Deep Dive

Luiz G. Mugnaini A.

Last modification: March 4, 2023

Contents

Todo list

Preface

First of all I should stress that the title page template was taken from the beautiful book [ADH21] by P. Haine et al. — the whole credit should be given to them. The following text is a compilation of notes for my studies in mathematics, that being said, it is clear that the collection of mathematical and linguistic errors is non-empty. My main references are

- Category Theory: [Rie16] and [KS06].
- Algebra: [MK89], [Kim20], [Alu09] and [Lan93].
- Topology: [Lee11], [BBT20], [Mun00] and [Eng89].
- Differential Structures: [Zor15], [Zor16], [Rud76] and [Jos06].
- Combinatorics: [Die16].

Part I

Category Theory

Chapter 1

Categories

1.1 Sets and Universes

In order to deal with size issues in the theory of categories, we shall work inside what are called *universes*.

Definition 1.1.1 (Universe)**.** A *universe* $\mathcal{U}$ is defined as a *set* satisfying the following properties

(U1) $\varnothing \in \mathcal{U}$.

(U2) If $u \in \mathcal{U}$, then $u \subseteq \mathcal{U}$.

(U3) If $u \in \mathcal{U}$, then $\{u\} \in \mathcal{U}$.

(U4) If $u \in \mathcal{U}$, then $2^u \in \mathcal{U}$.

(U5) If we have an indexing set $I \in \mathcal{U}$, for which we associate a collection $\{u_i \in \mathcal{U}\}_{i \in I}$, then $\bigcup_{i \in I} u_i \in \mathcal{U}$.

(U6) $\mathbf{N} \in \mathcal{U}$.

As a consequence of such properties, a universe also satisfies

(U7) If $u \in \mathcal{U}$, then $\bigcup_{x \in u} x \in \mathcal{U}$.

(U8) Given $u, v \in \mathcal{U}$, we have $u \times v \in \mathcal{U}$.

(U9) If $v \in \mathcal{U}$, and we have $u \subseteq v$, then $u \in \mathcal{U}$.

(U10) If $I \in \mathcal{U}$ is an idexing set with an associated collection $\{u_i \in \mathcal{U}\}_{i \in I}$, then $\prod_{i \in I} u_i \in \mathcal{U}$.

Axiom 1.1.2 (Grothendieck's axiom to ZF set theory)**.** For any set X, there exists an universe $\mathcal{U}$ for which $X \in \mathcal{U}$.

For some commonly used terminology, we say that a set x is a $\mathcal{U}$-set if $x \in \mathcal{U}$. Moreover, a set x is called $\mathcal{U}$-small if it is isomorphic to some set $s \in \mathcal{U}$.

Orderings are another important topic when dealing with sets, we shall define now some of these concepts.

Definition 1.1.3. Let I be a set. We define the following:

- An *order* on the set I is a relation $\leqslant$ satisfying the following properties:
    (a) The order is *reflexive* — that is, for all $i \in I$, we have $i \leqslant i$.
    (b) The order is *transitive* — given any three elements $i, j, k \in I$ such that $i \leqslant j$ and $j \leqslant k$, it follows that $i \leqslant k$.
    (c) The order is *anti-symmetric* — given any two elements $i, j \in I$, if $i \leqslant j$ and $j \leqslant i$, then $i = j$.
- An order is said to be *directed*, or *filtrant*, if I is non-empty and, for every $i, j \in I$, there exists $k \in I$ for which $i \leqslant k$ and $j \leqslant k$.
- An order is said to be *total* if, given any $i, j \in I$, necessarily have at least one of the following relations: $i \leqslant j$ or $j \leqslant i$.
- The set I is said to be *inductively ordered* if, for any totally ordered subset $J \subseteq I$, J has an upper bound $u \in I$ for which $j \leqslant u$ for all $j \in J$
- If I is ordered by the relation $\leqslant$, we define a strict relation $<$ as, given $i, j \in I$, we have $i \leqslant j$ if and only if $i \leqslant j$ and $i \neq j$.

1.2 Categories

Definition 1.2.1 (Category)**.** A category C consists of the following data

(C1) A collection of objects. We say that X is an object of C by writing $X \in \mathsf{C}$ or $X \in \mathrm{Obj}(\mathsf{C})$[1].

(C2) For every given pair of objects $X, Y \in \mathsf{C}$ there exists a collection of morphisms $\mathrm{Mor}_{\mathsf{C}}(X, Y)$ with source X and target Y. The collection of morphisms between objects of C is denoted $\mathrm{Mor}(\mathsf{C})$.

(C3) For every object $X \in \mathsf{C}$, there exists an identity morphism $\mathrm{id}_X \in \mathrm{Mor}_{\mathsf{C}}(X, X)$.

(C4) For every triple of given objects $X, Y, Z \in \mathsf{C}$, there exists a composition map

$$\mathrm{Mor}_{\mathsf{C}}(X, Y) \times \mathrm{Mor}_{\mathsf{C}}(Y, Z) \to \mathrm{Mor}_{\mathsf{C}}(X, Z).$$

So that for given morphisms $f \in \mathrm{Mor}_{\mathsf{C}}(X, Y)$ and $g \in \mathrm{Mor}_{\mathsf{C}}(Y, Z)$ there exists a uniquely defined map $gf \in \mathrm{Mor}_{\mathsf{C}}(X, Z)$ such that the following diagram commutes

$$\begin{tikzcd} Y \arrow[r, "g"] & Z \\ X \arrow[u, "f"] \arrow[ur, dashed, "gf"'] & \end{tikzcd}$$

[1]We shall adopt the former notation, which should not cause any confusion.

(C5) For every morphism $f\colon X \to Y$ we have that

$$\begin{tikzcd} X \arrow[loop left, "\mathrm{id}_X"] \arrow[r, "f"] & Y \arrow[loop right, "\mathrm{id}_Y"] \end{tikzcd}$$

so that $\mathrm{id}_Y f = f = f\,\mathrm{id}_X$.

(C6) Given objects $W, X, Y, Z \in \mathsf{C}$, the following diagram commutes

$$\begin{tikzcd} W \arrow[r, "f"] \arrow[rrr, bend left=50, "(hg)f"] \arrow[rr, bend right, "gf"'] \arrow[rrr, bend right=50, "h(gf)"'] & X \arrow[r, "g"] \arrow[rr, bend left, "hg"] & Y \arrow[r, "h"] & Z \end{tikzcd}$$

that is, $h(gf) = (hg)f$.

Notation 1.2.2 (On arrows and diagrams)**.** I'll adopt throughout this whole text a series of notations regarding morphisms, diagrams and so on, here I collect some of those: given a category C and objects $x, y, z \in \mathsf{C}$

- An arrow $x \rightarrowtail y$ denotes a *monomorphism* — to be seen in Definition 1.3.3.
- An arrow $x \twoheadrightarrow y$ denotes an *epimorphism* — to be seen in Definition 1.3.4.
- An arrow $x \xrightarrow{\simeq} y$ denoted an *isomorphism* — to be seen in Definition 1.2.6. We say that x is isomorphic to y, and write $x \simeq y$, if there exists an isomorphism $x \xrightarrow{\simeq} y$.
- Given arrows $f\colon x \to y$ and $g\colon y \to z$ in C, we denote the *composition* of f with g by the *juxtaposition* $gf\colon x \to z$[2].
- Let the following be a commutative diagram on C (that is, $fg = h$):

$$\begin{tikzcd} x \arrow[r, "f"] \arrow[dr, "h"'] & y \arrow[d, dashed, "g"] \\ & z \end{tikzcd}$$

  The *dashed* arrow $g\colon y \to z$ denotes that g is the *unique* morphism between y and z in the category C such that $fg = h$.
- The diagram in C

$$\begin{tikzcd} x \arrow[r, shift left, "g"] \arrow[r, shift right, "h"'] & y \arrow[r, "f"] & z \end{tikzcd}$$

  denotes that $fg = fh$.

Definition 1.2.3 (Small)**.** A category C is said to be small if it is composed of a set's worth of morphisms.

[2]When need be, we may use the symbol $g \circ f$ to denote the composition gf.

Definition 1.2.4 (Locally small)**.** A category C is said to be locally small if for any objects $A, B \in \mathsf{C}$ there exists a set's worth of morphisms A and B.

Corollary 1.2.5 (Unique identity)**.** Given a category C and an object $c \in \mathsf{C}$, the identity $\mathrm{id}_c \in \mathrm{Mor}(\mathsf{C})$ is unique.

Proof. Let $f : c \to c$ be an identity of c, then $f = f\,\mathrm{id}_c = \mathrm{id}_c$. ♮

Definition 1.2.6 (Isomorphism)**.** Given a category C and objects $A, B \in \mathsf{C}$, we define a morphism $f \in \mathrm{Mor}(A, B)$ to be an *isomorphism* if and only if it has a both sided inverse, so that exists $f^{-1} \in \mathrm{Mor}(B, A)$ such that $f^{-1}f = \mathrm{id}_A$ and $ff^{-1} = \mathrm{id}_B$.

Proposition 1.2.7. Given an isomorphism f, its inverse is unique.

Proof. Suppose for instance that there are two such functions, $g, h \in \mathrm{Mor}(B, A)$, that act as an inverse for $f \in \mathrm{Mor}(A, B)$. Note that

$$g = g\,\mathrm{id}_B = g(fh) = (gf)h = \mathrm{id}_A\,h = h$$

Thus $g = h$ and therefore the inverse is indeed unique. ♮

Definition 1.2.8 (Non-empty category)**.** A category is said to be *non-empty* if the collection of objects is non-empty.

Definition 1.2.9 (Discrete category)**.** A category is said to be *discrete* if all morphisms are the identity morphisms.

Example 1.2.10. Given a set I, one can view I as a category whose objects are the elements of I and morphisms are identities — that is, $\mathrm{Mor}_I(x, y)$ is the identity when $x = y$, while empty if $x \neq y$.

Definition 1.2.11 (Finite category)**.** A category is said to be *finite* if the collection of all morphisms is a finite set. This category will be of important use when dealing with limits.

Definition 1.2.12 (Connected category)**.** A category C is said to be connected if it is non-empty and, for every pair $x, y \in \mathsf{C}$, there exists a finite sequence of objects $(x_0, \dots, x_n)$, $x_j \in C$ for all $0 \leqslant j \leqslant n$, such that $x_0 = x$, $x_n = y$ and at least one of the collections of morphisms $\mathrm{Mor}_{\mathsf{C}}(x_j, x_{j+1})$ or $\mathrm{Mor}_{\mathsf{C}}(x_{j+1}, x_j)$ is non-empty for every $0 \leqslant j \leqslant n-1$

Definition 1.2.13 (Monoid)**.** A monoid is a set M equipped with a binary operation $\otimes : M \times M \to M$ and a neutral element $e \in M$. The binary operation is associative and obeys the right and left unit laws, that is

$$x \otimes (y \otimes z) = (x \otimes y) \otimes z \qquad \text{and} \qquad e \otimes x = x = x \otimes e.$$

Example 1.2.14. A monoid M defines a category with one object, denoted $\mathtt{BM}$, such that $\mathrm{Obj}(\mathtt{BM}) = \{*\}$ and $\mathrm{Mor}_{\mathtt{BM}}(*, *) = M$. Composition of morphisms $f, g : * \to *$ is defined as $gf = g \otimes f$. The identity morphism is $\mathrm{id}_* = e$. Hence we have $e * f = f = f * e$ for all morphisms $f \in \mathrm{Mor}(\mathtt{BM})$.

Definition 1.2.15 (Groupoids)**.** A *groupoid* is the name given to a category in which all of its morphisms are isomorphisms.

Definition 1.2.16 (Group)**.** A group is a groupoid with one object.

Definition 1.2.17 (Automorphism)**.** Given a category $\mathtt{C}$ and an object $A \in \mathtt{C}$, we define an automorphism of A to be an isomorphism $A \to A$. The set consisting of all such automorphisms of this object is denoted $\mathrm{Aut}(A)$, the automorphisms have properties:

i. The composition of two automorphisms is an automorphism.

ii. Composition is associative.

iii. $\mathrm{id}_A \in \mathrm{Aut}(A)$.

iv. Every automorphism $f \in \mathrm{Aut}(A)$ has an inverse $f^{-1} \in \mathrm{Aut}(A)$.

With this, the structure $\mathrm{Aut}(A)$ is a *group*, for all choices of objects $A \in \mathtt{C}$.

Definition 1.2.18 (Subcategory)**.** Let $\mathtt{C}$ be a category. We define $\mathtt{D} \subseteq \mathtt{C}$ as a subcategory of $\mathtt{C}$ if: $\mathrm{Obj}(\mathtt{D})$ is a restriction of $\mathrm{Obj}(\mathtt{C})$; for all $A \in \mathrm{Obj}(\mathtt{D})$ there exists $\mathrm{id}_A \in \mathrm{Mor}(\mathtt{D})$; for any $f \in \mathrm{Mor}(\mathtt{D})$ there exists $\mathrm{dom}(f), \mathrm{cod}(f) \in \mathrm{Obj}(\mathtt{D})$; for any composable pair of morphisms $f, g \in \mathrm{Mor}(\mathtt{D})$ there exists $fg \in \mathrm{Mor}(\mathtt{D})$.

Lemma 1.2.19. Any category $\mathtt{C}$ contains a subcategory containing all of the objects and whose morphisms are only the isomorphisms. Such subcategory is called a maximal groupoid.

Proof. We are going to prove that the maximal groupoid, name it $\mathtt{G}$ is indeed a subcategory of $\mathtt{C}$. Notice that $\mathrm{Obj}(\mathtt{G}) = \mathrm{Obj}(\mathtt{C})$, moreover every identity is an isomorphism, then $\mathrm{id}_* \in \mathrm{Mor}(\mathtt{G})$. Let $f \in \mathrm{Mor}(\mathtt{G})$ then in particular we have $f \in \mathrm{Mor}(\mathtt{C})$ and hence $\mathrm{dom}\, f, \mathrm{codom}\, f \in \mathrm{Obj}(\mathtt{C}) = \mathrm{Obj}(\mathtt{G})$. Consider $f \in \mathrm{Mor}_{\mathtt{G}}(A, B)$ and $g \in \mathrm{Mor}_{\mathtt{G}}(B, C)$, composable isomorphisms, and notice that $f^{-1}g^{-1} \in \mathrm{Mor}_{\mathtt{C}}(C, A)$ is such that $(f^{-1}g^{-1})(gf) = \mathrm{id}_A$ and $(gf)(f^{-1}g^{-1}) = \mathrm{id}_C$, hence we conclude that gf is an isomorphism and therefore $gf \in \mathrm{Mor}_{\mathtt{G}}(A, C)$. ♮

Definition 1.2.20 (Full subcategory)**.** A subcategory $\mathtt{D}$ of $\mathtt{C}$ (see Definition 1.2.18) is said to be a *full subcategory* if for all $x, y \in \mathtt{D}$ we have $\mathrm{Mor}_{\mathtt{D}}(x, y) = \mathrm{Mor}_{\mathtt{C}}(x, y)$. The full subcategory $\mathtt{D}$ is said to be *saturated* if $x \in \mathtt{C}$ is also an object of $\mathtt{D}$ whenever there exists an object $u \in \mathtt{D}$ such that $x \simeq u$ in C.

Definition 1.2.21 (Skeleton)**.** Given a category $\mathtt{C}$, we define the *skeleton* of $\mathtt{C}$, denoted by $\mathrm{sk}\,\mathtt{C}$, to be the full subcategory of $\mathtt{C}$ such that, for every $c \in \mathtt{C}$, there exists a *unique* object $s \in \mathrm{sk}\,\mathtt{C}$ such that $c \simeq s$ in $\mathtt{C}$. Furthermore, we say that a category $\mathtt{D}$ is skeletal if $\mathtt{D} = \mathrm{sk}\,\mathtt{D}$.

Example 1.2.22 (Set based categories)**.** The following are some important categories regarding sets:

- The category $\mathtt{Set}$ consists of $\mathcal{U}$-sets and set-maps between such objects.

- The category consisting of *finite* $\mathcal{U}$-sets and set-maps between them is a full subcategory of $\mathtt{Set}$, we denote it by $\mathtt{FinSet}$.
- We also define the category $\mathtt{pSet}$ of *pointed* $\mathcal{U}$-sets, that is, objects are pairs (X, x), where X is a $\mathcal{U}$-set and $x \in X$. Morphisms $f\colon (X, x) \to (Y, y)$ are defined to be maps $f\colon X \to Y$ such that $f(x) = y$.

Remark 1.2.23. The structure consisting of all sets and the set-maps between them does *not* shape a category, since the collection of all sets is not itself a set.

Example 1.2.24. Let $(I, \leqslant)$ be an ordered set. We define a category $\mathtt{I}$ associated with I to consist of the collection of objects contained in I, and

$$\mathrm{Mor}_{\mathtt{I}}(i, j) = \begin{cases} \{*\}, & \text{if } i \leqslant j \\ \varnothing, & \text{otherwise} \end{cases}$$

Definition 1.2.25 (Morphism category)**.** Let $\mathtt{C}$ be a category. We'll denote by $\mathrm{Mor}(\mathtt{C})$ the category whose objects are *morphisms* in $\mathtt{C}$ and whose morphisms between given objects $f\colon x \to y$ and $g\colon z \to w$ are pairs of morphisms (u, v), with $u\colon x \to z$ and $v\colon y \to w$ such that the following diagram commutes

$$\begin{tikzcd} x \arrow[r, "u"] \arrow[d, "f"'] & z \arrow[d, "g"] \\ y \arrow[r, "v"] & w \end{tikzcd}$$

That is, $\mathrm{Mor}_{\mathrm{Mor}(\mathtt{C})}(f, g) \coloneqq \{(u, v) \in \mathrm{Mor}_{\mathtt{C}}(x, z) \times \mathrm{Mor}_{\mathtt{C}}(y, w) : gu = vf\}$.

Proposition 1.2.26 (Slice category)**.** Given a category $\mathtt{C}$ and an object $c \in \mathtt{C}$. The following define categories:

(SC1) (Slice under c) A category $c/\mathtt{C}$, called slice category of $\mathtt{C}$ under c, whose objects are morphisms $f \in \mathrm{Mor}_{\mathtt{C}}(c, *)$. Given objects $f, g \in c/\mathtt{C}$ such that $f\colon c \to x$ and $g\colon c \to y$, we define a morphism $f \to g$ as a map $h\colon x \to y$ such that the following diagram commutes

$$\begin{tikzcd} & c \arrow[dl, "f"'] \arrow[dr, "g"] & \\ x \arrow[rr, "h"] & & y \end{tikzcd}$$

that is, $g = hf$.

(SC2) (Slice over c) A category $\mathtt{C}/c$, called the slice category of $\mathtt{C}$ over c, whose objects are morphisms $f \in \mathrm{Mor}_{\mathtt{C}}(*, c)$. Morphisms between objects $f, g \in \mathtt{C}/c$ such that $f\colon x \to c$ and $g\colon y \to c$ are maps $h\colon x \to y$ such that the following diagram commutes

$$\begin{tikzcd} x \arrow[rr, "h"] \arrow[dr, "f"'] & & y \arrow[dl, "g"] \\ & c & \end{tikzcd}$$

so that $f = gh$.

Proof. (SC1) Given objects $f, g \in c/\mathsf{C}$ we have from construction that

$$\mathrm{Mor}_{c/\mathsf{C}}(f, g) = \mathrm{Mor}_{\mathsf{C}}(\mathrm{cod}\, f, \mathrm{cod}\, g),$$

hence we are ensured of the existence of such morphisms between objects. Given an object $f\colon c \to x$, the morphism $\mathrm{id}_x \in \mathrm{Mor}(\mathsf{C})$ is such that $f = \mathrm{id}_x f$, so that id_x is the identity morphism for f. Let $f, g, u \in c/\mathsf{C}$ be objects such that $f\colon c \to x$, $g\colon c \to y$, $u\colon c \to z$, then there exists morphisms $h \in \mathrm{Mor}_{c/\mathsf{C}}(f, u)$ and $\ell\colon \mathrm{Mor}_{c/\mathsf{C}}(u, g)$ so that the following diagram commutes

$$\begin{tikzcd} & c \arrow[dl, "f"'] \arrow[d, "u"] \arrow[dr, "g"] & \\ x \arrow[r, "h"] & z \arrow[r, "\ell"] & y \end{tikzcd}$$

so that we have a uniquely defined morphism $\ell h \in \mathrm{Mor}_{\mathsf{C}}(x, y)$, where $\ell h \in \mathrm{Mor}_{c/\mathsf{C}}(f, g)$ and thus this defines a map $\mathrm{Mor}_{c/\mathsf{C}}(f, u) \times \mathrm{Mor}_{c/\mathsf{C}}(u, g) \to \mathrm{Mor}_{c/\mathsf{C}}(f, g)$. Since $\mathrm{id}_x, \mathrm{id}_z \in \mathrm{Mor}(c/\mathsf{C})$ for any $x, y \in \mathsf{C}$, then given $f, u \in c/\mathsf{C}$ just as above, the morphism $h\colon x \to z$ is such that $\mathrm{id}_z h = h = h\,\mathrm{id}_x$. In addition to the objects and morphisms named above, define $v\colon c \to w$ and the corresponding morphism $t \in \mathrm{Mor}_{c/\mathsf{C}}(g, v)$ so that $t\colon y \to w$ and $v = tg$. From the fact that C is a category, we find that $(t\ell)h = t(\ell h)$ and hence the same is true for c/C.

(SC2) Let $f, g \in \mathsf{C}/c$, then from definition we have $\mathrm{Mor}_{\mathsf{C}/c}(f, g) = \mathrm{Mor}_{\mathsf{C}}(\mathrm{dom}\, f, \mathrm{dom}\, g)$, which is well defined on C. Given an object $f\colon x \to c$, there exists $\mathrm{id}_x\colon x \to x$ so that $f = f\,\mathrm{id}_x$ and hence id_x is the identity morphism for f. Define objects $f, g, u \in \mathsf{C}/c$ such that $f\colon x \to c$, $g\colon y \to c$, $u\colon c \to z$, and morphisms $h \in \mathrm{Mor}_{\mathsf{C}/c}(u, f)$ and $\ell \in \mathrm{Mor}_{\mathsf{C}/c}(g, u)$. Then, the following diagram commutes

$$\begin{tikzcd} x & z \arrow[l, "h"'] & y \arrow[l, "\ell"'] \\ & c \arrow[ul, "f"] \arrow[u, "u"] \arrow[ur, "g"'] & \end{tikzcd}$$

hence the morphism $h\ell \in \mathrm{Mor}_{\mathsf{C}}(y, x)$ defines a morphism $h\ell \in \mathrm{Mor}_{\mathsf{C}/c}(g, f)$ so that we can construct a map well defined map $\mathrm{Mor}_{\mathsf{C}/c}(g, u) \times \mathrm{Mor}_{\mathsf{C}/c}(u, f) \to \mathrm{Mor}_{\mathsf{C}/c}(g, f)$. Moreover, given f and u as above, we have that there exists $\mathrm{id}_x, \mathrm{id}_z \in \mathrm{Mor}(\mathsf{C}/c)$ so that from the category C it follows that $\mathrm{id}_x h = h\,\mathrm{id}_z$. In addition to the above, define $v\colon w \to c$ and the morphism $t \in \mathrm{Mor}_{\mathsf{C}/c}(v, g)$ so that $t\colon w \to y$. Then we have $g = tv$ and since C is a category, we find that $(h\ell)t = h(\ell t)$. ♮

Initial & Final Objects

Definition 1.2.27 (Initial, final and zero objects)**.** Given a category C, we define the following objects:

(a) An object $I \in \mathsf{C}$ is said to be an *initial object* in the category C if, for all $A \in \mathsf{C}$, there exists exactly one morphism $f \in \mathrm{Mor}(I, A)$ so that $\mathrm{Mor}(I, A) = f$.

(b) An object $F \in \mathsf{C}$ is said to be a *final object* in C if there is exactly one morphism $g \in \mathrm{Mor}(A, F)$ for all given $A \in \mathsf{C}$ and thus $\mathrm{Mor}(A, F) = g$.

(c) An object $O \in \mathsf{C}$ is said to be a *zero* object if it is both initial and terminal.

Proposition 1.2.28. Let a category C, then initial and final objects are said to be unique up to a unique isomorphism:

I. If $I, I' \in \mathsf{C}$ are initial objects of the category, then $I \simeq I'$, where the isomorphism $\varphi_I \colon I \xrightarrow{\simeq} I'$ is unique.

II. If $F, F' \in \mathsf{C}$ are initial objects of the category, then $F \simeq F'$, where the isomorphism $\varphi_F \colon F \xrightarrow{\simeq} F'$ is unique.

Proof. Since I, I' are both initial objects of C then exists a unique morphism $f \in \mathrm{Mor}(I, I')$ and there exists a unique $g \in \mathrm{Mor}(I', I)$ from which we can compose and conclude that $fg = \mathrm{id}_{I'}$ and $gf = \mathrm{id}_I$ and thus f and g are isomorphisms and also $f^{-1} = g$ and $g^{-1} = f$. For final objects the same reasoning works and thus the proof will be omitted. ♮

We normally say that a given construction *satisfies a universal property* if such construction is a terminal object of the category.

1.3 Duality

Definition 1.3.1 (Opposite category)**.** Let C be a category. We define the opposite category of C as C^{op}, such that

(COP1) $\mathrm{Obj}(\mathsf{C}^{\mathrm{op}}) = \mathrm{Obj}(\mathsf{C})$.

(COP2) Given a morphism $f \colon x \to y$ in C, there exists a corresponding morphism $f^{\mathrm{op}} \colon y \to x$ in the C^{op}. That is $\mathrm{dom}\, f = \mathrm{cod}\, f^{\mathrm{op}}$ and $\mathrm{cod}\, f = \mathrm{dom}\, f^{\mathrm{op}}$, these form all of the morphisms in the category C^{op}.

(COP3) For all $A \in \mathsf{C}^{\mathrm{op}}$, there exists an identity morphism $\mathrm{id}_A^{\mathrm{op}} \in \mathrm{End}_{\mathsf{C}^{\mathrm{op}}}(A)$.

(COP4) A pair of morphisms $f^{\mathrm{op}}, g^{\mathrm{op}} \in \mathsf{C}^{\mathrm{op}}$ is composable, so that $\mathrm{dom}\, g^{\mathrm{op}} = \mathrm{cod}\, f^{\mathrm{op}}$, if and only if the pair $g, f \in \mathsf{C}$ is composable, that is $\mathrm{dom}\, f = \mathrm{cod}\, g$. Moreover, we define their composition as $g^{\mathrm{op}} f^{\mathrm{op}} = (fg)^{\mathrm{op}}$.

Lemma 1.3.2. The following propositions are equivalent

(a) $f \colon x \xrightarrow{\simeq} y$ is an isomorphism in C.

(b) For all $c \in \mathsf{C}$, the map

$$f_* \colon \mathrm{Mor}_{\mathsf{C}}(c, x) \to \mathrm{Mor}_{\mathsf{C}}(c, y), \qquad g \overset{f_*}{\longmapsto} fg$$

is a bijection.

(c) For all $c \in \mathsf{C}$, the map

$$f^*\colon \mathrm{Mor}_{\mathsf{C}}(x,c) \to \mathrm{Mor}_{\mathsf{C}}(y,c), \qquad g \overset{f^*}{\longmapsto} gf$$

is a bijection.

Proof. ((a) $\Rightarrow$ (b)) Let $f\colon x \xrightarrow{\simeq} y$ be an isomorphism and $c \in \mathsf{C}$ be an object and define $\ell\colon y \xrightarrow{\simeq} x$ to be its inverse. Given $c \in \mathsf{C}$, define $\ell_*\colon \mathrm{Mor}_{\mathsf{C}}(c,y) \to \mathrm{Mor}_{\mathsf{C}}(c,x)$. Notice that $f_*\ell_*\colon \mathrm{Mor}_{\mathsf{C}}(c,y) \to \mathrm{Mor}_{\mathsf{C}}(c,y)$ mapping $g \mapsto f_*(\ell_*(g)) = f\ell g = g$, hence $f_*\ell_* = \mathrm{id}_{\mathrm{Mor}_{\mathsf{C}}(c,y)}$. Moreover, we have $\ell_* f_*\colon \mathrm{Mor}_{\mathsf{C}}(c,x) \to \mathrm{Mor}_{\mathsf{C}}(c,x)$ mapping $h \mapsto \ell_*(f_*(h)) = hf\ell = h$, hence $\ell_* f_* = \mathrm{id}_{\mathrm{Mor}_{\mathsf{C}}(c,x)}$. This shows that ℓ_* is the inverse of f_* and therefore f_* is an isomorphism. ((b) $\Rightarrow$ (a)) Suppose the contrary, so that f_* is an isomorphism. In particular, we can take $c = y$ so that $f_*\colon \mathrm{Mor}_{\mathsf{C}}(y,x) \to \mathrm{Mor}_{\mathsf{C}}(y,y)$. From the isomorphism property, there exists $\ell \in \mathrm{Mor}_{\mathsf{C}}(y,x)$ such that $f_*(\ell) = f\ell = \mathrm{id}_y$. Consider now that $c = x$, then $f_*\colon \mathrm{Mor}_{\mathsf{C}}(x,x) \to \mathrm{Mor}_{\mathsf{C}}(x,y)$. Notice that $f_*(\ell f) = f\ell f = \mathrm{id}_y f = f$ and $f_*(\mathrm{id}_x) = f\,\mathrm{id}_x = f$. Since f_* is supposed to be an isomorphism, it follows that $\mathrm{id}_x = \ell f$. With this we conclude that ℓ is the inverse of f and hence f is an isomorphism.

((a) $\Leftrightarrow$ (c)) Suppose that $f^{\mathrm{op}}\colon y \to x \in \mathrm{Mor}(\mathsf{C}^{\mathrm{op}})$, then from the last paragraph we have that f^{op} is an isomorphism if and only if $(f^{\mathrm{op}})_*\colon \mathrm{Mor}_{\mathsf{C}^{\mathrm{op}}}(c,y) \to \mathrm{Mor}_{\mathsf{C}^{\mathrm{op}}}(c,x)$ is an isomorphism. Therefore the dual of such statement is that $f\colon x \to y \in \mathrm{Mor}(\mathsf{C})$ is an isomorphism if and only if $(f^{\mathrm{op}})_*^{\mathrm{op}} = f^*\colon \mathrm{Mor}_{\mathsf{C}}(x,c) \to \mathrm{Mor}_{\mathsf{C}}(y,c)$ is an isomorphism, since $\mathrm{Mor}_{\mathsf{C}^{\mathrm{op}}}(*,*') = \mathrm{Mor}_{\mathsf{C}}(*',*)$. ♮

Definition 1.3.3 (Monomorphism)**.** Let a category C. We say that $f \in \mathrm{Mor}_{\mathsf{C}}(x,y)$ is a *monomorphism* if for all $c \in \mathsf{C}$, and for all $\alpha, \beta \in \mathrm{Hom}(c,x)$ we have that $f\alpha = f\beta$ implies $\alpha = \beta$. Equivalently, for all $c \in \mathsf{C}$ the map $f_*\colon \mathrm{Mor}_{\mathsf{C}}(c,x) \to \mathrm{Mor}_{\mathsf{C}}(c,y)$ is injective.

Definition 1.3.4 (Epimorphism)**.** Let a category C. We say that morphism $g \in \mathrm{Hom}_{\mathsf{C}}(x,y)$ to be an *epimorphism* if for all $c \in \mathsf{C}$, and for all $\gamma, \delta \in \mathrm{Hom}_{\mathsf{C}}(y,c)$ we have that $\gamma g = \delta g$ implies $\gamma = \delta$. Equivalently, for all $c \in \mathsf{C}$ the map $g^*\colon \mathrm{Mor}_{\mathsf{C}}(y,c) \to \mathrm{Mor}_{\mathsf{C}}(x,c)$ is injective.

Proposition 1.3.5. Let C be a category. The following are properties regarding monomorphisms and epimorphisms in C:

(a) Every identity morphism is a monomorphism and an epimorphism

(b) The composite of two monomorphisms (or epimorphisms) is a monomorphism (or epimorphism).

(c) If the composition kf is a monomorphism, then f is a monomorphism. Conversely, if fk is an epimorphism, then f is an epimorphism.

Proof. We prove the assertions about monomorphisms, the respective ones for epimorphisms are dually true from the former.

(a) Identities are isomorphism, so clearly they are monomorphisms and epimorphisms.

(b) Let f and g be composable monomorphisms, then if $gfp = gfq$ for two given morphisms p and q, for since g is a monomorphism then $fp = fq$ — then using the fact that f is a monomorphism we obtain $p = q$.

(c) Suppose kf is a monomorphism and consider morphisms g and h such that $fg = fh$. Composing with k one sees that $kfg = kfh$ but since kf is a monomorphism, it follows that $g = h$.

$\natural$

Definition 1.3.6. Let $x \xrightarrow{s} y \xrightarrow{r} x$ be morphisms such that $rs = \mathrm{id}_x$. We define the following terms

(a) s is said to be a section of r. The morphism s is always a monomorphism, being called a split monomorphism.

(b) r is said to be the retraction of s. The morphism r is always an epimorphism, being called a split epimorphism.

(c) x is the retract of y.

Proposition 1.3.7. A morphism $f \in \mathrm{Mor}_{\mathsf{C}}(x, y)$ is a split epimorphism if and only if for all $c \in \mathsf{C}$ the map $f_*\colon \mathrm{Mor}_{\mathsf{C}}(c, x) \to \mathrm{Mor}_{\mathsf{C}}(c, y)$ is surjective. Dually, f is a split monomorphism if and only if for all $c \in \mathsf{C}$ the map $f^*\colon \mathrm{Mor}_{\mathsf{C}}(x, c) \to \mathrm{Mor}_{\mathsf{C}}(y, c)$ is surjective.

Proof. ($\Rightarrow$) Suppose $f\colon x \to y$ is a split epimorphism and define $g\colon y \to x$ as a section of f, that is $fg = \mathrm{id}_y$. Let $c \in \mathsf{C}$, and $\alpha \in \mathrm{Mor}_{\mathsf{C}}(c, y)$ be any morphism. Notice that $g\alpha \in \mathrm{Mor}_{\mathsf{C}}(c, x)$, hence we find that $f_*(g\alpha) = fg\alpha = \alpha$ and therefore f_* is surjective. ($\Leftarrow$) Suppose f_* is surjective, then in particular for $c = y$ we have that $\mathrm{id}_y \in \operatorname{im} f_*$ and hence f is a split epimorphism. (Dual) Let $f^{\mathrm{op}}\colon y \to x$, then from the above proposition f^{op} is a split epimorphism if and only if $(f^{\mathrm{op}})_*\colon \mathrm{Mor}_{\mathsf{C}^{\mathrm{op}}}(y, c) \to \mathrm{Mor}_{\mathsf{C}^{\mathrm{op}}}(x, c)$ is surjective. Dually we have that f is a monomorphism if and only if $(f^{\mathrm{op}})_*^{\mathrm{op}} = f^*\colon \mathrm{Mor}_{\mathsf{C}}(c, x) \to \mathrm{Mor}_{\mathsf{C}}(c, y)$, which proves the last part. $\natural$

Proposition 1.3.8. Let $f \in \mathrm{Mor}(\mathsf{C})$. If f is a monomorphism and also a split epimorphism, then f is an isomorphism. Dually, if f is an epimorphism and a split monomorphism, then f is an isomorphism.

Proof. Suppose $f\colon x \to y$ is a split epimorphism and $g\colon y \to x$ be a section of f, then $fg = \mathrm{id}_y$ and, moreover $fgf = \mathrm{id}_y f = f$ hence from if f is a monomorphism we conclude that $gf = \mathrm{id}_x$. Thus g is the inverse of f and hence f is an isomorphism. Dually, let $f^{\mathrm{op}}\colon y \to x$ be a monomorphism and split epimorphism, then f^{op} is an isomorphism, which dually means that f is an epimorphism and split monomorphism, then f is an isomorphism. $\natural$

Lemma 1.3.9. Let $f\colon x \to y$ and $g\colon y \to z$. The following propositions hold

(a). If f and g are monomorphisms, so is $gf\colon x \rightarrowtail z$. For the dual proposition, if f and g are epimorphisms, then $gf\colon x \twoheadrightarrow z$ is an epimorphism.

(b). If $gf\colon x \rightarrowtail z$ is a monomorphism, then f is also monomorphism. Dually, if $gf\colon x \twoheadrightarrow z$ is an epimorphism, then g is an epimorphism.

Proof. (a) Suppose that f, g are both monomorphisms. Given any $c \in \mathtt{C}$, let $\alpha, \beta \in \mathrm{Mor}_{\mathtt{C}}(c, x)$ be such that $gf\alpha = gf\beta$. In particular, since g is monic, then $f\alpha = f\beta$, but f is also monic, hence $\alpha = \beta$. For the dual part, suppose that $f^{\mathrm{op}}\colon y \to x$ and $g^{\mathrm{op}}\colon z \to y$ be monic, then from above we have $f^{\mathrm{op}} g^{\mathrm{op}} = (gf)^{\mathrm{op}}$ monic, which dually implies that gf is epic.

(b) Let gf be monic. Then given $c \in \mathtt{C}$ and $\alpha, \beta \in \mathrm{Mor}_{\mathtt{C}}(c, x)$ such that $f\alpha = f\beta$, then in particular $g(f\alpha) = g(f\beta)$, then from the monic property we have $\alpha = \beta$. Dually, suppose that $(gf)^{\mathrm{op}} = f^{\mathrm{op}} g^{\mathrm{op}} \in \mathrm{Mor}(\mathtt{C}^{\mathrm{op}})$ is monic, then from above we find that g^{op} is monic, which dually implies that g is epic. ♮

1.4 Functors

Definition 1.4.1 (Covariant functor)**.** Let $\mathtt{C}$ and $\mathtt{D}$ be categories. A covariant functor $F\colon \mathtt{C} \to \mathtt{D}$ has the following data:

(DF1) For all $c \in \mathtt{C}$ exists a corresponding $Fc \in \mathtt{D}$[3].

(DF2) For all $f\colon c \to c' \in \mathrm{Mor}(\mathtt{C})$ there exists a morphism $Ff\colon Fc \to Fc' \in \mathrm{Mor}(\mathtt{D})$.

Such data satisfies the two following axioms:

(AF1) For all composable $f, g \in \mathrm{Mor}(\mathtt{C})$, we have $Fg \circ Ff = F(g \circ f)$.

(AF2) For all $c \in \mathtt{C}$ we have $F\,\mathrm{id}_c = \mathrm{id}_{Fc} \in \mathrm{Mor}(\mathtt{D})$.

Definition 1.4.2 (Contravariant functor)**.** A contravariant functor from categories $\mathtt{C}$ to $\mathtt{D}$ is a functor $F\colon \mathtt{C}^{\mathrm{op}} \to \mathtt{D}$ together with the following data:

(DCF1) For all $c \in \mathtt{C}$ exists $Fc \in \mathtt{D}$.

(DCF2) For all $f\colon c \to c' \in \mathrm{Mor}(\mathtt{C})$ we have $Ff\colon Fc' \to Fc \in \mathrm{Mor}(\mathtt{D})$.

Moreover, a contravariant functor satisfies the following axioms:

(ACF1) For all composable $f, g \in \mathrm{Mor}(\mathtt{C})$ we have $Ff \circ Fg = F(g \circ f)$.

(ACF2) For all $c \in \mathtt{C}$ we have $F\,\mathrm{id}_c = \mathrm{id}_{Fc}$.

[3]When convenient, we may discard the use of parenthesis, but in occasions where the use of parenthesis brings more clarity to the situation, we shall use it.

This can all be comprised diagrammatically as:

$$
\begin{tikzcd}
\mathtt{C}^{\mathrm{op}} \arrow[r, "F"] & \mathtt{D}
\end{tikzcd}
$$

$$
\begin{tikzcd}
c \arrow[r, mapsto] \arrow[d, "f"] \arrow[dd, bend right=60, "gf"'] & Fc \\
c' \arrow[r, mapsto] \arrow[d, "g"] & Fc' \arrow[u, "Ff"'] \\
c'' \arrow[r, mapsto] & Fc'' \arrow[u, "Fg"'] \arrow[uu, bend right=60, "Ff\circ Fg=F(gf)"']
\end{tikzcd}
$$

Definition 1.4.3 (Composition of functors)**.** Let $F\colon \mathtt{C} \to \mathtt{D}$ and $G\colon \mathtt{D} \to \mathtt{B}$ be functors. We define their composition $GF\colon \mathtt{C} \to \mathtt{B}$ by $(GF)(x) = G(F(x))$, for all $x \in \mathtt{C}$, and $(GF)(f) = G(F(f))$ for all morphism $f \in \mathrm{Mor}(\mathtt{C})$.

Example 1.4.4. An important example of contravariant functor is $\mathrm{op}\colon \mathtt{C} \to \mathtt{C}^{\mathrm{op}}$, where $\mathtt{C}$ is some category, defined by the identity on objects and morphisms.

Definition 1.4.5 (Forgetful functor)**.** A functor is said to be forgetful if the functor "forgets" some object, structure or property of its domain category.

Example 1.4.6. We have some classical examples of forgetful functors, for instance, the following are functors that forget the structure of their domain categories:

- The functor $G\colon \mathtt{Grp} \to \mathtt{Set}$ mapping groups to its corresponding underlying set.

- The functor $T\colon \mathtt{Top} \to \mathtt{Set}$ maps any topological space to its corresponding set of points.

- The functor $V, E\colon \mathtt{Graph} \to \mathtt{Set}$ maps the vertices and edges of a graph to the set of such vertices an edges.

Example 1.4.7 ($\mathtt{Top}^{\mathrm{op}} \to \mathtt{Ring}$)**.** Let $C\colon \mathtt{Top}^{\mathrm{op}} \to \mathtt{Ring}$ be a contravariant functor such that for all $X \in \mathtt{Top}$, let CX be the ring of continuous functions $X \to \mathbf{R}$. The ring operations on CX are defined pointwise, that is, given $p, q\colon X \to \mathbf{R} \in CX$ we have $(p \cdot q)(x) = p(x) \cdot q(x)$ and $(p + q)(x) = p(x) + q(x)$ for all $x \in X$. Moreover, given a morphism $f\colon X \to Y \in \mathrm{Mor}(\mathtt{Top})$ we define $Cf\colon CY \to CX$ as the composition $(Cf)(q) = qf \in \mathrm{Mor}(CX)$ for all $q \in CY$, that is

$$
\begin{tikzcd}
X \arrow[r, "f"] \arrow[rr, bend right, "Cf(q)=qf"'] & Y \arrow[r, "q"] & \mathbf{R}
\end{tikzcd}
$$

We now show that the axioms for the contravariant functor are satisfied by C. Let $f\colon X \to Y$ and $g\colon Y \to Z$, then given any $p \in CZ$ we have

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[rrr, bend left=50, "Cf(Cg(p))"] \arrow[rrr, bend right=30, "C(gf)(p)"'] & Y \arrow[r, "g"] \arrow[rr, bend left, "Cg(p)"] & Z \arrow[r, "p"] & \mathbf{R}
\end{tikzcd}$$

hence $C(f)C(g) = C(gf)$. Moreover, given any $X \in \mathtt{Top}^{\mathrm{op}}$ we find that $C\,\mathrm{id}_X\colon CX \to CX$ is such that for all $q \in CX$, $C\,\mathrm{id}_X(q) = q\,\mathrm{id}_X = q$ hence $C\,\mathrm{id}_X = \mathrm{id}_{CX}$. This finishes the proof that $C\colon \mathtt{Top}^{\mathrm{op}} \to \mathtt{Ring}$ is a contravariant functor.

Definition 1.4.8 (Presheaf)**.** Let $\mathtt{C}$ be a $\mathcal{U}$-small category. A contravariant functor $\mathtt{C}^{\mathrm{op}} \to \mathtt{Set}$ is called a presheaf on $\mathtt{C}$.

Example 1.4.9 ($\mathcal{O}(X)^{\mathrm{op}} \to \mathtt{Set}$)**.** Let $X \in \mathtt{Top}$ we define $\mathcal{O}(X)$ to be the poset category whose objects are open sets of X. That is, for sets $U, U' \in \mathcal{O}(X)$, if $U \subseteq U'$, then there exists a morphism $U \to U'$ in $\mathrm{Mor}(\mathcal{O}(X))$. A presheaf on the category $\mathcal{O}(X)$ is a functor $F\colon \mathcal{O}(X)^{\mathrm{op}} \to \mathtt{Set}$ that assigns $FU = \{f\colon U \to \mathbf{R} : f \text{ continuous}\}$ for all $U \in \mathcal{O}(X)$. Moreover, for maps $g\colon U \to U'$ (that is $U \subseteq U'$) we have $Fg\colon FU' \to FU$ such that $Fg(f) = f|_U\colon U \to \mathbf{R}$ for all $f\colon U' \to \mathbf{R}$ continuous. Since the restriction of a continuous map is continuous, then $f|_U \in FU$

Example 1.4.10 (Simplex category)**.** The simplex category Δ comprises objects that are finite non-empty ordinals and order-preserving morphisms. Simplicial sets are defined as presheaves $\Delta^{\mathrm{op}} \to \mathtt{Set}$.

Lemma 1.4.11. Functors preserve isomorphisms. Let $\mathtt{C}$ and $\mathtt{D}$ be categories and $F\colon \mathtt{C} \to \mathtt{D}$ be a functor. Given an isomorphism $f\colon c \xrightarrow{\simeq} c' \in \mathrm{Mor}(\mathtt{C})$, we have that $Ff\colon Fc \xrightarrow{\simeq} c'$ is an isomorphism.

Proof. Denote by $f^{-1}\colon c' \xrightarrow{\simeq} c$ the inverse of f. By the composition axiom we have

$$\begin{aligned} F(f^{-1})F(f) &= F(f^{-1}f) = F\,\mathrm{id}_c = \mathrm{id}_{Fc}, \\ F(f)F(f^{-1}) &= F(ff^{-1}) = F\,\mathrm{id}_{c'} = \mathrm{id}_{Fc'}\,. \end{aligned}$$

This shows that Ff^{-1} is the right and left inverse of Ff, hence $Ff\colon Fc \xrightarrow{\simeq} Fc'$ is indeed an isomorphism. ♮

Example 1.4.12 (Group action)**.** Let G be any group and consider the category $\mathtt{B}G$ generated by G — that is, $\mathtt{B}G$ consists of a unique object $*$ and the morphisms of the category are automorphisms $* \to *$ given by the elements of G. Given a category $\mathtt{C}$, a functor $X\colon \mathtt{B}G \to \mathtt{C}$ — given by mapping $X* := X \in \mathtt{C}$ and each object $g \in G$ to an endomorphism $Xg := g_*\colon X \to X$ — is said to define a left group action on the object $X \in \mathtt{C}$. Moreover, the functor X has to obey

- Composition preserving: for any $h, g \in G$, we have $(hg)_* = h_* g_*$.
- Identity: $e_* = \mathrm{id}_X$.

Since functors preserve isomorphisms and every morphism of $\mathtt{B}G$ is an automorphism, it follows that, for all $g \in G$, the map $g_*: X \to X$ is an automorphism — in particular, this implies that $(g^{-1})_* = g_*^{-1}$.

Some particular cases of interest are the following:

- If $\mathtt{C} = \mathtt{Set}$, then the set X together with the actions $\{g_* : g \in G\}$ is called a G-set.
- If $\mathtt{C} = \mathtt{Vect}_k$, then the k-vector space X together with the actions generated by G is said to be a G-representation.
- If $\mathtt{C} = \mathtt{Top}$, then the topological space X endowed with the actions generated by G is called a G-space.

A right group action is nothing more than a contravariant functor $X: \mathtt{B}G^{\mathrm{op}} \to \mathtt{C}$ such that $X* := X$ and $Xg := g^*: X \to X$ are endomorphisms. The rules for such a functor are the contravariant preservation of compositions, that is, $(hg)_* = g_* h_*$, and that $e^* = \mathrm{id}_X$ as before.

Example 1.4.13 (Skeletal functor)**.** Let $\mathtt{C}$ be a category. The inclusion functor $J: \mathrm{sk}\,\mathtt{C} \to \mathtt{C}$ equivalence of categories, indeed, one can define a quasi-inverse functor $F: \mathtt{C} \to \mathrm{sk}\,\mathtt{C}$ by mapping $a \in \mathtt{C}$ to the unique object $Fa \in \mathrm{sk}\,\mathtt{C}$ for which there exists an isomorphism $a \simeq Fa$. Choosing a collection $(\alpha_a : a \xrightarrow{\simeq} Fa)_{a \in \mathtt{C}}$ of isomorphisms where $\alpha_a = \mathrm{id}_a$ whenever $Fa = a$, we can map each morphism $f: a \to b$ of $\mathtt{C}$ to the morphism $Ff := \alpha_b f \alpha_a^{-1}: Fa \to Fb$. From this construction one gets $FJ = \mathrm{id}$ and the collection $(\alpha_a)_{a \in \mathtt{A}}$ of chosen isomorphisms define a natural isomorphism $\alpha: \mathrm{id} \overset{\simeq}{\Longrightarrow} JF$.

Lemma 1.4.14. Functors preserve split monomorphisms and split epimorphisms.

Proof. Let $\mathtt{C}$ and $\mathtt{D}$ be categories and consider a functor $F: \mathtt{C} \to \mathtt{D}$. Define morphisms $x \xrightarrow{s} y \xrightarrow{r} x$ in $\mathrm{Mor}(\mathtt{C})$ such that $rs = \mathrm{id}_x$, that is, s is a split monomorphism and r is a split epimorphism. Consider the morphisms $Fs: Fx \to Fy$ and $Fr: Fy \to Fx$ in $\mathrm{Mor}(\mathtt{D})$. Notice that $F(s)F(r) = F(sr) = F(\mathrm{id}_x) = \mathrm{id}_{Fx}$. Hence Fs is a split monomorphism and Fr is a split epimorphism. ♮

Definition 1.4.15 (Mor functors)**.** Let $\mathtt{C}$ be a $\mathcal{U}$-category. Given any $c \in \mathtt{C}$, there exists a pair of covariant and contravariant functors, $\mathrm{Mor}(c, -)$ and $\mathrm{Mor}(-, c)$, respectively — represented by the object c. That is:

$$
\mathtt{C} \xrightarrow{\mathrm{Mor}(c,-)} \mathtt{Set} \qquad \mathtt{C}^{\mathrm{op}} \xrightarrow{\mathrm{Mor}(-,c)} \mathtt{Set}
$$

$$
\begin{tikzcd}
x \arrow[r, mapsto] \arrow[d, "f"'] & \mathrm{Mor}(c,x) \arrow[d, "f_*"] \\
y \arrow[r, mapsto] & \mathrm{Mor}(c,y)
\end{tikzcd}
\qquad
\begin{tikzcd}
x \arrow[r, mapsto] \arrow[d, "f"'] & \mathrm{Mor}(x,c) \\
y \arrow[r, mapsto] & \mathrm{Mor}(y,c) \arrow[u, "f^*"']
\end{tikzcd}
$$

We now prove that such definition indeed satisfies the axioms for covariant and contravariant functors. Given morphisms $f\colon x \to y$ and $g\colon y \to z$ in $\mathrm{Mor}(\mathtt{C})$, we see that $g_* f_* = (gf)_*$, moreover $f^* g^* = (fg)^*$. Let $x \in \mathtt{C}$ be any object, then $\mathrm{id}_{x*} = \mathrm{id}_{\mathrm{Mor}(c,x)} = \mathrm{id}_x^*$. This proves that $\mathrm{Mor}(c, -)$ is covariant and $\mathrm{Mor}(-, c)$ is contravariant.

Definition 1.4.16 (Faithful, full & its friends)**.** Let $\mathtt{C}$ and $\mathtt{D}$ be categories. A functor $F\colon \mathtt{C} \to \mathtt{D}$ is said to be

(a) *Faithful* if for all $x, y \in \mathtt{C}$ the map $\mathrm{Mor}_{\mathtt{C}}(x, y) \rightarrowtail \mathrm{Mor}_{\mathtt{D}}(Fx, Fy)$ is injective.

(b) *Full* if for all $x, y \in \mathtt{C}$ the map $\mathrm{Mor}_{\mathtt{C}}(x, y) \twoheadrightarrow \mathrm{Mor}_{\mathtt{D}}(Fx, Fy)$ is surjective.

(c) *Fully faithful* if for all $x, y \in \mathtt{C}$ the map $\mathrm{Mor}_{\mathtt{C}}(x, y) \xrightarrow{\simeq} \mathrm{Mor}_{\mathtt{D}}(Fx, Fy)$ is a bijection.

(d) *Essentially surjective* if for each $y \in \mathtt{D}$ there exists $x \in \mathtt{C}$ and an isomorphism $Fx \xrightarrow{\simeq} y$ in $\mathtt{D}$.

(e) *Conservative* if, given a morphism f in $\mathtt{C}$, if Ff is an isomorphism in $\mathtt{D}$, then f is an isomorphism in $\mathtt{C}$.

Proposition 1.4.17 (Fully faithful functors and isomorphisms)**.** Let $F\colon \mathtt{C} \to \mathtt{D}$ be a *fully faithful* functor. If $Fx \simeq Fy$ in $\mathtt{D}$, for some $x, y \in \mathtt{C}$, then $x \simeq y$ in $\mathtt{C}$ for a *unique* isomorphism.

Proof. If $Fx \simeq Fy$, let $\phi\colon Fx \xrightarrow{\simeq} Fy$ be an isomorphism in $\mathtt{D}$. Since F is fully faithful, there exists unique morphisms $f\colon x \to y$ and $g\colon y \to x$ in $\mathtt{C}$ for which $Ff = \phi$ and $Fg = \phi^{-1}$. In particular, one has

$$\mathrm{id}_{Fx} = \phi^{-1}\phi = Fg \circ Ff = F(gf).$$

From the faithfulness of F, since $F\,\mathrm{id}_x = \mathrm{id}_{Fx}$, then $gf = \mathrm{id}_x$. On the other hand, one also has

$$\mathrm{id}_{Fy} = \phi\phi^{-1} = Ff \circ Fg = F(fg),$$

therefore, since $F\,\mathrm{id}_y = \mathrm{id}_{Fy}$ we obtain $fg = \mathrm{id}_y$. We can now finally conclude that f is an isomorphism and its inverse is g. ♮

Definition 1.4.18 (Product & disjoint union categories)**.** Let I be an indexing set and $\{\mathtt{C}_i\}_{i \in I}$ be a collection of categories associated with I. We define the following categories:

(a) The *product* category $\prod_{i \in I} \mathtt{C}_i$ consists of objects $\mathrm{Obj}(\prod_{i \in I} \mathtt{C}_i) := \prod_{i \in I} \mathrm{Obj}(\mathtt{C}_i)$, and morphisms $\mathrm{Mor}_{\prod_{i \in I} \mathtt{C}_i}((x_i)_{i \in I}, (y_i)_{i \in I}) := \prod_{i \in I} \mathrm{Mor}_{\mathtt{C}_i}(x_i, y_i)$ between any two objects $(x_i)_{i \in I}$ and $(y_i)_{i \in I}$ in the category. Composable morphisms $(f_i)_{i \in I}$ and $(g_i)_{i \in I}$ have composition defined component-wise — that is, $(f_i)_{i \in I}(g_i)_{i \in I} := (f_i g_i)_{i \in I}$.

(b) The *disjoint union* category $\coprod_{i \in I} \mathtt{C}_i$ consists of objects

$$\mathrm{Obj}\left(\coprod_{i \in I} \mathtt{C}_i\right) := \coprod_{i \in I} \{(x, i) : i \in I \text{ and } x \in \mathtt{C}_i\},$$

and morphisms

$$\mathrm{Mor}_{\coprod_{i\in I}\mathtt{C}_i}((x,i),(y,j)) := \begin{cases}\mathrm{Mor}_{\mathtt{C}_i}(x,y), & \text{if } i=j\\ \varnothing, & \text{otherwise}\end{cases}$$

for any objects (x,i) and (y,j) in the category.

Moreover, if $\{\mathtt{D}_i\}_{i\in I}$ is another collection of categories, and $\{F_i\colon \mathtt{C}_i \to \mathtt{D}_i\}_{i\in I}$ is a collection of functors, we associate to each of the above categories the functors $\prod_{i\in I} F_i$ and $\coprod_{i\in I} F_i$.

Example 1.4.19 (Orbit category)**.** Let G be a group. We define the orbit category associated to G as $\mathcal{O}_G$ whose objects are subgroups $H \subseteq G$, identified by the left G-set G/H of left cosets of H. The morphisms $\phi\colon G/H \to G/Q$ are maps commuting with the left G-action, that is, $\phi(g_*h) = g_*\phi(h)$ — such maps are called G-equivariant.

Proposition 1.4.20 (Bifunctor)**.** Let $\mathtt{A}$, $\mathtt{B}$ and $\mathtt{C}$ be any categories. A functor

$$F\colon \mathtt{A}\times\mathtt{B} \longrightarrow \mathtt{C}$$

is called a *bifunctor*. The functor F is defined so that, given any objects $x \in \mathtt{A}$ and $y \in \mathtt{B}$,

$$F(x,-)\colon \mathtt{B} \longrightarrow \mathtt{C} \ \text{ and } \ F(-,y)\colon \mathtt{A} \longrightarrow \mathtt{C}$$

are both functors. Moreover, given any morphisms $f\colon x \to y$ in $\mathtt{A}$ and $g\colon x' \to y'$ in $\mathtt{B}$, the following diagram commutes

$$\begin{tikzcd} F(x,x') \arrow[r, "F(x,g)"] \arrow[d, "F(f,x')"'] & F(x,y') \arrow[d, "F(f,y')"] \\ F(y,x') \arrow[r, "F(y,g)"] & F(y,y') \end{tikzcd}$$

Notice that the product of small categories can also be understood as a bifunctor

$$\times\colon \mathtt{Cat}\times\mathtt{Cat} \longrightarrow \mathtt{Cat}$$

which associates to each pair of small categories $(\mathtt{A},\mathtt{B})$ the product category $\mathtt{A}\times\mathtt{B}$, and any pair of functors $(F\colon \mathtt{A}\to\mathtt{B}, G\colon \mathtt{C}\to\mathtt{D})$ is mapped to a bifunctor

$$F\times G\colon \mathtt{A}\times\mathtt{C} \to \mathtt{B}\times\mathtt{D}$$

such that $(F\times G)(f,g) := (Ff, Gg)$ and $(F\times G)(A,C) := (FA, GC)$, for any pair of morphisms $(f,g) \in \mathrm{Mor}(\mathtt{A})\times\mathrm{Mor}(\mathtt{C})$ and pair of objects $(A,C) \in \mathtt{A}\times\mathtt{C}$.

Definition 1.4.21 (Functor-induced categories)**.** Let $F\colon \mathtt{C} \to \mathtt{D}$ be a functor between categories $\mathtt{C}$ and $\mathtt{D}$, and let $y \in \mathtt{D}$. We define the following two categories:

(a) The category C_y is defined to consist of objects

$$\operatorname{Obj}(\mathsf{C}_y) \coloneqq \{(x,s) : x \in \mathsf{C} \text{ and } s\colon Fx \to y \text{ in } \mathsf{D}\},$$

and morphisms between any objects $(a,s),(b,t) \in \mathsf{C}_y$ are defined to be

$$\operatorname{Mor}_{\mathsf{C}_y}((a,s),(b,t)) \coloneqq \{f \in \operatorname{Mor}_{\mathsf{C}}(a,b) : s = tF(f) \text{ in } \mathsf{D}\},$$

that is, a morphism $f\colon (a,s) \to (b,t)$ makes to following diagram commute

$$\begin{tikzcd} Fa \arrow[r, "s"] \arrow[dr, "Ff"'] & y \\ & Fb \arrow[u, "t"'] \end{tikzcd}$$

Together with such category, we define a faithful functor $j_y\colon \mathsf{C}_y \to \mathsf{C}$ by $j_y(x,s) \coloneqq x$, acting as a projection.

(b) The category C^y is defined to consist of objects

$$\operatorname{Obj}(\mathsf{C}^y) \coloneqq \{(x,s) : x \in \mathsf{C} \text{ and } s\colon y \to Fx \text{ in } \mathsf{D}\},$$

and morphisms between any objects $(a,s),(b,t) \in \mathsf{C}^y$ are defined to be

$$\operatorname{Mor}_{\mathsf{C}^y}((a,s),(b,t)) \coloneqq \{f \in \operatorname{Mor}_{\mathsf{C}}(a,b) : t = F(f)s \text{ in } \mathsf{D}\},$$

that is, a morphism $f\colon (a,s) \to (b,t)$ makes to following diagram commute

$$\begin{tikzcd} y \arrow[r, "s"] \arrow[d, "t"'] & Fa \arrow[dl, "Ff"] \\ Fb & \end{tikzcd}$$

Together with such category, we define a faithful functor $j^y\colon \mathsf{C}^y \to \mathsf{C}$ by $j^y(x,s) \coloneqq x$, which acts as a projection.

Definition 1.4.22 (Equivalence classes)**.** Let C be a category and $\sim$ denote an equivalence relation on the objects of C — where $x \sim y$ if $\operatorname{Mor}_{\mathsf{C}}(x,y) \neq \varnothing$. We denote the collection of all equivalence classes on the objects of C by $\pi_0(\mathsf{C})$.

Corollary 1.4.23. A category C is *connected* if and only if $\pi_0(\mathsf{C})$ consists of a single element.

Proof. If $\pi_0(\mathsf{C})$ is a single object, every equivalence class on C is such that every element is equivalent to each other, which implies that the collection of morphisms $\operatorname{Mor}_{\mathsf{C}}(x,y)$, between any two elements $x,y \in \mathsf{C}$, is non-empty — thus clearly C is connected. ♮

Definition 1.4.24 (Isomorphisms of monomorphisms & epimorphisms)**.** Let C be a category and $x,y,z \in \mathsf{C}$ be any objects. We define the following concepts:

(a) Two monomorphisms $f\colon x \rightarrowtail z$ and $g\colon y \rightarrowtail z$ in C are said to be *isomorphic* in C if there exists an isomorphism $h\colon x \xrightarrow{\simeq} y$ for which the following diagram commutes in

$$\begin{tikzcd}
 & z & \\
x \arrow[ur, tail, "f"] \arrow[rr, "h", "\simeq"'] & & y \arrow[ul, tail, "g"']
\end{tikzcd}$$

This is equivalent to f and g being isomorphic in C_z.

(b) Two epimorphisms $f\colon x \twoheadrightarrow z$ and $g\colon y \twoheadrightarrow z$ in C are said to be *isomorphic* in C if there exists an isomorphism $h\colon x \xrightarrow{\simeq} y$ for which the following diagram commutes in

$$\begin{tikzcd}
x \arrow[rr, "h", "\simeq"'] \arrow[dr, two heads, "f"'] & & y \arrow[dl, two heads, "g"] \\
 & z &
\end{tikzcd}$$

This is equivalent to f and g being isomorphic in C^z.

Definition 1.4.25. Let C be a category and $x \in \mathsf{C}$ be any object. We define the following:

(a) An isomorphism class of a monomorphism with *target* x is called a *subobject* of x.

(b) An isomorphism class of an epimorphism with *source* x is called a *quotient* of x.

These isomorphism classes are given in the sense of Definition 1.4.24.

Example 1.4.26 (Ordering subobject)**.** One can *order* the collection of subobjects of a given object $x \in \mathsf{C}$ by defining a relation $[f\colon y \rightarrowtail x] \leqslant [g\colon z \rightarrowtail x]$ if there exists $h\colon y \to z$ such that

$$\begin{tikzcd}
y \arrow[r, tail, "f"] \arrow[dr, dashed, "h"'] & x \\
 & z \arrow[u, tail, "g"']
\end{tikzcd}$$

Moreover, *if* h exists, then it's unique.

1.5 Natural Transformations

Definition 1.5.1 (Natural transformation)**.** Let C and D be categories, and consider functors $F, G\colon \mathsf{C} \rightrightarrows \mathsf{D}$. A *natural transformation* $\alpha\colon F \Rightarrow G$ consists of morphisms $\alpha_x\colon Fx \to Gx$, for all $x \in \mathsf{C}$, such that, for any morphism $f\colon x \to y$ in C, the following diagram commutes

$$\begin{tikzcd}
Fx \arrow[r, "\alpha_x"] \arrow[d, "Ff"'] & Gx \arrow[d, "Gf"] \\
Fy \arrow[r, "\alpha_y"] & Gy
\end{tikzcd}$$

Moreover, if $L\colon \mathtt{C} \to \mathtt{D}$ is another functor, and $\beta\colon G \Rightarrow L$ is a natural transformation, we define the *composition* of natural transformations α and β as the map $\beta\alpha\colon F \Rightarrow L$ such that $(\beta\alpha)_x := \beta_x\alpha_x$ for every $x \in \mathtt{C}$.

Definition 1.5.2 (Functor category)**.** Let $\mathtt{C}$ and $\mathtt{D}$ be categories. We define a category $\mathrm{Fct}(\mathtt{C}, \mathtt{D})$ whose objects are functors $\mathtt{C} \to \mathtt{D}$, and morphisms are natural transformations between functors.

Proposition 1.5.3 (Horizontal composition)**.** Let $\mathtt{A}, \mathtt{B}$ and $\mathtt{C}$ be categories. Consider functors $F, F'\colon \mathtt{A} \rightrightarrows \mathtt{B}$, and $G, G'\colon \mathtt{B} \rightrightarrows \mathtt{C}$. Let $\alpha\colon F \Rightarrow F'$ and $\beta\colon G \Rightarrow G'$ be natural transformations — that is,

$$\begin{tikzcd} \mathtt{A} \arrow[r, bend left, "F"] \arrow[r, bend right, "F'"'] \arrow[r, phantom, "\Downarrow\alpha"] & \mathtt{B} \arrow[r, bend left, "G"] \arrow[r, bend right, "G'"'] \arrow[r, phantom, "\Downarrow\beta"] & \mathtt{C} \end{tikzcd}$$

There exists an induced natural transformation $\beta * \alpha\colon GF \Rightarrow G'F'$ — called the *horizontal composition*, also called *Godement product*, of α and β — such that, for all $a \in \mathtt{A}$,

$$(\beta * \alpha)_a = \beta_{F'a} G(\alpha_a) = G'(\alpha_a)\beta_{Fa}.$$

The horizontal composition can be depicted by the following diagram

$$\begin{tikzcd} \mathtt{A} \arrow[r, bend left, "GF"] \arrow[r, bend right, "G'F'"'] \arrow[r, phantom, "\Downarrow\beta*\alpha"] & \mathtt{C} \end{tikzcd}$$

Proof. Notice that the naturality of $\beta * \alpha$ solely depends on the naturality of both α and β. Indeed, for every $a \in \mathtt{A}$ and morphism $f\colon a \to a'$ in $\mathtt{A}$ the following diagram

$$\begin{tikzcd} GFa \arrow[r, "G\alpha_a"] \arrow[d, "GFf"'] \arrow[rr, bend left, "(\beta*\alpha)_a"] & GF'a \arrow[r, "\beta_{F'a}"] \arrow[d, "GF'f"] & G'F'a \arrow[d, "G'F'f"] \\ GFa' \arrow[r, "G\alpha_{a'}"'] \arrow[rr, bend right, "(\beta*\alpha)_{a'}"'] & GF'a' \arrow[r, "\beta_{F'a'}"'] & G'F'a' \end{tikzcd}$$

is commutative, which proves that $\beta * \alpha$ is a natural transformation. ♮

Notation 1.5.4. For the ease of notation, we denote the vertical composition $\alpha * \mathrm{id}_F$ by $\alpha * F$.

Definition 1.5.5 (Vertical composition)**.** Let $\mathtt{C}$ and $\mathtt{D}$ be categories, and consider functors and natural transformations given in the following diagram

$$\begin{tikzcd} \mathtt{C} \arrow[r, bend left=60, "F"{name=F}] \arrow[r, "H"{name=H, description}] \arrow[r, bend right=60, "G"'{name=G}] & \mathtt{D} \arrow[Rightarrow, from=F, to=H, "\alpha"] \arrow[Rightarrow, from=H, to=G, "\beta"] \end{tikzcd}$$

We define the *vertical composition* of β with α as the natural transformation $\beta \circ \alpha \colon F \Rightarrow G$ — diagrammatically,

$$\begin{tikzcd} \mathtt{C} \arrow[r, bend left=60, "F"{name=F}] \arrow[r, bend right=60, "G"'{name=G}] & \mathtt{D} \arrow[Rightarrow, from=F, to=G, "\beta\circ\alpha"] \end{tikzcd}$$

Proposition 1.5.6. Consider the following diagram, with categories $\mathtt{A}$, $\mathtt{B}$ and $\mathtt{C}$, functors F, G, H, K, L, M, and natural transformations α, β, γ and δ:

$$\begin{tikzcd} \mathtt{A} \arrow[r, bend left=60, "F"{name=F}] \arrow[r, "H"{name=H, description}] \arrow[r, bend right=60, "L"'{name=L}] & \mathtt{B} \arrow[r, bend left=60, "G"{name=G}] \arrow[r, "K"{name=K, description}] \arrow[r, bend right=60, "M"'{name=M}] & \mathtt{C} \arrow[Rightarrow, from=F, to=H, "\alpha"] \arrow[Rightarrow, from=H, to=L, "\gamma"] \arrow[Rightarrow, from=G, to=K, "\beta"] \arrow[Rightarrow, from=K, to=M, "\delta"] \end{tikzcd}$$

This diagram is such that the following equality holds:

$$(\delta * \gamma) \circ (\beta * \alpha) = (\delta \circ \beta) * (\gamma \circ \alpha)$$

Where by $\circ$ we denote the *vertical composition* and by $*$ we denote the *horizontal composition*.

Proof. The composition $(\delta * \gamma) \circ (\beta * \alpha) \colon GF \Rightarrow ML$ is given by the following diagrams

$$\begin{tikzcd} \mathtt{A} \arrow[r, bend left=60, "GF"{name=GF}] \arrow[r, "HK"{name=HK, description}] \arrow[r, bend right=60, "ML"'{name=ML}] & \mathtt{C} \arrow[Rightarrow, from=GF, to=HK, "\beta*\alpha"] \arrow[Rightarrow, from=HK, to=ML, "\delta*\gamma"] \end{tikzcd} \longmapsto \begin{tikzcd} \mathtt{A} \arrow[r, bend left=60, "GF"{name=GF}] \arrow[r, bend right=60, "ML"'{name=ML}] & \mathtt{C} \arrow[Rightarrow, from=GF, to=ML, "(\delta*\gamma)\circ(\beta*\alpha)" description] \end{tikzcd}$$

On the other hand, the composition $(\delta \circ \beta) * (\gamma \circ \alpha) \colon GF \Rightarrow ML$ is given by the following diagrams

$$\begin{tikzcd} \mathtt{A} \arrow[r, bend left=60, "F"{name=F}] \arrow[r, bend right=60, "L"'{name=L}] & \mathtt{B} \arrow[r, bend left=60, "G"{name=G}] \arrow[r, bend right=60, "M"'{name=M}] & \mathtt{C} \arrow[Rightarrow, from=F, to=L, "\gamma\circ\alpha"] \arrow[Rightarrow, from=G, to=M, "\delta\circ\beta"] \end{tikzcd} \longmapsto \begin{tikzcd} \mathtt{A} \arrow[r, bend left=60, "GF"{name=GF}] \arrow[r, bend right=60, "ML"'{name=ML}] & \mathtt{C} \arrow[Rightarrow, from=GF, to=ML, "(\delta\circ\beta)*(\gamma\circ\alpha)" description] \end{tikzcd}$$

This is sufficient to prove the assertion. ♮

Notice that, given a functor $\phi\colon \mathtt{C} \to \mathtt{D}$ between categories $\mathtt{C}$ and $\mathtt{D}$, for every category $\mathtt{A}$, there arises a natural functor

$$\phi^*\colon \mathrm{Fct}(\mathtt{A},\mathtt{C}) \longrightarrow \mathrm{Fct}(\mathtt{I},\mathtt{D}), \text{ mapping } F \longmapsto \phi F.$$

Lemma 1.5.7. If ϕ is a faithful functor (respectively, fully faithful), then so is the functor ϕ^* for any category $\mathtt{A}$.

Proof. Given any two functors $F, G\colon \mathtt{A} \rightrightarrows \mathtt{C}$, let $\eta\colon F \Rightarrow G$ be any natural transformation. If we apply ϕ^*, we get the following commutative diagram — for every pair $x, y \in \mathtt{A}$ and every morphism $f\colon x \to y$ in $\mathtt{A}$,

$$\begin{tikzcd}
\phi F x \arrow[r, "\phi\eta_x"] \arrow[d, "\phi F(f)"'] & \phi G x \arrow[d, "\phi G(f)"] \\
\phi F y \arrow[r, "\phi\eta_y"] & \phi G y
\end{tikzcd}$$

Since ϕ is faithful (or fully faithful), the mappings $\eta_x \mapsto \phi\eta_x$ and $\eta_y \mapsto \phi\eta_y$ are both injective (or bijective), thus the natural map

$$\mathrm{Mor}_{\mathrm{Fct}(\mathtt{A},\mathtt{C})}(F, G) \longrightarrow \mathrm{Mor}_{\mathrm{Fct}(\mathtt{A},\mathtt{D})}(\phi F, \phi G), \text{ mapping } \eta \mapsto \phi^*\eta,$$

is ensured to be injective (or bijective). ♮

We consider now the category consisting of $\mathcal{U}$-small categories and the morphisms are functors between them, we denote this category by $\mathcal{U}\text{-}\mathtt{Cat}$. Notice that, given any two $\mathcal{U}$-small categories $\mathtt{C}$ and $\mathtt{D}$, the collection of functors between them also forms a category $\mathrm{Mor}_{\mathcal{U}\text{-}\mathtt{Cat}}(\mathtt{C},\mathtt{D}) = \mathrm{Fct}(\mathtt{C},\mathtt{D})$ — this emergent structure gives birth to the concept of a 2-category.

Definition 1.5.8 (Isomorphism of categories)**.** Let $\mathtt{C}$ and $\mathtt{D}$ be categories. We say that $\mathtt{C}$ is isomorphic to $\mathtt{D}$ if there are morphisms $F\colon \mathtt{C} \to \mathtt{D}$ and $G\colon \mathtt{D} \to \mathtt{C}$ such that $GF = \mathrm{id}_{\mathtt{C}}$, and $FG = \mathrm{id}_{\mathtt{D}}$.

A weaker and even more important concept it that of an equivalence between categories.

Definition 1.5.9 (Equivalence of categories)**.** Let $\mathtt{C}$ and $\mathtt{D}$ be categories. We say that a functor $F\colon \mathtt{C} \xrightarrow{\simeq} \mathtt{D}$ is an *equivalence* of the categories $\mathtt{C}$ and $\mathtt{D}$ if there exists a functor $G\colon \mathtt{D} \to \mathtt{C}$, and two natural isomorphisms $\alpha\colon GF \overset{\cong}{\Longrightarrow} \mathrm{id}_{\mathtt{C}}$ and $\beta\colon FG \overset{\cong}{\Longrightarrow} \mathrm{id}_{\mathtt{D}}$. If this is the case, we say that F and G are *quasi-inverses* of each other.

Lemma 1.5.10. Let $F\colon \mathtt{C} \xrightarrow{\simeq} \mathtt{D}$ and $G\colon \mathtt{D} \xrightarrow{\simeq} \mathtt{C}$ be equivalences of given categories $\mathtt{C}$ and $\mathtt{D}$, and suppose that F and G are quasi-inverses. Then there are natural isomorphisms $\alpha\colon GF \overset{\cong}{\Longrightarrow} \mathrm{id}_{\mathtt{C}}$ and $\beta\colon FG \overset{\cong}{\Longrightarrow} \mathrm{id}_{\mathtt{D}}$ for which

$$F\alpha = \beta F \quad \text{and} \quad \alpha G = G\beta.$$

Proof. Let $x \in \mathtt{C}$ be any object. Notice that, since $\mathrm{id}_{\mathtt{C}}\, x = x$ and $\mathrm{id}_{\mathtt{D}}\, Fx = Fx$, it follows that

$$\begin{tikzcd} GFx \arrow[r, "\alpha_x", "\simeq"'] \arrow[d, "F"'] & x \arrow[d, "F"] \\ FG(Fx) \arrow[r, "\simeq", "\beta_{Fx}"'] & Fx \end{tikzcd}$$

is commutative — thus indeed $F\alpha = \beta F$. Now if we let $y \in \mathtt{D}$ be any other object, since $\mathrm{id}_{\mathtt{D}}\, y = y$ and $\mathrm{id}_{\mathtt{C}}\, Gy = Gy$, we get the following commutative diagram

$$\begin{tikzcd} GF(Gy) \arrow[r, "\alpha_{Gy}", "\simeq"'] & Gy \\ FGy \arrow[u, "G"] \arrow[r, "\simeq", "\beta_y"'] & y \arrow[u, "G"'] \end{tikzcd}$$

therefore $\alpha G = G\beta$ as wanted. ♮

Lemma 1.5.11. Let $F\colon \mathtt{C} \to \mathtt{D}$ be a functor, and $\mathtt{D}_0$ be a full subcategory of $\mathtt{D}$ such that, for all $x \in \mathtt{C}$, there exists $y \in \mathtt{D}_0$ and an isomorphism $Fx \simeq y$.

Denote by $\iota\colon \mathtt{D}_0 \hookrightarrow \mathtt{D}$ the canonical embedding functor. Then there exists a functor $F_0\colon \mathtt{C} \to \mathtt{D}_0$ and a natural isomorphism $\alpha\colon F \overset{\cong}{\Longrightarrow} \iota F_0$. Moreover, F_0 is unique up to unique isomorphism[4]. This can be diagrammatically expressed by the following quasi-commutative diagram[5]

$$\begin{tikzcd} \mathtt{C} \arrow[r, "F"] \arrow[dr, dashed, "F_0"'] & \mathtt{D} \\ & \mathtt{D}_0 \arrow[u, hook, "\iota"'] \end{tikzcd}$$

Proof. We first build the functor $F_0\colon \mathtt{C} \to \mathtt{D}_0$:

- Let $x \in \mathtt{C}$ be any object. By means of Zorn's Lemma (see Lemma 5.6.5), we choose $a \in \mathtt{D}_0$ for which exists an isomorphism $\phi_x\colon a \xrightarrow{\cong} F(x)$ in $\mathtt{D}$ — and consequently we let $F_0 x := a$.
- Given any morphism $f\colon x \to y$ in $\mathtt{C}$, we know from construction that the objects $F_0 x := a$ and $F_0 y := b$ are defined so that there exists isomorphisms $\phi_x\colon a \xrightarrow{\cong} Fx$ and $\phi_y\colon b \xrightarrow{\cong} Fy$ in the category $\mathtt{D}$. This allows us to define the morphism $F_0 f\colon F_0 x \to F_0 y$ in $\mathtt{D}$ as the mapping $Ff := \phi_y^{-1}(Ff)\phi_x$ — that is, so that the following diagram commutes

$$\begin{tikzcd} F_0 x \arrow[r, "F_0 f"] \arrow[d, "\phi_x"', "\simeq"] & F_0 y \arrow[d, "\simeq"', "\phi_y"] \\ Fx \arrow[r, "Ff"'] & Fy \end{tikzcd}$$

[4] We say that F_0 is *unique up to unique isomorphism* when, given another functor $G\colon \mathtt{C} \to \mathtt{D}_0$ and natural isomorphism $\beta\colon F \overset{\cong}{\Longrightarrow} \iota G$, there exists a *unique* natural isomorphism $\eta\colon G \overset{\cong}{\Longrightarrow} F_0$ for which $\alpha = \iota\eta\beta$.

[5] A diagram whose nodes are categories an arrows are morphisms is said to be *quasi-commutative* if it commutes up to natural isomorphism of functors.

Notice that from this definition we find naturally that the composition condition is met — given any other morphism $g: y \to z$ in C have that $F_0(gf) = (F_0 g)(F_0 f)$.

This proves the existence of F_0 as a functor. For the isomorphism α, we can define for each pair $x, y \in \mathsf{C}$ the morphisms $\alpha_x := \phi_x^{-1}$ and $\alpha_y := \phi_y^{-1}$, so that the following diagram commutes

$$\begin{tikzcd} Fx \arrow[r, "\alpha_x", "\simeq"'] \arrow[d, "Ff"'] & \iota F_0(x) = a \arrow[d, "\iota F_0(f)"] \\ Fy \arrow[r, "\simeq", "\alpha_y"'] & \iota F_0(y) = b \end{tikzcd}$$

For the uniqueness of F_0 up to unique isomorphism, let $G: \mathsf{C} \to \mathsf{D}_0$ be another functor, together with a natural isomorphism $\beta: F \overset{\cong}{\Longrightarrow} \iota G$. Define $\eta: G \overset{\cong}{\Longrightarrow} F_0$ so that, for each $x \in \mathsf{C}$, we have $\eta_x := \alpha_x \beta_x^{-1}$ — then, η is clearly an isomorphism and also uniquely defined, thus the proposition follows. ♮

Lemma 1.5.12. Let C be any category. There exists a *full subcategory* C_0 of C such that the embedding functor $\iota: \mathsf{C}_0 \hookrightarrow \mathsf{C}$ is a *category equivalence* and any two isomorphic elements in C_0 are equal to each other.

Proof. Let $\sim$ be the equivalence relation on the set $\mathrm{Obj}(\mathsf{C})$ where $x \sim y$ if and only if we have $x \simeq y$ in C. By means of Zorn's lemma, choose for each equivalence class a representative — and define C_0 as the full subcategory of C consisting of such representatives. In this way, if $x, y \in \mathrm{Obj}(\mathsf{C}_0)$ satisfy $x \simeq y$ in C, then necessarily $x = y$.

If we apply Lemma 1.5.11 to the identity functor id_{C}, we find the existence of a unique functor $F_0: \mathsf{C} \to \mathsf{C}_0$ and a natural isomorphism $\alpha: \mathrm{id}_{\mathsf{C}} \overset{\cong}{\Longrightarrow} \iota F_0$ for which $\iota F_0 \simeq \mathrm{id}_{\mathsf{C}}$. Moreover, we have the chain of isomorphisms:

$$\iota(F_0 \iota) = (\iota F_0)\iota \simeq \mathrm{id}_{\mathsf{C}}\, \iota \simeq \iota \simeq \iota\, \mathrm{id}_{\mathsf{C}_0} .$$

From the fact that ι is fully faithful, we obtain $F_0 \iota \simeq \mathrm{id}_{\mathsf{C}_0}$. ♮

Proposition 1.5.13. A functor $F: \mathsf{C} \to \mathsf{D}$ is an *equivalence* of categories if and only if F is *fully faithful* and *essentially surjective*.

Proof. Suppose F is an equivalence of categories, then exists a quasi inverse $G: \mathsf{D} \to \mathsf{C}$ and natural isomorphisms $\alpha: GF \overset{\cong}{\Longrightarrow} \mathrm{id}_{\mathsf{C}}$ and $\beta: FG \overset{\cong}{\Longrightarrow} \mathrm{id}_{\mathsf{D}}$. We now prove that F is both fully faithful and essentially surjective.

- (Fully Faithful) Let $g: Fx \to Fy$ be any morphism in D, thus if there exists an $f: x \to y$ in C for which $Ff = g$, it must be the case that $f = \alpha_y (GFf) \alpha_x^{-1} = \alpha_y (Gg) \alpha_x^{-1}$. Indeed, if $f = \alpha_y (Gg) \alpha_x^{-1}: x \to y$ is a morphism in C, from the distributivity over composition, we have

$$GFf = GF\alpha_y \circ GF(Gg) \circ GF\alpha_x^{-1}$$

then by the naturality of α on the map α_x^{-1} implies

$$\begin{tikzcd} GFx \arrow[r, "\alpha_x"'] \arrow[d, "GF\alpha_x^{-1}"] & x \arrow[d, "\alpha_x^{-1}"] \\ GF(GFx) \arrow[r, "\alpha_{GFx}"'] & GFx \end{tikzcd}$$

but since α_x^{-1} is an isomorphism, then $GF\alpha_x^{-1}\alpha_x^{-1} = \alpha_{GFx}^{-1}\alpha_x^{-1}$ implies in

$$GF\alpha_x^{-1} = \alpha_{GFx}^{-1}.$$

On the other hand, the same can be done for $GF\alpha_y$, yielding $GF\alpha_y = \alpha_{GFy}$, therefore

$$GFf = \alpha_{GFy}(GFGg)\alpha_{GFx}^{-1}. \tag{1.1}$$

The naturality of α on the map Gg gives

$$\begin{tikzcd} GF(GFx) \arrow[r, "\alpha_{GFx}"] \arrow[d, "GF(Gg)"'] & GFx \arrow[d, "Gg"] \\ GF(GFy) \arrow[r, "\alpha_{GFy}"'] & GFy \end{tikzcd}$$

that is, $GF(Gg)\alpha_{GFx}^{-1} = \alpha_{GFy}^{-1}Gg$, thus substituting in Eq. (1.1) we obtain finally that

$$GFf = Gg.$$

From this we conclude that F is fully faithful.

- (Essentially surjective) Let $a \in \mathtt{D}$ be any element and simply consider $x = Ga \in \mathtt{D}$ then from β we obtain that $Fx = FGa \simeq a$.

For the converse, suppose F is both fully faithful and essentially surjective. Using Lemma 1.5.12, let $\mathtt{C}_0$ be the full subcategory of $\mathtt{C}$ such that $\iota_{\mathtt{C}} \colon \mathtt{C}_0 \hookrightarrow \mathtt{C}$ is an equivalence and if $x \simeq y$ in $\mathtt{C}$ then $x = y$ in $\mathtt{C}_0$. Let $\kappa_{\mathtt{C}}$ be a quasi-inverse of $\iota_{\mathtt{C}}$. Apply Lemma 1.5.12 again, but now to the category $\mathtt{D}$, yielding a category $\mathtt{D}_0$, an embedding $\iota_{\mathtt{D}}$ and a quasi-inverse $\kappa_{\mathtt{D}}$. Notice that the composition of functors $\kappa_{\mathtt{D}}F\iota_{\mathtt{C}} \colon \mathtt{C}_0 \xrightarrow{\simeq} \mathtt{D}_0$ is an isomorphism. Let H be the inverse of $\kappa_{\mathtt{D}}F\iota_{\mathtt{C}}$, and define $G := \iota_{\mathtt{C}}H\kappa_{\mathtt{D}}$, then G is a quasi-inverse of F. ♮

Corollary 1.5.14. Let $F \colon \mathtt{C} \to \mathtt{D}$ be a *fully faithful* functor. Then there exists a *full subcategory* $\mathtt{D}_0$ of $\mathtt{D}$ and an *equivalence* of categories $G \colon \mathtt{C} \xrightarrow{\simeq} \mathtt{D}_0$ for which F is *isomorphic* to $\iota_{\mathtt{D}}G$ — where $\iota_{\mathtt{D}} \colon \mathtt{D}_0 \hookrightarrow \mathtt{D}$ is the embedding functor.

Proof. Define $\mathtt{D}_0$ to be the full category with $\mathrm{Obj}(\mathtt{D}_0) := \{Fx : x \in \mathtt{C}\}$. Now define $G \colon \mathtt{D} \to \mathtt{D}_0$ to be the functor $G = F$, we find that clearly $F \simeq \iota_{\mathtt{D}}G$. ♮

Example 1.5.15. For any two given categories $\mathtt{C}$ and $\mathtt{D}$, there is an isomorphism of categories $\mathrm{Fct}(\mathtt{C}, \mathtt{D})^{\mathrm{op}} \simeq \mathrm{Fct}(\mathtt{C}^{\mathrm{op}}, \mathtt{D}^{\mathrm{op}})$. Explicitly, such isomorphism maps $F \mapsto \mathrm{op}F\mathrm{op}$.

Definition 1.5.16 (Essentially $\mathcal{U}$-small)**.** A category C is said to be *essentially $\mathcal{U}$-small* if it is equivalent to a $\mathcal{U}$-small category. Equivalently, C is essentially $\mathcal{U}$-small if and only if C is a $\mathcal{U}$-category and there exists a subset $S \subseteq \mathrm{Obj}(\mathsf{C})$ that is $\mathcal{U}$-small for which, given any $x \in \mathsf{C}$, there exists $y \in S$ such that $x \simeq y$.

Definition 1.5.17 (Half-full)**.** We define the concepts of *half-fullness* for functors and subcategories:

(a) A functor $F \colon \mathsf{C} \to \mathsf{D}$ is said to be *half-full* if for any two $x, y \in \mathsf{C}$ for which $Fx \simeq Fy$ in D, there exists *an* isomorphism $x \simeq y$ in C.

(b) A subcategory C_0 of C is said to be *half-full* if the embedding functor $\iota \colon \mathsf{C}_0 \hookrightarrow \mathsf{C}$ is half-full.

