1).$$

Lemma 3.3.7. For $[\alpha] = [\beta] \in \pi_1(\mathbf{S}^1, p_0)$, we have

$$(3.3.3.5) \quad \det(\alpha) = \deg(\beta).$$

Hence we have a well-defined function $\deg : \pi_1(\mathbf{S}^1, p_0) \rightarrow \mathbf{Z}$ given by

$$(3.3.3.6) \quad \deg([\alpha]) := \det(\alpha).$$

PROOF. Let $F : \alpha \simeq \beta$ be a homotopy relative to $\partial[0, 1] = \{0, 1\}$ from α to β . Then F is a continuous map from $[0, 1] \times [0, 1] \rightarrow \mathbf{S}^1$. By **Lemma 3.3.6**, there is a unique continuous map $\tilde{F} : [0, 1] \times [0, 1] \rightarrow \mathbf{R}$ such that

$$\tilde{F}(0, 0) = 0, \quad r \circ \tilde{F} = F.$$

Because $\tilde{F}(0, t) : [0, 1] \rightarrow \mathbf{R}$ is continuous and $e(\tilde{F}(0, t)) = F(0, t) = p_0 = 1$, we obtain $\tilde{F}(0, t)$ is constant for all $t \in [0, 1]$ and then $\tilde{F}(0, t) = \tilde{F}(0, 0) = 0$. Similarly one has $\tilde{F}(1, t) = \tilde{F}(1, 1)$ for all $t \in [0, 1]$. Since

$$e(\tilde{F}(s, 0)) = F(s, 0) = \alpha(s), \quad e(\tilde{F}(s, 1)) = F(s, 1) = \beta(s)$$

the uniqueness in **Lemma 3.3.6** implies that $\tilde{\alpha}(s) = \tilde{F}(s, 0)$ and $\tilde{\beta} = \tilde{F}(s, 1)$. In particular,

$$\tilde{\alpha}(1) = \tilde{F}(1, 0) = \tilde{F}(1, 1) = \tilde{\beta}(1).$$

Thus $\deg(\alpha) = \deg(\beta)$. □

Theorem 3.3.8. The function \deg is a group isomorphism

$$(3.3.3.7) \quad \deg : \pi_1(\mathbf{S}^1, p_0) \longrightarrow \mathbf{Z}, \quad [\alpha] \longmapsto \deg(\alpha).$$

PROOF. For any $[\alpha], [\beta] \in \pi_1(\mathbf{S}^1, p_0)$, we have continuous maps $\tilde{\alpha}, \tilde{\beta} : [0, 1] \rightarrow \mathbf{R}$ such that

$$\tilde{\alpha}(0) = \tilde{\beta}(0) = 0, \quad e \circ \tilde{\alpha} = \alpha, \quad e \circ \tilde{\beta} = \beta.$$

Since $(\tilde{\alpha} + \tilde{\beta})(0) = 0$ and $e \circ (\tilde{\alpha} + \tilde{\beta}) = \alpha\beta$, it follows from **Lemma 3.3.6** that $\widetilde{\alpha\beta} = \tilde{\alpha} + \tilde{\beta}$. Therefore

$$\deg([\alpha] * [\beta]) = \deg([\alpha\beta]) = \widetilde{\alpha\beta}(1) = \tilde{\alpha}(1) + \tilde{\beta}(1) = \deg([\alpha]) + \deg([\beta])$$

where the multiplication is determined by (3.2.2.3). Hence \deg is a group homomorphism.

If $\deg([\alpha]) = 0$, then the lifted map $\tilde{\alpha} : [0, 1] \rightarrow \mathbf{R}$ is also a loop based at 0 and satisfies $e \circ \tilde{\alpha} = \alpha$. Since $\tilde{\alpha}$ can be viewed as a continuous map from \mathbf{S}^1 to \mathbf{R} and \mathbf{R} is contractible, it follows from **Corollary 3.1.28** that $\tilde{\alpha} \simeq c_0$ and hence $e \circ \tilde{\alpha} \simeq c_{p_0}$. Hence $\alpha \simeq c_{p_0}$. Thus \deg is injective.

For any $n \in \mathbf{Z}$, define $\tilde{\alpha}(s) := ns$ and $\alpha(s) := e(\tilde{\alpha}(s))$. Then

$$\alpha(0) = e(\tilde{\alpha}(0)) = e(0) = p_0, \quad \alpha(1) = e(\tilde{\alpha}(1)) = e(n) = p_0.$$

Hence $\deg(\alpha) = \tilde{\alpha}(1) = n$. □