An analogous but *less strict* proposition when compared to Corollary 1.5.14 is done by substituting the condition of fully faithfulness to only that of faithful and half-full. It goes as follows.

Proposition 1.5.18. Let $F \colon \mathsf{C} \to \mathsf{D}$ be a *faithful* and *half-full* functor. Then there exists a subcategory D_0 of D for which

$$F(\mathrm{Obj}(\mathsf{C})) \subseteq \mathrm{Obj}(\mathsf{D}_0), \text{ and } F(\mathrm{Mor}(\mathsf{C})) \subseteq \mathrm{Mor}(\mathsf{D}_0).$$

Moreover, F induces an *equivalence* of categories $\mathsf{C} \simeq \mathsf{D}_0$ and the embedding $\mathsf{D}_0 \hookrightarrow \mathsf{D}$ is *faithful* and *half-full*.

Proof. Let D_0 be the category with $\mathrm{Obj}(\mathsf{D}_0) := \{F(x) : x \in \mathsf{C}\}$, and for each two $x, y \in \mathsf{C}$ define $\mathrm{Mor}_{\mathsf{D}_0}(Fx, Fy) := F(\mathrm{Mor}_{\mathsf{C}}(x, y))$ — so that $\mathrm{Mor}_{\mathsf{D}_0}(Fx, Fy) \subseteq \mathrm{Mor}_{\mathsf{D}}(Fx, Fy)$. Since F is half-full, the definition of $\mathrm{Mor}_{\mathsf{D}_0}(Fx, Fy)$ is independent of the initial choice of $x, y \in \mathsf{C}$ — indeed, if $x', y' \in \mathsf{C}$ are such that $Fx' \simeq Fx$ and $Fy' \simeq Fy$ in D_0, then $x' \simeq x$ and $y' \simeq y$ in C, thus $\mathrm{Mor}_{\mathsf{C}_0}(x, y) \simeq \mathrm{Mor}_{\mathsf{C}_0}(x', y')$.

Restricting the codomain of F to D_0, we find that the induced functor $F \colon \mathsf{C} \to \mathsf{D}_0$ is fully faithful and essentially surjective — thus by Proposition 1.5.13 F is an equivalence of categories. ♮

1.6 Comma Categories

Definition 1.6.1 (Comma category)**.** Let $F \colon \mathsf{A} \to \mathsf{C}$ and $G \colon \mathsf{B} \to \mathsf{C}$ be two functors. The comma category $F \downarrow G$ induced by the functors F and G is defined as follows:

- The objects of $F \downarrow G$ are triples (a, f, b) — where $a \in \mathsf{A}$, $b \in \mathsf{B}$, and $f \colon Fa \to Gb$ is a morphism of C.
- A morphism $\phi \colon (a, f, b) \to (a', f', b')$ in $F \downarrow G$ is a pair of (α, β) where $\alpha \colon a \to a'$ is a morphism of A, while $\beta \colon b \to b'$ is a morphism of B — moreover, such morphisms

are such that the following diagram commutes

$$\begin{tikzcd}
Fa \arrow[r, "f"] \arrow[d, "F\alpha"'] & Gb \arrow[d, "G\beta"] \\
Fa' \arrow[r, "f'"'] & Gb'
\end{tikzcd}$$

- The composition of *compatible* morphisms (α, β) and (α', β') in $F \downarrow G$ is induced by the composition law of $\mathtt{A}$ and $\mathtt{B}$ as follows:

$$(\alpha', \beta') \circ (\alpha, \beta) \coloneqq (\alpha'\alpha, \beta'\beta).$$

Proposition 1.6.2 (Projection functors in $F \downarrow G$)**.** Let $F\colon \mathtt{A} \to \mathtt{C}$ and $G\colon \mathtt{B} \to \mathtt{C}$ be functors. There are two functors $A\colon F \downarrow G \to \mathtt{A}$ and $B\colon F \downarrow G \to \mathtt{B}$, and a canonical natural transformation $\eta\colon FA \Rightarrow GB$. The scenario can be depicted in the following diagram

$$\begin{tikzcd}
F \downarrow G \arrow[r, "B"] \arrow[d, "A"] \arrow[dr, bend right, "FA"'] \arrow[dr, bend left, "GB"] \arrow[dr, phantom, "\overset{\eta}{\Nearrow}"] & \mathtt{B} \arrow[d, "G"] \\
\mathtt{A} \arrow[r, "F"] & \mathtt{C}
\end{tikzcd}$$

It is to be noted that such diagram *does not commute in general* — that is, we may have comma categories where $FA \neq GB$.

Proof. Define the functors A and B as follows — for every object $(a, f, b) \in F \downarrow G$ and morphism $(\alpha, \beta) \in \operatorname{Mor}(F \downarrow G)$:

- Map $A(a, f, b) \coloneqq a$, and $A(\alpha, \beta) \coloneqq \alpha$.
- Map $B(a, f, b) \coloneqq b$, and $A(\alpha, \beta) \coloneqq \beta$.

For the natural transformation, simply define $\eta_{(a,f,b)} \coloneqq f\colon Fa \to Fb$ which must be natural from the construction of comma categories. ♮

Proposition 1.6.3 (Comma category property)**.** Let $F\colon \mathtt{A} \to \mathtt{C}$ and $G\colon \mathtt{B} \to \mathtt{C}$ be two functors. If there exists a category $\mathtt{D}$ with two functors $A'\colon \mathtt{D} \to \mathtt{A}$ and $B'\colon \mathtt{D} \to \mathtt{B}$, and a natural transformation $\eta'\colon FA' \Rightarrow GB'$ — then there *exists a unique* functor $W\colon \mathtt{D} \to F \downarrow G$ such that

$$\alpha * W = \alpha',$$

and that the following diagram commutes

$$\begin{tikzcd}
 & \mathtt{D} \arrow[dl, bend right, "A'"'] \arrow[d, dashed, "W"] \arrow[dr, bend left, "B'"] & \\
 & F \downarrow G \arrow[dl, "A"] \arrow[dr, "B"'] & \\
\mathtt{A} & & \mathtt{B}
\end{tikzcd}$$

Proof. For every $d \in D$, define $Wd := (A'd, \eta'_d, B'd)$, and for each morphism $f \in \mathrm{Mor}(\mathtt{D})$, we define $Wf := (A'f, B'f)$. This completely determines W and thus shows that, if it exists, it is unique. To show that $W\colon \mathtt{D} \to F \downarrow G$ is indeed a functor, we note that:

- Given morphisms $f\colon x \to y$ and $g\colon x \to z$ of $\mathtt{D}$, we have

$$\begin{aligned} Wg \circ Wf &= (A'g, B'g) \circ (A'f, B'f) \\ &= (A'g \circ A'f, B'g \circ B'f) \\ &= (A'(gf), B'(gf)) \\ &= W(gf). \end{aligned}$$

  Note that although we used the same symbol $\circ$ for composition, one should note that they have different laws.

- Moreover, for any $d \in D$ we have

$$W\,\mathrm{id}_d = (A'\,\mathrm{id}_d, B'\,\mathrm{id}_d) = (\mathrm{id}_{A'd}, \mathrm{id}_{B'd}) = \mathrm{id}_{Wd}\,.$$

♮

Definition 1.6.4 (Category of elements)**.** Let $F\colon \mathtt{C} \to \mathtt{Set}$ be a covariant functor. The *category of elements* of $\mathtt{C}$ associated to F is denoted by $\mathrm{El}_F(\mathtt{C})$ whose objects are pairs (c, s) for $c \in \mathtt{C}$ and $s \in Fc$, and morphisms between any two objects $(c, s), (c', s') \in \mathrm{El}_F(\mathtt{C})$ is defined as

$$\mathrm{Mor}_{\mathrm{El}_F(\mathtt{C})}((c, s), (c', s')) := \{u \in \mathrm{Mor}_{\mathtt{C}}(c, c') : F(u)(s) = s'\}.$$

For a contravariant $G\colon \mathtt{C}^{\mathrm{op}} \to \mathtt{Set}$, the category of elements of $\mathtt{C}$ associated to G is composed of objects (c, s) for $c \in \mathtt{C}$ and $s \in Gc$, and the collection of morphisms between objects $(c, s), (c', s') \in \mathrm{El}_G(\mathtt{C})$ is given by

$$\mathrm{Mor}_{\mathrm{El}_G(\mathtt{C})}((c, s), (c', s')) := \{u \in \mathrm{Mor}_{\mathtt{C}}(c', c) : F(u)(s') = s\}.$$

Corollary 1.6.5. Let $F\colon \mathtt{C} \to \mathtt{D}$ be a functor. The category of elements $\mathrm{El}_F(\mathtt{C})$ is exactly the comma category $1 \downarrow F$ — where $1\colon \mathbf{1} \to \mathtt{Set}$ is the functor from the discrete category $\mathbf{1}$ with a single object $\star$ to the category of sets, mapping $\star \mapsto \{*\}$.

1.7 Yoneda Lemma

Remark 1.7.1. I again stress that $\mathtt{Set}$, for us, is defined to be the category whose objects are $\mathcal{U}$-sets for a given universe $\mathcal{U}$ and set-functions between these sets — this is a relevant remark, since confusions with that would lead one to undesirable size issues.

Definition 1.7.2 (Category of presheaves & Yoneda functors)**.** Given a $\mathcal{U}$-category C, we define the *big* category of *presheaves* $\mathtt{Psh}(\mathsf{C})$ and a *big* category of functors $\mathtt{Psh}(\mathsf{C}^{\mathrm{op}}) = [C, \mathtt{Set}]$. Together with such categories we define functors

$$
\begin{aligned}
\text{よ}_{\mathsf{C}}\colon \mathsf{C} &\longrightarrow \mathtt{Psh}(\mathsf{C}), \text{ mapping } x \mapsto \mathrm{Mor}_{\mathsf{C}}(-, x) \text{ and } f \mapsto f_*, \\
\text{よ}'_{\mathsf{C}}\colon \mathsf{C} &\longrightarrow \mathtt{Psh}(\mathsf{C}^{\mathrm{op}}), \text{ mapping } x \mapsto \mathrm{Mor}_{\mathsf{C}}(x, -) \text{ and } f \mapsto f^*.
\end{aligned}
$$

The functors よ_{C} and $\text{よ}'_{\mathsf{C}}$ are called *Yoneda functors*.

Remark 1.7.3 (f_* and f^* as natural transformations)**.** The attentive reader may note that setting $\text{よ}_{\mathsf{C}} f = f_*$ is not quite right since f_* is not a morphism in the category of presheaves $\mathtt{Psh}(\mathsf{C})$ according to our definition that roots from Lemma 1.3.2 — you are right, but we are being sloppy here just to simplify how we treat our objects and notations. The arrow f_* (and f^*) is to be interpreted as the *natural transformation* $f_*\colon \text{よ}_{\mathsf{C}} x \Rightarrow \text{よ}_{\mathsf{C}} y$, where $f\colon x \to y$ — for which we define, for every object $z \in \mathsf{C}$, the morphism $f_*\colon \mathrm{Mor}_{\mathsf{C}}(z, x) \to \mathrm{Mor}_{\mathsf{C}}(z, y)$ mapping $g \mapsto fg$.

Remark 1.7.4 (Size issues)**.** Although C is a $\mathcal{U}$-category, its categories of presheaves described above *need not* be a $\mathcal{U}$-category — however, if C happens to be $\mathcal{U}$*-small*, then $\mathtt{Psh}(\mathsf{C})$ and $\mathtt{Psh}(\mathsf{C}^{\mathrm{op}})$ are both $\mathcal{U}$*-categories*.

Lemma 1.7.5 (Yoneda)**.** The Yoneda lemma states that:

(a) For all functors $F\colon \mathsf{C}^{\mathrm{op}} \to \mathtt{Set}$ and $x \in \mathsf{C}$, there is a natural isomorphism:

$$\mathrm{Mor}_{\mathtt{Psh}(\mathsf{C})}(\text{よ}_{\mathsf{C}} x, F) \simeq Fx.$$

(b) For all functors $G\colon \mathsf{C} \to \mathtt{Set}$ and $x \in \mathsf{C}$, there is a natural isomorphism:

$$\mathrm{Mor}_{\mathtt{Psh}(\mathsf{C}^{\mathrm{op}})}(G, \text{よ}'_{\mathsf{C}} x) \simeq Gx.$$

These natural isomorphisms have a *functorial* nature — they define, respectively, functors $\mathsf{C}^{\mathrm{op}} \times \mathtt{Psh}(\mathsf{C}) \to \mathtt{Set}$ and $\mathtt{Psh}(\mathsf{C}^{\mathrm{op}})^{\mathrm{op}} \times \mathsf{C} \to \mathtt{Set}$

Proof. By means of **??** one only needs to prove one of the two statements, since the other follows from duality. We prove the item (a).

Let ϕ be the morphism making

$$
\begin{tikzcd}
\mathrm{Mor}_{\mathtt{Psh}(\mathsf{C})}(\text{よ}_{\mathsf{C}} x, F) \arrow[r] \arrow[rr, bend left, "\phi"] & \mathrm{Mor}_{\mathtt{Set}}(\mathrm{Mor}_{\mathsf{C}}(x, x), Fx) \arrow[r] & Fx \\
\eta \arrow[r, mapsto] & \eta_x \arrow[r, mapsto] & \eta_x(\mathrm{id}_x)
\end{tikzcd}
$$

commute. Notice that the choice of $\mathrm{id}_x \in \mathrm{Mor}_{\mathsf{C}}(x,x) = \text{よ}_{\mathsf{C}}x(x)$ is done so that we can define a distinguished point $\eta_x(\mathrm{id}_x)$ from the set F_x.

Define now $\psi\colon Fx \to \mathrm{Mor}_{\mathsf{Psh(C)}}(\text{よ}_{\mathsf{C}}(x), F)$ to be the map $Fx \ni a \mapsto \eta^a\colon \text{よ}_{\mathsf{C}}x \Rightarrow F$, where the natural transformation η^a associated to a is defined, for every object $y \in \mathsf{C}$, by the map $\eta^a_y\colon \mathrm{Mor}_{\mathsf{C}}(y,x) \to Fx$ sending $f \mapsto F(f)(a)$.

Notice that any natural transformation $\eta \in \mathrm{Mor}_{\mathsf{Psh(C)}}(\text{よ}_{\mathsf{C}}x, F)$ is completely determined by $\eta_x(\mathrm{id}_x)$. Indeed, given $f \in \mathrm{Mor}_{\mathsf{C}}(y,x)$ one has the following diagram, which comes from the naturality of η:

$$\begin{tikzcd}
\mathrm{Mor}_{\mathsf{C}}(x,x) \arrow[r, "\eta_x"] \arrow[d, "f^*"'] & Fx \arrow[d, "Ff"] \\
\mathrm{Mor}_{\mathsf{C}}(y,x) \arrow[r, "\eta_y"] & Fy
\end{tikzcd}$$

Thus we have $\eta_y(f) = (Ff)(\eta_x(\mathrm{id}_x))$ since $f^*(\mathrm{id}_x) = \mathrm{id}_x f = f$ — which shows that η is determined by the distinguished map id_x. That is, ψ is determined, for each $a \in Fx$ by $\eta^a_x(\mathrm{id}_x)$. This is thus sufficient to prove that ϕ and ψ are mutual inverses and therefore stablish the wanted isomorphism. $\natural$

Corollary 1.7.6. The Yoneda functors よ and よ$'$ are *fully faithful*.

Proof. From duality, we only check that the map $\mathrm{Mor}_{\mathsf{C}}(x,y) \to \mathrm{Mor}_{\mathsf{Psh(C)}}(\text{よ}_{\mathsf{C}}x, \text{よ}_{\mathsf{C}}y)$ sending $f \mapsto \eta^f$, where $\eta^f_z(g) = f_*(g) = fg$ for every $z \in \mathsf{C}$ and map $g \in \mathrm{Mor}_{\mathsf{C}}(z,x)$ — is a bijection for all $x, y \in \mathsf{C}$. Such map is certainly injective, since given two distinct maps $f, h \in \mathrm{Mor}_{\mathsf{C}}(x,y)$ we have $f_* \neq h_*$. Moreover, we have the isomorphism

$$\mathrm{Mor}_{\mathsf{Psh(C)}}(\text{よ}_{\mathsf{C}}x, \text{よ}_{\mathsf{C}}y) \simeq \text{よ}_{\mathsf{C}}y(x) = \mathrm{Mor}_{\mathsf{C}}(x,y)$$

from Lemma 1.7.5. Let $\eta\colon \text{よ}_{\mathsf{C}}x \Rightarrow \text{よ}_{\mathsf{C}}y$ be any natural transformation — we wish to find $f \in \mathrm{Mor}_{\mathsf{C}}(x,y)$ for which $\eta_z(g) = fg$ for every $z \in \mathsf{C}$ and $g \in \mathrm{Mor}_{\mathsf{C}}(z,x)$. The natural choice is given by $f := \eta_x(\mathrm{id}_x)$. By the naturality of η we have, for any $w \in \mathsf{C}$

$$\begin{tikzcd}
\mathrm{Mor}_{\mathsf{C}}(w,x) \arrow[r, "\eta_w"] \arrow[d, "g^*"'] & \mathrm{Mor}_{\mathsf{C}}(w,y) \arrow[d, "g^*"] \\
\mathrm{Mor}_{\mathsf{C}}(z,x) \arrow[r, "\eta_z"] & \mathrm{Mor}_{\mathsf{C}}(z,y)
\end{tikzcd}$$

Thus given $h\colon w \to x$ in C we have the equality

$$\eta_w(h)g = \eta_z(hg).$$

Therefore, if we restrict $w = x$ and $h = \mathrm{id}_x$ we get $\eta_x(\mathrm{id}_x)g = fg = \eta_z(g)$ — which is what we wanted, because this means that f will have image η under our mapping. $\natural$

Remark 1.7.7 (Full subcategory of the presheaf category)**.** From Corollary 1.7.6 one can conclude that C can be viewed as a *full subcategory* of the presheaf category $\mathsf{Psh}(\mathsf{C})$ (or of the category $\mathsf{Psh}(\mathsf{C}^{\mathrm{op}})$).

The following corollary establishes the main idea of the Yoneda lemma — one knows about an object by knowing how it interacts with other objects.

Corollary 1.7.8. Given a category C and two objects $x, y \in \mathsf{C}$. Then there exists an isomorphism $x \simeq y$ if and only if there exists an isomorphism $\mathrm{Mor}_{\mathsf{C}}(-, x) \simeq \mathrm{Mor}_{\mathsf{C}}(-, y)$ — or, for the covariant version, $\mathrm{Mor}_{\mathsf{C}}(x, -) \simeq \mathrm{Mor}_{\mathsf{C}}(y, -)$.

Proof. Since functors preserve isomorphisms, if $x \simeq y$ for any two objects $x, y \in \mathsf{C}$, we obtain that $\text{よ}_{\mathsf{C}} x \simeq \text{よ}_{\mathsf{C}} y$. From Proposition 1.4.17 and since よ_{C} is fully faithful we find that $\text{よ}_{\mathsf{C}} x \simeq \text{よ}_{\mathsf{C}} y$ implies $x \simeq y$. ♮

Corollary 1.7.9. Let $F \colon \mathsf{C} \to \mathsf{D}$ be a functor of $\mathcal{U}$-categories and assume that C is $\mathcal{U}$-small. For every *functor* $A \in \mathsf{Psh}(\mathsf{C})$, the category C_A *associated*[6] with the functor

$$\mathsf{C} \xrightarrow{F} \mathsf{D} \xrightarrow{\text{よ}_{\mathsf{D}}} \mathsf{Psh}(\mathsf{C})$$

is $\mathcal{U}$*-small*. Analogously, given a *functor* $B \in \mathsf{Psh}(\mathsf{C}^{\mathrm{op}})$, the category C^B *associated* with the functor

$$\mathsf{C} \xrightarrow{F} \mathsf{D} \xrightarrow{\text{よ}'_{\mathsf{D}}} \mathsf{Psh}(\mathsf{C}^{\mathrm{op}})$$

is $\mathcal{U}$*-small*.

Proof. Recalling the definition, the category C_A associated with $\text{よ}_{\mathsf{C}} F$ is given by the following collections:

$$\mathrm{Obj}(\mathsf{C}_A) = \{(x, \phi) : x \in \mathsf{C} \text{ and } \phi \colon \text{よ}_{\mathsf{C}} Fx \to A \text{ in } \mathsf{D}\},$$
$$\mathrm{Mor}_{\mathsf{C}_A}((x, \phi), (y, \psi)) = \{f \in \mathrm{Mor}_{\mathsf{C}}(x, y) : \phi = \psi\, \text{よ}_{\mathsf{C}} F(f)\},$$

where $(x, \phi), (y, \psi) \in \mathsf{C}_A$ are any two objects, in other words, the collection of morphisms so that

$$\begin{tikzcd} \text{よ}_{\mathsf{C}} Fx \arrow[r, "\phi"] \arrow[dr, bend right, "\text{よ}_{\mathsf{C}} F(f)"'] & A \\ & \text{よ}_{\mathsf{C}} Fy \arrow[u, "\psi"'] \end{tikzcd}$$

commutes in D. Since C is assumed to be $\mathcal{U}$-small by hypothesis, it follows that the collection of arrows

$$\coprod_{x \in \mathsf{C}} \mathrm{Mor}_{\mathsf{C}}(Fx, A)$$

is a disjoint union of $\mathcal{U}$-small sets, thus is itself $\mathcal{U}$-small. The same proof can be analogously constructed for the covariant case C^B. ♮

[6] Recall Definition 1.4.21.

Corollary 1.7.10. Let C be a category and $f: x \to y$ a morphism in C. If for every $z \in \mathsf{C}$ the morphism

$$f_*: \mathrm{Mor}_{\mathsf{C}}(z, x) \to \mathrm{Mor}_{\mathsf{C}}(z, y)$$

is an *isomorphism* (or $f^*: \mathrm{Mor}_{\mathsf{C}}(x, z) \to \mathrm{Mor}_{\mathsf{C}}(y, z)$ for the covariant case), then f is an *isomorphism*.

Proof. If the condition is met, then $よ_{\mathsf{C}} f: よ_{\mathsf{C}} x \to よ_{\mathsf{C}} y$ is an isomorphism. Since $よ_{\mathsf{C}}$ is fully faithful, then f is an isomorphism. Notice that we already stated this proposition back when we were studying dual categories, with the Yoneda lemma we were able to prove it functorially — see Lemma 1.3.2. ♮

Functor Representation

Definition 1.7.11 (Representable functor)**.** A presheaf $F: \mathsf{C}^{\mathrm{op}} \to \mathtt{Set}$ (or a functor $\mathsf{C} \to \mathtt{Set}$) is said to be *representable* if there exists a natural isomorphism $よ_{\mathsf{C}} x \overset{\cong}{\Longrightarrow} F$ (or $F \overset{\cong}{\Longrightarrow} よ'_{\mathsf{C}} x$) for some $x \in \mathsf{C}$ — the object x is called a *representative* of F.

Corollary 1.7.12 (Representative uniqueness)**.** The representative of a representable functor is *unique up to unique isomorphism*.

Proof. Since $よ_{\mathsf{C}}$ (and $よ'_{\mathsf{C}}$) is fully faithful, it follows that, if $x, y \in \mathsf{C}$ are representatives of F, then there exists a natural isomorphism $よ_{\mathsf{C}} x \simeq よ_{\mathsf{C}} y$ hence, evoking Proposition 1.4.17, we find a unique isomorphism $x \simeq y$. ♮

Corollary 1.7.13 (Universal element)**.** Let $F: \mathsf{C}^{\mathrm{op}} \to \mathtt{Set}$ (or $\mathsf{C} \to \mathtt{Set}$) be any representable presheaf, with representative object $x_0 \in \mathsf{C}$. Then the natural isomorphism

$$\eta: よ_{\mathsf{C}} x_0 \overset{\cong}{\Longrightarrow} F$$

is uniquely determined by an element $s_0 \in F x_0$ — such an element is called a *universal element* of F.

Proof. Indeed, if F is represented by x_0, then by the Yoneda lemma we obtain an isomorphism

$$\mathrm{Mor}_{\mathtt{Psh}(\mathsf{C})}(よ_{\mathsf{C}} x_0, F) \simeq F x_0,$$

therefore for each $s_0 \in F x_0$ there exists a unique corresponding natural transformation $\eta: よ_{\mathsf{C}} x_0 \Rightarrow F$ (notice we didn't require it to be an isomorphism) and η is completely determined by s_0 in the sense that, for every $y \in \mathsf{C}$ and element $t \in F y$, there exists a unique morphism $f: x_0 \to y$ such that $F(f)(s_0) = t$. ♮

Corollary 1.7.14. Let $F: \mathsf{C} \to \mathtt{Psh}(\mathsf{D})$ be a functor. If, for every $c \in \mathsf{C}$, there exists an object $d \in \mathsf{D}$ such that $Fc \simeq d$, then there exists a *unique* — up to unique isomorphism — functor $F_0: \mathsf{C} \to \mathsf{D}$ for which

$$F \simeq よ_{\mathsf{D}} F_0.$$

Proof. From Remark 1.7.7 we know that $\mathtt{D}$ is a full subcategory of the presheaf category $\mathtt{Psh}(\mathtt{D})$. Since $よ_{\mathtt{D}} : D \to \mathtt{Psh}(\mathtt{D})$ is fully faithful, by means of Lemma 1.5.11 one finds that there exists a unique functor $F_0 : \mathtt{C} \to \mathtt{D}$ (up to unique isomorphism) and a natural isomorphism $F \simeq よ_{\mathtt{D}} F_0$. In other words, the following diagram is quasi-commutative

$$\begin{tikzcd} \mathtt{C} \arrow[r, "F"] \arrow[dr, dashed, bend right, "F_0"'] & \mathtt{Psh}(\mathtt{D}) \\ & \mathtt{D} \arrow[u, "よ_{\mathtt{D}}"'] \end{tikzcd}$$

♮

Properties of the Category of Elements

Proposition 1.7.15. Let $F : \mathtt{C}^{\mathrm{op}} \to \mathtt{Set}$ be a functor (or $F : \mathtt{C} \to \mathtt{Set}$ for the covariant case). Then F is *representable* if and only if the category of elements $\mathrm{El}_F(\mathtt{C})$ has a *final* object (or *initial* object for the covariant case).

Proof. We prove the proposition for the contravariant case, the covariant case has a completely analogous construction and can be obtained by duality. Let F be representable, with representative $x_0 \in \mathtt{C}$, and universal element s_0 — that is, defining a natural isomorphism $よ_{\mathtt{C}} x_0 \simeq F$. Consider any element $(x, s) \in \mathrm{El}_F(\mathtt{C})$ — the morphisms $(x, s) \to (x_0, s_0)$ in $\mathrm{El}_F(\mathtt{C})$ are arrows $f : x \to x_0$ in $\mathtt{C}$ for which $F(f)(s_0) = s$. Nonetheless, since $よ_{\mathtt{C}}$ is fully faithful the morphism f satisfying such condition is unique — therefore (x_0, s_0) is final in the category of elements $\mathrm{El}_F(\mathtt{C})$.

On the other hand, assume that $\mathrm{El}_F(\mathtt{C})$ has a final object (x_0, s_0). We construct a natural transformation $\theta : F \Rightarrow よ_{\mathtt{C}} x_0$ by the maps $\theta_x : Fx \to \mathrm{Mor}_{\mathtt{C}}(x, x_0)$, for $x \in \mathtt{C}$, sending $s \mapsto f$ where $F(f)(s_0) = s$ — this is well defined since (x_0, s_0) is final and thus $f : x \to x_0$ is the unique morphism in $\mathtt{C}$ with such property. Moreover, θ_x is clearly injective: if $s, s' \in Fx$ are any two elements such that $\theta_x(s) = \theta_x(s') = f$, then $F(f)(s_0) = s$ and also $F(f)(s_0) = s'$ — which can only be the case for $s = s'$. The surjectivity of θ_x is ensured by the fact that, given a morphism $f : x \to x_0$ in $\mathtt{C}$, one can choose the element $F(f)(s_0) \in Fx$ so that $\theta_x(F(f)(s_0)) = f$. We conclude that θ_x is a bijection and hence θ is a natural isomorphism $F \simeq よ_{\mathtt{C}} x_0$. ♮

Definition 1.7.16 (Two sided represented functor)**.** Let $\mathtt{C}$ be a $\mathcal{U}$-category, then there is a functor $\mathrm{Mor}(-,-) : \mathtt{C}^{\mathrm{op}} \times \mathtt{C} \to \mathtt{Set}$ defined by mapping objects $(x, y) \mapsto \mathrm{Mor}(x, y)$ and morphisms $(f, g) : (a, y) \to (x, b)$ to a set-function $(f^*, g_*) : \mathrm{Mor}(x, y) \to \mathrm{Mor}(a, b)$ given by the mapping $g \mapsto hgf$.

Chapter 2

Limits

2.1 Products

Definition 2.1.1 (Product)**.** Let J be a set, and $(C_j)_{j\in I}$ be a collection of objects in a category C.

A *product* of such collection, *if it exists*, is a pair $(P, (\pi_j)_{j\in J})$—where P is an object of C, and for every $j \in J$, $\pi_j\colon P \to C_j$ is a morphism of C.

Furthermore, this pair has to satisfy the following *universal property*: for every pair $(Q, (q_j)_{j\in J})$—where $Q \in \mathsf{C}$ and for every $j \in J$, $q_j\colon Q \to C_j$ is a morphism in C—there exists a *unique morphism* $f\colon Q \to P$ of C such that, for every $j \in J$ the following diagram commutes

$$\begin{tikzcd} Q \arrow[d, dashed, "f"'] \arrow[dr, bend left, "q_j"] & \\ P \arrow[r, "\pi_j"'] & C_j \end{tikzcd}$$

Proposition 2.1.2 (Uniqueness)**.** The product of a collection of objects, if existent, is *unique up to isomorphism*.

Proof. Let C be a category admitting the product of a collection $(C_j)_{j\in J}$ of objects of C, and J is a set. Let $(P, (p_j)_{j\in J})$ and $(Q, (q_j)_{j\in J})$ be products of $(C_j)_{j\in J}$ in the category C. Since P and Q are products, there exists unique morphisms $f\colon Q \to P$ and $g\colon P \to Q$ such that, for all $j \in J$, we have $q_j = p_j f$ and $p_j = q_j g$.

Moreover, one can apply the product property of P to P itself: there exists a unique morphism $h\colon P \to P$ such that, for all $j \in J$, we have $p_j = p_j h$. For that to be true, it is clear that we must have $h = \mathrm{id}_P$. Note, however, that for each $j \in J$,

$$p_j = q_j g = p_j f g.$$

Since h is unique, we obtain $fg = h = \mathrm{id}_P$. On the other hand, applying the product property of Q in Q will yield $gf = \mathrm{id}_Q$. This shows that $Q \simeq P$ in C, via f and g. ♮

Proposition 2.1.3 (Products are independent of ordering)**.** Let I be a set and $(J_k)_{k\in K}$ be a partition of I by disjoint subsets $J_k \subseteq I$. Let $(C_i)_{i\in I}$ be a collection of objects in a category C. Then, if all the products presented below exist in C, they are isomorphic:

$$\prod_{i \in I} C_i \simeq \prod_{k \in K} \left(\prod_{j \in J_k} C_j \right).$$

Proof. Define the following collections:

- Let $(q_k)_{k \in K}$ be the collection of morphisms $q_{k_0} \colon \prod_{k \in K} \left(\prod_{j \in J_k} C_j \right) \to C_{k_0}$, for each $k_0 \in K$, associated with the product $\prod_{k \in K} \left(\prod_{j \in J_k} C_j \right)$ in C.
- For all $k \in K$, let $(p_j)_{j \in J_k}$ be the collection of morphisms $p_{j_0} \colon \prod_{j \in J_k} C_j \to C_{j_0}$, for each $j_0 \in J_k$ associated with the product $\prod_{j \in J_k} C_j$ in C.

Let $(L, (\ell)_{i \in I})$ be a pair where $L \in \mathsf{C}$ and $\ell_i \colon L \to C_i$ is a morphism in C for each $i \in I$.

Define a collection $(g_k)_{k \in K}$ where, for each $k_0 \in K$, we let $g_{k_0} \colon L \to \prod_{j \in J_{k_0}} C_j$ be the unique morphism of C such that $\ell_i = g_{k_0} p_i$ for every $i \in J_k$. Let $f \colon \prod_{k \in K} \left(\prod_{j \in J_k} C_j \right)$ be the unique morphism of C such that $g_k = q_k f$.

We see that the following diagram commutes for every $k_0 \in K$ and $i \in J_k$:

$$\begin{tikzcd}
L \arrow[d, dashed, "f"'] \arrow[dr, dashed, "g_{k_0}"] \arrow[drr, bend left, "\ell_i"] & & \\
\prod_{k \in K} \left(\prod_{j \in J_k} C_j \right) \arrow[r, "q_{k_0}"'] & \prod_{j \in J_{k_0}} C_j \arrow[r, "p_i"'] & C_i
\end{tikzcd}$$

Since $I = \bigcup_{k \in K} J_k$, the diagram commutes for any $i \in I$—thus $\prod_{k \in K} \left(\prod_{j \in J_k} C_j \right)$ is a product of the collection $(C_i)_{i \in I}$. Now, Proposition 2.1.2 finishes the proof. ♮

2.2 Coproducts

Definition 2.2.1 (Coproduct)**.** Let J be a set, and $(C_j)_{j \in I}$ be a collection of objects in a category C.

A *coproduct* of such collection, *if it exists*, is a pair $(P, (\iota_j)_{j \in J})$—where P is an object of C, and for every $j \in J$, $\iota_j \colon P \to C_j$ is a morphism of C.

Furthermore, this pair has to satisfy the following *universal property*: for every pair $(Q, (q_j)_{j \in J})$—where $Q \in \mathsf{C}$ and for every $j \in J$, $q_j \colon C_j \to Q$ is a morphism in C—there exists a *unique morphism* $f \colon P \to Q$ of C such that, for every $j \in J$ the following diagram commutes

$$\begin{tikzcd} P \arrow[d, dashed, "f"'] & C_j \arrow[l, "\iota_j"'] \arrow[dl, bend left, "q_j"] \\ Q & \end{tikzcd}$$

We can see right away that coproducts are dual to products, yielding the following two dual properties.

Proposition 2.2.2 (Uniqueness)**.** If a coproduct of a collection of objects of a category exists, then this coproduct is *unique up to isomorphism*.

Proposition 2.2.3 (Coproducts are independent of ordering)**.** Let I be a set and $(J_k)_{k\in K}$ be a partition of I by disjoint subsets $J_k \subseteq I$. Let $(C_i)_{i\in I}$ be a collection of objects in a category C. Then, if all the coproducts presented below exist in C, they are isomorphic:

$$\coprod_{i\in I} C_i \simeq \coprod_{k\in K}\Bigg(\coprod_{j\in J_k} C_j\Bigg).$$

Remark 2.2.4 (Duality)**.** Coproducts are *dual* to products in the following sense: if P is a product in a category C, then P^{op} is a coproduct in the opposite category C^{op}. Notice however that one *cannot* simply reverses the arrows of a product in C and end up with a coproduct in C—in fact, there are categories where product do exist, while coproducts do not.

2.3 Equalizers and Coequalizers

Definition 2.3.1 (Equalizer)**.** Let $f, g\colon A \rightrightarrows B$ be parallel morphisms in a category C. An *equalizer of f and g* is a pair (K, k)—where $K \in \mathsf{C}$ is an object and $k\colon K \to A$ is a morphism of C for which $fk = gk$—such that, for any object $M \in \mathsf{C}$ and morphism $m\colon M \to A$ of C satisfying $fm = gm$, there exists a *unique* morphism $n\colon M \to K$ such that the following diagram commutes

$$\begin{tikzcd} M \arrow[d, dashed, "n"'] \arrow[dr, bend left, "m"] & & \\ K \arrow[r, "k"'] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \end{tikzcd}$$

Just as with products and coproducts, one has a dual notion of an equalizer, we call it a *coequalizer*. For the sake of later reference, we write down its definition.

Definition 2.3.2 (Coequalizer)**.** Let $f, g\colon A \rightrightarrows B$ be parallel morphisms in a category C. An *equalizer of f and g* is a pair (C, c)—where $C \in \mathsf{C}$ is an object and $c\colon B \to C$ is a morphism of C for which $cf = cg$—such that, for any object $M \in \mathsf{C}$ and morphism

$m\colon B \to M$ of C satisfying $mf = mg$, there exists a *unique* morphism $n\colon C \to M$ such that the following diagram commutes

$$\begin{tikzcd} C \arrow[d, dashed, "n"] & B \arrow[l, "c"'] \arrow[dl, bend left, "m"] & A \arrow[l, shift right, "f"'] \arrow[l, shift left, "g"] \\ M & & \end{tikzcd}$$

Proposition 2.3.3 (Uniqueness)**.** Given two parallel morphisms $f, g\colon A \rightrightarrows B$ in a category C, if the (co)equalizer of them exists, then it is *unique up to isomorphism*. We denote *the* equalizer of f and g by $\mathrm{eq}(f, g)$ and the coequalizer of f and g by $\mathrm{coeq}(f, g)$.

Proof. We prove for equalizers. Let (K, k) and (K', k') be equalizers of f and g. If we apply the equalizer property of (K', k') in (K, k) and vice versa, we obtain unique morphisms $n'\colon K \to K'$ and $n\colon K' \to K$ such that $k = k'n'$ and $k' = kn$.

Moreover, applying the equalizer property of (K, k) in itself we obtain a unique $t\colon K \to K$ such that $k = kt$—therefore $t = \mathrm{id}_K$. Since $k = k'n = k\,\mathrm{id}_K$ and id_K is unique with such property, it follows that, since $k = k'n' = (kn)n' = k(nn')$ we find that $nn' = \mathrm{id}_K$. On the other hand, doing the same for (K', k') implies in $n'n = \mathrm{id}_{K'}$. Therefore $K \simeq K'$ in C, via n and n'. ♮

Corollary 2.3.4. Given a morphism $f\colon A \to B$ in a category C, the (co)equalizer of f with itself always exists—in fact $\mathrm{eq}(f, f) = (A, \mathrm{id}_A)$ and $\mathrm{coeq}(f, f) = (B, \mathrm{id}_B)$.

Proposition 2.3.5. Let C be a category and $f, g\colon A \rightrightarrows B$ be parallel morphisms in C. The following are properties concerning equalizers and coequalizers:

1. If f and g have an equalizer $\mathrm{eq}(f, g)$ in C, the morphism $\mathrm{eq}(f, g) \rightarrowtail A$ associated to the equalizer is a *monomorphism*.
2. If f and g have an coequalizer $\mathrm{coeq}(f, g)$ in C, the morphism $B \twoheadrightarrow \mathrm{coeq}(f, g)$ associated to the coequalizer is an *epimorphism*.

Proof. The properties are dual to each other, thus we may simply prove the one concerning equalizers. Let $k\colon \mathrm{eq}(f, g) \to A$ be the morphism in C associated to the equalizer. Consider any two parallel morphisms $x, y\colon C \rightrightarrows \mathrm{eq}(f, g)$ such that $kx = ky$—therefore, the following diagram commutes

$$\begin{tikzcd} C \arrow[r, shift left, "x"] \arrow[r, shift right, "y"'] & \mathrm{eq}(f, g) \arrow[r, "k"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \end{tikzcd}$$

From the diagram we obtain the relation $f(kx) = g(kx)$. Since $kx\colon C \to A$, the universal property implies that x must be the unique—hence $y = x$. ♮

Proposition 2.3.6. Let $f\colon A \to B$ be a morphism of a category C. If f is both an *epimorphism and equalizer*[1], then f is an *isomorphism*.

[1]This is clearly an abuse of language, the equalizer is in fact A and f is the morphism associated with A.

Proof. Let $x, y\colon B \rightrightarrows C$ be parallel morphisms in C with $\mathrm{eq}(x, y) = f$, then

$$\begin{tikzcd} A \arrow[r, "f"] & B \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & C \end{tikzcd}$$

which implies in $xf = yf$. By hypothesis f is an epimorphism, thus $x = y$. Notice however that the equalizer of a morphism $B \to C$ with itself is the identity on B. Since equalizers are unique up to isomorphism, it follows that $f \simeq \mathrm{id}_B$—thus f itself is an isomorphism (and $A \simeq B$). ♮

2.4 Pullbacks & Pushouts

Definition 2.4.1 (Pullback)**.** Let $f\colon A \to C$ and $g\colon B \to C$ be any two morphisms in a category C. A *pullback of* (f, g) is a triple (P, f', g') where the diagram

$$\begin{tikzcd} P \arrow[r, "f'"] \arrow[d, "g'"'] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$

commutes in C, and for every other triple (Q, f'', g'') making the diagram

$$\begin{tikzcd} Q \arrow[r, "f''"] \arrow[d, "g''"'] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$

commute in C, there exists a unique morphism $\ell\colon Q \to P$ in C such that the following diagram commutes

$$\begin{tikzcd} Q \arrow[dr, dashed, "\ell"] \arrow[drr, bend left, "f''"] \arrow[ddr, bend right, "g''"'] & & \\ & P \arrow[r, "f'"] \arrow[d, "g'"'] & B \arrow[d, "g"] \\ & A \arrow[r, "f"'] & C \end{tikzcd}$$

We usually denote that a square is a pullback in C by marking it as follows

$$\begin{tikzcd} P \arrow[r, "f'"] \arrow[d, "g'"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, "g"] \\ A \arrow[r, "f"'] & C \end{tikzcd}$$

The *dual* notion of a pullback is that of a pushout, just as before, we'll write it down just for the sake of later reference.

Definition 2.4.2 (Pushout)**.** Let $f\colon C \to A$ and $g\colon C \to B$ be any two morphisms in a category C. A *pushout of* (f, g) is a triple (P, f', g') where the diagram

$$\begin{tikzcd} C \arrow[r, "f"] \arrow[d, "g"'] & A \arrow[d, "g'"] \\ B \arrow[r, "f'"'] & P \end{tikzcd}$$

commutes in C, and for every other triple (Q, f'', g'') making the diagram

$$\begin{tikzcd} C \arrow[r, "f"] \arrow[d, "g"'] & A \arrow[d, "g''"] \\ B \arrow[r, "f''"'] & Q \end{tikzcd}$$

commute in C, there exists a unique morphism $\ell\colon P \to Q$ in C such that the following diagram commutes

$$\begin{tikzcd} C \arrow[r, "f"] \arrow[d, "g"'] & A \arrow[d, "g'"] \arrow[ddr, bend left, "g''"] & \\ B \arrow[r, "f'"'] \arrow[drr, bend right, "f''"'] & P \arrow[dr, dashed, "\ell"] & \\ & & Q \end{tikzcd}$$

Analogously, if we want to visually say that a square is a pushout in C, we mark it as follows

$$\begin{tikzcd} C \arrow[r, "f"] \arrow[d, "g"'] & A \arrow[d, "g'"] \\ B \arrow[r, "f'"'] \arrow[ur, phantom, "\ulcorner", very near end] & P \end{tikzcd}$$

Proposition 2.4.3 (Uniqueness)**.** The pullback (or pushout) of two morphisms, if existent, is *unique up to isomorphism*.

Proof. Let $f\colon A \to C$ and $g\colon B \to C$ be two morphisms in a category C. Suppose there exists two pullbacks of (f, g), namely (P, f', g'') and (Q, f'', g''). Let $\phi\colon P \to Q$ and $\psi\colon Q \to P$ be the uniquely defined morphisms given by the universal property of the pullback. Consider the following two commutative diagrams

$$
\begin{tikzcd}
P \arrow[dr, dashed, "\phi"] \arrow[ddrrr, bend left, "f'"] \arrow[dddrr, bend right, "g'"'] & & & \\
& Q \arrow[dr, dashed, "\psi"] \arrow[drr, bend left, "f''"] \arrow[ddr, bend right, "g''"'] & & \\
& & P \arrow[r, "f'"] \arrow[d, "g'"'] & B \arrow[d, "g"] \\
& & A \arrow[r, "f"'] & C
\end{tikzcd}
\qquad
\begin{tikzcd}
Q \arrow[dr, dashed, "\psi"] \arrow[ddrrr, bend left, "f''"] \arrow[dddrr, bend right, "g''"'] & & & \\
& P \arrow[dr, dashed, "\phi"] \arrow[drr, bend left, "f'"] \arrow[ddr, bend right, "g'"'] & & \\
& & Q \arrow[r, "f''"] \arrow[d, "g''"'] & B \arrow[d, "g"] \\
& & A \arrow[r, "f"'] & C
\end{tikzcd}
$$

Notice, however, that since $\mathrm{id}_P : P \to P$ and $\mathrm{id}_Q : Q \to Q$ also make them commute—respectively, the left and right diagrams. From uniqueness we obtain $\phi\psi = \mathrm{id}_P$ and $\psi\phi = \mathrm{id}_Q$—therefore $P \simeq Q$ in C, via ϕ and ψ. ♮

Proposition 2.4.4. Let (P, f', g') be the pullback of a pair of morphisms (f, g) in a category C, then:

(a) If g is a monomorphism, then g' is also a monomorphism.

(b) If g is an isomorphism, then g' is also an isomorphism.

Dually, if (B, f'', g'') is the pushout of the pair (f, g), then:

(c) If g is an epimorphism, then g'' is also an epimorphism.

(d) If g is an isomorphism, then g'' is also an isomorphism.

Proof. We only prove items (a) and (b), since (c) and (d) are merely dual consequences of the former items.

(a) Let g be a monomorphism and consider parallel morphisms $u, v : Q \rightrightarrows P$ such that $g'u = g'v$—we want to prove that $u = v$. Take into account the following commutative diagram

$$
\begin{tikzcd}
Q \arrow[dr, shift left, "u"] \arrow[dr, shift right, "v"'] & & \\
& P \arrow[r, "f'"] \arrow[d, "g'"'] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, "g"] \\
& A \arrow[r, "f"'] & C
\end{tikzcd}
$$

Lets first consider the morphism u. Notice that $p := g'u$ and $q := f'u$ are such that $fp = gq$. From the universal property of the pullback we have that u is the unique morphism factorizing (p, q) through (g, f).

On the other hand, if we consider the arrow v, one can define $p' := g'v$ and $q' := f'v$ so that $fp' = gq'$—thus v is the unique factorization of (p', q') through (g, f). Notice, however, that from construction $p' = g'v = g'u = p$. Moreover,

$$gq' = fp' = f(g'v) = f(g'u) = fp = gq,$$

since g is a monomorphism, it follows that $gq' = gq$ implies $q' = q$. Therefore both u and v are factorizations of the same pair of morphisms—and from uniqueness, it can only be the case that $u = v$.

(b) Suppose g is an isomorphism and consider the following commutative diagram

$$\begin{tikzcd}
A \arrow[dr, dashed, "\ell"] \arrow[drr, bend left, "g^{-1}f"] \arrow[ddr, bend right, "\mathrm{id}_A"'] & & \\
& P \arrow[r, "f'"] \arrow[d, "g'"'] & B \arrow[d, "g"] \\
& A \arrow[r, "f"'] & C
\end{tikzcd}$$

Thus $g'\ell = \mathrm{id}_A$ is already given. We can now consider the diagram

$$\begin{tikzcd}
P \arrow[dr, dashed, "\ell g'"] \arrow[drr, bend left, "f'(\ell g')"] \arrow[ddr, bend right, "g'(\ell g')"'] & & \\
& P \arrow[r, "f'"] \arrow[d, "g'"'] & B \arrow[d, "g"] \\
& A \arrow[r, "f"'] & C
\end{tikzcd} \tag{2.1}$$

Moreover, we consider the following composition

$$g'(\ell g') = (g'\ell)g' = \mathrm{id}_A g' = g' = g' \mathrm{id}_P,$$

on the other hand we have

$$f'(\ell g') = (f'\ell)g' = (g^{-1}f)g' = g^{-1}(fg') = g^{-1}(gf') = f' = f' \mathrm{id}_P .$$

Therefore $g'(\ell g') = g' \mathrm{id}_P$ and $f'(\ell g') = f' \mathrm{id}_P$ but by uniqueness, since id_P also makes Eq. (2.1) commute, it follows that $\ell g' = \mathrm{id}_P$. Thus g' is an isomorphism with inverse ℓ.

♮

Definition 2.4.5 (Kernel & cokernel)**.** Let $f\colon A \to B$ be a morphism in a category C. The *(co)kernel* of f, if existent, is defined to be the pullback of f with itself—or pushout in the case of cokernels.

Proposition 2.4.6. Let C be a category and $f: A \to B$ be a morphism of C. If the (co)kernel of f exists, its associated morphisms are both epimorphisms (monomorphisms for the case of cokernels).

Proof. Let $\ker f := (K, \alpha, \beta)$ and consider A itself, together with the identity morphisms. Since K is a pullback of f with itself, there exists a unique morphism $\gamma: A \to K$ such that $\beta\gamma = \mathrm{id}_A = \alpha\gamma$—therefore γ is a split monomorphism and both α and β are split epimorphisms. ♮

Proposition 2.4.7. Let $f: A \to B$ be a morphism in a category C. Then the following properties are equivalent:

(a) The morphism f is monic (conversely, epic).

(b) The kernel of f exists, furthermore $\ker f = (A, \mathrm{id}_A, \mathrm{id}_A)$ (conversely we have $\operatorname{coker} f = (B, \mathrm{id}_B, \mathrm{id}_B)$).

(c) The (co)kernel (K, α, β) of f exists, and $\alpha = \beta$.

Proof.
- (a) $\Rightarrow$ (b): Since f is monic, if P is any object together with morphisms $\phi, \psi: P \rightrightarrows A$ such that $f\phi = f\psi$ then $\phi = \psi$ and we may simply take A to be the kernel of f together with the identity morphisms—while the unique morphism $P \to A$ is given by $\phi = \psi$.
- (b) $\Rightarrow$ (c): If (K, α, β) is a kernel for f then in particular there exists an isomorphism $\phi: A \to K$ and therefore $\alpha\phi = \mathrm{id}_A$ and $\beta\phi = \mathrm{id}_A$. Since ϕ is epic, it follows that $\alpha = \beta$.
- (c) $\Rightarrow$ (a): Suppose (K, α, α) is kernel for f, then for all objects P and morphisms $g, h: P \rightrightarrows A$ such that$fg = fh$, we have a unique $\phi: P \to K$ such that $\alpha\phi = g$ and $\alpha\phi = h$—thus $g = h$.

♮

Proposition 2.4.8 (Coequalizers & kernels)**.** Let C be a category. The following properties relate coequalizers and kernels:

(a) Consider parallel morphisms $x, y: X \rightrightarrows A$. If the coequalizer $f = \operatorname{coeq}(x, y)$ exists and has a kernel pair $\ker f = (\alpha, \beta)$, then $\operatorname{coeq}(x, y)$ is the coequalizer of $\ker f$.

(b) Consider a morphism $h: A \to B$. If the kernel $\ker h = (\varepsilon, \delta)$ exists and has a coequalizer $w = \operatorname{coeq}(\varepsilon, \delta)$, then $\ker h$ is the kernel of $\operatorname{coeq}(\varepsilon, \delta)$.

Proof. (a) Since f is the coequalizer of x and y, it cleartly satisfies $fx = fy$. Therefore, the triple (X, x, y) can be used to apply the universal property of the pullback of $\ker f$ to obtain a unique factorization $\phi: X \to \ker f$ such that $x = \alpha\phi$ and $y = \beta\phi$.

Notice that $\alpha, \beta: \ker f \rightrightarrows A$ are parallel morphisms, so one can be tempted to find its coequalizer—we shall prove that $\operatorname{coeq}(\alpha, \beta) = \operatorname{coeq}(x, y)$. Let C be an object together with a morphism $g: A \to C$ such that $g\alpha = g\beta$. We may precompose this morphism with ϕ, obtaining $g\alpha\phi = g\beta\phi$, but using the result fro the last paragraph we conclude that $gx = gy$. Using the universal property for the coequalizer of

x and y, we find a unique factorization morphism $\psi\colon \operatorname{coeq}(x,y) \to C$ such that $g = \psi f$. Therefore, $\operatorname{coeq}(x,y)$ satisfies the universal property for the coequalizer of the kernel pair (α, β) of f. The whole construction of this item's proof can be seen in the following commutative diagram:

$$\begin{tikzcd}
X \arrow[drr, bend left, "x"] \arrow[dr, dashed, "\phi"] \arrow[ddr, bend right, "y"'] & & & \\
& \ker f \arrow[r, "\alpha"] \arrow[d, "\beta"'] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[d, "f"] \arrow[ddr, bend left, "g"] & \\
& A \arrow[r, "f"'] \arrow[drr, bend right, "g"'] & \operatorname{coeq}(x,y) \arrow[dr, dashed, "\psi"] & \\
& & & C
\end{tikzcd}$$

(b) From the pullback definition, we know that $h\varepsilon = h\delta$—since w is the coequalizer of (ε, δ), it follows that there exists a unique morphism $\gamma\colon B \to \operatorname{coeq}(\varepsilon, \delta)$ such that $\gamma w = h$.

Consider now two parallel morphisms $a, b\colon Y \rightrightarrows A$ such that $wa = wb$. Then we have that $ha = (\gamma w)a$, and $hb = (\gamma w)b$ since $wb = wa$ by assumption, it follows that $ha = hb$. From the kernel property, there must exist a unique morphism $\eta\colon Y \to \ker h$ such that $\varepsilon\eta = a$ and $\delta\eta = b$. Therefore $\ker w = \ker h$. All constructions can be visualized in the following commutative diagram:

$$\begin{tikzcd}
Y \arrow[drr, bend left, "a"] \arrow[dr, dashed, "\eta"] \arrow[ddr, bend right, "b"'] & & & \\
& \ker h \arrow[r, "\varepsilon"] \arrow[d, "\delta"'] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[d, "h"] \arrow[ddr, bend left, "w"] & \\
& A \arrow[r, "h"'] \arrow[drr, bend right, "w"'] & B & \\
& & & \operatorname{coeq}(\varepsilon, \delta) \arrow[ul, dashed, "\gamma"']
\end{tikzcd}$$

♮

Proposition 2.4.9 (Associativity property)**.** Let C be a category and consider the following commutative diagram in C:

$$\begin{tikzcd}
A \arrow[r, "a"] \arrow[d, "c"'] & B \arrow[r, "b"] \arrow[d, "d"'] & C \arrow[d, "e"] \\
D \arrow[r, "f"'] & E \arrow[r, "g"'] & F
\end{tikzcd}$$

The following are properties which regard pullbacks of such commutative squares:

(a) If both squares are pullbacks, then the outer-square

$$\begin{tikzcd} A \arrow[r, "ba"] \arrow[d, "c"'] \arrow[dr, phantom, "\lrcorner", very near start] & C \arrow[d, "e"] \\ D \arrow[r, "gf"'] & F \end{tikzcd} \tag{2.2}$$

is a pullback.

(b) If the second square is a pullback and the outer diagram Eq. (2.2) is a pullback, then the first square is a pullback.

Proof. (a) Let $x\colon Z \to D$ and $y\colon Z \to C$ be two morphisms of $\mathtt{C}$ such that $gfx = ey$. From the pullback property of the second square, there exists a unique morphism $z\colon Z \to B$ such that $bz = y$ and $dz = fx$. Using the pullback property of the first square on the morphisms z and x we find a unique $w\colon Z \to A$ such that $aw = z$ and $cw = x$.

Notice that since $aw = z$ then $b(aw) = bz = y$—we'll show that w is the unique morphism of $\mathtt{C}$ such that $(ba)w = y$ and $cw = x$. Suppose the existence of another morphism $w'\colon Z \to A$ such that $(ba)w' = y$ and $cw' = x$. In particular, it follows that $b(aw') = b(aw)$, and

$$d(aw') = f(cw') = fx = f(cw) = d(aw).$$

By the uniqueness of the morphism $Z \to B$ from the second pullback square diagram, $aw' = z = aw$. Using this last equality and the fact that $cw' = x = cw$, by the pullback property of the first square we obtain $w' = w$—which finally settles that w is unique and the outer-square is indeed a pullback.

(b) We assume the existence of a pullback (A', c', a') of (f, d) and show that A' must be isomorphic to A. From the commutativity of the diagram, since $da = cf$, we use the pullback property of (A', c', a') to get a unique morphism $h\colon A \to A'$—we'll show that h is the required isomorphism. Via the result of item (a), we know that (A', c, ba') is a pullback for the outer-square, therefore it follows that $c'h = c$ and $ba'h = ba$. Hence from the last two equalities one concludes that h is a factorization between two pullbacks of the outer-square. Since h is unique and pullbacks are unique up to isomorphism, it must be the case that h is an isomorphism—thus $A \simeq A'$, then (A, c, a) is a pullback for the first square.

♮

2.5 Limits & Colimits

(Co)Cones & (Co)Limits

Definition 2.5.1 (Cone)**.** Let $F\colon \mathtt{D} \to \mathtt{C}$ be a functor. We define a *cone* on F to consist of the following data:

- An object $C \in \mathtt{C}$.
- For each object $D \in \mathtt{D}$, a corresponding morphism $p_D\colon C \to FD$ in $\mathtt{C}$. Moreover, for every morphism $d\colon D \to D'$ in $\mathtt{D}$, one has that $p_{D'} = Fd \circ p_D$.

Definition 2.5.2 (Limit of a functor)**.** Let $F\colon \mathtt{D} \to \mathtt{C}$ be a functor. We define a *limit* of F to be a *cone* $\lim F = (L, (p_D)_{D\in\mathtt{D}})$ with the property that, for any cone $(M, (q_D)_{D\in\mathtt{D}})$ on F, there exists a *unique* morphism $m\colon M \to L$ of $\mathtt{C}$ such that, for every $D \in \mathtt{D}$, the following diagram commutes

$$\begin{tikzcd} M \arrow[d, dashed, "m"'] \arrow[dr, bend left, "q_D"] & \\ \lim F \arrow[r, "p_D"'] & FD \end{tikzcd}$$

Proposition 2.5.3 (Uniqueness)**.** The limit of a functor, when existent, is unique up to isomorphism.

Proof. Let $F\colon \mathtt{D} \to \mathtt{C}$ be a functor admitting limit cones $(L, (p_D)_{D\in\mathtt{D}})$ and $(L', (p'_D)_{D\in\mathtt{D}})$. Then there exists unique morphisms $m\colon L' \to L$ and $m\colon L \to L'$ such that $p_D m = p'_D$ and $m' p'_D = p_D$. Stacking these two we find that $mm'\colon L \to L$ is such that $p_D(mm') = p_D$, but since id_L has the same property, by the uniqueness of the factoring morphism, it follows that $mm' = \mathrm{id}_L$. The same analogous argument goes for $m'm = \mathrm{id}_{L'}$. We conclude that $L \simeq L'$ in $\mathtt{C}$ via m and m'. ♮

Proposition 2.5.4 (Parallel factorizations)**.** Let $F\colon \mathtt{D} \to \mathtt{C}$ be a functor admiting a limit, and let M be any object in $\mathtt{C}$. Two parallel morphisms $f, g\colon M \rightrightarrows \lim F$ in $\mathtt{C}$ are *equal* if, for every $D \in \mathtt{D}$, one has $p_D f = p_D g$.

Proof. Notice that $(M, (p_D f)_{D\in\mathtt{D}})$ forms a cone on F, therefore, from the universal property of the limit of F, we obtain $f = g$. ♮

Definition 2.5.5 (Cocone)**.** Let $F\colon \mathtt{D} \to \mathtt{C}$ be a functor. We define a *cocone* on F to consist of the following data:

- An object $C \in \mathtt{C}$.
- For each object $D \in \mathtt{D}$, a corresponding morphism $s_C\colon FD \to C$ of $\mathtt{C}$ such that, for every morphism $d\colon D' \to D$ in $\mathtt{D}$, we have that $s_{D'} = s_D \circ Fd$.

Definition 2.5.6 (Colimit)**.** Let $F\colon \mathtt{D} \to \mathtt{C}$ be a functor. We define a *colimit* on F to be a *cocone* $\operatorname{colim} F = (L, (s_D)_{D\in\mathtt{D}})$ such that, for every cocone $(M, (t_D)_{D\in\mathtt{D}})$ on F, there exists a *unique* morphism $m\colon L \to M$ such that, for every $D \in \mathtt{D}$, the following diagram commutes

$$\begin{tikzcd} M & FD \arrow[l, "t_D"'] \arrow[dl, bend left, "s_D"] \\ L \arrow[u, dashed, "m"] & \end{tikzcd}$$

Example 2.5.7 (Products)**.** Consider a set I as a discrete category (as in Example 1.2.10), and any category C. The limit of a functor $F\colon I \to \mathsf{C}$, if existent, is simply a *product* in C, that is

$$\lim F \simeq \prod_{i \in I} Fi.$$

Example 2.5.8 (Inverse & direct limit)**.** The limit of a functor $F\colon \omega^{\mathrm{op}} \to \mathsf{C}$ is commonly refered to as an *inverse limit*:

$$\begin{tikzcd}
\lim F \arrow[d] \arrow[dr, bend left] \arrow[drr, bend left] \arrow[drrr, bend left] & & & \\
\dots \arrow[r] & F(2) \arrow[r] & F(1) \arrow[r] & F(0)
\end{tikzcd}$$

The dual of this concept is that of the *direct limit*, which is the colimit of a functor $G\colon \omega \to \mathsf{C}$, visually given by:

$$\begin{tikzcd}
G(0) \arrow[r] & G(1) \arrow[r] & G(2) \arrow[r] & \dots \\
\operatorname{colim} G \arrow[u] \arrow[ur, bend right] \arrow[urr, bend right] \arrow[urrr, bend right] & & &
\end{tikzcd}$$

Complete Categories

Proposition 2.5.9 (All limits)**.** Let C be a $\mathcal{U}$-category. If for all $\mathcal{U}$-categories D and functors $F\colon \mathsf{D} \to \mathsf{C}$, the limit of F exists in C—that is, the category *admits all limits*—then C is a *preorder class*.

Proof. For C to be a preorder, there must exist at most one morphism between every pair of objects of C, so this is what we settle to do. Let $A, B \in \mathsf{C}$ be any two objects and suppose there exists a pair of *distinct* parallel morphisms $f, g\colon A \rightrightarrows B$. Since every limit exists, then in particular the product $B^{|\operatorname{Mor}(\mathsf{C})|}$ is a well defined object of C. From f and g, one can create $2^{|\operatorname{Mor}(\mathsf{C})|}$ distinct collections of morphisms $(h\colon A \to B)_{|\operatorname{Mor}(\mathsf{C})|}$—where each h is either f or g—which are cones over an "inclusion" functor $F\colon \mathsf{A} \to \mathsf{C}$, where A is composed of objects A and B, and morphisms are $f, g\colon A \rightrightarrows B$. The admittance of a limit over this functor is to state the existence of $2^{|\operatorname{Mor}(\mathsf{C})|}$ distinct factorizations of $A \to B^{|\operatorname{Mor}(\mathsf{C})|}$ in C. Since all of these factorizations are morphisms of C, it should be the case that $2^{|\operatorname{Mor}(\mathsf{C})|} < |\operatorname{Mor}(\mathsf{C})|$, which contradicts Cantor's theorem (see Theorem B.1.1) since C is a $\mathcal{U}$-category. If follows that there cannot exist more than one morphism between the objects of C, proving that it is a preorder. ♮

Definition 2.5.10 (Completeness)**.** We define the following notions concerning categories and the existence of limits:

(a) A category C is said to be *(co)complete* if, for every *small category* D, any functor $F\colon \mathsf{D} \to \mathsf{C}$ *has a (co)limit* in C.

(b) A category C is said to be *finitely (co)complete* if, for every *finite category* D, any functor $F\colon \mathsf{D} \to \mathsf{C}$ *has a (co)limit* in C.

Existence Theorem for Limits

Theorem 2.5.11. A category C is *complete* if and only if each collection of objects, indexed by a set, has a *product* and each pair of parallel morphisms has an *equalizer*.

Proof. Let C be complete and I be any set, which defines a small category. If $(C_i)_{i\in I}$ is any collection of objects in C, one can define a functor $F\colon I \to \mathsf{C}$ by $Fi := C_i$ and, since F has a limit by hypothesis, it follows that $\prod_{i\in I} Fi = \prod_{i\in I} C_i$ is a product in C. For the equalizer, let $f, g\colon A \rightrightarrows B$ be two parallel morphisms in C. Define a category D whose objects are A and B, and morphisms are the identities together with both f and g. Defining $F\colon \mathsf{D} \to \mathsf{C}$ to be simply an inclusion, since F has a limit, then there exists $\lim F$ such that

$$\lim F \longrightarrow A \underset{g}{\overset{f}{\rightrightarrows}} B$$

is an equalizer in C, which proves the first proposition.

For the second proposition, suppose that C admits products and equalizers. Let D be a small category and $F\colon \mathsf{D} \to \mathsf{C}$ be any functor. Define the following:

- Consider two pairs of products
$$\Big(\prod_{D\in\mathsf{D}} FD, (r_D)_{D\in\mathsf{D}}\Big) \quad \text{and} \quad \Big(\prod_{f\in\mathrm{Mor}(\mathsf{D})} F(\operatorname{cod} f), (q_f)_{f\in\mathrm{Mor}(\mathsf{D})}\Big),$$
where $r_{D'}\colon \prod_{D\in\mathsf{D}} FD \to FD'$ and $q_{f'}\colon \prod_{f\in\mathrm{Mor}(\mathsf{D})} F(\operatorname{cod} f) \to F(\operatorname{cod} f')$ are the projections associated with the products.
- Let $\alpha, \beta\colon \prod_{D\in\mathsf{D}} FD \rightrightarrows \prod_{f\in\mathrm{Mor}(\mathsf{D})} F(\operatorname{cod} f)$ be the *unique* factorizations such that
$$q_f\alpha = r_{\operatorname{cod} f} \quad \text{and} \quad q_f\beta = Ff \circ r_{\operatorname{dom} f}$$
for all $f \in \mathrm{Mor}(\mathsf{D})$.
- Since equalizers always exist on C, let $(L, \ell) := \mathrm{eq}(\alpha, \beta)$. Define a collection $(p_D)_{D\in\mathsf{D}}$ for which $p_D := r_D\ell$. We shall prove that $\lim F = (L, (p_D)_{D\in\mathsf{D}})$.

We first prove that $(L, (p_D)_{D\in\mathsf{D}})$ is a cone on F. To that end, consider any morphism $f\colon D \to D'$ in D. From the construction of the collection of morphisms, we have that

$$\begin{aligned}
Ff \circ p_D &= Ff \circ (r_D\ell) = (Ff \circ r_D)\ell \\
&= (q_f\beta)\ell = q_f(\beta\ell) \\
&= q_f(\alpha\ell) = (q_f\alpha)\ell \\
&= r_{D'}\ell \\
&= p_{D'}
\end{aligned}$$

therefore, $(p_D)_{D\in\mathtt{D}}$ satisfies the conditions of a cone on F.

Let $(M,(h_D)_{D\in\mathtt{D}})$ be a cone on F. From the product universal property, there exists a *unique* factorization $\gamma\colon M\to\prod_{D\in\mathtt{D}}FD$ such that $r_D\gamma=h_D$. On the other hand, for any morphism $f\colon D\to D'$ of $\mathtt{D}$ we have

$$(q_f\alpha)\gamma = r_{D'}\gamma = h_{D'} = Ff\circ h_D = Ff\circ(r_D\gamma) = (Ff\circ r_D)\gamma = (q_f\beta)\gamma$$

Since the factorization on the product $\prod_{f\in\mathrm{Mor}(\mathtt{C})}F(\operatorname{cod} f)$ is unique, it follows that $\alpha\gamma=\beta\gamma$. Since L is the object of the equalizer of $\mathrm{eq}(\gamma,\beta)$, there must exist a unique factorization $u\colon M\to L$ such that $\ell u=\gamma$. Therefore, one has

$$p_D u = (r_D\ell)u = r_D(\ell u) = r_D\gamma = h_D,$$

showing that u is a factorization of of the cone M via the cone L on F.

We must show that the factorization u is unique, so that L is the limit of F. Suppose $v\colon M\to L$ is another factorization, that is, $p_Dv=h_D$ for all $D\in\mathtt{D}$. Notice that

$$r_D(\ell u) = (r_D\ell)u = p_Du = h_D = p_Dv = (r_D\ell)v = r_D(\ell v)$$

holds for all $D\in\mathtt{D}$, therefore both ℓu and ℓv are parallel factorizations—which by Proposition 2.5.4 implies $\ell u=\ell v$. From Proposition 2.3.5 we know that ℓ is monic, therefore $u=v$. This shows the uniqueness of u, thus we may conclude that

$$\lim F = (L,(p_D)_{D\in\mathtt{D}}).$$

♮

Proposition 2.5.12 (Finite completeness)**.** Let $\mathtt{C}$ be a category. The following properties are equivalent:

(a) The category $\mathtt{C}$ is *finitely complete*.

(b) The category $\mathtt{C}$ has a *terminal object, binary products and equalizers*.

(c) The category $\mathtt{C}$ has a *terminal object, and pullbacks*.

Proof.
- If $\mathtt{C}$ is finitely complete, then from definition items (b) and (c) hold.
- If (b) holds, then by Proposition 2.1.3 we find that any *finite* collection has a product. Since $\mathtt{C}$ has equalizers, by Theorem 2.5.11 and taking the particular case where $\mathtt{D}$ is a finite category, we conclude that $\mathtt{C}$ is finitely complete, therefore (b) implies (a).
- If (c) holds, let $1\in\mathtt{C}$ be the terminal object. For any two objects $A,B\in\mathtt{C}$, the pullback

$$\begin{tikzcd} P \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & A \arrow[d, dashed] \\ B \arrow[r, dashed] & 1 \end{tikzcd}$$

is the product of A with B in $\mathtt{C}$—thus binary products exist. Now consider two parallel morphisms $f, g \colon A \rightrightarrows B$ in $\mathtt{C}$. In this case, one can consider the pullback

$$\begin{tikzcd} E \arrow[r, "k"] \arrow[d, "\ell"'] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[d, "\mathrm{id}_A \times f"] \\ A \arrow[r, "\mathrm{id}_A \times f"'] & A \times B \end{tikzcd}$$

Then for E to be the equalizer of (f, g) it suffices to show that $k = \ell$ and $fk = gk$. For the former, notice that

$$k = \pi_A \circ (\mathrm{id}_A \times f) \circ k = \pi_A \circ (\mathrm{id}_A \times g) \circ \ell = \ell.$$

For the last equality, we proceed with a similar argument

$$fk = \pi_B \circ (\mathrm{id}_A \times f) \circ k = \pi_B \circ (\mathrm{id}_A \times g) \circ \ell g\ell = gk.$$

Therefore $\mathrm{eq}(f, g) = (E, k)$, and $\mathtt{C}$ has equalizers. This proves that (c) implies (b). ♮

Proposition 2.5.13. Let $F \colon \mathtt{D} \to \mathtt{C}$ be a functor, and suppose there exists a collection $(f_j)_{j \in J}$ of morphisms of $\mathtt{D}$ *generating* every other morphism of $\mathtt{D}$—that is, if g is a morphism in $\mathtt{D}$, then g can be written as the composition of *finitely* many arrows of $(f_j)_{j \in J}$. A *cone* on the functor F is a pair $(L, (p_D \colon L \to FD)_{D \in \mathtt{D}})$ such that, for any morphism $f_j \colon D \to D'$, we have $p_{D'} = Ff_j \circ p_D$.

Proof. To see that this is equivalent to the definition of a cone is simple since, given any $g \colon D \to \mathtt{D}'$ in $\mathtt{D}$, one has $g = f_{j_n} \dots f_{j_1}$ for a finite collection of maps $f_{j_i} \in (f_j)_{j \in J}$ where $f_{j_i} \colon D_{i-1} \to D_i$ with $D_0 = D$ and $D_n = D'$. Then

$$\begin{aligned} Fg \circ p_D &= F(f_{j_n} \dots f_{j_1}) \circ p_D \\ &= (Ff_{j_n} \circ \dots \circ Ff_{j_1}) \circ p_D \\ &= (Ff_{j_n} \circ \dots \circ Ff_{j_2}) \circ (Ff_{j_1} \circ p_D) \\ &= (Ff_{j_n} \circ \dots \circ Ff_{j_2}) \circ p_{D_1} \\ &= \dots \\ &= Ff_{j_n} \circ p_{D_{n-1}} \\ &= p_{D'}, \end{aligned}$$

which proves the equivalence. ♮

Definition 2.5.14 (Finitely generated categories)**.** A category $\mathtt{C}$ is said to be *finitely generated* if $\mathtt{C}$ is comprised of finitely many objects, and there exists a finite set of morphisms $\{f_1, \dots, f_n\}$ of $\mathtt{C}$ such that every morphism of $\mathtt{C}$ can be written as a composition of finitely many arrows f_j.

Proposition 2.5.15. Let $\mathtt{D}$ be *finitely generated* and $\mathtt{C}$ be *finitely complete*, then *any* functor $F \colon \mathtt{D} \to \mathtt{C}$ *has a limit*.

Proof. Since $\mathtt{C}$ is finitely complete, we can recycle the proof of Theorem 2.5.11 replacing the pair of products with the *finite* products $\prod_{D \in \mathtt{D}} FD$ and $\prod_{j=1}^{n} F(\operatorname{cod} f_j)$. ♮

Limit Preserving Functors

Definition 2.5.16 (Limit preserving functor)**.** A functor $F\colon \mathtt{B} \to \mathtt{C}$ is said to *preserve limits* if—for every small category $\mathtt{A}$ and functor $G\colon \mathtt{A} \to \mathtt{B}$—the limit of G *exists* and the limit of $FG\colon \mathtt{A} \to \mathtt{C}$ exists and is given by

$$\lim(FG) = F(\lim G).$$

To put more concretely, if $(L, (p_A)_{A\in\mathtt{A}})$ is the limit of G, then the limit of FG is given by $(FL, (Fp_A)_{A\in\mathtt{A}})$.

Proposition 2.5.17. Let $\mathtt{D}$ be a (finitely) complete category and $\mathtt{C}$ be *any* category. A functor $F\colon \mathtt{D} \to \mathtt{C}$ preserves (finite) limits if and only if it preserves (finite) products and equalizers.

Proof. Suppose F preserves (finite) limits, then in particular it preserves both (finite) products and equalizers. Now, we assume that F preserves both (finite) products and equalizers. Since $\mathtt{D}$ is (finitely) complete, for any small (finite) category $\mathtt{A}$ and functor $G\colon \mathtt{A} \to \mathtt{D}$, there exists $\lim G$ in $\mathtt{D}$. Moreover, since $\lim G$ can be build out of products and equalizers (see Theorem 2.5.11), it follows that, since F preserves products and equalizers, it must also preserve the limit of G—and $\lim(FG) = F(\lim G)$. ♮

Proposition 2.5.18. A functor that preserves *pullbacks* does preserve *monomorphisms*.

Proof. Let $F\colon \mathtt{B} \to \mathtt{C}$ be a functor preserving pullbacks, and consider any monomorphism $f\colon B \to B'$ in $\mathtt{B}$. From Proposition 2.4.7 we find that $\ker f = (B, \mathrm{id}_B, \mathrm{id}_B)$. Define $\mathtt{A}$ to be a category whose objects are B and B', and morphisms are identities together with $f\colon B \to B'$. The limit of the inclusion functor $I\colon \mathtt{A} \to \mathtt{D}$ is simply the pullback of f with itself, namely, $\ker f$. Since F preserves pullbacks, then $FI\colon \mathtt{A} \to \mathtt{C}$ has a limit $\lim(FI) = F\ker f$, which is the the pullback of $FIf = Ff\colon FB \to FB'$. This implies in $\ker(Ff) = (FB, \mathrm{id}_{FB}, \mathrm{id}_{FB})$, which is equivalent to Ff being a monomorphism in $\mathtt{C}$. Therefore F preserves monomorphisms. ♮

Proposition 2.5.19. Let $\mathtt{C}$ be a category and $C \in \mathtt{C}$ be any object. The covariant functor $\mathrm{Mor}_{\mathtt{C}}(C, -)\colon \mathtt{C} \to \mathtt{Set}$ *preserves all existing limits, including large ones*. In particular it *preserves monomorphisms*.

Proof. Let $F\colon \mathtt{D} \to \mathtt{C}$ be a functor admitting a limit $(L, (p_D)_{D\in\mathtt{D}})$. Our goal is to show that the functor $\mathrm{Mor}_{\mathtt{C}}(C, F(-))\colon \mathtt{D} \to \mathtt{Set}$—mapping $D \mapsto \mathrm{Mor}_{\mathtt{C}}(C, FD)$ and morphisms $(f\colon D \to D') \mapsto ((Ff)_*\colon \mathrm{Mor}_{\mathtt{C}}(C, FD) \to \mathrm{Mor}_{\mathtt{C}}(C, FD'))$—has a limit

$$\lim \mathrm{Mor}(C, F(-)) = (\mathrm{Mor}_{\mathtt{C}}(C, L), (\mathrm{Mor}_{\mathtt{C}}(C, p_D))_{D\in\mathtt{D}}).$$

Consider any cone $(q_D\colon S \to \mathrm{Mor}_{\mathtt{C}}(C, FD))_{D\in\mathtt{D}}$, on $\mathtt{Set}$, over the said composite functor. Notice that every element $s \in S$ induces a cone over F given by $(q_D(s)\colon C \to FD)_{D\in\mathtt{D}}$. Since F has a limit, there must exist a *unique* morphism $q_s\colon C \to L$ in $\mathtt{C}$ such that $p_D q_s = q_D(s)$ for all objects $D \in \mathtt{D}$. Since $s \in S$ was chosen arbitrarily, we can define a uniquely induced set-function $q\colon S \to \mathrm{Mor}_{\mathtt{C}}(C, L)$ given by $q(s) := q_s$. By the properties of q_s, the map q also satisfies $(Fp_D)_* \circ q = q_D$ for all $D \in \mathtt{D}$. ♮

Corollary 2.5.20. Given a category $\mathtt{C}$ and an object $C \in \mathtt{C}$, the contravariant functor $\mathrm{Mor}_{\mathtt{C}}(-, C)\colon \mathtt{C} \to \mathtt{Set}$ *transforms existing colimits into limits*. In particular, it transforms *epimorphisms into monomorphisms*.

Limit Reflection

Definition 2.5.21. A functor $F\colon \mathtt{B} \to \mathtt{C}$ is said to *reflect limits* if—for all small categories $\mathtt{A}$, and all functors $G\colon \mathtt{A} \to \mathtt{B}$, and every cone $(L, (g_A)_{A \in \mathtt{A}})$ over G—the limit of $FG\colon \mathtt{A} \to \mathtt{C}$ is given by $\lim(FG) = (FL, (Fg_A)_{\mathtt{A}})$, then $\lim G = (L, (g_A)_{A \in \mathtt{A}})$.

Proposition 2.5.22. Let $F\colon \mathtt{A} \to \mathtt{B}$ be a *limit preserving* functor. If $\mathtt{A}$ is *complete*, and F *reflects isomorphisms*, then F reflects limits.

Proof. Let $G\colon \mathtt{D} \to \mathtt{A}$ be a functor with $\lim G = (L, (g_D)_{D \in \mathtt{D}})$. Since F preserves limits, then

$$\lim(FG) = (FL, (Fg_D)_{D \in \mathtt{D}}).$$

Suppose that $(FM, (Ff_D)_{D \in \mathtt{D}})$ is also a limit of the composite FG. We want to show that it must be the case that $(L, (g_D)_{D \in \mathtt{D}}) \simeq (M, (f_D)_{D \in \mathtt{D}})$.

From the limit property in $\mathtt{A}$, there exists a unique factorization $u\colon M \to L$ such that for all $D \in \mathtt{D}$ we have $g_D u = f_D$. Notice that $Fu\colon FM \to FL$ is a factorization between limits of the composite FG. From the uniqueness of the factorization and since $FM \simeq FL$ in $\mathtt{B}$, it follows that Fu is an isomorphism. Since F reflects isomorphisms, it must be the case that u is an isomorphism and hence $M \simeq L$, showing that $(L, (g_D)_{D \in \mathtt{D}}) \simeq (M, (f_D)_{D \in \mathtt{D}})$ via u. ♮

Proposition 2.5.23. Let $F\colon \mathtt{A} \to \mathtt{B}$ be a functor between *finitely generated complete* categories $\mathtt{A}$ and $\mathtt{B}$. If the functor F *preserves* (or *reflects*) *finite limits*, then F preserves (or reflects) *finitely generated limits*.

Proof. By Proposition 2.5.15 we have that finitely generated limits can be expressed as finite products and equalizers. Therefore F preserves (or reflects) all existent finitely generated limits. ♮

Proposition 2.5.24. A *fully faithful* functor $F\colon \mathtt{C} \to \mathtt{D}$ *reflects limits*.

Proof. Let $G\colon \mathtt{E} \to \mathtt{C}$ be any functor from a small category $\mathtt{E}$. Since F is fully faithful, there exists a unique collection of morphisms $(g_E\colon L \to GE)_{E \in \mathtt{E}}$—which is a cone over G—such that the collection $(Fg_E\colon FL \to FGE)_{E \in \mathtt{E}}$ is a *limit* cone over FG. From the limit property, if $(f_E\colon M \to GE)_{E \in \mathtt{E}}$ is another cone over G, the corresponding cone $(Ff_E\colon FM \to FGE)_{E \in \mathtt{E}}$ over FG has a unique factorization $u\colon FM \to FL$ in $\mathtt{D}$ such that for all $E \in \mathtt{E}$ we have $Fg_E \circ u = Ff_E$. From the fully faithfulness of F, there exists a unique morphism $\ell\colon M \to L$ in $\mathtt{E}$ such that $F\ell = u$—which in this case satisfies $g_E \ell = f_E$, showing that ℓ is the unique factorization between the cones $(g_E)_{E \in \mathtt{E}}$ and $(f_E)_{E \in \mathtt{E}}$. This implies in $\lim G = (L, (g_E)_{E \in \mathtt{E}})$ as wanted. ♮

Absolute (Co)Limits

Definition 2.5.25 (Absolute (co)limit)**.** Let $G\colon \mathtt{A} \to \mathtt{B}$ be a functor admitting a colimit $\operatorname{colim} G$. We say that such colimit is *absolute* if, for every functor $F\colon \mathtt{B} \to \mathtt{C}$, we have

$$\operatorname{colim}(FG) = F(\operatorname{colim} G).$$

Proposition 2.5.26. Let $\mathtt{C}$ be a category and consider the following commutative diagram in $\mathtt{C}$:

$$\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \arrow[r, "q"] & C \end{tikzcd}$$

that is, $qf = qg$. If g and q are both *split epimorphisms* with *sections* s_g and s_q, respectively, and the diagram

$$\begin{tikzcd} B \arrow[r, "q"] \arrow[d, "s_g"'] & C \arrow[d, "s_q"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

commutes in $\mathtt{C}$—that is, $s_q q = f s_g$—then

$$\operatorname{coeq}(f, g) = (C, q).$$

Moreover, this coequalizer is *absolute*.

Proof. Let $p\colon B \to D$ be any morphism in $\mathtt{C}$ such that $pf = pg$. Define a morphism $\ell := ps_q\colon C \to D$, then

$$\ell q = (ps_q)q = p(s_q q) = p(f s_g) = (pf)s_g = (pg)s_g = p(gs_g) = p\,\mathrm{id}_A = p.$$

To show the uniqueness of ℓ, suppose $k\colon C \to D$ satisfies $kq = p$. Then we have

$$k = k(qs_q) = (kq)s_q = ps_q = \ell,$$

which proves that $\operatorname{coeq}(f, g) = (C, q)$. Since any functor preserves retracts and sections, it follows that the coequalizer is preserved under any given functor. ♮

Proposition 2.5.27 (Absolute pushout)**.** Let $\mathtt{C}$ be a category, and consider the following commutative square in $\mathtt{C}$:

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, two heads, "g"'] & B \arrow[d, two heads, "r"] \\ C \arrow[r, "h"'] & D \end{tikzcd}$$

If g and r are both *split epimorphisms* with corresponding *sections* α and β that are compatible with the square—meaning, $f\alpha = \beta h$— then the square is an *absolute pushout*.

Proof. We first show that the square is indeed a pushout. Let $X \in \mathsf{C}$ be any object and consider morphisms $p\colon B \to X$ and $q\colon C \to X$ such that $pf = qg$. Define a morphism $\ell\colon D \to X$ by $\ell := p\beta$. Note that

$$\ell h = (p\beta)h = p(\beta h) = p(f\alpha) = (pf)\alpha = (qg)\alpha = q(g\alpha) = q\,\mathrm{id}_A = q,$$
$$\ell r = (p\beta)r = p(\beta r) = p\,\mathrm{id}_B = p.$$

To show uniqueness of ℓ, suppose $k\colon D \to X$ is another morphism such that $kh = q$ and $kr = p$—then in particular $kr = \ell r$ but since r is an epimorphism it follows that $k = \ell$. With this we have proved that the square is indeed a pushout. Since split epimorphisms and commutative diagrams are preserved by any functor, it follows that the pushout is absolute. ♮

Chapter 3

Adjoint Functors

3.1 Reflections

Definition 3.1.1 (Reflection along a functor)**.** Let $F\colon \mathtt{C} \to \mathtt{D}$ be any functor and $D \in \mathtt{D}$. We define a *reflection of D along F* to be a pair (R_D, η_D) where $R_D \in \mathtt{C}$ is an object, and $\eta_D\colon D \to FR_D$ is a morphism of $\mathtt{D}$ such that: if $C \in \mathtt{C}$ is any object and $\delta\colon D \to FC$ is a morphism of $\mathtt{D}$, then there exists a *unique morphism* $\varepsilon\colon R_D \to C$ in $\mathtt{C}$ such that the diagram

$$\begin{tikzcd} D \arrow[rr, "\delta"] \arrow[dr, "\eta_D"'] & & FC \\ & FR_D \arrow[ur, dashed, "F\varepsilon"'] & \end{tikzcd}$$

commutes in the category $\mathtt{D}$.

Proposition 3.1.2 (Uniqueness of reflections)**.** Let $F\colon \mathtt{C} \to \mathtt{D}$ be a functor and $D \in \mathtt{D}$ be an object. If the reflection of D along F exists, then it is *unique up to isomorphism*.

Proof. Suppose the reflection of D along F exists and let (R, η) and (R', η') be two reflections of D. Considering morphisms $\delta\colon D \to FR$

Continue

$\natural$

Chapter 4

Monoidal Categories

4.1 Monoidal Categories

Definition 4.1.1. A *monoidal category* is a tuple $(\mathtt{M}, \otimes, 1, \alpha, \lambda, \rho)$ consisting of:

- A *category* $\mathtt{M}$.
- A bifunctor $\otimes\colon \mathtt{M} \times \mathtt{M} \longrightarrow \mathtt{M}$
- A distinguished object $1 \in \mathtt{M}$ that is *unitary* with respect to $\otimes$, that is:

$$m \otimes 1 = m = 1 \otimes m$$

  for any object $m \in \mathtt{M}$.
- A *natural isomorphism*

$$\alpha\colon (- \otimes (- \otimes -)) \xRightarrow{\cong} ((- \otimes -) \otimes -).$$

  called *associator*, in the sense that given any triple of objects (a, b, c) of $\mathtt{M}$, the image

$$a \otimes (b \otimes c) \xrightarrow[\alpha(a,b,c)]{\cong} (a \otimes b) \otimes c$$

  is an isomorphism in $\mathtt{M}$.
- Two *natural isomorphisms*

$$\lambda\colon (1 \otimes -) \xRightarrow{\cong} (-) \quad \text{and} \quad \rho\colon (- \otimes 1) \xRightarrow{\cong} (-)$$

  called *left and right unitors*, respectively. In other words, given any object $a \in \mathtt{M}$ the morphisms $\lambda a\colon 1 \otimes a \xrightarrow{\cong} a$ and $\rho a\colon a \otimes 1 \xrightarrow{\cong} a$ are *isomorphisms* in $\mathtt{M}$.

This data should satisfy the following two conditions:

- (Triangle identity) Given any pair (a, b) of objects in $\mathtt{M}$, the diagram

$$\begin{tikzcd} a \otimes (1 \otimes b) \arrow[rr, "\alpha(a,1,b)"] \arrow[dr, "\mathrm{id}_a \otimes \rho b"'] & & (a \otimes 1) \otimes b \arrow[dl, "\lambda a \otimes \mathrm{id}_b"] \\ & a \otimes b & \end{tikzcd}$$

  commutes in $\mathtt{M}$.

- (Pentagon identity) Given any tuple (a, b, c, d) of objects in $\mathtt{M}$, the diagram

$$\begin{tikzcd}
 & (a \otimes b) \otimes (c \otimes d) \arrow[dr, "\alpha(a\otimes b, c, d)"] & \\
a \otimes (b \otimes (c \otimes d)) \arrow[ur, "\alpha(a,b,c\otimes d)"] \arrow[d, "\mathrm{id}_a \otimes \alpha(b,c,d)"'] & & ((a \otimes b) \otimes c) \otimes d \\
a \otimes ((b \otimes c) \otimes d) \arrow[rr, "\alpha(a, b\otimes c, d)"'] & & (a \otimes (b \otimes c)) \otimes d \arrow[u, "\alpha(a,b,c)\otimes \mathrm{id}_d"']
\end{tikzcd}$$

is commutative in $\mathtt{M}$.

The tuple $(\mathtt{M}, \otimes, 1, \alpha, \lambda, \rho)$ is said to be a *strict monoidal category* if the three natural isomorphisms are naturally isomorphic to the identity—if this is the case, we shall refer to the category simply by the triple $(\mathtt{M}, \otimes, 1)$.

Definition 4.1.2 (Monoidal functor)**.** Let $(\mathtt{M}, \otimes, 1, \alpha, \lambda, \rho)$ and $(\mathtt{N}, \widehat{\otimes}, \widehat{1}, \widehat{\alpha}, \widehat{\lambda}, \widehat{\rho})$ be two (strict) monoidal categories. We say that a functor $F \colon \mathtt{M} \to \mathtt{N}$ is a (*strict*) *monoidal functor* if it preserves the actions of the natural isomorphisms. To put concretely, we have:

- The unit of $\mathtt{M}$ is mapped to the unit of $\mathtt{N}$, that is, $Fe = \widehat{e}$.
- For any $a \in \mathtt{M}$ one has $F(\lambda a) = \widehat{\lambda}(Fa)$ and $F(\rho a) = \widehat{\rho}(Fa)$.
- For any pair (a, b) of objects in $\mathtt{M}$ there exists an isomorphism $F(a \otimes b) \simeq Fa \widehat{\otimes} Fb$ in $\mathtt{N}$—in the strict case, the isomorphism is replaced by an equality.
- For any triple (a, b, c) of objects in $\mathtt{M}$ we have $F\alpha(a, b, c) = \widehat{\alpha}(Fa, Fb, Fc)$.
- For every two maps f and g in $\mathtt{M}$ there exists an isomorphism $F(f \otimes g) \simeq Ff \widehat{\otimes} Fg$ in $\mathtt{N}$—in the strict case, the isomorphism is replaced by an equality.

Definition 4.1.3 (Monoidal natural transformation)**.** Let $(\mathtt{M}, \otimes, 1, \alpha, \lambda, \rho)$ and $(\mathtt{N}, \widehat{\otimes}, \widehat{1}, \widehat{\alpha}, \widehat{\lambda}, \widehat{\rho})$ be two (strict) monoidal categories, and consider a pair of parallel (strict) functors $F, G \colon \mathtt{M} \rightrightarrows \mathtt{N}$. A natural transformation $\eta \colon F \Rightarrow G$ is said to be *monoidal* if $\eta_1 = \widehat{1}$, and for any pair of objects $a, b \in \mathtt{M}$ the diagram

$$\begin{tikzcd}
F(a \otimes b) \arrow[r, "\eta_{a\otimes b}"] \arrow[d, "\simeq"'] & G(a \otimes b) \arrow[d, "\simeq"] \\
Fa \widehat{\otimes} Fb \arrow[r, "\eta_a \widehat{\otimes} \eta_b"'] & Ga \widehat{\otimes} Gb
\end{tikzcd}$$

commutes in the monoidal category $\mathtt{N}$.

Theorem 4.1.4 (Strictification of monoidal categories)**.** Every monoidal category is *monoidally equivalent* to a *strict* monoidal category.

Proof. Let $(\mathtt{M}, \otimes, 1, \alpha, \lambda, \rho)$ be a monoidal category. We shall construct a strict monoidal category out of $\mathtt{M}$. To that end, define a category $\mathtt{N}$ where:

- The objects of $\mathtt{N}$ are pairs (F, η) where F is an *endofunctor* of $\mathtt{M}$ and

  $$\eta \colon F(- \otimes -) \overset{\cong}{\Longrightarrow} F(-) \otimes (-)$$

  is a *natural isomorphism* such that, for any triple (a, b, c) of objects of $\mathtt{M}$, the pentagonal diagram

  $$\begin{tikzcd}
  & (F(a) \otimes b) \otimes c & \\
  F(a \otimes b) \otimes c \arrow[ur, "\eta_{(a,b)} \otimes \mathrm{id}_c"] & & F(a) \otimes (b \otimes c) \arrow[ul, "\alpha(Fa,b,c)"'] \\
  F((a \otimes b) \otimes c) \arrow[u, "\eta_{(a \otimes b,c)}"] & & F(a \otimes (b \otimes c)) \arrow[ll, "F\alpha(a,b,c)"] \arrow[u, "\eta_{(a,b \otimes c)}"']
  \end{tikzcd}$$

- A morphism $\varepsilon \colon (F, \eta) \to (F', \eta')$ is a natural transformation $\varepsilon \colon F \Rightarrow F'$ such that, given any pair (a, b) of objects of $\mathtt{M}$, the diagram

  $$\begin{tikzcd}
  F(a \otimes b) \arrow[r, "\varepsilon_{a \otimes b}"] \arrow[d, "\eta_{(a,b)}"'] & F'(a \otimes b) \arrow[d, "\eta'_{(a,b)}"] \\
  F(a) \otimes b \arrow[r, "\varepsilon_a \otimes \mathrm{id}_b"'] & F'(a) \otimes b
  \end{tikzcd} \tag{4.1}$$

  commutes in $\mathtt{M}$. Moreover, we define the composition of morphisms in $\mathtt{N}$ to be given by the vertical composition of natural transformations.

- Define a bifunctor $\widehat{\otimes} \colon \mathtt{N} \times \mathtt{N} \to \mathtt{N}$ as $(F, \eta) \widehat{\otimes} (F', \eta') := (FF', \widehat{\eta})$, where

  $$\widehat{\eta} \colon FF'(- \otimes -) \Rightarrow FF'(-) \otimes (-)$$

  is the natural transformation given by the composition

  $$\begin{tikzcd}
  FF'(a \otimes b) \arrow[r, "F\eta'_{(a,b)}"'] \arrow[rr, bend left, "\widehat{\eta}_{(a,b)}"] & F(F'(a) \otimes b) \arrow[r, "\eta_{(F'a,b)}"'] & FF'(a) \otimes b
  \end{tikzcd}$$

  for any pair of objects (a, b) of $\mathtt{M}$.

From this construction we find that the triple $(\mathtt{N}, \widehat{\otimes}, (\mathrm{id}_{\mathtt{M}}, I))$—where the natural transformation $I \colon (- \otimes -) \overset{\cong}{\Longrightarrow} (- \otimes -)$ is the identity morphism $I_{(a,b)} := \mathrm{id}_{a \otimes b}$ in $\mathtt{M}$ for any two $a, b \in \mathtt{M}$—is a *strict monoidal category*, since:

- The bifunctor $\widehat{\otimes}$ satisfies *equality* for both left and right unitors: given an object $(F,\eta) \in \mathtt{N}$, consider any two objects $a,b \in \mathtt{N}$ then by the definition of $(F,\eta)\widehat{\otimes}(\mathrm{id}_{\mathtt{M}}, I) = (F,\widehat{\eta})$ and $(\mathrm{id}_{\mathtt{M}}, I)\widehat{\otimes}(F,\widehat{\eta}')$ one has

$$\begin{tikzcd}
F(a\otimes b) \arrow[r, "F\,\mathrm{id}_{a\otimes b}=\mathrm{id}_{F(a\otimes b)}"'] \arrow[rr, bend left, "\widehat{\eta}_{(a,b)}"] & F(a\otimes b) \arrow[r, "\eta_{(a,b)}"'] & F(a)\otimes b
\end{tikzcd}$$

$$\begin{tikzcd}
F(a\otimes b) \arrow[r, "\mathrm{id}_{\mathtt{M}}\,\eta_{(a,b)}=\eta_{(a,b)}"] \arrow[rr, bend right, "\widehat{\eta}'_{(a,b)}"'] & F(a)\otimes b \arrow[r, "I_{(Fa,b)}=\mathrm{id}_{F(a)\otimes b}"] & F(a)\otimes b
\end{tikzcd}$$

  therefore $\widehat{\eta} = \eta = \widehat{\eta}'$. Moreover, this also shows that the triangle identity is satisfied.

- Associativity follows from the associativity of morphisms and functors.

We now prove that $\mathtt{M}$ and $\mathtt{N}$ are equivalent categories. In order to do that, define a functor $E\colon \mathtt{M} \longrightarrow \mathtt{N}$ mapping objects $a \mapsto (a\otimes(-), \alpha(a,-,-))$ and morphisms $f \mapsto f\otimes(-)$. We now show that E is an equivalence of categories:

- (Essentially surjective) Notice that, given any object $(F,\eta) \in \mathtt{N}$, we can define a morphism

$$\varepsilon\colon (F1\otimes(-), \alpha(F1,-,-)) \longrightarrow (F,\eta)$$

  by constructing a natural transformation $\varepsilon\colon F1\otimes(-) \Rightarrow F$ where $\varepsilon_a := \lambda_a \eta^{-1}_{(1,a)}$, which is an isomorphism $F(1)\otimes a \simeq Fa$ for any $a \in \mathtt{M}$—showing that ε is a natural isomorphism, defining an isomorphism $E(F1) \simeq (F,\eta)$.

- (Full) Let $a,b \in \mathtt{M}$ be any two objects, and $\varepsilon\colon Ea \longrightarrow Eb$ be any morphism of $\mathtt{N}$—that is, a natural transformation $\varepsilon\colon (a\otimes -) \Rightarrow (b\otimes -)$ satisfying the coherence diagram Eq. (4.1). Define $f\colon a \longrightarrow b$ to be the morphism in $\mathtt{M}$ given by $f := (\lambda b)\circ\varepsilon_1\circ(\lambda^{-1}a)$. By the definition of E, one has $Ef = f\otimes(-)$—we wish to show that this agrees with ε. Given any $c \in \mathtt{M}$ the diagram

$$\begin{tikzcd}
a\otimes c \arrow[r, "\mathrm{id}_a\otimes\rho(c)^{-1}"] \arrow[d, "\varepsilon_c"'] & a\otimes(1\otimes c) \arrow[r, "\alpha(a{,}1{,}c)"] \arrow[d, "\varepsilon_{e\otimes c}"'] \arrow[rr, bend left, "\mathrm{id}_a\otimes\rho c"] & (a\otimes 1)\otimes c \arrow[r, "\lambda a\otimes\mathrm{id}_c"] \arrow[d, "\varepsilon_1\otimes\mathrm{id}_c"] & a\otimes c \arrow[d, "f\otimes\mathrm{id}_c"] \\
b\otimes c \arrow[r, "\mathrm{id}_b\otimes\rho(c)^{-1}"'] & b\otimes(1\otimes c) \arrow[r, "\alpha(b{,}1{,}c)"'] \arrow[rr, bend right, "\mathrm{id}_b\otimes\rho c"'] & (b\otimes 1)\otimes c \arrow[r, "\lambda b\otimes\mathrm{id}_c"'] & b\otimes c
\end{tikzcd}$$

is commutative in $\mathtt{M}$: the left and center squares commute by the naturallity of ε, the up and down wings commute by the triangle identities, the right square commutes by the definition of f. It follows from commutativity that $\varepsilon_c = f \otimes \mathrm{id}_c$, therefore $Ef = \varepsilon$.

- (Faithful) Let f and g be morphisms of $\mathtt{M}$ such that $Ef = Eg$, so that in particular $f \otimes \mathrm{id}_1 = g \otimes \mathrm{id}_1$—hence $f = g$, proving injectivity on the morphism collections of $\mathtt{M}$ and $\mathtt{N}$.
- (Monoidal) First, it is clear that $E1 = (1 \otimes (-), \alpha(1, -, -)) \simeq (\mathrm{id}_{\mathtt{M}}, I)$. Moreover, for any pair of morphisms f and g of $\mathtt{M}$ one has

$$E(f \otimes g) = (f \otimes g) \otimes (-) \simeq f \otimes (g \otimes -) = Ef \widehat{\otimes} Eg.$$

Given any two $a, b \in \mathtt{M}$, from definition:

$$E(a \otimes b) = ((a \otimes b) \otimes (-), \alpha(a \otimes b, -, -)) \simeq (a \otimes (b \otimes -), \alpha(a \otimes b, -, -)),$$

also we know that if

$$(a \otimes (b \otimes -), \beta) \coloneqq (a, \alpha(a, -, -)) \widehat{\otimes} (b, \alpha(b, -, -)) = Ea \widehat{\otimes} Eb,$$

then β is defined so that the up wing of the diagram

$$\begin{tikzcd}
a \otimes (b \otimes (c \otimes d)) \arrow[r, "a \otimes \alpha(b,c,d)"'] \arrow[d, "\alpha(a,b,c \otimes d)"'] \arrow[rr, bend left, "\beta_{(c,d)}"] & a \otimes ((b \otimes c) \otimes d) \arrow[r, "\alpha(a, b \otimes c, d)"'] & (a \otimes (b \otimes c)) \otimes d \arrow[d, "\alpha(a,b,c) \otimes \mathrm{id}_d"] \\
(a \otimes b) \otimes (c \otimes d) \arrow[rr, "\alpha(a \otimes b, c, d)"'] & & ((a \otimes b) \otimes c) \otimes d
\end{tikzcd}$$

commutes in $\mathtt{M}$ for any two $c, d \in \mathtt{M}$—the square commutes by the pentagon identity. This shows that

$$\alpha(a, b, -) \colon (a \otimes (b \otimes -), \beta) \xrightarrow{\simeq} (a \otimes (b \otimes -), \alpha(a \otimes b, -, -))$$

is an isomorphsim in $\mathtt{N}$. Therefore $E(a \otimes b) \simeq Ea \widehat{\otimes} Eb$. For the left and right unitor isomorphisms $E(\lambda a) \simeq \widehat{\lambda}(Ea)$ and $E(\rho a) \simeq \widehat{\rho}(Ea)$ we shall simply argue that they both come straight from the triangle identity of $\mathtt{M}$. Similarly,

$$E(\alpha(a, b, c)) \simeq \widehat{\alpha}(Ea, Eb, Ec)$$

works via a reduction to the pentagon identity in $\mathtt{M}$.

This proves that $E \colon \mathtt{M} \to \mathtt{N}$ is indeed a monoidal equivalence of categories. ♮

From now on, in view of the equivalence given by Theorem 4.1.4, whenever possible we shall adress any monoidal category as a strict monoidal category.

Definition 4.1.5 ((Co)monoid objects in $\mathtt{Mon}$)**.** Let $(\mathtt{M}, \otimes, 1, \alpha, \lambda, \rho)$ be a monoidal category. We define the following objects:

(a) A *monoid* in $\mathtt{M}$ is a triple (m, μ, θ)—where m is an object of $\mathtt{M}$, a bifunctor $\mu\colon m \otimes m \to m$ refered to as a *multiplication*, and a functor $\theta\colon 1 \to m$ called *unit*—such that both diagrams

$$\begin{tikzcd}
m \otimes (m \otimes m) \arrow[r, "\alpha(m,m,m)"] \arrow[d, "\mathrm{id}_m \otimes \mu"'] & (m \otimes m) \otimes m \arrow[r, "\mu \otimes \mathrm{id}_m"] & m \otimes m \arrow[d, "\mu"] \\
m \otimes m \arrow[rr, "\mu"'] & & m
\end{tikzcd}$$

$$\begin{tikzcd}
1 \otimes m \arrow[r, "\theta \otimes \mathrm{id}_m"] \arrow[dr, "\rho"'] & m \otimes m \arrow[d, "\mu"] & m \otimes 1 \arrow[l, "\mathrm{id}_m \otimes \theta"'] \arrow[dl, "\lambda"] \\
& m &
\end{tikzcd}$$

commute in $\mathtt{M}$. A morphism of monoids $\phi\colon (m, \mu, \theta) \to (m', \mu', \theta')$ is a morphism $\phi\colon m \to m'$ in $\mathtt{M}$ satisfying $\phi\mu = \mu'(\phi \otimes \phi)$, and $\phi\theta = \theta'$. We then define the subcategory $\mathtt{Mon}_{\mathtt{M}}$ of $\mathtt{M}$ composed of monoidal objects in $\mathtt{M}$.

(b) A *comonoid* in $\mathtt{M}$ is a triple (c, κ, σ) c is an object of $\mathtt{M}$, a bifunctor $\kappa\colon c \to c \otimes c$ refered to as a *comultiplication*, and a functor $\sigma\colon c \to 1$ called *counit*—such that both diagrams

$$\begin{tikzcd}
c \otimes (c \otimes c) & (c \otimes c) \otimes c \arrow[l, "\alpha(c,c,c)^{-1}"'] & c \otimes c \arrow[l, "\kappa \otimes \mathrm{id}_c"'] \\
c \otimes c \arrow[u, "\mathrm{id}_c \otimes \kappa"] & & c \arrow[ll, "\kappa"] \arrow[u, "\kappa"']
\end{tikzcd}$$

$$\begin{tikzcd}
1 \otimes c & c \otimes c \arrow[l, "\sigma \otimes \mathrm{id}_c"'] \arrow[r, "\mathrm{id}_c \otimes \sigma"] & c \otimes 1 \\
& c \arrow[ul, "\rho^{-1}"] \arrow[u, "\kappa"'] \arrow[ur, "\lambda^{-1}"'] &
\end{tikzcd}$$

commute in $\mathtt{M}$. A morphism of comonoids $\psi\colon (c, \kappa, \sigma) \to (c', \kappa', \sigma')$ is a morphism $\psi\colon c \to c'$ in $\mathtt{M}$ satisfying $\kappa'\psi = (\psi \otimes \psi)\kappa$, and $\sigma = \sigma'\psi$. We then define the subcategory $\mathtt{coMon}_{\mathtt{M}}$ of $\mathtt{M}$ composed of comonoidal objects in $\mathtt{M}$.

Definition 4.1.6 (Monoid actions)**.** Let $(\mathtt{M}, \otimes, 1)$ be a monoidal category, and $(m, \mu, \theta) \in \mathtt{Mon}_{\mathtt{M}}$. A *left-action* of the monoid (m, μ, θ) on an object $a \in \mathtt{M}$ is a bifunctor $\sigma\colon m \otimes a \to a$ such that

$$\begin{tikzcd}
m \otimes (m \otimes a) \arrow[r, "\alpha(m,m,a)"] \arrow[d, "\mathrm{id}_m \otimes \sigma"'] & (m \otimes m) \otimes a \arrow[r, "\mu \otimes \mathrm{id}_a"] & m \otimes a \arrow[d, "\sigma"'] & 1 \otimes a \arrow[l, "\theta \otimes \mathrm{id}_a"'] \arrow[dl, "\lambda"] \\
m \otimes a \arrow[rr, "\sigma"'] & & a &
\end{tikzcd}$$

commutes in $\mathtt{M}$. Right-actions are defined analogously.

Given any two left-actions $\sigma \colon m \otimes a \to a$ and $\lambda \colon m \otimes b \to b$, we define a *morphism of left-actions* $\phi \colon \sigma \to \lambda$ to be a morphism $\phi \colon a \to b$ in $\mathtt{M}$ such that

$$\begin{tikzcd}
m \otimes a \arrow[r, "\mathrm{id}_m \otimes \phi"] \arrow[d, "\sigma"'] & m \otimes b \arrow[d, "\lambda"] \\
a \arrow[r, "\phi"'] & b
\end{tikzcd}$$

is a commutative diagram in $\mathtt{M}$. With these notions we are able to define two categories $\mathtt{rActMon}_{(\mathtt{M},m)}$ and $\mathtt{lActMon}_{(\mathtt{M},m)}$, composed of right and left actions of m on objects of $\mathtt{M}$, respectively, and morphisms between them.

4.2 Braided & Symmetric Monoidal Categories

Definition 4.2.1 (Braiding)**.** Given a monoidal category $(\mathtt{M}, \otimes, 1, \alpha, \lambda, \rho)$, we define a *braiding* of $\mathtt{M}$ to be a natural isomorphism[1]

$$\gamma \colon (- \otimes -') \overset{\cong}{\Longrightarrow} (-' \otimes -),$$

that is coherent with associativity and unitors of $\mathtt{M}$, in the sense that the diagrams

$$\begin{tikzcd}
(a \otimes b) \otimes c \arrow[r, "\gamma_{(a \otimes b, c)}"] \arrow[d, "\alpha^{-1}(a,b,c)"'] & c \otimes (a \otimes b) \arrow[d, "\alpha(c,a,b)"] \\
a \otimes (b \otimes c) \arrow[d, "\mathrm{id}_a \otimes \gamma_{(b,c)}"'] & (c \otimes a) \otimes b \arrow[d, "\gamma_{(c,a)} \otimes \mathrm{id}_b"] \\
a \otimes (c \otimes b) \arrow[r, "\alpha(a,c,b)"'] & (a \otimes c) \otimes b
\end{tikzcd}
\qquad
\begin{tikzcd}
a \otimes (b \otimes c) \arrow[r, "\gamma_{(a, b \otimes c)}"] \arrow[d, "\alpha(a,b,c)"'] & (b \otimes c) \otimes a \arrow[d, "\alpha(b,c,a)^{-1}"] \\
(a \otimes b) \otimes c \arrow[d, "\gamma_{(a,b)} \otimes \mathrm{id}_c"'] & b \otimes (c \otimes a) \arrow[d, "\gamma_{(c,a)} \otimes \mathrm{id}_b"] \\
(b \otimes a) \otimes c \arrow[r, "\alpha(b,a,c)"'] & b \otimes (a \otimes c)
\end{tikzcd}$$

$$\begin{tikzcd}
a \otimes 1 \arrow[rr, "\gamma_{(a,1)}"] \arrow[dr, "\lambda"'] & & 1 \otimes a \arrow[dl, "\rho"] \\
& a &
\end{tikzcd}$$

should commute for all triples (a, b, c) of objects of $\mathtt{M}$.

Definition 4.2.2 (Braided monoidal category)**.** A monoidal category $(\mathtt{M}, \otimes, 1)$ is said to be *braided* if it is endowed with a braiding γ. We shall denote this data shortly by $(\mathtt{M}, \gamma)$.

Corollary 4.2.3. For any given pair of morphisms $f \colon a \to b$ and $g \colon c \to d$ in a braided monoidal category $\mathtt{M}$, we have

$$(g \otimes f)\gamma_{(a,c)} \simeq \gamma_{(b,d)}(f \otimes g).$$

[1]That is, for any two $a, b \in \mathtt{M}$ one has an isomorphism $\gamma_{(a,b)} \colon a \otimes b \xrightarrow{\simeq} b \otimes a$.

Definition 4.2.4 (Braided monoidal functor)**.** A monoidal functor $F\colon(\mathtt{A},\gamma)\to(\mathtt{B},\widehat{\gamma})$ between braided monoidal categories is said to be a *braided monoidal functor* if for every pair of objects $a,b\in\mathtt{A}$ the braiding coherence square

$$\begin{tikzcd}
Fa\otimes Fb \arrow[r, "\widehat{\gamma}"] \arrow[d, "\simeq"'] & Fb\otimes Fa \arrow[d, "\simeq"] \\
F(a\otimes b) \arrow[r, "F\gamma"'] & F(b\otimes a)
\end{tikzcd}$$

commutes in $\mathtt{B}$.

Corollary 4.2.5. The composition of braided monoidal functors is again a braided monoidal functor.

Proof. Indeed, if $(\mathtt{A},\gamma)\xrightarrow{F}(\mathtt{B},\widehat{\gamma})\xrightarrow{G}(\mathtt{C},\widetilde{\gamma})$ are braided monoidal functors, then for every pair $a,b\in\mathtt{A}$ one has the following commutative diagram in $\mathtt{C}$:

$$\begin{tikzcd}
GFa\otimes GFb \arrow[r, "\widetilde{\gamma}"] \arrow[d, "\simeq"'] & GFb\otimes GFa \arrow[d, "\simeq"] \\
G(Fa\otimes Fb) \arrow[r, "G\widehat{\gamma}"] \arrow[d, "\simeq"'] & G(Fb\otimes Fa) \arrow[d, "\simeq"] \\
GF(a\otimes b) \arrow[r, "GF\gamma"'] & GF(b\otimes a)
\end{tikzcd}$$

♮

Definition 4.2.6. We define a category $\mathtt{BrMonCat}$ composed of braided monoidal categories and braided monoidal functors between them.

Definition 4.2.7 (Symmetric monoidal category)**.** A braided monoidal category $(\mathtt{M},\gamma)$ is said to be *symmetric* if for any two $a,b\in\mathtt{M}$ the triangle

$$\begin{tikzcd}
a\otimes b \arrow[rr, "\mathrm{id}_{a\otimes b}"] & & a\otimes b \arrow[dl, shift left, "\gamma_{(a,b)}"] \\
& b\otimes a \arrow[ur, shift left, "\gamma_{(b,a)}"] &
\end{tikzcd}$$

commutes in $\mathtt{M}$. A morphism between symmetric monoidal categories is a braided monoidal functor between them.

Part II

Algebra

Chapter 5

Vector Spaces

5.1 Vector Spaces and Subspaces

Vector Spaces

Definition 5.1.1 (Vector Space)**.** A set V is called a *vector space* over a field k (or k-space) if it is equipped with an internal operation $+\colon V \times V \to V$ where $(a, b) \mapsto a + b$ and the external operation $\cdot\colon k \times V \to V$ where $(r, a) \mapsto r \cdot a$. We normally call the elements of V as vectors, and the elements of k as scalars. Also, these operations satisfy

I. $(V, +)$ is an abelian group.

II. Multiplication of vectors by scalars is associative and distributive and is unitary (that is, $1 \cdot a = a$ for every $a \in V$).

Subspaces

Definition 5.1.2 (Subspaces)**.** Let V be a vector space. A set $S \subseteq V$ is called a *subspace* of V if it satisfies all properties of a vector space. Also S is called a *proper subspace* if it is not equal to the original vector space.

Theorem 5.1.3 (Cover Avoidance)**.** A non-zero vector space V over an infinite field k is not the union of a finite number of proper subspaces.

Proof. Suppose, for the sake of contradiction, that $V = \bigcup_{i\in I} S_i$ where I is a finite indexing set and S_i are all proper subspaces of V. Then, let S_1 be such that it is not contained in any other subspace S_i. Define elements $a \in S_1 \setminus \bigcup_{i\in I\setminus\{1\}} S_i$ and $b \in V \setminus S_1$ and construct the set $A := \{ra + b : r \in k\}$. Notice that if we have that one element $ra+b \in S_1$, the fact that $a \in S_1$ makes $(ra+b)-ra = b \in S_1$, contradicting the assumption of $b \notin S_1$ thus, we can't have $ra + b \in S_1$. Suppose now that we let $ra + b, r'a + b \in S_i$ different elements, for some $i > 1$, then $(ra + b) - (r'a + b) = (r - r')a \in S_i$ but since k is a field, then $a \in S_i$ contradicting again the construction and therefore we cannot have more than one element of A contained in S_i. Now, since A is an infinite set and is contained in V, then we cannot have the equality between V and the finite union $\bigcup_{i\in I} S_i$. ♮

Definition 5.1.4 (Lattice)**.** A poset P is called a *lattice* if for every pair of elements of P there exists a *join* (or a least upper bound) and a *meet* (greatest lower bound). The set P is called a *complete lattice* if there exists a join and a meet for every collection of sets and also every collection contains smallest and larger elements under the partial order.

Proposition 5.1.5 (Intersection of subspaces)**.** The intersection of any collection of subspaces of a given vector space is a subspace of the original vector space. This intersection contains the greatest lower bound of subspace that is contained in every subspace of the intersection, then, we denote it by

$$\bigcap_{i \in I} S_i = \mathrm{Glb}\{S_i : i \in I\}.$$

Proof. Let V be a k-space and S_i be an arbitrary subspace. Notice that $0 \in \bigcup_{i \in I} S_i$. Let elements $u, v \in \bigcap_{i \in I} S_i$ so that for all subspace we have that the element $u + vt \in S_i$ for any given $t \in k$, thus $u + vt \in \bigcap_{i \in I} S_i$, which makes the union a vector space by itself. The claim follows. $\natural$

Definition 5.1.6 (Sum of subspaces)**.** Let V be a vector space and S_i be subspaces of V. We define the sum of such subspaces as

$$\sum_{i \in I} S_i \coloneqq \left\{ \sum_j s_j : s_j \in \bigcup_{i \in I} S_i \right\}.$$

Therefore the least upper bound under set inclusion as

$$\mathrm{Lub}\{S_i : i \in I\} = \sum_{i \in I} S_i.$$

Theorem 5.1.7 (Subspaces form a complete lattice)**.** The set containing all subspaces of V, denoted by $\mathcal{S}(V)$, is a complete lattice under set inclusion (partial order of the set), with smallest element $\{0\}$ (the zero subspace) and largest element V. The meet of any collection of sets $\{S_i : i \in I\}$, where I is a finite indexing set, is

$$\bigcap_{i \in I} S_i = \mathrm{Glb}\{S_i : i \in I\}$$

and the join is defined as

$$\sum_{i \in I} S_i = \mathrm{Lub}\{S_i : i \in I\}.$$

Morphisms of Vector Spaces

Definition 5.1.8 (Morphisms)**.** Let V, L be k-vector spaces. We say that $\varphi \colon V \to L$ is a morphism of vector spaces if it satisfies

I. $\varphi(0) = 0$, that is $V \ni 0 \mapsto 0 \in L$.

II. For all $u, v \in V$, $\varphi(u + b) = \varphi(u) + \varphi(v)$.

III. For all $a \in k$ and $u \in V$ we have $\varphi(au) = a\varphi(u)$.

Notice in fact that the first property of the morphism is in fact redundant. By means of item two we can see that $\varphi(0) = \varphi(0) + \varphi(0)$, for which the item one is obtained.

Definition 5.1.9. A k-linear morphism $f\colon V \to V$ is called a linear operator.

Proposition 5.1.10. The k-vector spaces, together with morphism between such vector spaces, form a category, of which we'll denote by $\mathtt{Vect}_k$.

Proof. Certainly, the identity map $\mathrm{id}_V\colon V \to V$ is a morphism of k-vector spaces and all of the properties come directly from the fact that V is a vector space.

Another important feature of morphism between vector spaces is that they are closed under composition. Let $\varphi\colon V \to L$ and $\psi\colon L \to U$ be morphism between k-vector spaces, then the composition $\psi\varphi\colon V \to U$ is such that, given any $u, v \in V$ then

$$\psi(\varphi(u + v)) = \psi(\varphi(u) + \varphi(v)) = \psi(\varphi(u)) + \psi(\varphi(v))$$

and also, being $a \in k$, we have

$$\psi(\varphi(au)) = \psi(a\varphi(u)) = a\psi(\varphi(u)).$$

Which proves the closure under composition.

We now prove that the composition of morphism of vector spaces is associative. Let the morphisms be as before and define yet another morphism $\ell\colon W \to V$, then

$$(\psi(\varphi\ell))(w) = \psi(\varphi(\ell(w)) = (\psi\varphi)(\ell(w)) = ((\psi\varphi)\ell)(w).$$

Consider the morphism $\varphi\colon V \to L$, then clearly $\varphi\,\mathrm{id}_V = \varphi = \mathrm{id}_L\,\varphi$. Together with the fact that the collection of morphisms between k-vector spaces V, L and the collection of morphisms between k-vector spaces U, W, these two collections are clearly disjoint for $V \neq U$ and $L \neq W$. This finishes the proof of the properties needed for a category. ♮

Proposition 5.1.11 (Initial and Final object)**.** The k-vector space 0 is a initial and final object of $\mathtt{Vect}_k$.

Proof. Essentially, we need to prove that for all $V \in \mathrm{Obj}(\mathtt{Vect}_k)$ there exists unique morphisms φ and ψ where

$$\begin{tikzcd} 0 \arrow[r, dashed, bend left, "\psi"] & V \arrow[l, dashed, bend left, "\varphi"] \end{tikzcd}$$

Notice that, since φ is a morphism of k-vector spaces, we'll need to impose $\varphi(0) = 0$ and thus this morphism is clearly unique and satisfies the properties needed for a morphism. Moreover, the morphism has a unique target element, thus the image of ψ is the singleton $\{0\}$ which is also clearly unique and satisfies the properties of morphism. ♮

Definition 5.1.12 (Isomorphism)**.** Let $V, L \in \mathrm{Obj}(\mathtt{Vect}_k)$. We say that L and V are isomorphic, that is $L \simeq V$, if there exists an isomorphism $L \to V$ in $\mathrm{Mor}(L, V)$.

5.2 Matrices

Definition 5.2.1 (Matrix)**.** We define a $m \times n$ matrix with entries in k as morphism $k^n \to k^m$ in the category $\mathtt{Vect}_k$, that is, regarding k^n, k^m as k-vector spaces.

Classifying Matrices

Let e_j be defined the be a tuple whose j-th element is $1 \in k$ and all of the other elements of the tuple are $0 \in k$, moreover, if $e_j \in k^n$ it is an n-tuple. Let the morphism $\varphi \in \mathrm{Mor}_{\mathtt{Vect}_k}(k^n, k^m)$ and define for all $1 \leqslant j \leqslant n$ the image

$$\varphi(e_j) := (t_{1j}, t_{2j}, \dots, t_{mj}) = \sum_{i=1}^{m} t_{ij} e_i \in k^m.$$

We now prove that in fact the mn elements $t_{ij} \in k$ determine completely the behaviour of φ, since for any element $(a_j)_{j=1}^n \in k^n$ we have

$$\varphi((a_j)_{j=1}^n) = \varphi\left(\sum_{j=1}^{n} a_j e_j\right) = \sum_{j=1}^{n} a_j \varphi(e_j) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_j t_{ij} e_i\right) \in k^m.$$

Moreover, it can trivially be seen that the mapping $(a_j)_{j=1}^n \mapsto \sum_{j=1}^n \sum_{i=1}^m a_j t_{ij} e_i$is indeed a morphism of vector spaces since, for another $(b_j)_{j=1}^n \in k^n$, we have

$$\varphi((a_j + b_j)_{j=1}^n) = \varphi((a_j)_{j=1}^n) + \varphi((b_j)_{j=1}^n)$$

also, given $c \in k$ we have

$$\varphi((ca_j)_{j=1}^n) = c\varphi((a_j)_{j=1}^n).$$

Since the matrix $k^n \to k^m$ can be identified and completely determined with elements $(t_{ij})_{i,j}$ where $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$, then we visually can represent it by

$$\varphi = \begin{pmatrix} t_{11} & \dots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{m1} & \dots & t_{mn} \end{pmatrix}$$

Also, we can regard the fact that $\varphi((a_i)_{i=1}^n) = (b_i)_{i=1}^m$ visually as a system of equations

$$\begin{pmatrix} t_{11} & \dots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{m1} & \dots & t_{mn} \end{pmatrix}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Definition 5.2.2 (Square diagonal Matrix)**.** We define a $n \times n$ matrix A to be diagonal if for all $i \neq j$ indices we have $a_{ij} = 0$.

Definition 5.2.3 (Matrix for a linear map)**.** Suppose $V \simeq k^n$ and $W \simeq k^m$ are k-vector spaces and $L\colon V \to W$ is a linear morphism. Let also $\{v_j\}_{j=1}^n$ and $\{w_i\}_{i=1}^m$ be basis for the respective given finite dimensional vector spaces. By the isomorphism, we can represent L as a matrix $k^n \to k^m$ whose components $t_{i,j}$ with $1 \leqslant i \leqslant m$ and $1 \leqslant i \leqslant n$ are defined with respect to the given basis as

$$Lv_j = \sum_{i=1}^{m} t_{i,j} w_i$$

Matrix multiplication

Definition 5.2.4 (Multiplication of matrices)**.** Let $A = [a_{i,j}]\colon k^n \to k^m$ and $B = [b_{i,j}]\colon k^\ell \to k^n$ be matrices. Then, the product of the matrices A and B is defined as $AB\colon k^n \to k^\ell$ with coefficients

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

Proposition 5.2.5 (Composition of morphisms matrix)**.** Let k-linear morphisms $V \xrightarrow{g} W \xrightarrow{f} L$ of finite k-dimensional vector spaces, and choose basis $\{v_j\}, \{w_k\}, \{l_i\}$ to be basis of V, W, L respectively, and let A_g and A_f be the matrix representation of the morphisms g and f with respect to the given basis. Then the matrix representation of the composition $fg\colon V \to L$ is given by $A_{fg} = A_f A_g$.

Proof. Let $A_f := [a_{i,k}]$, and $A_g := [b_{k,j}]$, and $A_{fg} = [c_{i,j}]$. Then, since from definition we have $g(v_j) = \sum_k b_{k,j} w_k$, hence

$$fg(v_j) = \sum_k b_{k,j} f(w_k) = \sum_k \left(b_{k,j} \left(\sum_i a_{i,k} l_i \right) \right) = \sum_i \left(\sum_k a_{i,k} b_{k,j} \right) l_i = \sum_i c_{i,j} l_i$$

thus $c_{i,j} = \sum_k a_k i, k b_{k,j}$ and thus $A_{fg} = A_f A_g$ as wanted. ♮

Definition 5.2.6 (Conjugation)**.** Let $M_n(k)$ be the collection of matrices $k^n \to k^n$. We define a matrix conjugation as the linear morphism $M_n(k) \to M_n(k)$ with the mapping $A \mapsto B^{-1}AB$, where B is an invertible matrix.

Proposition 5.2.7. Every conjugation is an automorphism of the matrix algebra $M_n(k)$.

5.3 Product and Direct Sum of Vector Spaces

Free Vector Spaces

Proposition 5.3.1. Let S be a set and define the power set k^S with addition and scalar multiplication, that is, given $f, g \in k^S$ and $a \in k$ we have $(f+g)(x) = f(x) + g(x)$ and $(af)(x) = af(x)$. Then the set k^S is a k-vector space.

Proof. Notice that clearly $(k^S, +)$ is an abelian group from the construction of the additive structure, moreover, since k is a field, it inherit the associativity and distributivity of scalar multiplication. The 0 vector can be regarded as the map whose image is the singleton $\{0\}$. ♮

Proposition 5.3.2 (Functoriality of k^S)**.** Let sets S, S' and a map $\alpha\colon S \to S'$, then

$$\alpha^*\colon k^{S'} \to k^S \text{ mapping } (f\colon S' \to k) \mapsto (f\alpha\colon S \to k)$$

is a morphism of k-vector spaces. Moreover if $\beta\colon S' \to S''$ then $(\beta\alpha)^* = \alpha^*\beta^*$.

Proof. Let maps $f, g\colon S' \to k$, then

$$(f + g\colon S' \to k) \xmapsto{\alpha^*} ((f + g)\alpha\colon S \to k)\,,$$

but for every $s \in S$ we have $(f + g)(\alpha(s)) = f(\alpha(s)) + g(\alpha(s))$, which shows the first property. For the second property, let $a \in k$, then we get the map

$$((a \cdot f)\colon S' \to k) \xmapsto{\alpha^*} ((a \cdot f)\alpha\colon S \to k),$$

but for each $s \in S$ we have $(a \cdot f)(\alpha(s)) = a \cdot f(\alpha(s))$, proving the second property, which finishes the proof. ♮

Definition 5.3.3 (Free vector space)**.** Let S be a set, we define the free vector space on S to be the object

$$k^{\oplus S} = \{f \in k^S : f(s) \neq 0 \text{ only for finitely many } s \in S\},$$

together with an additive structure and scalar multiplication, satisfying for all $f, g \in k^{\oplus S}$, $s \in S$ and $a \in k$:

$$(f + g)(s) = f(s) + g(s) \text{ and } (a \cdot f)(s) = a \cdot f(s).$$

Proposition 5.3.4. The object $k^{\oplus S}$ is a k-vector space.

Proof. Let the maps $f, g \in k^{\oplus S}$, so that the specification of finiteness of non-zero values is satisfied. Denote A, B the finite sets containing the non-zero elements of S under, respectively, f and g. Notice now that the map $f + g \in k^S$ is such that for all $s \in A \cup B$ then $(f + g)(s) = f(s) + g(s) \neq 0$ thus, denoting C as the set containing all non-zero elements of S under $f + g$, we see that $A \cup B \subseteq C$. The converse is trivial since if $s \in C$ then $f(s) + g(s) \neq 0$ and thus at least one of the images is non-zero, thus $s \in A \cup B$ and then $C \subseteq A \cup B$, which implies that $A \cup B = C$ and, in particular, C is finite. Therefore $f + g \in k^{\oplus S}$. Note now that for any $c \in \mathbf{R} \setminus \{0\}$, then $c \cdot f$ has non-zero values only in $A \subseteq S$, in the case $c = 0$ then $c \cdot f$ is zero all over S, which makes $c \cdot f \in k^{\oplus S}$. Thus indeed $k^{\oplus S}$ is a vector space. ♮

For each $s \in S$ we can define a map $\mathbf{s}\colon S \to k$ such that

$$\mathbf{s}(x) := \begin{cases} 1, & \text{for } x = s \\ 0, & \text{for } x \neq s \end{cases}$$

Also, this notion comes together with the natural monomorphism

$$\iota\colon S \hookrightarrow k^{\oplus S} \text{ mapping } s \mapsto \mathbf{s}$$

which allow us to write any element $v \in k^{\oplus S}$ as a linear combination, with non-zero scalars $a_i \in k$, of the form

$$v = \sum_{i=1}^{n} a_i \mathbf{s}_i \in k^{\oplus S}, \text{ mapping } s \mapsto \begin{cases} a_i, & \text{if } s = s_i \text{ for some } i \\ 0, & \text{if } s \neq s_i \text{ for all } i \end{cases}$$

Proposition 5.3.5 (Universal property of the free vector spaces)**.** Let S be a set, and V be a k-vector space, and a function $f\colon S \to V$. Then there exists a unique k-linear morphism $\ell\colon k^{\oplus S} \to V$ such that the diagram commutes:

$$\begin{tikzcd} S \arrow[r, "f"] \arrow[d, hook, "\iota"] & V \\ k^{\oplus S} \arrow[ur, dashed, "\ell"'] & \end{tikzcd}$$

Proof. (Uniqueness) Notice that for all $\alpha \in k^{\oplus S}$ we can write it as a finite sum $\alpha = \sum_{s\in S} \alpha(s)\mathbf{s}$ (the finiteness comes from the fact that only finitely may $\mathbf{s}$ will actually appear, since there are only finite non-zero values under α). Let now $L\colon k^{\oplus S} \to V$ be a morphism of k-vector spaces such that $f = L\iota$ for a given $f\colon S \to V$. Then

$$L(\alpha) = \sum_{s\in S} L(\alpha(s)\mathbf{s}) = \sum_{s\in S} \alpha(s)L(\mathbf{s}) = \sum_{s\in S} \alpha(s)(L\iota)(s) = \sum_{s\in S} \alpha(s)f(s),$$

which implies that the morphism L is indeed unique.

To prove the existence of L, consider its definition as above. Notice that for any $\alpha, \beta \in k^{\oplus S}$ we have

$$L(\alpha + \beta) = \sum_{s\in S} (\alpha + \beta)(s)f(s) = \sum_{s\in S} \alpha(s)f(s) + \sum_{s\in S} \beta(s)f(s) = L(\alpha) + L(\beta).$$

Moreover, if $a \in k$, then $k\alpha = k\sum_{s\in S} \alpha(s)\mathbf{s}$, thus indeed

$$L(k\alpha) = \sum_{s\in S} (k\alpha)(s)f(s) = k\sum_{s\in S} \alpha(s)f(s) = kL(\alpha),$$

proving the last condition in order to be a morphism of k-vector spaces, ♮

This way we find that actually

$$\mathrm{Mor}_{\mathtt{Set}}(S, V) \simeq \mathrm{Mor}_{\mathtt{Vect}_k}(k^{\oplus S}, V).$$

We now want in some way to construct a morphism that maps $k^{\oplus S'} \to k^{\oplus S}$, where S, S' are sets, just as we've done for the case $k^{S'} \to k^S$. To do so, we can first consider the maps $\alpha \colon S \to S'$, and the inclusion maps $\iota_S \colon S \to k^{\oplus S}$ and $\iota_{S'} \colon S' \hookrightarrow k^{\oplus S'}$, and construct the map

$$\ell = \iota_{S'}\alpha \colon S \to k^{\oplus S'}.$$

Then, by means of the universal property for free vector spaces, we find that there exists a unique morphism of k-vector spaces $\alpha_* \colon k^{\oplus S} \to k^{\oplus S'}$ such that $\ell = \alpha_* \iota_S$, that is, the following diagram commutes

$$\begin{tikzcd} S \arrow[r, "\alpha"] \arrow[d, "\iota_S"] \arrow[dr, "\ell"] & S' \arrow[d, "\iota_{S'}"] \\ k^{\oplus S} \arrow[r, dashed, "\alpha_*"] & k^{\oplus S'} \end{tikzcd}$$

Product

Definition 5.3.6 (Pair product)**.** Let L, V be k-vector spaces. Then the product

$$L \times V = \{(\ell, v) : \ell \in L,\ v \in V\},$$

equipped with pairwise sum and product by scalar is a k-vector space.

Definition 5.3.7 (General product)**.** Let $\{V_i\}_{i \in I}$ be a set of k-vector spaces, with an indexing set I. The product of these vector spaces is defined as

$$\prod_{i \in I} V_i = \{(v_i)_{i \in I} : v_i \in V_i\}$$

equipped with addition and multiplication by scalar as

$$(v_i)_{i \in I} + (u_i)_{i \in I} = (v_i + u_i)_{i \in I} \text{ and } a(v_i)_{i \in I} = (av_i)_{i \in I}$$

is a k-vector space.

Proposition 5.3.8. Let S be a set. Then there exists a natural isomorphism

$$k^S \simeq \prod_{s \in S} k,$$

so that $\prod_{s \in S} k$ generalizes the power set.

Proof. For any $f \in k^S$ map $(f \colon S \to k) \mapsto (f(s))_{s \in S}$. Notice that the mapping is obviously a monomorphism, since the maps $S \to k$ are well defined, and is also an epimorphism because we just need to create a function fitting such an image. ♮

Proposition 5.3.9 (Universal property for products)**.** Let the indexed set of k-vector spaces $\{V_i\}_{i\in I}$. For any given k-vector space L, together with a morphism $\varphi_j\colon L \to V$, there exists a unique morphism $\ell\colon L \to \prod_{i\in I} V_i$ such that the diagram commutes:

$$\begin{tikzcd} L \arrow[r, dashed, "\ell"] \arrow[dr, "\varphi_j"'] & \prod_{i\in I} V_i \arrow[d, two heads, "\pi_j"] \\ & V_j \end{tikzcd}$$

Proof. Notice that this is simply the product universal product restricted for only one branch of the product $V_j \times \prod_{i\neq j} V_i$. For the uniqueness, notice that since we want $\varphi_j = \pi_j \ell$, then it ought to be the case that

$$L \ni x \overset{\ell}{\longmapsto} (\varphi_j(x))_{i\in I} \in \prod_{i\in I} V_i$$

since the composition needs to be satisfied for any index $j \in I$. Therefore the morphism ℓ is indeed unique, if it exists. Now we show its existence. Notice that since φ_j is a k-linear map, then indeed for any $x, y \in L$ and $a \in k$ we have $\ell(x + ay) = (\varphi_i(x + ay))_{i\in I} = (\varphi_i(x))_{i\in I} + a(\varphi_i(y))_{i\in I} = \ell(x) + a\ell(y)$, which shows the linearity and the existence. ♮

Definition 5.3.10 (Direct sum)**.** Given an indexed collection of k-vector spaces $\{V_i\}_{i\in I}$, we define the direct sum of them as

$$\bigoplus_{i\in I} V_i := \{(v_i)_{i\in I} : v_i \in V_i, \text{ where } v_i \neq 0 \text{ finitely many times in } I\}$$

together with addition and scalar multiplication being defined component-wise.

Proposition 5.3.11. The direct sum of k-vector spaces is again a k-vector space.

Proposition 5.3.12. Let a set S. Then, the exists natural isomorphism $\bigoplus_{s\in S} k \simeq k^{\oplus S}$.

Proof. Notice that if we make the morphism $\ell\colon \bigoplus_{s\in S} k \to k^{\oplus S}$ defined as the mapping

$$\bigoplus_{s\in S} k \ni (a_s)_{s\in S} \overset{\ell}{\longmapsto} (s \overset{f}{\mapsto} a_s) \in k^{\oplus S}$$

then we see that no information is lost since this is surely injective, because the morphisms of $k^{\oplus S}$ are well defined, and given any morphism $f \in k^{\oplus S}$ we know that its non-zero values are finite, thus we can construct its image as a tuple in $\bigoplus_{s\in S} k$, which allow us to say that such mapping is also surjective. Therefore, the morphism ℓ is an isomorphism. We now show that it is indeed k-linear. Let the tuples $(a_s)_{s\in S}$ and $(b_s)_{s\in S}$, then

$$(a_s)_s + (b_s)_s = (a_s + b_s)_s \overset{\ell}{\longmapsto} (s \overset{f}{\mapsto} a_s + b_s) = (s \overset{g}{\mapsto} a_s) + (s \overset{h}{\mapsto} b_s).$$

Moreover, if $c \in k$ is any scalar, then $(ca_s)_s \longmapsto (s \mapsto ca_s) = c(s \mapsto a_s)$, which shows the last property for ℓ being a k-linear morphism. ♮

Proposition 5.3.13 (Universal property for the direct sum)**.** Let the collection of k-vector spaces $\{V_i\}_{i\in I}$ and for every $j \in I$ define the inclusion $\iota_j \colon V_j \to \bigoplus_{i\in I} V_i$, where

$$v \xmapsto{\iota_j} (v_i)_{i\in I}, \text{ where } v_i = \begin{cases} v, & i = j \\ 0, & i \neq j \end{cases}.$$

Then, for any arbitrary k-vector space L, together with k-linear morphisms $(f_i \colon V_i \to L)_{i\in I}$, there exists a unique k-linear morphism $\ell \colon \bigoplus_{i\in I} V_i \to L$ such that the diagram commutes for every $j \in I$:

$$\begin{tikzcd} L & \bigoplus_{i\in I} V_i \arrow[l, dashed, "\ell"'] \\ & V_j \arrow[ul, "f_j"] \arrow[u, hook, "\iota_j"'] \end{tikzcd}$$

Proof. Since we want to have $\ell\iota_j = f_j$, it must be the case for any $v \in v_j$ that, being $\iota_j(v) := (v_i)_i$, then

$$\ell((v_i)_{i\in i}) = \ell\left(\sum_{i\in i} \iota_i(v_i)\right) = \sum_{i\in i} \ell(\iota_i(v_i)) = \sum_{i\in i} f_i(v_i)$$

thus indeed necessarily, if such a map exists, then ℓ is unique. We now check that ℓ is indeed a k-linear morphism.

$$\ell\left((v_i + w_i)_i\right) = \sum_i f_i(v_i + w_i) = \sum_i f_i(v_i) + \sum_i f_i(w_i)$$
$$\ell((av_i)_i) = \sum_i f_i(av_i) = a\sum_i f_i(v_i)$$

Which proves that ℓ is indeed a morphism of k vector spaces as wanted. ♮

Proposition 5.3.14. Let $\{S_i\}_{i\in S}$ be a collection of sets. There exists a canonical isomorphism

$$\bigoplus_{i\in I} k^{\oplus S_i} \simeq k^{\oplus(\coprod_{i\in I} S_i)}$$

Proof. Just use both universal properties for coproduct and free vector spaces, this establishes a two way unique morphism, which proves the canonical isomorphism. ♮

Proposition 5.3.15. Let the k-linear morphism $T \colon k^n \to k^m$ and $S \colon k^n \to k^\ell$. If we consider their matrix representation, they induce the k-linear morphism

$$M \colon k^n \to k^{m+\ell} \text{ represented by } M = \begin{pmatrix} T \\ S \end{pmatrix}$$

Proof. Notice that from the universal property of products we have

$$\begin{tikzcd} k^n \arrow[r, "T"] \arrow[d, "S"'] \arrow[dr, "M"] & k^m \\ k^\ell & k^m \times k^\ell \simeq k^{m+\ell} \arrow[l, "\pi_\ell"] \arrow[u, "\pi_m"'] \end{tikzcd}$$

Where the isomorphism $k^m \times k^\ell \simeq k^{m+\ell}$ is induced by the mapping $((a_i)_{i=1}^m, (b_i)_{i=1}^\ell) = (c_i)_{i=1}^{m+\ell}$ where $c_i \coloneqq a_i$ for all $1 \leqslant i \leqslant m$ and $c_i \coloneqq b_i$ for all $m+1 \leqslant i \leqslant m+\ell$. Notice then that the morphism M is induced by the morphism T, S so that $T = \pi_m M$ and $S = \pi_\ell M$, so that we need to have $M(v) = (T(v), S(v))$ for all $v \in k^m$, but notice that $M \colon k^n \to k^m \times k^\ell$ is in fact isomorphic to a k-linear map $k^n \to k^{m+\ell}$, so that we can encode the transformation of M, without losing, information as follows. Suppose that

$$T = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{\ell 1} & \dots & b_{\ell n} \end{bmatrix}$$

then we can say that M can be isomorphically represented by

$$M(v) = \left((u_i)_{i=1}^m, (u_i)_{i=m+1}^{m+\ell}\right) \simeq \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \\ b_{(m+1)1} & \dots & b_{(m+1)n} \\ \vdots & \ddots & \vdots \\ b_{(m+\ell)1} & \dots & b_{(m+\ell)n} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ u_{m+\ell} \end{bmatrix}$$

where $(u_i)_{i=1}^m = T(v)$ and $(u_i)_{i=m+1}^{m+\ell} = S(v)$. ♮

Proposition 5.3.16. Let the k-linear morphism $T \colon k^n \to k^\ell$ and $S \colon k^m \to k^\ell$. They induce the k-linear morphism

$$M \colon k^{n+m} \simeq k^n \oplus k^m \to k^\ell \text{ represented by } M = \begin{pmatrix} T & S \end{pmatrix}$$

Proof. From the universal property of the coproduct we have that

$$\begin{tikzcd} k^n \arrow[d, hook, "\iota_n"] \arrow[dr, "T"] & \\ k^{n+m} \simeq k^n \oplus k^m \arrow[r, "M"] & k^\ell \\ k^m \arrow[u, hook, "\iota_m"'] \arrow[ur, "S"'] & \end{tikzcd}$$

Where k^n, k^m are both finite, then $k^n \oplus k^m \simeq k^n \times k^m \simeq k^{n+m}$, which proves the isomorphism written. Therefore $M \colon k^n \oplus k^m \to k^\ell$ is isomorphic to a k-linear morphism

$k^{n+m} \to k^\ell$. Moreover, if T and S are represented by

$$T = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{\ell 1} & \dots & a_{\ell n} \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{\ell 1} & \dots & b_{\ell m} \end{bmatrix}$$

Then from the construction $T = M\iota_n$ and $S = M\iota_m$ we conclude that for all element $a = (a_i)_{i=1}^n \in k^n$ and every $b = (b_i)_{i=1}^m \in k^m$, will be such that

$$k^n \ni (a_i)_{i=1}^n \overset{\iota_n}{\longmapsto} (a_1, \dots, a_n, 0, \dots, 0) \overset{M}{\longmapsto} (a_i')_{i=1}^\ell = T(a) \in k^\ell$$
$$k^m \ni (b_i)_{i=1}^m \overset{\iota_m}{\longmapsto} (b_1, \dots, b_m, 0, \dots, 0) \overset{M}{\longmapsto} (b_i')_{i=1}^\ell = S(b) \in k^\ell$$

therefore, if we have the element

$$(a_1, \dots, a_n, b_1, \dots, b_m) \simeq ((a_1, \dots, a_n), (b_1, \dots, b_m)) \overset{M}{\longmapsto} (a_i' + b_i')_{i=1}^\ell = T(a) + S(b)$$

which shows that indeed

$$M = \begin{bmatrix} a_{11} & \dots & a_{1n} & b_{1(n+1)} & \dots & b_{1(n+m)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{\ell 1} & \dots & a_{\ell n} & b_{\ell(n+1)} & \dots & b_{\ell(n+m)} \end{bmatrix} = \begin{bmatrix} T & S \end{bmatrix}.$$

$\natural$

5.4 Quotients and Subspaces

Definition 5.4.1 (Subspace)**.** Let V be k-vector space. We say that $W \subseteq V$ is a subspace of V if

I. $0 \in W$.

II. If $v, u \in W$ then $v + u \in W$.

III. For all $c \in k$ and for all $v \in W$, we have $cv \in W$.

Definition 5.4.2 (Quotient)**.** Let V be a k-vector space and $W \subseteq V$ a subspace. The quotient of V by W is defined as $V/W = V/\sim$ where $v \sim u \Leftrightarrow v - u \in W$.

Theorem 5.4.3 (Universal property for quotients)**.** Let V be a k-vector space and $W \subseteq V$ be a subspace. Let also L be any k-vector space and a k-linear morphism $f \colon V \to L$ such that for all $w \in W$ we have $f(w) = 0$, that is $W \subseteq \ker f$. There exists a unique k-linear morphism $\ell \colon V/W \to L$ such that the following diagram commutes

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & L \\ V/W \arrow[ur, dashed, "\ell"'] & \end{tikzcd}$$

Moreover, $\ker \ell = \ker f / W$ and $\operatorname{im}(\ell) = \operatorname{im}(f)$.

Proof. Since we need $\ell\pi = f$ then $\ell\pi\colon v \overset{\pi}{\mapsto} [v] \overset{\ell}{\mapsto} f(v)$. We first show that it is indeed well defined. For that, suppose $v \sim u$, that is $[v] = [u]$, then we know that since they belong to the same class, they must have the same image under the mapping ℓ. Notice that since $v \sim u \Leftrightarrow v - u \in W \subseteq \ker f$ then $f(v-u) = f(v) - f(u) = 0$ and therefore indeed $f(v) = f(u)$ whenever $[v] = [u]$. The uniqueness of ℓ comes from the fact that its image depends strictly on the unique image of f.

Now we show the last two propositions. For the first one, notice that $\ker \ell = \{[v] \in V/W : v \in \ker f\} = \ker f/W$. For the last statement of the theorem we have that since $\ell\pi = f$ then $\operatorname{im}(f) \subseteq \operatorname{im}(\ell)$ and also π is surjective, thus $\operatorname{im}(\ell) \subseteq \operatorname{im}(f)$, which proves what is proposed at last. ♮

Theorem 5.4.4 (First Isomorphism)**.** Let $f\colon V \to W$ be a k-linear morphism, then the injective k-linear morphism $\ell\colon V/\ker f \hookrightarrow W$ is such that

$$V/\ker f \simeq \operatorname{im}(f).$$

is an isomorphism of k-vector spaces.

Proof. This is just a special case of the universal property. ♮

Theorem 5.4.5 (Second Isomorphism)**.** Let V be a k-vector space and $U, W \subseteq V$ be subspaces. Then there is a natural isomorphism

$$(W+U)/U \simeq W/(W \cap U).$$

Proof. Notice that this isomorphism needs to be the mapping

$$(W+U)/U \ni [w+u] \longmapsto [w] \in W/(W \cap U)$$

Notice that we have the mappings

$$\begin{gathered}[w+u] + [w'+u'] = [(w+w') + (u+u')] \longmapsto [w+w'] = [w] + [w'] \\ c[w+u] = [cw+cu] \longmapsto [cw] = c[w]\end{gathered}$$

which shows the conditions for the mapping to be a k-linear morphism. Now notice that the mapping is clearly surjective, since given any class $[w] \in W/(W \cap U)$ we have the element $[w+u] \in (W+U)/U$, where u is any element of U and moreover $[w+u] \mapsto [w]$ from construction, showing the surjective character of the mapping. Moreover, notice that $[w] \in W/(W \cap U) \Leftrightarrow w \in U$, thus the kernel of the mapping is equal to U. Now by means of the first isomorphism theorem shows that the mapping is also injective. ♮

Theorem 5.4.6 (Third Isomorphism)**.** Let V be a k-vector space and let the subspaces $U \subseteq W \subseteq V$. Then there exists an embedding $W/U \hookrightarrow V/U$ with a mapping $[w] \mapsto [w]$ such that W/U can be regarded as a subspace of V/U. Moreover, we have a natural isomorphism

$$V/W \simeq (V/U)\big/(W/U).$$

Proof. Consider the morphism $\varphi\colon V/U \to V/W$ with the mapping $[v] \overset{\varphi}{\mapsto} [v]$. Since $U \subseteq W$ then we have that $[v] = [v'] \in V/U$ implies $[v] = [v'] \in V/W$, which shows that φ is well defined. Moreover, given a class $[v] \in V/W$ we have the corresponding class $[v] \in V/U$ such that $[v] \overset{\varphi}{\mapsto} [v]$, thus φ is surjective.

Moreover, notice that if $[w] \in W/U$ then from the embedding we have the corresponding class $[w] \in V/U$, we'll have two possibilities, if $[w] = [0]$, then the k-linear morphism φ will map it to $[0]$, on the other hand, if $[w] \neq [0]$ then $w \in W$ implies that $[w] \overset{\varphi}{\mapsto} [w] = [0] \in V/W$ so that in both cases we have $[w] \in W/U \Rightarrow [w] \in \ker\varphi$ so that $W/U \subseteq \ker\varphi$. Moreover, if $V/U \ni [v] \in \ker\varphi$, then $[v] = [0] \in V/W$ which means that $v \in W$ and therefore we have the corresponding class $[v] \in W/U$, which sums up to $\ker\varphi \subseteq W/U$ and thus $\ker\varphi = W/U$. From the first isomorphism theorem we have that

$$(V/U)\big/(W/U) \simeq V/W$$

since the image of φ is equal to V/W, settling the proof. ♮

5.5 Vector Spaces from Linear Morphism

Proposition 5.5.1. Given k-vector spaces V and W, consider the k-linear morphism $\varphi\colon V \to W$. Then φ is injective if and only if $\ker\varphi = \{0\}$.

Proof. ($\Rightarrow$) Suppose φ is injective, then, since $0 \overset{\varphi}{\mapsto} 0$, it follows that $\ker\varphi = \{0\}$. ($\Leftarrow$) Suppose now that $\ker\varphi = \{0\}$, then from the universal property of quotients we find that $V/\ker\varphi \rightarrowtail W$ is injective, since the only possible element in V that generates the zero class is $0 \in V$, therefore, we can regard the above mapping as $[v] \mapsto \varphi(v)$, therefore we conclude that φ is injective. ♮

Theorem 5.5.2 (Universal property for kernels)**.** Let V, W be k-vector spaces and consider the k-linear morphism $\varphi\colon V \to W$. Denote $\iota\colon \ker\varphi \hookrightarrow V$ the embedding morphism. Then, for any given k-vector space L and k-linear morphism $\psi\colon L \to V$ such that $\varphi\psi = 0$, there exists a unique k-linear morphism $\ell\colon L \to \ker\varphi$ such that the following diagram commutes

$$\begin{tikzcd}
L \arrow[d, dashed, "\ell"'] \arrow[dr, "\psi"] \arrow[drr, bend left, "0"] & & \\
\ker\varphi \arrow[r, hook, "\iota"] & V \arrow[r, "\varphi"] & W
\end{tikzcd}$$

Proof. The condition for the commutativity of the diagram is $\psi = \iota\ell$. Notice that since the definition of ψ depends uniquely in the condition $\varphi\psi = 0$, we find that since $\mathrm{im}(\ell) \subseteq \ker\varphi$ it follows that indeed $\varphi(\iota\ell) = 0$ and therefore the uniqueness of ℓ comes by defining it as $\ell(l) = \psi(l)$ and the existence comes merely by the fact that ψ is already a k-linear morphism. ♮

Definition 5.5.3 (Cokernel)**.** Let k-vector spaces V and W and consider the linear morphism $\varphi\colon V \longrightarrow W$. We define the cokernel of φ as

$$\operatorname{coker}\varphi = W/\operatorname{im}(\varphi).$$

Theorem 5.5.4 (Universal property for cokernels)**.** Let k-vector spaces V and W and the linear morphism $\varphi\colon V \longrightarrow W$. Denote by $\pi\colon W \longrightarrow \operatorname{coker}\varphi$ the projection morphism. For any given k-vector space L and linear morphism $\psi\colon W \longrightarrow L$ for which $\psi\varphi = 0$, there exists a unique linear morphism $\ell\colon \operatorname{coker}\varphi \longrightarrow L$ such that the diagram commutes

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[drr, bend right, "0"'] & W \arrow[r, "\pi"] \arrow[dr, "\psi"] & \operatorname{coker}\varphi \arrow[d, dashed, "\ell"] \\ & & L \end{tikzcd}$$

Proof. The condition for the diagram to commute is that $\ell\pi = \psi$. For that, notice that since $\operatorname{coker}\varphi = W/\operatorname{im}(\varphi)$ it follows that $\ell\pi\varphi(v) = \ell([\varphi(v)])$ but notice that obviously $\varphi(v) \in \operatorname{im}(\varphi)$ thus $[\varphi(v)] = 0 \in \operatorname{coker}\varphi$ thus indeed $\ell\pi\varphi = 0$. Moreover, in order to define ℓ we ought to have $\ell([w]) = \psi(w)$, which is well defined since $[w] = w + \operatorname{im}(\varphi)$ and $\forall w' \in \operatorname{im}(\varphi)$ we have $\psi(w') = 0$ and therefore $\psi(w + w') = \psi(w)$, which ends the proof since ℓ gets the linear structure of ψ and also is unique from construction. ♮

Proposition 5.5.5. Let k-vector spaces V and W and define any linear morphism $\varphi\colon V \longrightarrow W$, then there are natural isomorphisms

$$\ker(W \xrightarrow{\pi} \operatorname{coker}\varphi) \simeq \operatorname{coker}(\ker\varphi \xrightarrow{\iota} V) \simeq \operatorname{im}(\varphi)$$

Proof. First we show the existence of the isomorphism $\ker\pi \simeq \operatorname{im}(\varphi)$, which is obtained by taking the mapping $w \mapsto w$, since $\ker\pi = \operatorname{im}(\pi)$, establishing an obvious isomorphism between the two given objects. Now we focus on showing the isomorphism $\operatorname{coker}\iota = V/\ker\varphi \simeq \operatorname{im}(\varphi)$, which can be obtained by considering the mapping $[v] \mapsto \varphi(v)$, which is well defined from the same reasoning as in the above theorem. ♮

Consider the collections of k-vector spaces $\{V_i\}_i$ and $\{W_i\}_i$ with a corresponding collection of morphism $\{f_i\colon V_i \longrightarrow W_i\}_i$ and $\{\varphi_i\colon V_i \longrightarrow V_{i+1}\}_i$ and $\{\psi_i\colon W_i \longrightarrow W_{i+1}\}$ such that for any index i the following diagram commutes

$$\begin{tikzcd} V_i \arrow[r, "f_i"] \arrow[d, "\varphi_i"] & W_i \arrow[d, "\psi_i"] \\ V_{i+1} \arrow[r, "f_{i+1}"] & W_{i+1} \end{tikzcd}$$

By defining the morphisms $\iota_i\colon \ker f_i \longrightarrow V$ and $\pi_i\colon W \longrightarrow \operatorname{coker} f_i$ we find that since

$$f_{i+1}\varphi_i\iota_i = \psi_i f_i\iota_i = \psi_i 0 = 0$$

then we can use the universal property of kernels in order to find the existence of a unique morphism $\ker f_i \to \ker f_{i+1}$, that is

$$\begin{tikzcd}
\ker f_i \arrow[d, dashed] \arrow[dr, "\varphi_i \iota_i"] \arrow[drr, bend left, "0"] & & \\
\ker f_{i+1} \arrow[r, hook, "\iota_{i+1}"] & V_{i+1} \arrow[r, "f_{i+1}"] & W_{i+1}
\end{tikzcd}$$

On the other hand we have that

$$\pi_{i+1}\psi_i f_i = \pi_{i+1} f_{i+1} \varphi_i = 0\varphi_i = 0$$

thus by means of the universal property for cokernels we find that there exists a unique morphism $\operatorname{coker} f_i \to \operatorname{coker} f_{i+1}$, that is

$$\begin{tikzcd}
V_i \arrow[r, "f_i"] \arrow[drr, bend right, "0"'] & W_i \arrow[r, "\pi_i"] \arrow[dr, "\pi_{i+1}\psi_i"] & \operatorname{coker} f_i \arrow[d, dashed] \\
& & \operatorname{coker} f_{i+1}
\end{tikzcd}$$

Binding both results together we find that in general we have the following diagram

$$\begin{tikzcd}
\vdots \arrow[d, dashed] & \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d, dashed] \\
\ker f_i \arrow[r, "\iota_i"] \arrow[d, dashed] & V_i \arrow[r, "f_i"] \arrow[d, "\varphi_i"] & W_i \arrow[r, "\pi_i"] \arrow[d, "\psi_i"] & \operatorname{coker} f_i \arrow[d, dashed] \\
\ker f_{i+1} \arrow[r, "\iota_{i+1}"] \arrow[d, dashed] & V_{i+1} \arrow[r, "f_{i+1}"] \arrow[d] & W_{i+1} \arrow[r, "\pi_{i+1}"] \arrow[d] & \operatorname{coker} f_{i+1} \arrow[d, dashed] \\
\vdots & \vdots & \vdots & \vdots
\end{tikzcd}$$

Definition 5.5.6 (Exact sequence)**.** Given a collection of k-vector spaces $\{V_i\}_i$ and k-linear morphisms $\{f_i \colon V_i \to V_{i+1}\}$, we say that the sequence

$$V_0 \xrightarrow{f_0} V_1 \xrightarrow{f_1} \dots \xrightarrow{f_{n-1}} V_n$$

is exact if $\ker f_i = \operatorname{im}(f_{i-1})$ for all index.

Proposition 5.5.7. Let the k-linear morphism $f \colon V \to W$. Then f is injective if and only if the sequence $0 \longrightarrow V \xrightarrow{f} W$ is exact. On the other hand, the morphism f is surjective if and only if the sequence $V \xrightarrow{f} W \longrightarrow 0$ is exact.

Proof. ($\Rightarrow$) Suppose f is injective, then $\ker f = \{0\}$, moreover, surely $\mathrm{im}(0) = \ker f$ thus the sequence is exact. ($\Leftarrow$) Suppose the sequence is exact, so that $\mathrm{im}(0) = \ker f$ but then $\ker f = \{0\}$ which, as we already have proven above, implies in f injective.

($\Rightarrow$) Suppose f is surjective, then $\mathrm{im}(f) = W = \ker 0$ and thus the sequence is exact. ($\Leftarrow$) Suppose the sequence is exact, so that $\mathrm{im}(f) = W = \ker 0$, then given any element $w \in W$ there exists at least one corresponding $v \in V$ such that $f(v) = w$ and therefore the morphism is surjective. ♮

Definition 5.5.8 (Short exact sequence)**.** Let the k-vector spaces V, W and L. A exact sequence

$$0 \longrightarrow W \overset{f}{\rightarrowtail} V \overset{g}{\twoheadrightarrow} L \longrightarrow 0$$

that is, with $\ker g = \mathrm{im}(f)$, is said to be a short exact sequence of k-vector spaces.

Proposition 5.5.9. In the short exact sequence $0 \longrightarrow W \overset{f}{\rightarrowtail} V \overset{g}{\twoheadrightarrow} L \longrightarrow 0$ we have $L \simeq V/W$.

Proof. Firstly, notice that since $W \rightarrowtail V$ is injective, we can regard W as a subspace of V, so that taking the quotient V/W is possible. Moreover, notice that since g is surjective, from the first isomorphism theorem we find that $V/\ker g \simeq L$ which from the fact that the sequence is exact, is the same as $V/\mathrm{im}\, f \simeq L$ but this is exactly what we meant by V/W since W was regarded as a subspace of V via the embedding f. ♮

Proposition 5.5.10. Let $f\colon V \to W$ be a k-linear morphism, then the following is true

(a). The sequence $0 \to \ker f \to V \to \mathrm{im}\, f \to 0$ is exact.

(b). The sequence $0 \to \mathrm{im}\, f \to W \to \mathrm{coker}\, f \to 0$ is exact.

(c). The sequence $0 \to \ker f \to V \to W \to \mathrm{coker}\, f \to 0$ is exact.

Proof. For (I) notice that trivially $\ker(\ker f \to V) = \mathrm{im}(0) = \{0\}$, since $\ker f \subseteq V$; moreover, $\mathrm{im}(\ker f \to V) = \ker f = \ker(V \to \mathrm{im}(f))$ from the mere definition of kernel; finally, it is to be noticed that $\mathrm{im}(V \to \mathrm{im}(f)) = \ker(\mathrm{im}(f) \to 0)$ because the zero morphism has its whole domain as its kernel. For (II) we have that $\mathrm{im}(0 \to \mathrm{im}(f)) = \{0\} \subseteq W$, thus indeed $\ker(\mathrm{im}(f) \to W) = \mathrm{im}(0 \to \mathrm{im}(f))$; after that, notice $\mathrm{im}(\mathrm{im}(f) \to W) = \mathrm{im}(f)$ but from definition we have $\mathrm{coker}\, f = W/\mathrm{im}(f)$ thus we really get $\mathrm{im}(f) = \ker(\mathrm{coker}\, f)$; the next morphism trivially obeys $\mathrm{im}(W \to \mathrm{coker}\, f) = \ker(\mathrm{coker}\, f \to 0)$. For (III) the part $0 \to \ker f \to V$ is already exact, also from definition $\mathrm{im}(\ker f \to V) = \ker f = \ker(V \to W)$; then, $\mathrm{im}(V \to W) = \mathrm{im}(f) = \ker(\mathrm{coker}\, f)$; and finally $\mathrm{im}(W \to \mathrm{coker}\, f) = \ker(\mathrm{coker}\, f \to 0)$. ♮

Definition 5.5.11 (The Mor space)**.** Let the k-vector spaces V, W. Then the collection of morphisms $\mathrm{Mor}_{\mathtt{Vect}_k}(V, W)$ is a k-vector space with structure given by

$$(f+g)(v) = f(v) + g(v),\ \forall f, g \in \mathrm{Mor}_{\mathtt{Vect}_k}(V, W)$$
$$(cf)(v) = cf(v),\ \forall f \in \mathrm{Mor}_{\mathtt{Vect}_k}(V, W)\ \forall c \in k$$

where 0 is the map $0\colon V \to 0 \to W$.

Remark 5.5.12. For the remaining this section we'll be dealing with the category $\mathtt{Vect}_k$, unless said otherwise, therefore I'll omit the category for the sake of notation.

Proposition 5.5.13. For any k-vector space V there exists a natural isomorphism

$$\mathrm{Mor}(k, V) \simeq V$$

Proof. Map the zero morphism $(0: k \to V) \longmapsto 0 \in V$. Then we could take the morphism $1_v: k \to V$ being defined as $k \ni 1 \overset{1_v}{\longmapsto} v \in V$, not restricting the other mappings whatsoever, then we could trivially make the mappings $(1_v: k \to V) \longmapsto v \in V$ and then we are essentially done, since this is trivially an isomorphism. ♮

Proposition 5.5.14. Let the collection of k-vector spaces $\{V_i\}_{i\in I}$, then, for any L, k-vector space, then:

I. There is a natural isomorphism

$$\mathrm{Mor}\left(L,\ \prod_{i\in I} V_i\right) \simeq \prod_{i\in I} \mathrm{Mor}(L,\ V_i).$$

II. There is a natural isomorphism

$$\mathrm{Mor}\left(\bigoplus_{i\in I} V_i,\ L\right) \simeq \prod_{i\in I} \mathrm{Mor}(V_i,\ L)$$

and therefore, given a set S, there is also a natural isomorphism $\mathrm{Mor}(k^{\oplus S}, V) \simeq V^S$.

Proof. For the first proposition, we can assign the map of zero mappings $(0: L \to \prod V_i) \longmapsto (0: L \to V_i)_i$. Next we may consider the functions $g: L \to \prod V_i$, defined as $g(l) = (f_1(l), f_2(l), \dots)$ for all $l \in L$, where $f_i: L \to V_i$; then we can simply make the mapping $g \longmapsto (\pi_i f_i)_i$.

For the second proposition, consider the morphism $\psi: \mathrm{Mor}\left(\bigoplus_{i\in I} V_i,\ L\right)$ with the mapping $f \mapsto (f\iota_i)_{i\in I}$ where $\iota_j: V_j \hookrightarrow \bigoplus_{i\in I} V_i$ is the inclusion map. For the injectivity, consider $f \in \mathrm{Mor}\left(\bigoplus_{i\in I} V_i,\ L\right), f \neq 0$ be a morphism, then $f\iota_i \neq 0$ and therefore $\ker(\psi) = 0$, which implies in the injectivity of ψ. Now, let any tuple of morphisms $(g_i)_{i\in I} \in \prod_{i\in I} \mathrm{Mor}(V_i, L)$, then we can define a morphism $f \in \mathrm{Mor}\left(\bigoplus_{i\in I} V_i,\ L\right)$ such that $f\iota_i = g_i$ then definitely $\psi(f) = (g_i)_{i\in I}$. ♮

Definition 5.5.15 (Induced Mor k-linear morphism)**.** Given a k-linear morphism $f: V \to L$ and any k-vector space W, there are induced, uniquely defined, k-linear morphisms

(a) (Pushforward) $f_*: \mathrm{Mor}(W, V) \to \mathrm{Mor}(W, L)$ with the mapping $\alpha \mapsto f\alpha$, so that the diagram commutes

$$\begin{tikzcd} V \arrow[r, "f"] & L \\ W \arrow[u, "\alpha"] \arrow[ur, dashed, "f_*(\alpha):=f\alpha"'] & \end{tikzcd}$$

(b) (Pullback) $f^*\colon \mathrm{Mor}(L, W) \longrightarrow \mathrm{Mor}(V, W)$ with the mapping $\alpha \longmapsto \alpha f$, so that the diagram commutes

$$\begin{tikzcd} V \arrow[r, "f"] & L \arrow[d, "\alpha"] \\ & W \arrow[ul, dashed, "f^*(\alpha):=\alpha f"] \end{tikzcd}$$

Proposition 5.5.16. Given V_1, V_2, V_3, k-vector spaces, the following holds for any given k-vector space L:

I. If the sequence $0 \to V_1 \to V_2 \to V_3$ is exact, then

$$\mathrm{Mor}(L, 0) = 0 \to \mathrm{Mor}(L, V_1) \to \mathrm{Mor}(L, V_2) \to \mathrm{Mor}(L, V_3)$$

is an exact sequence, that is, covariant Mor is left exact.

II. If the sequence $V_1 \to V_2 \to V_3 \to 0$ is exact, then

$$\mathrm{Mor}(0, L) = 0 \to \mathrm{Mor}(V_3, L) \to \mathrm{Mor}(V_2, L) \to \mathrm{Mor}(V_1, L)$$

is an exact sequence, that is, contravariant Mor is left exact.

Proof. For each of the propositions, we'll denote the morphisms $f\colon V_1 \to V_2$ and $g\colon V_2 \to V_3$ such that $\ker g = \operatorname{im} f$.

For the first proposition, let $0 \to V_1 \to V_2 \to V_3$ be a an exact sequence, and consider the sequence $0 \to \mathrm{Mor}(L, V_1) \xrightarrow{f_*} \mathrm{Mor}(L, V_2) \xrightarrow{g_*} \mathrm{Mor}(L, V_3)$. Let $\beta \in \operatorname{im} f_*$, then there exists $\alpha \in \mathrm{Mor}(L, V_1)$ such that $f_*(\alpha) = f\alpha = \beta$, hence $g_*(\beta) = g\beta = gf\alpha = 0$ since $\ker g = \operatorname{im} f$, which implies in $\operatorname{im} f_* \subseteq \ker g_*$. Suppose now that $\beta \in \ker g_*$, so that $g_*(\beta) = g\beta = 0$, then we find that $\operatorname{im}\beta \subseteq \ker g = \operatorname{im} f$ and hence $\beta \in \operatorname{im} f_*$, which implies in $\ker g_* \subseteq \operatorname{im} f_*$. Therefore $\ker g_* = \operatorname{im} f_*$.

For the second proposition, let $V_1 \to V_2 \to V_3 \to 0$ be a an exact sequence and consider the sequence $0 \to \mathrm{Mor}(V_3, L) \xrightarrow{g^*} \mathrm{Mor}(V_2, L) \xrightarrow{f^*} \mathrm{Mor}(V_1, L)$. Let $\gamma \in \operatorname{im} g_*$ and consider $\lambda \in \mathrm{Mor}(V_3, L)$ such that $g^*(\lambda) = \lambda g = \gamma$. Then we find that $f^*(\gamma) = \gamma f = \lambda g f = \lambda 0 = 0$ hence $\gamma \in \ker f^*$, which implies in $\operatorname{im} g^* \subseteq \ker f^*$.

Let $\gamma \in \ker f^*$ so that $f^*(\gamma) = \gamma f = 0$, this implies in $\operatorname{im} f = \ker g \subseteq \ker \gamma$. From Theorem 5.4.3 we find that γ induces a morphism $\overline{\gamma}$ such that the following diagram commutes

$$\begin{tikzcd} V_2 \arrow[r, "\gamma"] \arrow[d, "\pi"] & L \\ V/\ker g \arrow[ur, "\overline{\gamma}"'] & \end{tikzcd}$$

so that $\gamma = \overline{\gamma}\pi$. Moreover, since g is surjective we have $\operatorname{im} g = V_3$, and from Theorem 5.4.4 we find that g induces an isomorphism $\overline{g}\colon V_2/\ker g \xrightarrow{\simeq} V_3$, so that the following diagram commutes

$$\begin{tikzcd}
V_2 \arrow[r, "g"] \arrow[d, "\pi"] & V_3 \\
V_2/\ker g \arrow[ur, "\overline{g}"'] &
\end{tikzcd}$$

so that $g = \overline{g}\pi$ and since $\overline{g}$ is an isomorphism, then $\pi = g\overline{g}^{-1}$. Notice now that $g^*(\overline{\gamma g}^{-1}) = \overline{\gamma g}^{-1} g = \overline{\gamma}\pi = \gamma$ and therefore $\gamma \in \operatorname{im} g^*$. This shows that $\ker f^* \subseteq \operatorname{im} g^*$ and therefore $\ker f^* = \operatorname{im} g^*$. ♮

5.6 Bases and Dimensions

Definition 5.6.1. Let V a k-vector space and $S \subseteq V$ a subset. By the universal property of free vector spaces, the inclusion $S \hookrightarrow V$ induces the unique k-linear morphism $f\colon k^{\oplus S} \to V$. Then we can classify the set S in terms of f as:

I. Linearly independent, in the case where f is injective, and linearly dependent otherwise.

II. Generating (or spanning) if f is surjective.

III. Basis if f is an isomorphism.

Proposition 5.6.2. A set $S \subseteq V$ is linearly independent if and only if for all subset $\{v_i\}_i \subseteq S$ and scalars $\{a_i\}_i \subseteq k$, the equation $\sum_i a_i v_i = 0$ implies $a_i = 0$ for all index i.

Definition 5.6.3 (Subspace spanned)**.** Let V be a k-vector space. Given a set $S \subseteq V$, we define the subspace spanned by S to be

$$\operatorname{span}(S) = \operatorname{im}(k^{\oplus S} \to V)$$

Proposition 5.6.4. Let V be a k-vector space and the set $S \subseteq V$. Then S is a basis for V if and only if for all $v \in V$ there exists a unique tuple $(a_s)_{s\in S} \subseteq k$ such that there are only finitely many $a_s \neq 0$ and $\sum_{s\in S} a_s s = v$.

Lemma 5.6.5 (Zorn's Lemma)**.** Let X be a set and consider the partial order relation $(\vdash) \subseteq X \times X$, so that $\vdash$ is reflexive for all $x \in X$, antisymmetric for all $x, y \in X$, and transitive for all $x, y, z \in X$. Moreover, if $S \subseteq X$ is such that only one of the propositions: $x \vdash y$ or $y \vdash x$ is true for every $x, y \in S$; then there exists an element $z \in X$ for which $x \vdash z$ for all choices of $x \in X$. Then there exists an element $m \in X$ for which $m \vdash x$ if and only if $x = m$.

Theorem 5.6.6. Every vector space has a basis.

Proof. Let V be a k-vector space. Consider the set $\mathcal{A} := \{X \subseteq V : X \text{ linearly independent}\}$ partially order under inclusion, and let $C \subseteq \mathcal{A}$ be a *totally* ordered subset by inclusion.

We now show that $\mathcal{B}$ satisfies the hypothesis needed for the use of Zorn's lemma. Define Z to be the *union* of all sets in C—we'll show that Z is linearly independent. For the sake of contradiction, suppose Z is linearly dependent, then one must be able to pick a finite collection $(v_j)_{j=1}^n$ of members of Z such that there exists an associated collection $(a_j)_{j=1}^n$ of elements $a_j \in k$, not all zero, such that

$$a_1 v_1 + \cdots + a_n v_n = 0. \tag{5.1}$$

Since each $v_j \in S_j$ for some set $S_j \in C$, then by the total ordering of C we may assume that

$$S_1 \subseteq S_2 \subseteq \cdots \subseteq S_{n-1} \subseteq S_n$$

up to change of indexing. Since $v_j \in S_n$ for all $1 \leqslant j \leqslant n$, then in particular Eq. (5.1) shows that S_n is linearly dependent, which is false by hypothesis that $S_n \in \mathcal{A}$. We conclude that Z must be linearly independent and thus $Z \in \mathcal{A}$.

Since $\mathcal{A}$ is ordered by inclusion, we can apply Zorn's lemma to conclude that $\mathcal{A}$ admits a maximal element $M \in \mathcal{A}$. We'll show that M does generate V. Suppose, for the sake of contradiction, that there exists $v \in V$ such that v is not generated by M. Then it must be the case that $M \cup \{v\}$ is linearly-independent hence $M \cup \{v\} \in \mathcal{A}$ and $M \subseteq M \cup \{v\}$, which is a contradiction since M is maximal. Therefore $\operatorname{span}(M) = V$, which shows[1] that M is a basis for V. ♮

Corollary 5.6.7. Every vector space is isomorphic to $k^{\oplus S}$ for some set S.

Proposition 5.6.8. Let B be a basis of the k-vector space V. Then for every k-vector space W there is a natural isomorphism of vector spaces

$$\operatorname{Mor}_{\mathtt{Vect}_k}(V, W) \simeq \prod_{b \in B} W$$

Proof. We show that the map

$$(f : V \to W) \longmapsto (f(b))_{b \in B}$$

is an isomorphism. First, it is injective: let $f, g \in \operatorname{Mor}(V, W)$; suppose that $f = g$ so that obviously $(f(b))_{b\in B} = (g(b))_{b \in B}$ since $f(b) = g(b)$ for all $b \in B$; suppose now that $(f(b))_{b \in B} = (g(b))_{b\in B}$, since B is a basis on V then, given any $v \in V$, we can write it as $v = \sum_{b \in B} a_b b$ where $a_b \neq 0$ only finitely many and $a_b \in k$, then

$$f(v) = f\left(\sum_{b\in B} a_b b\right) = \sum_{b \in B} a_b f(b) = \sum_{b\in B} a_b g(b) = g\left(\sum_{b \in B} a_b b\right) = g(v)$$

thus indeed the mapping is injective. Now we show that it is surjective: notice that $|W|^{|B|} = |\prod_{b \in B} W|$ and since the morphism of $\operatorname{Mor}(V, W)$ are completely determined by its image under $B \subseteq V$ then we conclude that $|\operatorname{Mor}(V, W)| = |W|^{|B|}$, thus we can create a mapping which satisfies the surjectivity. ♮

[1] This proof is also true for modules over division rings.

Proposition 5.6.9. Let $\alpha\colon S \to T$ be a map between sets. Then the induced map $\alpha_*\colon k^{\oplus S} \to k^{\oplus T}$ is such that

i. α_* is an monomorphism if α is injective.

ii. α_* is an epimorphism if α is surjective.

iii. α_* is an isomorphism if α is bijective.

Proof. For the first, let α be injective, then given $f \in k^{\oplus S}$ and define the finite set $S_f := S \smallsetminus \ker f$. Now we construct a unique function $g \in k^{\oplus T}$ defined such that $g(t) \neq 0$ if and only if $t \in \alpha(S_f)$ (well defined since $\alpha(S_f)$ is finite by the injectivity of α), and moreover the condition $f(s) = g(\alpha(s))$ (uniqueness), then our mapping α_* can be defined as $f \mapsto g$, which is surely a monomorphism and well defined.

For the second proposition, let α be surjective, then from the same map as before, given any function $g \in k^{\oplus T}$, let $T_g := T \smallsetminus \ker g$ then for every $t \in T_g$ we chose one $s \in \alpha^{-1}(t)$ so that for some function $f \in k^{\oplus T}$ defined by $f(s) = g(\alpha(s)) = g(t)$ and moreover $f(s) \neq 0$ if and only if $s \in \alpha^{-1}(t)$ for some $t \in T_g$ (thus f is well defined, since T_g is finite and we are taking only one corresponding s for each $t \in T_g$). Thus indeed $f \xmapsto{\alpha_*} g$ is surjective.

The last proposition comes trivially from the last two. ♮

Proposition 5.6.10. Let the surjective k-linear morphism $f\colon V \twoheadrightarrow W$. There exists an injective k-linear morphism $g\colon W \rightarrowtail V$ such that $fg = \mathrm{id}_W$.

Proof. Just take the mapping $w \xmapsto{g} v \in f^{-1}(w) \in V$, then $w \overset{g}{\mapsto} v \overset{f}{\mapsto} w$ and thus $fg = \mathrm{id}_W$. ♮

Proposition 5.6.11. Let the injective k-linear morphism $f\colon V \rightarrowtail W$. There exists a surjective morphism $g\colon W \twoheadrightarrow V$ such that $gf = \mathrm{id}_V$.

Proof. Given any $v \in V$, we want $g(f(v)) = g(w) = v$, but since $f^{-1}(w)$ is a singleton, we can simply make the well defined mapping $w \overset{g}{\mapsto} v$ for all $w \in \mathrm{im}(f)$. With this unique condition we already have the wanted $gf = \mathrm{id}_V$. Moreover, given any $v \in V$, there must exist $w \in W$ (in fact we could specify $w \in \mathrm{im}(f)$, but here we can be more general) such that $g(w) = v$, thus g is surjective. ♮

Proposition 5.6.12. Let k-vector spaces V and W. There exists injective linear morphism $V \rightarrowtail W$ if and only if there exists a surjective morphism $W \twoheadrightarrow V$.

Proof. The last two propositions. ♮

Theorem 5.6.13. Let sets S and T. The following propositions are equivalent:

I. There exists injective linear morphism $k^{\oplus S} \rightarrowtail k^{\oplus T}$.

II. There exists surjective linear morphism $k^{\oplus T} \twoheadrightarrow k^{\oplus S}$.

III. There exists an injection $S \rightarrowtail T$.

IV. There exists a surjective map $T \twoheadrightarrow S$ or $S = \varnothing$.

Lemma 5.6.14. Let V a k-vector space, $B \subseteq V$ a basis and $S \subseteq V$ a spanning set. Then for all $b \in B$ there exists $a \in A$ such that $(B \setminus \{b\}) \cup \{a\}$ is a basis.

Proof. Let $b \in B$ be any element. Suppose, for the sake of contradiction, that $A \subseteq \operatorname{span}(B \setminus \{b\})$ then $V = \operatorname{span}(A) \subseteq \operatorname{span}(B \setminus \{b\})$, which is a contradiction because B is said to be a basis for V; thus $A \nsubseteq \operatorname{span}(B \setminus \{b\})$. Let $a \in A \setminus \operatorname{span}(B \setminus \{b\})$, then consider the set $(B \setminus \{b\}) \setminus \{a\}$. Let then $a = cb + c_1 b_1 + \cdots + c_n b_n$ with $c \neq 0$ from the construction of a (that is, $a \notin \operatorname{span}(B \setminus \{b\})$). This way we can make

$$b = a/c - c_1/cb_1 - \cdots - c_n/cb_n \quad \text{therefore} \quad b \in \operatorname{span}((B \setminus \{b\}) \cup \{a\})$$

which makes $(B \setminus \{b\}) \cup \{a\}$ a basis for V. ♮

Lemma 5.6.15. Let B and B' be bases the k-vector space V, then there exists a bijection $B \simeq B$, and thus $|B| = |B'|$.

Proof. From Theorem 5.6.13 we see that since exists injective k-linear morphism $V \rightarrowtail V$, then exists injective map $B \rightarrowtail B'$ and also surjective map $B \twoheadrightarrow B'$. From Cantor-Schröder-Bernstein theorem we see that there exists bijection $B \simeq B'$. ♮

Proposition 5.6.16. Let V be a k-vector space and $S \subseteq V$ be any linearly independent set. There exists a set $B \supseteq S$ such that B is a basis of V.

Proof. Let $\mathcal{S}$ be the non-empty collection of linearly independent sets of V containing S. Notice that a the union of a chain of elements of $\mathcal{S}$ is again a linearly independent set containing S — thus $\mathcal{S}$ is closed under arbitrary unions. In other words, every chain of elements has an upper bound in $\mathcal{S}$. By Zorn's lemma, it follows that the collection $\mathcal{S}$ has a maximal element — call it B. We now prove that B is a basis for V.

Let $v \in V$ be an element such that $v \notin B$. Since B is maximal, it follows that the set $B \cup \{v\}$ must not be linearly independent — since it contains B. This implies in the existence of a collection $(c_b)_{b \in B}$ of finitely many non-zero elements $c_b \in k$, and $c_0 \in k$ such that

$$c_0 v + \sum_{b \in B} c_b b = 0$$

with not all coefficients equal to zero. Since v isn't a member of B, it follows that $c_0 \neq 0$, otherwise B wouldn't be linearly independent. It follows that

$$v = \sum_{b \in B} -\frac{c_b}{c_0} b,$$

therefore v belongs to the span of B in V. Therefore B indeed generates V, making it a basis. ♮

Lemma 5.6.17. Let V be a k-vector space. If $B \subseteq V$ is a *minimal* generating set for V, then B is a basis of V.

Proof. We must show that B is linearly independent. If, on the contrary, B where to be linearly dependent, let $(c_b)_{b\in B}$ be a collection of finitely many non-zero scalars, not all zero, such that $\sum_{b\in B} c_b b = 0$. Choose a non-zero scalar $c_{b'}$ from this family, then

$$b' = \sum_{b\in B\smallsetminus\{b'\}} -\frac{c_b}{c_{b'}}b,$$

which implies that b' can be written as a linear combination of the remaining elements of B. This in particular implies that $B \smallsetminus \{b'\}$ is a generating set of V, which contradicts the hypothesis that B is minimal. ♮

Matrices and changes of base

Definition 5.6.18 (Change of basis operator)**.** Let $C\colon V \longrightarrow V$ be a linear operator of a finite dimensional k-vector space, $V \simeq k^n$, defined by the mapping $[v]_{B_1} \overset{C}{\mapsto} [v]_{B_2}$, where $[v]_{B_1}$ and $[v]_{B_2}$ are the representations of the vector v in the basis B_1 and B_2 of V. We define the matrix representation $C_{B_1,B_2}\colon k^n \longrightarrow k^n$ of the linear operator C to be the *change of basis matrix*. Then if $B_1 = \{e_i\}_{i=1}^n$, and $B_2 = \{e'_j\}_{j=1}^n$, and $C_{B_1,B_2} = [c_{i,j}]$, then the coefficients $c_{i,j}$ of C_{B_1,B_2} must be such that

$$v = \sum_{i=1}^n a_i e_i = \sum_{j=1}^n b_j e'_j = \sum_{j=1}^n b_j\left(\sum_{i=1}^n c_{i,j}e_i\right) = \sum_{i=1}^n\left(\sum_{j=1}^n c_{i,j}b_j\right)e_i$$

Proposition 5.6.19 (Change of basis, linear morphism)**.** Let $f\colon V \longrightarrow W$ be a k-linear morphism of finite dimensional spaces, $V \simeq k^n$ and $W \simeq k^m$. Let $M_{B,S}\colon k^n \longrightarrow k^m$ and $M_{B',S'}$ be the matrix representation of f in the basis B and S, where B is a basis of V and S a basis of W. Let now B' and S' be basis for V and S. Then, the matrix representation of f in the basis B', S', that is, $M_{B',S'}\colon k^n \longrightarrow k^m$, is given by the conjugation

$$M_{B',S'} = C_{S,S'}^{-1}M_{B,S}C_{B,B'}$$

where $C_{S,S'}$ and $C_{B,B'}$ are the change of basis matrix representations defined in Definition 5.6.18.

Proof. Consider the function $C_{S,S'}^{-1}fC_{B,B'}\colon V \longrightarrow W$, then we have the mapping

$$[v]_B \xmapsto{C_{B,B'}} [v]_{B'} \overset{f}{\mapsto} [f([v]_{B'})]_S \xmapsto{C_{S,S'}} [f([v]_{B'})]_{S'}$$

hence the statement follows directly from Proposition 5.2.5. ♮

Corollary 5.6.20. In particular, if $f\colon V \longrightarrow V$ is a k-linear operator, and B, B' are basis of V, then given the matrix representation of f in base B, namely, M_B, we have that the matrix representation of f in base B' is given by

$$M_{B'} = C^{-1}M_BC$$

where C is the matrix representation of the change of basis operator from base B to B'.

Dimension and Rank plus Nullity Theorem

Definition 5.6.21. Let V be a k-vector space. We define the dimension of V as

$$\dim_k(V) = |B|$$

for $B \subseteq V$ basis of V. In particular, if $\dim_k(V) < \infty$ we say that V is a finite dimensional k-vector space.

Lemma 5.6.22. Let V be a k-vector space and $W \subseteq V$ a k-subspace, then

$$\dim_k(V) = \dim_k(W) + \dim_k(V/W).$$

Proof. Let $B_W = \{w_i\}_{i\in I}$ be a basis for W and $B_{V/W} = \{q_j\}_{j\in J}$ a basis for V/W. Then for all $j \in J$ there exists $v_j \in V$ such that $q_j = [v_j] \in V/W$. Define the set $A := \{v_j : j \in J\}$, then the map $\varphi\colon J \to A$ with the mapping $j \mapsto v_j$ is surely surjective; moreover, if $j, t \in J$ and $j \neq t$, then $q_j \neq q_t$ since $B_{V/W}$ is a basis, hence $[v_j] \neq [v_t]$ and the map φ is injective. Then φ is a bijection and thus $|A| = |J|$. Now we show that A is a linearly independent set: let $J_t \subseteq J$ be a finite subset of J with t elements and $\alpha_1, \dots, \alpha_t \in k$ be such that $\sum_{\ell=1}^t \alpha_\ell v_{j_\ell} = 0 \in V$ then $\sum_{\ell=1}^t \alpha_\ell [v_{j_\ell}] = \left[\sum_{\ell=1}^t \alpha_\ell v_{j_\ell}\right] = [0] \in V/W$ but notice that the set of classes $[v_j]\colon v_j \in A$ is linearly independent (from the fact that this corresponds to the basis $B_{V/W}$), hence we must have $\alpha_\ell = 0$ for all $1 \leqslant \ell \leqslant t$, thus A is linearly independent.

Consider now the set $A \cup B_W$, we'll show that this is a basis for V. First we show that $A \cup B_W$ is linearly independent: let $I_s \subseteq I$ be the finite set with s indices, and $J_t \subseteq J$ be the finite set with t indices, and $\alpha_1, \dots \alpha_s, \beta_1, \dots, \beta_t \in k$, be such that $\sum_{\ell=1}^s \alpha_\ell w_{i_\ell} + \sum_{r=1}^t \beta_r v_{j_r} = 0 \in V$ if we now take the module with respect to W (as we did before with A) we see that

$$[0] = \sum_{\ell=1}^{s} \alpha_\ell [w_{i_\ell}] + \sum_{r=1}^{t} \beta_r [v_{j_r}] = \left[\sum_{r=1}^{t} \beta_r v_{j_r}\right] \in V/W$$

since $w_{i_\ell} \in W$ and thus $[w_{i_\ell}] = [0]$; since $B_{V/W}$ (and $[v_{j_r}] = q_{j_r}$) is linearly independent, we see that necessarily $\beta_r = 0$ for all $1 \leqslant r \leqslant t$. Therefore, we conclude that $\sum_{\ell=1}^s \alpha_\ell w_{i_\ell} = 0$ and since B_W is linearly independent, we conclude that $\alpha_\ell = 0$ for all $1 \leqslant \ell \leqslant s$. For the final part, we need to show that $V = \operatorname{span}(A \cup B_W)$: let $v \in V$ be any element, then $[v] \in V/W = \operatorname{span}(B_{V/W})$ then $[v] = \sum_{\ell=1}^t \alpha_\ell [v_{j_\ell}]$ for $J_t \subseteq J$ and $\alpha_1, \dots, \alpha_t \in k$; this implies that $[v] - \left[\sum_{\ell=1}^t \alpha_\ell v_{j_\ell}\right] = [0]$ and thus $v - \sum_{\ell=1}^t \alpha_\ell v_{j_\ell} \in W = \operatorname{span}(B_W)$; hence for some $\beta_1, \dots, \beta_s \in k$ and $I_s \subseteq I$ we have $v - \sum_{\ell=1}^t \alpha_\ell v_{j_\ell} = \sum_{r=1}^s \beta_r w_{i_r}$ hence $v \in \operatorname{span}(A \cup B_W)$ and thus $A \cup B_W$ is a basis for V.

We'll now show that $A \cap B_W = \varnothing$: suppose $u \in A \cap B_W$, then there exists $i \in I$ and $j \in J$ such that $u = w_i$ and $u = v_j$ — thus $q_j = [v_j] = [w_i] = [0] \in V/W$ which cannot be true since $B_{V/W}$ is a basis. Therefore $|A \cup B_W| = |A| + |B_W| = |J| + |B_W| = |B_{V/W}| + |B_W| = \dim_k(V)$ as wanted. ♮

Theorem 5.6.23 (Rank plus nullity)**.** Let $L\colon V \to W$ be a k-linear morphism, then

$$\dim_k(V) = \dim_k(\ker(L)) + \operatorname{rank}(L).$$

Proof. Notice that from the first isomorphism theorem we have $V/\ker(L) \simeq \operatorname{im}(L)$ then it follows that $\dim_k(V/\ker(L)) = \dim_k(\operatorname{im}(L)) = \operatorname{rank}(L)$ from the fact that $\ker(L)$ is a k-subspace of V, we conclude from Lemma 5.6.22 that $\dim_k(V) = \dim_k(\ker(L)) + \dim_k(V/\ker(L)) = \dim_k(\ker(L)) + \operatorname{rank}(L)$. ♮

Corollary 5.6.24. Let V and W be k-vector spaces and $L\colon V \longrightarrow W$ be a k-linear morphism. Then we have

(a). If L is injective, then $\dim_k(V) \leqslant \dim_k(W)$.

(b). If L is surjective, then $\dim_k(V) \geqslant \dim_k(W)$.

(c). If L is an isomorphism then $\dim_k(V) = \dim_k(W)$.

Proof. (a) If L is injective, then $\ker(L) = 0$ and thus $\dim_k(\ker(L)) = 0$ and from Theorem 5.6.23 we have $\dim_k(V) = \operatorname{rank}(L)$ and since $\operatorname{rank}(L) \leqslant \dim_k(W)$ since $\operatorname{im}(L) \subseteq W$ then we conclude that the proposition is true. (b) Since L is surjective, then $\operatorname{im}(L) = W$ and in particular $\operatorname{rank}(L) = \dim_k(W)$, thus Theorem 5.6.23 implies that $\dim_k(V) = \dim_k(\ker(L)) + \dim_k(W) \geqslant \dim_k(W)$. (c) Already proved before, but also comes from the last two propositions. ♮

Corollary 5.6.25. Let $L\colon V \longrightarrow W$ be a k-linear morphism of finite vector spaces V and W such that $\dim_k(V) = \dim_k(W)$. Then the following propositions are equivalent:

(a). L is an isomorphism.

(b). L is injective.

(c). L is surjective.

Proof. (b $\Rightarrow$ c) Suppose L injective, then from Theorem 5.6.23 we have $\dim_k(V) = \operatorname{rank}(L)$ since $\dim_k(\ker(L)) = 0$ then we have $\dim_k(W) = \operatorname{rank}(L)$ but then $\operatorname{im}(L) = W$ and therefore L is a surjective linear map. (c $\Rightarrow$ a) Suppose that L is surjective, from from Theorem 5.6.23 we have $\dim_k(V) = \dim_k(\ker(L)) + \operatorname{rank}(L) = \dim_k(\ker(L)) + \dim_k(W)$, but from hypothesis we have $\dim_k(V) = \dim_k(W)$, then, if the given vector spaces V and W are finite dimensional we have $\dim_k(\ker(L)) = 0$ and thus $\ker(L)$ is the zero space, which implies that L is injective, hence an isomorphism. ♮

5.7 Dual Vector Spaces

Definition 5.7.1. Let $V \in \operatorname{Obj}(\mathtt{Vect}_k)$, then, we define the dual vector space of V to be

$$V^* = \operatorname{Mor}_{\mathtt{Vect}_k}(V, k)$$

and $f \in V^*$ is called a linear functional on V.

Proposition 5.7.2. For all sets S there exists a natural isomorphism between the dual free vector space generated by S and the function space $S \longrightarrow k$, that is

$$(k^{\oplus S})^* \simeq k^S.$$

Proof. This result comes directly from the free vector space universal property, that is

$$\begin{tikzcd} S \arrow[r, "f"] \arrow[d, tail] & k \\ k^{\oplus S} \arrow[ur, dashed, "\ell"'] & \end{tikzcd}$$

so that we can assign for each function $f \in k^S$ the corresponding unique k-linear morphism $\ell \in (k^{\oplus S})^*$. ♮

Proposition 5.7.3. Let V be finite dimensional, then there exists a non-canonical isomorphism

$$V \simeq V^*$$

Proof. The proof here is rather arbitrary, we'll just construct explicitly an isomorphism by defining an ordered basis $B = (v_1, \dots, v_n)$ of V, but we could build other isomorphisms (this is why the proposition states "non-canonical"). Define the k-linear morphism $\varphi(v_i) = e_i$ where e_i is the i-th standard basis on k^n, then φ is an isomorphims $V \simeq k^n$. Moreover, from the standard inner product, we have that $k^n \simeq (k^n)^*$, thus we are done, since

$$V \xrightarrow{\varphi} k^n \simeq (k^n)^* \xrightarrow{f \mapsto f\varphi} V^*$$

thus indeed $V \simeq V^*$. ♮

Proposition 5.7.4. If V is finite-dimensional, then V^* is also finite dimensional and

$$\dim_k(V) = \dim_k(V^*).$$

Proposition 5.7.5. Given k-vector spaces V and W, there exists a natural isomorphism

$$(V \oplus W)^* \simeq V^* \oplus W^*.$$

Proof. Consider the mapping $\Phi\colon (V \oplus W)^* \to V^* \oplus W^*$ defined by $\Phi(f^*) = (f_V^*, f_W^*)$, where $f_V^* = f^*|_{V \oplus \{0\}}$ and $f_W^* = f^*|_{\{0\} \oplus W}$. We now show that Φ is an isomorphism. (Injective) Suppose $g^* \in \ker \Phi$, then necessarily $g^*|_{V \oplus \{0\}} = 0$ and $g^*|_{\{0\} \oplus W} = 0$, but since $g^* = g^*|_{V \oplus \{0\}} + g^*|_{\{0\} \oplus W}$ we conclude that $g^* = 0$. (Surjective) Let any $(\ell^*, t^*) \in V^* \oplus W^*$, then define the function $f^* \in (V \oplus W)^*$ such that $f^*(v, w) = \ell^*(v) + t^*(w)$, then $f^*|_{V \oplus \{0\}} = \ell^*$ and $f^*|_{\{0\} \oplus W} = t^*$, therefore $\Phi(f^*) = (\ell^*, t^*)$. ♮

Definition 5.7.6. Let $f\colon V \to W$ be a k-linear morphism. We define the dual (or transpose) k-linear morphism of f as $f^* = f^T\colon W^* \to V^*$. This is a particular case of the induced Mor k-linear morphism, by taking the k-vector space k, so that

$$f^*\colon \mathrm{Mor}(W, k) \to \mathrm{Mor}(V, k),\ \alpha \mapsto \alpha f.$$

It is immediate from this definition that $(gf)^* = f^* g^*$ for every composable pair of functionals f and g.

Proposition 5.7.7. Let $f\colon V \to W$ be a k-linear morphism and consider its dual $f^*\colon W^* \to V^*$. If f is injective, then f^* is surjective, on the other hand, if f is surjective, then f^* is injective. Hence, if f is an isomorphism, then f^* is an isomorphism.

Proposition 5.7.8. Let $V_1 \to V_2 \to V_3$ be an exact sequence, then $V_3^* \to V_2^* \to V_1^*$ is also exact.

Proof. Let $f\colon V_1 \to V_2$ and $g\colon V_2 \to V_3$ be such that $\operatorname{im}(f) = \ker(g)$. Then, given any $\alpha \in V_3^*$ and consider $g^*(\alpha) = \alpha g$, we then have $f^*(\alpha g) = (\alpha g)f = \alpha(gf) = 0$ from the construction of f, g, thus $\operatorname{im}(g^*) \subseteq \ker(f^*)$. On the other hand, let $\beta \in \ker(f^*) \subseteq V_2^*$ so that $f^*(\beta) = \beta f = 0$, which is the same thing as requiring $\beta|_{\operatorname{im}(f)} = 0$; this way we can see that for any element $\alpha \in V_3^*$ we have $g^*(\alpha) = \alpha g$ be such that $\alpha g|_{\operatorname{im}(f)} = 0$ since $\ker(g) = \operatorname{im}(f)$ and thus already satisfies as an element of the kernel of f^*, therefore $\ker(f^*) \subseteq \operatorname{im}(g^*)$. ♮

Proposition 5.7.9. Let $V_1 \to V_2 \to V_3$ be exact sequence of k-vector spaces, then for any given k-vector space W we have that

$$\operatorname{Mor}(V_3, W) \longrightarrow \operatorname{Mor}(V_2, W) \longrightarrow \operatorname{Mor}(V_1, W)$$

is exact (contravariant Mor).

Proof. The proof is almost identical to the latter. ♮

Proposition 5.7.10. Let $V_1 \to V_2 \to V_3$ be exact sequence of k-vector spaces, then for any given k-vector space W we have that

$$\operatorname{Mor}(W, V_1) \longrightarrow \operatorname{Mor}(W, V_2) \longrightarrow \operatorname{Mor}(V_3, W)$$

is exact (covariant Mor).

Proof. The proof only diverges from the above because we have the mappings $g_*(\alpha) = g\alpha$ and $f_*(\beta) = f\beta$. ♮

Proposition 5.7.11. Consider a k-linear morphism $f\colon V \to W$ such that $\operatorname{rank}(f)$ is finite (that is, $\dim_k(\operatorname{im}(f))$ exists). Then $\operatorname{rank}(f) = \operatorname{rank}(f^*)$.

Proof. Notice that if $f\colon V \twoheadrightarrow U \rightarrowtail W$ is the decomposition of f, then we can consider the dual $f^*\colon W^* \to V^*$ as having a decomposition $W^* \twoheadrightarrow U^* \rightarrowtail V^*$, since the dual of an injective map is surjective and the dual of a surjective map is injective. Therefore, given a basis B for U, since $\forall b \in B$ there exists $v \in V$ for which $f(v) = b$ we can say that $b \in \operatorname{im}(f)$ and, moreover, $\operatorname{rank}(f) \geqslant \dim_k(U)$. Also, since $\operatorname{im}(f) \subseteq U$, then we have directly that $\operatorname{rank}(f) \leqslant \dim_k(U)$, thus $\operatorname{rank}(f) = \dim_k(U)$. The same shows that $\operatorname{rank}(f^*) = \dim_k(U^*)$ and since $U \simeq U^*$ we find that $\dim_k(U) = \dim_k(U^*)$ and hence $\operatorname{rank}(f) = \operatorname{rank}(f^*)$. ♮

Proposition 5.7.12. Let $\ell \in \operatorname{Mor}(V, W)$, where $V \simeq k^n$ and $W \simeq k^m$, so that ℓ can be written as a matrix $L\colon k^n \to k^m$. Then the dual ℓ^* is represented by the matrix L^T.

Proof. Define $(v_j)_{j=1}^n$ an ordered basis for V and $(v_i^*)_{i=1}^n$ be its dual ordered basis; define also $(w_i)_{i=1}^m$ to be an ordered basis for W and $(w_j^*)_{j=1}^m$ be its dual ordered basis. Notice that the dual of the matrix, that is, $L^*\colon (k^m)^* \longrightarrow (k^n)^*$ can be written as $L^*\colon k^m \longrightarrow k^n$, since $(k^m)^* \simeq k^m$ and $(k^n)^* \simeq k^n$. Define the representations

$$L = \begin{bmatrix} t_{1,1} & \dots & t_{1,n} \\ \vdots & \ddots & \vdots \\ t_{m,1} & \dots & t_{m,n} \end{bmatrix} \text{ and define } L^* = \begin{bmatrix} d_{1,1} & \dots & d_{1,m} \\ \vdots & \ddots & \vdots \\ d_{n,1} & \dots & d_{n,m} \end{bmatrix}$$

By definition of the dual linear morphism, we have $L^*(w_j^*) = w_j^* L$. When this transformation is thought of as a $k^n \longrightarrow k^m$ matrix, the definition of a matrix for a linear transformation with respect to the given dual basis for W^* and V^* is defined such that $L^*(w_j^*) = \sum_{r=1}^n d_{r,j} v_r^*$. Hence, given any element $v_k \in (v_j)_{j=1}^n \subseteq V$, from the first definition:

$$w_j^* L(v_k) = w_j^* \left(\sum_{r=1}^m t_{r,k} w_r \right) = \sum_{r=1}^m t_{r,k} w_j^*(w_r) = t_{j,k} \tag{5.2}$$

from the fact that we defined w_j^* so that $w_j^*(w_j) = 1$ and $w_j^*(w_i) = 0$ for $i \neq j$. Now, from the second definition that we discussed, we have (applying the same v_k):

$$\sum_{r=1}^n d_{r,j} v_r^*(v_k) = d_{k,j} \tag{5.3}$$

from the same fact. Thus we conclude that $d_{k,j} = t_{j,k}$ for all $1 \leqslant k \leqslant n$ and $1 \leqslant j \leqslant m$, thus indeed $L^* = L^T$ ♮

Definition 5.7.13 (Column and Row Rank)**.** Let $L\colon k^n \longrightarrow k^m$ be a matrix. Define $(v_j)_{j=1}^n$ to be the vectors whose components are given by the ordered j-th column elements of the matrix representation of L. Analogously, define $(w_i)_{i=1}^m$ to be the vectors whose components are the ordered i-th row elements of the matrix representation of L. We define the column rank of L to be $\dim_k(\operatorname{span}(v_j)_{j=1}^n)$ and the row rank of L to be $\dim_k(\operatorname{span}(w_i)_{i=1}^m)$.

Proposition 5.7.14. Let $L\colon V \longrightarrow W$ be a linear morphism with $V \simeq k^n$ and $W \simeq k^m$, its row and column rank of its matrix representation $k^n \longrightarrow k^m$ are both equal to the rank of L.

Proof. Let $(v_j)_{j=1}^n$ and $(w_i)_{i=1}^m$ be the column and row vectors of the matrix representation of L. (Rank equals column rank) Let B_V and S_W be ordered basis for V and W respectively and any element $w \in \operatorname{im}(L) = \operatorname{span}(Lb_k\colon b_k \in B_V)$ notice that since $Lb_k = \sum_{i=1}^m t_{i,k} s_i$, where $t_{i,j}$ are the elements of the matrix representation of L, then we see that Lb_k is a linear combination of the k-th column vector of L, hence $Lb_k \in \operatorname{span}(v_j)_{j=1}^n$ for all $1 \leqslant k \leqslant m$. Therefore we conclude that $w \in \operatorname{span}(v_i)_{i=1}^n$. Moreover, we have $\operatorname{span}(v_i)_{i=1}^n \subseteq \operatorname{im}(L)$, thus $\operatorname{span}(v_i)_{i=1}^n = \operatorname{im}(L)$, hence we conclude finally that $\dim_k(\operatorname{im}(L)) = \operatorname{rank}(L) = \dim_k(\operatorname{span}(v_i)_{i=1}^n)$ which states that the rank of L equals the column rank. (Rank equals column rank) Notice that $L^*\colon (k^m)^* \to (k^n)^*$ is isomorphic to the matrix $k^m \to k^n$ and the matrix representation of the dual of L is equal to its transpose by 5.7.12, moreover we proved lastly that the column rank of a matrix equals its rank, thus it should be true that $\operatorname{rank}(L^*) = \dim_k(\operatorname{span}(w_i)_{i=1}^m)$. Since $L^* \simeq L$, it follows that $\operatorname{rank}(L) = \dim_k(\operatorname{span}(w_i)_{i=1}^m)$. Hence we conclude that

$$\operatorname{rank}(L) = \operatorname{rank}(\text{column}) = \operatorname{rank}(\text{row}).$$

♮

Definition 5.7.15 (Bilinear map)**.** A map $V^* \times V \to k$ that has the mapping $(\alpha, v) \mapsto \alpha(v)$ is called a bilinear map. This map is linear only when we fix the second component to some $v \in V$ so that $f_v\colon V^* \to k$ with the mapping $\alpha \mapsto \alpha(v)$ is a k-linear morphism.

This defines a rather interesting canonical k-linear morphism

$$\Psi_V\colon V \to (V^*)^*, \ \text{ mapping } \ v \mapsto (f_v = (\alpha \mapsto \alpha(v))).$$

Proposition 5.7.16. The morphism Ψ_V is an injection for every given vector space V.

Proposition 5.7.17. A finite dimensional space V is isomorphic to its double dual V^{**}. The same can't be said if V is infinite dimensional.

Proof. Let $\{v_i\}_{i=1}^n$ be a basis for V, $\{v_i^*\}_{i=1}^n$ be a basis for V^* dual to the first given basis, and $\{v_i^{**})_{i=1}^n\}$ be a basis for V^{**} dual to the second given basis. We'll show that $\Psi_V(v_i) = v_i^{**}$. Let v_i be an element of the given basis of V, then, we know from the definition of the dual of the basis that $v_k^*(v_i) = \delta_{i,k}$ and $v_t^{**}(v_k^*) = \delta_{k,t}$. Moreover, since Ψ_V is an injection, we see that if $\Psi_V(v_i) \neq 0$, hence $v_i \mapsto v_i^{**}$, which proves the statement. ♮

Definition 5.7.18 (Annihilator)**.** Let V be a finite dimensional k-vector space and W be a subspace of V. The annihilator of W is defined as the subspace of V^* given by

$$W^0 = \{\alpha \in V^* : \alpha(W) = 0\}.$$

Proposition 5.7.19. Let V be a finite dimensional space and $W \subseteq V$ a subspace. Then

$$\dim_k(V) = \dim_k(W) + \dim_k(W^0).$$

Proof. Let the inclusion $\iota\colon W \hookrightarrow V$, then its dual $\iota^*\colon V^* \twoheadrightarrow W^*$ is surjective. Moreover, let $\alpha \in W^0$ then, $\iota^*(\alpha) = \alpha\iota$, since $\operatorname{im}(\iota) \subseteq W$ we conclude that $\iota^*(\alpha) = 0$ and thus $\alpha \in \ker(\iota^*)$, so that $W^0 \subseteq \ker(\iota^*)$. Moreover, if $\alpha \in \ker(\iota^*)$ then, in particular, $\alpha(W) = 0$ hence $\alpha \in W^0$. Then we find that $W^0 = \ker(\iota^*)$. We now use the theorem 5.6.23 on the morphism ι^* so that $\dim_k(V^*) = \dim_k(\ker(\iota^*)) + \operatorname{rank}(\iota^*)$. Since ι^* is surjective, then $\operatorname{im}(\iota^*) = W^*$ and therefore $\operatorname{rank}(\iota^*) = \dim_k(W^*)$; moreover, $\dim_k(\ker(\iota^*)) = \dim_k(W^0)$; and finally $\dim_k(V^*) = \dim_k(V)$, and $\dim_k(W^*) = \dim_k(W)$. Thus indeed $\dim_k(V) = \dim_k(W) + \dim_k(W^0)$. ♮

Complexification and decomplex

Chapter 6

Multilinear Algebra

Add content about bilinear maps

6.1 Multilinear Maps and Tensor Products

The goal of this section is to introduce another concept of linearity, one of which is induced by a kind of map that is called multilinear. This is a very important generalization of what we've been studying so far.

Definition 6.1.1 (Multilinear map). Let $\{V_i\}_{i=1}^n$ be a collection of k-vector spaces and W be a k-vector space. We say that a map

$$\varphi\colon \prod_{i=1}^{n} V_i \to W$$

is a multilinear map if for a given $1 \leqslant t \leqslant n$ we have that, for all collections $\{v_i : v_i \in V_i, i \neq t\}$, the map

$$\varphi(v_1, \dots, v_{t-1}, -, v_{t+1}, \dots, v_n)\colon V_t \longrightarrow W, \text{ mapping } x \overset{\varphi}{\longmapsto} \varphi(v_1, \dots, v_{t-1}, x, v_t, \dots, v_n)$$

is a k-linear morphism.

Multilinear structures seems interesting and what is even better is that there is a way of working with multilinear maps by means of the machinery acquired when we where studying linear structures — we can simply linearize a multilinear map! That is a great idea, one that is accomplished via the introduction of a structure called tensor product. We'll now focus on the construction of it.

Given a finite collection of k-vector spaces $\{V_i\}_{i \in I}$, we define the free k-vector space $\mathcal{M} = k^{\oplus \prod_{i \in I} V_i}$ — that is, the space of set-functions $\prod_{i \in I} V_i \to k$ with finite support. A map $f \in \mathcal{M}$ is entirely described by the set of all tuples of its domain. For each tuple $(v_i)_{i \in I} \in \prod_{i \in I} V_i$, we define a characteristic map $\delta_{(v_i)_{i \in I}}\colon \prod_{i \in I} V_i \to \{0, 1\}$ — which happens to be an element of the space $\mathcal{M}$. Since $f(v) = \sum_{u \in \prod_{i \in I} V_i} f(u)\delta_v(u) =$

$f(v)\delta_v(v) = f(v)$, then we conclude that the collection $\{\delta_v \in \mathcal{M} : v \in V\}$ is a basis for the space $\mathcal{M}$.

For the sake of brevity, we'll simply denote $\delta_v = v$ for every $v \in \prod_{i \in I} V_i$. The space $\mathcal{M}$ can then be viewed as the space of formal linear combination of the tuples of $\prod_{i \in I} V_i$, that is

$$\mathcal{M} = \left\{ \sum a_v v : v \in \prod_{i \in I} V_i \text{ and } a_v \in k \right\}.$$

Thus, if $\operatorname{char} k = 0$ and exists $i \in I$ such that $\dim_k(V_i) > 0$, then $\mathcal{M}$ is an infinite-dimensional vector space.

The next step in our construction is to refine the space $\mathcal{M}$ by means of taking the quotient of $\mathcal{M}$ by some subspace $\mathcal{M}_0$ so that the maps $f \in \mathcal{M}/\mathcal{M}_0$ satisfy the properties arising from the multilinear structure induced by the maps Definition 6.1.1. We can do so by considering the collection of maps of the following form

$$f_0 = (v_1, \dots, v_j + a v_j', \dots, v_n) - (v_1, \dots, v_j, \dots, v_n) - a(v_1, \dots, v_j', \dots, v_n) \in \mathcal{M}.$$

We may simply define $\mathcal{M}_0 \subseteq \mathcal{M}$ to be the subspace spanned by the maps of the same form as $f_0 \in \mathcal{M}$. Now, when we take the quotient $\mathcal{M}/\mathcal{M}_0$ we essentially classify the maps of the form f_0 as the zero-maps — so that every element of $\mathcal{M}/\mathcal{M}_0$ satisfy an internal version of the properties of the multilinear maps. We are now ready to define what we call a tensor product — we actually already constructed it, is the space $\mathcal{M}/\mathcal{M}_0$!

Definition 6.1.2 (Tensor product)**.** We define the tensor product of the finite collection of k-vector spaces $\{V_i\}_{i \in I}$ as the quotient space

$$\bigotimes_{i \in I} V_i = \mathcal{M}/\mathcal{M}_0,$$

where the elements of $\bigotimes_{i \in I} V_i$ are called tensors, and in special, the elements of the form

$$\otimes_{i \in I} v_i = (v_i)_{i \in I} + \mathcal{M}_0 \in \bigotimes_{i \in I} V_i$$

are called factorizable tensors.

Lemma 6.1.3. Let $\{V_i\}_{i \in I}$ be a finite collection of k-vector spaces. The canonical map

$$\otimes\colon \prod_{i \in I} V_i \to \bigotimes_{i \in I} V_i, \text{ mapping } (v_i)_{i \in I} \longmapsto \otimes_{i \in I} v_i$$

is multilinear.

Proof. Let $|I| = n$, and $a \in k$ be any constant, then

$$\begin{aligned}
\otimes(v_1, \dots, v_j + a v_j', \dots, v_n) &= v_1 \otimes \cdots \otimes (v_j + a v_j') \otimes \cdots \otimes v_n \\
&= (v_1, \dots, v_j + a v_j', \dots, v_n) + \mathcal{M}_0 \\
&= [(v_1, \dots, v_j, \dots, v_n) + \mathcal{M}_0] + [a(v_1, \dots, v_j', \dots, v_n) + \mathcal{M}_0] \\
&= v_1 \otimes \cdots \otimes v_j \otimes \cdots \otimes v_n + a(v_1 \otimes \cdots \otimes v_j' \otimes \cdots \otimes v_n) \\
&= \otimes(v_1, \dots, v_j, \dots, v_n) + a \otimes (v_1, \dots, v_j', \dots, v_n).
\end{aligned}$$

Where we used the fact that

$$(v_1, \dots, v_j + av'_j, \dots, v_n) - (v_1, \dots, v_j, \dots, v_n) - a(v_1, \dots, v'_j, \dots, v_n) = 0 \in \mathcal{M}/\mathcal{M}_0$$

♮

Theorem 6.1.4 (Universal property of tensor products)**.** Let $\{V_i\}_{i\in I}$ be a finite collection of k-vector spaces, and L be any k-vector space, and $f\colon \prod_{i\in I} V_i \to L$ be any multilinear map. Then there exists a unique k-linear morphism $\ell\colon \bigotimes_{i\in I} V_i \to L$ such that the following diagram commutes

$$\begin{tikzcd} \prod_{i\in I} V_i \arrow[r, "f"] \arrow[d, "\otimes"'] & L \\ \bigotimes_{i\in I} V_i \arrow[ur, dashed, bend right, "\ell"'] & \end{tikzcd}$$

Proof. Let $|I| = n$. (Uniqueness) Suppose that for any L and f, the morphism ℓ exists, so that $f = \ell \circ \otimes$ and therefore

$$\ell \otimes (v_1, \dots, v_n) = \ell(v_1 \otimes \cdots \otimes v_n) = f(v_1, \dots, v_n)$$

then certainly ℓ is uniquely defined by the pointwise image of f, since $\{(v_1, \dots, v_n) \in \prod_{i=1}^n V_i\}$ generates $\bigotimes_{i=1}^n V_i$.

(Existence) Let $g\colon \mathcal{M} \to L$ be defined as $g(v_1, \dots, v_n) = f(v_1, \dots, v_n)$ so that f completely determines the image of g. Let now $(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n) \in \mathcal{M}_0$, then

$$g(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n) = f(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n) = 0$$

because f is a multilinear map, then $(v_1, , \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n) \in \ker(g)$. Moreover, for the second type of element of $\mathcal{M}_0$

$$\begin{aligned} g\big(a(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n)\big) &= f\big(a(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n)\big) \\ &= f\big(a(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n)\big) \\ &= af(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n) \\ &= 0 \end{aligned}$$

thus $a(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n) \in \ker(g)$ and hence $\mathcal{M}_0 \subseteq \ker(g)$. From the universal property of quotients Theorem 5.4.3 we have that g induces the uniquely defined k-linear morphism $\ell\colon \mathcal{M}/\mathcal{M}_0 \to L$ such that the diagram commutes

$$\begin{tikzcd} \mathcal{M} \arrow[r, "g"] \arrow[d, "\pi"] & L \\ \mathcal{M}/\mathcal{M}_0 \arrow[ur, dashed, "\ell"'] & \end{tikzcd}$$

then $g = \ell\pi$ and therefore for all $(v_1, \dots, v_n) \in \prod_{i=1}^n V_i$ we have

$$f(v_1, \dots, v_n) = g(v_1, \dots, v_n) = \ell\pi(v_1, \dots, v_n) = \ell(v_1 \otimes \cdots \otimes v_n)$$

then indeed $\ell\otimes = f$ as wanted. ♮

Corollary 6.1.5 (Multilinear maps are isomorphic to linear maps)**.** Let $\{V_i\}_{i=1}^p$ be a finite collection of k-vector spaces, and $\mathrm{Mor}(V_1, \dots, V_p; L)$ denote the space of multilinear maps $\prod_{i=1}^p V_i \to L$, where L is a given k-vector space. The k-linear morphism

$$\psi \colon \mathrm{Mor}\left(V_1, \dots, V_p; L\right) \longrightarrow \mathrm{Mor}\left(\bigotimes_{i=1}^p V_i, L\right), \; f \overset{\psi}{\longmapsto} \ell$$

where $f = \ell \otimes$ (as in Theorem 6.1.4), is an isomorphism, so that

$$\mathrm{Mor}\left(V_1, \dots, V_p; L\right) \simeq \mathrm{Mor}\left(\bigotimes_{i=1}^p V_i, L\right).$$

Proof. Let $\ell \in \mathrm{Mor}(\bigotimes_{i=1}^p V_i, L)$, then $f = \ell\otimes \in \psi^{-1}(\ell)$ thus ψ is surjective. On the other hand, if $f \neq 0$, then $\ell\otimes \neq 0$ and in particular $\ell \neq 0$ hence $\psi(f) = \ell \neq 0$, therefore $\ker(\psi) = 0$. Since ψ is linear, then ψ is injective. ♮

Dimensions and Bases of Tensor Products

Proposition 6.1.6. Let $\{V_i\}_{i\in I}$ be a finite collection of k-vector spaces. If exists $j \in I$ such that $V_j = 0$ is the null space then $\bigotimes_{i\in I} V_i = 0$, that is, the tensor product is the null space.

Proof. Let $V_j = 0$ for some $j \in I$. Consider any $f \in \mathrm{Mor}(\prod_{i\in I} V_i,\ L)$, where L is a k-vector space. Since f is linear on V_j, consider any element $(v_i)_{i\in I} \in \prod_{i\in I} V_i$ then necessarily $v_j = 0$ and hence $f((v_i)_{i\in I}) = 0$, since V_j has only one element, every element of the domain of f is mapped to zero, which implies in $f = 0$. Consider in particular the mapping $\otimes \in \mathrm{Mor}(\prod_{i\in I} V_i,\ \bigotimes_{i\in I} V_i)$ (Lemma 6.1.3) then from the latter discussion we have $\otimes = 0$, since $\bigotimes_{i\in I} V_i = \mathrm{span}(\mathrm{im}(\otimes))$, then we conclude that $\bigotimes_{i\in I} V_i = 0$. ♮

Proposition 6.1.7 (Tensor product dimension)**.** The dimension of the tensor product of a finite collection of finite k-vector spaces $\{V_i\}_{i\in I}$ is equal to the product of the dimensions of the vector spaces, that is

$$\dim_k \bigoplus_{i\in I} V_i = \prod_{i\in I} \dim_k V_i.$$

Proof. Let $|I| = n$. If exists a null vector space in the collection, then the dimension of the tensor product is null, thus the proposition follows. Suppose there is no such null vector space in the collection, then since $\bigotimes_{i\in I} V_i \simeq \left(\bigotimes_{i\in I} V_i\right)^* = \mathrm{Mor}(\bigotimes_{i\in I} V_i, k)$. From Corollary 6.1.5 we find that $\left(\bigotimes_{i\in I} V_i\right)^* \simeq (\prod_{i\in I} V_i)^*$. Let $B_i = \{e_j^{(i)}\}_{j=1}^{\dim_k V_i}$ be a basis for V_i for every $i \in I$, so that for every element $v_i \in V_i$ we can write it as a linear combination $v_i = \sum_{j=1}^{|B_1|} x_j^{(i)} e_j^{(i)}$ where $x_j^{(i)} \in k$. Then, for any multilinear map $f \in (\prod_{i\in I} V_i)^*$ we have

$$f(v_1, \dots, v_n) = f\left(\sum_{j_1=1}^{|B_1|} x_{j_1}^{(1)} e_{j_1}^{(1)}, \dots, \sum_{j_n=1}^{|B_n|} x_{j_n}^{(n)} e_{j_n}^{(n)}\right) = \sum_{\substack{1\leqslant j_i\leqslant |B_i|,\\ 1\leqslant i\leqslant n}} x_{j_1}^{(1)} \dots x_{j_n}^{(n)} f\left(e_{j_1}^{(1)}, \dots, e_{j_n}^{(n)}\right)$$

so that the collection $\prod_{i\in I} B_i$ forms a base for $(\prod_{i\in I} V_i)^*$ and therefore

$$\dim_k \left(\prod_{i\in I} V_i\right)^* = \prod_{i\in I} |B_i| = \prod_{i\in I} \dim_k V_i$$

Hence, since $\bigoplus_{i\in I} V_i \simeq (\prod_{i\in I} V_i)$, we conclude that $\dim_k \bigoplus_{i\in I} V_i = \prod_{i\in I} \dim_k V_i$. ♮

Lemma 6.1.8 (Tensor basis)**.** Given a finite collection of k-vector spaces $\{V_i\}_{i\in I}$ and suppose none of them are null. The basis for the tensor product $\bigotimes_{i\in I} V_i$ is given by the collection of factorizable tensors $\{\otimes_{i\in I} e^{(i)}_{j_i}\}$, where $e^{(i)}_{j_i} \in B_i$ and B_i is a basis for V_i.

What about infinite dimensional vector spaces and infinite tensor products?

Tensor Product of Function Spaces

Proposition 6.1.9. Let $\{S_i\}_{i\in I}$ be a finite collection of sets. There exists a canonical isomorphism

$$k^{\oplus(\prod_{i\in I} S_i)} \simeq \bigotimes_{i\in I} k^{\oplus S_i}.$$

Proof. We show that the k-linear morphism $\phi\colon k^{\oplus(\prod_{i\in I} S_i)} \to \bigoplus_{i\in I} k^{\oplus S_i}$ mapping $\delta_{(s_i)_{i\in I}} \mapsto \otimes_{i\in I}\delta_{s_i}$, where δ is the Kronecker delta function, is an isomorphism. First, from Proposition 5.3.5 we find that $\{\delta_{(s_i)_{i\in I}} : (s_i)_{i\in I} \in \prod_{i\in I} S_i\}$ is a base for the space $k^{\oplus(\prod_{i\in I} S_i)}$ and, in particular, $\{\delta_{s_i} : s_i \in S_i\}$ is a base for the space $k^{\oplus S_i}$. From Lemma 6.1.8 we find that $\{\otimes_{i\in I}\delta_{s_i} : s_i \in S_i\}$ is a basis for $\bigotimes_{i\in I} k^{\oplus S_i}$. ♮

6.2 Canonical Isomorphisms and Tensor Products

Associativity and Commutativity

Proposition 6.2.1 (Associativity)**.** Let k-vector spaces V, W, L, the map

$$(V \otimes W) \otimes L \to V \otimes (W \otimes L) \ \text{ mapping } \ (v \otimes w) \otimes \ell \mapsto v \otimes (w \otimes \ell)$$

is a canonical isomorphism. Hence for any collection of k-vector spaces $\{V_i\}_{i=1}^p$ we can write any arrangement of parenthesis for their tensor product as canonically isomorphic to $V_1 \otimes \cdots \otimes V_p$.

Proof. Let B_V, B_W, B_L be basis for the vector spaces V, W, L respectively. Notice that $\{(v \otimes w) \otimes \ell : (v, w, \ell) \in B_V \times B_W \times L\}$ is a basis for $(V \otimes W) \otimes L$ and $\{v \otimes (w \otimes \ell) : (v, w, \ell) \in B_V \times B_W \times B_L\}$ is a basis for $V \otimes (W \otimes L)$ (from Lemma 6.1.8). Therefore, the map $(v \otimes w) \otimes \ell \mapsto v \otimes (w \otimes \ell)$ transforms one base into another, which implies that it is a canonical isomorphism of the considered spaces. ♮

Proposition 6.2.2 (Commutativity)**.** Let $\{V_i\}_{i=1}^p$ be a collection of k-vector spaces, and σ be any permutation of the numbers $\{1, \ldots, p\}$. Define the k-linear morphism

$$f_\sigma \colon \bigotimes_{i=1}^{p} V_i \to \bigotimes_{i=1}^{p} V_{\sigma(i)} \text{ mapping } v_1 \otimes \cdots \otimes v_p \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(p)}$$

where $f_{\tau\sigma} = f_\tau f_\sigma$, for any permutation τ on $\{1, \ldots, p\}$. Then f_σ is a canonical isomorphism.

Proof. Let the map $g_\sigma \colon \prod_{i=1}^p V_i \to \bigotimes_{i=1}^p V_{\sigma(i)}$ such that $(v_1, \ldots, v_p) \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(p)}$. Then the following diagram commutes

$$\begin{tikzcd}
\prod_{i=1}^p V_i \arrow[r, "g_\sigma"] \arrow[d, "\otimes"'] & \bigotimes_{i=1}^p V_{\sigma(i)} \\
\bigotimes_{i=1}^p V_i \arrow[ur, "f_\sigma"'] &
\end{tikzcd}$$

Now, by means of Theorem 6.1.4, we find that the morphism f_σ is unique. Notice that f_σ maps the base of $\bigotimes_{i=1}^p V_i$ to the base of $\bigotimes_{i=1}^p V_{\sigma(i)}$, hence f_σ is an isomorphism. ♮

Duality

Proposition 6.2.3. Let $\{V_i\}_{i=1}^p$ be a collection of finite dimensional k-vector spaces. Then the k-linear morphism

$$\bigoplus_{i=1}^{p} V_i^* \xrightarrow{\sim} \left[\bigoplus_{i=1}^{p} V_i\right]^* \text{ mapping } f_1 \otimes \cdots \otimes f_p \longmapsto (v_1 \otimes \cdots \otimes v_p \mapsto f_1(v_1) \ldots f_p(v_p))$$

is a natural isomorphism.

Proof. Since V_i is finite dimensional for all i, then $V_i \simeq V_i^*$, which in particular yield $\dim_k(V_i) = \dim_k(V_i^*)$ (Proposition 5.7.3), then $\dim_k(\bigoplus_i V_i^*) = \dim_k(\bigoplus_i V_i)$. Now, since $\bigotimes_i V_i$ is finite, then $\bigotimes_i V_i \simeq (\bigotimes_i V_i)^*$, which implies in $\dim_k(\bigotimes_i V_i) = \dim_k(\bigotimes_i V_i)^*$. From this, $\dim_k(\bigoplus V_i^*) = \dim_k(\bigoplus_i V_i)^*$. From Corollary 5.6.25 it suffices to show that the map is surjective or injective. Let $f_1 \otimes \cdots \otimes f_p \neq 0$, then its image maps to $f_1(v_1) \ldots f_p(v_p)$, which cannot be zero if $v_1 \otimes \cdots \otimes v_p \neq 0$. This implies that the kernel of the k-linear map shown is zero, hence it's an isomorphism. ♮

Proposition 6.2.4. Let V, L be finite dimensional k-vector spaces. Then the k-linear morphism

$$V^* \otimes L \xrightarrow{\sim} \mathrm{Mor}(V, L) \text{ mapping } f \otimes \ell \longmapsto (v \mapsto f(v)\ell)$$

is a canonical isomorphism.

Proof. Name the above morphism ϕ. Let $\dim_k(V) = n$ and $\dim_k(L) = m$. Let basis $\{v_j\}_{j=1}^n$ and $\{\ell_i\}_{i=1}^m$ of V and L, respectively. Then we find the corresponding dual basis

$\{v_j^*\}_{j=1}^n$ and $\{\ell_i^*\}_{i=1}^m$. Notice that $\phi\colon v_j^* \otimes \ell_i \mapsto (v \mapsto v_j^*(v)\ell_i)$, hence, given $g \in \mathrm{im}(\phi) \subseteq \mathrm{Mor}(V, L)$ we can write its matrix representation $k^n \to k^m$ with factors a_{ij} defined by

$$g(v_k) = \sum_{i=1}^m a_{ik}\ell_i = v_j^*(v_k)\ell_i = \begin{cases} \ell_i, & k = j \\ 0, & \text{otherwise} \end{cases}$$

for some $(v_j^*, \ell_i) \in \{v_j^*\}_{j=1}^m \times \{\ell_i\}_{i=1}^m$, so that $a_{ik} = 0$ for all $k \neq j$ and $a_{ij} = 1$. Notice that this makes $\mathrm{im}(\phi)$ a basis for $\mathrm{Mor}(V, L)$, transforming a basis into other, which qualifies ϕ as an isomorphism. ♮

Notice that if V is finite dimensional k-vector space, we can consider the special case of the endomorphism $\mathrm{End}(V) \simeq V^* \otimes V$. Notice that $\mathrm{id}_V \in \mathrm{End}(V)$ is such that, given $v_j \in \{v_i\}_{i=1}^n$, where the last is a basis for V, we have $\mathrm{id}_V(v_j) = \sum_{i=1}^n \delta_{ij} v_i = \sum_{i=1}^n v_i^*(v_j) v_i$. This implies in the mapping $\mathrm{id}_V \mapsto \sum_{i=1}^n v_i^* \otimes v_i$ for the canonical isomorphism.

Definition 6.2.5 (Trace)**.** Given a finite k-vector space V, we define the canonical linear functional

$$\mathrm{tr}\colon V^* \otimes V \to k,\ \alpha \otimes v \mapsto \alpha(v)$$

Since $V^* \otimes V \simeq \mathrm{End}(V)$, then we can view the above definition in terms of endomorphisms, so that $\mathrm{tr}\colon \mathrm{End}(V) \to k$. For a more computational interpretation, let $\dim_k V = n$ and $\{v_i\}_{i=1}^n$ be a basis for V. Given a linear endomorphism $f \in \mathrm{End}(V)$, suppose that it's coefficients in the matrix representation are a_{ij}, so that for any $v_k \in \{v_i\}_{i=1}^n$ we have $f(v_k) = \sum_{i=1}^n a_{ij} v_i = \sum_{i,j=1}^n a_{ij} v_j^*(v_k) v_i$. From the canonical map described in Proposition 6.2.4 we find that $f \mapsto \sum_{i,j=1}^n a_{ij} v_j^* \otimes v_i$. Now, from the definition of the trace it follows that

$$f \longmapsto \sum_{i,j=1}^n a_{ij} v_j^* \otimes v_i \overset{\mathrm{tr}}{\longmapsto} \sum_{i=1}^n a_{ii}.$$

As one can note, this is the sum of the diagonal of the matrix representation of f (which happens to be independent of the base, as can be asserted from the more abstract definition).

Corollary 6.2.6. Given finite dimensional vector spaces V, L, W we have that

$$\mathrm{Mor}(V \otimes L, W) \simeq \mathrm{Mor}(V, \mathrm{Mor}(L, W)).$$

Proof. Notice that from Proposition 6.2.4 we find

$$\begin{aligned} \mathrm{Mor}(V \otimes L, W) \simeq (V \otimes L)^* \otimes W &\simeq (V^* \otimes L^*) \otimes W \\ &\simeq V^* \otimes (L^* \otimes W) \\ &\simeq V^* \otimes \mathrm{Mor}(L, W) \\ &\simeq \mathrm{Mor}(V, \mathrm{Mor}(L, W)) \end{aligned}$$

♮

Corollary 6.2.7 (Tensor product of morphisms)**.** Let $\{V_i\}_{i\in I}$ and $\{L_i\}_{i\in I}$ be a finite collection of finite k-vector spaces. Then the k-linear morphism

$$\bigotimes_{i\in I} \mathrm{Mor}(V_i, L_i) \to \mathrm{Mor}\left(\bigotimes_{i\in I} V_i, \bigotimes_{i\in I} L_i\right), \ \otimes_{i\in I} f_i \mapsto (\otimes_{i\in I} v_i \mapsto \otimes_{i\in I} f_i(v_i))$$

is a canonical isomorphism.

Proof. First notice that

$$\begin{aligned}\mathrm{Mor}\left(\bigotimes_i V_i, \bigotimes_i L_i\right) \simeq \left(\bigotimes_i V_i\right)^* \otimes \left(\bigotimes_i L_i\right) &\simeq \left(\bigotimes_i V_i^*\right) \otimes \left(\bigotimes_i L_i\right)\\ &\simeq \bigotimes_i V_i^* \otimes L_i\\ &\simeq \bigotimes_i \mathrm{Mor}(V_i, L_i)\end{aligned}$$

from propositions 6.2.4, and 6.2.3, and 6.2.2. Moreover, the map is clearly injective and surjective, hence an isomorphism. ♮

Contractions, and Raising (Lowering) of Indices

Definition 6.2.8 (Contraction)**.** Let $\{V_i\}_{i\in I}$ be a finite collection of finite k-vector spaces such that for some $k, j \in I$ we have $V_k = V$ and $V_j = V^*$. We define the contraction of the tensor product $\bigotimes_{i\in I} V_i$ as the linear mapping

$$\bigotimes_{i\in I} V_i \longrightarrow \bigotimes_{\substack{i\in I\\ i\neq i,j}} V_i \ \text{ mapping } \ \otimes_{i\in I} v_i \longmapsto v_j^*(v_k)(\otimes_{\substack{i\in I\\ i\neq j,k}} v_i)$$

where $v_j^*(v_k) = \delta_{jk}$.

Definition 6.2.9 (Raising and Lowering)**.** Let $\{V_i\}_{i=1}^p$ be a finite collection of finite k-vector spaces and $g\colon V_i \to V_i^*$ be an isomorphism. Then we define the lowering of the index i as the linear morphism

$$\mathrm{id} \otimes \cdots \otimes g \otimes \cdots \otimes \mathrm{id}\colon \bigotimes_{i=1}^p V_i \longrightarrow V_1 \otimes \cdots \otimes V_i^* \otimes \cdots \otimes V_p$$

Moreover, the raising of the index i is just defined as the inverse of the above linear morphism.

Tensor Multiplication Functor

Definition 6.2.10 (Tensor multiplication functor)**.** Let $\mathtt{FinVect}_k$ be the category of finite k-vector spaces together with linear morphisms between them. Given $M \in \mathtt{FinVect}_k$, we define the functor of tensor multiplication on M as the mapping of objects $L \overset{F}{\mapsto} L\otimes M$ and the mapping of morphisms $f \overset{F}{\mapsto} f \otimes \mathrm{id}_M$. Hence we have $\mathrm{id}_L \mapsto \mathrm{id}_L \otimes \mathrm{id}_M = \mathrm{id}_{L\otimes M}$ and given composable morphisms $f, g \in \mathrm{Mor}(\mathtt{FinVect}_k)$ we have $fg \mapsto (fg) \otimes \mathrm{id}_M = (f \otimes \mathrm{id}_M)(g \otimes \mathrm{id}_M)$ as wanted.

Proposition 6.2.11 (Exactness)**.** Let $0 \to V \overset{f}{\rightarrowtail} L \overset{g}{\twoheadrightarrow} W \to 0$ be a short exact sequence, where $V, L, W \in \mathtt{FinVect}_k$. Let $M \in \mathtt{FinVect}_k$, then following sequence is exact

$$0 \to V \otimes M \xrightarrow{f\otimes \mathrm{id}_M} L \otimes M \xrightarrow{g\otimes \mathrm{id}_M} W \otimes W \to 0.$$

Proof. ($f \otimes \mathrm{id}_M$ is injective) Let $v \otimes m \in V \otimes M$ be a non-zero factorizable tensor. Then in particular $f(v) \neq 0$, since f is injective, then $f(v) \otimes m$ is also non-zero, since $m \neq 0$, hence $\ker(f \otimes \mathrm{id}_M) = 0$. ($g \otimes \mathrm{id}_M$ is surjective) Let $w \otimes m \in W \otimes M$ be any factorizable tensor, then in particular exists $\ell \in L$ such that $g(\ell) = w$, from the fact that g is surjective. Hence we find that $\ell \otimes m \xmapsto{g\otimes \mathrm{id}_M} g(\ell) \otimes m = w \otimes m$, since the collection of factorizable tensors form a base for the tensor product, we can conclude that $g \otimes \mathrm{id}_M$ is surjective. ($\mathrm{im}(f \otimes \mathrm{id}_M) = \ker(g \otimes \mathrm{id}_M)$) Suppose $\ell \otimes m \in \mathrm{im}(f \otimes \mathrm{id}_M)$, then in particular we have $\ell \in \mathrm{im}\, f$ and hence $\ell \in \ker g$ since $\mathrm{im}\, f = \ker g$. This implies in $\ell \otimes m \in \ker(g \otimes \mathrm{id}_M)$ and hence $\mathrm{im}(f \otimes \mathrm{id}_M) \subseteq \ker(g \otimes \mathrm{id}_M)$. Take now any tensor $\ell' \otimes m' \in \ker(g \otimes m)$ if $\ell' \otimes m' = 0$ then clearly $\ell' \otimes m' \in \mathrm{im}(f \otimes \mathrm{id}_M)$, suppose on the contrary that $\ell' \otimes m' \neq 0$, then certainly $\ell' \in \ker g$ and in particular $\ell' \in \mathrm{im}\, f$. Then it follows that $\ell' \otimes m' \in \mathrm{im}(f \otimes \mathrm{id}_M)$ and hence $\ker(g \otimes \mathrm{id}_M) \subseteq \mathrm{im}(f \otimes \mathrm{id}_M)$. This finishes the proof. ♮

6.3 Tensor Algebra

Definition 6.3.1 (Mixed tensor product)**.** Let V be a finite dimensional k-vector space. We define the p-covariant and q-contravariant mixed tensor product on V as

$$T_p^q(V) = V^{*\otimes p} \otimes V^{\otimes q}$$

Elements of such object are called tensors of type (p, q) and rank $p + q$ on V. We define $T_0^0(V) = k$.

Example 6.3.2. The following examples illustrate some specific type of mixed tensor product on V, showing that such a construction generalizes various objects in linear algebra.

(a) Tensors of type $(0, 0)$ are called scalar tensors of rank 0.

(b) Tensors of type $(1, 0)$ are linear functionals on V.

(c) Tensors of type $(0, 1)$ are vectors of V.

(d) Tensors of type $(1, 1)$ are elements of $V^* \otimes V \simeq \mathrm{End}_{\mathtt{Vect}_k}(V)$, that is, linear operators.

(e) Tensors of type $(2,0)$ are elements of $V^* \otimes V^* \simeq \mathrm{Mor}(V^{**}, V^*) \simeq \mathrm{Mor}(V, V^*)$ for a finite dimensional V. Moreover $V^* \otimes V^* \simeq (V \otimes V)^* \simeq \mathrm{Mor}(V, V; k)$ of multilinear maps $V \times V \to k$, that is, the mixed tensors of type $(2,0)$ are inner products.

Definition 6.3.3 (Mixed tensor multiplication)**.** Let V be a finite k-vector space, then

$$T_p^q(V) = V^{*\otimes p} \otimes V^{\otimes q} \simeq (V^{\otimes p} \otimes V^{*\otimes q})^*,$$

which in turn is isomorphic to the space of multilinear maps $V^p \times V^{*q} \to k$. Let $f: V^p \times V^{*q} \to k$ and $g: V^{p'} \times v^{*q'} \to k$ be multilinear maps, then we define their tensor multiplication as

$$f \otimes g: V^{p+p'} \times V^{*(q+q')} \to k$$

mapping $(v_1, \dots, v_p, v'_1, \dots, v'_{p'}, u^*_1, \dots, u^*_p, u'^*_1, \dots, u'^*_{q'})$ into

$$f(v_1, \dots, v_p, u^*_1, \dots, u^*_p) g(v'_1, \dots, v'_{p'}, u'^*_1, \dots, u'^*_{q'}),$$

where $v_i, v'_i \in V$ and $u^*_j, u'^*_j \in V^*$. This shows that clearly this tensor multiplication is, in general, non-commutative. However, it is

- (Bilinear) For all $a, b \in k$, then

  $$(af_1 + bf_2) \otimes g = a(f_1 \otimes g) + b(f_2 \otimes g) \text{ and } f \otimes (ag_1 + bg_2) = a(f \otimes g_1) + b(f \otimes g_2).$$

- (Associative) $(f \otimes g) \otimes h = f \otimes (g \otimes h)$.

Definition 6.3.4 (Tensor algebra)**.** We define the tensor algebra of the k-vector space V to be the infinite dimensional k-vector space

$$T(V) = \bigotimes_{p,q=1}^{\infty} T_p^q(V).$$

6.4 Symmetric Tensors and Symmetric Algebra

Definition 6.4.1 (Symmetric tensor)**.** Let V be a k-vector space and consider the mixed tensor product $T_0^q(V)$. For every permutation $\sigma \in S_q$ (where S_q denotes the symmetry group of q elements), define the linear transformation

$$f_\sigma: T_0^q(V) \to T_0^q(V),\ \otimes_{i=1}^q v_i \xmapsto{f_\sigma} \otimes_{i=1}^q v_{f_\sigma(i)}.$$

We say that a tensor $T \in T_0^q(V)$ is symmetric if for all $\sigma \in S_q$

$$f_\sigma(T) = T.$$

Definition 6.4.2 (Symmetric tensor space)**.** We denote the subspace of $T_0^q(V)$ consisting of symmetric tensors as the symmetric space $\mathrm{Sym}^q(V)$, also called symmetric power.

Definition 6.4.3 (Symmetrization map)**.** Let V be a k-vector space. We define the projection operator $S\colon T_0^q(V) \twoheadrightarrow \mathrm{Sym}^q(V)$, which maps factorizable tensors to symmetric tensors.

Proposition 6.4.4 (Symmetric space universal property)**.** Let V be a k-vector space. Given any k-vector space L and a multilinear symmetric map $\mu\colon V^q \to L$, there exists a unique linear map $\ell\colon \mathrm{Sym}^q(V) \to L$ such that $\ell \circ S \circ \otimes = \mu$. That is, the following diagram commutes

$$\begin{tikzcd}
V^q \arrow[r, "\mu"] \arrow[d, "\otimes"] & L \\
T_0^q(V) \arrow[r, two heads, "S"] & \mathrm{Sym}^q(V) \arrow[u, dashed, "\ell"']
\end{tikzcd}$$

Proof. Since $\mathrm{Sym}^q(V)$ is a subspace of $T_0^q(V)$, then we can use Theorem 6.1.4. ♮

Proposition 6.4.5. For $\mathrm{char}\, k \nmid q!$, we have that

$$S(T) = \frac{1}{q!} \sum_{\sigma \in S_q} f_\sigma(T)$$

and hence $S^2 = S$, and $S(T_0^q(V)) = \mathrm{Sym}^q(V)$.

Proof. Let $\{e_i\}_{i=1}^n$ be a basis for V. First, notice that if $T = T^{i_1,\dots,i_q} \in \mathrm{Sym}^q(V)$, then

$$S(T) = \frac{1}{q!} \sum_{\sigma \in S_q} f_\sigma(T^{i_1,\dots,i_q}) = \frac{1}{q!}(q! T^{i_1,\dots,i_q}) = T$$

where we assumed that $\mathrm{char}\, k \nmid q!$ in order to obtain a non-zero tensor after the summation. Hence $S|_{\mathrm{Sym}^q(T)} = \mathrm{id}_{\mathrm{Sym}^q(V)}$ therefore we find that $\mathrm{Sym}^q(V) \subseteq \mathrm{im}\, S$. Moreover, let $T = T^{i_1,\dots,i_q} \in T_0^q(V)$ be any factorizable tensor, then its image under S is clearly a symmetric tensor, that is $S(T) \in \mathrm{Sym}^q(V)$, hence $\mathrm{im}\, S \subseteq \mathrm{Sym}^q(V)$. This implies that $\mathrm{im}\, S = \mathrm{Sym}^q(V)$ for $\mathrm{char}\, k \nmid q!$. Moreover, $S(S(T)) = S(T)$ from the first argument, since $S(T) \in \mathrm{Sym}^q(V)$. ♮

Remark 6.4.6. From now on we are going to assume that $\mathrm{char}\, k \nmid q!$

Notation 6.4.7. Notice that since $\mathrm{im}\, S = \mathrm{Sym}^q(V)$ then we can write any symmetrized tensor in any permutation that we choose, hence, we sometimes use the dot notation $\mathrm{Sym}^q(V) \ni v_1 \otimes \cdots \otimes v_q = v_1 \cdot \ldots \cdot v_q$. Moreover, if $\{e_i\}_{i=1}^n$ is a basis for V, one can even adopt an exponential notation $e_1 \cdot \ldots \cdot e_q = e_1^{a_1} \cdot \ldots \cdot e_q^{a_n}$, where a_i denotes the number of times the component e_i appears in the factorizable tensor, and $a_1 + \cdots + a_q = q$.

Proposition 6.4.8 (Basis for $\mathrm{Sym}^q(V)$)**.** Let $\{e_i\}_{i=1}^n$ be a basis for the k-vector space V. Then, the collection of tensors $\{e_1^{a_1} \cdot \ldots \cdot e_n^{a_n} : a_1 + \cdots + a_n = q\} \subseteq \mathrm{Sym}^q(V)$ form a basis for the symmetric space $\mathrm{Sym}^q(V)$. This implies that $\mathrm{Sym}^q(V)$ is the subspace of $k[e_1, \dots, e_n]$ of homogeneous polynomials of total degree q.

Proof. We know from Lemma 6.1.8 that $\{\otimes_{j=1}^q e_{i_j} : 1 \leqslant i_j \leqslant n\}$ forms a basis for $T_0^q(V)$. Since S is multilinear, then $S(\{\otimes_{j=1}^q e_{i_j}\}) = \mathrm{Sym}^q(V)$, that is $\{e_{i_1}^{a_1} \cdot \ldots \cdot e_n^{a_n}\} = \{S(\otimes_{j=1}^q e_{i_j})\}$ generates the space $\mathrm{Sym}^q(V)$. Hence, to show that the collection of symmetric tensors of the basis elements of V forms a basis for the symmetric space, we just need to show their linear independence.

To do that, fix a sequence of indices $I = (i_1, \ldots, i_q)$ such that $1 \leqslant i_1 \leqslant \ldots \leqslant i_q \leqslant n$. For all $v_j = \sum_{i=1}^n a_{ij} e_i \in V$ where $1 \leqslant j \leqslant q$, let $\mu_I \colon V^q \to k$ be the mapping

$$\mu_I(v_1, \ldots, v_n) = \sum_{\sigma \in S_q} \prod_{j=1}^q e^*_{i_{\sigma(j)}}(v_j).$$

From the universal property of tensor products we find a unique functional $f_I \colon \mathrm{Sym}^q(V) \to k$ such that the diagram commutes:

$$\begin{tikzcd} V^q \arrow[r, "\mu_I"] \arrow[d, "S\circ\otimes"'] & k \\ \mathrm{Sym}^q(V) \arrow[ur, dashed, "f_I"'] & \end{tikzcd}$$

hence for all $(v_1, \ldots, v_q) \in T_0^q(V)$ we have $\mu(v_1, \ldots, v_q) = f(v_1 \cdot \ldots \cdot v_n)$. Notice that if we have a monotonically increasing sequence of indices $I' = (i'_1, \ldots, i'_q) \neq I$, there must be an index $i'_{j_0} \in I'$ such that $i'_{j_0} \neq i_j$ for all $i_j \in I$. This way we find that for all $\sigma \in S_q$ the product $\prod_{j=1}^q e^*_{i_{\sigma(j)}}(e_{i'_j}) = 0$ because $e^*_{i_{\sigma(j)}}(e_{i'_{j_0}}) = 0$. In particular, this implies in

$$\mu_I(e_{i'_1}, \ldots, e_{i'_q}) = f_I(e_{i'_1} \cdot \ldots \cdot e_{i'_q}) = 0.$$

Lets compute the value of $f_I(e_{i_1} \cdot \ldots \cdot e_{i_q})$. To do this, consider again the fixed monotone increasing sequence of indices that we started with, that is, I. Let $Q \leqslant q$ denote the number of distinct values of $i_j \in I$ for $1 \leqslant j \leqslant q$. Define now, for all $1 \leqslant r \leqslant Q$ the index sets $J_r := \{j : e_{i_j} \text{ have equal values}\}$. Define $n_r := |J_r|$, this construction yields a total of $\prod_{r=1}^Q (n_r!)$ permutations. Notice that all permutations of J_r's are such that $\sigma_{J_r}(I) = I$ (they change the position of equal elements, causing no alteration of the sequence), we'll denote such permutations by $\sigma_J \in S_q$. Moreover, these are all the possible permutations that leave I unaltered. Hence

$$f_I(e_{i_1} \cdot \ldots \cdot e_{i_q}) = \mu_I(e_{i_1}, \ldots, e_{i_q}) = \sum_{\sigma \in S_q} \prod_{j=1}^q e^*_{i_{\sigma(j)}}(e_j) = \sum_{\sigma_J \in S_q} \prod_{j=1}^q e^*_{i_{\sigma_J(j)}}(e_j) = \prod_{r=1}^Q (n_r!) \in k.$$

Therefore, if $\operatorname{char} k \nmid q!$, we find $f_I(e_{i_1} \cdot \ldots \cdot e_{i_q}) \neq 0 \in k$.

If there exists a linear relation

$$\sum_{i'_1, \ldots, i'_q} c_{i'_1, \ldots, i'_q}(e_{i'_1} \cdot \ldots \cdot e_{i'_q}) = 0$$

where the indices are arranged in monotonic increasing order, then we conclude that

$$0 = f_I\left(\sum_{i'_1,\dots,i'_q} c_{i'_1,\dots,i'_q}(e_{i'_1}\cdot\dots\cdot e_{i_q})\right) = c_{i'_1,\dots,i'_q}\sum_{\sigma\in S_q}\prod_{j=1}^{q} e^*_{i_{\sigma(j)}}(e_{i'_j})$$
$$= c_{i_1,\dots,i_q}\sum_{\sigma\in S_q}\prod_{j=1}^{q} e^*_{i_{\sigma(j)}}(e_j)$$
$$= c_{i_1,\dots,i_q} f_I(e_{i_1}\cdot\dots\cdot e_{i_q})$$

hence $c_{i_1,\dots,i_q} = 0$ since, as proved above, $f_I(e_{i_1}\cdot\dots\cdot e_{i_q}) \neq 0$. Now, passing I through all the monotone increasing sequences of indices, we prove that all of the coefficients $c_{i'_1,\dots,i'_c}$ vanish, as wanted. This proves that $\{e_1^{a_1}\cdot\dots\cdot e_n^{a_n} : a_1+\dots+a_n = q\}$ is indeed a linearly independent collection. ♮

Corollary 6.4.9. If V is a finite n-dimensional k-vector space, then

$$\dim_k(\mathrm{Sym}^q(V)) = \binom{n+q-1}{q}$$

Proof. Notice that from Proposition 6.4.8 we have that $\mathrm{Sym}^q(V)$ is the subspace of $k[e_1,\dots,e_n]$ consisting of homogeneous polynomials of degree q. In order to find the dimension of such space, it is sufficient to find the cardinality $|\{a_1+\dots+a_n = q : a_i \geqslant 0\}|$. A combinatorial way to do so is to define q symbols $|$ (which represent the unity 1), which can be divided into n different groups. In order to do so, we introduce $n-1$ symbols (we are going to use +), which are arranged in order to separate the bars into the n requested groups (example: for $n = 3$ and $q = 5$, one possible arrangement is $|+||+||$). We have a total of $n+q-1$ symbols, and we want q of those (the bars) in order to create the groupings, that is, we have a total of $\binom{n+q-1}{q}$ ways of doing so. This proves the corollary. ♮

Definition 6.4.10. We define the symmetric algebra over a finite n-dimensional k-linear space V to be the object $\mathrm{Sym}\, V = \bigoplus_{q=1}^{\infty}\mathrm{Sym}^q V$. This is the space of all polynomials in $e_1,\dots,e_n$ variables (where $e_1,\dots,e_n$ form a basis for V).

We define multiplication in $\mathrm{Sym}\, V$ as the map

$$\mathrm{Sym}^q V\times\mathrm{Sym}^d V\longrightarrow\mathrm{Sym}^{q+d} V,\ \text{ mapping }\ (T_1,T_2)\mapsto S(T_1\otimes T_2) =: T_1T_2 \tag{6.1}$$

Proposition 6.4.11. The multiplicative structure defined by Eq. (6.1) makes $\mathrm{Sym}\, V$ a commutative associative k-algebra. Under the isomorphism of tensors and homogeneous polynomials as described in Proposition 6.4.8, this multiplicative structure preserves that of the multiplication of polynomials, making such isomorphism an isomorphism of algebras.

$$\{\text{polynomials in } e_1,\dots,e_n\}\simeq\mathrm{Sym}\, V$$

where $e_1,\dots,e_n$ forms a basis for V.

Proof. Let $T_1 \in T_0^q(V)$ and $T_2 \in T_0^d(V)$, then we find that

$$S(T_1) \otimes T_2 = \left(\frac{1}{q!} \sum_{\sigma \in S_q} f_\sigma(T_1) \right) \otimes T_2$$

This way we find that

$$\begin{aligned} S(S(T_1) \otimes T_2) = S\left(\frac{1}{q!} \sum_{\sigma \in S_q} f_\sigma(T_1) \otimes T_2 \right) &= \frac{1}{q!} \sum_{\sigma \in S_q} S(f_\sigma(T_1) \otimes T_2) \\ &= \frac{1}{q!} \sum_{\sigma \in S_q} S(T_1 \otimes T_2) \\ &= S(T_1 \otimes T_2) \end{aligned}$$

where we used the fact that $S(f_\sigma(T_1) \otimes T_2) = S(T_1 \otimes T_2)$. Hence, we can conclude in general that

$$S(S(T_1) \otimes T_2) = S(T_1 \otimes S(T_2)) = S(T_1 \otimes T_2)$$

If we extend such argument for tensors T_1, T_2, T_3, we find that

$$(T_1T_2)T_3 = S(S(T_1 \otimes T_2) \otimes T_3) = T_1(T_2T_3) = S(T_1 \otimes (T_1 \otimes T_3)) = T_1T_2T_3 = S(T_1 \otimes T_2 \otimes T_3)$$

which proves associativity for the multiplication. From tensor commutativity we find that

$$T_1T_2 = S(T_1 \otimes T_2) = S(T_2 \otimes T_1) = T_2T_1$$

Finally, notice that from this multiplicative structure we have that the multiplication of polynomials is given by $(e_1^{a_1} \cdots e_n^{a_n})(e_n^{b_1} \cdots e_n^{b_n}) = e_1^{a_1+b_1} \cdots e_n^{a_n+b_n}$, which shows the algebra isomorphism. ♮

6.5 Alternating Tensors and Exterior Powers

Definition 6.5.1 (Alternating map)**.** Let V and W be vector spaces. A multilinear map $f\colon V^n \to W$ is called alternating if for any $(v_1, \ldots, v_n) \in V^n$ such that exists $i < j$ for which $v_i = v_j$ then

$$f(v_1, \ldots, v_n) = 0.$$

Proposition 6.5.2. Let $f\colon V^n \to W$ be an alternating multilinear map. Then if $\tau \in S_n$ is a transposition, we have

$$f(v_{\tau(1)}, \ldots, v_{\tau(n)}) = -f(v_1, \ldots, v_n).$$

In general, if $\sigma \in S_n$ is any permutation, then

$$f(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \mathrm{sign}(\sigma) f(v_1, \ldots, v_n).$$

Proof. Let τ be a transposition between indices $i < j$, then consider

$$\begin{aligned}0 = f(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) &= f(v_1, \dots, v_i, \dots, v_i + v_j, \dots, v_n) \\ &\quad + f(v_1, \dots, v_j, \dots, v_i + v_j, \dots, v_n)\end{aligned}$$

therefore we find

$$\begin{aligned}f(v_1, \dots, v_i, \dots, v_i + v_j, \dots, v_n) &= -f(v_1, \dots, v_j, \dots, v_i + v_j, \dots, v_n) \\ f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) &= -f(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \\ f(v_1, \dots, v_n) &= f(v_{\tau(1)}, \dots, v_{\tau(n)})\end{aligned}$$

where we used that $f(v_1, \dots, v_i, \dots, v_i, \dots, v_n) = f(v_1, \dots, v_j, \dots, v_j, \dots, v_n) = 0$. This shows the first proposition. For the second proposition, by means of Proposition 7.2.14 and Proposition 7.2.8 and the above proposition for transpositions, we conclude the proof. ♮

Remark 6.5.3. Notice that Proposition 6.5.2 is not sufficient to characterize alternating maps, it is but a necessary property. A counterexample to that is the field with characteristic 2.

Alternating Tensor and Exterior Powers

Definition 6.5.4 (Alternating tensor)**.** Let V be a k-vector space we define a tensor $T \in T_0^q(V)$ to be alternating if for all $\sigma \in S_q$ permutation we have

$$f_\sigma(T) = \operatorname{sign}(\sigma) T.$$

We denote by $\Lambda^q V$ the subspace of $T_0^q(V)$ of alternating tensors and call it the q-exterior power of V.

Proposition 6.5.5 (Alternating tensor projection)**.** Let V be a k-vector space and $\operatorname{char} k \nmid q!$. We define the linear projection operator $A \colon T_0^q(V) \twoheadrightarrow T_0^q(V)$ where

$$A(T) = \frac{1}{q!} \sum_{\sigma \in S_q} \operatorname{sign}(\sigma) f_\sigma(T)$$

Then $A^2 = A$ and $\operatorname{im} A = \Lambda^q V$.

Proof. First we show that $\operatorname{im} A \subseteq \Lambda^q V$ (we already know that clearly $\Lambda^q V \subseteq \operatorname{im} A$). Notice that

$$\begin{aligned}f_\sigma(A(T)) &= f_\sigma\left(\frac{1}{q!} \sum_{\tau \in S_q} \operatorname{sign}(\tau) f_\tau(T)\right) \\ &= \frac{1}{q!} \sum_{\tau \in S_q} \operatorname{sign}(\tau) f_{\sigma\tau}(T) \\ &= \operatorname{sign}(\sigma)\left(\frac{1}{q!} \sum_{\tau \in S_q} \operatorname{sign}(\sigma\tau) f_{\sigma\tau}(T)\right) \\ &= \operatorname{sign}(\sigma) A(T)\end{aligned}$$

where we've used Proposition 7.2.14. Hence we conclude that $\operatorname{im} A = \Lambda^q V$. For the second proposition, notice that

$$A^2 = \frac{1}{q!^2} \sum_{\sigma,\tau \in S_q} \operatorname{sign}(\sigma\tau) f_{\sigma\tau} = \frac{1}{q!} \sum_{\rho \in S_q} \operatorname{sign}(\rho) f_\rho = A$$

since any permutation $\rho \in S_q$ can be represented in $q!$ different ways as a form of a product $\sigma\tau$. This concludes the proof. ♮

Definition 6.5.6 (Exterior multiplication)**.** Let V be a k-vector space and $v_1 \otimes \cdots \otimes v_q \in T_0^q(V)$. We define the exterior multiplication to be the map

$$A(v_1 \otimes \cdots \otimes v_q) \coloneqq v_1 \wedge \cdots \wedge v_q$$

Proposition 6.5.7 (Universal property for exterior power)**.** Let V be a k-vector space. For all k-vector spaces L together with an alternating multilinear map $\mu\colon V^q \to L$, there exists a unique k-linear map $\ell\colon \Lambda^q V \to L$ for which the diagram commutes

$$\begin{tikzcd} V^q \arrow[r, "\mu"] \arrow[d, "\otimes"'] & L \\ T_0^q(V) \arrow[r, two heads, "A"'] & \Lambda^q V \arrow[u, dashed, "\ell"'] \end{tikzcd}$$

Proof. Since $\Lambda^q V$ is a subspace of $T_0^q(V)$, then we can use the universal property Theorem 6.1.4. ♮

Now we get some geometric motivation behind the algebraic structure of the exterior power $\Lambda^q V$. Alternating tensors $v_1 \wedge \cdots \wedge v_q \in \Lambda^q V$ can be seen as q-dimensional oriented volume elements, where by oriented we mean that the transposition of two edges, say v_j and v_i, implies in a change up to a minus sign of the value (the sign is what creates the bridge between orientation and wedge product). Notice that when there are equal edges, the volume becomes malformed and $v_1 \wedge \cdots \wedge v_q = 0$.

Proposition 6.5.8. Let V be a n-dimensional k-vector space, where $\operatorname{char} k \neq 2$. Define $\{e_i\}_{i=1}^n$ to be a base for V. Then the exterior product $e_{i_1} \wedge \cdots \wedge e_{i_q} = 0$ if there exists $i_a = i_b$ for some $1 \leqslant a, b \leqslant q$.

Proof. Since A is an alternating map, we can use Definition 6.5.1 and conclude the proof. ♮

For the next proposition, we proceed in a similar fashion as in Proposition 6.4.8.

Proposition 6.5.9 (Exterior power $\Lambda^q V$ basis)**.** Let $\{e_i\}_{i=1}^n$ be a basis for the k-vector space V. The factorizable tensors

$$A(e_{i_1} \otimes \cdots \otimes e_{i_q}) = e_{i_1} \wedge \cdots \wedge e_{i_q}$$

with $1 \leqslant i_1 < \cdots < i_q \leqslant n$ form a basis for the subspace $\Lambda^q V$.

Proof. Let $B := \{e_{i_1} \wedge \cdots \wedge e_{i_q} : 1 \leqslant i_1 < \cdots < i_q \leqslant n\}$. Since $\{e_{i_1} \otimes \cdots \otimes e_{i_q} : 1 \leqslant i_1 < \dots i_q \leqslant n\}$ generates $T_0^q(V)$ (see Lemma 6.1.8), then clearly B does generate the subspace $\Lambda^q V$. We now show that B is linearly independent.

Denote by $\mathcal{I} := \{I = (i_j)_{j=1}^q : 1 \leqslant i_1 < \cdots < i_q \leqslant n\}$ the set of strictly increasing q-tuples. For each $I := (i_1, \dots, i_q) \in \mathcal{I}$ we define the alternating multilinear map $\mu_I : V^q \to k$ as

$$\mu_I(v_1, \dots, v_q) = \sum_{\sigma \in S_q} \operatorname{sign}(\sigma) \prod_{j=1}^q e^*_{i_{\sigma(j)}}(v_{i_j})$$

where $\{e^*_{i_j}\}_{i_j \in I}$ is the dual of $\{e_{i_j}\}_{i_j \in I}$. Using Proposition 6.5.7 we can define a unique linear map $f_I : \Lambda^q V \to k$ such that the diagram commutes

$$\begin{tikzcd} V^q \arrow[r, "\mu_I"] \arrow[d, "A \circ \otimes"'] & k \\ \Lambda^q V \arrow[ur, dashed, "f_I"'] & \end{tikzcd}$$

which implies in $\mu_I(v_1, \dots, v_q) = f_I(v_1 \wedge \cdots \wedge v_q)$ for all $(v_1, \dots, v_q) \in V^q$.

Let $I' = (i'_1, \dots, i'_q) \in \mathcal{I}$ such that $I' \neq I$. From the strictly ordering of the indices, we conclude that there must exists some $1 \leqslant j_0 \leqslant q$ such that $i'_{j_0} \neq i_j$ for all $1 \leqslant j \leqslant q$. In particular, this implies that

$$f_I(e_{i'_1} \wedge \cdots \wedge e_{i'_q}) = \mu_I(e_{i'_1}, \dots, e_{i'_q}) = \sum_{\sigma \in S_q} \operatorname{sign}(\sigma) \prod_{j=1}^q e^*_{i_{\sigma(j)}}(e_{i'_j}) = 0$$

since $e^*_{i_{\sigma(j)}}(e_{i_{j_0}}) = 0$ for all permutations $\sigma \in S_q$ and $1 \leqslant j \leqslant q$. On the other hand, we have that

$$f_I(e_{i_1} \wedge \cdots \wedge e_{i_q}) = \mu_I(e_{i_1}, \dots, e_{i_q}) = \sum_{\sigma \in S_q} \operatorname{sign}(\sigma) \prod_{j=1}^q e^*_{i_{\sigma(j)}}(e_{i_j}) = \operatorname{sign}(\mathrm{id}) \prod_{j=1}^q e^*_{i_{\mathrm{id}(j)}}(e_{i_j}) = 1$$

Let $c_P \in k$ for all $P \in \mathcal{I}$ and consider the vanishing linear combination

$$\sum_{P \in \mathcal{I}} c_P(e_{p_1} \wedge \cdots \wedge e_{p_q}) = 0. \tag{6.2}$$

Then, if we look at its image under the map f_I for all $I \in \mathcal{I}$ we conclude that

$$0 = f_I\left(\sum_{P \in \mathcal{I}} c_P(e_{p_1} \wedge \cdots \wedge e_{p_q})\right) = \sum_{P \in \mathcal{I}} c_P f_I(e_{p_1} \wedge \cdots \wedge e_{p_q}) = c_I$$

hence we conclude that the linear combination Eq. (6.2) vanishes if and only if each coefficient vanishes. This concludes that B is linearly independent. Hence we've proved that B is a basis for $\Lambda^q V$. ♮

Proposition 6.5.10. Let V be a n-dimensional k-vector space. Then we have

$$\dim_k \Lambda^q V = \binom{n}{q}$$

Proof. Since we want our indices to be strictly increasing, we are left with n elements of which we want to arrange in groups of q elements. Hence the number of possible arrangements is $\binom{n}{q}$ and this concludes the proof. $\natural$

Exterior Algebra

Definition 6.5.11 (Exterior algebra)**.** Let V be a k-vector space. We define the exterior algebra on V as

$$\Lambda^\bullet V = \bigoplus_{q=0}^{\infty} \Lambda^q V$$

together with a multiplicative structure $\wedge\colon \Lambda^d V \otimes \Lambda^q V \longrightarrow \Lambda^{d+q} V$ mapping

$$(v_1 \wedge \cdots \wedge v_d) \otimes (w_1 \wedge \cdots \wedge w_q) \overset{\wedge}{\mapsto} v_1 \wedge \cdots \wedge v_d \wedge w_1 \wedge \cdots \wedge w_q.$$

We interpret $\Lambda^0 V = k$.

Proposition 6.5.12. The multiplicative structure of the tensor algebra $\Lambda^\bullet V$ is associative and skew-commutative, that is, for all $\alpha \in \Lambda^d V$ and $\beta \in \Lambda^q V$ we have $\alpha \wedge \beta = (-1)^{dq} \beta \wedge \alpha$.

Proof. First, we prove that if $T_1 \in T_0^q(V)$ and $T_2 \in T_0^d(V)$ then

$$A(A(T_1) \otimes T_2) = A(T_1 \otimes A(T_2)) = A(T_1 \otimes T_2) \tag{6.3}$$

Notice that

$$A(T_1) \otimes T_2 = \frac{1}{q!} \sum_{\sigma \in S_q} \operatorname{sign}(\sigma) f_\sigma(T_1) \otimes T_2$$

therefore we find

$$A(A(T_1) \otimes T_2) = \frac{1}{q!} \sum_{\sigma \in S_q} \operatorname{sign}(\sigma) A(f_\sigma(T_1) \otimes T_2) \tag{6.4}$$

Moreover, we can construct an injection of the symmetry groups $S_q \rightarrowtail S_{q+d}$ via the mapping $\sigma \mapsto \overline{\sigma}$ where $\overline{\sigma}(i) = \sigma(i)$ for all $i \in \{1, \ldots, q\}$ and $\overline{\sigma}(i) = i$ for all $i \in \{q+1, \ldots, q+d\}$. In particular, we find that $f_\sigma(T_1) \otimes T_2 = f_{\overline{\sigma}}(T_1 \otimes T_2)$ and clearly $\operatorname{sign}(\overline{\sigma}) = \operatorname{sign}(\sigma)$. This implies that

$$A(f_\sigma(T_1) \otimes T_2) = A(f_{\overline{\sigma}}(T_1 \otimes T_2)) = f_{\overline{\sigma}}(A(T_1 \otimes T_2)) = \operatorname{sign}(\overline{\sigma}) A(T_1 \otimes T_2) = \operatorname{sign}(\sigma) A(T_1 \otimes T_2) \tag{6.5}$$

If we substitute Eq. (6.5) in Eq. (6.4) we find

$$A(A(T_1) \otimes T_2) = \frac{1}{q!} \sum_{\sigma \in S_q} \operatorname{sign}^2(\sigma) A(T_1 \otimes T_2) = A(T_1 \otimes T_2).$$

In the same manner we can show that $A(T_1 \otimes A(T_2)) = A(T_1 \otimes T_2)$. Hence Eq. (6.3) holds.

Let $\alpha \in \Lambda^d V, \beta \in \Lambda^q V, \gamma \in \Lambda^p V$. Notice that from the construction of the product map $\wedge$ and Eq. (6.3) we find

$$(\alpha \wedge \beta) \wedge \gamma = \wedge(\wedge(\alpha \otimes \beta) \otimes \gamma) = \wedge(\alpha \otimes \wedge(\beta \otimes \gamma)) = (\alpha \wedge \beta) \wedge \gamma$$

which proves associativity of the tensor algebra.

We now prove skew-commutativity. Let $\alpha \in \Lambda^d V$ and $\beta \in \Lambda^q V$, and a permutation $\sigma \in S_{q+d}$ such that $\sigma(i) = q + d - i + 1$, consisting of dq transpositions. Hence we find that

$$\beta \wedge \alpha = f_\sigma(\alpha \wedge \beta) = \operatorname{sign}(\sigma)(\alpha \wedge \beta) = (-1)^{dq}(\alpha \wedge \beta).$$

♮

Proposition 6.5.13. Let V be a k-vector space and $\{v_i\}_{i=1}^d \subseteq V$. The vectors $v_1, \ldots, v_d$ are linearly independent if and only if

$$\Lambda^d V \ni v_1 \wedge \cdots \wedge v_d \neq 0.$$

Proof. Let $\{v_i\}_{i=1}^d$ be a set of linearly dependent vectors and choose a set of not-all zero scalars $\{c_i\}_{i=1}^d \subseteq k$ such that $\sum_{i=1}^d c_i v_i = 0$. Choose $1 \leqslant j \leqslant d$ such that $c_j \neq 0$ and write $v_j = -\sum_{i=1}^d \frac{c_i}{c_j} v_i$. Then we find that (recall $\mathcal{M}_0$ from our construction of the tensor product)

$$\begin{aligned}
v_1 \wedge \cdots \wedge v_j \wedge \cdots \wedge v_d &= v_1 \wedge \cdots \wedge \left(-\sum_{i=1}^d \frac{c_i}{c_j} v_i\right) \wedge \cdots \wedge v_d \\
&= \left(v_1, \ldots, -\sum_{i=1}^d \frac{c_i}{c_j} v_i, \ldots, v_d\right) + \mathcal{M}_0 \\
&= -\sum_{i=1}^d \frac{c_i}{c_j}(v_1, \ldots, v_i, \ldots, v_d) + \mathcal{M}_0 &(6.6)\\
&= -\sum_{i=1}^d \frac{c_i}{c_j}(v_1 \wedge \cdots \wedge v_i \wedge \ldots v_d) &(6.7)\\
&= 0 &(6.8)
\end{aligned}$$

Where Eq. (6.8) comes from the fact that we have a repeated v_i in the wedge product Eq. (6.7).

Suppose $\{v_i\}_{i=1}^d$ is a linearly independent set. Then from Proposition 5.6.16 we can build $B \supseteq \{v_i\}_{i=0}^d$ such that B is a basis for V. From Proposition 6.5.9, we find that $\mathcal{B} = \{v_{i_1} \wedge \cdots \wedge v_{i_d} : v_{i_j} \in B \text{ and } i_1 < \cdots < i_d\}$ is a basis for $\Lambda^d V$. In particular, $v_1 \wedge \cdots \wedge v_d \in \mathcal{B}$, hence necessarily $v_1 \wedge \cdots \wedge v_d \neq 0$. ♮

Therefore one can trivially see that the kernel of the alternating projection A is simply the collection of all decomposable tensors $v_1 \otimes \cdots \otimes v_d$ such that the collection $\{v_1, \ldots, v_d\} \subseteq V$ is linearly dependent.

Corollary 6.5.14. The kernel of the alternating projection $A\colon T^d_0(V) \twoheadrightarrow \Lambda^d V$ is given by the collection $\{v_1 \otimes \cdots \otimes v_d \in T^d_0(V)\}$ such that $\{v_1 \otimes \cdots \otimes v_d\}$ is linearly dependent on V.

Theorem 6.5.15. Let V be a finite dimensional k-vector space. The map $\phi\colon \Lambda^d V^* \xrightarrow{\simeq} (\Lambda^d V)^*$ given by the mapping

$$f_1 \wedge \cdots \wedge f_d \overset{\phi}{\longmapsto} \left(v_1 \wedge \cdots \wedge v_d \mapsto \sum_{\sigma \in S_d} \operatorname{sign}(\sigma) \prod_{j=1}^{d} f_j(v_{\sigma(j)}) \right)$$

is a k-linear isomorphism.

Proof. First of all, we show that ϕ is a k-linear map. For each $f \in \Lambda^d V^*$, define the map $\psi_f \in (\Lambda^d V)^*$ for which $\phi(f) = \psi_f$. Notice that if $a \in k$ and $f, g \in \Lambda^d(V^*)$, then $\phi(f + ag) = \psi_{f+ag}$. Notice that

$$\begin{aligned}
\psi_{f+ag}(v_1 \wedge \cdots \wedge v_d) &= \sum_{\sigma \in S_d} \operatorname{sign}(\sigma) \prod_{j=1}^{d} (f_j + a g_j)(v_{\sigma(j)}) \\
&= \sum_{\sigma \in S_d} \operatorname{sign}(\sigma) \prod_{j=1}^{d} f_j(v_{\sigma(j)}) + a g_j(v_{\sigma(j)}) \\
&= \sum_{\sigma \in S_d} \operatorname{sign}(\sigma) \prod_{j=1}^{d} f_j(v_{\sigma(j)}) + a \sum_{\sigma \in S_d} \operatorname{sign}(\sigma) \prod_{j=1}^{d} g_j(v_{\sigma(j)}) \\
&= \psi_f + a\psi_g
\end{aligned}$$

Thus $\phi(f + ag) = \phi(f) + a\phi(g)$ and therefore ϕ is indeed linear.

Suppose $\dim V = n$ and let $\{e_j\}_{j=1}^n$ be a base for V. Define the collection of strictly increasing d-tuples $\mathcal{I} = \{I = (i_j)_{j=1}^d : 1 \leqslant i_1 < \cdots < i_d \leqslant n\}$. From Proposition 6.5.9, we know that $\{e_{i_1} \wedge \cdots \wedge e_{i_d} \in \Lambda^d V : i_j \in I \text{ and } I \in \mathcal{I}\}$ is a basis for $\Lambda^d V$. Moreover, $\{e_j^*\}_{j=1}^n$ is a basis for the dual space V^*, thus the collection $\{e_{i_1}^* \wedge \cdots \wedge e_{i_d}^* \in \Lambda^d V : i_j \in I \text{ and } I \in \mathcal{I}\}$ is a basis for $\Lambda^d V^*$.

Let $I \in \mathcal{I}$ be any strictly increasing d-tuple, and define the notation $e_I^* := e_{i_1}^* \wedge \cdots \wedge e_{i_d}^*$, so that

$$e_I^* \overset{\phi}{\longmapsto} \psi_{e_I^*}(v_1 \wedge \cdots \wedge v_d) = \sum_{\sigma \in S_d} \operatorname{sign}(\sigma) \prod_{j=1}^{d} e_{i_j}^*(v_{\sigma(j)}) = \begin{cases} 1, & \text{if } v_1 \wedge \cdots \wedge v_d = e_{i_1} \wedge \cdots \wedge e_{i_d} \\ 0, & \text{otherwise} \end{cases} \tag{6.9}$$

Notice that we mapped a basis of $\Lambda^d V^*$ into a basis of $(\Lambda^d V)^*$, since any map $\psi \in (\Lambda^d V)^*$ can be obtained by a unique linear combination of maps of the collection $\{\psi_{e_I^*} \in (\Lambda^d V)^* : I \in \mathcal{I}\}$ — which, together with Eq. (6.9), makes it a basis. This shows that the morphism ϕ establishes an isomorphism $\Lambda^d V^* \simeq (\Lambda^d V)$. ♮

Add cool propositions for the exterior algebra: the Hodge Star Operator; the isomorphisms

$$\mathrm{Sym}^d(V \oplus W) \simeq \bigoplus_{i=0}^{d} \mathrm{Sym}^i(V) \otimes \mathrm{Sym}^{d-i}(W)$$

$$\Lambda^d(V \oplus W) \simeq \bigoplus_{i=0}^{d} \Lambda^i(V) \otimes \Lambda^{d-i}(W)$$

After creating sections on affine and projective geometry, introduce grassmanian varieties.

6.6 Determinants

Definition 6.6.1. Let $f\colon V \to L$ be a k-linear map. The map f naturally induces a k-linear pullback $f^{\wedge d}\colon \Lambda^d V \to \Lambda^d L$ which is defined by

$$v_1 \wedge \cdots \wedge v_d \xmapsto{f^{\wedge d}} f(v_1) \wedge \cdots \wedge f(v_d).$$

Proposition 6.6.2. Let $f\colon V \to W$ and $g\colon W \to L$ be k-linear maps. Then the composition $gf\colon V \to L$ satisfy

$$(gf)^{\wedge d} = g^{\wedge d} f^{\wedge d}.$$

Proof. Notice that for any given $v = v_1 \wedge \cdots \wedge v_d \in \Lambda^d V$ we have

$$(gf)^{\wedge d}(v) = (gf)(v_1) \wedge \cdots \wedge (gf)(v_d) = g(f(v_1)) \wedge \cdots \wedge g(f(v_2)) = (g^{\wedge d} f^{\wedge d})(v).$$

Therefore $(gf)^{\wedge d} = g^{\wedge d} f^{\wedge d}$. ♮

Definition 6.6.3 (Determinant)**.** Let V be an n-dimensional k-vector space and $f \in \mathrm{End}(V)$. Since $\Lambda^n V$ is 1-dimensional, the induced map $f^{\wedge n} \in \mathrm{End}(\Lambda^n V)$ is a multiplication by some scalar in $\Lambda^n V$. We define the determinant of the k-linear endomorphism f as the element $\det f \in \Lambda^n V$ such that

$$f^{\wedge n}(\omega) = \det(f)\omega.$$

Proposition 6.6.4 (Composition determinant)**.** Let $f, g \in \mathrm{End}(V)$ where V is a n-dimensional vector space, then $\det(gf) = \det g \det f$.

Proof. From Proposition 6.6.2, $(gf)^{\wedge n}(\omega) = g^{\wedge n}(f^{\wedge n}(\omega))$. From Definition 6.6.3,

$$\det(gf)\omega = \det(g)\det(f)\omega,$$

thus $\det(gf) = \det(g)\det(f)$. ♮

Proposition 6.6.5 (Identity determinant)**.** Let V be an n-dimensional vector space. The identity map $\mathrm{id}_V \in \mathrm{End}(V)$ is such that $\det(\mathrm{id}_V) = 1$.

Proof. Notice that $\mathrm{id}_V^{\wedge n}(\omega) = \omega$ and from Definition 6.6.3 we have $\mathrm{id}_V^{\wedge n}(\omega) = \det(\mathrm{id}_V)\omega$. Hence $\det(\mathrm{id}_V) = 1$. ♮

Proposition 6.6.6 (Dual determinant)**.** Let V be a finite dimensional vector space and $f \in \mathrm{End}(V)$. Then $\det f = \det f^*$, where f^* is the dual map of f.

Proof. ♮

Dual determinant proof

Lemma 6.6.7. Let V be an n-dimensional vector space. If $f \in \mathrm{End}(V)$ is not surjective, then $\det f = 0$.

Proof. Let $\operatorname{im} f = W \subsetneq V$. Since f is not surjective, then W is a subspace with $\dim(W) < n$, hence $\Lambda^n W = 0$ because any collection with a number of elements greater than $\dim(W)$ is linearly dependent on W, and from Proposition 6.5.13 their wedge product is equal to zero. Given any non-zero $\omega = v_1 \wedge \cdots \wedge v_n \in \Lambda^n V$ we find that $f^{\wedge n}(\omega) = f(v_1) \wedge \cdots \wedge f(v_n) \in \Lambda^n W = 0$, since $\operatorname{im} f = W$, and therefore $f^{\wedge n}(\omega) = 0$. Moreover, since $f^{\wedge n}(\omega) = \det(f)\omega$ and $\omega \neq 0$, it follows that $\det(f) = 0$. ♮

Proposition 6.6.8 (Isomorphism determinant)**.** Let $f \in \mathrm{End}(V)$ where V is an n-dimensional vector space. Then f is an isomorphism if and only if $\det f \neq 0$.

Proof. Let $\{v_1, \dots, v_n\}$ be a basis for V. From Proposition 6.5.13, $v = v_1 \wedge \cdots \wedge v_n \neq 0$. Suppose first that f is an isomorphism, then $f^{\wedge n}(v) = \det(f)v$ implies $\det f \neq 0$.

Suppose now that $\det f \neq 0$. From Lemma 6.6.7 we find that f is surjective. Let $\omega \in \Lambda^n V$ be such that $f^{\wedge n}(\omega) = 0$, then from the fact that $f^{\wedge n}(\omega) = \det(f)\omega$ and $\det f \neq 0$, it follows that $\omega = 0$. Hence $\ker f^{\wedge n} = 0$ and therefore $f^{\wedge n}$ is injective. This proves that $f^{\wedge n}$ is an isomorphism. ♮

Proposition 6.6.9 (Matrix determinant)**.** Let V be an n-dimensional k-vector space and $f \in \mathrm{End}(V)$. Let $A \colon k^n \to k^n$ be the matrix representation of f and $a_{i,j} \in k$ be the entries of A, where $1 \leqslant i, j \leqslant n$. Then

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \dots a_{\sigma(n),n}.$$

Proof. Let me show you something interesting. Suppose $f \colon V \to L$ (the specified case for Proposition 6.6.9 is $L = V$). I want to show you that in order to define the determinant we'll need f to be an endomorphism in V, otherwise the determinant cannot be fully well-defined. Let $A = [a_{i,j}]$ with respect to the basis $\{v_j\}_{j=1}^n$ of V. From Definition 6.6.3 we have that

$$\det(A) v_1 \wedge \cdots \wedge v_n = A v_1 \wedge \cdots \wedge A v_n$$

since $Av_j = \sum_{i=1}^{n} a_{i,j}v_i$ (see Definition 5.2.3) then we can substitute to the previous equation to obtain

$$\det(A)v_1 \wedge \cdots \wedge v_n = \sum_{i_1=1}^{n} a_{i_1,1}v_{i_1} \wedge \cdots \wedge \sum_{i_n=1}^{n} a_{i_n,n}v_{i_n} = \sum_{1\leqslant i_1,\ldots,i_n\leqslant n} \prod_{j=1}^{n} a_{i_j,j}(v_{i_1} \wedge \cdots \wedge v_{i_n}). \tag{6.10}$$

However, notice that if $i_j = i_{j'}$ for some $1 \leqslant j, j' \leqslant n$ then $v_{i_1} \wedge \cdots \wedge v_{i_n} = 0$, from antisymmetry. Therefore we can write Eq. (6.10) as a sum of permutations of the set $\{i_1, \ldots, i_n\} = \{1, \ldots, n\}$, that is

$$\begin{aligned}\det(A)v_1 \wedge \cdots \wedge v_n &= \sum_{\sigma\in S_n} a_{\sigma(1),1} \ldots a_{\sigma(n),n}(v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)}) \\ &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)a_{\sigma(1),1} \ldots a_{\sigma(n),n}(v_1 \wedge \cdots \wedge v_n).\end{aligned}$$

Since $v_1 \wedge \cdots \wedge v_n \neq 0$, it follows that

$$\det A = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)a_{\sigma(1),1}a_{\sigma(2),2} \ldots a_{\sigma(n),n}.$$

♮

Corollary 6.6.10. Another equivalent way of writing the determinant of A is

$$\det A = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)a_{1,\sigma(1)}a_{2,\sigma(2)} \ldots a_{n,\sigma(n)}.$$

Proposition 6.6.11. Let V be an n-dimensional k-vector space, and $(v_1, \ldots, v_n)$ be a basis for V, and $\omega \in \Lambda^n V$. Define, for every $1 \leqslant i \leqslant n$ the elements $w_i := \sum_{j=1}^{n} \alpha_{ij}v_j$. Then

$$\omega(w_1, \ldots, w_n) = \det[\alpha_{ij}]\omega(v_1, \ldots, v_n)$$

where $[\alpha_{ij}]$ is the matrix composed of the coefficients $\alpha_{ij} \in k$ for $1 \leqslant i, j \leqslant n$.

Proof. This is simply an application of the pullback operation, notice that every i-th argument can be seen as the i-th row of the following resulting vector

$$\begin{bmatrix} \sum_{j=1}^{n} \alpha_{1j}v_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{nj}v_j \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

Thus $\omega(w_1, \ldots, w_n) = [\alpha_{ij}]^{\wedge n}(w(v_1, \ldots, v_n)) = \det[\alpha_{ij}]\omega(v_1, \ldots, v_n)$. ♮

Definition 6.6.12 (Orientation)**.** Let V be an n-dimensional k-vector space, and let $\mathcal{B}$ be the collection of all basis for the vector space V. We define the following equivalence relation on $\mathcal{B}$: two basis $B, B' \in \mathcal{B}$ are equivalent said to be equivalent, $B \sim B'$, if and only if the change of basis matrix C from B to B' (or vice versa) has $\det C > 0$.

Since every change of basis matrix is an isomorphism, $\det C$ can only be either strictly positive or strictly negative, therefore, the quotient $\mathcal{B}/\sim$ splits the basis of V into two distinguished classes, both classes are said to define an *orientation* for V. If $[v]$ is an orientation class for V, we commonly refer to the opposite orientation as $-[v]$.

If $\phi\colon V \xrightarrow{\cong} L$ is an isomorphism of k-vector spaces, and $[v]$ is an orientation for V, then $[\phi v]$ is the *induced orientation* for W.

An *oriented k-vector space* is a pair $(V,[v])$, where V is a k-vector space and $[v]$ is an orientation class for V. An isomorphism between oriented k-vector spaces $\psi\colon (V,[v]) \xrightarrow{\cong} (L,[\ell])$ isa said to be *orientation preserving* if $[\psi v] = [\ell]$, otherwise we say that ψ is *order reversing*.

Definition 6.6.13 (Standard euclidean orientation)**.** The *standard orientation* on the euclidean space $\mathbf{R}^n$ is given by the basis $[e_1,\dots,e_n]$ where $e_j = (\delta_{ij})_{i=1}^n$ is the j-th unit vector of $\mathbf{R}^n$.

Corollary 6.6.14. Let V be an n-dimensional k-vector space endowed with an inner product $\langle -,-\rangle\colon V\times V \to k$. Let $\omega \in \Lambda^n V$ be non-zero. Then, there is a *unique orientation* μ for V, for which $\mu = [v_1,\dots,v_n]$ if and only if $\omega(v_1,\dots,v_n) > 0$ — where $(v_1,\dots,v_n)$ is a basis for V.

Proof. Let $B := (v_j)_{j=1}^n$ and $B' := (w_j)_{j=1}^n$ be orthonormal basis for V with respect to the given inner product, and C be the change of basis matrix from B to B' — that is, C is composed of entries $a_{ij} \in k$ such that $w_i = \sum_{j=1}^n a_{ij}v_j$. Since the given basis are orthonormal, we have

$$\delta_{ij} = \langle w_i, w_j\rangle = \sum_{k,\ell=1}^n a_{ik}a_{j\ell}\langle v_k, v_\ell\rangle = \sum_{k=1}^n a_{ik}a_{jk} := b_{ij}.$$

Notice that $[b_{ij}]_{i,j=1}^n$ is simply the matrix resulting from the product of C with its transpose C^* — therefore, $CC^* = \mathrm{id}$ and hence $\det C = \pm 1$. Regarding Proposition 6.6.11, we find that, if $\omega \in \Lambda^n V$ is such that $\omega(v_1,\dots,v_n) = \pm 1$, then necessarily $\omega(w_1,\dots,w_n) = \pm 1$ — for instance, if $(v_1,\dots,v_n)$ is chosen so that $\omega(v_1,\dots,v_n) = 1$, for some $\omega \in \Lambda^n V$, then this ω is unique.

If μ is an orientation for V, it's clear that if $\mu = [v_1,\dots,v_n]$ if and only if $\omega(v_1,\dots,v_n) > 0$ (the $\omega \in \Lambda^n V$ is unique by our last discussion), then μ is necessarily a unique orientation with such property. ♮

Definition 6.6.15 (Volume element)**.** If V is a n-dimensional vector space endowed with an inner product $\langle -,-\rangle\colon V\times V \to k$. Let μ be an orientation for V. The form $\omega \in \Lambda^n V$, described in Corollary 6.6.14, is called the *volume element* of V determined by the inner product $\langle -,-\rangle$ and orientation μ.

Useful determinant theorems, matrix determinant, etc.

Write about inverse matrix calculations

Chapter 7

Group Theory

7.1 Welcome to the Group

Lets recall Definition 1.2.16 and demystify it in the following definition.

Definition 7.1.1 (Group)**.** A group G is a groupoid G with one object $*$. The elements of the group are the morphisms $\mathrm{Aut}_{\mathsf{G}}(*)$. From the axioms contained in Definition 1.2.1, the graph G contains the following data:

(G1) An identity element $e = \mathrm{id}_*$.

(G2) An associative binary operation $G \times G \to G$, commonly denoted by juxtaposition.

(G3) For all $g \in G$ there exists an inverse element $g^{-1} \in G$ such that $gg^{-1} = e = g^{-1}g$.

Some Basic Laws

The following three propositions where already proved in a more general setting, but I'll rewrite them in this particular context for the sake of completeness.

Proposition 7.1.2. The identity element of a group is unique.

Proof. See Corollary 1.2.5. ♮

Proposition 7.1.3. The inverse of an element of a group is unique.

Proof. See Proposition 1.2.7. ♮

Proposition 7.1.4. Let G be a group and $g, h \in G$, then $(gh)^{-1} = h^{-1}g^{-1}$.

Proof. $(h^{-1}g^{-1})(gh) = h^{-1}(g^{-1}g)h = h^{-1}eh = h^{-1}h = e$ hence $h^{-1}g^{-1} = (gh)^{-1}$. ♮

Proposition 7.1.5 (Cancellation)**.** Let G be a group and elements $a, b, c \in G$. If $ac = bc$ then $a = b$.

Proof. Notice that $a = ae = (ac)c^{-1} = (bc)c^{-1} = be = b$. ♮

Proposition 7.1.6. Let G be a group. If $g \in G$, then the collection $\{gh : h \in G\}$ is equal to G.

Proof. Denote $G' := \{gh : h \in G\}$. Clearly we have $G' \subseteq G$. On the other hand, if $\ell \in G$, the element $g(g^{-1}\ell) = e\ell = \ell \in G'$, hence $G \subseteq G'$. Thus $G = G'$. ♮

Definition 7.1.7 (Commutative group)**.** A group G is said to be commutative (or abelian) if for all $g, h \in G$ we have $gh = hg$.

Corollary 7.1.8. Let G be a group such that for all $g \in G$, $g^2 = e$. Then G is a commutative group.

Proof. Let $g, h \in G$ be any elements, then $(gh)(hg) = gh^2g = geg = g^2 = e = (gh)(gh)$ and from cancellation law we find that $gh = hg$. ♮

Subgroups

Definition 7.1.9 (Subgroup)**.** Let G be a group. A subgroup H of G is a collection of elements $H \subseteq G$ containing the identity element, and such that composition of elements of H and inverses are closed in H — that is, given $x, y \in H$, then $xy \in H$ and $x^{-1}, y^{-1} \in H$. We say that a subgroup is trivial if it consists only of the identity element.

Corollary 7.1.10 (Intersection of subgroups)**.** Let G be a group and consider a non-empty collection of subgroups $\{G_j\}_{j\in J}$. Then, the intersection $\bigcap_{j\in J} G_j$ together with the group operation inherited from G is a subgroup of G.

Proof. Let $g \in \bigcap_{j\in J} G_j$ be any element, then $g \in G_j$ for every $j \in J$, and since G_j is a subgroup, then $g^{-1} \in G_j$ — that is, $g^{-1} \in \bigcap_{j\in J} G_j$. Let $g, h \in \bigcap_{j\in J} G_j$ be any two elements, then $gh \in G_j$ for each $j \in G$ and hence $gh \in \bigcap_{j\in J} G_j$. ♮

The following notion will accompany us for the whole journey, the idea of taking elements of any set and inducing a group structure from it, using the set elements to *generate* this new group.

Definition 7.1.11 (Groups and generators)**.** Let S be a set and G be a group. If every element of G can be written as the product of finitely many powers of the elements of S, then we say that G is generated by S, and the elements $s \in S$ are called generators of G. Moreover, if S is finite, we naturally say that G is finitely generated. Sometimes we denote this by $G = \langle S \rangle$. On the other hand, if $\varphi\colon S \to G$ is a set-function, G is said to be generated by φ if the image $\operatorname{im}\varphi$ generates the group G.

Orders

Definition 7.1.12 (Order of an element)**.** Let G be a group and $g \in G$ be any element. We say that g has finite order if there exists $n \in \mathbf{Z}_{>0}$ such that $g^n = e$. The order $|g|$ of the element g is defined as the smallest such positive integer. If g does not have a finite order, it is common to write $|g| = \infty$.

Lemma 7.1.13. Let G be a group and $g \in G$ be an element with finite order. Then $g^n = e$ for some $n \in \mathbf{Z}_{>0}$ if and only if $|g|$ divides n.

Proof. Since $|g| \leqslant n$, define $m \in \mathbf{Z}_{>0}$ such that $n - m|g| \geqslant 0$ and $n - (m+1)|g| < 0$. Define $r = n - m|g|$ to be the remainder, hence $r < |g|$. Our goal is to show that $r = 0$. Notice that $g^r = g^n g^{-|g|m} = ee^{-m} = e$, which can only be the case for $r = 0$, since $|g|$ is defined to be the least positive integer such that $g^{|g|} = e$. This proves that $m|g| = n$.

For the second part, suppose n is a multiple of $|g|$ and denote it by $m|g| = n$. Then $g^{m|g|} = e^m = e$. ♮

Definition 7.1.14 (Order of a group)**.** Let G group of finite number of elements. We define the order of G to be the number of its elements and denote it by $|G|$. If G is an infinite group, then $|G| = \infty$.

Order of Products

Proposition 7.1.15 (Order of the power)**.** Let G be a group and $g \in G$ be an element with finite order. Then for all $m \in \mathbf{Z}_{\geqslant 0}$ the element g^m has finite order. For $m = 0$ we have $|g^m| = 1$, for $m > 0$ we have

$$|g^m| = \frac{\operatorname{lcm}(m, |g|)}{m} = \frac{|g|}{\gcd(m, |g|)}.$$

Proof. From divisibility arguments, we have $\operatorname{lcm}(a,b)\gcd(a,b) = ab$ for integers a and b, hence the second equality is justified. We prove the equality $|g^m| = \frac{\operatorname{lcm}(m,|g|)}{m}$. For the sake of notation, let $d := |g^m|$. Notice that $g^{md} = e$ and hence $|g|$ divides md. Since d is the least positive integer with such property, it follows that md is the least common multiple of m and $|g|$. Therefore $m|g^m| = \operatorname{lcm}(m, |g|)$, which proves the equation. ♮

Proposition 7.1.16. Let G be a group, then for any $g, h \in G$ we have $|gh| = |hg|$.

Proof. First, let $x, y \in G$ be any elements, we prove that $|yxy^{-1}| = |x|$. Notice that for any $n \geqslant 1$ we have $(yxy^{-1})^n = yx^n y^{-1}$, hence the least element that annihilates the product yxy^{-1} is the order of x — that is, $|yxy^{-1}| = x$. In particular, $hg = g^{-1}ghg$, hence $|hg| = |g^{-1}(gh)g| = |gh|$. ♮

Proposition 7.1.17. Let G be a group and $g, h \in G$ be such that $gh = hg$. Then $|gh|$ divides $\operatorname{lcm}(|g|, |h|)$.

Proof. Let n be a common multiple of $|g|$ and $|h|$, then $g^n = h^n = e$ from Lemma 7.1.13. Notice that the commutative property $gh = hg$ allow us to permute the terms of $g^n h^n = e$ in order to obtain $g^n h^n = (gh)^n = e$. In particular, since $\operatorname{lcm}(|g|, |h|)$ is a common multiple of $|g|$ and $|h|$, we find that $(gh)^{\operatorname{lcm}(|g|,|h|)} = e$ and again from Lemma 7.1.13 we find that $|gh|$ divides $\operatorname{lcm}(|g|, |h|)$ ♮

Lemma 7.1.18. Let G be a group and $g, h \in G$ commute — $gh = hg$. If $\gcd(|g|, |h|) = 1$, then $|gh| = |g||h|$.

Proof. Let $|gh| = \ell, |g| = m$ and $|h| = n$. From Proposition 7.1.17, we have that $\ell \mid \operatorname{lcm}(m,n)$, and since $mn = \operatorname{lcm}(m,n)\gcd(m,n) = \operatorname{lcm}(m,n)$, then $\ell \mid mn$, which implies that $\ell \leqslant mn$. Moreover, since the elements commute, $(gh)^\ell = g^\ell h^\ell = e$ hence $g^\ell = (h^\ell)^{-1}$. From Proposition 7.1.15, $|g^\ell| = \frac{|g|}{\gcd(\ell,|g|)} = \frac{m}{m} = 1$ and equivalently for h we have $|h^\ell| = 1$. This shows that $g^\ell = h^\ell = e$ and therefore both m and n divide ℓ, hence so does the product mn, thus $mn \leqslant \ell$. This completes the proof that $\ell = mn$. ♮

Definition 7.1.19 (Maximal finite order)**.** Let G be a group. An element $g \in G$ is said to be of maximal finite order if its order is finite and for all $h \in G$ with finite order, we have $|h| \leqslant |g|$.

Proposition 7.1.20. Let G be a commutative group and $g \in G$ be of maximal finite order. If $h \in G$ has finite order, then $|h|$ divides $|g|$.

Proof. Define the notation $|g| = m$ and $|h| = n$. Let $P = (p_j)_j$ be a finite sequence containing all primes such that less than or equal to m and define a finite sequence of integers of same length $A = (a_j)_j$ such that $m = \prod_j p_j^{a_j}$. Since $n \leqslant m$ it follows that there also exists a finite sequence of integers $B = (b_j)_j$ such that $n = \prod_j p_j^{b_j}$. Suppose for the sake of contradiction that n doesn't divide m, so that there exists an index k such that $a_k < b_k$ — from the fact that $\frac{m}{n} = \prod_j p_j^{a_j - b_j}$. Consider now the order — following from Lemma 7.1.18:

$$|g^{\left(p_k^{a_k}\right)} h^{\left(n/p_k^{b_k}\right)}| = |g^{\left(p_k^{a_k}\right)}||h^{\left(n/p_k^{b_k}\right)}| = \frac{m}{\gcd(m, p_k^{a_k})} \frac{n}{\gcd(n, n/p_k^{b_k})} = \frac{m}{p_k^{a_k}} \frac{n}{n/p_k^{b_k}} = m p_k^{b_k - a_k}$$

and since $b_k > a_k$, it follows that $m p_k^{b_k - a_k} > m$ and hence there is a contradiction since we assumed that m was the the maximal finite order of the group. We conclude that there does not exist k for which b_k is less than a_k — thus n divides m. ♮

Finite Groups and Elements of Order 2

Lemma 7.1.21 (Order 2 elements implies commutative)**.** Let G be a group such that all non-identity elements have order 2. Then G is commutative.

Proof. Let $g, h \in G$, notice that $|gh| = 2$ from Proposition 7.1.17 hence $gh \in G$, then $(gh)^2 = e$ and therefore $gh = g^{-1}h^{-1} = g^{-1}(ghgh)h^{-1} = hg$ commutes. ♮

Lemma 7.1.22. Let G be a finite group such that any element has order at most 2. Let H be any subgroup of G, for any $t \in G \setminus H$ consider the collection $T = H \cup \{ht : h \in H\}$, then T is a subgroup of G with $|T| = 2|H|$.

Proof. Since $t \notin H$, any element $ht \in T$ with $h \in H$ is such that $ht \notin H$. The number of distinct elements of the form is $|H|$ — one for each $h \in H$ — hence $|T| = |H| + |H| = 2|H|$. ♮

Lemma 7.1.23 (Order 2^n)**.** Let G be a finite group such that any element has order at most 2. The order of G is of the form 2^n for some $n \geqslant 0$. Moreover, if $|G| > 1$, then there exists a subgroup H of G with order $|H| = 2^{n-1}$.

Proof. We create a recursive algorithm to find the collection of elements of G that generate any other element contained in G:

1. (Base case) If $G = H_j$, return H_j.
2. (Recursion) Let $g \in G \setminus H_j$ and construct $H_{j+1} = H_j \cup \{hg : h \in H_j\}$, recursively call the algorithm with H_{j+1}.

Such algorithm is certain to terminate since $|G|$ is finite. Notice that the second step always doubles the order of the set H_j, so that — if H_n is the result of the algorithm — then $|H_n| = 2^n$. This shows that $|G| = |H_n| = 2^n$. The second part of the statement follows immediately from the construction of the algorithm. ♮

Proposition 7.1.24. Let G be a commutative group, if there exists exactly one element of order 2 — say $f \in G$ — then the product of all elements of the group is

$$\prod_{g \in G} g = f.$$

Otherwise, we have $\prod_{g \in G} g = e$.

Proof. Since G is finite, let $G = \{e, g_1, g_1^{-1}, \dots, g_n, g_n^{-1}\}$. Suppose there exists one and only one element of order 2 and denote it by f — that is, $f = f^{-1}$. Since the group is commutative, we can rearrange the product of elements so that we can take the pairwise product of each element and its inverse — this being possible only in the case where $g \neq g^{-1}$ — taking the product we find

$$\prod_{g \in G} g = e(g_1 g_1^{-1}) \dots (g_j g_j^{-1}) f \dots (g_{j+2} g_{j+2}^{-1})(g_n g_n^{-1}) = e^{j+1} f e^{n-(j+1)} = f$$

since f has no inverse in $G \setminus \{f\}$.

Let there be no element with order 2 in G, then the rearrangement of pairwise element and respective inverse is possible with each $g \in G$, making

$$\prod_{g \in G} g = e(g_1 g_1^{-1}) \dots (g_j g_j^{-1}) \dots (g_n g_n^{-1}) = e^{n+1} = e.$$

Suppose there exists $m > 1$ distinct elements in G with order 2 and define $T = \{f \in G : |f| \leqslant 2\}$ and $S = G \setminus T$. From the last case we know that $\prod_{s \in S} s = e$. Note that T forms a subgroup of G:

- $e \in T$.
- If $g, h \in T$, then from Proposition 7.1.17 $|gh|$ divides $\operatorname{lcm}(|g|, |h|) \leqslant 2$ hence $|gh| \leqslant 2$ and $gh \in T$.

- Since $g \in T$ implies $gg = e$ then $g = g^{-1}$ and $|g^{-1}| = 2$, so $g^{-1} \in T$.

Moreover, the order of the group T has to be of the form 2^k for some $k \geqslant 2$, from Lemma 7.1.23. From the same lemma, choose a subgroup H with order 2^{k-1} and take $u \in T \smallsetminus H$, so that $T = H \cup \{hu : h \in H\}$ — such u exists from the algorithm constructed in the lemma. Since T is commutative (see Lemma 7.1.21), we can write

$$\prod_{t \in T} t = \prod_{h \in H} h \prod_{h \in H} hu = \prod_{h \in H} uh^2 = \prod_{h \in H} u = u^{2^{k-1}} = (u^2)^{2^{k-2}} = e.$$

Therefore we can finally conclude that

$$\prod_{g \in G} g = \prod_{s \in S} s \prod_{t \in T} t = e.$$

♮

7.2 Examples of Groups

Symmetry Group

Definition 7.2.1 (Symmetric groups)**.** Let $A \in \mathtt{Set}$. The symmetric group of A — also referred to as the permutation group of A — $\mathrm{Aut}_{\mathtt{Set}}(A)$, is denoted by S_A. The symmetric group of the range set $\{1, \dots, n\}$ is denoted S_n.

Notation 7.2.2 (Permutations)**.** A permutation $\sigma \in S_n$ is denoted by a table of the form

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n-1) & \sigma(n) \end{pmatrix}$$

Remark 7.2.3 (Convention)**.** If $\sigma, \tau \in S_n$, we write the composition $\sigma\tau$ to denote the permutation

$$i \xmapsto{\sigma\tau} \sigma(\tau(i)).$$

That is, it follows the same order of composition for maps.

Proposition 7.2.4. There exists an embedding $f \colon S_n \hookrightarrow M_{n \times n}(\{0, 1\})$ — where $M_{n\times n}(\{0,1\})$ is the collection of $n \times n$ matrices with entries assuming the value of 1 or 0. Moreover, if $f(\sigma) = M_\sigma$ and $f(\tau) = M_\tau$ then $f(\sigma\tau) = M_\sigma M_\tau$.

Proof. Let $\sigma \in S_n$ be any permutation. We define f as the mapping

$$\sigma \xmapsto{f} M_\sigma = \begin{bmatrix} \delta_{1\sigma(1)} & \dots & \delta_{1\sigma(n)} \\ \vdots & \ddots & \vdots \\ \delta_{n\sigma(1)} & \dots & \delta_{n\sigma(n)} \end{bmatrix}$$

It should be noted that since σ is injective, $\delta_{i\sigma(j)}$ assumes the value 1 only once in the j-th column. If $\sigma = \tau$ then clearly $\delta_{i\sigma(j)} = \delta_{i\tau(j)}$ for every $1 \leqslant i, j \leqslant n$, thus $f(\sigma) = f(\tau)$ — f is injective.

Consider now the composition of permutations $\sigma\tau \in S_n$. Suppose we want to check where the j-th element goes to from the action of $\sigma\tau$. Notice that for all possible new positions $1 \leqslant i \leqslant n$ we have $\delta_{i\sigma\tau(j)} = \sum_{k=1}^{n} \delta_{i\sigma(k)}\delta_{k\tau(j)}$ — that is, j belongs to i after $\sigma\tau$ if and only if $\tau(j) = k$ and $\sigma(k) = i$ for some $1 \leqslant k \leqslant n$. We have

$$f(\sigma\tau) = M_{\sigma\tau} = \begin{bmatrix} \delta_{1\sigma\tau(1)} & \dots & \delta_{1\sigma\tau(n)} \\ \vdots & \ddots & \vdots \\ \delta_{n\sigma\tau(1)} & \dots & \delta_{n\sigma\tau(n)} \end{bmatrix} = \begin{bmatrix} \delta_{1\sigma(1)} & \dots & \delta_{1\sigma(n)} \\ \vdots & \ddots & \vdots \\ \delta_{n\sigma(1)} & \dots & \delta_{n\sigma(n)} \end{bmatrix} \begin{bmatrix} \delta_{1\tau(1)} & \dots & \delta_{1\tau(n)} \\ \vdots & \ddots & \vdots \\ \delta_{n\tau(1)} & \dots & \delta_{n\tau(n)} \end{bmatrix} = M_\sigma M_\tau.$$

♮

Proposition 7.2.5. The symmetry group S_n contains elements of all orders d for $1 \leqslant d \leqslant n$.

Proof. Let d be any integer in the range $1 \leqslant d \leqslant n$. Consider the subgroup $S_d \subseteq \mathrm{Aut}_{\mathsf{Set}}(\mathbf{Z}/n\mathbf{Z})$, composed of permutations $\sigma \in S_d$ such that $\sigma([i]_n) = [i]_n$ for all $i > d$. Notice that the permutation τ given by $\tau([i]_n) = [i+1]_n$ for all $i \leqslant d$ and $\tau([j]_n) = [j]_n$ is an element of S_d and $|\tau| = d$. Thus τ is an element of order d in S_n. ♮

Corollary 7.2.6. For every $n \in \mathbf{N}$, there exists an element $x \in S_{\mathbf{N}}$ with order $|x| = n$.

Proof. Let any $n \in \mathbf{N}$ and choose $m \in \mathbf{N}$ with $m \geqslant n$. From Proposition 7.2.5 we find that the group S_m contains an element $\sigma \in S_m$ with order $|\sigma| = n$. We can now construct an element $\tau \in S_{\mathbf{N}}$ — defined by

$$\tau(a) = \begin{cases} \sigma(a), & a \leqslant m \\ a, & a > m \end{cases}$$

So clearly $|\tau| = n$. ♮

Permutations, Transpositions and Sign

Definition 7.2.7 (Transposition)**.** We define a transposition on a collection $\{1, \dots, n\}$ to be a map $\tau \in S_n$ such that exists indices $1 \leqslant i < j \leqslant n$ for which $\tau(i) = j$ and $\tau(j) = i$, and $\tau(k) = k$ for all $k \neq i, j$.

Proposition 7.2.8. Every permutation can be written as a composition of finitely many transpositions.

Proof. We proceed via induction on the number of points of $\{1, \dots, n\}$. For the base case $n = 2$ the composition is trivial. Assume as the induction hypothesis that for $n-1 > 2$ the statement is true. Now, consider $\sigma \in S_n$ and $i \in \{1, \dots, n\}$ be any element in the ordered collection $I_n := \{1, \dots, n\}$. Denote $\tau_{i,j}$ the transposition that changes i with j and maintains unchanged the remainder of the points. Assume that $\sigma(i) = j$, then clearly $\tau_{i,j}\sigma(i) = \tau_{i,j}(j) = i$. Notice that since $\tau_{i,j}\sigma$ maintains i unchanged, we can see it as a permutation of $n-1$ points (by simply ignoring the point i), hence it can be written as a composition of finitely many transpositions by the induction hypothesis. Notice that since $\sigma = \tau_{i,j}(\tau_{i,j}\sigma)$ we find that σ can also be written as a composition of finitely many transpositions. ♮

Definition 7.2.9 (Elementary transpositions)**.** Let $\tau \in S_n$ be a transposition. We say that τ is an elementary transposition if exists $i \in \{1, \dots, n\}$ for which $\tau(i) = i+1$ and $\tau(i+1) = i$, and for all $j \neq i$ we have $\tau(j) = j$.

Proposition 7.2.10. Every transposition can be written as a composition of finitely many elementary transpositions.

Proof. Let $\sigma \in S_n$ be any transposition. Denote by τ_k the elementary transposition $\tau_k(k) = k+1$ and $\tau(k+1)k$. Let x be the transposed element of σ, and $\sigma(x) = y$. Without loss of generality we can assume that $y > x$. Now we write σ as the composition — beware of Remark 7.2.3

$$\sigma = \left(\tau_x \tau_{x+1} \cdots \tau_{y-3} \tau_{y-2}\right) \left(\tau_{y-1} \tau_{y-1} \cdots \tau_{x+1} \tau_x\right) = \left(\prod_{k=x}^{y-2} \tau_k\right) \left(\prod_{k=0}^{(y-1)-x} \tau_{(y-1)-k}\right)$$

Which follows from the fact that

$$\left(\prod_{k=0}^{(y-1)-x} \tau_{(y-1)-k}\right)(i) = \begin{cases} y, & i = x \\ i-1, & x < i \leqslant y \\ i, & i < x \text{ or } y < i \end{cases}$$

$$\left(\prod_{k=x}^{y-2} \tau_k\right)(i) = \begin{cases} x, & i = y \\ i+1, & x \leqslant i < y \\ i, & i < x \text{ or } y < i \end{cases}$$

Hence the composition of both gives

$$\left(\prod_{k=x}^{y-2} \tau_k\right) \left(\prod_{k=0}^{(y-1)-x} \tau_{(y-1)-k}\right)(i) = \begin{cases} y, & i = x \\ x, & i = y \\ i, & i \notin \{x, y\} \end{cases}$$

which is equivalent to the transposition σ. ♮

Corollary 7.2.11. Every permutation can be written as a composition of finitely many elementary transpositions.

Definition 7.2.12 (Sign)**.** Let $\sigma \in S_n$ be a permutation on n objects. We define the sign of σ as a map $\operatorname{sign}: S_n \longrightarrow \{1, -1\}$ such that

$$\operatorname{sign}(\sigma) = (-1)^m, \text{ where } m := |\{(i, j) : 1 \leqslant i < j \leqslant n,\ \sigma(i) > \sigma(j)\}|.$$

Equivalently, if $x_1, \dots, x_n \in k$ then we can define the sign of σ as

$$\operatorname{sign}(\sigma) = \prod_{i<j} \frac{x_{\sigma(i)} - x_{\sigma(j)}}{x_i - x_j}.$$

Corollary 7.2.13. If σ can be written as an odd number of transpositions, then $\operatorname{sign}(\sigma) = -1$. Otherwise, if it is an even number of transpositions, we have $\operatorname{sign}(\sigma) = 1$.

Proposition 7.2.14. Let $\sigma, \tau \in S_n$, then

$$\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau).$$

This implies that $\operatorname{sign}\colon S_n \to \{1, -1\}$ is a group homomorphism.

Proof. Consider the composition $\sigma\tau$, then we have

$$\operatorname{sign}(\sigma\tau) = \prod_{i<j} \frac{x_{\sigma\tau(i)} - x_{\sigma\tau(j)}}{x_i - x_j}. \tag{7.1}$$

Define the following: when $\tau(i) < \tau(j)$ then $\tau(i) := p$ and $\tau(j) := q$; on the other hand, when $\tau(j) < \tau(i)$ let $\tau(i) := q$ and $\tau(j) := p$. This way we find that

$$\frac{x_{\sigma\tau(i)} - x_{\sigma\tau(j)}}{x_{\tau(i)} - x_{\tau(j)}} = \frac{x_{\sigma(p)} - x_{\sigma(q)}}{x_p - x_q}. \tag{7.2}$$

From construction of the above, Eq. (7.2) yields

$$\operatorname{sign}(\sigma) = \prod_{p<q} \frac{x_{\sigma(p)} - x_{\sigma(q)}}{x_p - x_q}. \tag{7.3}$$

Notice that we can write Eq. (7.1) (using Eq. (7.3)) as the product

$$\operatorname{sign}(\sigma\tau) = \prod_{i<j} \frac{x_{\sigma\tau(i)} - x_{\sigma\tau(j)}}{x_{\tau(i)} - x_{\tau(j)}} \frac{x_{\tau(i)} - x_{\tau(j)}}{x_i - x_j} = \operatorname{sign}(\sigma)\operatorname{sign}(\tau)$$

as wanted. ♮

Dihedral Group

Definition 7.2.15 (Dihedral group)**.** A dihedral group is defined as the group of isometric symmetries of regular polygons — which are rotations and reflections about a line. Given a n-sided regular polygon, its group of symmetries have $2n$ elements — n rotations and n reflections — and we denote it by D_{2n}.

Proposition 7.2.16. There exists an embedding $D_{2n} \hookrightarrow S_n$.

Proof. Label the vertices of the n-gon by $V = \{[1]_n, \dots, [n]_n\} = \mathbf{Z}/n\mathbf{Z}$. Notice that any element $x \in D_{2n}$ can be described as $x \in \operatorname{Aut}_{\mathsf{Set}}(V)$ — where the automorphism is restricted to the adjacency of the vertices, that is, if $x([i]_n) = [k]_n$, then $x([i-1]_n), x([i+1]_n) \in \{[k-1]_n, [k+1]_n\}$. This shows the existence of the embedding $D_{2n} \hookrightarrow \operatorname{Aut}_{\mathsf{Set}}(V) = S_V = S_n$. ♮

Proposition 7.2.17. Any symmetry $x \in D_{2n}$ can be written as $y^a z^b$ — where we choose any $y, z \in D_{2n}$ which are, respectively, a rotation and a reflection about a line — with $0 \leqslant a < 2$ and $0 \leqslant b < n$.

Proof. Notice that if y is any rotation, then $|y| = n$ and if z is any reflection, then $|z| = 2$.

We first show that y and z are independent. We'll work with the injection $D_{2n} \rightarrowtail \mathrm{Aut}_{\mathsf{Set}}(V)$, where we have the collection of vertices $V = \mathbf{Z}/n\mathbf{Z}$. Let $j \in V$ be any vertex and suppose $z(j) = k$, then by the adjacency of the vertices are maintained, implying in $z(j-1) = k-1$ and $z(j+1) = k+1$. On the other hand, if $y(j) = k$, then the adjacency of the vertices is inverted, that is, $y(j-1) = k+1$ and $y(j+1) = k-1$. Hence clearly reflections and rotations cannot be dependent.

The only possible symmetries of D_{2n} involve the maintenance or the inversion of the adjacency of each vertex — we cannot break nor deform the edges that connect each of the vertices — thus if $x \in D_{2n}$, then $x = \prod_{(\alpha,\beta)\in I} z^\alpha y^\beta$ for some finite set I with $0 \leqslant \alpha < 2$ and $0 \leqslant \beta < n$. We now show that such product can be reduced. Let $j \in V$ be any vertex and assume

$$z(j) = k \quad \text{and} \quad y(k) = \ell,$$

We now analyse the symmetries yz and zy^m — for $0 \leqslant m < n$.

- The adjacent vertices of j — when subjected to the symmetry yz — are obtained as
$$yz(j-1) = y(k+1) = \ell + 1 \quad \text{and} \quad yz(j+1) = y(k-1) = \ell - 1$$
- Since $y(k) = \ell$, y is the action rotating the vertex k an amount of $\ell - k$ times — where we'll adopt the convention that if $\ell - k < 0$, the rotation is counter-clockwise, if $\ell + k > 0$ the rotation is clockwise. Therefore
$$y(j) = j + (\ell - k).$$
Notice that this implies in $y^m(j) = j + m(\ell - k)$ — we now define $t = y^m(j)$. Since $z(j) = k$, then $z(t) = k - m(\ell - k)$. Therefore, adjacent vertices of j — when subjected to the symmetry zy^m — are given by
$$\begin{aligned} &\text{Vertex } j-1: \quad zy^m(j-1) = z(j + m(\ell-k) - 1) = k - m(\ell - k) + 1 \\ &\text{Vertex } j+1: \quad zy^m(j+1) = z(j + m(\ell-k) + 1) = k - m(\ell - k) - 1 \end{aligned}$$
In particular, for the case $m = n-1$ — beware of the modularity of the vertices, they lie on $\mathbf{Z}/n\mathbf{Z}$ — we get
$$\begin{aligned} &\text{Vertex } j-1: zy^{n-1}(j-1) = k - (n-1)(\ell - k) + 1 = kn - n\ell + \ell + 1 = \ell + 1 \\ &\text{Vertex } j: \quad\quad zy^{n-1}(j) = k - (n-1)(\ell - k) = kn - n\ell + \ell = \ell \\ &\text{Vertex } j+1: zy^{n-1}(j+1) = k - (n-1)(\ell - k) - 1 = kn - n\ell + \ell - 1 = \ell + 1 \end{aligned}$$

From this analysis we conclude the general relation $yz = zy^{n-1}$.

Lets analyse the product $(z^a y^b)(z^c y^d)$ for the different cases of $0 \leqslant a, c < 2$ and $0 \leqslant b, d < n$.

- If $a = 1$ and $c = 0$ then $(z^a y^b)(z^c y^d) = zy^{b+d}$.

- If $a, c = 0$ then
$$(z^a y^b)(z^c y^d) = y^{b+d}.$$

- If $a = 0$ and $c = 1$ then $(z^a y^b)(z^c y^d) = y^b z^c y^d = y^{b-1} z y^{d+(n-1)}$, by recurrence we find find finally that
$$(z^a y^b)(z^c y^d) = z y^{d+b(n-1)}.$$

- If $a, c = 1$ then $(z^a y^b)(z^c y^d) = z y^b z y^d = z y^{b-1} z y^{d+(n-1)}$ and, by recurrence
$$(z^a y^b)(z^c y^d) = z^2 y^{d+b(n-1)} = y^{d+b(n-1)}.$$

This shows that the finite product $x = \prod_{(\alpha,\beta)\in I} z^\alpha y^\beta$ can be reduced — for some $0 \leqslant a < 2$ and $0 \leqslant b < n$ — to $x = z^a y^b$. ♮

Corollary 7.2.18. Let $y, z \in D_{2n}$ be, respectively, a rotation and a reflection. The following relation are satisfied:

$$|y| = n,\ |z| = 2, \text{ and } yz = zy^{n-1}.$$

The set $\{y, z\}$ generates any element of D_{2n}.

Cyclic Group

We'll initially define a cyclic group as a group generated by one element x with a relation $x^n = e$ for some $n \in \mathbf{N}$. We'll denote such groups by C_n — for a formal definition, see Definition 7.3.20.

Example 7.2.19. An example of a cyclic group is $(\mathbf{Z}/n\mathbf{Z}, +)$, where $[1]_n \in \mathbf{Z}/n\mathbf{Z}$ is the element generating any elements of $\mathbf{Z}/n\mathbf{Z}$.

Proposition 7.2.20. Given an element $y \in C_n$ — where x is the generating element of C_n — assume $y = x^m$. We have that

$$|y| = \frac{n}{\gcd(m, n)}.$$

Proof. Notice that $|y| = |x^m| = \frac{|x|}{\gcd(m,|x|)} = \frac{n}{\gcd(m,n)}$. ♮

Corollary 7.2.21. The element $x \in C_n$ is the generator of the cyclic group if and only if

$$\gcd(|x|, n) = 1.$$

7.3 Grp Category

Definition 7.3.1 (Group morphism)**.** Let $(G, \cdot_G)$ and $(H, \cdot_H)$ be groups together with their binary operation. A group morphism — also called homomorphism — is a map $\varphi\colon (G, \cdot_G) \to (H, \cdot_H)$ such that the that the following diagram commutes

$$\begin{tikzcd} G \times G \arrow[r, "\varphi\times\varphi"] \arrow[d, "\cdot_G"'] & H \times H \arrow[d, "\cdot_H"] \\ G \arrow[r, "\varphi"] & H \end{tikzcd}$$

Where $\varphi \times \varphi$ is uniquely defined in $\mathtt{Set}$ by **??** — mapping $(g, \ell) \overset{\varphi\times\varphi}{\longmapsto} (\varphi(g), \varphi(\ell))$. The commutativity of such diagram can be viewed as the requirement that φ preserves the structure coming from the binary operations — that is, for any $g, \ell \in G$

$$\varphi(g \cdot_G \ell) = \varphi(g) \cdot_H \varphi(\ell).$$

Definition 7.3.2 (Category of groups)**.** The category of groups $\mathtt{Grp}$ consists of the collection of objects — called groups — and group morphisms between them.

Proposition 7.3.3. $\mathtt{Grp}$ is a category.

Proof. Let $(G, \cdot_G)$, $(H, \cdot_H)$ and $(K, \cdot_K)$ be any groups. The identity $\mathrm{id}_G\colon G \to G$ is a group morphism since $\mathrm{id}_G(g \cdot_G \ell) = g \cdot_G \ell$ for any $g, \ell \in G$. Moreover, we can define a map

$$f\colon \mathrm{Mor}_{\mathtt{Grp}}(G, H) \times \mathrm{Mor}_{\mathtt{Grp}}(H, K) \to \mathrm{Mor}_{\mathtt{Grp}}(G, K)$$

with the mapping $(\psi, \varphi) \overset{f}{\mapsto} \psi\varphi$ — since the following diagram commutes

$$\begin{tikzcd} G \times G \arrow[r, "\varphi\times\varphi"] \arrow[d, "\cdot_G"'] \arrow[rr, bend left, "(\psi\varphi)\times(\psi\varphi)"] & H \times H \arrow[r, "\psi\times\psi"] \arrow[d, "\cdot_H"] & K \times K \arrow[d, "\cdot_K"] \\ G \arrow[r, "\varphi"] \arrow[rr, bend right, "\psi\varphi"'] & H \arrow[r, "\psi"] & K \end{tikzcd}$$

In other words, for any $g, \ell \in G$ we have

$$\psi\varphi(g \cdot_G \ell) = \psi(\varphi(g \cdot_G \ell)) = \psi(\varphi(g) \cdot_H \varphi(\ell)) = \psi(\varphi(g)) \cdot_K \psi(\varphi(\ell)) = \psi\varphi(g) \cdot_K \psi\varphi(\ell).$$

Therefore $\psi\varphi \in \mathrm{Mor}_{\mathtt{Grp}}(G, K)$. For the other part of the diagram, we have

$$(\psi\varphi) \times (\psi\varphi)(g, \ell) = (\psi\varphi(g), \psi\varphi(\ell)) = \psi \times \psi(\varphi(g), \varphi(\ell)) = (\psi \times \psi)(\varphi \times \varphi)(g, \ell)$$

Since group morphisms are maps in $\mathtt{Set}$, we have that associativity is inherited. ♮

Proposition 7.3.4. There exists a covariant forgetful functor $F\colon \mathtt{Grp} \to \mathtt{Set}$.

Proof. For objects, define F as $F(G, \cdot_G) = G$ — where we denoted G together with its binary operation only to express that the multiplicative structure is lost in the process. Let $\varphi\colon (G, \cdot_G) \to (H, \cdot_H)$ be a group morphism, denote by $\overline{\varphi} \in \mathrm{Mor}(\mathtt{Set})$ the function $\overline{\varphi}\colon G \to H$ such that $\overline{\varphi}(g) = \varphi(g)$ for all $g \in G$. For such morphisms we define F as $F\varphi = \overline{\varphi}\colon F(G, \cdot_G) \to F(H, \cdot_H)$.

Let $\psi \in \mathrm{Mor}_{\mathtt{Grp}}(HK)$, then we have $\overline{\psi\varphi} = \overline{\psi}\overline{\varphi}\colon H \to K$. Thus

$$F(\psi\varphi) = \overline{\psi\varphi} = \overline{\psi}\overline{\varphi} = F\psi F\varphi.$$

This shows that F is a covariant forgetful functor. ♮

Proposition 7.3.5. The trivial group $* \in \mathtt{Grp}$ is the initial and final object of $\mathtt{Grp}$. That is, for any $G \in \mathtt{Grp}$ the diagram

$$\begin{tikzcd} * \arrow[r, bend right, dashed, "\varphi"'] & G \arrow[l, bend right, dashed, "\psi"'] \end{tikzcd}$$

commutes for uniquely defined group morphisms φ and ψ.

Proof. Let $G \in \mathtt{Grp}$ be any group. We define maps $\varphi\colon * \to G$ mapping $e \overset{\varphi}{\mapsto} e_G$, where e is the only element of $*$ — being unique possible map $* \to G$ that preserves the group structure. Clearly, φ is a group morphism since $\varphi(ee) = \varphi(e) = e_G = \varphi(e)\varphi(e)$ — this shows that $*$ is the initial object of $\mathtt{Grp}$. Let $\psi\colon G \to *$ be a map defined by $g \overset{\psi}{\mapsto} e$ — which is clearly unique. Then ψ is a morphism of groups, because $\psi(gh) = e = \psi(g)\psi(h)$ — showing that $*$ is the final object of $\mathtt{Grp}$. ♮

Properties of Morphisms

Proposition 7.3.6 (Commuting on inverses)**.** Let $(G, \cdot_G), (H, \cdot_H)$ be groups and consider $\varphi \in \mathrm{Mor}_{\mathtt{Grp}}(G, H)$. Define $\mathrm{inv}_G\colon G \xrightarrow{\simeq} G$ and $\mathrm{inv}_H\colon H \xrightarrow{\simeq} H$ as the maps $g \xmapsto{\mathrm{inv}_G} g^{-1}$ and $h \xmapsto{\mathrm{inv}_H} h^{-1}$. Then the following diagram commutes

$$\begin{tikzcd} G \arrow[r, "\varphi"] \arrow[d, "\mathrm{inv}_G"'] & H \arrow[d, "\mathrm{inv}_H"] \\ G \arrow[r, "\varphi"] & H \end{tikzcd}$$

That is, $\varphi(g^{-1}) = \varphi(g)^{-1}$ for every $g \in G$.

Proof. Let $g \in G$ be any element, then

$$\varphi(g^{-1}) = \varphi(g^{-1}e_G) = \varphi(g^{-1} \cdot_G g \cdot_G g^{-1}) = \varphi(g^{-1}) \cdot_H \varphi(g) \cdot_H \varphi(g^{-1}),$$

applying cancellation law on the equation above we find

$$e_H = \varphi(g) \cdot_H \varphi(g^{-1}).$$

Hence $\varphi(g^{-1}) = \varphi(g)^{-1}$. Moreover, this implies that $e_G \overset{\varphi}{\mapsto} e_H$ since

$$\varphi(e_G) = \varphi(g \cdot_G g^{-1}) = \varphi(g) \cdot_H \varphi(g^{-1}) = \varphi(g) \cdot_H \varphi(g)^{-1} = e_H.$$

♮

Proposition 7.3.7 (Generators and unique extension). Let G be a group and S be a generator set for G. Let H be any group and $f\colon S \to H$ be a set-function. If there exists a morphism $\phi\colon G \to H$ such that $\phi|_S = f$, then ϕ is unique.

Proof. Let $\psi\colon G \to H$ be another morphism satisfying the condition specified above — then clearly $\psi|_S = f = \phi|_S$, that is, ϕ and ψ agree on S. Since S generates G, every element $g \in G$ can be written as a finite product $g = \prod_j s_j \in \langle S \rangle$ thus

$$\psi(g) = \psi\Big(\prod_j s_j\Big) = \prod_j \psi(s_j) = \prod_j \phi(s_j) = \phi\Big(\prod_j s_j\Big) = \phi(g),$$

which implies in $\psi = \phi$. ♮

Proposition 7.3.8 (Image subgroup). Let $\phi\colon G \to H$ be a morphism of groups, then $\operatorname{im}\phi \subseteq H$ is a subgroup of H.

Proof. Let $h \in \operatorname{im}\phi$ be any element and consider $g \in \phi^{-1}(h)$, from Proposition 7.3.6 we see that $\phi(g^{-1}) = \phi(g)^{-1} = h^{-1} \in \operatorname{im}\phi$, thus $\operatorname{im}\phi$ is closed under inverses. Moreover, given another $h' \in \operatorname{im}\phi$, there exists $g' \in \phi^{-1}(h')$ and $\phi(gg') = \phi(g)\phi(g') = hh' \in \operatorname{im}\phi$ — hence ϕ is closed under products. ♮

Definition 7.3.9 (Kernel). We define the kernel of a morphism of groups $\phi \in \operatorname{Mor}_{\mathtt{Grp}}(G, H)$ as the collection $\ker\phi = \{g \in G : \phi(g) = e_H\}$.

Lemma 7.3.10 (Kernel subgroup). Let $\phi\colon G \to H$ be a group morphism. The kernel $\ker\phi \subseteq G$ is a subgroup of G.

Proof. Let $g \in \ker\phi$ be any element, then since $\phi(g^{-1}) = \phi(g)^{-1} = e_H^{-1} = e_H$, then $g^{-1} \in \ker\phi$. Also, if $u \in \ker\phi$ is another element, then $\phi(gu) = \phi(g)\phi(u) = e_H e_H = e_H$ and hence $gu \in \ker\phi$ — thus $\ker\phi$ is a subgroup of G. ♮

Proposition 7.3.11 (Monomorphisms, kernels and injectivity). Let $\phi\colon G \to H$ be a morphism of groups. Then the following properties are equivalent:

(a) ϕ is a monomorphism in $\mathtt{Grp}$.

(b) $\ker\phi = \{e_G\}$.

(c) ϕ is injective in $\mathtt{Set}$.

Proof.
- (a) $\Rightarrow$ (b). Suppose ϕ is a monormophism and consider the following commutative diagram

$$\begin{tikzcd} \ker\phi \arrow[r, hook, shift left, "\iota"] \arrow[r, shift right, "e"'] & G \arrow[r, tail, "\phi"] & H \end{tikzcd}$$

From the monomorphism definition we find $\iota = e$ and therefore $\ker\phi = \{e_G\}$.

- (b) $\Rightarrow$ (c). If $\ker \phi = \{e_G\}$, then, given elements $g, g' \in G$ such that $\phi(g) = \phi(g')$, then, multiplying both sides by $\phi(g')^{-1}$ we get

$$\phi(g)\phi(g')^{-1} = \phi(g)\phi(g'^{-1}) = \phi(gg'^{-1}) = e_G.$$

  That is, $gg'^{-1} \in \ker \phi$, but since $\ker \phi$ is trivial, then $gg'^{-1} = e_G$ and hence $g = g'$ — in other words, ϕ is injective.
- (c) $\Rightarrow$ (a). Let ϕ be injective — that is, for any pair of *set*-functions $f, g \colon A \rightrightarrows G$ from a set A, then $\phi f = \phi g$ if and only if $f = g$. In particular, if we impose a group structure in the set A and that both f and g preserve the group structure of A, we find that the injectivity in $\mathtt{Set}$ implies in the monomorphism ϕ as a morphism in $\mathtt{Grp}$.

$\natural$

Proposition 7.3.12. Let G be a group and H be an abelian group. If there exists an injective group morphism $\iota \colon G \rightarrowtail H$, then G is abelian.

Proof. Let $x, y \in G$ be any pair of elements, then $\phi(xy) = \phi(x)\phi(y) = \phi(y)\phi(x) = \phi(yx)$, thus $xy = yx$ — since ϕ is injective. $\natural$

The following proposition is a trivial one, but it is a really useful tool to prove the non-existence of non-trivial morphisms between certain groups — it relies on arguments based on the order of elements of each group.

Proposition 7.3.13 (Morphisms and orders)**.** Let $G \in \mathtt{Grp}$ be a group admitting an element $g \in G$ of finite order $|g| \in \mathbf{N}$. Let H be any group and consider the morphism $\phi \colon G \to H$. We have that $|\phi(g)|$ divides the order of $|g|$.

Proof. Notice that $\phi(g)^{|g|} = \phi(g^{|g|}) = \phi(e_G) = e_H$, hence $|g|$ is a multiple of the order of $\phi(g) \in H$. $\natural$

Example 7.3.14. Consider for example the collection of morphisms $\mathrm{Mor}_{\mathtt{Grp}}(C_7, C_{15})$. Let ϕ be any such morphism. Consider any element $g \in C_7$ and recall that $|g|$ must be a divisor of 7 — see Proposition 7.2.20. On the other hand, if $h \in C_{15}$ then $|h|$ divides 15. From Proposition 7.3.13 we see that if $\phi(g) = h$, then $|h|$ must divide $|g|$, but $\gcd(7, 15) = 1$, hence $|h| = 1$ — that is $h = e_H$ and ϕ is the trivial morphism $\phi(g) = e_H$ for all $g \in C_7$. This shows that there is no non-trivial morphism between C_7 and C_{15}.

Isomorphism of Groups

Proposition 7.3.15 (Isomorphisms are bijections)**.** Let $\phi \in \mathrm{Mor}(\mathtt{Grp})$. Then ϕ is an isomorphism if and only if it is a bijection.

Proof. Consider an isomorphism $\phi \colon G \to H$. Using the forgetful functor $F \colon \mathtt{Grp} \to \mathtt{Set}$ we see that $F\phi$ is a bijection of sets in $\mathtt{Set}$ — recall Lemma 1.4.11 — thus ϕ defines a bijection between the elements of G and H.

On the other hand, if ϕ is a bijection, consider its set-function inverse $(F\phi)^{-1}: H \to G$. We now show that $(F\phi)^{-1}$ preserves the structures of groups. Since $\phi(e_G) = e_H$, then $(F\phi)^{-1}(e_H) = e_G$. Moreover, for any $h, h' \in H$ — since ϕ is surjective — take elements $g, g' \in G$ such that $\phi(g) = h$ and $\phi(g') = h'$, then we find that $(F\phi)^{-1}(hh') = gg' = (F\phi)^{-1}(h) \cdot (F\phi)^{-1}(h')$. This implies in the existence of a naturally induced morphism of groups $\phi^{-1}: H \to G$ defined by $\phi^{-1}(h) = (F\phi)^{-1}(h)$. It is clear that ϕ^{-1} is the right and left inverse of ϕ, thus ϕ^{-1} is the inverse of ϕ in Grp and ϕ is an isomorphism. ♮

Definition 7.3.16 (Embedding)**.** Let $\phi: G \xrightarrow{\simeq} H$ be an isomorphism of groups. We define the group $\operatorname{im}\phi \subseteq H$ as an embedding of G on H.

Proposition 7.3.17. Let $\phi: G \xrightarrow{\simeq} H$ be an isomorphism of groups. Then:

- For all $g \in G$, we have $|g| = |\phi(g)|$.
- G is commutative if and only if H is commutative.

Proof. Let $g \in G$ be any element, then from Proposition 7.3.13 we have that $|\phi(g)|$ divides $|g|$. Since ϕ^{-1} exists and is a morphism of groups, it also follows that $|g|$ divides $|\phi(g)|$ — hence $|g| = |\phi(g)|$.

Let G be a commutative group and $G \simeq H$. Let $h, h' \in H$ be any elements and consider $\phi^{-1}(h) = g$ and $\phi^{-1}(h') = g'$. From the structure preserving property of ϕ we have

$$hh' = \phi(g)\phi(g') = \phi(gg') = \phi(g'g) = \phi(g')\phi(g) = h'h.$$

That is, H is commutative. The counter-implication is equivalent and will be omitted. ♮

Lemma 7.3.18 (Inner automorphism)**.** Let $G \in$ Grp. For each $g \in G$, the map $\gamma_g: G \to G$ given by $\gamma_g(a) = gag^{-1}$ is an automorphism — called inner automorphism of G.

Proof. Let $g \in G$ be any element. Suppose $a \in \ker\gamma_g$, then $\gamma_g(a) = gag^{-1} = e_G$, hence, $a = g^{-1}e_G g = g^{-1}g = e_G$, that is $\ker\gamma_g = e_G$ — γ_g is injective. Let $g' \in G$ be any element, then, $\gamma_g(g^{-1}g'g) = g(g^{-1}g'g)g^{-1} = g'$, that is, γ_g is surjective. We conclude that γ_g is a bijection — hence an isomorphism, so $\gamma_g \in \mathrm{Aut}_{\mathtt{Grp}}(G)$. ♮

Lemma 7.3.19 (Inner automorphism correspondence)**.** Let $G \in$ Grp. The map $\phi: G \to \mathrm{Aut}_{\mathtt{Grp}}(G)$ defined by the mapping $\phi(g) = \gamma_g$ (where γ_g is defined in Lemma 7.3.18) is a morphism of groups.

Proof. Let $g, g' \in G$ be any elements, then

$$\gamma_{gg'}(a) = (gg')a(gg')^{-1} = (gg')a(g'^{-1}g^{-1}) = g(g'ag'^{-1})g^{-1} = \gamma_g\gamma_{g'}(a).$$

This being said, its easy to see that ϕ preserves the group structure: $\phi(gg') = \gamma_{gg'} = \gamma_g\gamma_{g'} = \phi(g)\phi(g')$. Thus ϕ is a morphism of groups. ♮

More Thoughts On Cyclic Groups

We can now state the definition of a cyclic group in a formal manner, it goes as follows:

Definition 7.3.20 (Cyclic group)**.** A group G is said to be cyclic if $G \simeq \mathbf{Z}$ or $G \simeq \mathbf{Z}/n\mathbf{Z}$ for some $n \in \mathbf{N}$.

Proposition 7.3.21. A finite group of order $n \in \mathbf{N}$ is cyclic if and only if it contains an element of order n.

Proof. Let G be a cyclic group of order n. Since G is finite, then there exists an isomorphism $\phi\colon G \xrightarrow{\simeq} \mathbf{Z}/n\mathbf{Z}$. Consider the element $g = \phi^{-1}([1]_n) \in G$, from Proposition 7.3.17 we see that $|g| = n$.

Let G be a finite group of order n and $x \in G$ be such that $|x| = n$. Let $\phi\colon G \to \mathbf{Z}/n\mathbf{Z}$ be any morphism of groups sending $x \mapsto [1]_n$. Consider the collection $G' = \{e_G, x, x^2, \dots, x^{n-1}\} \subseteq G$, and notice that, together with the binary operation of G, G' becomes a group of n elements — that is, $G' = G$ and every element $g \in G$ can be written as $g = x^k$ for some $1 \leqslant k \leqslant n$. This implies in ϕ injective — thus a bijection. From Proposition 7.3.15 we see that ϕ is an isomorphism $G \simeq \mathbf{Z}/n\mathbf{Z}$ — G is a cyclic group, which finishes our proof. ♮

Proposition 7.3.22. The order of the cyclic group C_n is equal to $\phi(n)$, where ϕ is the Euler totient function — that is, the number of positive integers less than n that are relatively prime to n.

Proof. Let $x \in C_n$ be a generator of the group — that is, $x^n = e$ — then for all $d < n$ such that $\gcd(d, n) = 1$ we have $x^d \neq x$ and $(x^d)^n = e$ thus x^d is a generator of C_n. Therefore the number of distinct elements of C_n is the same as the number of positive integers coprime of n. ♮

Proposition 7.3.23 (Subgroup)**.** Any subgroup of a cyclic group is cyclic.

Proof. Let $G = \langle g \rangle$ be a cyclic group and $H \subseteq G$ be any subgroup of G. If $h \in H$ then there exists $n \in \mathbf{Z}$ such that $h = g^n$, thus $g^n \in H$ — but then, given any other $h' \in H$, there must exist $m \in \mathbf{Z}$ such that $h' = g^{n+m} = h^m$, thus $H = \langle h \rangle$ is cyclic. ♮

Lemma 7.3.24. Given a group G, the collection of inner automorphisms, which we'll denote by $\mathrm{Inn}(G)$, is a subgroup of $\mathrm{Aut}(G)$. Moreover, the following statements are equivalent:

(a) $\mathrm{Inn}(G)$ is cyclic.

(b) $\mathrm{Inn}(G)$ is trivial.

(c) G is abelian.

Therefore, if $\mathrm{Aut}(G)$ is cyclic, the group G is abelian.

Proof. From Lemma 7.3.18 and Lemma 7.3.19 we find that $\mathrm{Inn}(G) \subseteq \mathrm{Aut}(G)$ is indeed a subgroup. Now we prove the equivalences:

- (a) $\Leftrightarrow$ (b). If $\operatorname{Inn}(G)$ is cyclic, then there must exist $g \in G$ for which $\operatorname{Inn}(G) = \langle \gamma_g \rangle$. Therefore, given any $h \in G$, there exists $n \in \mathbf{Z}$ for which $\gamma_h = \gamma_g^n$ — therefore $hgh^{-1} = g^n g g^{-n} = g$, thus $hg = gh$. We can thus conclude that γ_g is the identity morphism $h \mapsto h$, which implies in $\operatorname{Inn}(G)$ being trivial. Now, if we assume $\operatorname{Inn}(G)$ to be trivial by hypothesis, then clearly $\operatorname{Inn}(G)$ is cyclic.
- (b) $\Leftrightarrow$ (c). If $\operatorname{Inn}(G)$ is trivial, then given any pair of elements $g, h \in G$ we have $\gamma_g = \mathrm{id}$ and thus $ghg^{-1} = h$, which implies in $gh = hg$, thus G is commutative. Now, if G is commutative by hypothesis, we find that $ghg^{-1} = (gg^{-1})h = h$ thus $\operatorname{Inn}(G)$ is trivial.

Thus, if $\operatorname{Aut}(G)$ is a cyclic group, we find that the subgroup $\operatorname{Inn}(G) \subseteq \operatorname{Aut}(G)$ is also cyclic (from Proposition 7.3.23), hence G is abelian. ♮

Some Matrix Groups

Example 7.3.25 (Important matrix groups)**.** Let k be a field and $\mathrm{Mat}_{n\times n}(k) = \mathrm{End}_{\mathtt{Vect}_k}(k^n)$ be the collection of all $n \times n$ matrices over k, for any $n \in \mathbf{Z}_{>0}$. We define the following important groups of square matrices:

1. General linear group: $\mathrm{GL}_n(k) := \{T \in \mathrm{Mat}_{n\times n}(k) : \det T \neq 0\}$, the group of invertible matrices.
2. Special linear group: $\mathrm{SL}_n(k) := \{T \in \mathrm{GL}_n(k) : \det T = 1\}$.
3. Orthogonal group: $\mathrm{O}_n(k) := \{T \in \mathrm{GL}_n(k) : TT^* = T^*T = \mathrm{id}_n\}$[1].
4. Special orthogonal group: $\mathrm{SO}_n(k) := \{T \in \mathrm{O}_n(k) : \det T = 1\}$.
5. Unitary group: $\mathrm{U}_n(\mathbf{C}) := \{T \in \mathrm{GL}_n(\mathbf{C}) : TT^\dagger = T^\dagger T = \mathrm{id}_n\}$[2].
6. Special unitary group: $\mathrm{SU}_n(\mathbf{C}) := \{T \in \mathrm{U}_n(\mathbf{C}) : \det T = 1\}$.
7. Lie algebra of the general linear group: $\mathfrak{gl}_n(k) \subseteq \mathrm{GL}_n(k)$ such that
$$[T, M] := TM - MT \in \mathfrak{gl}_n(k)$$
for all $T, M \in \mathfrak{gl}_n(k)$.
8. Lie algebra of the special linear group: $\mathfrak{sl}_n(k) := \{T \in \mathfrak{gl}_n(k) : \operatorname{tr} T = 0\}$.
9. Lie algebra of the orthogonal group: $\mathfrak{o}_n(k) := \{T \in \mathfrak{gl}_n(k) : T + T^* = 0\}$
10. Lie algebra of the unitary group: $\mathfrak{u}_n(\mathbf{C}) := \{T \in \mathfrak{gl}_n(\mathbf{C}) : T + T^\dagger = 0\}$.
11. Lie algebra of the special unitary group: $\mathfrak{su}_n(\mathbf{C}) := \{T \in \mathfrak{gl}_n(\mathbf{C}) : \operatorname{tr} T = 0\}$.

The proof that such examples are indeed groups come immediately from Proposition 6.6.4 and Definition 5.7.6. Notice that all of the above groups are subgroups of the general linear group $\mathrm{GL}_n(k)$. It should be noted, however, that the Lie groups presented do *not* form a group under multiplication.

[1] The matrix T^* denotes the transpose of T, which is the same as the dual of T.

[2] The matrix $T^\dagger$ denotes the complex transpose of T.

Example 7.3.26 (Upper triangular matrices)**.** The collection of upper triangular matrices over a field k is a subgroup of $\mathrm{GL}_n(k)$. Let A be any upper triangular matrix, then from definition $[a_{ij}]_{1\leqslant i,j\leqslant n}$ is such that $a_{ij} = 0$ for all $i > j$ and $\prod_{j=1}^{n} a_{jj} \neq 0$ — therefore, any permutation $\sigma \in S_n$ other than the identity will have at least one element $i_0 > j$. Hence

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{j=1}^{n} a_{\sigma(j)\,j} = \prod_{j=1}^{n} a_{jj} \neq 0,$$

which implies in $A \in \mathrm{GL}_n(k)$. Moreover, since the sum of two upper triangular matrices is upper triangular, clearly the sum is in $\mathrm{GL}_n(k)$ — which proves that the collection of upper triangular matrices is a subgroup, since the existence of inverses is immediate.

Example 7.3.27. Notice that any matrix of the form

$$A := \begin{bmatrix} a+bi & c+di \\ -c+di & a-bi \end{bmatrix} \tag{7.4}$$

Is such that

$$\begin{aligned} AA^{\dagger} &= \begin{bmatrix} a+bi & c+di \\ -c+di & a-bi \end{bmatrix} \begin{bmatrix} a-bi & -c-di \\ c-di & a+bi \end{bmatrix} = \begin{bmatrix} a^2+b^2+c^2+d^2 & 0 \\ 0 & a^2+b^2+c^2+d^2 \end{bmatrix}, \\ A^{\dagger}A &= \begin{bmatrix} a-bi & -c-di \\ c-di & a+bi \end{bmatrix} \begin{bmatrix} a+bi & c+di \\ -c+di & a-bi \end{bmatrix} = \begin{bmatrix} a^2+b^2+c^2+d^2 & 0 \\ 0 & a^2+b^2+c^2+d^2 \end{bmatrix}. \end{aligned}$$

Therefore, imposing $a^2+b^2+c^2+d^2 = 1$ in Eq. (7.4) makes A into a matrix of the group $\mathrm{SL}_2(\mathbf{C})$. Furthermore, if $T \in \mathrm{GL}_n(\mathbf{C})$ is a matrix with a form other than that of Eq. (7.4), then it cannot be the case that $\det T$ equal 1 — thus, every $\mathrm{SL}_2(\mathbf{C})$ matrix has the form of A.

Proposition 7.3.28. The group $\mathrm{SL}_2(\mathbf{Z})$ is generated by the pair matrices

$$A := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad B := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Proof. It should be noted that $B^n = \left[\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right]$ and $A^2 = \left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ — thus, given any matrix $X = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbf{Z})$ we have

$$AX = \begin{bmatrix} -c & -d \\ a & b \end{bmatrix} \quad \text{and} \quad B^n X = \begin{bmatrix} a+nc & b+nd \\ c & d \end{bmatrix}.$$

If we assume that $c \neq 0$, and $|a| \geqslant |c|$ (otherwise, simply multiply from the left by A), we can apply the division algorithm to find $q_0, r_0 \in \mathbf{Z}$ for which $a = q_0 c + r_0$, satisfying $0 \leqslant r_0 \leqslant |c|$. Then we can apply

$$AB^{-q_0}X = A \begin{bmatrix} r_0 & b-q_0 d \\ c & d \end{bmatrix} = \begin{bmatrix} -c & -d \\ r_0 & b-q_0 d \end{bmatrix} := X_0,$$

Now, if $r_0 \neq 0$, we can recursively apply the division algorithm (which I'll carry again just for the sake of clarity): we find $q_1, r_1 \in \mathbf{Z}$ such that $-c = q_1 r_0 + r_1$, where $0 \leqslant r_1 \leqslant |r_0|$ — we thus find a new power q_1 for which

$$AB^{-q_1} X_0 = \begin{bmatrix} -r_0 & q_0 d - b \\ r_1 & (q_0 q_1 - 1)d - q_1 b \end{bmatrix} = X_1.$$

This process is ensured to terminate at some point with a zero lower left entry. Since the group is acting on the left of $\mathrm{SL}_2(\mathbf{Z})$ and the determinant is always zero, the resulting matrix will be of the form $\left[\begin{smallmatrix} \pm 1 & m \\ 0 & \pm 1 \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbf{Z})$ — which equals either B^m or $-B^{-m}$. We conclude that there must exist some $g \in \langle A, B \rangle$ such that $gX = \pm B^t$ for some $t \in \mathbf{Z}$ — and since $A^2 = -I_2$, we obtain $X = \pm g^{-1} B^t \in \mathrm{SL}_2(\mathbf{Z})$ and $\pm g^{-1} B^t \in \langle A, B \rangle$. ♮

Proposition 7.3.29 (Union of subgroups)**.** Let G be any group, then:

(a) Given subgroups $H, Q \subseteq G$, the union $H \cup Q$ is a subgroup of G if and only if either $H \subseteq Q$ or $Q \subseteq H$.

(b) If $H_0 \subseteq H_1 \subseteq \dots$ is a collection of subgroups of G, then the union $\bigcup_{j \geqslant 0} H_j \subseteq G$ is a subgroup of G.

Proof. (a) If $H \subseteq Q$ or $Q \subseteq H$, then $H \cup Q$ is clearly a subgroup. On the other hand, if $H \cup Q$ is a subgroup, then for every $h \in H$ and $q \in Q$ we must have $hq \in H \cup Q$ — which implies that $hq \in H$ or $hq \in Q$, for the first case, we have $h^{-1}(hq) = q \in H$ thus $Q \subseteq H$, for the second case, $(hq)q^{-1} = h \in Q$ implying in $H \subseteq Q$.

(b) Since the composition of elements is only defined for finitely many elements, the proposition follows immediately.

♮

Group Products

Let $(G, \cdot_G), (H, \cdot_H) \in \mathtt{Grp}$ be any objects. We define a binary operation $\cdot \colon (G \times H)^2 \to G \times H$ as the mapping

$$(g, h) \cdot (g', h') = (g \cdot_G g', h \cdot_H h'). \tag{7.5}$$

Such binary operation defines a group structure on $G \times H$. Notice that, given an element $(g, h) \in G \times H$, there exists an element $(g^{-1}, h^{-1}) \in G \times H$ such that $(g, h) \cdot (g^{-1}, h^{-1}) = (e_G, e_H)$. Moreover, clearly $(e_G, e_H) \in G \times H$ is the identity element of the structure. Hence $(G \times H, \cdot) \in \mathtt{Grp}$.

Also, the natural projections $\pi_G \colon G \times H \to G$ and $\pi_H \colon G \times H \to H$ define morphisms of groups.

Definition 7.3.30 (Direct product)**.** Let $\{G_j\}_{j \in J}$ be a collection of groups. We define the direct product of this family as the group $\prod_{j \in J} G_j$ given by elements $(x_j)_{j \in J}$ such that $x_j \in G_j$. The composition of elements of the direct product is defined component-wise, that is, if $(x_j)_{j \in J}, (y_j)_{j \in J} \in \prod_{j \in J} G_j$, then $(x_j)_{j \in J} (y_j)_{j \in J} := (x_j y_j)_{j \in J}$. Moreover, inverses are also defined component-wise, $(x_j)_{j \in J}^{-1} := (x_j^{-1})_{j \in J}$.

Proposition 7.3.31. The direct products of groups are products on the category of groups, $\mathtt{Grp}$. That is, for all group W and group morphisms $f \in \mathrm{Mor}_{\mathtt{Grp}}(W, G)$ and $g \in \mathrm{Mor}_{\mathtt{Grp}}(W, H)$, there exists a unique morphism $\varphi \in \mathrm{Mor}_{\mathtt{Grp}}(W, G \times H)$ such that the following diagram commutes

$$\begin{tikzcd} & W \arrow[d, dashed, "\varphi"] \arrow[ddl, bend right, "f"'] \arrow[ddr, bend left, "g"] & \\ & G \times H \arrow[dl, "\pi_G"] \arrow[dr, "\pi_H"'] & \\ G & & H \end{tikzcd}$$

Proof. We just take $\varphi\colon W \to G \times H$ as the mapping $w \overset{\varphi}{\mapsto} (f(w), g(w))$. We show that φ exists in $\mathtt{Grp}$: let $x, y \in W$ be any elements, then, since f and g are group morphisms, we find that

$$\varphi(xy) = (f(xy), g(xy)) = (f(x)f(y), g(x)g(y)) = (f(x), g(x))(f(y), g(y)) = \varphi(x)\varphi(y).$$

That is, φ is a group morphism. The uniqueness comes from the covariant functor $F\colon \mathtt{Grp} \to \mathtt{Set}$, since $\mathtt{Set}$ allows for products and hence the set-function $F\varphi$ is unique. ♮

Remark 7.3.32. We now show that if $G, H \in \mathtt{Grp}$ are such that $G \simeq H \times G$, *it does not follow that* H is trivial. Let $G := \bigoplus_{j=0}^{\infty} \mathbf{Z}$ and $H := \mathbf{Z}$ be abelian groups under the natural structure of addition. Notice that $G \simeq H \times G$ by the natural assignment of each element of G to itself in $H \times G$. Since H is non-trivial, we found a counterexample.

Proposition 7.3.33. Let G be a group and $H, Q \subseteq G$ be subgroups for which $H \cap Q = e$ and $HQ = G$ — that is, for every $g \in G$, there exists $h \in H$ and $q \in Q$ such that $g = hq$ — we also impose that $hq = qh$ for every $h \in H$ and $q \in Q$. Then, the morphism of groups $H \times Q \xrightarrow{\simeq} G$ defined by the mapping $(h, q) \mapsto hq$ is an isomorphism.

Proof. Notice that $(h, q)(h', q') = (hh', qq') \mapsto (hh')(qq') = (hq)(h'q')$ thus the map is indeed a morphism of groups. Moreover, since every element of G can be written as a product HQ, it follows that the map is surjective. Now, let (h, q) be in the kernel of the morphism, then $hq = e$ which implies in $h = q^{-1}$ but then $h \in H \cap Q$ and by hypothesis $h = e$ — thus the morphism is injective. ♮

7.4 Quotient Groups — The Birth of Normal Subgroups

Cosets

Definition 7.4.1 (Coset)**.** Let G be a group and H be a subgroup of G. Given any $g \in G$, a left coset of H in G induced by g and denoted by gH is a set whose elements have the form gh for each $h \in H$. A right coset of H in G induced by g is denoted Hg and is a set consisting of elements of the form hg for each $h \in H$. An element of a coset is commonly called *coset representative*.

Corollary 7.4.2. Let G be a group and H be subgroup of G, then, for every $h \in H$, we have

$$hH = Hh = H.$$

Proof. Notice that, given $x \in H$, the element $h(h^{-1}x) = x \in hH$, thus $H \subseteq hH$, on the other hand, it's clear that $hH \subseteq H$, since hH is composed of product of elements of H, which itself is closed under products. The same analogous proof goes for Hh so I won't bother to write it down. ♮

Corollary 7.4.3 (Equal cosets)**.** Let G be a group and $H \subseteq G$ be a subgroup. Given $x, y \in G$, if the cosets xH and yH share any common element, then $xH = yH$.

Proof. Let $g \in xH \cap yHk$, then there exists $xh \in xH$ and $yh' \in yH$ such that $xh = g = yh'$, then, in particular, $x = yh'h^{-1}$ moreover, since H is a subgroup, it is clear that $h'h^{-1} \in H$ then $xH = (yh'h^{-1})H = y(h'h^{-1})H = yH$. ♮

Definition 7.4.4 (Index)**.** Let G be a group and $H \subseteq G$ be a subgroup. The number of left cosets of H in G is denoted by $[G : H]$, which will be commonly referred to as the *index* of H in G.

Corollary 7.4.5. If we denote by $*$ the trivial group, the *order* of a group G is the same as $[G : *]$ — that is, $|G| = [G : *]$.

Proof. One can view the trivial group $*$ as a subgroup of G containing only the identity. Notice that the number of left cosets of $*$ will be exactly the number of elements of G, that is $[G : *] = |G|$. ♮

Proposition 7.4.6. Let G be a group, and $H \subseteq G$ be a subgroup, and $Q \subseteq H$ be a subgroup. Then, if any two of the quantities $\{[G : H], [H : Q], [G : Q]\}$ is finite, the third is also finite and the following equality holds

$$[G : H][H : Q] = [G : Q].$$

Proof. Let $\{x_i\}_{i\in I} \subseteq H$ be coset representatives of Q, and $\{y_j\}_{j\in J} \subseteq G$ be coset representatives of H — that is, each one of the collections $\{x_i Q\}_{i\in I}$ and $\{y_j H\}_{j\in J}$ have pairwise disjoint elements, and $H = \bigcup_{i\in I} x_i Q$, and $G = \bigcup_{j\in J} y_j H$. Then we have that $G = \bigcup_{(i,j)\in I\times J} y_j x_i Q$, and our goal will be to prove that $y_j x_i Q \cap y_{i'} x_{j'} Q = \varnothing$. Suppose on the contrary that their intersection is non-empty, which by Corollary 7.4.3 implies $y_j x_i Q = y_{j'} x_{i'} Q$. Since $x_j, x_{j'} \in H$, we have $y_j x_i Q H = y_j x_i H = y_j H$ and analogously $y_{j'} x_{i'} Q H = y_{j'} H$ — thus $y_j H = y_{j'} H$, which implies in $y_j = y_{j'}$. This shows that the collection $\{y_j x_i\}_{(i,j)\in I\times J} \subseteq G$ are coset representatives for Q and therefore $[G : H][H : Q] = [G : Q]$. ♮

Corollary 7.4.7 (Lagrange's theorem)**.** Let G be a finite group, then the order of any subgroup H of G divides the order $|G|$.

Proof. Since $|G| = [G : *]$ is finite, any subgroup H of G is also finite and therefore $[G : H][H : *] = [G : *]$, which is exactly the same as $[G : H]|H| = |G|$. ♮

Example 7.4.8 (Prime order)**.** Let G be a group with order $|G| := p$ prime. Choose any $g \in G$ with $g \neq e$, and consider the subgroup $H := \langle g \rangle$. From Proposition 7.4.6 we find that $[G : H]|H| = p$, hence $|H|$ divides p, but since $|H| \leqslant p$, then $|H| = p$ and therefore $H = G$. This implies that any non-identity element of G generates the whole group, which is the same as to say that G is cyclic.

Normal Subgroups

Definition 7.4.9 (Normal subgroup)**.** Let G be a group. We define a *normal subgroup* to be the kernel of some morphism of groups in $\mathrm{Mor}_{\mathsf{Grp}}(G, -)$[3]. In other words, a subgroup $N \subseteq G$ is normal if there exists a morphism of groups $\phi\colon G \to H$, for some group H, for which $\ker \phi = N$.

Definition 7.4.10 (Quotient group)**.** Let G be a group and N be a normal subgroup of G. We denote by G/N the collection of all left cosets of N in G, on the other hand, $G\backslash N$ denotes the collection of all right cosets of N in G. Moreover, we view G/N (and $G\backslash N$) as groups where:

- The product of two cosets xN and yN (or, respectively, Nx and Ny) is given by $(xN)(yN) := (xy)N$ which is again a left coset in G/N (conversely, $(Nx)(Ny) := N(xy) \in G\backslash N$).
- Given any coset xN (respectively, Nx), its inverse is given by $x^{-1}N$ (respectively, Nx^{-1}).
- The identity of the group is N.

The group G/N is commonly referred to as the *quotient group* of G by H.

Proposition 7.4.11. Let G be a group. A subgroup $N \subseteq G$ is normal if and only if, for every $g \in G$, we have $gNg^{-1} = N$.

Proof. First, suppose that N is a normal group and $\phi\colon G \to H$ is a morphism such that $\ker \phi = N$, then if $g \in G$ is any element, we see that for any $n \in N$ we have $\phi(gn) = \phi(g)\phi(n) = \phi(g)$ and analogously $\phi(ng) = \phi(n)\phi(g) = \phi(g)$, thus in general $gN = Ng = \phi^{-1}(\phi(g))$. Note that if we multiply both groups on the right by g^{-1} we get $gNg^{-1} = N$, as wanted.

On the other hand, let $N \subseteq G$ be a subgroup such that $gNg^{-1} = N$ for every $g \in G$, then in particular left cosets are equal to right cosets because, multiplying on the right by g we obtain $gN = Ng$. Consider the group of left cosets G/N (since right and left cosets are equivalent in this specific case, we could also have considered $G\backslash N$) and define the morphism of groups $\pi\colon G \to G/N$ by the mapping $\pi(g) = gN$. Notice that if $g \in \ker \pi$, then $gN = N$, which implies that $g \in N$, moreover, if $n \in N$ is any element, then $\pi(n) = nN = N$ — thus $\ker \pi = N$. ♮

[3]In this case, we are using $\mathrm{Mor}_{\mathsf{Grp}}(G, -)$ to denote the same as the collection of all group morphisms whose source is G.

In fact the morphism $\pi\colon G \twoheadrightarrow G/N$, defined above by the map $g \overset{\psi}{\mapsto} gN$, is so important we are even going to distinguishably call it the *canonical projection map* of G onto the factor group G/N. It is trivial that such canonical projection π is surjective.

Corollary 7.4.12. Let G be a group and $H \subseteq G$ be a subgroup. Then H is normal in G if and only if, for all given $\gamma \in \operatorname{Inn}(G)$, we have $\gamma(H) \subseteq H$. Therefore, the morphism $\operatorname{Inn}(G) \to \operatorname{Aut}(H)$ mapping $\gamma_g \mapsto \gamma_g|_H$ is well defined.

Corollary 7.4.13 (Intersection of normal subgroups is normal)**.** Let $\{N_j\}_{j\in J}$ be any collection of normal subgroups of a given group G. Then $N := \bigcap_{j\in J} N_j$ is a normal subgroup of G.

Proof. Let $n \in N$ and $g \in G$ be any two elements, then $n \in N_j$ for all $j \in J$ and from the normal condition we obtain that $gng^{-1} \in N_j$ for all $j \in J$ as well — which implies in $gng^{-1} \in N$. ♮

Proposition 7.4.14. If G is a group and $N \subseteq G$ is a subgroup with index $[G : N] = 2$, then N is normal in G.

Proof. Suppose $N \subseteq G$ is a subgroup with index 2 — that is, G/N consists only of two cosets. Thus there must exist $g \in G$ such that $gN \neq N$ and consequently $Ng \neq N$ — moreover, since $G = N \cup gN = N \cup Ng$, one concludes that $gN = Ng$. ♮

Definition 7.4.15 (Normal closure)**.** Given a group G and a subset $S \subseteq G$, the *normal closure* of S is the subgroup of G generated by all elements of the form $g^{-1}sg$ — for $g \in G$ and $s \in S$.

Quotient Group Properties

We now study the properties of factorization of maps between groups and sequences of maps from the viewpoint of quotientings. One interesting immediate example is when $H \subseteq G$ is a subgroup of a group G, then

$$H \overset{\iota}{\rightarrowtail} G \overset{\pi}{\twoheadrightarrow} G/H$$

which is a short exact sequence. Moreover, given any groups G, Q and K, if we have an exact sequence

$$* \longrightarrow Q \xrightarrow{f} G \xrightarrow{g} K \longrightarrow *$$

we can conclude that f is injective, and g is surjective — this comes from the fact that $\ker f = e_Q$ and, since the kernel of $K \to 0$ is the whole group K, we necessarily have $g(G) = K$. Moreover, if we define $H := \ker g$, there exists a natural identification

$$\begin{tikzcd}
* \arrow[r] & Q \arrow[r, tail, "f"] \arrow[d, "\simeq"'] & G \arrow[r, two heads, "g"] \arrow[d, "\simeq"] & K \arrow[r] \arrow[d, "\simeq"] & * \\
* \arrow[r] & H \arrow[r] & G \arrow[r] & G/H \arrow[r] & *
\end{tikzcd}$$

Proposition 7.4.16 (Universal property of quotient groups)**.** Let G be a group and N be a normal subgroup of G. Then, given a group Q together with a morphism of groups $\phi\colon G \to Q$ such that N is the kernel of ϕ, there exists a unique morphism $\phi_*\colon G/N \to Q$ such that the diagram

$$\begin{tikzcd} G \arrow[r, "\phi"] \arrow[d, two heads, "\pi"'] & Q \\ G/N \arrow[ur, tail, dashed, bend right, "\phi_*"'] & \end{tikzcd}$$

is commutative. Moreover, the map ϕ_* is injective. The morphism ϕ_* induces an isomorphism of groups $\overline{\phi}\colon G/N \xrightarrow{\simeq} \operatorname{im}\phi$ and therefore we have the following factorization

$$\begin{tikzcd} G \arrow[r, "\phi"] \arrow[d, two heads] & Q \\ G/N \arrow[r, "\simeq", "\overline{\phi}"'] & \operatorname{im} f \arrow[u, hook] \end{tikzcd}$$

Proof. Define ϕ_* simply as the mapping $gH \to \phi(g)$, then clearly $\phi_*\pi = \phi$. Moreover, given any $x \in G/N$, the fiber $\pi^{-1}(x)$ is non-empty, thus ϕ_* cannot assume any other values beside those specified by ϕ, which implies in its uniqueness. Moreover, since $\ker\phi = N$, if $n \in N$ then $\phi_*\pi(n) = \phi_*(nN) = e_Q$ but $nN = N$ thus $N \in \ker\phi_*$, moreover, if $xN \in \ker\phi_*$ it follows that $\phi(x) = e_Q$ then $x \in N$ and therefore $xN = N$ — this implies that $\ker\phi_* = N$, which is the identity element of G/N, thus ϕ_* is injective. ♮

Corollary 7.4.17 (First isomorphism)**.** Let $\phi\colon G \twoheadrightarrow H$ be a *surjective* morphism of groups. There exists a canonical isomorphism of groups

$$G/\ker\phi \simeq H.$$

Proof. By the universal property, the induced map $\phi_*\colon G/\ker\phi \rightarrowtail H$ is already injective. From hypothesis, ϕ being surjective implies that ϕ_* is surjective since $\phi = \pi\phi_*$. ♮

Corollary 7.4.18. Let G be a group and H be a subgroup of G. Define S to be the subgroup of G consisting of the intersection of all *normal* subgroups of G containing H. Then S is normal in G and is the smallest normal subgroup of G containing H. Let Q be a group and $\phi\colon G \to Q$ be a morphism of groups such that $H \subseteq \ker\phi$. Then we have $S \subseteq \ker\phi$, and there exists a *unique* morphism $\phi_*\colon G/S \rightarrowtail Q$ such that the following diagram commutes

$$\begin{tikzcd} G \arrow[r, "\phi"] \arrow[d, two heads, "\pi"'] & Q \\ G/S \arrow[ur, tail, dashed, bend right, "\phi_*"'] & \end{tikzcd}$$

moreover, ϕ_* is *injective*.

Proof. As before, we just define $\phi_*(xS) := \phi(x)$. The commutativity follow from construction, uniqueness follows from the universal property. Moreover, S is given by the intersection of arbitrarily many normal subgroups of G containing H, thus in particular $H \subseteq S$ and, for every $g \in G$, we have $gSg^{-1} = S$ — thus S is indeed normal. ♮

Corollary 7.4.19 (Third isomorphism)**.** Let G be a group and $H \subseteq G$ be a *normal* subgroup of G. Let $Q \subseteq G$ be a subgroup *containing* H. Then, Q/H is normal in G/H if and only if Q is normal in G, if that is the case, then we have a canonical isomorphism

$$\frac{G/H}{Q/H} \simeq G/Q.$$

Proof. Suppose Q is normal in G. Since H is contained in Q, and $\ker(G \twoheadrightarrow G/Q) = Q$, we can apply the universal property of quotients Corollary 7.4.18 to the subgroup H of the kernel and find a unique induced injective morphism $G/H \rightarrowtail G/Q$ — whose kernel, on the other hand, is Q/H, thus Q/H is normal in G/H.

For the converse, let Q/H be normal in G/H, and consider the morphism given by the composition of the canonical projections

$$G \longrightarrow\!\!\!\!\!\!\rightarrow G/H \longrightarrow\!\!\!\!\!\!\rightarrow \frac{G/H}{Q/H}$$

which has a kernel given by Q — thus Q is normal in G. With that, using Corollary 7.4.18, the induced map is the wanted canonical isomorphism. ♮

In the context of the last corollary, we can visualize the propositions by the following diagram, which commutes

$$\begin{tikzcd}
* \arrow[r] & Q \arrow[r, tail] \arrow[d, "\simeq"'] & G \arrow[r, two heads] \arrow[d, "\simeq"] & G/Q \arrow[r] \arrow[d, "\mathrm{id}"] & * \\
* \arrow[r] & Q/H \arrow[r, tail] & G/H \arrow[r, two heads] & G/Q \arrow[r] & *
\end{tikzcd}$$

Corollary 7.4.20 (Second isomorphism)**.** Let G be a group, and let H and Q be subgroups of G such that $H \subseteq N_G(Q)$ — that is, H is contained in the normalizer of Q. Then we have the following canonical isomorphism

$$\frac{H}{H \cap Q} \simeq \frac{HQ}{Q}.$$

Proof. Since $H \subseteq N_G(Q)$, then $hQh^{-1} = Q$ for every $h \in H$, so that Q is normal in H. Since both H and Q are normal in H, their intersection $H \cap Q$ is also normal in H. Moreover, $HQ = QH$ and, given $hq \in HQ$, the element $q^{-1}h^{-1} \in QH = HQ$ exists, thus HQ is closed under inverses, and clearly closed under products, thus HQ is a subgroup of G. Consider now the surjective morphism $H \twoheadrightarrow HQ/Q$ given by the mapping $h \mapsto hQ$ — whose kernel is $H \cap Q$, and therefore, by the quotient universal property, there exists a unique injective morphism $H/(H \cap Q) \rightarrowtail HQ/Q$, which is also surjective by the construction of $H \twoheadrightarrow HQ/Q$. ♮

Proposition 7.4.21 (Morphism preimage preserve normality)**.** Let G and H be groups and $N \subseteq H$ be a normal subgroup. If $f \colon G \to H$ is a morphism, then $f^{-1}(N) \subseteq G$ is a normal subgroup of G.

Proof. Let $x \in G$ and $y \in f^{-1}(N)$ be any two element of G. We have $f(xyx^{-1}) = f(x)f(y)f(x)^{-1} \in N$, thus $xyx^{-1} \in f^{-1}(N)$. $\natural$

Notice that the previous proposition gets us the following commutative diagram

$$\begin{tikzcd}
G \arrow[r, "f"] & H \arrow[r, two heads, "\pi"] & H/N \\
f^{-1}(N) \arrow[u, hook] \arrow[r, "f"] & N \arrow[u, hook] &
\end{tikzcd}$$

Moreover, since $\ker \pi f = f^{-1}(N)$, we can use the universal property to obtain

$$\begin{tikzcd}
G \arrow[r, "\pi f"] \arrow[d, two heads] & H/N \\
G/f^{-1}(N) \arrow[ur, tail, dashed, bend right, "\phi"'] &
\end{tikzcd}$$

If we now impose surjectivity to the morphism f, we find that ϕ must also be surjective, giving rise to an isomorphism $G/f^{-1}(N) \simeq H/N$. The last two diagrams can be encoded in a single commutative diagram:

$$\begin{tikzcd}
0 \arrow[r] & f^{-1}(N) \arrow[r, tail] \arrow[d, "f"] & G \arrow[r, two heads] \arrow[d, "f"] & G/f^{-1}(N) \arrow[r] \arrow[d, dashed, tail, two heads, "\phi"] & 0 \\
0 \arrow[r] & N \arrow[r, tail] & H \arrow[r, two heads] & H/N \arrow[r] & 0
\end{tikzcd}$$

Centralizers & Normalizers

Definition 7.4.22 (Centralizers and normalizers)**.** Let G be a group and $S \subseteq G$ be any set of elements. We define the following groups:

(a) The *normalizer* of S is defined as the group $N(S) := \{g \in G : gSg^{-1} = S\}$.

(b) A *centralizer* of S is defined as the group $Z(S) := \{g \in G : gsg^{-1} = s, \text{ for all } s \in S\}$. The centralizer $Z(G)$ is commonly called the *center* of G.

The normalizer and centralizer are indeed groups, notice that if $g \in N(S)$, then $g^{-1} \in N(S)$ since $gSg^{-1} = S$ it follows, by multiplying on the left by g^{-1} and on the right by g, that $S = g^{-1}Sg$. Moreover, given any two elements $g, h \in N(S)$, we have

$$(gh)^{-1}N(gh) = (h^{-1}g^{-1})N(gh) = (h^{-1}Nh)(g^{-1}Ng) = S.$$

For the case of the centralizer the proof is analogous as the one just made for normalizers, where instead of S we would be considering any element $s \in S$.

In the following lemmas, let G be a group and H be a subgroup of G.

Lemma 7.4.23. If $Q \subseteq G$ is any subgroup containing H, and H is normal in Q, then $Q \subseteq N_G(H)$ — that is, the normalizer $N_G(H)$ of H is the largest subgroup of G such that H is normal.

Proof. Let K be a group and $\phi\colon Q \to K$ be a morphism of groups such that $\ker\phi = H$, then given any $q \in Q$, and any $h \in H$, we have $\phi(qhq^{-1}) = \phi(q)\phi(h)\phi(q)^{-1} = \phi(q)\phi(q)^{-1} = e_K$ — thus $qhq^{-1} \in H$ and since $qHq^{-1} = H$ from hypothesis, then $q \in N(H)$. The last statement follows clearly from the proposition. ♮

Lemma 7.4.24. If $Q \subseteq N_G(H)$ is a subgroup, then QH is a group and H is normal in QH.

Proof. First we verify that QH is indeed a group. Let $qh \in QH$ be any element. Since $Q \subseteq N_G(H)$, then $QHQ^{-1} = H$ and therefore $QH = HQ$ — hence there exists $q'h' \in QH$ such that $q'h' = q^{-1}h^{-1}$ so that QH is closed under inverses. Moreover, given any two elements $qh, q'h' \in QH$, we have $(qh)(q'h') = q(hq')h'$, but since there exists $q''h'' \in QH$ such that $q''h'' = hq'$ then $q(hq')h' = q(q''h'')h' = (qq'')(h''h') \in QH$ — thus QH is also closed under products, and therefore QH is a group.

Let $qh \in QH$ be any element and consider the group $(qh)H(h^{-1}q^{-1})$. Given any element $h' \in H$, we have that $(q^{-1}h^{-1})h'(qh) \in H$, thus $(qh)[(q^{-1}h^{-1})h'(qh)](h^{-1}q^{-1}) = h'$ thus $H \subseteq (qh)H(h^{-1}q^{-1})$. Moreover, clearly $(qh)H(h^{-1}q^{-1}) \subseteq H$ thus the equality holds — which implies that H is normal on the group QH. ♮

Epimorphisms

We lay out the most difficult part of Proposition 7.4.27 in the following lemma.

Lemma 7.4.25 (Epimorphisms are surjections)**.** In the category of groups, an epimorphism is a surjective set-function.

Proof. Let $\phi\colon H \twoheadrightarrow G$ be an epimorphism in `Grp` — our goal will be to prove that $\operatorname{im}\phi = G$.

Suppose, for the sake of contradiction, that $[G : \operatorname{im}\phi] = 2$ — so that, by Proposition 7.4.14 $\operatorname{im}\phi$ is normal in G. Now, if we consider the canonical projection and the trivial morphism $\pi, 0\colon G \rightrightarrows G/\operatorname{im}\phi$ we see that $\pi\phi = 0\phi$ — however, since ϕ is epic, this would imply in $\pi = 0$ which cannot possibly be true. Hence $[G : \operatorname{im}\phi] \geqslant 3$ necessarily and thus $\operatorname{im}\phi$ is .

Define the *set*-function $\sigma\colon G/\operatorname{im}\phi \to G/\operatorname{im}\phi$ to be a permutation of the cosets of $G/\operatorname{im}\phi$ such that $\sigma(\operatorname{im}\phi) := \operatorname{im}\phi$ is the *only* fixed point of σ — which can occur only because $[G : \operatorname{im}\phi] > 2$. We now consider the *set*-functions given by the canonical projection $\pi\colon G \twoheadrightarrow G/\operatorname{im}\phi$ and a map $\eta\colon G/\operatorname{im}\phi \to G$ defined so that $\pi\eta = \mathrm{id}_{G/\operatorname{im}\phi}$ and $\eta(\operatorname{im}\phi) := e_G$ — which is possible since π is surjective and therefore has a right-inverse.

One can define a *set*-function $\alpha\colon G \to \operatorname{im}\phi$ for which every $g \in G$ can be written uniquely as

$$g = \alpha(g)\eta\pi(g). \tag{7.6}$$

Since G is a group, the existence of α is ensured — simply define $\alpha(g) := g(\eta\pi(g))^{-1}$. The unicity of α comes from the fact that if $\beta\colon G \to \operatorname{im}\phi$ is another map such that $g = \beta(g)\eta\pi(g)$, then by the injectivity of η we conclude that $\beta = \alpha$. In the special case of $g \in \operatorname{im}\phi$ then $\eta\pi(g) = \eta(\operatorname{im}\phi) = e_G$, and thus $\alpha(g) := g$.

Define $\lambda\colon G \to G$ to be the *set*-function given by $\lambda(g) := \alpha(g)\eta(\sigma\pi(g))$, for all $g \in G$. Notice that λ is nothing but a permutation on G since every element of G is uniquely written as in Eq. (7.6) — and σ merely permutes the cosets of $G/\operatorname{im}\phi$. Let P be the group of all permutations of G and consider group morphisms $k, \ell\colon G \rightrightarrows P$ defined by

$$k(g)(x) := gx, \text{ for all } g, x \in G, \tag{7.7}$$

$$\ell(g) := \lambda^{-1}k(g)\lambda, \text{ for all } g \in G. \tag{7.8}$$

Indeed, given $g, g' \in G$ we have

$$k(gg')(x) = (gg')x = g(g'x) = k(g)(k(g')(x)),$$
$$\ell(gg') = \lambda^{-1}k(gg')\lambda = \lambda^{-1}k(g)k(g')\lambda = \ell(g)\ell(g')$$

thus both are group morphisms. The condition for $k(g) = \ell(g)$ is the same as $\lambda k(g) = k(g)\lambda$, in turn if $x \in G$ is any element, the requirement that

$$\lambda(gx) = g\lambda(x) \tag{7.9}$$

is also an equivalent condition.

For the case where $g \in \operatorname{im}\phi$ and $x \in G$ is any element, one has $\pi(gx) = \pi(g)\pi(x) = \pi(x)$ and $\alpha(gx) = g\alpha(x)$. Therefore,

$$\lambda(gx) = \alpha(gx)\eta(\sigma\pi(gx)) = g\alpha(x)\eta(\sigma\pi(x)) = g\lambda(x).$$

Using Eq. (7.9) one concludes that $k(g) = \ell(g)$ for all $g \in \operatorname{im}\phi$ — which in turn implies in $k\phi = \ell\phi$. Note however that since ϕ is an epimorphism of groups by hypothesis, we conclude that $k = \ell$ in general — thus $\lambda(gx) = g\lambda(x)$ is true for all choices of $g, x \in G$. Fix $g_0 \in G$ and consider the particular case where $x := e_G$, we thus have $\lambda(g_0) = g_0\lambda(e_G) = g_0$ — therefore, $g_0 = \lambda(g_0) = \alpha(g_0)\eta(\sigma\pi(g_0))$ from the definition of λ, and $g_0 = \alpha(g_0)\eta\pi(g_0)$ from Eq. (7.6). Combining both equations for g_0 we conclude that $\eta(\sigma\pi(g_0)) = \eta\pi(g_0)$, which in turn implies in $\sigma\pi(g_0) = \pi(g_0)$ since η is injective. Furthermore, σ was constructed so that $\operatorname{im}\phi$ was its only fixed point, that is, $\pi(g_0) = \operatorname{im}\phi$ and therefore $g_0 \in \operatorname{im}\phi$. This shows that $\operatorname{im}\phi = G$ as wanted. ♮

Remark 7.4.26. The clever reader may be tempted to think that every epimorphism in the category of groups is a split epimorphism — so that the forgetful functor `Grp` $\to$ `Set` still preserves it. However, things are not as bright as one might think.

Consider for instance the natural projective group morphism $\mathbf{Z}/5\mathbf{Z} \twoheadrightarrow \mathbf{Z}/2\mathbf{Z}$ sending $[x]_5 \mapsto [x]_2$. Although an epimorphism, such projection is not a split epimorphism — the lack of elements of order 2 in $\mathbf{Z}/5\mathbf{Z}$ does not allow the existence of non-trivial morphisms of groups $\mathbf{Z}/2\mathbf{Z} \to \mathbf{Z}/5\mathbf{Z}$.

Proposition 7.4.27 (Epimorphisms in `Grp`)**.** Let $\phi\colon G \to H$ be a group morphism. The following propositions are equivalent:

(a) The morphism ϕ is an epimorphism in `Grp`.

(b) The set function ϕ is surjective in `Set`.

Proof. For (a) $\Rightarrow$ (b), we have Lemma 7.4.25. On the other hand, the proof that (b) $\Rightarrow$ (a) is straightforward: let $f_1, f_2\colon H \rightrightarrows Q$ be any two group morphisms such that $f_1\phi = f_2\phi$, then given any $h \in H$ there exists $g \in G$ such that $\phi(g) = h$ and hence $f_1\phi(g) = f_1(h) = f_2(h) = f_2\phi(g)$ — since h was chosen arbitrarily over H, we conclude that $f_1 = f_2$. ♮

Examples & Consequences

Proposition 7.4.28 (Modular group $\mathrm{PSL}_2(\mathbf{Z})$)**.** Let $\sim$ be an equivalence relation on $\mathrm{SL}_2(\mathbf{Z})$ for which $A \sim B$ if and only if $A = \pm B$. We define the *modular group* as the quotient of $\mathrm{SL}_2(\mathbf{Z})$ by this equivalence relation, that is

$$\mathrm{PSL}_2(\mathbf{Z}) := \mathrm{SL}_2(\mathbf{Z})/\sim$$

The modular group so constructed is generated by the cosets of the matrices

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

Proof. Let's consider the canonical projection morphism $\pi\colon \mathrm{SL}_2(\mathbf{Z}) \twoheadrightarrow \mathrm{PSL}_2(\mathbf{Z})$ and let $g := \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ and $h := \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, then since g and $gh = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right]$ generate $\mathrm{SL}_2(\mathbf{Z})$, the classes $\pi(g)$ and $\pi(gh)$ generate $\mathrm{PSL}_2(\mathbf{Z})$. Moreover, since $g^2 = -I_2$, the order of $\pi(g)$ is 2 — on the other hand, $(gh)^3 = -I_2$ then $\pi(gh)$ has order 3. ♮

Proposition 7.4.29. Let G be a *finite abelian group*. If p is a prime divisor of the order of G, then there exists an element of G whose order is p.

Proof. Since p divides $|G|$, let $m \in \mathbf{Z}$ such that $|G| = pm$. We proceed via strong induction on m.

- The base case $m = 1$, one has $|G| = p$. If $g \in G$ is any element, then $\langle g \rangle$ is a subgroup of G and, by Lagrange's theorem, we have that $|\langle g \rangle|$ divides $|G|$ — thus it must be the case that $|\langle g \rangle| = p$. Therefore g is an element of order p, since p is prime.
- For the inductive hypothesis, let $m > 1$ and assume that that the proposition is true for all $m' < m$.
- We now prove the case for m, that is, $|G| = pm$. Let $g \in G$ be any element and consider the subgroup $\langle g \rangle$ of G.

  If p divides the order of g, then there exists $a \in \mathbf{Z}$ such that $pa = |g|$, hence the element $ag \in G$ has order p — if this is the case, we are done.

Otherwise, since G is abelian, then $\langle g \rangle$ is a subgroup and we may consider the quotient $G/\langle g \rangle$. By Proposition 7.4.6 we find that $|G/\langle g \rangle| = |G|/|\langle g \rangle| = pm/|g|$. Since p does not divide $|g|$ by hypothesis, then $|g|$ must divide m, therefore $m/|g| < m$. By the inductive hypothesis, the proposition is true for positive integers less than m, hence there exists $h + \langle g \rangle \in G/\langle g \rangle$ with order p. If we consider the natural projection $\pi \colon G \twoheadrightarrow G/\langle g \rangle$, by Proposition 7.3.13 we have that $|\pi(h)| = |h + \langle g \rangle| = p$ divides the order of h. If $|h| = pb$ for some $b \in \mathbf{Z}$, then the element $bh \in G$ has order p and we are done.

$\natural$

Solve Aluffi problems from section 8 (quotient groups)

7.5 Group Towers & Solvability

Change from "towers" to "series"

Definition 7.5.1 (Tower of subgroups)**.** Let G be a group. A finite sequence of groups $(G_0, \dots, G_n)$ is said to be a *tower of subgroups* of G if

$$G = G_0 \supsetneq G_1 \supsetneq \cdots \supsetneq G_n.$$

The tower may be classified as a *normal tower* if G_{j+1} is normal in G_j for every $0 \leqslant j \leqslant n-1$. The tower is said to be *abelian* (respectively, *cyclic*) if it is a normal tower and each quotient G_j/G_{j+1} is abelian (respectively, cyclic).

A direct corollary of Proposition 7.4.21 goes as follows.

Corollary 7.5.2. Let $f \colon G \to H$ be a group morphism, and $H \supsetneq H_0 \supsetneq \cdots \supsetneq H_n$ be a normal tower in H. Then, if we define subgroups $G_j \coloneqq f^{-1}(H_j)$ for every $0 \leqslant j \leqslant n$, the sequence $(G_0, \dots, G_n)$ forms a normal tower in G.

Moreover, if $(H_0, \dots, H_n)$ is an abelian tower (respectively, cyclic tower) in H, then $(G_0, \dots, G_n)$ is an abelian tower (respectively, cyclic tower) in G.

Proof. The first part is clear. For the second, notice that since for all $0 \leqslant j \leqslant n-1$ we have the following commutative diagram

$$\begin{tikzcd}
G_j \arrow[r, "f"] \arrow[d, two heads] & H_j \arrow[r, two heads] & H_j/H_{j+1} \\
G_j/G_{j+1} \arrow[urr, tail, dashed, bend right] & &
\end{tikzcd}$$

then the injection $G_j/G_{j+1} \rightarrowtail H_j/H_{j+1}$ allow us to view G_j/G_{j+1} as a subgroup of H_j/H_{j+1} — thus necessarily abelian (respectively, cyclic). $\natural$

Definition 7.5.3 (Refinement)**.** Given a tower $G = G_0 \supsetneq \cdots \supsetneq G_n$ of subgroups of G, a *refinement* of such tower is given by the insertion of a finite sequence of subgroups $(G_{n+1}, \dots, G_m)$ to the end of the tower — that is,

$$G = G_0 \supsetneq \cdots \supsetneq G_n \supsetneq G_{n+1} \supsetneq \cdots \supsetneq G_m.$$

Definition 7.5.4 (Solvable group)**.** A group G is said to be *solvable*, if there exists an *abelian tower* in G such that the last element is the trivial subgroup.

Proposition 7.5.5. The following propositions regard implications of the solvability of finite groups:

(a) Let G be a *finite* group. An abelian tower of G admits a *cyclic refinement*.

(b) Let G be a *finite solvable* group. Then G admits a *cyclic tower* whose last element is the trivial subgroup.

Proof. Let G be a finite abelian group. We'll set out to prove that G admits a cyclic tower ending with the trivial subgroup. We proceed by induction on the order of G. If $|G| = 1$, then the proposition follows trivially. Suppose, as the hypothesis of induction, that the proposition is true for $1 \leqslant |G| < n$. Let now $|G| = n$ and consider any element $x \neq e$ of G. Define the group $X := \langle x \rangle$, and $H := G/X$ — since $|H| < |G| = n$ it must be true, from hypothesis, that there exists a cyclic tower

$$H \supsetneq H_0 \supsetneq \cdots \supsetneq H_m = \{e\}.$$

Let $\pi\colon G \twoheadrightarrow H$ be the canonical projection. From Corollary 7.5.2, the sequence of subgroups $G_0 := G$, and $G_j := f^{-1}(H_j)$ for $1 \leqslant j \leqslant m$ form a cyclic tower in G, thus

$$G = G_0 \supsetneq G_1 \supsetneq \cdots \supsetneq G_m \supsetneq \{e\}.$$

is the desired cyclic tower on G. This proves both the first statement and the second. ♮

For the time being, the following theorem is a classic but probably still beyond the scope of this text, so I shall only mention it.

Theorem 7.5.6 (Feit-Thompson)**.** All finite groups of odd order are solvable.

Theorem 7.5.7. Let $N \subseteq G$ be a normal subgroup of the group G. Then G is solvable if and only if both N and G/N are solvable.

Proof. ($\Rightarrow$) Suppose G is solvable, and let $G = G_0 \supsetneq \cdots \supsetneq G_n = \{e\}$ be an abelian tower in G. For each $1 \leqslant j \leqslant n$, define $N_j := N \cap G_j$, and let $\phi_j\colon G_j \to Q$ be a group morphism with kernel G_{j+1}, then the induced morphism $\phi_j|_{N \cap G_j}\colon N \cap G_j \to Q$ has kernel $N \cap G_{j+1} = N_{j+1}$ — which proves that N_{j+1} is normal in N_j. Moreover, there exists a canonical embedding $N_j/N_{j+1} \hookrightarrow G_j/G_{j+1}$, which implies that N_j/N_{j+1} is abelian by Proposition 7.3.12. Therefore $N = N_0 \supsetneq \cdots \supsetneq N_n = \{e\}$ is an abelian tower and hence N is solvable. For the group G/N we can simply consider the tower $G/N = (G_0N)/N \supsetneq \cdots \supsetneq (G_nN)/N = (\{e\}N)/N = N$ — where $(G_{j+1}N)/N$ is surely normal in $(G_jN)/N$. Moreover, $[(G_jN)/N]\big/[(G_{j+1}N)/N] \simeq (G_jN)/(G_{j+1}N)$ which inherits the commutative structure from G_j/G_{j+1} — thus the tower is abelian and G/N is solvable.

($\Leftarrow$) Let both N and G/N be solvable and consider the abelian towers $G/N = H_0 \supsetneq \cdots \supsetneq H_n = N$, and $N = Q_0 \supsetneq \cdots \supsetneq Q_m = \{e\}$. Since every H_j is isomorphic to a subgroup H'_j of G containing N, we see that

$$G = H'_0 \supsetneq \cdots \supsetneq H'_n = N$$

is an abelian tower. Moreover, we can append the abelian tower of N to obtain the abelian tower

$$G = H'_0 \supsetneq \cdots \supsetneq H'_n = Q_0 \supsetneq \cdots \supsetneq Q_m = \{e\}.$$

This shows that G is a solvable group. $\natural$

Commutator Group

Definition 7.5.8 (Commutator group)**.** Let G be a group, we define the *commutator* group $[G, G]$ of G to be the subgroup of G generated by all elements of the form $xyx^{-1}y^{-1}$, for $x, y \in G$.

Lemma 7.5.9. The commutator group $[G, G]$ is *normal* in G.

Proof. We show that $g[G, G]g^{-1} = [G, G]$ for all $g \in G$. Let $g, x, y \in G$ be any triple of elements. If we let $p := gxg^{-1}$ and $q := gyg^{-1}$ we obtain

$$pqp^{-1}q^{-1} = (gxg^{-1})(gyg^{-1})(gxg^{-1})(gyg^{-1}) = gxyx^{-1}y^{-1}g^{-1},$$

therefore $g[G, G]g^{-1} \subseteq [G, G]$. Moreover, if $a := g^{-1}xg$ and $b := g^{-1}yg$, then

$$\begin{aligned} g[aba^{-1}b^{-1}]g^{-1} &= g[(g^{-1}xg)(g^{-1}yg)(g^{-1}xg)(g^{-1}x^{-1}g)(g^{-1}y^{-1}g)]g^{-1} \\ &= g[g^{-1}xyx^{-1}y^{-1}g]g^{-1} \\ &= xyx^{-1}y^{-1}, \end{aligned}$$

which implies in $xyx^{-1}y^{-1} \in g[G, G]g^{-1}$ and hence $[G, G] \subseteq g[G, G]g^{-1}$. $\natural$

Lemma 7.5.10. The quotient group $G/[G, G]$ is commutative.

Proof. Notice that $xy = yx$ is equivalent to $xy(yx)^{-1} = xyx^{-1}y^{-1} = e$, thus, for any given $x, y \in G$, since $xyx^{-1}y^{-1} \in [G, G]$, then $[xy] = [yx] \in G/[G, G]$ — which shows the commutativity. $\natural$

Lemma 7.5.11. Let G be any group and H be a commutative group. Any morphism of groups $f : G \to H$ has the commutator $[G, G]$ in its kernel. Therefore the following diagram commutes

$$\begin{tikzcd} G \arrow[r, "f"] \arrow[d, two heads] & H \\ G/[G,G] \arrow[ur, tail, dashed, bend right] & \end{tikzcd}$$

Proof. If H is commutative, then given any pair $x, y \in G$ we have

$$f(xyx^{-1}y^{-1}) = f(x)f(y)f(x^{-1})f(y^{-1}) = f(y)f(x)f(x)^{-1}f(y)^{-1} = e_H,$$

thus indeed $[G, G] \subseteq \ker f$. ♮

7.6 Ab — Category of Abelian Groups

Definition 7.6.1 (Category of abelian groups)**.** We define the category of abelian groups, denoted Ab, to be the category whose objects are abelian groups and group morphisms between them.

In the following propositions we are going to describe ways of telling if the group you might be interested is abelian or not.

Proposition 7.6.2. A group G is abelian if and only either one of the following conditions are satisfied:

- The map $\phi\colon G \to G$ sending $g \mapsto g^{-1}$ is a morphism of groups.
- The map $\psi\colon G \to G$ sending $g \mapsto g^2$ is a morphism of groups.

Proof. If G is abelian, then for any $g, g' \in G$ we have:

$$\phi(gg') = (gg')^{-1} = g'^{-1}g^{-1} = g^{-1}g'^{-1} = \phi(g)\phi(g'),$$
$$\psi(gg') = (gg')^2 = (gg')(gg') = g^2g'^2 = \psi(g)\psi(g').$$

That is, ϕ and ψ are morphisms. Now suppose ϕ and ψ are morphisms and consider any elements $g, g' \in G$ — we get the following relations from each of the morphisms:

$$gg' = \phi(g^{-1})\phi(g'^{-1}) = \phi(g^{-1}g'^{-1}) = \phi((g'g)^{-1}) = g'g,$$
$$(gg')^2 = \psi(gg') = \psi(g)\psi(g') = g^2g'^2.$$

Using cancellation law for the second relation we find $g'g = gg'$. This shows that G is abelian. ♮

Proposition 7.6.3. Let $\phi \in \mathrm{Mor}_{\mathtt{Grp}}(G, \mathrm{Aut}_{\mathtt{Grp}}(G))$ be the morphism of groups mapping $g \overset{\phi}{\mapsto} \gamma_g$ — as defined in Lemma 7.3.19. G is an abelian group if and only if ϕ is trivial.

Proof. Let G be an abelian group, then for all $g \in G$ the corresponding inner automorphism γ_g maps $a \overset{\gamma_g}{\longmapsto} gag^{-1} = a(gg^{-1}) = a$ hence $\gamma_g = \mathrm{id}_G$, thus ϕ is indeed trivial. Let ϕ be trivial, then for all $g \in G$ the map $\gamma_g = \mathrm{id}_G$ and therefore for all $a \in G$ we have $gag^{-1} = a$, which implies in $gag^{-1} = agg^{-1} = ag^{-1}g = gg^{-1}a = g^{-1}ga$ — that is, the group G is abelian. ♮

Now that we know some ways of identifying abelian groups, we dive deep again into the categorical foundations of the category of abelian groups Ab.

Proposition 7.6.4 (Mor$_{\mathtt{Ab}}$ abelian group)**.** Let $G, H \in \mathtt{Ab}$ be any two commutative groups. The collection of morphisms $\mathrm{Mor}_{\mathtt{Ab}}(G, H)$ forms an abelian group with a binary operation defined by $(\phi + \psi)(g) = \phi(g) +_H \psi(g)$ — where $+_H$ is the binary operation of H.

Proof. Let $\phi, \psi \in \mathrm{Mor}_{\mathtt{Ab}}(G, H)$ be any morphisms and consider elements $g, g' \in G$. From the commutativity of H it follows that

$$\begin{aligned}(\phi + \psi)(g +_G g') &= \phi(g +_G g') +_H \psi(g +_G g') \\ &= \big(\phi(g) +_H \phi(g')\big) +_H (\psi(g) +_H \psi(g')) \\ &= \big(\phi(g) +_H \psi(g)\big) +_H \big(\phi(g') +_H \psi(g')\big) \\ &= (\phi + \psi)(g) +_H (\phi + \psi)(g').\end{aligned}$$

That is, $\phi + \psi \in \mathrm{Mor}_{\mathtt{Ab}}(G, H)$. Moreover, given a morphism $f \in \mathrm{Mor}_{\mathtt{Ab}}(G, H)$, define the map $k \colon G \to H$ mapping $k(g) = f(g)^{-1}$. Notice that k is a morphism:

$$\begin{aligned}k(g +_G g') &= f(g +_G g')^{-1} \\ &= (f(g) +_H f(g'))^{-1} \\ &= f(g')^{-1} +_H f(g)^{-1} \\ &= f(g)^{-1} +_H f(g')^{-1} \\ &= k(g) +_H k(g').\end{aligned}$$

Moreover, $(f + k)(g) = f(g) +_H k(g) = f(g) +_H f(g)^{-1} = e_H$ — that is, $f + k$ is the trivial morphism $g \mapsto e_H$, i.e. k is the inverse of f. This finishes the proof that $\mathrm{Mor}_{\mathtt{Ab}}(G, H)$ is a group. For the commutativity, it follows directly from the commutativity of H: for all $\phi, \psi \in \mathrm{Mor}_{\mathtt{Ab}}(G, H)$ we have

$$(\phi + \psi)(g) = \phi(g) +_H \psi(g) = \psi(g) +_H \phi(g) = (\psi + \phi)(g).$$

♮

The following corollaries follow immediately from the construction of the category.

Corollary 7.6.5. Let $G \in \mathtt{Grp}$ and $H \in \mathtt{Ab}$. The collection of morphisms $\mathrm{Mor}_{\mathtt{Grp}}(G, H)$ forms a group under the binary operation defined above.

Corollary 7.6.6. Let $A \in \mathtt{Set}$ and $H \in \mathtt{Ab}$. The collection of morphisms $\mathrm{Mor}_{\mathtt{Set}}(A, FH)$ forms a group under the binary operation defined above — where $F \colon \mathtt{Grp} \to \mathtt{Set}$ is a forgetful functor.

Proposition 7.6.7 (Cokernel in $\mathtt{Ab}$)**.** The category of abelian groups has cokernels.

Proof. Let $\phi \colon G \to H$ be a morphism of abelian groups and Q be any other abelian group and $\alpha \colon G \to Q$ be any morphism for which $\alpha\phi = 0$ — the trivial morphism.

Since G is abelian, it follows that any subgroup of G is normal — in particular, $\operatorname{im}\phi$ is normal in G and from Corollary 7.4.18 we see that

$$\begin{tikzcd}
H \arrow[r, "\phi"] \arrow[rr, bend left, "0"] & G \arrow[r, "\alpha"] \arrow[d, two heads, "\pi"] & Q \\
& G/\operatorname{im}\phi \arrow[ur, dashed, bend right] &
\end{tikzcd}$$

thus we conclude that

$$G/\operatorname{im}\phi \simeq \operatorname{coker}\phi.$$

♮

A much more trivial proof of Proposition 7.4.27 can be achieved easily in the category of abelian groups with the help of Proposition 7.6.7 — for expository purposes we shall not use our results concerning the equivalence of epimorphisms and surjection in the category of groups, as we proved above.

Proposition 7.6.8. Let $\phi\colon G \to H$ be a morphism of *abelian groups*. The following propositions are equivalent in $\mathtt{Ab}$:

(a) The morphism ϕ is an *epimorphism*.

(b) The $\operatorname{coker}\phi$ is *trivial*.

(c) The set-function ϕ is *surjective*.

Proof.
- (a) $\Rightarrow$ (b): Let ϕ be an epimorphism and consider both the canonical projection and the trivial morphism $\pi, 0\colon G \rightrightarrows \operatorname{coker}\phi$. Notice that, from the definition of the cokernel, both maps are trivial when composed with ϕ, that is, $\pi\phi = 0\phi$ — but since ϕ is an epimorphism, then $\pi = 0$ and such a thing can only occur when $\operatorname{coker}\phi$ itself is trivial.
- (b) $\Rightarrow$ (c): Let $\operatorname{coker}\phi$ be trivial, that is, $G/\operatorname{im}\phi = \operatorname{im}\phi$, which implies that $G = \operatorname{im}\phi$ and hence ϕ is surjective.
- (c) $\Rightarrow$ (a): If ϕ is surjective, then since $\mathtt{Ab}$ is a concrete category, ϕ is an epimorphism.

♮

Proposition 7.6.9 (Coequalizers in $\mathtt{Ab}$)**.** The category of abelian groups have coequalizers. Moreover, given group morphisms $f, g\colon A \rightrightarrows B$ between abelian groups, we have

$$\operatorname{coeq}(f, g) = B/\operatorname{im}(f - g).$$

Proof. Given any abelian group G and a map $m\colon B \to G$ such that $mf = mg$, we have that, for all $a \in A$,

$$m((f-g)(a)) = m(f(a) - g(a)) = mf(a) - mg(a) = 0$$

therefore $\operatorname{im}(f-g) \subseteq \ker m$. Thus using the universal property Corollary 7.4.18 we obtain a unique morphism $n\colon B/\operatorname{im}(f-g) \to G$ such that the following diagram commutes

$$\begin{tikzcd}
B/\operatorname{im}(f-g) \arrow[d, dashed, "n"'] & B \arrow[l, two heads, "\pi"] \arrow[dl, bend left, "m"] & A \arrow[l, shift right, "f"'] \arrow[l, shift left, "g"] \\
G & &
\end{tikzcd}$$

Therefore $B/\operatorname{im}(f-g) = \operatorname{coeq}(f,g)$. ♮

Coproduct and Fiber Product

Proposition 7.6.10 (Coproduct in $\mathtt{Ab}$)**.** The direct product of abelian groups is a coproduct in $\mathtt{Ab}$. That is, for any $G, H, W \in \mathtt{Ab}$ and morphisms $f \in \operatorname{Mor}_{\mathtt{Ab}}(W, G)$ and $k \in \operatorname{Mor}_{\mathtt{Ab}}(W, H)$, there exists a unique morphism $\varphi \in \operatorname{Mor}_{\mathtt{Ab}}(W, G \times H)$ such that the following diagram commutes

$$\begin{tikzcd}
G \arrow[dr, "\iota_G"] \arrow[ddr, bend right, "f"'] & & H \arrow[dl, "\iota_H"'] \arrow[ddl, bend left, "k"] \\
& G \times H \arrow[d, dashed, "\varphi"] & \\
& W &
\end{tikzcd}$$

Where we define inclusion morphisms $g \overset{\iota_G}{\longmapsto} (g, e_H)$ and $h \overset{\iota_H}{\longmapsto} (e_G, h)$.

Proof. Since $\mathtt{Ab} \subset \mathtt{Grp}$ — that is, the category of abelian groups is a subcategory of $\mathtt{Grp}$ — then, from Proposition 7.3.4 there exists a functor $\mathtt{Ab} \to \mathtt{Set}$. Since coproducts exists in $\mathtt{Set}$ and are unique, a set-function φ exists and is unique, commuting the diagram in $\mathtt{Set}$. We now show that we can extend such set-function into a morphism of groups. Let $\varphi\colon G \times H \to W$ be the mapping $(g, h) \overset{\varphi}{\mapsto} f(g)k(h)$. Notice that

$$\begin{aligned}
\varphi((g,h)(g',h')) = \varphi(gg', hh') = f(gg')k(hh') = f(g)f(g')k(h)k(h') &= f(g)k(h)f(g')k(h') \\
&= \varphi(g,h)\varphi(g',h')
\end{aligned}$$

that is, φ is a morphism of groups. ♮

Remark 7.6.11 (Coproducts in $\mathtt{Grp}$)**.** Proposition 7.6.10 is not at all true for the category of groups. Consider for instance the cyclic groups $C_2 = \{e_x, x\}$ and $C_3 = \{e_y, y, y^2\}$. Let $\sigma_k \in S_3$, for $0 \leqslant k \leqslant 2$ be the rotation of $\{1,2,3\}$ by k, that is, the permutations represented by

$$M_{\sigma_0} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad M_{\sigma_1} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad M_{\sigma_2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Consider the embeddings $f\colon C_2 \rightarrowtail S_3$ mapping $x^k \overset{f}{\mapsto} \sigma_k$ for $k \in \{0,1\}$, and $g\colon C_3 \rightarrowtail S_3$ mapping $y^k \overset{g}{\mapsto} \sigma_k$ for $k \in \{0,1,2\}$.

Suppose, for the sake of contradiction, that $C_2 \times C_3$ is a coproduct in `Grp`, that is, exists a unique morphism $\varphi\colon C_2 \times C_3 \to S_3$ such that $f = \varphi\iota_{C_2}$ and $g = \varphi\iota_{C_3}$. Since φ is supposedly a morphism of groups,

$$\varphi(x,y) = \varphi(x,e_y)\varphi(e_x,y) = \sigma_1\sigma_1 = \sigma_2$$
$$\varphi(x,y^2) = \varphi(x,e_y)\varphi(e_x,y^2) = \sigma_1\sigma_2 = \sigma_0$$

However, $\varphi(e_x,e_y) = \sigma_0$ and on the other hand we have $\varphi(x,y)\varphi(x,y^2) = \sigma_2$, which contradicts the properties of a group morphism. This shows that there exists no such φ in `Grp` and hence $C_2 \times C_3$ is not a coproduct in `Grp`.

Although $C_2 \times C_3$ is not a coproduct in `Grp`, that doesn't mean that `Grp` has no coproducts, they just behave differently when compared with `Ab`. For instance, let $C_2 * C_3 \in$ `Grp` be defined to be the group generated by elements x and y, such that $x^2 = e$ and $y^3 = e$. We'll now show that $C_2 * C_3$ is a coproduct of C_2 and C_3 in `Grp`. Let G be any group and consider morphisms $f\colon C_2 \to G$ and $k\colon C_3 \to G$. The inclusions $\iota_{C_2}\colon C_2 \to C_2 * C_3$ and $\iota_{C_3}\colon C_3 \to C_2 * C_3$ will be naturally given maps by taking each element to itself.

Let $q \in C_2 * C_3$ be any element. We know that there exists a finite collection of coefficients $I = \{(a,b) \in \mathbf{Z}^2\}$ such that $q = \prod_{(a,b)\in I} x^a y^b$. Define $\phi\colon C_2 * C_3 \to G$ as the mapping

$$\phi(q) = \phi\left(\prod_{(a,b)\in I} x^a y^b\right) = \prod_{(a,b)\in I} f(x^a)k(y^b) = \prod_{(a,b)\in I} f(x)^a k(y)^b$$

It should be clear that this definition implies $\phi\iota_{C_2} = f$ and $\phi\iota_{C_3} = g$. Notice that $\phi(e) = f(e)k(e) = e_G$ and for all $q,p \in C_2 * C_3$ — with respective coefficients $I = \{(a,b) \in \mathbf{Z}^2\}$ and $J = \{(c,d) \in \mathbf{Z}^2\}$ — we have

$$\begin{aligned}
\phi(q)\phi(p) = \phi\left[\prod_{(a,b)\in I} x^a y^b\right]\phi\left[\prod_{(c,d)\in J} x^c y^d\right] &= \prod_{(a,b)\in I} f(x^a)k(y^b) \prod_{(c,d)\in J} f(x^c)k(y^d) \\
&= \prod_{(\alpha,\beta)\in A} f(x^\alpha)g(y^\beta) \\
&= \phi\left(\prod_{(\alpha,\beta)\in A} f(x^\alpha)g(y^\beta)\right) \\
&= \phi\left(\prod_{(a,b)\in I} x^a y^b \prod_{(c,d)\in J} x^c y^d\right) = \phi(qp)
\end{aligned}$$

that is, ϕ is a morphism of groups — where we define A as the concatenation of the coefficients I and J. We have shown that the following diagram commutes

$$\begin{tikzcd}
C_2 \arrow[dr, "\iota_{C_2}"] \arrow[ddr, bend right, "f"'] & & C_3 \arrow[dl, "\iota_{C_3}"'] \arrow[ddl, bend left, "g"] \\
& C_2 * C_3 \arrow[d, dashed, "\phi"] & \\
& G &
\end{tikzcd}$$

We can then conclude that $C_2 * C_3$, as defined above, is the coproduct of C_2 and C_3 in `Grp`.

Proposition 7.6.12 (Fiber products)**.** Fiber products exist in `Ab`. That is, given abelian groups $G, H, W \in \mathtt{Ab}$ and group morphisms $\phi \in \mathrm{Mor}_{\mathtt{Ab}}(G, W)$ and $\psi \in \mathrm{Mor}_{\mathtt{Ab}}(H, W)$. Let

$$G \times_W H \in \mathtt{Ab}_W$$

be the fiber product of ϕ and ψ in the category $\mathtt{Ab}_W \subseteq \mathtt{Ab}$[4] for which exists natural projections π_G and π_H. Let $Q \in \mathtt{Ab}$ be any abelian group and consider any morphisms $f \in \mathrm{Mor}_{\mathtt{Ab}}(Q, G)$ and $k \in \mathrm{Mor}_{\mathtt{Ab}}(Q, H)$. Then there exists a unique morphism $\varphi \in \mathrm{Mor}_{\mathtt{Ab}}(Q, P)$ such that the following diagram commutes

$$\begin{tikzcd}
Q \arrow[dr, dashed, "\varphi"] \arrow[drr, bend left, "k"] \arrow[ddr, bend right, "f"'] & & \\
& G \times_W H \arrow[r, "\pi_H"] \arrow[d, "\pi_G"'] & H \arrow[d, "\psi"] \\
& G \arrow[r, "\phi"'] & W
\end{tikzcd}$$

Proof. Define $G \times_W H = \{(g, h) \in G \times H : \phi(g) = \psi(h)\}$. Since ϕ and ψ are group morphisms, given $(g, h) \in G \times_W H$, we have

$$\phi(g^{-1}) = \phi(g)^{-1} = \psi(h)^{-1} = \psi(h^{-1}),$$

that is, $(g^{-1}, h^{-1}) \in G \times_W H$ exists and is the inverse of (g, h). Moreover, given any elements $(g, h), (g', h') \in G \times_W H$, the product $(g, h)(g', h') = (gg', hh')$ is such that

$$\phi(gg') = \phi(g)\phi(g') = \psi(h)\psi(h') = \psi(hh'),$$

thus $(gg', hh') \in G \times_W H$. This shows that $G \times_W H$ is indeed a group and since $G \times H$ is abelian, so is the fiber product defined above.

[4]As a reminder: $\mathtt{Ab}_W$ is the category whose objects are morphisms $f \in \mathrm{Mor}_{\mathtt{Ab}}(-, W)$. If $f \colon G \to W$ and $g \colon H \to W$ are objects of $\mathtt{Ab}_W$, a morphism $h \in \mathrm{Mor}_{\mathtt{Ab}_W}(f, g)$ is such that $hg = f$.

From the forgetful functor $\mathtt{Ab} \to \mathtt{Set}$ we know that there exists a unique set-function φ such that the diagram commutes in $\mathtt{Set}$. If we define φ as the mapping $q \xmapsto{\varphi} (f(q), k(q))$ we can see that it preserves the group structure, since

$$\begin{aligned} \varphi(qq') &= (f(qq'), k(qq')) \\ &= (f(q)f(q'), k(q)k(q')) \\ &= (f(q), k(q))(f(q'), k(q')) \\ &= \varphi(q)\varphi(q'), \end{aligned}$$

that is, $\varphi \in \operatorname{Mor}(\mathtt{Ab})$. ♮

Direct Sums and Free Groups

Definition 7.6.13 (Direct sum of abelian groups)**.** Let $\{G_j\}_{j\in J}$ be an indexed collection of abelian groups. We define their direct sum as the collection of tuples $(g_j)_{j\in J}$ such that $g_j \neq e_{G_j}$ for only finitely many indexes $j \in J$ — that is, set of tuples with finite support. We denote the direct sum of $\{G_j\}_{j\in J}$ as $\bigoplus_{j\in J} G_j$ and the group structure of the direct sum is defined naturally as $(x_j)_{j\in J} + (y_j)_{j\in J} = (x_j +_{G_j} y_j)_{j\in J}$.

Proposition 7.6.14 (Direct sum universal property)**.** The direct sum defined in Definition 7.6.13 satisfies the universal property of direct sums. In other words, let $\{G_j\}_{j\in J}$ be a collection of abelian groups and H be any abelian group, in addition, consider the collection of morphisms $\{\phi_j \in \operatorname{Mor}_{\mathtt{Grp}}(G_j, H)\}$. There exists a unique morphism $\psi\colon \bigoplus_{j\in J} G_j \to H$ such that the following diagram commutes

$$\begin{tikzcd} G_j \arrow[d, tail, "\iota"] \arrow[dr, bend left, "\phi_j"] & \\ \bigoplus_{j\in J} G_j \arrow[r, dashed, "\psi"] & H \end{tikzcd}$$

for every $j \in J$ — where $\iota_j\colon G_j \to \bigoplus_{j\in J} G_j$ is the natural inclusion, mapping $x_{j_0} \xmapsto{\iota_j} (x_j)_{j\in J}$ such that $x_j = x_{j_0}$ for $j = j_0$ and $x_j = e_{G_j}$ for $j \neq j_0$.

Proof. Let $\psi\colon \bigoplus_{j\in J} G_j \to H$ be the map defined by $\psi(x) = \prod_{j\in J} \phi_j(x_j)$, where $x = (x_j)_{j\in J} \in \bigoplus_{j\in J} G_j$ is any element. Notice that $\psi(x)$ is therefore a finite product of elements of H and from the group structure of the direct sum, we find that ψ is clearly a group morphism. Moreover, ψ satisfies the commutativity $\psi\iota_j = \phi_j$ for each $j \in J$. Since such definition of ψ defines the image of every element of its domain, ψ is the unique morphism making the diagram commute. ♮

If G is an abelian group and $A, B \subseteq G$ are subgroups such that $A \cap B = 0$ and $A + B = G$ — that is, every $g \in G$ can be written as $g = a + b$ for some $a \in A$ and $b \in B$ — then, in the context of abelian groups, we'll denote this fact shortly by $G = A \oplus B$.

Just like a vector space, we can define a basis of an abelian group by means of the ring $\mathbf{Z}$. Moreover, if an abelian group has a basis, then we say that it is free.

Definition 7.6.15 (Basis)**.** Let G be an abelian group. We say that a non-empty collection $\{g_j\}_{j\in J}$ is a basis for G if, given an element $g \in G$, there exists a unique tuple of coefficients $(a_j)_{j\in J} \in \bigoplus_{j\in J} \mathbf{Z}$ such that $g = \sum_{j\in J} a_j g_j$.

Therefore the existence of a basis allow us to couple the coefficients coming from $\bigoplus_{j\in J} \mathbf{Z}$ to any element of G in a unique way, which induces a natural isomorphism

$$G \simeq \bigoplus_{j\in J} \mathbf{Z}.$$

Definition 7.6.16 (Abelian free group)**.** An abelian group is said to be free if it allows a basis.

Equivalently as we did with free vector spaces, we can build free abelian groups out of sets, say S, by analysing maps $S \to \mathbf{Z}$ with finite support — the collection of those will be likewise denoted by $\mathbf{Z}^{\oplus S}$, which is a group under pointwise addition. For each $s \in S$ we define a map $\mathbf{s} \in \mathbf{Z}^{\oplus}$ by the mapping

$$\mathbf{s}(x) := \begin{cases} 1, & \text{for } x = s \\ 0, & \text{otherwise} \end{cases}$$

Then, any element $\phi \in \mathbf{Z}^{\oplus S}$ can be written as a linear combination of finitely many maps $\mathbf{s}$ (which is possible because of the finite support of ϕ) that is, for some $a_j \in \mathbf{Z}$ for each $1 \leqslant j \leqslant n$, we have

$$\phi = \sum_{j=1}^{n} a_j \mathbf{s}_j, \text{ mapping } s \overset{\phi}{\mapsto} \begin{cases} a_j, & \text{if } s = s_j \text{ for some } 1 \leqslant j \leqslant n \\ 0, & \text{otherwise} \end{cases}$$

Moreover, such choice of coefficients $a_j \in \mathbf{Z}$ is unique. Let $\{b_j\}_{j=1}^n \subseteq \mathbf{Z}$ be another set of coefficients with the same property, then in particular $\sum_{j=1}^n (a_j - b_j)\mathbf{s}_j = 0$ and since $\mathbf{s}_j$ are all non-zero maps, we find that $b_j = a_j$. We also define a map $\iota_S\colon S \rightarrowtail \mathbf{Z}^{\oplus S}$ by pairing $s \mapsto \mathbf{s}$.

Proposition 7.6.17 (Free `Ab` universal property)**.** Let S be a set. Consider G to be any abelian group and a set-function $g\colon S \to G$, then there exists a unique morphism of groups $g_*\colon \mathbf{Z}^{\oplus S} \to G$ such that

$$\begin{tikzcd} S \arrow[r, "g"] \arrow[d, hook, "\iota_S"'] & G \\ \mathbf{Z}^{\oplus S} \arrow[ur, dashed, "g_*"'] & \end{tikzcd}$$

is a commutative diagram.

Proof. Define g_* by $g_*(\sum_{s\in S} a_s \mathbf{s}) := \sum_{s\in S} a_s g(s)$. Then $g_* \iota_S(s) = g_*(\mathbf{s}) = g(s)$. Moreover, it follows from the construction that $g_*(\sum_{s\in S} a_s \mathbf{s}) = \sum_{s\in S} a_s g_*(\mathbf{s})$ thus g_* is a group morphism. If we let $f\colon \mathbf{Z}^{\oplus S} \to G$ be a morphism satisfying such commutativity, we'll see that, since $f_*(\mathbf{s}) = g(s) = g_*(\mathbf{s})$, since any element of $\mathbf{Z}^{\oplus S}$ can be written uniquely as a linear combination of the basis $\{\mathbf{s}\}_{s\in S}$, then f_* and g_* agree in every point of the domain — hence $f_* = g_*$. ♮

Corollary 7.6.18. Let $f\colon X \to Y$ be a set-function between sets X and Y. Then, there exists a unique morphism $\overline{f}\colon \mathbf{Z}^{\oplus X} \to \mathbf{Z}^{\oplus Y}$ such that the following diagram commutes

$$\begin{tikzcd} X \arrow[r, hook, "\iota_X"] \arrow[d, "f"'] & \mathbf{Z}^{\oplus X} \arrow[d, dashed, "\overline{f}"'] \\ Y \arrow[r, hook, "\iota_Y"] & \mathbf{Z}^{\oplus Y} \end{tikzcd}$$

Proof. Let $\overline{f}$ be defined by $\overline{f}(\sum a_x \mathbf{x}) := \sum a_x \iota_Y f(x)$ — so that, in particular, $\overline{f}(\mathbf{x}) = \iota_Y f(x)$, that is, the diagram commutes. Let $\overline{g}\colon \mathbf{Z}^{\oplus X} \to \mathbf{Z}^{\oplus Y}$ be a morphism such that the diagram commutes, then necessarily $\overline{g}(\mathbf{x}) = \iota_Y f(x) = \overline{f}(\mathbf{x})$, that is, $\overline{g}$ and $\overline{f}$ agree on the basis of $\mathbf{Z}^{\oplus X}$, thus $\overline{g} = \overline{f}$. ♮

Notation 7.6.19. When it's not confusing, we can even drop the notation $\mathbf{s}$ and instead identify the elements as $s \in \mathbf{Z}^{\oplus S}$, so that $\sum a_s s := \sum a_s \mathbf{s}$. Moreover, for now on, we'll refer to the *free abelian group generated by* S, that is, $\mathbf{Z}^{\oplus S}$, as $F_{\mathtt{Ab}}(S)$ — this is motivated by the fact that the notation $\mathbf{Z}^{\oplus S}$, although cool, may be a rather obscure way of talking about a group. The elements of S, in particular, will be referred to as *free generators*.

Proposition 7.6.20. Let G be any abelian group

Continue: mini exercises on free ab grp

Continue direct products

7.7 Free Groups

Let S be a set. We define the category C as composed of objects that are set-functions $f\colon S \to G$, denoted $(f, G) \in \mathsf{C}$, where $G \in \mathtt{Grp}$. Given another object $(g, H) \in \mathsf{C}$, a morphism between $f \to g$ is a morphisms of groups $\phi\colon G \to H$ such that the following diagram commutes

$$\begin{tikzcd} & S \arrow[dl, "f"'] \arrow[dr, "g"] & \\ G \arrow[rr, "\phi"] & & H \end{tikzcd}$$

Proposition 7.7.1 (Free group universal property)**.** Given a set S we define the free group of S to be the initial object $(\iota, F(S))$ in the category C defined above. In other words, $F(S)$ is said to be the free group of S if there exists a set-functions ι such that for all $(f, G) \in \mathsf{C}$ there exists a unique morphism $f \to \iota$ given by a morphism of groups $\phi\colon F(S) \to G$ such that the diagram commutes

$$\begin{tikzcd} S \arrow[r, "f"] \arrow[d, hook, "\iota"'] & G \\ F(S) \arrow[ur, dashed, bend right, "\phi"'] & \end{tikzcd}$$

Although we defined the universal property of free groups, there is still no certainty that such objects shall in fact exist. We now proceed by proving that any set has a corresponding free group — which is unique up to isomorphism. To do that, we first prove the following lemma.

Lemma 7.7.2. Let G be a group generated by the set-function $f\colon S \to G$, where S is a set. Then there exists an indexing set I and a collection of indexed groups $\{G_i\}_{i \in I}$ for which there exists $i \in I$ such that $G \simeq G_i$.

Proof. Since $G = \langle \operatorname{im} f \rangle$ then, if S is a finite set, it follows $\operatorname{im} f$ is also finite, hence G is either finite or enumerably infinite, since every element of G is given by a finite product of elements in $\operatorname{im} f$ and their inverses. On the other hand, if S is infinite, then $\operatorname{im} f$ can be either finite or infinite, on both of these cases we find that G can be either finite or infinite — that is, in both of these cases we have $|G| \leqslant |S|$.

Define X to be an infinitely enumerable set if S is finite, on the other hand, if S is infinite, let $|X| = |S|$. Let $A \subseteq X$ be non-empty and define $\Gamma(A)$ to be the collection of binary operations $\gamma\colon A \times A \to A$ such that (A, γ) is a group. Define $\mathcal{A} = \{(A, \gamma) : A \subseteq X, \gamma \in \Gamma(A)\}$, the collection of groups on subsets of the set X.

We now show that $\mathcal{A}$ satisfies the proposition. The set X is defined so that every possible product of elements of S can be injectively assigned to an element of X, that is, for every collection B of finite sequences of elements of S, there exists an injective set-function $i\colon B \rightarrowtail X$. Since $\langle \operatorname{im} f \rangle$ can be seen set-wise as a collection of finite sequences of elements of S, then there exists a bijection $j\colon \langle \operatorname{im} f \rangle \xrightarrow{\simeq} B$ — for some B as defined above. Since there exists an injection $i\colon B \rightarrowtail X$, the induced map $i\colon B \xrightarrow{\simeq} \operatorname{im} i \subseteq X$ is a bijection and hence the composition $ij\colon \langle \operatorname{im} f \rangle \xrightarrow{\simeq} \operatorname{im} i$ is also bijective. Then, given $\gamma \in \Gamma(\operatorname{im} i)$ the group $(\operatorname{im} i, \gamma)$ is isomorphic to $\langle \operatorname{im} f \rangle = G$. ♮

Proposition 7.7.3 (Every set has a unique free group)**.** Let S be any set. Then there exists a free group $(\iota, F(S))$, unique up to isomorphism, such that $F = \langle \operatorname{im} \iota \rangle$ and ι is injective.

Proof. Let I be an indexing set and $\{G_i\}_{i \in I}$ be an indexed collection of groups. Define

$$F_0 = \prod_{i \in I} \prod_{\ell \in \operatorname{Mor}_{\mathsf{Set}}(S, G_i)} G_i \times \{\ell\},$$

and consider the set-function $\iota_0\colon S \to F_0$ defined by mapping the elements $s \in S$ to a tuple of pairs whose j-th component would be $(\ell_j(s), \ell_j) \in \prod_{\ell \in \operatorname{Mor}_{\mathsf{Set}}(S, G_j)} G_j \times \{\ell\}$.

Consider G to be a group and let $g: S \to G$ generate G. From Lemma 7.7.2 we find that there exists a group $G_j \in \{G_i\}_{i \in I}$ such that there is an isomorphism $\phi: G \xrightarrow{\simeq} G_j$. Moreover, the set-function $\psi = \phi g: S \to G_j$ is an element of $\bigcup_{i \in I} \mathrm{Mor}_{\mathtt{Set}}(S, G_i)$. Define the projection map $\pi_{j,\psi}: F_0 \twoheadrightarrow G_j \times \{\psi\}$. We can now define a map $\psi_* = \phi^{-1}\pi_{j,\psi}: F_0 \to G$ such that the following diagram commutes

$$\begin{tikzcd} S \arrow[rr, "\iota_0"] \arrow[d, "g"'] & & F_0 \arrow[d, two heads, "\pi_{j,\psi}"] \arrow[dll, "\psi_*"'] \\ G \arrow[rr, "\simeq", "\phi"'] & & G_j \times \{\psi\} \end{tikzcd}$$

Define the subgroup $F = \langle \operatorname{im} \iota_0 \rangle$ of F_0 and let $\iota: S \to F$ be the set-function mapping $\iota(s) = \iota_0(s)$. Also, let $g_*: F \to G$ be the restriction $g_* = \psi_*|_F = \phi^{-1}\pi_{j,\psi}|_F$ — which immediately implies in the uniqueness of g_*. Summarizing, we may view this construction as the following diagram

$$\begin{tikzcd} S \arrow[r, hook, "\iota"] \arrow[d, "g"'] & F \arrow[dl, dashed, "g_*"] \\ G & \end{tikzcd}$$

We see that (ι, F) satisfies Proposition 7.7.1 and thus is a free group of S. Moreover, since the construction works with no restriction on the choice of group G and set-function g, given any $s, s' \in S$ distinct elements, define $G = \langle s, s' \rangle$ and let $g: S \to G$ be any set-function such that $g(s) = s$ and $g(s') = s'$. We see that for the diagram to commute, it is necessary that $\iota(s) \neq \iota(s')$, otherwise g would not be equal to the composition $g_*\iota$. This shows that for any two distinct elements of S, their image under ι is also distinct — that is, ι is injective on S, which proves the last assertion of the proposition. ♮

Corollary 7.7.4. Given a set S, its corresponding free group is unique up to isomorphism.

Proof. It suffices to see that if F is a free group of S, then it is the initial object of the category $\mathcal{C}$, we defined above. From this, we can use Proposition 1.2.28 to obtain the immediate conclusion that F is unique up to isomorphism in $\mathcal{C}$. ♮

Notice since every set has a corresponding free group via the latter theorem, we can view the free group as a covariant functor $F: \mathtt{Set} \to \mathtt{Grp}$ such that $F(X)$ is the free group of $X \in \mathtt{Set}$ — which is unique up to isomorphism — and, given a set-function $\phi: X \to Y$ between sets X and Y, we assign $F(f) = \phi_*: F(X) \to F(Y)$. This whole construction is such that the following diagram commutes

$$\begin{tikzcd} X \arrow[r, "\iota_X"] \arrow[d, "\phi"'] \arrow[dr] & F(X) \arrow[d, "\phi_*"] \\ Y \arrow[r, "\iota_Y"'] & F(Y) \end{tikzcd} \tag{7.10}$$

In the above diagram we see that $F(X)$ is the free group of X generated by the image of set-function ι_X, which induces an injection of the set-elements of X into a group structure. These elements are called the free generators of $F(X)$. The same remark is true for $F(Y)$. Moreover, the functor F preserves injectivity and surjectivity of the maps ϕ, which is stated in the following lemma.

Lemma 7.7.5. If the map $\phi\colon X \to Y$ is a surjective (or injective) set-function, then $\phi_*\colon F(X) \to F(Y)$ is a surjective (or injective) morphism of groups.

Proof. From Eq. (7.10) we have $\phi_* \iota_X = \iota_Y \phi$, hence, given any element $\gamma \in \operatorname{im} \iota_Y \subseteq F(Y)$, there exists $x \in X$ for which $\iota_Y \phi(x) = \gamma$ — which is true only because ϕ is surjective — thus $\phi_* \iota_X(x) = \gamma$. Since $F(Y) = \langle \operatorname{im} \iota_Y \rangle$, it follows that, for any given element $\gamma \in F(Y)$, there exists a finite sequence of elements $(\gamma_j)_{j=1}^n$ with $\gamma_j \in \operatorname{im} \iota_Y$ for which $\gamma = \prod_{j=1}^n \gamma_j$. From the fact that $\gamma_j \in \operatorname{im} \iota_Y$, we are able to a corresponding sequence $(x_j)_{j=1}^n$ of elements $x_j \in X$ such that $\iota_Y \phi(x_j) = \gamma_j$ for each $1 \leqslant j \leqslant n$ — which is always possible since $\gamma_j \in \operatorname{im} \iota_Y$. Then we obtain

$$\phi_* \left(\prod_{j=1}^n \chi_j \right) = \prod_{j=1}^n \phi_*(\chi_j) = \prod_{j=1}^n \gamma_j = \gamma$$

and therefore $\prod_{j=1}^n \chi_j \in \phi_*^{-1}(\gamma)$, which proves that ϕ_* is surjective.

Now we prove the second part of the lemma, let $\phi\colon X \rightarrowtail Y$ be injective. Given distinct elements $a, b \in X$, since ι_X and ι_Y are injective, then $\iota_Y \phi(a) \neq \iota_Y \phi(b)$ and thus $\phi_* \iota_X(a) \neq \phi_* \iota_X(b)$ — that is, ϕ_* is injective on the set $\operatorname{im} \iota_X$ and since ϕ_* is a morphism of groups and multiplication is preserved, it follows that ϕ_* is in fact injective in $\langle \operatorname{im} \iota_X \rangle = F(X)$. Thus ϕ_* is injective. ♮

Proposition 7.7.6. If the cardinality of the sets X and Y are equal, the free groups $F(X)$ and $F(Y)$ are isomorphic.

Proof. Notice that if $|X| = |Y|$ then there exists a bijection $f\colon X \to Y$ between such sets. From Lemma 7.7.5 we see that the induced morphism of groups $f_*\colon F(X) \to F(Y)$ is both injective and surjective — thus bijective. Since bijections are isomorphisms in $\mathtt{Grp}$, it follows that $F(X) \simeq F(Y)$. ♮

Proposition 7.7.7 (Canonical factorization)**.** Every group is the image of a morphism on a free group. That is, given a group G, there exists a canonical surjective morphism of groups $F \twoheadrightarrow G$ such that F is a free group.

Proof. Let $S = \{g \in G\}$ be a set and consider the identity set-function $f\colon S \to G$. The induced group morphism $f_*\colon F(S) \to G$ is simply a projection, establishing a surjection between $F(S)$ and G and hence $f_*(F(S)) = G$. ♮

Our construction of free groups got rather abstract and avoided any concrete identification of the elements of a free group — the intention being of creating the most general setting possible. In fact, our construction is equivalent to that of the standard notion regarding the free group $F(S)$ as composed of finite strings of elements of a set S together with their inverses — forming the so called "*words*". This idea boils down to the concept of the elements of S being *generators* of the free group $F(S)$, while the collection of elements of $F(S)$ itself are the *relations* between the generator elements — which provide the structure of a group to $F(S)$. In that view, we define the following concept.

Definition 7.7.8 (Groups from generators and relations)**.** Let S be a set and $R \subseteq F(S)$ be any subset of the elements of the free group. Let N be the smallest subgroup of $F(S)$ containing R. Then we define $F(S)/N$ to be the group determined by the generators S and the relations R.

Coproducts in $\mathtt{Grp}$

We now get back to the idea of constructing coproducts, not only in $\mathtt{Ab}$, but also in $\mathtt{Grp}$. We exemplified, in the infamous example coming from Remark 7.6.11, that such construction doesn't come trivially without the requirement of commutativity. Now we are ready for the construction of coproducts in $\mathtt{Grp}$, which will be supported by the ideas developed on free groups.

Proposition 7.7.9. The category of groups $\mathtt{Grp}$ has coproducts.

Proof. Let J be an indexing set for which there exists a corresponding indexed collection of groups $\{G_j\}_{j\in J}$. Similarly, let $\{S_j\}_{j\in J}$ be a collection of sets for which $S_j = G_j$ if G_j is an infinite group, otherwise, if G_j happens to be finite, we let S_j be any infinitely enumerable set (for instance, $\mathbf{N}$). We define now S to be any set such that $|S| = |\coprod_{j\in J} S_j|$.

Define Γ as the collection of all binary operation $\gamma\colon S\times S \to S$ for which the pair (S,γ) is a group — which we'll shortly denote by S_γ. For each relation $\gamma\in\Gamma$, define the collection

$$\Phi_\gamma = \{\phi = \{\varphi_j \in \mathrm{Mor}_{\mathtt{Grp}}(G_j, S_\gamma) : j \in J\}\},$$

that is, the collection composed of elements ϕ which are themselves collections of group morphisms $G_j \to S_\gamma$ for each $j\in J$. For each $\phi\in\Phi_\gamma$, the product $S_\gamma\times\{\phi\}$ is a group with the operation defined as $\gamma^*((x,\phi),(y,\phi)) = (\gamma(x,y),\phi)$.

Now we construct a group F_0 defined as

$$F_0 = \prod_{\gamma\in\Gamma}\prod_{\phi\in\Phi_\gamma} S_\gamma\times\{\phi\},$$

which has a binary operation $F_0\times F_0 \to F_0$ which is defined naturally as applying γ^* to the corresponding $S_\gamma\times\{\phi\}$ factor of the tuples.

Notice that we carried over the collection ϕ to be able to relate F_0 to the groups $\{G_j\}_{j\in J}$ that we started with — to realize that, define the collection of morphisms

$$I = \{\iota_j \in \mathrm{Mor}_{\mathtt{Grp}}(G_j, F_0) : j\in J\},$$

such that the mapping ι_j is defined by mapping elements $g \in G_j$ to the tuple of F_0 whose $S_\gamma \times \{\phi\}$ factor is given by $(\varphi_j(g), \phi)$ — recall that $\varphi_j \in \phi$ is a morphism of groups $\varphi_j\colon G_j \to S_\gamma$. So far, what we have is the following scenario

$$\begin{tikzcd} G_j \arrow[r, "\iota_j"] \arrow[dr, bend right, "\varphi_j"'] & F_0 \\ & S_\gamma \end{tikzcd}$$

We set out to construct a group that would satisfy the coproduct universal property. We need to somehow modify both groups F_0 and S_γ so that we have a unique arrow for any choice of a group G in place of S_γ.

We start out by letting G be any group and considering the collection

$$\Psi = \{\psi_j \in \mathrm{Mor}_{\mathtt{Grp}}(G_j, G) : j \in J\},$$

with the purpose of replacing both S_γ and ϕ, respectively. Let $H \subseteq G$ be the subgroup of G generated by the images the morphisms of Ψ — that is, $H = \langle \bigcup_{j\in J} \operatorname{im} \psi_j \rangle$. It is natural to see that since H can be interpreted as the group whose elements are finite product of elements from $\{G_j\}_{j\in J}$ — and some of those can even be finite groups — then $|H| \leqslant |S|$, since the latter was constructed so that S had a cardinality bigger than or equal to the set-theoretic coproduct $\coprod_{j\in J} G_j$. From Lemma 7.7.2 we find that there exists $S_\gamma \in \{S_\gamma\}_{\gamma\in\Gamma}$ and an isomorphism of groups $\eta\colon S_\gamma \xrightarrow{\simeq} H$.

Let $i\colon H \hookrightarrow G$ be the canonical inclusion sending $H \ni g \mapsto g \in G$. Let $\phi \in \Phi_\gamma$ be the collection of morphisms such that $i\eta\varphi_j = \psi_j$, that is, the following diagram commutes for all $j \in J$

$$\begin{tikzcd} G_j \arrow[r, "\psi_j"] \arrow[d, "\varphi_j"'] & G \\ S_\gamma \arrow[r, "\sim", "\eta"'] & H \arrow[u, hook, "i"'] \end{tikzcd}$$

Define $p_{(\gamma,\phi)}\colon F_0 \twoheadrightarrow S_\gamma$ to be a projection mapping the $(s, \phi) \in S_\gamma \times \{\phi\}$ factor into $s \in S_\gamma$. Define $\psi_*\colon F_0 \to G$ to be the morphism of groups such that the following diagram commutes

$$\begin{tikzcd} F_0 \arrow[r, "\psi_*"] \arrow[d, two heads, "p_{(\gamma,\phi)}"'] & G \\ S_\gamma \arrow[r, "\sim", "\eta"'] & H \arrow[u, hook, "i"'] \end{tikzcd}$$

Notice that, by Section 7.7, we obtain — for all $g \in G_j$

$$i\eta p_{(\gamma,\phi)}\iota_j(g) = i\eta p_{(\gamma,\phi)}((\dots, (\varphi_j(g), \phi), \dots)) = i\eta\varphi_j(g) = \psi_j(g)$$

But, from Section 7.7, since $i\eta p_{(\gamma,\phi)} = \psi_*$, then we conclude that $\psi_* \iota_j = \psi_j$, that is, the following diagram commutes

$$\begin{tikzcd} G_j \arrow[r, "\iota_j"] \arrow[dr, bend right, "\psi_j"'] & F_0 \arrow[d, "\psi_*"] \\ & G \end{tikzcd}$$

Let $F \subseteq F_0$ be the subgroup $F = \langle \bigcup_{j \in J} \operatorname{im} \iota_j \rangle$. Then, if we consider the restriction $\psi_* \colon F \to G$, the morphism ψ_* is now uniquely defined by the collection of morphisms ψ, that is, the diagram

$$\begin{tikzcd} G_j \arrow[r, "\iota_j"] \arrow[dr, bend right, "\psi_j"'] & F \arrow[d, dashed, "\psi_*"] \\ & G \end{tikzcd}$$

commutes for every $j \in J$ — hence F is the coproduct in the category of groups. ♮

Exercise 7.12 Aluffi

7.8 Group Actions on Sets

Definition 7.8.1 (Group action)**.** Let G be a group and $A \in \mathsf{C}$ be an object in some category C. An *action* of G on the object A is a *group morphism*

$$G \longrightarrow \operatorname{Aut}_{\mathsf{C}}(A),$$

that is, the elements of G, which are automorphisms in the category $\mathsf{B}G$, define automorphisms in A—this shows that group actions are functors $\mathsf{B}G \to \mathsf{C}$ (see Example 1.4.12 for more). Left actions are covariant functors, while right actions are contravariant. Naturally the action is said to be *faithful* if the functor is faithful, that is, if $G \rightarrowtail \operatorname{Aut}_{\mathsf{C}}(A)$ is injective.

Definition 7.8.2. Let G be a group acting on a *set* X. We define the following concepts concerning such action:

(a) The action is said to be *free* if the identity e_G is the only element fixing *any* of the elements of X. In other words, G acts freely on X if given any $g \in G$ for which there exists $x \in X$ with $g \cdot x = x$, then $g = e$ is the identity element.

(b) An action is said to be *effective* if the *only* member of G acting trivially is the identity e_G.

Remark 7.8.3. Mind you, free actions are effective, but not all effective actions are free.

Definition 7.8.4 (Conjugation)**.** Let G be a group. We define a *conjugation* to be a group action of G on itself, $G \times G \to G$, mapping

$$(g, h) \longmapsto ghg^{-1}.$$

Theorem 7.8.5 (Cayley)**.** Every group acts faithfully on *some* set. That is, there exists a set A and an injective action $G \rightarrowtail \mathrm{Aut}_{\mathsf{Set}}(A) = S_A$—where S_A is the symmetry group of A—making G a subgroup of S_A.

Proof. Simply consider the action $G \to \mathrm{Aut}_{\mathsf{Grp}}(G)$ mapping $g \mapsto f_g$, where $f_g(x) = gx$ is the left-multiplication by g. This action is certainly faithful, making G isomorphic to a subgroup of $\mathrm{Aut}_{\mathsf{Grp}}(G)$. ♮

Definition 7.8.6 (Opposite group)**.** Given a group G, we define its *opposite group* G^{op} to be composed of the elements of G, and endowed with a contravariant action $G \times G \to G$ mapping $(g, h) \mapsto g \cdot h \coloneqq hg$.

Corollary 7.8.7. The following are properties relating a group G with its opposite group:

(a) The set-function $\phi\colon G^{\mathrm{op}} \to G$ mapping $g \mapsto g$ is an isomorphism of groups $G^{\mathrm{op}} \simeq G$ if and only if G is commmutative.

(b) There exists a natural isomorphism of groups $G \simeq G^{\mathrm{op}}$ even when G is non-commutative.

Proof. We first prove item (a). If the said map is an isomorphism, then for all $g, h \in G$ we have

$$gh = \phi(gh) = \phi(h \cdot g) = \phi(h)\phi(g) = hg,$$

thus G is commutative. Conversely, if we assume that G is commutative then the map, besides being bijective, is also a group morphism, since

$$\phi(g \cdot h) = \phi(hg) = hg = gh = \phi(g)\phi(h).$$

For item (b), define a map $\psi\colon G \to G^{\mathrm{op}}$ to be given by $g \mapsto g^{-1}$, which is bijective since inverses are unique. The set-function ψ is also a group morphism since

$$\psi(gh) = (gh)^{-1} = h^{-1}g^{-1} = g^{-1} \cdot h^{-1} = \psi(g) \cdot \psi(h).$$

Therefore $G \simeq G^{\mathrm{op}}$ via ψ. ♮

Proposition 7.8.8 (Left to right & back)**.** Let G be a group and A be a set. Given any *left-action* $\sigma\colon G \to \mathrm{Aut}_{\mathsf{C}}(A)$, we can turn σ into a unique corresponding *right-action* $\sigma^{\mathrm{op}}\colon G \to \mathrm{Aut}_{\mathsf{C}}(A)$ given by $\sigma^{\mathrm{op}}(g)(a) = \sigma(g^{-1})(a)$ for any $g \in G$ and $a \in A$. The conversion of a right-action into a left-action is also unique.

Proof. Let $g, h \in G$ and $a \in A$ be any elements. The map σ^{op} is indeed a right action:

$$\sigma^{\mathrm{op}}(gh)(a) = \sigma((gh)^{-1})(a) = \sigma(h^{-1}g^{-1})(a) = \sigma(h^{-1})(\sigma(g^{-1})(a)) = \sigma^{\mathrm{op}}(h)(\sigma^{\mathrm{op}}(g)(a)).$$

The uniqueness comes from the fact that inverses are unique. ♮

Definition 7.8.9 (Transitive action)**.** A group G is said to act *transitively* on a set A if for all pairs of elements $a, b \in A$, there exists a group element g such that $ga = b$

Definition 7.8.10 (Orbit & Stabilizer)**.** Given a group action $G \to \mathrm{Aut}_{\mathtt{Set}}(A)$ on a set A, the *orbit* of an element $a \in A$ under the action of G is defined to be the set

$$\mathrm{Orb}_G(a) \coloneqq \{ga : g \in G\} \subseteq A.$$

The *stabilizer* of a under the action of G is the subgroup

$$\mathrm{Stab}_G(a) \coloneqq \{g \in G : ga = a\} \subseteq G.$$

Let σ be any action of G on A. Stabilizers indeed define a subgroup of G, notice that if $g, h \in \mathrm{Stab}_G(a)$, then $\sigma(gh)(a) = \sigma(g)(\sigma(h)(a)) = \sigma(g)(a) = a$ thus $gh \in \mathrm{Stab}_G(a)$. Moreover, since σ is a group morphism, we have $\sigma(g^{-1})(a) = \sigma(g)^{-1}(a) = a$ thus $g^{-1} \in \mathrm{Stab}_G(a)$.

Definition 7.8.11 (G-$\mathtt{Set}$ category)**.** Given a group G, we define a category G-$\mathtt{Set}$ whose objects are pairs (σ, A)—where A is a set, and $\sigma \colon G \to \mathrm{Aut}_{\mathtt{Set}}(A)$ is a group action on A—and morphisms $\phi \colon (\sigma, A) \to (\rho, B)$ are set-functions $\phi \colon A \to B$ such that the following diagram commutes

$$\begin{tikzcd}
G \times A \arrow[r, "\mathrm{id}_G \times \phi"] \arrow[d, "\sigma"'] & G \times B \arrow[d, "\rho"] \\
A \arrow[r, "\phi"'] & B
\end{tikzcd}$$

That is, $\phi(\sigma(g)(a)) = \rho(g)(\phi(a))$—or put even more simply as $\phi(ga) = g\phi(a)$ when there is no chance of confusion. These functions are called *G-equivariant*.

Proposition 7.8.12. Let $(\sigma, A) \in G$-$\mathtt{Set}$, where A is non-empty and σ is a *transitive* left-action on A. Then there exists an *isomorphism*

$$(\sigma, A) \simeq (\ell, G/\mathrm{Stab}_\ell(a)),$$

in G-$\mathtt{Set}$, where ℓ is the left-multiplication of G on $G/\mathrm{Stab}_\ell(a)$, and $a \in A$ is *any* element.

Proof. For the sake of brevity, define $S_a \coloneqq \mathrm{Stab}_\ell(a)$. Let $a \in A$ be any element, and define a *set*-function $\phi \colon G/S_a \to A$ given by $\phi(gS_a) \coloneqq ga$. This function is well defined since, for any $gS_a = g'S_a$ we have $g'^{-1}g \in S_a$ thus $(g'^{-1}g)a = a$, which implies in $ga = g'a$. To show that ϕ is bijective, we construct its inverse: define $\psi \colon A \to G/S_a$ by mapping $ga \mapsto gS_a$, which is well defined because if $ga = g'a$ then $g'^{-1}ga = a$ thus $g'^{-1}g \in S_a$ so that $gS_a = g'S_a$ by definition. One sees easely that ϕ and ψ are inverses of each other. Finally, the map ϕ is also equivariant since

$$\phi(g(g'S_a)) = \phi((gg')S_a) = (gg')a = g(g'a) = g\phi(g'S_a).$$

♮

Corollary 7.8.13. Let G be a *finite group* and A be a set. For *any* action of G on a group A, the orbit $\mathrm{Orb}_G(a)$ of *any* element $a \in A$ is a *finite subset of A*. Moreover the we have that $|\mathrm{Orb}_G(a)|$ *divides* $|G|$.

Proof. Given a G-set (σ, A) for any action σ, if $a \in A$ is any element, then the restriction

$$\overline{\sigma}\colon G \longrightarrow \mathrm{Aut}_{\mathtt{Set}}(\mathrm{Orb}_G(a)),$$

is a transitive action by the definition of the orbit—where $\overline{\sigma}(g)(g'a) := \sigma(g)(ga)$ for all $g \in G$ and $g'a \in \mathrm{Orb}_G(a)$.

Therefore by Proposition 7.8.12 we have $(\overline{\sigma}, \mathrm{Orb}_G(a)) \simeq (\ell, G/\mathrm{Stab}_G(ga))$ in G-$\mathtt{Set}$, for any $ga \in \mathrm{Orb}_G(a)$—which means the existence of a G-equivariant bijection between the sets $\mathrm{Orb}_G(a)$ and $G/\mathrm{Stab}_G(ga)$. From Proposition 7.4.6 we have $|G/\mathrm{Stab}_G(ga)| = |G|/|\mathrm{Stab}_G(ga)|$ and therefore

$$|\mathrm{Orb}_G(a)| \cdot |\mathrm{Stab}_G(ga)| = |G|.$$

$\natural$

Theorem 7.8.14. Let G be a group acting on a set A. Consider elements $a \in A$, $g \in G$, and define $b := ga$. It follows that

$$\mathrm{Stab}_G(b) = g\,\mathrm{Stab}_G(a) g^{-1}.$$

Proof. Let $h \in \mathrm{Stab}_G(a)$ be any element, then

$$(ghg^{-1})b = (gh)(g^{-1}b) = (gh)a = g(ha) = ga = b,$$

therefore $ghg^{-1} \in \mathrm{Stab}_G(a)$. Now if $\ell \in \mathrm{Stab}_G(b)$, we have $\ell b = b$ but since $b = ga$ then $\ell(ga) = ga$ thus multiplying by g^{-1} in both sides we obtain $(g^{-1}\ell g)a = a$. Therefore, $\mathrm{Stab}_G(b) = g\,\mathrm{Stab}_G(a)g^{-1}$. $\natural$

7.9 Topological Groups

Construction and Properties

Study topological groups on Dieck

Actions on Topological Spaces

Definition 7.9.1 (Orbit space)**.** Let X be a topological space, and a topological group G together with a left action on X. We define the *orbit space* of the G-space X to be the set

$$X/G := X/x \sim \mathrm{Orb}_G(x)$$

together with the quotient topology given by the canonical projection $X \twoheadrightarrow X/G$.

Definition 7.9.2 (G-stable)**.** Given a G-space X, a subset $A \subseteq X$ is said to be G-stable if for every pair $(g, a) \in G \times A$ we have $ga \in A$.

7.10 Group Objects

Definition 7.10.1 (Group object)**.** Let C be a category with (finite) products and with a terminal object 1. A triple (G, m, e, i) is said to be a *group object* of C if:

- G is an object of C.
- $m\colon G \times G \to G$ is a morphism of C satisfying the commutativity of

$$\begin{tikzcd}
(G \times G) \times G \arrow[r, "m \times \mathrm{id}_G"] \arrow[d, "\simeq"'] & G \times G \arrow[r, "m"] & G \arrow[d, equal] \\
G \times (G \times G) \arrow[r, "\mathrm{id}_G \times m"] & G \times G \arrow[r, "m"] & G
\end{tikzcd}$$

  in C. This arrow defines the notion of multiplication of group members.
- $e\colon 1 \to G$ is a morphism of C making the diagram

$$\begin{tikzcd}
1 \times G \arrow[r, "e \times \mathrm{id}_G"] \arrow[dr, "\simeq"'] & G \times G \arrow[d, "m"] & G \times 1 \arrow[l, "\mathrm{id}_G \times e"'] \arrow[dl, "\simeq"] \\
& G &
\end{tikzcd}$$

  commute in C. The map e defines the notion of a neutral member of the group object.
- $i\colon G \to G$ is a morphism of C such that—if $\Delta := \mathrm{id}_G \times \mathrm{id}_G$ is the diagonal morphism of G—the diagram

$$\begin{tikzcd}
G \arrow[r, "\Delta"] \arrow[d, dashed] & G \times G \arrow[r, "\mathrm{id}_G \times i"] & G \times G \arrow[d, "m"] & G \times G \arrow[l, "i \times \mathrm{id}_G"'] & G \arrow[l, "\Delta"'] \arrow[d, dashed] \\
1 \arrow[rr, "e"'] & & G & & 1 \arrow[ll, "e"]
\end{tikzcd}$$

  is commutative in C. The map i defines the notion of inversion of members of the group.

Chapter 8

Rings & Modules

8.1 Introduction

Definition 8.1.1 (Ring). A ring R is a set endowed with an additive structure $+\colon R\times R \to R$ such that $(R,+)$ is an *abelian group*, moreover, R is also endowed with a second multiplicative structure $\cdot\colon R \times R \to R$, making $(R,\cdot)$ into a monoid — that is, it is both associative and has a two-sided identity:

- Given $r, s, t \in R$, we have $r(st) = (rs)t$.
- There exists a unitary element $1_R \in R$ for which all $r \in R$ is such that

$$r1_R = 1_R r = r.$$

Furthermore, we impose that the product structure is distributive over the additive structure, that is, given any three $r, t, s \in R$, one has

$$(r+t)s = rs + ts \text{ and } r(t+s) = rt + rs.$$

Remark 8.1.2. Notice that we impose the following in the definition of a ring — which may vary from author to author — the ring is always associative and has a unity.

Definition 8.1.3. A ring R is said to be *commutative* if for all $r, s \in R$ we have $rs = sr$.

Corollary 8.1.4. Given a ring R, its identity is unique.

Proof. Let $1, 1' \in R$ be both identities of R, then $11' = 1$ but $11' = 1'$, thus $1 = 1'$. ♮

Corollary 8.1.5. Given a ring R, we have for all $r \in R$ that $0_R r = r0_R = 0_R$.

Proof. Note that for both cases, since $0 = 0 + 0$, we are able to write the following

$$\begin{aligned} r\cdot 0 &= r\cdot(0+0) = r\cdot 0 + r\cdot 0,\\ 0\cdot r &= (0+0)\cdot r = 0\cdot r + 0\cdot r. \end{aligned}$$

Further, since $(R,+)$ is a group, we evoke the cancellation to obtain $r \cdot 0 = 0$ and $0 \cdot r = 0$. ♮

Corollary 8.1.6. Let R, then $(-1) \cdot r = r \cdot (-1) = -r$.

Proof. Notice that since $1 \cdot r = r \cdot 1 = r$ we may write

$$
\begin{aligned}
r + (-1) \cdot r = (1-1) \cdot r = 0 \cdot r = 0, \\
r + r \cdot (-1) = r \cdot (1-1) = r \cdot 0 = 0.
\end{aligned}
$$

Thus we arrive at the desired conclusion. ♮

Examples of Rings

Example 8.1.7 (Trivial ring)**.** The trivial group $*$ can be endowed with a ring structure by imposing that $* \cdot * = *$ and $* + * = *$, so that $(*, +, \cdot)$ is a ring — we call such ring *trivial* or *zero-ring*. Moreover, in such ring we have the equality $0 = 1$.

Corollary 8.1.8. A ring R is a zero-ring if and only if $0 = 1$.

Proof. If R is a zero-ring, it's obvious that $0 = 1$. On the other hand, assume that R is a ring such that $0 = 1$, then if $r \in R$ is any element, we have $r = 1 \cdot r = 0 \cdot r = 0$, thus R is indeed a zero ring. ♮

Example 8.1.9 (Power set ring)**.** Let S be any set and 2^S be its power set. If we define additive and multiplicative structures in 2^S given by $A + B := (A \cup B) \smallsetminus (A \cap B)$ and $A \cdot B := A \cap B$, for all $A, B \in 2^S$, then $(2^S, +, \cdot)$ is a ring.

In order to prove that, we first show that $(2^S, +)$ forms an abelian group. Since $A \cup B = B \cup A$ and $A \cap B = B \cap A$, then clearly $A + B = B + A$. Moreover, the empty set $\varnothing$ is the additive identity, since $A \cup \varnothing = A$ and $A \cap \varnothing = \varnothing$ and $A \smallsetminus \varnothing = A$ — thus $A + \varnothing = A$. Every set is its own inverse: notice that $A \cup A = A$ and $A \cup A = A$, hence $A + A = A \smallsetminus A = \varnothing$. We therefore conclude that $(2^S, +)$ is an abelian group.

Now we show that $(2^S, \cdot)$ is a monoid with distributivity over addition. Let $A, B, C \in 2^S$ be any three sets. For associativity we have

$$
(A \cdot B) \cdot C = (A \cap B) \cap C = A \cap (B \cap C) = A \cdot (B \cdot C).
$$

The unitary element is S itself, since $A \cap S = S \cap A = A$ for all $A \in 2^S$. Moreover,

$$
\begin{aligned}
(A + B) \cdot C &= [(A \cup B) \smallsetminus (A \cap B)] \cap C \\
&= [(A \cap C) \cup (B \cap C)] \smallsetminus [(A \cap C) \cap (B \cap C)] \\
&= A \cdot C + B \cdot C.
\end{aligned}
$$

We can analogously show the same for $A \cdot (B + C) = A \cdot B + A \cdot C$.

Example 8.1.10. Let R be a ring and S be a set. The set of set-functions R^S, of the form $S \to R$, together with the pointwise addition and multiplication make R^S into a ring.

Example 8.1.11. Let $M_n(R)$ denote the collection of all $n \times n$ matrices with entries in the ring R. If we endow $M_n(R)$ with component-wise addition and matrix multiplication, then $(M_n(R), +, \cdot)$ is a ring.

Zero-Divisors

Definition 8.1.12 (Zero-divisor)**.** Let R be a ring. We say that $a \in R$ is a *left-zero-divisor* if there exists $b \neq 0$ in R for which $ab = 0$. On the other hand, a is said to be a *right-zero-divisor* if $ba = 0$.

Remark 8.1.13. The zero-ring is the *only* ring with *no* zero-divisors.

If R is a non-zero ring, then $0 \in R$ and since 0 is clearly a zero-divisor, R has at least one zero-divisor. On the other hand, if $*$ is a zero-ring, then zero is the only element of $*$ — thus we have no element different than zero for there to be a zero-divisor.

Proposition 8.1.14. Let R be a ring. An element $a \in R$ is *not* a left-zero-divisor (or right-zero-divisor) if and only if the map $R \to R$ given by the *left-multiplication* by a (or right-multiplication) is *injective*.

Proof. We only concern ourselves with the left case, the right case is clearly analogous. Let $a \in R$ be a non-left-zero-divisor, then if $ab = ac$ for some $b, c \in R$, then $ab - ac = a(b - c) = 0$ — and, since a is not a zero-divisor, then necessarily $b - c = 0$, that is $b = c$.

On the other hand, if a is a left-zero-divisor then let $b \neq 0$ be any element of R such that $ab = 0$. Notice then that both b and zero have the same image under the multiplication map, thus $R \to R$ is definitely not injective. ♮

Remark 8.1.15. Notice that Proposition 8.1.14 simply states that for an element to be a zero-divisor the cancellation law must be true, that is, $a \in R$ is not a zero-divisor if $ab = ac$ implies $b = c$ for every $b, c \in R$.

Definition 8.1.16 (Domain)**.** An object R is said to be a *domain* if R is non-zero ring and for every two elements $a, b \in R$ such that $ab = 0$ we must have either $a = 0$ or $b = 0$.

Definition 8.1.17 (Integral domain)**.** An object R is said to be an *integral domain* if R is a non-zero commutative ring such that, for all $a, b \in R$ such that $ab = 0$, then either $a = 0$ or $b = 0$.

That is, in an integral domain *every non-zero* element is a non-zero-divisor. Classic examples of integral domains are $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$.

Corollary 8.1.18. In an integral domain R, if $r, a, b \in R$ are such that $ra = rb$ and $r \neq 0$, then $a = b$ — that is, cancellation by non-zero elements holds.

Proof. This is immediate from Proposition 8.1.14. ♮

Corollary 8.1.19. If R is an integral domain, an element $x \in R$ is such that $x^2 = 1$, if and only if $x \in \{-1, 1\}$.

Proof. If $x \in \{-1, 1\}$, clearly $x^2 = 1$, since this is true for any given ring. For the other implication, suppose now that $x^2 = 1$, then $x^2 - 1 = (x + 1)(x - 1) = 0$ but since R is an integral domain, either $x = 1$ or $x = -1$. ♮

Nilpotent Elements

Definition 8.1.20 (Nilpotent)**.** Let R be a ring. An element $a \in R$ is said to be *nilpotent* if there exists some $n \in \mathbf{Z}$ such that $a^n = 0$.

Lemma 8.1.21. Let R be a ring. If $a, b \in R$ are nilpotent elements with $ab = ba$, then $a + b$ is nilpotent.

Proof. Suppose $n, m \in \mathbf{Z}$ are such that $a^n = b^m = 0$. Since a and b commute, we can write the binomial equation $(a+b)^{nm} = \sum_{j=0}^{nm} \binom{nm}{j} a^{nm-j} b^j$ — for $j < m$ we have $nm - j > nm - m = n(m-1)$, greater than a multiple of n, so that $a^{nm-j} = 0$; for $j \geqslant m$, we have $b^j = 0$. Thus we arrive at $(a+b)^{nm} = 0$. ♮

Lemma 8.1.22. A class $[m]_n$ is nilpotent in $\mathbf{Z}/n\mathbf{Z}$ if and only if m shares all prime factors of n.

Proof. Suppose that there exists a prime p in the prime factorization of n such that p does not divide m. If $\ell \in \mathbf{Z}$ is any exponent, m^ℓ won't be divisible by n since it lacks the p factor — hence m fails to be nilpotent.

If m shares every prime factor of n, let k be defined to be the maximum exponent in the prime factorization of n (which is ensured to exist since the collection is finite). Then clearly $[m]_n^k = [m^k]_n = 0$ since m^k is divisible by n. ♮

Lemma 8.1.23. Let R be a *commutative* ring and $x \in R$ be a nilpotent element. The following statements hold:

(a) The element x is either zero or zero-divisor.

(b) For every $r \in R$, the product rx is nilpotent.

(c) For every invertible element $u \in R$, the sum $x + u$ is invertible.

Proof. (a) If x is non-zero, both the left and right multiplications of x have the same image under the distinct elements x^{n-1} and 0 — hence x is a zero-divisor.

(b) From commutativity $(rx)^n = r^n x^n = r^n \cdot 0 = 0$.

(c) First we deal with the case where $u = 1$. Remember that given any $a \in R$ and $m \in \mathbf{N}$ we have

$$(1+a)(1 - a + a^2 - \cdots + (-1)^m a^m) = 1 + (-1)^m a^{m+1}.$$

Therefore, for every $m > n$, we have $(1+x)\sum_{j=0}^{m}(-1)^j a^j = 1$ — which shows that $1 + x$ is invertible. For the general case, if v is the inverse of u, one sees that for any $m > n$ we have

$$(u+x)\sum_{j=0}^{m}(-1)^j v x^j = 1 + (-1)^m x^{m+1} = 1.$$

♮

Units

Definition 8.1.24 (Unit)**.** Let R be a ring. An element $u \in R$ is said to be a *left-unit* (or *right-unit*) if there exists $a \in R$ for which $ua = 1$ (or $au = 1$). An element is said to be a *unit* if it is both a left and right unit.

Proposition 8.1.25. Let R be a ring and $u \in R$ be an element. The following propositions hold.

(a) The element u is a left-unit (or right-unit) if and only if the left-multiplication (or right-multiplication) map $R \to R$ is *surjective*.

(b) If u is a *left*-unit (or right-unit), then u is a non-*right*-zero-divisor (or non-*left*-zero-divisor).

(c) The *inverse* of a unit is *unique*.

(d) Units form a *group* under multiplication — such group is denoted by $R^{\times}$.

Proof. (a) Suppose first that u is a left-unit. Let $a \in R$ be any element and suppose that $v \in R$ is such that $uv = R$ (which is ensured to exist) — then $u(va) = (uv)a = a$, thus left-multiplication is surjective.

On the other hand, if the left-multiplication by u is surjective, we can conclude that there must exist $v \in R$ for which $uv = 1$ — hence u is a left-unit of R.

The proof for the right-unit case is completely analogous and won't be included.

(b) If u is a left-unit, assume $v \in R$ is such that $uv = 1$. Notice that the *right*-multiplication maps $f_u, f_v \colon R \rightrightarrows R$ by u and v, respectively, are such that — for all $a \in R$,
$$f_v f_u(a) = f_v(au) = (au)v = a(uv) = a \cdot 1 = a.$$
Therefore, f_v is the left-inverse of f_u and thus f_u is injective. The equivalent proof can be written for the right-unit.

(c) Let u be a unit and $v_1, v_2 \in R$ be such that $v_1 u = 1$ and $uv_2 = 1$ — notice that
$$v_1 = v_1 \cdot 1 = v_1(uv_2) = (v_1 u)v_2 = 1 \cdot v_2 = v_2.$$

(d) Let U denote the collection of units of R (such set is ensured to be non-empty since 1 is a unit) — zero-rings are not a problem since zero is a unit in such case. If $a, b \in U$ are units, then there are $x, y \in U$ such that $ax = xa = 1$ and $by = yb = 1$ (from the last item, inverses are unique) — therefore, $(ab)(yx) = a(by)x = ax = 1$, that is, $ab \in U$ and yx its inverse.

♮

Lemma 8.1.26. Let R be a ring and $a \in R$ be an element. If a is a right-unit (or left-unit) and has two or more left-inverses (or right-inverses), then

- The element a is *not* a left-zero-divisor (or right-zero-divisor).
- The element a *is* a right-zero-divisor (or left-zero-divisor).

Proof. We shall restrict ourselves with the right-unit case — the other is completely analogous. Since a is a right-unit, there must exist $u \in R$ such that $ua = 1$, then if $x, y \in R$ are such that $ax = ay$, we can multiply both left-sides by u and obtain $(ua)x = x = y = (ua)y$ — thus left-multiplication by a is injective and a is not a left-zero-divisor.

For the other item, notice that if u and v are distinct left-inverses of a, then $ua = va = 1$ have the same image under right-multiplication by a — hence a is a right-zero-divisor. ♮

Division Rings & Fields

Definition 8.1.27 (Division ring)**.** A ring R is said to be a *division ring* if *every* non-zero element is a *unit*.

Definition 8.1.28 (Field)**.** A non-zero ring R is said to be a field if R is a *commutative* division ring.

Remark 8.1.29. By the item (b) of Proposition 8.1.25, we see that a unit is a non-zero-divisor, thus, if every non-zero element is a unit, then every non-zero element is also a non-zero-divisor. This implies directly that *every field is an integral domain*. On the other hand, $\mathbf{Z}$ is an integral domain but fails to be a field, hence *not every integral domain is a field*. The concepts can however coincide in a special case, as we see in Proposition 8.1.31.

Proposition 8.1.30. Any subring of a field is an integral domain.

Proof. Let k be a field and $R \subseteq k$ be a non-zero subring. Certainly R is commutative. Moreover if $a, b \in R$ are such that $ab = 0$, suppose, for the sake of contradiction, that a and b are both non-zero. From the field properties, there are inverses $a^{-1}, b^{-1} \in k$ so that $a^{-1}(ab)b^{-1} = (a^{-1}a)(bb^{-1}) = 1 \cdot 1 = 1$, but $ab = 0$ — this implies that $1 = 0$, which is a contradiction, thus a or b are zero. ♮

Proposition 8.1.31. A finite and commutative ring R is a field if and only if R is an integral domain.

Proof. If R is a finite commutative integral domain, let $a \in R$ be a non-zero-divisor of R — so that multiplication by a is an injective map. Since the ring is finite, the multiplication by a is also surjective — thus by Proposition 8.1.25 we find that a is a unit. ♮

Proposition 8.1.32 (Particular case of Weddenburn's theorem)**.** Let R be a division ring with p^2 elements and p be a prime, then R is a field.

Proof. For the proof, we shall use the concepts of centre and centralizer, developed in Definition 8.2.15 and Definition 8.2.18. If we assume that R is non-commutative, the centre $Z(R)$ will be a proper subring of R. Moreover, since $Z(R)$ is also a subgroup of R under multiplication, by Corollary 7.4.7 we conclude that $|Z(R)| = p$. Let now $r \in R \setminus Z(R)$, clearly the centralizer must contain both the centre of the ring — since by definition it must commute with r — and by symmetry, $r \in Z(r)$, thus $\{r\} \cup Z(R) \subseteq Z(r)$. Since the centralizer is a subgroup of R and $|Z(r)| > p$, one concludes by Corollary 7.4.7 that $|Z(r)| = p^2$ — therefore $Z(r) = R$. That, however gives us a contradiction because if the centralizer of r is the whole ring, then r commutes with every element and thus $r \in Z(R)$ — this cannot be the case by hypothesis, thus one concludes that R must be commutative and therefore a field. ♮

Monoid Ring

Definition 8.1.33 (Monoid ring)**.** Given a monoid $(M, \cdot)$ and a ring R, we define a new ring called *monoid ring* $R[M] := R^{\oplus M}$ — that is, the ring whose elements are formal linear combinations $\sum_{m \in M} a_m \cdot m$, where the coefficients $a_m \in R$ are non-zero for at most finitely many $m \in M$.

The unitary element of $R[M]$ is defined to be $1_R \cdot 1_M$. The additive and multiplicative structures are defined as follows

$$\sum_{m \in M} a_m \cdot m + \sum_{m \in M} b_m \cdot m := \sum_{m \in M} (a_m + b_m) \cdot m,$$
$$\sum_{m \in M} a_m \cdot m + \sum_{m \in M} b_m \cdot m := \sum_{m \in M} \sum_{m_1 \cdot m_2 = m} (a_{m_1} b_{m_2}) \cdot m.$$

Polynomial Ring

Definition 8.1.34 (Polynomial)**.** Let R be a ring. A *polynomial* is a map $f \colon R \to R$ given by a *finite* linear combination of non-negative powers of the indeterminate variable with coefficients in R

$$f(x) = \sum_{j \geqslant 0} a_j x^j, \text{ where } a_j = 0 \text{ for } j \gg 0.$$

The collection of all polynomials over R with variable x is denoted by $R[x]$ — this collection forms a ring under point-wise addition and product, that is, given another element $g(x) := \sum_{j \geqslant 0} b_j x^j$, we define

$$f(x) + g(x) := \sum_{j \geqslant 0} (a_j + b_j) x^j,$$
$$f(x) \cdot g(x) := \sum_{k \geqslant 0} \sum_{i+j=k} a_i b_j x^{i+j}.$$

It should be noted immediately that $R[x] \simeq R[\mathbf{N}]$, where $\mathbf{N}$ is viewed as a monoid under addition.

Two polynomials are said to be equal if the sequence of coefficients match — that is, if $f(x) := \sum_{j \geqslant 0} a_j x^j$ and $g(x) := \sum_{j \geqslant 0} b_j x^j$, then $f = g$ if and only if $a_j = b_j$ for all $j \geqslant 0$ — since the collection of non-zero coefficients is necessarily finite, this equality relation is well defined.

Corollary 8.1.35. If R is a commutative ring, then $R[x]$ is commutative.

Lemma 8.1.36. A ring R is an integral domain if and only if $R[x]$ is an integral domain.

Proof. If R is an integral domain, then clearly $R[x]$ is both non-zero and commutative. Moreover, suppose that $f(x) := \sum_{j\geqslant 0} a_j x^j$ and $g(x) := \sum_{j\geqslant 0} b_j x^j$ are elements of $R[x]$ such that $f \cdot g = 0$ — notice that from definition we have

$$f(x)g(x) = \sum_{k\geqslant 0} \sum_{i+j=k} a_i b_j x^{i+j},$$

which is merely a finite combination with coefficients, so that if $f \cdot g = 0$, we necessarily have $a_j = 0$ for all $j \geqslant 0$ or $b_j = 0$ for all $j \geqslant 0$ since R is an integral domain — thus either $f = 0$ or $g = 0$.

On the other hand, if $R[x]$ is an integral domain, let $a, b \in R$ be elements such that $ab = 0$ in R, then we can consider the constant polynomials $f(x) := a$ and $g(x) := b$ to conclude that, since $f \cdot g = 0$ then $a = 0$ or $b = 0$. ♮

Definition 8.1.37. The *degree* of a *non-zero* polynomial $f(x) := \sum_{j\geqslant 0} a_j x^j$ is defined as the maximum index for which the corresponding coefficient is non-zero, that is

$$\deg f := \max\{j : a_j \neq 0\}.$$

This is well defined because $a_j \neq 0$ for only finitely many indices. By convention, the zero-polynomial has degree $-\infty$.

Polynomials of multiple variables are obtained simply as an iteration of the process of construction of $R[x]$, that is,

$$R[x_1, \dots, x_n] := R[x_1] \dots [x_n].$$

The ordering of the list $x_1, \dots, x_n$ is irrelevant in the construction of the ring, if we swap any elements of the list we still end up with isomorphic rings.

One can consider a ring of polynomials over infinitely many variables, say $R[x_1, x_2, \dots]$, but still every polynomial of such ring consists of only finitely many terms — as we impose that the coefficient indexing set is finite.

Definition 8.1.38 (Power series ring)**.** Let R be a ring. The ring of *power series* with variable x and coefficients in R, denoted by $R[\![x]\!]$, is defined to be the ring whose elements are formal sums $\sum_{j=0}^{\infty} a_j x^j$.

8.2 The Category of Rings

Definition 8.2.1. Let R and S be rings. We define a morphism of rings $\phi\colon R \to S$ to be a map such that, for all $a, b \in R$, we have $\phi(a + b) = \phi(a) + \phi(b)$ and $\phi(ab) = \phi(a)\phi(b)$. Moreover, we also impose that morphisms are unitary, that is, $\phi(1_R) = 1_S$.

Definition 8.2.2. The category of rings is defined to consist of rings and morphisms between them — such category will be denoted by `Ring`.

Remark 8.2.3. Curiously, the zero-ring is *not* a zero-object in the category of rings, it *is* a *final* object but *fails* to be *initial* — this is due to the fact that we imposed morphisms to be unitary, hence for a non-zero ring R, there exists no ring morphism from the zero-ring to R.

One should, however, not be afraid, because `Ring` does have initial objects. In fact, $\mathbf{Z}$ is initial in `Ring` — for any ring R we can define a unique ring morphism $\mathbf{Z} \to R$ mapping $n \mapsto n \cdot 1_R$ for any $n \in \mathbf{Z}$.

Corollary 8.2.4 (Unit preservation)**.** If $\phi\colon R \to S$ is a ring morphism and $u \in R$ is a left-unit (or right-unit), then $\phi(u) \in S$ is again a left-unit (or right-unit).

Proof. Let $v \in R$ be such that $vu = 1_R$, then $\phi(vu) = \phi(v)\phi(u)$ but $\phi(vu) = \phi(1_R) = 1_S$ thus $\phi(v)\phi(u) = 1_S$ and $\phi(u)$ is indeed a left-unit in S. ♮

Proposition 8.2.5 (Image of morphism is a subring)**.** The image of a ring morphism $R \to S$ is a subring of S.

Proof. Let $\phi\colon R \to S$ be any ring morphism. Since it is a ring morphism, $\phi(1_R) = 1_S \in \operatorname{im}\phi$, moreover, if $a, b \in R$ then $\phi(a+b) = \phi(a) + \phi(b) \in \operatorname{im}\phi$ and $\phi(ab) = \phi(a)\phi(b) \in \operatorname{im}\phi$ — thus $\operatorname{im}\phi$ is a subring of S. ♮

Remark 8.2.6. The image of non-zero-divisors may be a zero-divisor in the new ring — for instance, $3 \in Z$ is a non-zero-divisor but the ring morphism $Z \twoheadrightarrow \mathbf{Z}/6\mathbf{Z}$ (canonical projection) maps $3 \mapsto [3]_6$, which is a zero divisor since $[2]_6[3]_6 = [0]_6$.

Proposition 8.2.7. Let $\phi\colon R \to S$ be a set-function between rings R and S such that the additive and multiplicative structures are preserved. If either one of the following propositions hold, then ϕ is a morphism of rings:

(a) The set-function ϕ is surjective.

(b) The set-function $\phi \neq 0$ and S is an integral domain.

Proof. We simply need to show that each of the propositions yield $\phi(1_R) = 1_S$ — which is the only condition left for ϕ to be a morphism of rings.

(a) If ϕ is surjective, then there exists $r \in R$ such that $\phi(a) = 1_S$ but then

$$1_S = \phi(a) = \phi(a \cdot 1_R) = \phi(a)\phi(1_R) = 1_S \cdot \phi(1_R) = \phi(1_R).$$

(b) If $\phi(1_R)$ where to be zero, the whole image of ϕ would also evaluate to zero, thus 1_R has a non-zero image under ϕ. Let $r \in R$ be any element such that $\phi(r) \neq 0$ — then $\phi(r) = \phi(r \cdot 1_R) = \phi(r)\phi(1_R)$ thus by means of Corollary 8.1.18 we can conclude that $\phi(1_R) = 1_S$.

♮

Universal Property of Polynomial Rings

Let $A := \{a_1, \dots, a_n\}$ be a set and define a category $\mathtt{R}_A$ consisting of pairs (α, R) — where $\alpha\colon A \to \mathbf{R}$ is a set-function and R is a ring such that $\alpha(a_i)$ commutes with every element of R for each $1 \leqslant i \leqslant n$ (one may restrict this further by imposing that R is commutative). Intuitively, morphisms between two objects $(\alpha, R) \to (\beta, S)$ are ring morphisms $\phi\colon R \to S$ such that the following diagram commutes in the category of sets:

$$\begin{tikzcd} R \arrow[rr, "\phi"] & & S \\ & A \arrow[ul, "\alpha"] \arrow[ur, "\beta"'] & \end{tikzcd}$$

In the same context of what is explained above, we now state a universal property concerning the ring $\mathbf{Z}[x_1, \dots, x_n]$ in the category $\mathtt{R}_A$.

Proposition 8.2.8. Let R be a ring and define a set-function $\alpha\colon A \to R$ mapping $a_j \mapsto x_j$. The object $(\alpha, \mathbf{Z}[x_1, \dots, x_n])$ is initial in $\mathtt{R}_A$.

Proof. Let (β, R) be any object in $\mathtt{R}_A$. We now construct a morphism

$$\phi\colon (\alpha, \mathbf{Z}[x_1, \dots, x_n]) \to (\beta, R).$$

We can impose that $\phi(x_j) := \beta(a_j)$ for every $1 \leqslant j \leqslant n$. For ϕ to be a ring morphism, one must also impose that is preserves both the additive and multiplicative structures of the rings. Since $\mathbf{Z}$ is initial in $\mathtt{Ring}$, we can uniquely determine that $\phi(r) = \psi(r)$ — where $\psi\colon \mathbf{Z} \to \mathbf{R}$ is a unique morphism. That is, we defined morphism ϕ given by

$$\begin{aligned} \phi\left(\sum m_{i_1 \dots i_n} x_1^{i_1} \dots x_n^{i_n} \right) &= \sum \phi(m_{i_1 \dots i_n}) \phi(x_1^{i_1}) \dots \phi(x_n^{i_n}) \\ &= \sum \psi(m_{i_1 \dots i_n}) \beta(a_1)^{i_1} \dots \beta(a_n)^{i_n}, \end{aligned}$$

which is surely both a ring morphism and unique — thus $(\alpha, \mathbf{Z}[x_1, \dots, x_n])$ is initial, and

$$\begin{tikzcd} \mathbf{Z}[x_1, \dots, x_n] \arrow[rr, dashed, "\phi"] & & R \\ & A \arrow[ul, "\alpha"] \arrow[ur, "\beta"'] & \end{tikzcd}$$

$\natural$

Proposition 8.2.9 (Universal property of polynomial rings)**.** Let $f\colon R \to S$ be any ring morphism and $s_0 \in S$ be a fixed element commuting with $f(r)$ for all $r \in R$. There exists a *unique* ring morphism $\phi\colon R[x] \to S$ which *extends* f and maps $x \mapsto s_0$.

Proof. In the construction of ϕ we first impose that $\phi(x) := s_0$. Moreover, for ϕ to extend f one has to define $\phi(r) := f(r)$ for all given $r \in R$. In order for ϕ to be a ring morphism it also needs to preserve the additive and multiplicative structures of the rings — hence we end up with

$$\phi\left(\sum_{j\geqslant 0} r_j x^j\right) = \sum_{j\geqslant 0} \phi(r_j)\phi(x)^j = \sum_{j\geqslant 0} f(r_j)s_0^j.$$

Since ϕ is completely defined by the image under constant polynomials and x, we find that ϕ is the unique ring morphism sending x to s_0 for which the following diagram commutes

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, hook] & S \\ R[x] \arrow[ur, dashed, "\phi"'] & \end{tikzcd}$$

♮

In fact, in the case where R is *commutative* and $f: R \to R$ is an *endomorphism*, the unique ring morphism $\phi: R[x] \to R$ defined above is called the *evaluation map* at the given fixed point. We normally denote such map by eval_r — where $r \in R$ is the chosen fixed point for evaluation.

Monomorphisms

Definition 8.2.10 (Kernel)**.** The kernel of a ring morphism $\phi: R \to S$ is the set

$$\ker \phi := \{r \in R : \phi(r) = 0_S\}.$$

Proposition 8.2.11 (Ring monomorphisms)**.** Let $\phi: R \to S$ be a ring morphism. The following properties are equivalent:

(a) The ring morphism ϕ is a monomorphism.

(b) The kernel of ϕ is trivial — that is, $\ker\phi = \{0_R\}$.

(c) The set-function ϕ injective.

Proof. (a) ⇒ (b): Suppose ϕ is monic. Let $r \in \ker\phi$ be any element and consider the uniquely defined morphisms $\mathrm{eval}_r, \mathrm{eval}_{0_R}: \mathbf{Z}[x] \rightrightarrows R$ provided by Proposition 8.2.8 by fixing $\mathrm{eval}_r(x) = r$ and $\mathrm{eval}_{0_R}(x) = 0_S$. Notice that $\phi\,\mathrm{eval}_r(n) = \phi\,\mathrm{eval}_{0_R}(n) = \phi(n)$ for any $n \in \mathbf{Z}$, while $\phi\,\mathrm{eval}_r(x) = \phi(r) = 0_S$ and $\phi\,\mathrm{eval}_{0_R}(x) = \phi(0_R) = 0_S$ — thus $\mathrm{eval}_r = \mathrm{eval}_{0_S}$, which in turn implies that $r = \mathrm{eval}_r(x) = \mathrm{eval}_{0_R}(x) = 0_S$. Therefore indeed $\ker\phi = \{0_R\}$.

(b) ⇒ (c): Suppose ϕ has trivial kernel and let $r, t \in R$ be such that $\phi(r) = \phi(t)$, then $\phi(r) - \phi(t) = \phi(r - t) = 0$ and $r - t \in \ker\phi$, which implies in $r = t$ — that is, ϕ is injective.

(c) ⇒ (a): Suppose ϕ is an injective set-function. Since injections are monomorphisms in `Set`, we conclude that the set-function $\phi: R \to S$ is a monomorphism in `Set` when R and S are viewed as sets. If we now endows R and S with their respective ring structures, we obtain that ϕ is a monic in `Ring`. ♮

Subrings

Definition 8.2.12 (Subring)**.** Let R be a ring. A subring S of R is a ring whose elements are contained in R and the canonical inclusion map $S \hookrightarrow R$ is a ring morphism.

Remark 8.2.13. It should be noted that S is a ring with unity where $1_S = 1_R$.

Proposition 8.2.14 (Intersection of subrings)**.** Let R be a ring and $\{S_j\}_{j\in J}$ be a collection of subrings of R. The intersection $\bigcap_{j\in J} S_j$ is also a subring of R.

Proof. Since each S_j is a subring, if $a, b \in \bigcap_{j\in J} S_j$ then $a, b \in S_j$ for all $j \in J$, and $a + b, ab \in S_j$. Since every subring contains the unity, so does the intersection — hence $\bigcap_{j\in J} S_j$ is a subring of R. ♮

Center

Definition 8.2.15 (Center)**.** Given a ring R, we define its centre to be

$$Z(R) := \{r \in R : rx = xr \text{ for all } x \in R\}.$$

Corollary 8.2.16. The centre of a ring is a subring.

Proof. Let R be a ring. Clearly, $1 \in Z(R)$. Moreover, if $r, s \in Z(R)$ then for any $x \in R$ we have

$$\begin{aligned}(r + s)x &= rx + sx = xr + xs = x(r + s),\\ (rs)x &= r(sx) = (sx)r = (xs)r = x(sr) = x(rs).\end{aligned}$$

Therefore $Z(R)$ is indeed a subring of R. ♮

Corollary 8.2.17. If R is a division ring, then its centre $Z(R)$ is a field.

Proof. Since $Z(R)$ inherits the division ring structure and every element of $Z(R)$ commutes, it is indeed a field. ♮

Definition 8.2.18 (Centralizer)**.** Given a ring R and an element $r \in R$, the centralizer of r is defined to be the collection of elements $x \in R$ such that $xr = rx$ — we shall denote the centralizer of r as $Z(r)$.

Corollary 8.2.19. The centralizer of an element is a subring.

Proof. This is simply a straightforward particular case of Corollary 8.2.16. ♮

Corollary 8.2.20. In a division ring a centralizer is also a division ring.

Proof. Let R be a division ring and $r \in R$ any element. If $x \in Z(r)$ is any element, then since $rx = xr$, we have $(r^{-1}x^{-1})rx = (r^{-1}x^{-1})xr = 1$ thus $r^{-1}x^{-1}r$ is a left-inverse of r — moreover, we equivalently see that $rx^{-1}r^{-1}$ is a right-inverse of r. Since $Z(r)$ is a subring, such inverses are contained in $Z(r)$ and thus $Z(r)$ is a division ring.

Notice however that the centralizer may not be a field since there can be non-commuting elements in $Z(r)$. ♮

Corollary 8.2.21. The centre of a ring is the intersection of all centralizers of the ring.

Proof. Let R be a ring. Certainly, if $x \in Z(R)$ then it commutes with every element of R — which is equivalent to $x \in \bigcap_{r \in R} Z(r)$. Moreover, if an element belongs to the intersection of all centralizers, every element of the ring commutes with it and thus such element is also present in the centre of the ring. Therefore $Z(R) = \bigcap_{r \in R} Z(r)$. ♮

Epimorphisms

Remark 8.2.22 (Epimorphisms and surjection in `Ring`)**.** In `Ring`, epimorphisms are *not necessarily* surjective set-functions.

A classical counterexample is the ring morphism given by the canonical inclusion $\iota\colon \mathbf{Z} \hookrightarrow \mathbf{Q}$. Let R be any ring and consider ring morphisms $f, g\colon \mathbf{Q} \rightrightarrows R$ such that the following diagram commutes

$$\mathbf{Z} \xhookrightarrow{\iota} \mathbf{Q} \underset{g}{\overset{f}{\rightrightarrows}} R$$

Since $f|_{\mathbf{Z}} = g|_{\mathbf{Z}}$ we have, for any $p/q \in \mathbf{Q}$, that $f(p/q) = f(p)f(q)^{-1}$ and $g(p/q) = g(p)g(q)^{-1}$, which implies in $f(p/q) = g(p/q)$ — that is, $f = g$ in general, which implies that ι is an epimorphism in `Ring`. For the shock of the reader, the same is obviously not the case in `Set`. If we take the rings $\mathbf{Z}$ and $\mathbf{Q}$ as abelian groups, one sees that coker ϕ is non-trivial and by Proposition 7.6.8 we arrive at the fact that ϕ is not an epimorphism in `Ab` neither.

Remark 8.2.23. With the caution given by Remark 8.2.22 one can rightly observe that in the category of rings a morphism may be both monic and epic but yet lack the conditions for being an isomorphism.

Remark 8.2.24. Cokernels in `Ring` are not what one would normally expect of a good category, notice that given a ring morphism $\phi\colon R \to S$ and, if $\alpha\colon S \to Q$ is any ring morphism such that $\alpha\phi = 0$ — as is required by the universal property of cokernels — then $\alpha\phi(1) = \alpha(1) = 0$, which can only be the case for $Q = 0$, the zero-ring. We therefore conclude that coker ϕ must be the zero-ring.

Products

Proposition 8.2.25 (Product)**.** Products exist in the category of rings.

Proof. Let R and S be rings. We'll define on $R \times S$ additive and multiplicative structures naturally as follows:

$$(x, a) +_{R\times S} (y, b) := (x +_R y, a +_S b),$$
$$(x, a) \cdot_{R\times S} (y, b) := (x \cdot_R y, a \cdot_S b),$$

for any $x, y \in R$ and $a, b \in S$. We now check that $R \times S$ is a product in `Ring`.

Let Z be any ring and let $\phi: Z \to R$ and $\psi: Z \to S$ be any two ring morphisms. Considering the canonical projections $\pi_R: R \times S \twoheadrightarrow R$ and $\pi_S: R \times S \twoheadrightarrow S$, we can construct a map $\ell: Z \to R \times S$ sending $z \mapsto (\phi(z), \psi(z))$. Such a map inherits the preservation of both the multiplicative and additive structures of Z and $R \times S$ since ϕ and ψ do so — therefore ℓ is a morphism of rings. Moreover, ℓ is completely determined by the image of both ϕ and ψ, hence ℓ is the unique morphism of rings such that the following diagram commutes

$$\begin{tikzcd}
& Z \arrow[dl, bend right, "\phi"'] \arrow[d, dashed, "\ell"] \arrow[dr, bend left, "\psi"] & \\
& R \times S \arrow[dl, "\pi_R"] \arrow[dr, "\pi_S"'] & \\
R & & S
\end{tikzcd}$$

♮

Remark 8.2.26. Coproducts on the other hand, although present in `Ring`, are not as easy to construct. Lets work out a special case: consider the commutative ring $\mathbf{Z}[x, y]$ *in* the *category of commutative rings*. We'll show that $\mathbf{Z}[x, y]$ is the coproduct of two copies of $\mathbf{Z}[x]$. Let R be any commutative ring together with morphisms of rings $f, g: \mathbf{Z}[x] \rightrightarrows R$. We define a map $\phi: \mathbf{Z}[x, y] \to R$ for which $x \mapsto f(x)$, while $y \mapsto g(x)$ and finally $n \mapsto n$ for every $n \in \mathbf{Z}$. One immediately sees that ϕ indeed satisfies every condition of being a ring morphism. Moreover, since the image of x, y and $\mathbf{Z}$ completely defines ϕ, we conclude that ϕ is the unique morphism of rings such that the following diagram commutes

$$\begin{tikzcd}
\mathbf{Z}[x] \arrow[dr, hook] \arrow[ddr, bend right, "f"'] & & \mathbf{Z}[x] \arrow[dl, hook'] \arrow[ddl, bend left, "g"] \\
& \mathbf{Z}[x, y] \arrow[d, dashed, "\phi"] & \\
& R &
\end{tikzcd}$$

thus $\mathbf{Z}[x, y]$ is a coproduct of two copies of $\mathbf{Z}[x]$ in the category of commutative rings.

The Ring $\mathrm{End}_{\mathtt{Ab}}(G)$

The ring of endomorphisms of an abelian group $\mathrm{End}_{\mathtt{Ab}}(G)$ is a structure that pops up in plenty of situations — module theory will explore this ring structure vastly — so in this subsection we take some time to consider a small collection of interesting and useful facts about it. Such ring, has its additive structure defined by point-wise addition and the multiplicative structure is given by composition of morphisms.

Proposition 8.2.27. In the category of rings, there exists a natural isomorphism

$$\mathbf{Z} \simeq \mathrm{End}_{\mathtt{Ab}}(\mathbf{Z}).$$

Proof. Let $\phi\colon \mathrm{End}_{\mathtt{Ab}}(\mathbf{Z}) \to \mathbf{Z}$ sending $f \mapsto f(1)$ it is clear that

$$\phi(f+g) = (f+g)(1) = f(1) + g(1) = \phi(f) + \phi(g),$$

moreover the multiplicative structure is also preserved:

$$\phi(fg) = (fg)(1) = f(g(1)) = f(1 \cdot g(1)) = f(1)g(1) = \phi(f)\phi(g).$$

Also, $\phi(\mathrm{id}) = \mathrm{id}(1) = 1$, therefore ϕ is a ring morphism. Since the image of the identity element over a group morphism completely determines the map, one can be certain that ϕ is a bijection — thus an isomorphism of rings. ♮

Lets agree for the time being that, given a ring R and an element $r \in R$, the morphisms of rings ${}_r m, m_r\colon R \rightrightarrows R$ are the left and right, respectively, multiplication of elements of R by r.

Proposition 8.2.28. Let R be a ring. The map $m\colon R \rightarrowtail \mathrm{End}_{\mathtt{Ab}}(R)$ sending $r \mapsto {}_r m$ is an injective ring morphism.

Proof. We first show that m is a ring morphism: let $a, b \in R$ be any three elements, then for any $r \in R$ we have

$$m(a+b)(r) = {}_{a+b}m(r) = (a+b)r = ar + br = m(a)(r) + m(b)(r).$$

On the other hand, multiplication yields composition, as expected

$$m(ab)(r) = {}_{ab}m(r) = (ab)r = a(br) = {}_a m({}_b m(r)) = m(a)(m(b)(r)).$$

Also, $m(1)(r) = 1 \cdot r = r = \mathrm{id}_R(r)$ and hence $m(1) = \mathrm{id}_R$. We conclude that m is indeed a ring morphism. The injectivity comes from the fact that the only element that yields a zero-map in $\mathrm{End}_{\mathtt{Ring}}(R)$ is zero — thus the kernel is trivial. ♮

Remark 8.2.29. Notice that one cannot further extend Proposition 8.2.28 for right-multiplications, this comes from the fact that

$$m_{ab}(r) = r(ab) = (ra)b = m_b(m_a(r)).$$

That is, the multiplicative structure has its order reversed when transitioning to the compositional structure.

Proposition 8.2.30. Up to isomorphism, there exists a unique ring (with identity) structure whose underlying group is $(\mathbf{Z}, +)$.

Proof. Let R be any ring with underlying group $\mathbf{Z}$, and fix any $r \in R$. Consider the ring morphism $m\colon R \to \mathrm{End}_{\mathtt{Ab}}(R)$ as defined in Proposition 8.2.28. Let $f \in \mathrm{End}_{\mathtt{Ab}}(R)$ is completely defined by $f(1_R)$, thus ${}_{f(1_r)}m = f$. Therefore m is surjective and hence a bijection — m thus establishes an isomorphism $R \simeq \mathrm{End}_{\mathtt{Ab}}(R)$. Moreover, by Proposition 8.2.27 we have $\mathrm{End}_{\mathtt{Ab}}(R) \simeq \mathbf{Z}$, thus

$$R \simeq \mathrm{End}_{\mathtt{Ab}}(R) \simeq \mathbf{Z},$$

which proves the statement. ♮

Proposition 8.2.31. Let R be a ring and. There is a subring $S \subseteq Z(R)$ of the centre of R such that there exists a *ring* isomorphism

$$S \simeq Z(\mathrm{End}_{\mathsf{Ab}}(R)).$$

Proof. Let $f \in Z(\mathrm{End}_{\mathsf{Ab}}(R))$ be any group morphism. Since f commutes with every group endomorphism on R, in particular f commutes with every right-multiplication by an element of R — that is, given $r \in R$, we have $f(xr) = f(x)r$, which can only be the case if f was a left-multiplication by an element of R. Moreover, by Proposition 8.2.28 the map $m: R \to \mathrm{End}_{\mathsf{Ab}}(R)$ is injective — restricting m to the subring of $Z(R)$ given by $S := \{r \in Z(R) : {}_r m \in Z(\mathrm{End}_{\mathsf{Ab}}(R))\}$ makes $m|_S$ into a surjective morphism. Therefore the morphism $\overline{m}: S \simeq Z(\mathrm{End}_{\mathsf{Ab}}(R))$, where $\overline{m}(r) = m(r)$ for all $r \in S$, is an isomorphism of rings. ♮

Corollary 8.2.32. Let $n \in \mathbf{Z}_{>0}$ be a positive integer. There exists a *ring* isomorphism

$$\mathbf{Z}/n\mathbf{Z} \simeq \mathrm{End}_{\mathsf{Ab}}(\mathbf{Z}/n\mathbf{Z}).$$

8.3 Ideals & Quotients of Rings

Ideals

Definition 8.3.1 (Ideal)**.** Let R be a ring. A subgroup $\mathfrak{a}$ of $(R, +)$ is said to be a *left-ideal* of R if for all $r \in R$ we have $r\mathfrak{a} \subseteq \mathfrak{a}$. On the other hand, $\mathfrak{a}$ is a *right-ideal* if for all $r \in R$ we have $\mathfrak{a}r \subseteq \mathfrak{a}$. Furthermore, if $\mathfrak{a}$ is both a left and right ideal, we say that $\mathfrak{a}$ is a *two-sided-ideal* — or simply an *ideal*, without qualifiers, which will be the preferred nomenclature.

Remark 8.3.2 (Ideals and 1)**.** The only ideal of a ring R containing the unity 1_R is R itself — thus ideals need not be subrings.

Corollary 8.3.3. If $\phi: R \to S$ is a ring morphism such that $\operatorname{im}\phi$ is an ideal of S, then ϕ is surjective.

Proof. Since $\operatorname{im}\phi$ is a subring, it contains 1_S — by Remark 8.3.2 we have $\operatorname{im}\phi = S$. ♮

Corollary 8.3.4. The collection of ideals of a ring is closed under addition and intersection.

Proof. Let R be a ring and $\mathfrak{a}, \mathfrak{b} \subseteq R$ be ideals. Let $a, b \in R$ be any two elements, then $(a+b)(\mathfrak{a}+\mathfrak{b}) = (a+b)\mathfrak{a}+(a+b)\mathfrak{b}$ but since $(a+b)\mathfrak{a}, (a+b)\mathfrak{b} \subseteq \mathfrak{a}+\mathfrak{b}$, then $(a+b)(\mathfrak{a}+\mathfrak{b}) \subseteq \mathfrak{a}+\mathfrak{b}$ — the same analogous arguments can be used to show that $(\mathfrak{a}+\mathfrak{b})(a+b) \subseteq \mathfrak{a}+\mathfrak{b}$. If $a, b \in \mathfrak{a} \cap \mathfrak{b}$, then $(ab)\mathfrak{a} \cap \mathfrak{b}, \mathfrak{a} \cap \mathfrak{b}(ab) \subseteq \mathfrak{a}, \mathfrak{b}$ thus also contained in $\mathfrak{a} \cap \mathfrak{b}$. ♮

Corollary 8.3.5 (Kernel is an ideal)**.** Given a ring morphism $\phi: R \to S$, then $\ker\phi$ is a ring ideal of R.

Proof. From group theoretic considerations, we already know that $\ker \phi$ is a subring of R. On the other hand, let $r \in R$ be any element and $a \in \ker \phi$, then $\phi(ra) = \phi(r)\phi(a) = 0$ and $\phi(ar) = \phi(a)\phi(r) = 0$, thus indeed both $ra, ar \in \ker \phi$. ♮

Corollary 8.3.6 (Preimage of ideal is an ideal)**.** Let $\phi\colon R \to S$ be a ring morphism and $\mathfrak{a}$ be an ideal of S, then the preimage $\phi^{-1}(\mathfrak{a})$ is an ideal of R.

Proof. Let $r \in R$ and $a \in \phi^{-1}(\mathfrak{a})$ be any two elements. Notice that $\phi(ra) = \phi(r)\phi(a)$ and $\phi(ar) = \phi(a)\phi(r)$, and since $\phi(a) \in \mathfrak{a}$ then $\phi(ra), \phi(ar) \in \mathfrak{a}$ — therefore $ra, ar \in \phi^{-1}(\mathfrak{a})$. ♮

Remark 8.3.7 (Image of ideals)**.** Although the preimage of ideals is an ideal of the morphism's source, the *image* of a given ideal *need not be an ideal*. For instance, let $\mathfrak{a} \subseteq R$ be an ideal of a ring R and $\phi\colon R \to S$ be a *non-surjective* morphism, then there exists $s \in S$ whose preimage is the empty set, therefore $s\phi(\mathfrak{a})$ is not contained in $\phi(\mathfrak{a})$.

Corollary 8.3.8 (Surjective morphisms preserve ideals)**.** If $\phi\colon R \twoheadrightarrow S$ is a surjective ring morphism, then the image of any ideal $\mathfrak{a} \subseteq R$ is an ideal of S.

Proof. This is immediate from Remark 8.3.7, for all $s \in S$ we have both $s\phi(\mathfrak{a}), \phi(\mathfrak{a})s \subseteq \phi(\mathfrak{a})$ because there will exist $r \in R$ whose image is s. ♮

Example 8.3.9. For every element $r \in R$ of a ring R, the objects rR and Rr are, respectively, *right* and *left* ideals of R. Indeed, given any $a \in R$, we have $(rR)a = r(Ra) \subseteq rR$, the analogous being true for the left ideal. If R is commutative, we have $rR = Rr$ and such ideal is commonly denoted by (r) — the *principal ideal* generated by r.

Proposition 8.3.10. Let R be a ring and $(\mathfrak{a}_j)_{j\in J}$ be a collection of ideals of R, then the direct sum $\bigoplus_{j\in J} \mathfrak{a}_j$ is an ideal of R.

Proof. Let $r \in R$ be any element and consider any formal sum $\sum_{j\in J} a_j \in \bigoplus_{j\in J} \mathfrak{a}_j$ where $a_j \neq 0$ for only finitely many $j \in J$, then $r(\sum_{j\in J} a_j) = \sum_{j\in J} ra_j$ bus since $ra_j \in \mathfrak{a}_j$ for each $j \in J$, then $\sum_{j\in J} ra_j \in \bigoplus_{j\in J} \mathfrak{a}_j$ — the same argument can be used for right-multiplication. ♮

Lemma 8.3.11. Given a collection $(\mathfrak{a}_j)_{j\in J}$ of ideals of a ring R, the ideal $\bigoplus_{j\in J} \mathfrak{a}_j$ is the smallest ideal of R containing each ideal $\mathfrak{a}_j$ for $j \in J$.

Proof. Let $\mathfrak{b}$ be an ideal of R containing every ideal $\mathfrak{a}_j$ for $j \in J$. Then in particular, given any element $\sum_{j\in J} a_j \in \bigoplus_{j\in J} \mathfrak{a}_j$, since $a_j \in \mathfrak{b}$ for every $j \in J$ only finitely many such a_j are non-zero, we find that the sum $\sum_{j\in J} a_j \in \mathfrak{b}$ — thus indeed $\bigoplus_{j\in J} \mathfrak{a} \subseteq \mathfrak{b}$. ♮

Definition 8.3.12 (Finitely generated ideals)**.** An ideal $\mathfrak{a}$ of a *commutative* ring R is said to be *finitely generated* if there exists a finite set of elements $A \subseteq R$ such that $\mathfrak{a} = (A)$.

Division Rings and its Ideals

Proposition 8.3.13 (Units and its ideals)**.** Let R be a ring and $a \in R$. The element a is a *left-unit* (or *right-unit*) of R if and only if $R = aR$ (or $R = Ra$)

Proof. We prove the proposition only for left-units, for right-units the proof is completely analogous. Suppose a is a left-unit of R and let $u \in R$ be such that $au = 1$, then aR contains 1 which implies in $aR = R$. Now if $R = aR$, then $1 \in aR$, which means that there must exist $u \in R$ for which $au = 1$, thus a is a left-unit. ♮

Proposition 8.3.14 (Division ring ideals)**.** Let R be a ring. Then R is a *division ring* if and only if its *only* left-ideals and right-ideals are $\{0\}$ and R.

Proof. Let R be a division ring. Suppose $\mathfrak{a}$ is a left-ideal (or right-ideal) of R, then given any $r \in R$ and $a \in \mathfrak{a}$ we have $ra \in \mathfrak{a}$ (or $ar \in \mathfrak{a}$) — in particular, if $a \neq 0$, then $a^{-1} \in R$ thus $a^{-1}a = 1 \in \mathfrak{a}$ (or $aa^{-1} = 1 \in \mathfrak{a}$), thus $\mathfrak{a} = R$ or $\mathfrak{a} = 0$.

Suppose 0 and R are the only left and right ideals. Given any non-zero element $r \in R$, the ideal rR (or Rr) is non-zero, thus must be equal to R, which is equivalent of $1 \in rR$ (or $1 \in Rr$) — therefore r must be a unit of R, proving that R is a division ring. ♮

Corollary 8.3.15 (Field ideals)**.** The only left or right ideals of a field k are 0 and k.

Proposition 8.3.16. Let k be a field and R be a ring, then any ring morphism $k \to R$ is injective.

Proof. Let $\phi\colon k \to R$ be a ring morphism and let $a, b \in k$ be any two elements such that $\phi(a) = \phi(b)$, then $a - b \in \ker \phi$. Remember that $\ker \phi$ is an ideal of k but this means that $\ker \phi$ is either 0 or k — by Proposition 8.3.14. Since $\phi(1_k) = 1_R$ for ϕ to be a ring morphism, then $\ker \phi \neq k$ and we are left with $\ker \phi = 0$ — thus $a = b$. ♮

Nilpotent Ideal

Definition 8.3.17 (Nilpotent ideal)**.** Let R be a ring. We say that an ideal $\mathfrak{a} \subseteq R$ is a *nilpotent ideal* if there exists $k \in \mathbf{Z}_{>0}$ such that the product of any k elements of $\mathfrak{a}$ equals zero.

Nilradicals

Definition 8.3.18 (Nilradical)**.** Given a ring R, we define the *nilradical* of R to be the object N composed of every nilpotent element of R.

Corollary 8.3.19. The nilradical N of a *commutative ring* R is an *ideal* of R.

Proof. Let $a \in R$ and $x \in N$ be any two elements — suppose $n \in \mathbf{Z}_{>0}$ is such that $x^n = 0$. Since R is commutative then $(ax)^n = a^n x^n = a^n \cdot 0 = 0$, thus $ax \in N$. ♮

Definition 8.3.20 (Reduced ring)**.** A ring R is said to be reduced if it contains no non-zero nilpotent elements

Corollary 8.3.21. Let R be a *commutative ring* and N be its nilradical. The quotient ring R/N is *reduced*.

Proof. Suppose $a + N \in R/N$ is a nilpotent element and $a^n + N = N$ for some $n \in \mathbf{Z}_{>0}$ — which implies in $a^n \in N$ thus there must exist $m \in \mathbf{Z}_{>0}$ such that $(a^n)^m = a^{nm} = 0$, which implies in $a \in N$ and hence $a + N = N$. ♮

Lemma 8.3.22. A ring R is *reduced* if and only if for all $r \in R$ such that $r^2 = 0$ implies $r = 0$.

Proof. If R is reduced then clearly $r^2 = 0$ implies $r = 0$. Let $r \in R$ be any element such that $r^n = 0$ for some $n \in \mathbf{Z}_{>0}$, then proceed by the following algorithm — start with r^n, then:

- If $n = 1$, return r — since $r = 0$.
- If n is odd, $r^{n+1} = 0$ and thus $r^{(n+1)/2}r^{(n+1)/2} = 0$ which by hypothesis implies $r^{(n+1)/2} = 0$. Continue the algorithm for $r^{(n+1)/2}$.
- If n is even, $r^n = r^{n/2}r^{n/2} = 0$ implies $r^{n/2} = 0$. Continue the algorithm for $r^{n/2} = 0$.

Such algorithm is ensured to terminate and will always result in $r = 0$, which implies in R being a reduced ring. ♮

Quotient Ring

We define now, for every ideal $\mathfrak{a}$ of a given ring R, a ring $R/\mathfrak{a}$ together with an additive and multiplicative structure: for every $a + \mathfrak{a}, b + \mathfrak{a} \in R/\mathfrak{a}$, we define

$$(a + \mathfrak{a}) + (b + \mathfrak{a}) := (a + b) + \mathfrak{a},$$
$$(a + \mathfrak{a}) \cdot (b + \mathfrak{a}) := ab + \mathfrak{a}.$$

Let's show that both operations are well defined. Addition is clearly well defined. For the multiplication, suppose $a_1 + \mathfrak{a} = a_2 + \mathfrak{a}$ and $b_1 + \mathfrak{a} = b_2 + \mathfrak{a}$ — that is, both differences $a_1 - a_2$ and $b_1 - b_2 \in \mathfrak{a}$ are elements of $\mathfrak{a}$. Notice that

$$\begin{aligned} a_1b_1 - a_2b_2 &= a_1b_1 - a_2b_2 + (a_1b_2 - a_1b_2) \\ &= (a_1b_1 - a_1b_2) + (a_1b_2 - a_2b_2) \\ &= a_1(b_1 - b_2) + (a_1 - a_2)b_2 \end{aligned}$$

and therefore $a_1b_1 - a_2b_2 \in \mathfrak{a}$, implying in $a_1b_1 + \mathfrak{a} = a_2b_2 + \mathfrak{a}$ — thus multiplication is well defined.

Proposition 8.3.23 (Universal property of quotient rings)**.** Let R be a ring. For every ring S together with a ring morphism $\psi\colon R \to S$ and ideal $\mathfrak{a} \subseteq \ker\psi$, there exists a *unique* ring morphism $\phi\colon R/\mathfrak{a} \to S$ such that the following diagram commutes

$$\begin{tikzcd} R \arrow[r, "\psi"] \arrow[d, two heads, "\pi"'] & S \\ R/\mathfrak{a} \arrow[ur, dashed, bend right, "\phi"'] & \end{tikzcd}$$

Proof. Let S be any ring together with a ring morphism $\psi\colon R \to S$. Let $\pi\colon R \twoheadrightarrow R/\mathfrak{a}$ be the canonical projection morphism. Define a map $\phi\colon R/\mathfrak{a} \to S$ by sending $r+\mathfrak{a} \mapsto \psi(r)$. We show now that ϕ is indeed well defined, consider elements $x, y \in R/\mathfrak{a}$ to be such that $\phi(x) = \phi(y)$, then $\psi(x) = \psi(y)$ which is the same as $\psi(x)-\psi(y) = \psi(x-y) = 0$, therefore $x - y \in \ker \psi$, then $x = y$ in $R/\mathfrak{a}$. Moreover, ϕ inherits from ψ the preservation of both additive and multiplicative structures, thus ϕ is a morphism of rings and $\phi\pi = \psi$. If ϕ' is another morphism such that $\phi'\pi = \psi$, since π is surjective then $\phi\pi = \phi'\pi$ implies in $\phi = \phi'$, thus ϕ is unique. ♮

Corollary 8.3.24 (Every ideal is a kernel)**.** Given a ring R, for every ideal $\mathfrak{a} \subseteq R$ is the kernel of some ring morphism $R \to S$. Therefore an additive subgroup of a ring R is an ideal if and only if it is a kernel of some ring morphism.

Proof. Simply let $S = R/\mathfrak{a}$ and consider the canonical projection $\pi\colon R \to R/\mathfrak{a}$, whose kernel is clearly $\mathfrak{a}$. ♮

Definition 8.3.25 (Characteristic)**.** Let R be a ring and $\phi\colon \mathbf{Z} \to R$ be the unique ring morphism mapping $r \mapsto r \cdot 1_R$. We define the *characteristic* of R to be the non-negative integer $n \in \mathbf{Z}_{>0}$ such that $\ker \phi = n\mathbf{Z}$ — we denote such a property by $\operatorname{char} R = n$.

Proposition 8.3.26 (Characteristic of integral domains)**.** The characteristic of an *integral domain* is either *zero* or a *prime number*.

Proof. Let R be an integral domain and let $\operatorname{char} R = n$. Suppose that n is non-zero and that there exists an integer $m < n$ dividing n — that is, for some integer q we have $n = qm$. Then in particular $n \cdot 1_R = (q \cdot 1_R)(m \cdot 1_R)$ but $n \cdot 1_R = 0$ and since $q, m < n$ it follows that $q \cdot 1_R, m \cdot 1_R \neq 0$ in R — thus we obtained a contradiction since R is said to be an integral domain, hence $\operatorname{char} R$ is either prime or zero. ♮

Definition 8.3.27 (Boolean ring)**.** A ring R is said to be *boolean* if for all $r \in R$ we have $r^2 = r$.

Corollary 8.3.28. A non-zero *boolean* ring is *commutative* and has characteristic 2.

Proof. Let $x \in R$ be any element of a boolean ring R. Then we have $x+x = (x+x)^2 = 4x^2$ and $x + x = x^2 + x^2$, therefore, $2x^2 = 4x^2$ which implies in $x^2 + x^2 = 0$ but since $x^2 + x^2 = x + x$ then we conclude that $x + x = 0$ and thus $\operatorname{char} R = 2$.

If $x, y \in R$ are any two elements, then consider the sum

$$x + y = (x + y)^2 = x^2 + xy + yx + y^2 = x + xy + yx + y,$$

hence cancelling the common terms we obtain $xy + yx = 0$, now since $\operatorname{char} R = 2$ one can use the fact that $yx + yx = 0$ to make

$$0 = xy + yx = xy + yx - (yx + yx) = xy - yx,$$

thus $xy = yx$. ♮

Corollary 8.3.29. If R is a boolean ring and also an integral domain, then there exists a canonical isomorphism of rings $R \simeq \mathbf{Z}/2\mathbf{Z}$.

Proof. Let $r in R$ be any element, since $r^2 = r$ then $r^2 - r = r(1_R - r) = 0_R$ but since R is an integral domain, either $r = 0_R$ or $1_R - r = 0_R$, that is, $r = 1_R$. Thus the map $R \to \mathbf{Z}/2\mathbf{Z}$ sending $0_R \mapsto [0]_2$ and $1_R \mapsto [1]_2$ is an isomorphism of rings. ♮

Decompositions

Theorem 8.3.30 (First isomorphism)**.** Every ring morphism $\phi\colon R \to S$ can be decomposed into the commutative diagram

$$\begin{tikzcd} R \arrow[r, two heads] \arrow[rrr, bend left, "\phi"] & R/\ker\phi \arrow[r, "\simeq", "\overline{\phi}"'] & \operatorname{im}\phi \arrow[r, hook] & S \end{tikzcd}$$

where $\overline{\phi}\colon R/\ker\phi \xrightarrow{\simeq} \operatorname{im}\phi$ is the natural ring morphism induced by ϕ.

Proof. The morphism $\overline{\phi}$ is obtained by the quotient ring universal property — that is, $\overline{\phi}(r) = \phi(r) + \ker\phi$, which is a morphism $R/\ker\phi \to S$. Restricting the codomain of $\overline{\phi}$ we obtain the claimed isomorphism. The rest follows trivially. ♮

Corollary 8.3.31. Let $\phi\colon R \twoheadrightarrow S$ be a *surjective* ring morphism, then there exists a natural isomorphism

$$S \simeq R/\ker\phi$$

Proof. Indeed, if ϕ is surjective, then $\operatorname{im}\phi = S$ and by the first isomorphism theorem we obtain the natural isomorphism $S \simeq R/\ker\phi$. ♮

Proposition 8.3.32 (Ideal of a quotient)**.** Let R be a ring and $\mathfrak{a} \subseteq R$ be an ideal. If $\mathfrak{b}$ is an ideal of R and $\mathfrak{a} \subseteq \mathfrak{b}$, then $\mathfrak{b}/\mathfrak{a}$ is an ideal of $R/\mathfrak{a}$ and there exists a natural isomorphism

$$\frac{R/\mathfrak{a}}{\mathfrak{b}/\mathfrak{a}} \simeq R/\mathfrak{b}.$$

Proof. By Proposition 8.3.23, let $\phi\colon R/\mathfrak{a} \to R/\mathfrak{b}$ be the morphism making the following diagram commute

$$\begin{tikzcd} R \arrow[r, two heads] \arrow[d, two heads] & R/\mathfrak{b} \\ R/\mathfrak{a} \arrow[ur, dashed, bend right, "\phi"'] & \end{tikzcd}$$

That is, ϕ is defined by mapping $r + \mathfrak{a} \mapsto r + \mathfrak{b}$. Notice that the kernel of ϕ is composed of those elements $r + \mathfrak{a} \in R/\mathfrak{a}$ for which $\phi(r + \mathfrak{a}) = \mathfrak{b}$, that is, $r \in \mathfrak{b}$ for $r + \mathfrak{b} = \mathfrak{b}$ — thus $\ker\phi = \mathfrak{b}/\mathfrak{a}$, which is ensured to be an ideal. Since ϕ is surjective, by means of Corollary 8.3.31 we obtain a natural isomorphism $R/\mathfrak{b} \simeq \frac{R/\mathfrak{a}}{\mathfrak{b}/\mathfrak{a}}$. ♮

Corollary 8.3.33. Any surjective morphism $\phi\colon R \twoheadrightarrow S$ can be bijectively identified as a canonical projection $R \twoheadrightarrow R/\ker\phi$.

Example 8.3.34. If R is a *commutative* ring and we consider principal ideals (a) and (b) of R. If $[b] \in R/(a)$ is the class of b, then we have $([b]) = (a,b)/(a)$. By Proposition 8.3.32 we find a canonical isomorphism

$$\frac{R/(a)}{([b])} \simeq R/(a,b).$$

Generation of Ideals

Definition 8.3.35 (Noetherian ring)**.** A *commutative* ring R is said to be *Noetherian* if *every* ideal of R is *finitely* generated.

Proposition 8.3.36 (Image of Noetherian ring)**.** Let R be a Noetherian ring and S be a ring. If there exists a *surjective* ring morphism $R \twoheadrightarrow S$ then S is Noetherian.

Proof. Suppose there exists $\phi\colon R \twoheadrightarrow S$, a surjective ring morphism. Let $\mathfrak{s}$ be any ideal of S, since the preimage of an ideal is an ideal, then $\phi^{-1}(\mathfrak{s})$ is an ideal of R — thus finitely generated. Now, since ϕ is surjective, $\phi^{-1}(\mathfrak{s})$ is non-empty and there exists a finite set $A \subseteq \phi^{-1}(\mathfrak{s})$ such that $\phi^{-1}(\mathfrak{s}) = (A)$, since R is Noetherian. Therefore $\phi(A) = B \subseteq \mathfrak{s}$ is a finite set, hence $\phi((A)) = (B) = \mathfrak{s}$ — proving that $\mathfrak{s}$ is finitely generated. ♮

Definition 8.3.37 (Principal ideal domain)**.** An *integral domain* R is said to be a *principal ideal domain* (which we'll shortly name PID) if *every* ideal of R is *principal*.

Example 8.3.38. The ring of integers $\mathbf{Z}$ is a PID. Indeed, if $\mathfrak{a}$ is an ideal of $\mathbf{Z}$, then there exists $n \in \mathbf{Z}$ for which $\mathfrak{a}$ is a subgroup of $n\mathbf{Z} = (n)$, thus $\mathfrak{a}$ itself is principal.

Notice also that, given any two integers $m, n \in \mathbf{Z}$, if $d := \gcd(m,n)$ then $m, n \in (d)$, which implies that $(m,n) \subseteq (d)$. Moreover, from Bézout's identity[1], we have the existence of $a, b \in \mathbf{Z}$ for which $am + bn = d$, thus $d \in (m,n)$ — proving that $(m,n) = (d)$.

Example 8.3.39. The ring $\mathbf{Z}[x]$ is *not* a PID.

We consider the ideal $(2,x)$ and show that it isn't principal. Suppose there exists $f(x) \in \mathbf{Z}[x]$ for which $(f(x)) = (2,x)$, so that there exists $q(x) \in \mathbf{Z}[x]$ such that $q(x)f(x) = 2$. Notice however that since $\mathbf{Z}$ is an integral domain, the product of the leading term coefficients of $f(x)$ and $q(x)$ is necessarily non-zero (for non-zero polynomials), thus $\deg f(x)g(x) = \deg f(x) + \deg g(x)$. Notice that $\deg q(x)f(x) = 0$

[1] Using the well order on (m,n), let $\ell := a_0 m + b_0 n$ be the smallest element of (m,n). Notice that given any other $u = am + bn \in (m,n)$, by the euclidean division algorithm there exists two integers $q, r \in \mathbf{Z}$ for which $u = q\ell + r$ and $0 \leqslant r < \ell$. Therefore, one can write

$$r = u - q\ell = (am + bn) - q(a_0 m + b_0 n) = (a - qa_0)m + (b - qb_0)n$$

so that $r \in (m,n)$ — since ℓ is the smallest element, then $r = 0$ and thus ℓ divides u.

In general, we have shown that ℓ divides every element of (m,n), and in particular ℓ also divides both m and n. Now, if $c \in \mathbf{Z}$ is any common divisor of m and n then in particular c divides $a_0 m + b_0 n = \ell$ — showing that $c \leqslant \ell$ and hence $\ell = \gcd(m,n)$.

thus $\deg f(x) = 0$ and this can't be the case since we would have $x \notin (f(x))$ — this is a contradiction, such polynomial $f(x)$ cannot exist and therefore $(2, x)$ isn't principal.

Proposition 8.3.40. Given a field k, the ring of polynomials $k[x]$ is a PID.

Proof. The zero ideal is obviously principal, thus let $\mathfrak{a}$ be any non-zero proper ideal of $k[x]$ and, by the well-ordering of the positive degree polynomials in $k[x]$, let $f(x) \in \mathfrak{a}$ be a non-zero monic polynomial with least degree. We can indeed be certain that we can choose an $f(x)$, because every non-zero coefficient of a polynomial of $k[x]$ is a unity — since $\mathfrak{a}$ is non-zero, we are ensured that there will exist a polynomial with at least one non-zero coefficient.

Let $g(x) \in k[x]$ be any polynomial and, since k is a field, let $q(x), r(x) \in k[x]$ be polynomials such that $g(x) = q(x)f(x) + r(x)$, where $\deg r < \deg f(x)$. Notice then that the last condition implies in $\deg r(x) \leqslant 0$, otherwise $f(x)$ wouldn't be the monic polynomial with least degree among polynomials of positive degree in $k[x]$. Now, if $r(x) = 0$, then $g(x) = q(x)f(x)$ and $g(x) \in (f(x))$ — on the other hand, if $r(x) = a$ for some $a \in k$, then $a \in \mathfrak{a}$ but since k is a field, this implies in $1 \in \mathfrak{a}$ thus $\mathfrak{a} = k[x]$ and hence a contradiction to the hypothesis that $\mathfrak{a}$ is proper. Assuming the latter does not hold, we find that $(f(x)) = \mathfrak{a}$ and hence $k[x]$ is a PID. ♮

Definition 8.3.41 (Product of ideals)**.** Given a ring R and ideals $\mathfrak{a}$ and $\mathfrak{b}$ of R, we denote by $\mathfrak{ab}$ the ideal generated by all products ab for $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$.

Lemma 8.3.42. Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals of a *commutative* ring R. If either one of the following properties is true:

(a) The ideals $\mathfrak{a}$ and $\mathfrak{b}$ are *comaximal*, that is, $\mathfrak{a} + \mathfrak{b} = R$.

(b) The quotient $R/(\mathfrak{ab})$ is a *reduced* ring.

Then it follows that

$$\mathfrak{ab} = \mathfrak{a} \cap \mathfrak{b}.$$

Proof. If $ab \in \mathfrak{ab}$ is any element, then $a, b \in \mathfrak{a} \cap \mathfrak{b}$ and hence $ab \in \mathfrak{a} \cap \mathfrak{b}$ — thus in either cases we have $\mathfrak{ab} \subseteq \mathfrak{a} \cap \mathfrak{b}$. We now prove the other side of the inclusion for each of the properties — let $\ell \in \mathfrak{a} \cap \mathfrak{b}$ be any element:

(a) Since $\mathfrak{a} + \mathfrak{b} = R$, then there are $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$ such that $a + b = 1$. Moreover, since $\ell \in \mathfrak{a} \cap \mathfrak{b}$, it follows that $a\ell + b\ell = \ell$ is an element of $\mathfrak{ab}$, thus $\mathfrak{a} \cap \mathfrak{b} \subseteq \mathfrak{ab}$.

(b) Notice that we have $\ell^2 \in \mathfrak{ab}$, thus $\ell^2 + \mathfrak{ab} = \mathfrak{ab}$. Since $R/(\mathfrak{ab})$ is reduced, it follows that $\ell + \mathfrak{ab} = \mathfrak{ab}$ (see Lemma 8.3.22) — thus $\ell \in \mathfrak{ab}$ and $\mathfrak{a} \cap \mathfrak{b} \subseteq \mathfrak{ab}$.

♮

Lemma 8.3.43. Let R be a ring and $f(x) \in R[x]$ be a *monomial*. Then $f(x)$ is *not* a left or right zero-divisor.

Proof. Since $f(x)$ is is a monomial and there is no element $r \in R$ other than zero for which $r \cdot 1_R = 0$ (where 1_R is the coefficient of the leading term of $f(x)$), then it follows that the only polynomial $z(x) \in R[x]$ for which $f(x)z(x) = 0$ (or $z(x)f(x) = 0$) is the zero polynomial $z(x) = 0$. ♮

Lemma 8.3.44 (Degree of the product of polynomials)**.** If $f(x) \in R[x]$ is *monic* a polynomial and R is a ring, then for every $g(x) \in R[x]$ we have

$$\deg(f(x)g(x)) = \deg f(x) + \deg g(x).$$

One can extend the euclidean division algorithm to the ring of polynomials, where we'll be able to divide polynomials by monomials.

Lemma 8.3.45 (Division of polynomials)**.** Given a *monic* polynomial $f(x) \in R[x]$, for some ring R, then for all $g(x) \in R[x]$ there exists two polynomials $q(x), r(x) \in R[x]$ for which

$$g(x) = f(x)q(x) + r(x),$$

and $\deg r(x) < \deg f(x)$. Moreover, such polynomials $q(x)$ and $r(x)$ are unique.

Proof. Indeed, if $q'(x), r'(x) \in R[x]$ are polynomials — and, for the sake of contradiction, *distinct* from the respective $q(x)$ and $r(x)$ — for which $g(x) = f(x)q'(x) + r'(x)$, and $\deg r'(x) < \deg f(x)$, we find that $f(x)q(x) + r(x) = f(x)q'(x) + r'(x)$ and thus

$$r(x) - r'(x) = f(x)(q'(x) - q(x)). \tag{8.1}$$

Notice however that by hypothesis $\deg r(x), \deg r'(x) < \deg f(x)$, thus $\deg(r(x) - r'(x)) \leqslant \max(\deg r(x), \deg r'(x)) < \deg f(x)$ — which is in contradiction with Eq. (8.1) since for such equation to be true we should have $\deg(r(x) - r'(x)) = \deg(f(x)(q'(x) - q(x))) = \deg f(x) + \deg(q'(x) - q(x))$. Therefore $r(x) - r'(x)$ is necessarily the zero polynomial and $f(x)q(x) = f(x)q'(x)$, now since $f(x)$ is a monomial and by Lemma 8.3.43 is not a zero-divisor, we use Proposition 8.1.14 to conclude that $q(x) = q'(x)$. ♮

In what follows, one should regard $R^{\oplus d}$, for a ring R, to be the ring of polynomials with degree less than or equal to d — this observation is based on the fact that one can map injectively

$$\Psi \colon R^{\oplus d} \rightarrowtail R[x] \text{ sending } (r_0, \dots, r_{d-1}) \mapsto \sum_{j=0}^{d-1} r_j x^j, \tag{8.2}$$

which is a morphism of *abelian groups*. Therefore, the restriction of the codomain to the image of such map induces an *isomorphism of abelian groups*

$$R^{\oplus} \simeq \operatorname{im} \Psi \subseteq R[x]. \tag{8.3}$$

Proposition 8.3.46. Let R be a *commutative* ring, and $f(x) \in R[x]$ a *monic* polynomial with degree $d \in \mathbf{Z}_{>0}$. Define a map $\phi \colon R[x] \to R^{\oplus d}$ sending $g(x) \mapsto (r_0, \dots, r_{d-1})$, where $r(x) := \Psi(r_0, \dots, r_{d-1}) \in R[x]$ is the *remainder* of the division of $g(x)$ by $f(x)$. Such map ϕ induces a natural isomorphism of *abelian groups*

$$R[x]/(f(x)) \simeq R^{\oplus d}.$$

Proof. Notice that every polynomial $g(x)$ can be divided by a monic polynomial $f(x)$ and the remainder $r(x)$ of such division is unique by Lemma 8.3.45. Moreover, the isomorphism in Eq. (8.3) is a right inverse of ϕ, thus ϕ is surjective. Moreover, $\ker \phi = (f(x))$ since every polynomial in the principal ideal $(f(x))$ is divisible by $f(x)$ and hence has a zero remainder when divided by $f(x)$.

Let's check that ϕ is indeed a morphism of abelian groups. Let $g(x), g'(x) \in R[x]$ be two polynomials and let $q(x), r(x) \in R[x]$ and $q'(x), r'(x) \in R[x]$ be their respective pair of quotient and remainder in the division by $f(x)$ — where $\deg r(x), \deg r'(x) < \deg f(x)$. Therefore, the sum of $g(x)$ with $g'(x)$ can be written as

$$g(x) + g'(x) = f(x)(q(x) + q'(x)) + (r(x) + r'(x)).$$

Since $\deg(r(x) + r'(x)) \leqslant \max(\deg r(x), \deg r'(x)) < \deg f(x)$, then $r(x) + r'(x)$ is the remainder of the division of $g(x) + g'(x)$ by $f(x)$. With this in our hands we can rightly see that

$$\begin{aligned} \phi(g(x) + g'(x)) &= (r_0 + r'_0, \dots, r_{d-1} + r'_{d-1}) \\ &= (r_0, \dots, r_{d-1}) + (r'_0, \dots, r'_{d-1}) \\ &= \phi(g(x)) + \phi(g'(x)), \end{aligned}$$

where $\Psi(r_0, \dots, r_{d-1}) \coloneqq r(x)$ and $\Psi(r'_0, \dots, r'_{d-1}) \coloneqq r'(x)$ — thus ϕ is a morphism of abelian groups. By Proposition 7.4.16 we find

$$R[x]/\ker \phi = R[x]/(f(x)) \simeq \operatorname{im} \phi = R^{\oplus d}.$$

♮

Example 8.3.47 (Evaluation)**.** Let R be a commutative ring. The evaluation morphism $\mathrm{eval}_a \colon R[x] \to R$ (see Proposition 8.2.9), mapping $f(x) \mapsto f(a)$, induces a natural isomorphism of *abelian groups*

$$R[x]/(x - a) \simeq R.$$

Such isomorphism, name it ϕ, is given by the map $g(x) + (x - a) \mapsto r$, where $r \in R$ is the remainder of the division of $g(x)$ by $x - a$.

First of all, if $f(x) \in R[x]$ then the division by $x - a$ yields $q(x) \in R[x]$ and $r \in R$ such that $f(x) = (x - a)q(x) + r$ since the degree of the remainder should be less than $\deg(x-a) = 1$. Therefore, evaluating $\mathrm{eval}_a(f(x)) = (a-a)q(x) + r = r$ implies in $f(a) = r$ — thus $f(x) \in \ker \phi$ if and only if $f(x) \in (x - a)$, hence $\ker \phi = (x - a)$. Moreover, we can now rethink of the ϕ as the mapping $f(x) + (x - a) \mapsto f(a)$.

If $u(x) = u(a) + (x-a)$ and $v(x) = v(a) + (x-a)$ are any two polynomials in $R[x]/(x-a)$, one finds that $u(x) + v(x) = (u(a) + v(a)) + (x - a)$ and therefore $\phi(u(x) + v(x)) = u(a) + v(a) = \phi(u(x)) + \phi(v(x))$, thus ϕ is a morphism of abelian groups. Therefore, by means of Corollary 8.3.31 we find that ϕ is an isomorphism.

Example 8.3.48 (Constructing $\mathbf{C}$)**.** Consider, as in Proposition 8.3.46, the map $\phi \colon \mathbf{R}[x] \to \mathbf{R} \oplus \mathbf{R}$ sending $f(x) \mapsto (r_0, r_1)$, where $r(x) = r_0 + r_1 x$ is the remainder of the division of $f(x)$ by the polynomial $x^2 + 1 \in \mathbf{R}[x]$.

For any element $f(x) := a_0 + a_1 x \in \mathbf{R}[x]$ we have $\deg f(x) < \deg(x^2+1) = 2$, then $\phi(f(x)) = (a_0, a_1)$. Hence, given another $g(x) = b_0 + b_1 x \in \mathbf{R}[x]$, one has that

$$\begin{aligned} f(x)g(x) &= (a_0 + a_1 x)(b_0 + b_1 x) \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0)x + a_1 b_1 x^2 \\ &= (x^2+1)a_1 b_1 + ((a_0 b_0 - a_1 b_1) + (a_0 b_1 + a_1 b_0)x). \end{aligned}$$

Thus $\phi(f(x)g(x)) = (a_0 b_0 - a_1 b_1, a_0 b_1 + a_1 b_0)$, this induces a multiplicative structure $\cdot\colon (\mathbf{R} \oplus \mathbf{R})^2 \to \mathbf{R} \oplus \mathbf{R}$ defined by

$$(a_0, a_1) \cdot (b_0, b_1) := (a_0 b_0 - a_1 b_1, a_0 b_1 + a_1 b_0).$$

Now, since $(\mathbf{R} \oplus \mathbf{R}, \cdot) \simeq \mathbf{C}$, by identifying $(a, b) \mapsto a + bi$, one obtains the following natural isomorphism

$$\mathbf{R}[x]/(x^2+1) \simeq \mathbf{C},$$

hence we've constructed the complex numbers out of the ring $\mathbf{R}[x]$.

Example 8.3.49. Let R be a *commutative* ring and elements $a_1, \dots, a_n \in R$. Then there exists a canonical isomorphism of *abelian groups*

$$\frac{R[x_1, \dots, x_n]}{(x_1 - a_1, \dots, x_n - a_n)} \simeq R.$$

Denote $S := \frac{R[x_1,\dots,x_n]}{(x_1-a_1,\dots,x_n-a_n)}$. Consider the evaluation map $\mathrm{eval}\colon S \to R$ given by $[p(x_1, \dots, x_n)] \mapsto p(a_1, \dots, a_n)$, which is clearly a group morphism. Moreover, the map $\rho\colon R \to S$ mapping $r \mapsto [r]$ is both a group morphism and right inverse of the evaluation map, $\mathrm{eval} \circ \rho = \mathrm{id}_R$. Since for all $1 \leqslant j \leqslant n$ we have $[x_j] = [a_j]$ as classes in the quotient ring S, we find that, for any given $p(x_1, \dots, x_n) \in R[x]$

$$[p(x_1, \dots, x_n)] = [p(a_1, \dots, a_n)],$$

implying in $\rho \circ \mathrm{eval} = \mathrm{id}_S$.

Prime & Maximal Ideals

Definition 8.3.50 (Prime and maximal ideal)**.** Let R be a *commutative* ring and $\mathfrak{a}$ be a *proper* ideal of R. We define the following:

(a) If $R/\mathfrak{a}$ is an *integral domain*, we call $\mathfrak{a}$ a *prime ideal*.

(b) If $R/\mathfrak{a}$ is a *field*, we call $\mathfrak{a}$ a *maximal ideal*.

Example 8.3.51. Given a commutative ring R and any monomial $x - a \in R[x]$, then:

- The ideal $(x - a)$ is prime if and only if R is an integral domain.
- The ideal $(x - a)$ is maximal if and only if R is a field.

Such propositions are direct implications of the isomorphism $R[x]/(x-a) \simeq R$ given by Proposition 8.3.46.

Some may want to avoid the machinery of quotient rings when defining prime and maximal ideals. Such ambition is palpable, the following proposition exposes this old fashioned alternative.

Proposition 8.3.52. Let $\mathfrak{a}$ be a *proper* ideal of a *commutative* ring R — then:

(a) The ideal $\mathfrak{a}$ is *prime* if and only if for every given $a, b \in R$ such that $ab \in \mathfrak{a}$, we have either $a \in \mathfrak{a}$ or $b \in \mathfrak{a}$.

(b) The ideal $\mathfrak{a}$ is *maximal* if and only the only ideals that contain $\mathfrak{a}$ is either $\mathfrak{a}$ itself or R.

Proof. 1. The quotient ring $R/\mathfrak{a}$ is an integral domain if and only if for all $a+\mathfrak{a}, b+\mathfrak{a} \in R/\mathfrak{a}$ such that $(a+\mathfrak{a})(b+\mathfrak{a}) = \mathfrak{a}$ we have $a+\mathfrak{a} = \mathfrak{a}$ or $b+\mathfrak{a} = \mathfrak{a}$ — which is equivalent to the requirement that either $a \in \mathfrak{a}$ or $b \in \mathfrak{b}$.

2. The quotient ring $R/\mathfrak{a}$ is a field if and only if $\mathfrak{a}$ and $R/\mathfrak{a}$ are the only ideals of $R/\mathfrak{a}$ — that is, if $\mathfrak{b}$ is an ideal of R with $\mathfrak{a} \subseteq \mathfrak{b}$, then $\mathfrak{b}$ is also an ideal of $R/\mathfrak{a}$ (see Proposition 8.3.32), thus either $\mathfrak{b} = \mathfrak{a}$ or $\mathfrak{b} = R/\mathfrak{a}$.

♮

Corollary 8.3.53. Maximal ideals are prime ideals.

Proof. Let R be a commutative ring and $\mathfrak{m}$ be maximal. Let $a, b \in R$ be any elements such that $ab \in \mathfrak{m}$, then consider the principal ideals (a) and (b). Since $\mathfrak{m}$ is maximal, such principal ideals can be either R or $\mathfrak{m}$ — for the former case, one of them equals 1_R, thus the other is contained in $\mathfrak{m}$, on the other hand, for the latter case the element whose ideal equal to $\mathfrak{m}$ belongs to $\mathfrak{m}$. This shows that either $a \in \mathfrak{m}$ or $b \in \mathfrak{m}$. ♮

Proposition 8.3.54. Let $\mathfrak{a}$ be an ideal of a *commutative* ring R. If $R/\mathfrak{a}$ is a *finite* ring, then $\mathfrak{a}$ is *prime* if and only if it is *maximal*.

Proof. Since a finite commutative ring is an integral domain if and only if it is a field, the proposition follows — see Proposition 8.1.31. ♮

Example 8.3.55. A good example of Proposition 8.3.54 in action is the case of the ring $\mathbf{Z}$ and the principal ideals (n) for $n \in \mathbf{Z}_{>0}$. Since $\mathbf{Z}/(n)$ is always finite, (n) is a prime ideal if and only if (n) is maximal — on the other hand, (n) is maximal if and only if in the case where n is a prime number.

Definition 8.3.56 (Ring spectrum)**.** Given a *commutative* ring R, we define the following:

(a) The collection of all *prime ideals* of R is called the *prime spectrum* of R and is denoted $\operatorname{Spec} R$.

(b) The collection of all *maximal ideals* of R is called the *maximal spectrum* of R and is denoted $\operatorname{Spec}_{\mathrm{m}} R$.

Proposition 8.3.57 (Prime & maximal ideals in PIDs)**.** Let R be a *PID*, and let $\mathfrak{a}$ be a *non-zero* ideal of R. Then the ideal $\mathfrak{a}$ is *prime* if and only if it is *maximal*.

Proof. Suppose $\mathfrak{a}$ is prime. Since R is PID, we can be certain that there exists a non-zero element $a \in R$ such that $\mathfrak{a} = (a)$. Now, consider $\mathfrak{b} := (b)$ to be any ideal of R containing $\mathfrak{b}$, then in particular there exists $q \in \mathbf{R}$ such that $a = bq$ — since $\mathfrak{a}$ is prime and $bq \in \mathfrak{a}$ then either $b \in \mathfrak{a}$ or $q \in \mathfrak{a}$:

- For the case where $b \in \mathfrak{a}$ then $\mathfrak{b} \subseteq \mathfrak{a}$ — thus from hypothesis that $\mathfrak{b}$ contained $\mathfrak{a}$ we conclude that $\mathfrak{b} = \mathfrak{a}$.
- If $q \in \mathfrak{a}$, then there exists $d \in R$ such that $q = ad$, therefore $a = bq = b(ad) = b(da) = (bd)a$ — since a PID is an integral domain and $a \neq 0$, one can cancel a from both sides to obtain $bd = 1$, thus b is a unit element of R. In particular $bd = 1 \in \mathfrak{b}$, thus $\mathfrak{b} = R$.

By Proposition 8.3.52 we conclude that $\mathfrak{a}$ is maximal. ♮

Example 8.3.58 (Prime ideals of $k[x]$ are maximal)**.** Let k be a field. Then non-zero prime ideals of $k[x]$ are maximal. Indeed, from Proposition 8.3.40 we know that $k[x]$ is PID, thus this is all but a consequence of Proposition 8.3.57.

Proposition 8.3.59 (Maximal ideals in algebraically closed fields)**.** Let k be an *algebraically closed field* and $\mathfrak{a}$ be an ideal of $k[x]$. Then $\mathfrak{a}$ is maximal if and only if there exists some $a \in k$ for which $\mathfrak{a} = (x - a)$.

Proof. From Example 8.3.51 we know that $(x - a)$ is a maximal ideal of $k[x]$, since k is a field.

Suppose $\mathfrak{a}$ is maximal, since $k[x]$ is a PID, it follows that there exists an $f(x) \in k[x]$ for which $\mathfrak{a} = (f(x))$. Moreover, since k is algebraically closed, there exists $a \in R$ for which $f(a) = 0$, therefore $f(x) \in (x - a)$ and then $\mathfrak{a} \subseteq (x - a)$. From the hypothesis that $\mathfrak{a}$ is maximal, we conclude that either $(x - a) = \mathfrak{a}$ or $(x - a) = k[x]$ — notice that the latter cannot be the case since the collection of constant polynomials is not contained in $(x - a)$, thus $\mathfrak{a} = (x - a)$. ♮

Example 8.3.60. Let K be a *compact* topological space and $C(K)$ be the ring of continuous maps $K \to \mathbf{R}$, with addition and multiplication defined point-wise. We have the following facts:

(a) For every $p \in K$, the collection $\mathfrak{m}_p := \{f \in C(K) : f(p) = 0\}$ is a *maximal ideal* of R.

(b) If $f_1, \dots, f_n \in C(K)$ have no common zeros, then $(f_1, \dots, f_n) = R$.

(c) For every maximal ideal $\mathfrak{m} \in \operatorname{Spec}_{\mathrm{m}} C(K)$, there exists a point $p \in K$ for which $\mathfrak{m} = \mathfrak{m}_p$.

(d) If in addition K is Hausdorff, then there exists a *bijective set-function* $K \xrightarrow{\simeq} \operatorname{Spec}_{\mathrm{m}} C(K)$ given by $p \mapsto \mathfrak{m}_p$.

Proof. We now prove each one of these statements.

(a) Consider any map $f + \mathfrak{m}_p \in C(K)/\mathfrak{m}_p$ such that $f(p) \neq a$ for some $a \neq 0$ and therefore $f \notin \mathfrak{m}_p$. Notice that the unit of $C(K)/\mathfrak{m}_p$ is composed of every element $g \in C(K)$ such that $g(p) = 1$. In particular, since a is non-zero, then there exists an inverse $a^{-1} \in \mathbf{R}$. If $g \in C(K)$ is any map assuming $g(p) = a^{-1}$, then $(f + \mathfrak{m}_p)(g + \mathfrak{m}_p) = 1 + \mathfrak{m}_p$.

(b) Since $f_1, \dots, f_n$ share no zeros, the function $f := \sum_{j=1}^n f_j^2$ must be strictly positive, therefore one can define a continuous map $g \in C(K)$ as $g(x) := \frac{1}{f(x)}$ for all $x \in K$. Therefore $gf = 1$, which implies in $1 \in (f_1, \dots, f_n)$ — where we denote by 1 the constant map assuming value 1.

(c) Let $\mathfrak{m}$ be maximal and suppose, for the sake of contradiction, that for all $p \in K$ we have $\mathfrak{m} \neq \mathfrak{m}_p$. Fix any $p \in K$ and let $f_p \in \mathfrak{m}$ be such that $f_p(p) \neq 0$. By the continuity of f_p there exists a neighbourhood $U_p \subseteq K$ of p such that $f(x) \neq 0$ for all $x \in U_p$. Let $\mathcal{U} := \{U_p\}_{p \in K}$ be an open cover of K, where each neighbourhood U_p is associated with a map f_p as described above. Since K is compact, there exists a finite collection $U_{p_1}, \dots, U_{p_n} \in \mathcal{U}$ covering K. From construction, we obtain $(f_{p_1}, \dots, f_{p_n}) \subseteq \mathfrak{m}$. Notice that the maps f_{p_j} share no common zero — if $q \in K$ is such that $f_{p_j}(q) = 0$ for each $1 \leqslant j \leqslant n$, then $q \notin \bigcup_{j=1}^n U_{p_j}$, which is a contradiction — therefore $1 \in \mathfrak{m}$ and $\mathfrak{m} = C(K)$. We conclude that there must exist $p \in K$ such that $\mathfrak{m} = \mathfrak{m}_p$.

Prove last item, requires Urysohn's lemma

$\natural$

Definition 8.3.61 (Krull dimension)**.** The *Krull dimension* of a commutative ring R is defined as the length of the longest chain of prime ideals in R.

8.4 Modules Over Rings

Modules

Definition 8.4.1 (Ring action)**.** Let R be a ring and A an abelian group. A left-R-action on A is given a ring morphism

$$\mu \colon R \to \mathrm{End}_{\mathtt{Ab}}(M).$$

Less compactly, given $r, s \in R$ and a group endomorphism $\phi \colon A \to A$, the ring morphism satisfies, for every $a, b \in A$:

(a) $\mu(r)(a + b) = \mu(r)(a) + \mu(r)(b)$.

(b) $\mu(r + s)(a) = \mu(r)(a) + \mu(s)(a)$.

(c) $\mu(rs)(a) = \mu(r)(\mu(s)(a))$.

(d) $\mu(1)(a) = a$.

Definition 8.4.2 (Module)**.** Given a ring R, a *left-R-module* is an *abelian group* M endowed with a *left-R-action* $R \times M \to M$ mapping $(r, m) \mapsto rm$ satisfying the following properties, for every $r, s \in R$ and $m, n \in M$:

(a) $r(m + n) = rm + rn$.

(b) $(r + s)m = rm + sm$.

(c) $(rs)m = r(sm)$.

(d) $1m = m$.

Right-R-modules are defined completely analogous, with a right-R-action.

Definition 8.4.3 (Opposite ring)**.** Given a ring $(R, +, \cdot)$, we define its *opposite ring* R^{op} to be ring inheriting the elements and additive structure of R, while its multiplicative structure is given by a map $*\colon R \times R \to R$ defined by $a * b := ba \in R$.

Corollary 8.4.4. The identity map $\mathrm{id}\colon R \to R^{\mathrm{op}}$ is an *isomorphism* of rings if and only if R is *commutative*.

Proof. If R is commutative, clearly the identity map is an isomorphism. Conversely, if id is a morphism of rings, then $\mathrm{id}(rs) = \mathrm{id}(r) * \mathrm{id}(s) = \mathrm{id}(s)\,\mathrm{id}(r) = sr$ but $\mathrm{id}(rs) = rs$ thus $rs = sr$. ♮

Example 8.4.5. The ring of real n-square matrices, $M_n(\mathbf{R})$, is isomorphic as a ring to its opposite ring $M_n(\mathbf{R})^{\mathrm{op}}$.

Indeed, if $\phi\colon M_n(\mathbf{R}) \to M_n(R)^{\mathrm{op}}$ is the map sending $A \mapsto A^T$ then, given $A, B \in M_n(\mathbf{R})$, we have

$$\phi(AB) = (AB)^T = B^T A^T = A^T * B^T = \phi(A) * \phi(B).$$

That is, ϕ is a morphism of rings. Injectivity and surjectivity are clear, thus ϕ is an isomorphism of rings.

Example 8.4.6. For a *commutative* ring R, left-R-modules can be bijectively assigned to a right-R-modules. To see that given an abelian group M, let ${}_RM$ be a left-R-module structure on M, we proceed by constructing a bijection ${}_RM \xrightarrow{\simeq} M_R$, where M_R denotes a right-R-module on M. For every element $r \in R$ and $m \in M$, map $rm \mapsto mr$. Since R is commutative, indeed

$$(rs)m = r(sm) \longmapsto m(rs) = m(sr) = (ms)r,$$

translating the multiplication by rs from a left-module to a right-module accordingly.

Vector Spaces

Example 8.4.7 (Vector spaces)**.** A module over a field k is nothing more than a k-vector space, thus $k\text{-}\mathtt{Mod} = \mathtt{Vect}_k$.

Example 8.4.8 ($k[x]$-module structure on V)**.** Let k be a field, V be a k-vector space, and $\phi\colon V \to V$ be an endomorphism. We consider the ring action $\mu\colon k \hookrightarrow \mathrm{End}_{\mathtt{Vect}_k}(V)$ given by $a \mapsto a\,\mathrm{id}_V$, which by Proposition 8.2.9 implies in the existence of a unique morphism of rings $\Psi\colon k[x] \to \mathrm{End}_{\mathtt{Vect}_k}(V)$ such that the following diagram commutes

$$\begin{tikzcd} k \arrow[r, "\mu"] \arrow[d, hook] & \mathrm{End}_{\mathtt{Vect}_k}(V) \\ k[x] \arrow[ur, dashed, bend right, "\Psi"'] & \end{tikzcd}$$

and that $x \overset{\Psi}{\mapsto} \phi$. Since k and x generate $k[x]$, for any $f(x) \coloneqq \sum_{j=1}^n a_j x^j \in k[x]$, we have a mapping

$$f(x) = \textstyle\sum_{j=1}^n a_j x^j \overset{\Psi}{\longmapsto} a_0\,\mathrm{id}_V + a_1\phi + \cdots + a_{n-1}\phi^{n-1} + a_n\phi^n \coloneqq f(\phi).$$

The ring action $\Psi\colon k[x] \to \mathrm{End}_{\mathtt{Vect}_k}(V)$ induces the structure of a left-$k[x]$-module on V as

$$p(x) \cdot v \coloneqq p(\phi)(v),$$

for all $p(x) \in k[x]$ and $v \in V$.

Corollary 8.4.9. Given a field k, there exists a bijection

$$\{k[x]\text{-modules}\} \xrightarrow{\simeq} \{(V, \phi) : V \in \mathtt{Vect}_k \text{ and } \phi \in \mathrm{End}_{\mathtt{Vect}_k}(V)\}.$$

In other words, $k[x]$-modules are equivalent to k-vector spaces together with a uniquely determined k-linear endomorphism.

Proof. As constructed in Example 8.4.8, for every k-vector space V and k-linear endomorphism $\phi\colon V \to V$, one has a unique left-$k[x]$-module structure induced on V. On the other hand, let M be a left-$k[x]$-module and $\psi\colon M \to M$ be the group morphism given by $m \mapsto xm$ for all $m \in M$. Notice that for any $a \in k$ and $m \in M$ we have

$$\psi(am) = x(am) = (xa)m = (ax)m = a(xm) = a\psi(am),$$

therefore ψ is k-linear. Thus (M, ϕ) is the corresponding uniquely defined k-vector space together with a k-linear endomorphism. ♮

Category of Modules

Definition 8.4.10 (Morphism of modules)**.** Let R be a ring, and both M and N be abelian groups. Let M_R and N_R denote right-R-modules on M and N, while ${}_RM$ and ${}_RN$ denote left-R-modules on M and N. We define the the following:

(a) A *morphism* between *right*-R-modules $\phi\colon M_R \to N_R$ is a morphism of abelian groups such that, for all $r \in R$ and $m \in M$ we have

$$\phi(mr) = \phi(m)r.$$

(b) A *morphism* between *left-R*-modules $\psi\colon {}_R M \to {}_R N$ is a morphism of abelian groups such that, for all $r \in R$ and $m \in M$ we have

$$\psi(rm) = r\psi(m).$$

Morphisms of R-modules can also be compactly named R-linear morphisms.

Definition 8.4.11 (Category of R-modules)**.** Given a ring R, we denote by $\mathtt{Mod}_R$ the category whose objects are *right-R-modules* and morphisms between them. Analogously, we define ${}_R\mathtt{Mod}$ to be the category whose objects are *left-R-modules* and morphisms between them.

If R is a *commutative* ring, we simply denote the category of modules over R and morphisms between them by R-$\mathtt{Mod}$.

Proposition 8.4.12 ($\mathbf{Z}$-modules)**.** Abelian groups are $\mathbf{Z}$-modules in exactly one way. Therefore ${}_{\mathbf{Z}}\mathtt{Mod}$ and $\mathtt{Ab}$ are isomorphic categories.

Proof. Let G be an abelian group. Since $\mathbf{Z}$ is initial in the category of rings and $\mathrm{End}_{\mathtt{Ab}}(G)$ forms a ring, we find that there exists a unique morphism of rings

$$\mathbf{Z} \dashrightarrow \mathrm{End}_{\mathtt{Ab}}(G)$$

defining an action of $\mathbf{Z}$ on the group G. ♮

Example 8.4.13 ($\mathbf{Q}$-vector space)**.** Let G be an abelian group. If there exists a $\mathbf{Q}$-vector space structure on G, this structure is *unique*.

Let $\mu, \sigma\colon \mathbf{Q} \rightrightarrows \mathrm{End}_{\mathtt{Ab}}(G)$ be two Q-module structures on G. Since the inclusion $\iota\colon \mathbf{Z} \hookrightarrow \mathbf{Q}$ is an epimorphism of *rings*, it follows that, since there exists a *unique* $\mathbf{Z}$-module structure $\mathbf{Z} \to \mathrm{End}_{\mathtt{Ab}}(G)$ (see Proposition 8.4.12), it follows that $\mu\iota = \sigma\iota$. On the other hand, since ι is an epimorphism, then $\mu = \sigma$.

Proposition 8.4.14 (Zero object)**.** Let R be a ring. The trivial group 0 has a *unique* R-module structure and defines a *zero object* in the category of (left or right) R-modules.

Proof. Notice that the only R-module structure on the trivial group 0 is given by $r \cdot 0 := 0$ for all $r \in R$. Moreover, for any R-module M, we have unique R-linear morphisms $0 \to M$ mapping $0 \mapsto 0_M$ and $M \to 0$ mapping $m \mapsto 0$ for all $m \in M$. ♮

Proposition 8.4.15 (Isomorphisms)**.** Let R be a ring. A morphism of (left or right) R-modules is an *isomorphism* if and only if it is a *bijective* set-function.

Proof. We prove for right-R-modules, the proof for left-R-modules is completely analogous. Let $\phi\colon M \to N$ be an isomorphism of R-modules and $\psi\colon N \to M$ be its inverse. Since $\phi\psi = \mathrm{id}_N$, then $\mathrm{im}\,\phi = N$, that is, ϕ is surjective. On the other hand, since $\psi\phi = \mathrm{id}_M$ and ψ is a well defined set-function, it follows that ϕ is injective.

Conversely, let $\phi\colon M \to N$ be an R-linear morphism and a bijective set-function. Let $\psi\colon N \to M$ be its inverse as a set-function — we shall prove that ψ is an R-linear morphism. Let $n, n' \in N$ be any elements, if $\phi(m) = n$ and $\phi(m') = n'$ then $\phi(m+m') =$

$\phi(m) + \phi(m') = n + n'$ — thus from construction $\psi(n + n') = m + m' = \psi(n) + \psi(n')$. Now, if $r \in R$ is any ring element, then since $\phi(mr) = \phi(m)r = nr$, we find that $\psi(nr) = mr = \psi(n)r$ — therefore ψ is indeed a morphism of R-modules and hence an inverse morphism for ϕ. ♮

Example 8.4.16. Let R be an integral domain, and $(a) \subseteq R$ be a non-zero principal ideal of R. There exists a natural isomorphism of R-modules $R \simeq (a)$.

Consider the map $\phi\colon R \to (a)$ given by $r \mapsto ra$. Then for any $r, s \in R$ we have

$$\phi(r+s) = (r+s)a = ra + sa = \phi(r) + \phi(s)$$
$$\phi(rs) = (rs)a = r(sa) = r\phi(s)$$

thus ϕ is an R-module morphism. The morphism is also clearly surjective by the definition of an ideal. Moreover, one should note that $\phi(1) = a$, which is, by hypothesis, non-zero. Given any two $r, r' \in R$ such that $\phi(r) = \phi(r')$ we obtain

$$0 = \phi(r) - \phi(r') = r\phi(1) - r'\phi(1) = (r - r')\phi(1) = (r - r')a,$$

and since R is an integral domain, it follows that $r = r'$ — therefore, ϕ is injective. We conclude that ϕ is a bijection between R-modules and therefore establishes an isomorphism $R \simeq (a)$.

Example 8.4.17 (Morphisms form an R-module)**.** Given a ring R and left-R-modules M and N, the collection of R-linear morphisms $\mathrm{Mor}_{{}_R\mathtt{Mod}}(M, N)$ can be endowed with the structure of an *right-R-module*.

Since $\mathrm{Mor}({}_R\mathtt{Mod}) \subseteq \mathrm{Mor}(\mathtt{Ab})$, it follows that $\mathrm{Mor}_{{}_R\mathtt{Mod}}(M, N)$ is an abelian group given by

$$(f + g)(m) := f(m) + g(m)$$

for any morphisms $f, g\colon M \rightrightarrows N$ and element $m \in M$. Moreover, one can endow $\mathrm{Mor}_{{}_R\mathtt{Mod}}(M, N)$ with the *right* ring action $\mathbf{R} \times \mathrm{Mor}_{{}_R\mathtt{Mod}}(M, N) \to \mathrm{Mor}_{{}_R\mathtt{Mod}}(M, N)$ given by

$$(f \cdot r)(m) := f(rm)$$

for every morphism $f\colon M \to N$, and any elements $r \in R$ and $m \in M$.

Notice that we've emphasized that, in general, we can *only* give a *right* R-module structure to $\mathrm{Mor}_{{}_R\mathtt{Mod}}(M, N)$, if on the contrary we defined an left-R-action by $(r \cdot f)(m) := f(rm)$, one would suffer from the following problem:

$$((rs)f)(m) = (r(sf))(m) = (sf)(rm) = f(s(rm)) = f((sr)m) = ((sr)f)(m) \tag{8.4}$$

where $f\colon M \to N$ is a morphism, and $r, s \in R$ and $m \in M$ are any elements. Mind that Eq. (8.4) *does not yield a left-R-module structure unless R is commutative.*

Proposition 8.4.18. Let R be a *commutative* ring. Then there exists a canonical isomorphism of R-modules

$$\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(R, M) \simeq M.$$

Proof. Notice that every R-module morphism $R \to M$ has to be of the form $r \mapsto rm$ for some fixed $m \in M$ — for convenience, name this morphism f_m. We define a map $\phi\colon \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(R, M) \to M$ by sending $f_m \mapsto m$, which is certainly surjective. Moreover, if $f_m = f_n$ then in particular $m = f_m(1) = f_n(1) = n$, therefore ϕ is injective. Notice that for any $m, n \in M$ we have

$$f_m(r) + f_n(r) = rm + rn = r(m + n) = f_{m+n}(r),$$

thus $f_m + f_n = f_{m+n}$. Also, given any $s \in R$ we have

$$sf_m(r) = s(rm) = (sr)m = (rs)m = r(sm) = f_{sm}(r),$$

then $sf_m = f_{sm}$. The bijection ϕ is also an R-module morphism since, given any two $f_m, f_n \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(R, M)$ we have

$$\phi(f_m + f_n) = \phi(f_{m+n}) = m + n = \phi(f_m) + \phi(f_n),$$

and given $r \in R$,

$$\phi(rf_m) = \phi(f_{rm}) = rm = r\phi(f_m).$$

Since bijective morphisms are isomorphisms in R-$\mathtt{Mod}$, it follows that $\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(R, M) \simeq M$ via ϕ. ♮

Proposition 8.4.19. Let M and N be R-modules, for some ring R. If $M \simeq N$, then there exists a natural isomorphism between *abelian groups*

$$\mathrm{End}_{R\text{-}\mathtt{Mod}}(M) \simeq \mathrm{End}_{R\text{-}\mathtt{Mod}}(N).$$

Proof. Let $\phi\colon M \xrightarrow{\simeq} N$ be an isomorphism. We define a map $\Phi\colon \mathrm{End}_{R\text{-}\mathtt{Mod}}(M) \to \mathrm{End}_{R\text{-}\mathtt{Mod}}(N)$ given by the conjugation $f \mapsto \phi f \phi^{-1}$. This uniquely defines an R-module morphism for each endomorphism $f\colon M \to M$. Notice that, since ϕ is a bijection, given any endomorphism $g\colon N \to N$ we may define an R-morphism $f\colon M \to M$ given by $f := \phi^{-1} g \phi$, so that $\Phi(f) = \phi(\phi^{-1} g \phi)\phi^{-1} = g$. Therefore Φ establishes an isomorphism of abelian groups via conjugation. ♮

Examples on Nilpotency

Lemma 8.4.20 (Nakayama's lemma, a particular case)**.** Let R be a *commutative* ring, and $a \in R$ be a *nilpotent* element. Then $M = 0$ if and only if $aM = M$.

Proof. If $a = 0$ then the statement is true. Suppose that a is non-zero, and let $n \in \mathbf{Z}_{>0}$ be the minimal positive integer such that $a^n = 0$. If $M = 0$ then obviously $aM = a \cdot 0 = 0 = M$. On the converse, if we assume that $aM = M$, let $m \in M$ be any element. Let $m_1 \in M$ be an element such that $m = am_1$. By induction, let $m_j \in M$ be an element such that $m_{j-1} = am_j$, for $1 < j \leqslant k$. If we consider the collection $(m_j)_{j=1}^k$ we get a chain of equalities

$$m = am_1 = a^2 m_2 = \cdots = a^k m_k = 0.$$

Therefore $M = 0$. ♮

Proposition 8.4.21. Let R be a commutative ring and $\mathfrak{a}$ a *nilpotent ideal* of R. Let $\phi\colon M \to N$ be an R-module morphism. If the induced R-module morphism

$$\overline{\phi}\colon \frac{M}{\mathfrak{a}M} \longrightarrow \frac{N}{\mathfrak{a}N} \quad \text{mapping} \quad m + \mathfrak{a}M \longmapsto \phi(m) + \mathfrak{a}N$$

is a surjection, then so is ϕ.

Proof. Since $\overline{\phi}$ is surjective, it follows that

$$\overline{\phi}(M/\mathfrak{a}M) = \frac{\phi(M) + \mathfrak{a}N}{\mathfrak{a}N} = \frac{N}{\mathfrak{a}N},$$

therefore $N = \phi(M) + \mathfrak{a}N$. Suppose that $k \in \mathbf{Z}_{>0}$ is such that $\mathfrak{a}^k = 0$ (which exists since $\mathfrak{a}$ is a nilpotent ideal). Then via induction we find

$$\begin{aligned} N = \phi(M) + \mathfrak{a}N &= \phi(M) + \mathfrak{a}(\phi(M) + \mathfrak{a}N) \\ &= \phi(M) + \mathfrak{a}\phi(M) + \mathfrak{a}^2N \\ &= \phi(M) + \mathfrak{a}^2N \\ &= \dots \\ &= \phi(M) + \mathfrak{a}^kN \\ &= \phi(M). \end{aligned}$$

Therefore ϕ is surjective, since $N = \phi(M)$. ♮

Kernels & Cokernels

Lemma 8.4.22. Kernels and cokernels exist in R-`Mod`.

Proof. Let $\phi\colon M \to N$ be any R-module and consider $\ker\phi := \{m \in M : \phi(m) = 0\}$, together with the canonical inclusion $\iota\colon \ker\phi \hookrightarrow M$. Notice that $\phi\iota = 0\iota = 0$. We prove that $(\ker\phi, \iota)$ is the equalizer of $(\phi, 0)$: let P be any other R-module and $f\colon P \to M$ be any R-module morphism such that $\phi f = 0f = 0$. We define a map $\overline{f}\colon P \to \ker\phi$ given by $\overline{f} = f$. Indeed, since $\phi f = 0$, then $\operatorname{im} f \subseteq \ker\phi$, and $\overline{f}$ is a well defined and unique R-module morphism making the following diagram commute

$$\begin{tikzcd} \ker\phi \arrow[r, hook, "\iota"] & M \arrow[r, shift left, "\phi"] \arrow[r, shift right, "0"'] & N \\ & P \arrow[u, "f"] \arrow[ul, dashed, bend left, "\overline{f}"] & \end{tikzcd}$$

Thus $(\ker\phi, \iota)$ is indeed the equalizer of $(f, 0)$.

Now we prove that $\operatorname{coker}\phi := N/\operatorname{im}\phi$ together with the natural projection $\pi\colon N \twoheadrightarrow \operatorname{coker}\phi$ is the coequalizer of $(\phi, 0)$. Let C be an R-module and $g\colon N \to C$ be an R-module morphism such that $g\phi = g0 = 0$. We define a map $\overline{g}\colon \operatorname{coker}\phi \to C$ to be given by $n + \operatorname{im}\phi \mapsto g(n)$. This map is well defined since $g(n) = 0$ for all $n \in \operatorname{im}\phi$.

Also, since g is a morphism, $\overline{g}$ is trivially a morphism of R-modules too. Moreover, since π is an epimorphism, $\overline{g}$ is the unique morphism for which the following diagram commutes

$$\begin{tikzcd}
M \arrow[r, shift left, "\phi"] \arrow[r, shift right, "0"'] & N \arrow[r, two heads, "\pi"] \arrow[d, "g"] & \operatorname{coker}\phi \arrow[dl, dashed, bend left, "\overline{g}"] \\
& P &
\end{tikzcd}$$

That is, $(\operatorname{coker}\phi, \pi)$ is the coequalizer of $(\phi, 0)$. $\natural$

Proposition 8.4.23 (Properties of kernels & cokernels)**.** Kernels and cokernels exist in R-`Mod`. Let $\phi\colon M \to N$ be an R-module morphism, then:

(a) The following propositions are equivalent:

- The R-module morphism ϕ is a monomorphism.
- The kernel of ϕ is trivial, that is, $\ker\phi = 0$.
- The set-function induced by ϕ is injective.

(b) The following propositions are equivalent:

- The R-module morphism ϕ is an epimorphism.
- The cokernel of ϕ is trivial, that is, $\operatorname{coker}\phi = 0$.
- The set-function induced by ϕ is surjective.

Proof. (a)

- Suppose ϕ is a monomorphism, and consider $\iota\colon \ker\phi \hookrightarrow M$ and 0, then $\phi\iota = \phi 0 = 0$, thus $\iota = 0$ — which implies in $\ker\phi = 0$.
- If we now suppose that $\ker\phi = 0$, given any two elements $m, m' \in M$ such that $\phi(m) = \phi(m')$ we obtain $0 = \phi(m) - \phi(m') = \phi(m - m')$ then $m - m' \in \ker\phi$, implying in $m = m'$.
- Lastly, if ϕ is injective, then given any two morphisms $\alpha, \beta\colon P \rightrightarrows M$ such that $\phi\alpha = \phi\beta$, we have for all $p \in P$ that $\phi(\alpha(p)) = \phi(\beta(p))$ — which implies in $\alpha(p) = \beta(p)$, thus $\alpha = \beta$. We conclude that ϕ is a monomorphism.

(b)

- Suppose that ϕ is an epimorphism, then if we consider the canonical inclusiont $\iota\colon \operatorname{im}\phi \hookrightarrow N$ and the identity morphism $\mathrm{id}_N\colon N \to N$, one has that $\iota\phi = \mathrm{id}_N\,\phi$. Since ϕ is an epimorphism, then $\iota = \mathrm{id}_N$ — which is only possible if $\operatorname{im}\phi = N$. Therefore $\operatorname{coker}\phi$ is trivial.
- If we suppose that $\operatorname{coker}\phi$ is trivial, then $\operatorname{im}\phi = N$ and ϕ is therefore surjective.
- Suppose that ϕ is surjective, and let $\alpha, \beta\colon N \to P$ be R-module morphisms such that $\alpha\phi = \beta\phi$. Then, for every $n \in N$, there exists $m \in M$ such that $\phi(m) = n$, thus $\alpha(n) = \alpha(\phi(m)) = \beta(\phi(m)) = \beta(n)$ — therefore $\alpha = \beta$ and ϕ is an epimorphism.

$\natural$

Example 8.4.24 (Right & left inverses)**.** One should not be deceived by the ideas permeating `Set`, just as in general categories, R-`Mod` monomorphisms and epimorphisms need *not* have left-inverse and right-inverse, respectively. An example of a monomorphism without a left-inverse is $\mathbf{Z} \rightarrowtail \mathbf{Z}$ given by $a \mapsto 2a$. On the other hand, the projection $\mathbf{Z} \twoheadrightarrow \mathbf{Z}/2\mathbf{Z}$ is an epimorphism, although it does not have a right-inverse.

R-Algebras

Example 8.4.25 (R-modules from ring morphisms)**.** Let $\phi\colon R \to S$ be any ring morphism. We can induce on S a structure of left-R-module by defining a map $\rho\colon R \times S \to S$ to map $(r, s) \mapsto \phi(r)s$ for every $r \in R$ and $s \in S$. In particular, this shows that we can endow R itself with a left-R-module structure. These constructions can be analogously done for right-R-modules as well.

If R happens to be commutative and $\operatorname{im}\phi \subseteq Z(S)$, then by Example 8.4.6 we find that the left and right R-module structures on S induced by ϕ coincide. Moreover, notice that, given any $s, s' \in S$ and $r, r' \in R$, we have that

$$\begin{aligned}\rho(r,s)\rho(r',s') &= (\phi(r)s)(\phi(r')s') = \phi(r)(s\phi(r'))s' = \phi(r)(\phi(r')s)s' \\ &= (\phi(r)\phi(r'))(ss') = \phi(rr')(ss') \\ &= \rho(rr', ss').\end{aligned}$$

This shows that the R-module structure induced by ϕ is compatible with the ring structure of S. This kind of R-module receive the name of R-algebra, which we now define for later reference.

Definition 8.4.26 (R-algebra)**.** Let R be a *commutative* ring. We define an *R-algebra* to be a *ring morphism* $\phi\colon R \to S$ such that the image of ϕ is contained in the centre of S — inducing an R-module structure on S. If the ring S itself is commutative, we say that it has a structure of a *commutative R-algebra*.

We define a morphism $\gamma\colon \alpha \to \beta$ of R-algebras $\alpha\colon R \to S$ and $\beta\colon R \to Q$ to be a ring morphism $\gamma\colon S \to Q$ such that for all $r \in R$ and $s \in S$ we have $\phi(rs) = r\phi(s)$, and that the following diagram commutes in `Ring`

$$\begin{tikzcd} & R \arrow[dl, "\alpha"'] \arrow[dr, "\beta"] & \\ S \arrow[rr, "\gamma"'] & & Q \end{tikzcd}$$

We denote by R-`Alg` the category consisting of R-algebras and morphisms between them. A particularly important subcategory is that of the commutative R-algebras, which we shall denote by R-`CAlg`.

Corollary 8.4.27. Given a commutative ring R, then R-`CAlg` is a subcategory of $R/$`CRing` — where $R/$`CRing` denotes the slice category under R.

Example 8.4.28. The category R-`CAlg`, however, is *not a full subcategory* of `CRing`. Notice for instance that the map $\mathbf{C} \to \mathbf{C}$ given by the complex conjugation $z \mapsto \overline{z}$ is a *ring* automorphism, nonetheless it isn't a morphism of $\mathbf{C}$-modules.

Proposition 8.4.29 (Initial object)**.** Given a ring R, the module structure of R over itself is the *initial* object of R-`Alg`.

Proof. Given any R-algebra $\phi\colon R \to A$ we have, for all $r, r' \in R$, that $\phi(rr') = r\phi(r')$ — thus ϕ is an R-algebra morphism. In particular, for $r' = 1_R$ we obtain $\phi(r) = r\phi(1_R) = r1_A = r$ — thus ϕ is the unique R-algebra morphism $R \to A$. ♮

Definition 8.4.30 (Field extension)**.** Let k be a field. We define a *field extension* of k to be a *commutative k-algebra K* with injective ring morphisms $\mathrm{inj}, \mathrm{inv}\colon k \rightrightarrows K$ such that $\mathrm{inj}(a)\,\mathrm{inv}(a) = 1$ for all $a \in k$.

Example 8.4.31. Let k be a subfield $k \subseteq \ell$ of a field ℓ. Then ℓ has a natural structure of field extension of k. Indeed, we have a natural inclusion $k \hookrightarrow \ell$ mapping $a \mapsto a$ and an inversion map $k \hookrightarrow \ell$ sending $a \mapsto a^{-1}$.

Definition 8.4.32 (Rees algebra)**.** Let R be a *commutative* ring and $\mathfrak{a}$ ideal of R. We define a ring

$$\mathrm{Rees}_R(\mathfrak{a}) := \bigoplus_{j \geqslant 0} \mathfrak{a}^j,$$

where $\mathfrak{a}^0 := R$, with a multiplication given by

$$(a_j)_{j\geqslant 0} \cdot (b_j)_{j \geqslant 0} := \left(\sum_{i+k=j} a_i b_k \right)_{j \geqslant 0} \in \mathrm{Rees}_R(\mathfrak{a}).$$

The ring morphism $R \to \mathrm{Rees}_R(\mathfrak{a})$ mapping $r \mapsto (r, 0, \dots, 0, \dots)$ is called the *Rees algebra of* $\mathfrak{a}$.

Proposition 8.4.33. Let R be a *commutative* ring and $a \in R$ be a non-zero-divisor. There exists a natural *isomorphism of R-algebras*

$$\mathrm{Rees}_R((a)) \simeq R[x]$$

Proof. Notice that $R[x]$ is realized as an R-algebra by the inclusion $\iota\colon R \hookrightarrow R[x]$ mapping to constant polynomials. Denote, for every $j \geqslant 0$, by $e_j \in \mathrm{Rees}_R((a))$ the element whose j-th coordinate is a^j and zero elsewhere. Define a map $\phi\colon R[x] \to \mathrm{Rees}_R((a))$ by mapping $\phi(r) := re_0$ and $\phi(x) := e_1$. Since $\{R, x\}$ generate $R[x]$, this completely defines ϕ as a ring morphism, since $\phi(rx^j) = r\phi(x)^j = re_1^j = re_j$. Moreover, given any $(a_j)_{j\geqslant 0} \in \mathrm{Rees}_R((a))$, since R is commutative, we have $a_j = r_j a^j$ for some $r_j \in R$. If we let $(r_j)_{j \geqslant 0}$ be the collection of these associated R terms, we can build a polynomial $p(x) := \sum_{j \geqslant 0} r_j x^j$ so that

$$\phi(p(x)) = \sum_{j \geqslant 0} \phi(r_j x^j) = \sum_{j\geqslant 0} r_j e_j = (r_j a^j)_{j \geqslant 0} = (a_j)_{j \geqslant 0}.$$

Therefore ϕ is surjective. Moreover, it is simple to see that if $p(x) := \sum_{j\geqslant 0} b_j x^j$ and $q(x) := \sum_{j \geqslant 0} c_j x^j$ are polynomials in $R[x]$ such that $\phi(p(x)) = \phi(q(x))$, then their coefficients match — that is, $b_j = c_j$ for all $j \geqslant 0$, and then $p(x) = q(x)$, making ϕ injective. Therefore ϕ is an isomorphism of R-algebras. ♮

Proposition 8.4.34. Let R be a commutative ring, $a \in R$ be a non-zero-divisor, and $\mathfrak{b} \subseteq R$ be any ideal. Then there exists a natural isomorphism of R-algebras

$$\mathrm{Rees}_R(a\mathfrak{b}) \simeq \mathrm{Rees}_R(\mathfrak{b}).$$

Proof. Define a map $\phi\colon \mathrm{Rees}_R(\mathfrak{b}) \to \mathrm{Rees}_R(a\mathfrak{b})$ by $(b_j)_{j\geqslant 0} \mapsto (a^j b_j)_{j\geqslant 0}$ — since a is a non-zero-divisor, $a^j b_j = 0$ if and only if $b_j = 0$, therefore ϕ is bijective. Notice that ϕ satisfies $\phi(ru) = r\phi(u)$ for any $r \in R$ and $u \in \mathrm{Rees}_R(\mathfrak{b})$. Moreover for any two $(u_j)_{j\geqslant 0}, (v_j)_{j\geqslant 0} \in \mathrm{Rees}_R(\mathfrak{b})$, we have

$$\begin{aligned}
\phi((u_j)_j \cdot (v_j)_j) &= \phi\Bigg(\sum_{i+k=j} u_i v_k\Bigg)_j = \Bigg(a^j \sum_{i+k=j} u_i v_k\Bigg)_j \\
&= \Bigg(\sum_{i+k=j} a^{i+k} u_i v_k\Bigg)_j = (a^j u_j)_j \cdot (a^j v_j)_j \\
&= \phi((u_j)_j) \cdot \phi((v_j)_j).
\end{aligned}$$

Thus ϕ is an R-algebra isomorphism between $\mathrm{Rees}_R(a\mathfrak{b})$ and $\mathrm{Rees}_R(\mathfrak{b})$. ♮

Submodules & Quotients

Submodules

Definition 8.4.35 (Submodule)**.** Let R be a ring and M be an R-module. An R-module N is said to be a *submodule* of R if $N \subseteq M$ and the inclusion map $N \hookrightarrow M$ is a morphism of R-modules.

Example 8.4.36 (Ideals are the submodules of R)**.** Let R be a ring endowed with the canonical left-R-module structure. The submodules of R are correspond exactly to the left-ideals of R.

Example 8.4.37 (Kernel and image are submodules)**.** Given a morphism $\phi\colon M \to N$ of R-modules, the kernel of ϕ is a submodule of M, while the image of ϕ is a submodule of N.

Suppose M and N are right-R-modules, for left-R-modules the proof is analogous. If $m \in \ker\phi$, then for all $r \in R$ we have $\phi(mr) = \phi(m)r = 0 \cdot r = 0$ — thus $mr \in \ker\phi$, making $\ker\phi \hookrightarrow M$ a morphism of right-R-modules. Now, if $n \in \operatorname{im}\phi$, there must exist $m \in M$ such that $\phi(m) = n$, therefore for all $r \in R$ we have $\phi(mr) = \phi(m)r = nr \in \operatorname{im}\phi$. We conclude that, indeed, $\ker\phi$ and $\operatorname{im}\phi$ are submodules of M and N, respectively.

Example 8.4.38 (Intersection & sum of submodules)**.** Let M be an R-module and $(N_j)_{j\in J}$ be a collection of submodules of M. We have the following:

(a) The intersection $\bigcap_{j\in J} N_j$ is a submodule of M.

(b) The sum

$$\sum_{j\in J} N_j \coloneqq \Bigg\{ \sum_{j\in F} n_j : F \subseteq J \text{ is finite, and } n_j \in N_j \text{ for all } j \in F \Bigg\}$$

is a submodule of M.

Proof. For the intersection, given any two $a, b \in \bigcap_{j\in J} N_j$ we have, for each $j \in J$, that $a, b \in N_j$ and therefore $a + b \in N_j$ — this implies in $a + b \in \bigcap_{j\in J} N_j$. Moreover, given any element $r \in R$ it is equally clear that $ar \in N_j$ for every $j \in J$, which implies in $ar \in \bigcap_{j\in J} N_j$. Therefore $\bigcap_{j\in J} N_j$ is indeed a submodule of M.

For the ease of notation, define $N := \sum_{j\in J} N_j$. Notice that since every sum in N is finite, then $N \subseteq M$. Given any two elements $\sum_{j\in F} a_j, \sum_{j\in F'} b_j \in N$, we have that $F \cup F' \subseteq J$ is finite, hence we have an element $\sum_{j\in F\cup F'} m_j \in N$ given by

$$m_j := \begin{cases} a_j, & j \in F \setminus F', \\ b_j, & j \in F' \setminus F, \\ a_j + b_j, & j \in F \cap F'. \end{cases}$$

It is easy to see that this constructed element satisfies

$$\sum_{j\in F\cup F'} m_j = \Big(\sum_{j\in F} a_j\Big) + \Big(\sum_{j\in F'} b_j\Big),$$

therefore N is closed under finite addition. Since multiplication is distributive over finite sums, it follows that for any $r \in R$ we have

$$\Big(\sum_{j\in F} a_j\Big) r = \sum_{j\in F} a_j r$$

and since $a_j r \in N_j$ for each $j \in F$, it follows that $\big(\sum_{j\in F} a_j\big) r \in N$. Thus N is a submodule of M. ♮

Example 8.4.39 (Union of submodules)**.** Let M be an R-module, for a ring R. The following are two propositions concerning submodules of M:

1. Given submodules $S, T \subseteq M$, the union $S \cup T$ is a submodule of M if and only if $S \subseteq T$ or $T \subseteq S$.
2. Let $(N_j)_{j\in \mathbf{N}}$ be an ascending chain of submodules of M—that is, $N_j \subseteq N_{j+1}$ for all $j \in \mathbf{N}$. Then the union $\bigcup_{j\in\mathbf{N}} N_j$ is a submodule of M.

Proof. For item (a), if $S \cup T$ is a submodule of M, let $s \in S$ and $t \in T$ be any two elements, then $s + t \in S \cup T$. This implies that either $s + t \in S$ (which would imply in $t \in S$) or $s + t \in T$ (which would imply in $s \in T$), therefore either $S \subseteq T$ or $T \subseteq S$. For the converse, suppose, without loss of generality, that $S \subseteq T$, then $S \cup T = S$, therefore $S \cup T$ is a submodule of M.

We now prove item (b). For the sake of notation, let $N := \bigcup_{j\in\mathbf{N}} N_j$. If $n, n' \in N$ are any two elements, there must exist indices $j, j' \in \mathbf{N}$ such that $n \in N_i$ and $n' \in N_{i'}$ for all $i > j$ and for all $i' > j'$. Define $k := \max(j, j')$, then $n, n' \in N_i$ for all $i > k$—implying in $n + n' \in N_i$ and $rn \in N_i$ for any $r \in R$. Therefore $n + n' \in N$ and $rn \in N$, hence N is a submodule of M. ♮

Example 8.4.40 (Totally ordered submodules)**.** Let M be a finite $\mathbf{Z}$-module such that the collection of its submodules is totally ordered with respect to inclusion. Then there exists a prime p such that the number of elements of M is a power of p.

Proof. Let $|M| \coloneqq d$ and suppose there exists two primes p and q dividing d. By Proposition 7.4.29 we know that there must exist $m, n \in M$ with order p and q, respectively. We now consider the submodules $\langle m \rangle$ and $\langle n \rangle$ of M. Since M has a totally ordered set of submodules, we may assume without loss of generality that $\langle m \rangle \subseteq \langle n \rangle$—therefore there exists $a \in \mathbf{Z}$ such that $m = an$. Notice that $pm = pan = 0$, therefore pa must be a divisor of q—but since q is prime, pa is either 1 or q. Since p is also prime, it must be the case that $a = 1$ and $p = q$. Thus there exists a unique prime divisor of d—hence $d = p^\alpha$ for some $\alpha \in \mathbf{Z}$. ♮

Simple R-Modules

Definition 8.4.41 (Simple module)**.** Let R be a ring. An R-module M is said to be *simple* if its only submodules are $\{0\}$ and M itself.

Lemma 8.4.42 (Schur's)**.** Let M and N be *simple* R-modules. If $\phi \colon M \to N$ is an R-module morphism, then either $\phi = 0$ or ϕ is an *isomorphism*.

Proof. Since $\ker \phi$ is a submodule of M, it can either be 0 or M. If $\ker \phi = M$, then $\phi = 0$. On the other hand, if $\ker \phi = 0$ then ϕ is an injective morphism. Moreover, since $\operatorname{im} \phi$ is a submodule of N, it's either 0 or N — since $\ker \phi$ in trivial, then the only possibility is that $\operatorname{im} \phi = N$, thus ϕ is surjective. ♮

Corollary 8.4.43. Let M and N be right-R-modules and consider $\operatorname{Mor}_{\mathtt{Mod}_R}(M, N)$ as a *ring*. If M is *simple*, then $\operatorname{Mor}_{\mathtt{Mod}_R}(M, N)$ is a *division ring*.

Proof. Let $\mathfrak{a}$ be a left-ideal (or right-ideal) of $\operatorname{Mor}_{\mathtt{Mod}_R}(M, N)$. If there exists an isomorphism $\phi \in \mathfrak{a}$, then $\phi^{-1}\phi = \mathrm{id}_M \in \mathfrak{a}$ and therefore, for all elements $\psi \in \operatorname{Mor}_{\mathtt{Mod}_R}(M, N)$ we have $\psi \, \mathrm{id}_M = \psi \in \mathfrak{a}$ — thus $\mathfrak{a} = \operatorname{Mor}_{\mathtt{Mod}_R}(M, N)$ (the same equivalent arguement can be used for right-ideals, where instead we get $\phi\phi^{-1} = \mathrm{id}_N$ in the ideal). If there exist no isomorphism in $\mathfrak{a}$, by Lemma 8.4.42 it is only composed of the zero morphism, thus $\mathfrak{a} = 0$. From Proposition 8.3.14 we conclude that $\operatorname{Mor}_{\mathtt{Mod}_R}(M, N)$ is a division ring. ♮

Example 8.4.44. We now show a counterexample ilustrating why the opposite of Corollary 8.4.43 does not hold in general. Let k be a field and

$$A \coloneqq \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad\qquad B \coloneqq \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

be matrices. Define R to be the k-algebra generated by A and B. Consider the left-R-module $M = k^2$, with left-multiplication by matrices. We'll show that $\operatorname{Mor}_{{}_R\mathtt{Mod}}(M, M)$ is a division ring, while M is not simple.

Consider the principal ideal $\mathfrak{a} \coloneqq \left(\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]\right)$ of M. Notice that for any $a, b \in k$ we have

$$(aA + bB) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a & a \\ 0 & b \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ 0 \end{bmatrix},$$

thus $\mathfrak{a} \neq M$ and $\mathfrak{a} \neq 0$, and M isn't simple.

Let $\mathfrak{h}$ be a left-ideal (or right-ideal) of the ring $\mathrm{Mor}_{_R\mathtt{Mod}}(M, M)$. If $\mathfrak{h} \neq 0$, consider a non-zero morphism $M = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathfrak{h}$. Since M is a left-R-module morphism, it must be the case that, for any $x = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in M$, we have $M(Ax) = AM(x)$ and $M(Bx) = BM(x)$, but

$$M(Ax) = \begin{bmatrix} a(x_1 + x_2) \\ c(x_1 + x_2) \end{bmatrix} \tag{8.5}$$

$$M(Bx) = \begin{bmatrix} bx_2 \\ dx_2 \end{bmatrix} \tag{8.6}$$

while on the other hand we have

$$AM(x) = \begin{bmatrix} (a + c)x_1 + (b + d)x_2 \\ 0 \end{bmatrix} \tag{8.7}$$

$$BM(x) = \begin{bmatrix} 0 \\ cx_1 + dx_2 \end{bmatrix} \tag{8.8}$$

Since Eq. (8.5) equals Eq. (8.7), while Eq. (8.6) equals Eq. (8.8) — and since M is non-zero by hypothesis — we obtain a solution $a = d = 1$ and $b = c = 0$, yielding $M = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, which implies in $\mathfrak{h} = \mathrm{Mor}_{_R\mathtt{Mod}}(M, M)$.

Proposition 8.4.45. Let M be an R-module. Then M is simple if and only if $M \simeq R/\mathfrak{m}$ for a maximal ideal $\mathfrak{m}$ of R.

Proof. Suppose $\phi\colon M \xrightarrow{\simeq} R/\mathfrak{m}$ is an isomorphism of R-modules. If $N \subsetneq M$ is a *proper* submodule, then $\phi(N) \subseteq R/\mathfrak{m}$ must also be a submodule of $R/\mathfrak{m}$—that is, an ideal. Since $\mathfrak{m}$ is maximal, then the only ideals of the field $R/\mathfrak{m}$ are either the field itself or the zero ideal $\mathfrak{m}$. Since ϕ is an isomorphism and N is proper, it must be the case that $\phi(N) = \mathfrak{m}$. But since $\phi(0) = \mathfrak{m}$, then $N = \{0\}$—thus M is simple.

For the converse, assume that M is simple, so that every non-zero element of M generates the whole module. In particular $1 \in M$ generates M, therefore the morphism of R-modules $\psi\colon R \to M$ mapping $a \mapsto a \cdot 1$ is surjective, therefore

$$M \simeq R/\ker\psi.$$

By Corollary 8.4.53 if $\mathfrak{a} \subsetneq R$ is any proper ideal (submodule) containing $\ker\psi$, then its corresponding submodule is $\mathfrak{a}/\ker\psi \subseteq R/\ker\psi \simeq M$. Since M is simple and $\mathfrak{a}$ is proper, it follows that $\mathfrak{a}/\ker\psi$ must be the zero ideal, therefore $\mathfrak{a} = \ker\psi$—that is $\ker\psi$ is maximal. ♮

Quotient modules

Definition 8.4.46 (Quotient module)**.** Let R be a ring, and M be a left-R-module. If $N \subseteq M$ is a submodule, then in particular N is a *normal* subgroup of M and therefore M/N is an *abelian group*. We endow the group M/N with a left-R-action $R \times (M/N) \to M/N$ given by

$$r(m + N) := rm + N,$$

for all $r \in R$ and $m \in M$. This action turns M/N into a left-R-module and the canonical projection $\pi\colon M \twoheadrightarrow M/N$ into an R-module morphism. Therefore, the submodule N is the *kernel* of the canonical projection π.

Example 8.4.47. Let R be a *non-commutative* ring and $\mathfrak{a}$ a left-submodule, then the *set* of equivalence classes $R/\mathfrak{a}$ is *not a ring* — since $\mathfrak{a}$ is be required to be a two-sided-ideal for the quotient to be a ring. Although not a ring, $R/\mathfrak{a}$ is is an abelian group and one can endow $R/\mathfrak{a}$ with the canonical left-R-action given by $r(a + \mathfrak{a}) = ra + \mathfrak{a}$ for every $r \in R$ and $a + \mathfrak{a} \in R/\mathfrak{a}$.

Theorem 8.4.48 (Universal property of quotient modules)**.** Let R be a ring and N be a submodule of an R-module M. For every module Z together with a morphism $\psi\colon M \to Z$ of R-modules such that $N \subseteq \ker\psi$, there exists a *unique* morphism of rings $\phi\colon M/N \to Z$ for which the following diagram commutes

$$\begin{tikzcd} M \arrow[r, "\psi"] \arrow[d, two heads] & Z \\ M/N \arrow[ur, dashed, bend right, "\phi"'] & \end{tikzcd}$$

Proof. In particular, the existence and uniqueness of ϕ as a *group morphism* is shown in Proposition 7.4.16 — we only show that ϕ is a morphism of R-modules. Let's assume that we are working with right-R-modules, the same analogous proof would work for the left modules. Let $m + N \in M/N$ and $r \in R$ be any elements, then $\psi(mr) = \psi(m)r$ and since $\psi = \phi\pi$ it follows that

$$\phi((m + N)r) = \phi(mr + N) = \psi(mr) = \psi(m)r = \phi(m + N)r,$$

which shows that ϕ is indeed a morphism of right-R-modules. ♮

Theorem 8.4.49 (Factorization of morphisms)**.** Every R-module morphism $\phi\colon M \to N$ factors as follows

$$\begin{tikzcd} M \arrow[r, "\phi"] \arrow[d, two heads] & N \\ M/\ker\phi \arrow[r, "\simeq", "\overline{\phi}"'] & \operatorname{im}\phi \arrow[u, hook] \end{tikzcd}$$

Proof. From the universal property Theorem 8.4.48 we find that ϕ induces a unique morphism $\overline{\phi}\colon M/\ker\phi \to N$ — being defined as $\overline{\phi}(a + \ker\phi) = \phi(a)$ for any $a + \ker\phi \in M/\ker\phi$. We simply restrict the codomain of $\overline{\phi}$ to $\operatorname{im}\phi$ so that it becomes surjective. Moreover, given any two classes $a + \ker\phi, b + \ker\phi \in M/\ker\phi$, such that $\overline{\phi}(a + \ker\phi) = \overline{\phi}(b + \ker\phi)$ then $\phi(a) = \phi(b)$ and thus $a - b \in \ker\phi$ — therefore $a + \ker\phi = b + \ker\phi$ and $\overline{\phi}$ is injective. ♮

Corollary 8.4.50 (First isomorphism)**.** Let $\phi\colon M \twoheadrightarrow N$ be a surjective R-module morphism. There exists a canonical isomorphism of R-modules

$$N \simeq M/\ker\phi$$

Proof. Since ϕ is surjective, $\operatorname{im}\phi = N$ and from Theorem 8.4.49 we conclude that the induced morphism $\overline{\phi}\colon M/\ker\phi \xrightarrow{\simeq} N$ establishes the wanted canonical isomorphism. ♮

Proposition 8.4.51 (Second isomorphism)**.** Let M be an R-module and $P, N \subseteq M$ be submodules. There exists a canonical isomorphism

$$\frac{N+P}{N} \simeq \frac{P}{N \cap P}$$

Proof. Since $N \hookrightarrow N + P$ mapping $n \mapsto n + 0 = n$ is clearly an R-module morphism, N is a submodule of $N + P$. Moreover, the inclusion $N \cap P \hookrightarrow P$ mapping $p \mapsto p$ is also a morphism since $p \in P$ for all $p \in N \cap P$.

Consider the map $\phi\colon N + P \to \frac{P}{N \cap P}$ given by $n + p \mapsto p + N \cap P$, which is clearly an R-module morphism. Given any class $p + N \cap P \in \frac{P}{N\cap P}$, one can choose a representative $p \in P$ so that $\phi(p) = p + N \cap P$ — thus ϕ is surjective. Moreover, $\phi(n + p) = N \cap P$ if, and only if $p \in N \cap P$, which in this case implies in $n + p \in N$. Therefore $\ker\phi = N$. By Corollary 8.4.50 we obtain $\frac{N+P}{\ker\phi} \simeq \frac{P}{N\cap P}$, which is the required isomorphism. ♮

Proposition 8.4.52 (Third isomorphism)**.** Let M be an R-module and $N \subseteq M$ be a submodule. If $P \subseteq M$ is a submodule *containing* N — then P/N is a *submodule* of M/N, and there exists a canonical isomorphism

$$\frac{M/N}{P/N} \simeq M/P$$

Proof. Notice that since $N \hookrightarrow P$ is clearly an R-module morphism, thus N is a submodule of P. Moreover, if we consider the inclusion $i\colon P/N \hookrightarrow M/N$ by mapping $p + N \mapsto p + N$, since $P \subseteq M$ this is well defined and is a morphism of groups — furthermore, given any $r \in R$, one has

$$i((p+N)r) = i(pr + N) = pr + N = (p+N)r = i(p+N)r$$

for any $p + N \in P/N$, thus i is an R-module morphism. From the last inclusion we conclude that P/N is a submodule of M/N.

Let $\phi\colon M/N \to M/P$ be defined by mapping $m + N \mapsto m + P$ — since $N \subseteq P$, this map is well defined and and surjective. Moreover, for any $p + N \in M/N$ where $p \in P$, we have $\phi(p + N) = p + P = P$. Moreover, if $m + N \in \ker\phi$ then necessarily $m \in P$. Therefore $\ker\phi = P/N$ and by Corollary 8.4.50 ϕ induces an isomorphism $\frac{M/N}{\ker\phi} \simeq M/P$. ♮

Corollary 8.4.53 (Submodule correspondence)**.** Let M be an R-module and $N \subseteq M$ be a submodule. There exists a bijection

$$\begin{aligned} \{\text{submodule } P \text{ of } M \text{ containing } N\} &\longrightarrow \{\text{submodules of } M/N\} \\ P &\longmapsto P/N \end{aligned}$$

Moreover, given submodules $P, P' \subseteq M$, we have that $N \subseteq P \subseteq P'$ if and only if $P/N \subseteq P'/N$ in M/N—that is, the bijection preserves inclusions.

Proposition 8.4.54. Let R be a *commutative* ring, and $\mathfrak{a}, \mathfrak{b} \subseteq R$ be ideals. Then, there exists a natural *R-module isomorphism*

$$\mathfrak{a} \cdot \frac{R}{\mathfrak{b}} \simeq \frac{\mathfrak{a} + \mathfrak{b}}{\mathfrak{b}}$$

Proof. Define a map $\phi\colon \mathfrak{a} + \mathfrak{b} \to \mathfrak{a} \cdot \frac{R}{\mathfrak{b}}$ given by $\phi(a + b) = a + \mathfrak{b}$ for every $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$. Clearly such a map establishes an R-module morphism. Moreover, if $p := \sum_{j=1}^{n} a_j r_j + \mathfrak{b} \in \mathfrak{a} \cdot \frac{R}{\mathfrak{b}}$ is any element, since R is commutative then $\mathfrak{a}$ is a two-sided ideal, hence one can choose, for every $1 \leqslant j \leqslant n$, an element $a'_j \in \mathfrak{a}$ such that $a_j r_j = a'_j$. Therefore if we consider $q := \sum_{j=1}^{n} a'_j \in \mathfrak{a} + \mathfrak{b}$, one has $\phi(q) = p$ — thus ϕ is surjective. Furthermore, an element $a + b \in \mathfrak{a} + \mathfrak{b}$ belongs to $\ker \phi$ if and only if $a + b \in \mathfrak{b}$ — therefore $\ker \phi = \mathfrak{b}$. By Corollary 8.4.50 we obtain an isomorphism $\frac{\mathfrak{a}+\mathfrak{b}}{\ker \phi} \simeq \mathfrak{a} \cdot \frac{R}{\mathfrak{b}}$ as wanted. ♮

Products & Coproducts

Definition 8.4.55 (Direct product)**.** Let R be a ring and $(M_j)_{j \in J}$ be a collection of R-modules (either right or left modules). We define the *direct product* of $(M_j)_{j \in J}$ as the set

$$\prod_{j \in J} M_j := \{(m_j)_{j \in J} : m_j \in M_j\},$$

together with coordinate-wise addition and multiplication by elements of R. This structure comes naturally with canonical projections $\pi_i\colon \prod_{j \in J} M_j \twoheadrightarrow M_i$ mapping $(m_j)_{j \in J} \mapsto m_i$ for all $i \in J$.

Definition 8.4.56 (Direct sum)**.** Let R be a ring and $(M_j)_{j \in J}$ be a collection of R-modules (either right or left modules). We define the direct product of this family to be a module

$$\bigoplus_{j \in J} M_j := \{(m_j) : m_j \in M_j \text{ and } m_j \neq 0 \text{ for finitely many } j \in J\},$$

with a natural component-wise addition, and multiplication by elements of R. Such a structure also naturally induces canonical inclusions $\iota_i\colon M_i \hookrightarrow \bigoplus_{j \in J} M_j$ given by $m \mapsto (m_j)_{j \in J}$ where $m_i = m$ and $m_j = 0$ for $j \neq i$.

Theorem 8.4.57. In the category of modules over a given ring, *direct products* are *products* and *direct sums* are *coproducts*.

Proof. Let $(M_j)_{j \in J}$ be a collection of R-modules (either left or right).

- (Product) Define $M := \prod_{j \in J} M_j$. Given an R-module N and a family of morphisms $(\phi_j\colon N \to M_j)_{j \in J}$, we define a map $\phi\colon N \to M$ to be given by $n \mapsto (\phi_j(n))_{j \in J}$. Notice that since each ϕ_j is a morphism, then given any $n, n' \in N$ we have

$$\begin{aligned}\phi(n + n') &= (\phi_j(n + n'))_{j \in J} = (\phi_j(n) + \phi_j(n'))_{j \in J} = (\phi_j(n))_{j \in J} + (\phi_j(n'))_{j \in J} \\ &= \phi(n) + \phi(n'),\end{aligned}$$

moreover, if $r \in R$ is any element then

$$\phi(rn) = (\phi_j(rn))_{j\in J} = (r\phi_j(n))_{j\in J} = r(\phi_j(n))_{j\in J} = r\phi(n).$$

That is, ϕ is an R-module morphism and clearly $\pi_j\phi = \phi_j$ for all $j \in J$. Moreover, uniqueness comes from the fact that the natural projections are epimorphisms.

- (Coproduct) Define $M := \bigoplus_{j\in J} M_j$. Given an R-module N and a family of morphisms $(\psi_j : M_j \to N)_{j\in J}$, we define a map $\psi: M \to N$ by $(m_j)_{j\in J} \mapsto \sum_{j\in J} \psi_j(m_j)$ — which is well defined since $m_j \neq 0$ only for finitely many $j \in J$ and thus $\sum_{j\in J} \psi_j(m_j)$ constitutes only of finitely many terms, since $\psi_j(0) = 0$ for any $j \in J$. This map clearly defines a morphism of R-modules, and since ι_j are monomorphisms, ψ is the unique morphism such that $\psi\iota_j = \psi_j$.

♮

Proposition 8.4.58. Let $(M_j)_{j\in J}$ be a collection of R-modules and $(N_j)_{j\in J}$ be a corresponding collection of submodules $N_j \subseteq M_j$. Then $\bigoplus_{j\in J} N_j$ is naturally identified as a submodule of $\bigoplus_{j\in J} M_j$, and there exists a natural isomorphism

$$\frac{\bigoplus_{j\in J} M_j}{\bigoplus_{j\in J} N_j} \simeq \bigoplus_{j\in J} M_j/N_j.$$

Proof. For the sake of notation, define $M := \bigoplus_{j\in J} M_j$ and $N := \bigoplus_{j\in J} N_j$. Since the inclusion $N \hookrightarrow M$ is an R-module morphism, it follows that N is a submodule of M. Consider the natural projections $(\pi_j : M_j \twoheadrightarrow M_j/N_j)_{j\in J}$, with kernels $\ker \pi_j = N_j$. Notice that the map $\pi: M \to \bigoplus_{j\in J} M_j/N_j$ given by $m \mapsto (\pi_j(m))_{j\in J}$ has a kernel $\ker\pi = N$, and is naturally surjective from its construction. Therefore by the universal property of quotients we find $M/\ker\pi = M/N \simeq \bigoplus_{j\in J} M_j/N_j$. ♮

Proposition 8.4.59 (Internal sum)**.** Given a ring R, let M be an R-module (either right or left module) and $(N_j)_{j\in J}$ be a collection of submodules of M. The following propositions are equivalent:

(a) The map $\bigoplus_{j\in J} N_j \to \sum_{j\in J} N_j$ given by $x \mapsto \sum_{j\in J} \pi_j(x)$ is a *unique isomorphism* of R-modules.

(b) For every $i \in J$, we have $N_i \cap \sum_{j\in J\setminus i} N_j = 0$.

(c) If $\sum_{j\in J} x_j \in \sum_{j\in J} N_j$ is *zero*, then $x_j = 0$ for all $j \in J$.

(d) Every element $x \in \sum_{j\in J} N_j$ can be *uniquely* expressed as a sum $x = \sum_{j\in J} x_j$ for $x_j \in N_j$ and $x_j \neq 0$ for only finitely many $j \in J$.

Proof.
- (a) ⇒ (b). Let $n_i \in N_i \cap \sum_{j\in J\setminus i} N_j$ be any element, then one can express n_i as $n_i = \sum_{j\in J\setminus i} a_j n_j$ for, therefore $n_i - \sum_{j\in J\setminus i} a_j n_j = 0$. By (a) we find that such an element has preimage $0 \in \bigoplus_{j\in J} N_j$, which can only be the case if $n_i = 0$.
- (b) ⇒ (c). If $\sum_{j\in J} x_j = 0$, then for every $i \in J$ we have $x_i = -\sum_{j\in J\setminus i} x_j$ — but assuming (b) we conclude that $x_i = 0$.

- (c) $\Rightarrow$ (d). The condition of finitely many non-zero terms is already satisfied from the definition of $\sum_{j\in J} N_j$. Now, if $x = \sum_{j\in J} x_j = \sum_{j\in J} y_j$ are two representations of the same element $x \in \sum_{j\in J} N_j$, then $\sum_{j\in J}(x_j - y_j) = 0$. By (c) this implies in $x_j = y_j$ for all $j \in J$.
- (d) $\Rightarrow$ (a). Unicity of expression implies that the given map is injective and unique. Moreover, the map is clearly surjective — therefore it establishes a unique isomorphism.

$\natural$

Proposition 8.4.60. Let R be a ring, and consider a right-R-module M, and a collection $(N_j)_{j\in J}$ of right-R-modules. The following propositions hold:

(a) There exists a canonical isomorphism of *abelian groups*

$$\mathrm{Mor}_{\mathtt{Mod}_R}\left(\bigoplus_{j\in J} N_j, M\right) \simeq \prod_{j\in J} \mathrm{Mor}_{\mathtt{Mod}_R}(N_j, M).$$

(b) There exists a canonical isomorphism of *abelian groups*

$$\mathrm{Mor}_{\mathtt{Mod}_R}\left(M, \prod_{j\in J} N_j\right) \simeq \prod_{j\in J} \mathrm{Mor}_{\mathtt{Mod}_R}(M, N_j).$$

(c) If $(M_i)_{i=1}^n$ and $(N_j)_{j=1}^m$ are *finite* collections of right-R-modules, then there exists a canonical isomorphism of *abelian groups*

$$\mathrm{Mor}_{\mathtt{Mod}_R}\left(\bigoplus_{i=1}^n M_i, \bigoplus_{j=1}^m N_j\right) \simeq \bigoplus_{i=1}^n \bigoplus_{j=1}^m \mathrm{Mor}_{\mathtt{Mod}_R}(M_i, N_j).$$

Proof. (a) Define a map $\Phi\colon \mathrm{Mor}_{\mathtt{Mod}_R}(\bigoplus_{j\in J} N_j, M) \to \prod_{j\in J} \mathrm{Mor}_{\mathtt{Mod}_R}(N_j, M)$ by sending $f \mapsto (f\iota_j)_{j\in J}$, where $\iota_j\colon N_j \hookrightarrow \bigoplus_{j\in J} N_j$ is the canonical inclusion. Notice that Φ is indeed a group morphism. Also, morphisms with equal image must agree on every element $\iota_j(n_j)$ for any $j \in J$ and $n_j \in N_j$, therefore, since these elements generate $\bigoplus_{j\in J} N_j$, it follows that the morphisms need to be equal — thus Φ is injective. Moreover, if $(g_j\colon N_j \to M)_{j\in J}$ is any collection of R-linear maps, then by the coproduct universal property there exists a unique $f\colon \bigoplus_{j\in J} N_j \to M$ such that $f\iota_j = g_j$ for all $j \in J$.

(b) Define a map $\Psi\colon \mathrm{Mor}_{\mathtt{Mod}_R}(M, \prod_{j\in J} N_j) \to \prod_{j\in J} \mathrm{Mor}_{\mathtt{Mod}_R}(M, N_j)$ by $f \mapsto (\pi_j f)$ — again this product preserves additive structure. Also, this map is injective since, given $f, g\colon M \rightrightarrows \prod_{j\in J} N_j$ such that $\Psi(f) = \Psi(g)$, we must have $\pi_j f(m) = \pi_j g(m)$ for every $j \in J$ and $m \in M$, therefore $f(m) = g(m)$ and $f = g$. Surjectivity comes from the product universal property: given a collection $(g_j\colon M \to N_j)_{j\in J}$ of morphisms, there exists a unique morphism $f\colon M \to \prod_{j\in J} N_j$ such that $\pi_j f = g_j$ for all $j \in J$.

(c) For finite indexing sets direct sums and products are isomorphic, thus the isomorphism follows directly from the last two items.

♮

Example 8.4.61. Let $M_1, \dots, M_n$ be right-R-modules and define a right-R-module $M := \bigoplus_{j=1}^n M_j$ and a ring

$$H := \left\{ \begin{bmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & \ddots & \vdots \\ f_{n1} & \cdots & f_{nn} \end{bmatrix} : f_{ij} \in \mathrm{Mor}_{\mathtt{Mod}_R}(M_j, M_i) \right\}.$$

Then there exists a natural isomorphism of rings

$$H \simeq \mathrm{End}_{\mathtt{Mod}_R}(M).$$

First we show that H is indeed a ring. Given any $[f_{ij}], [g_{ij}] \in H$, we have entrywise $f_{ij} + g_{ij} \in \mathrm{Mor}_{\mathtt{Mod}_R}(M_j, M_i)$, thus $[f_{ij} + g_{ij}] \in H$, and H is an abelian group via matrix addition. Now, if we define a product on H as the matrix product $[f_{ij}] \cdot [g_{ij}] := [\sum_{k=1}^n f_{ik} g_{kj}]$, then since $f_{ik} g_{kj} \colon M_j \to M_i$ thus the matrix product $[f_{ij}] \cdot [g_{ij}]$ is also an element of H—which shows that H is indeed a ring.

If $[f_{ij}] \in H$ is any matrix, define a morphism of right-R-modules $f \colon M \to M$ with projections $(f_i)_{i=1}^n$, where for every $1 \leqslant i \leqslant n$ the map $f_i \colon M \to M_i$ is the unique morphism such that the diagram

$$\begin{tikzcd} M_j \arrow[r, hook] \arrow[dr, bend right, "f_{ij}"'] & M \arrow[d, "f_i"] \\ & M_i \end{tikzcd}$$

commutes for all $1 \leqslant j \leqslant n$. We now simply define a map $\Phi \colon H \to \mathrm{End}_{\mathtt{Mod}_R}(M)$ by $[f_{ij}] \mapsto f$.

We now show that Φ is a ring morphism. Let $[g_{ij}] \in H$ be any other matrix and let $g := \Phi([g_{ij}]) \colon M \to M$. If $[h_{ij}] := [f_{ij}] \cdot [g_{ij}]$ then from definition we have $h_{ij} = \sum_{k=1}^n f_{ik} g_{kj}$. Notice that the composition $fg \colon M \to M$ with projections

$$(fg)_i = f_i\left(\sum_{k=1}^n g_k\right) = f_i\left(\sum_{k=1}^n \sum_{j=1}^n g_{kj}\right) = \sum_{j=1}^n \sum_{k=1}^n f_{ik} g_{kj}$$

for all $1 \leqslant i \leqslant n$, therefore $(fg)_i = h_i$. This proves that ϕ is a ring morphism:

$$\phi([f_{ij}] \cdot [g_{ij}]) = h = fg = \phi([f_{ij}])\phi([g_{ij}]).$$

For the injectivity of ϕ it suffices to see that the image of an element of H is uniquely defined by the universal property of the coproduct. For the surjectivity of ϕ, let $f \colon M \to M$ be any endomorphism, then again from the universal property each of its projections f_i are uniquely defined by a family of morphisms of R-modules $(f_{ij} \colon M_j \to M_i)_{j=1}^n$, therefore, the matrix $[f_{ij}] \in H$ has image f under ϕ. Therefore we conclude that $H \simeq \mathrm{End}_{\mathtt{Mod}_R}(M)$.

Example 8.4.62. Let R be a ring and consider R as a *right*-R-module over itself. We'll show that there exists a natural isomorphism of rings

$$M_n(R) \simeq \mathrm{End}_{\mathtt{Mod}_R}\Big(\bigoplus_{j=1}^{n} R\Big).$$

First, we shall show that $R \simeq \mathrm{End}_{\mathtt{Mod}_R}(R)$ as rings. For that end, define a map $\Psi\colon R \to \mathrm{End}_{\mathtt{Mod}_R}(R)$ given by $r \mapsto {}_r m$, where ${}_r m\colon R \to R$ is the *left*-multiplication by r, that is ${}_r m(a) = ra$. It is easy to notice that if $r, s \in R$, then ${}_{rs} m = {}_r m {}_s m$, therefore Ψ is a ring morphism. For the injectivity of Ψ, if $r, s \in R$ are elements with equal image, then in particular $r = {}_r m(1) = {}_s m(1) = s$. For surjectivity, if $f\colon R \to R$ is an endomorphism of right-R-modules, then given $r \in R$ we have $f(r) = f(1 \cdot r) = f(1)r$, therefore $f = {}_{f(1)} m$—and $\Psi(f(1)) = f$. We thus conclude that

$$R \simeq \mathrm{End}_{\mathtt{Mod}_R}(R).$$

As in Example 8.4.61, we can define a ring H of $n \times n$ matrices whose entries are elements of the ring $\mathrm{End}_{\mathtt{Mod}_R}(R)$. From our previous result, given a matrix $[f_{ij}] \in H$ there exists a unique matrix $[r_{ij}] \in M_n(R)$ such that $\Psi(r_{ij}) = f_{ij}$ for all $1 \leqslant i, j \leqslant n$. Therefore, Ψ induces an isomorphism $H \simeq M_n(R)$. From Example 8.4.61 we know that $H \simeq \mathrm{End}_{\mathtt{Mod}_R}(\bigoplus_{j=1}^n R)$, therefore we conclude that $M_n(R) \simeq \mathrm{End}_{\mathtt{Mod}_R}(\bigoplus_{j=1}^n R)$.

As a corollary of this result, together with Corollary 8.4.43, we find that if S is a *simple right-R-module*, then there exists a natural isomorphism of rings

$$M_n(D) \simeq \mathrm{End}_{\mathtt{Mod}_R}\Big(\bigoplus_{j=1}^{n} S\Big),$$

where D is a division ring—in particular, we have $D \simeq \mathrm{End}_{\mathtt{Mod}_R}(S)$.

Pullbacks & Pushouts of R-Modules

Proposition 8.4.63 (Pullback)**.** Let M, N, and Z be R-modules. Given R-module morphisms $\mu\colon M \to Z$ and $\lambda\colon N \to Z$, we define a triple $(M \times_Z N, \pi_M, \pi_N)$ — where

$$M \times_Z N := \{(m, n) \in M \times N : \mu(m) = \lambda(n)\},$$

with the natural R-module structures inherited from $M \times N$, and $\pi_M\colon M \times_Z N \twoheadrightarrow M$ and $\pi_N\colon M \times_Z N \twoheadrightarrow N$ are canonical projections. We claim that the following commutative square is a pullback

$$\begin{tikzcd}
M \times_Z N \arrow[r, two heads, "\pi_N"] \arrow[d, two heads, "\pi_M"'] \arrow[dr, phantom, "\lrcorner", very near start] & N \arrow[d, "\lambda"] \\
M \arrow[r, "\mu"'] & Z
\end{tikzcd}$$

Proof. From construction we have $\mu\pi_M = \lambda\pi_N$, thus the diagram commutes. Let W be any other R-module, and consider any two R-module morphisms $m\colon W \to M$ and $n\colon W \to N$ such that $\mu m = \lambda n$. We define a map $\phi\colon W \to M \times_Z N$ to be given by $\phi(w) = (m(w), n(w))$ — since $\mu m(w) = \mu n(w)$ then indeed $\phi(w) \in M \times_Z N$. Moreover, since m and n are morphisms, it follows trivially that ϕ is an R-module morphism. Uniqueness of ϕ comes from the fact that both π_M and π_N are epimorphisms. ♮

Proposition 8.4.64 (Pushout)**.** Let M, N, and P be three R-modules, and consider morphisms $n\colon M \to N$ and $p\colon M \to P$. Consider the following:

- Define two R-modules:

$$L := \{(n(m), -p(m)) : m \in M\} \subseteq N \oplus P \quad \text{and} \quad Q := \frac{N \oplus P}{L}$$

- Define morphisms $\varepsilon_N\colon N \to Q$ and $\varepsilon_P\colon P \to Q$ given by

$$\varepsilon_N := \pi\iota_N \quad \text{and} \quad \varepsilon_P := \pi\iota_P.$$

  Where $\pi\colon N \oplus P \twoheadrightarrow Q$ is the canonical projection and ι_N and ι_P are the canonical inclusions of N and P, respectively, into $N \oplus P$.

Then the triple $(Q, \varepsilon_N, \varepsilon_P)$ is the *pushout* of the pair (n, p), that is

$$\begin{tikzcd} M \arrow[r, "p"] \arrow[d, "n"'] & P \arrow[d, "\varepsilon_P"] \\ N \arrow[r, "\varepsilon_N"'] & Q \arrow[ul, phantom, "\ulcorner", very near start] \end{tikzcd}$$

Proof. First we show that $\varepsilon_N n = \varepsilon_P p$. Notice that for any $m \in M$ we have

$$\begin{aligned} \varepsilon_P p(m) &= (\pi\iota_P)p(m) = \pi(0, p(m)) = (0, p(m)) + L, \\ \varepsilon_N n(m) &= (\pi\iota_N)n(m) = \pi(n(m), 0) = (n(m), 0) + L, \end{aligned}$$

therefore $\varepsilon_N n(m) - \varepsilon_P p(m) = (n(m), -p(m)) + L = L$, thus $\varepsilon_N n = \varepsilon_P p$.

Now, let X be any R-module, and consider morphisms $p'\colon P \to X$ and $n'\colon N \to X$ such that $p'p = n'n$. Define a map $q\colon Q \to X$ given by $(x, y) + L \mapsto n'(x) + p'(y)$, where $(x, y) \in N \oplus P$. We now show that q is well defined: suppose $(x, y) + L = (x', y') + L$— that is, $(x - x', y - y') \in L$—then since $q((x - x', y - y') + L) = q(L) = 0$ and thus their images are the same. Also, from construction, q is a uniquely determined morphism of R-modules and satisfies $p' = q\varepsilon_P$ and $n' = q\varepsilon_N$. ♮

Chapter 9

Integral Domains

9.1 Torsion

Definition 9.1.1 (Torsion)**.** Let R be an integral domain, and M be an R-module. We define the *torsion* of M to be the submodule

$$\mathtt{tor}\, M := \{m \in M : rm = 0 \text{ for some } r \in R \setminus 0\}.$$

If $\mathtt{tor}\, M = 0$, we say that M is of *free-torsion*. On the contrary, if $\mathtt{tor}\, M = M$, then M is said to be of *torsion*.

Definition 9.1.2 (Rank)**.** Given an integral domain R and an R-module M, we define the *rank* of M to be the cardinality of the maximal R-linearly independent subset of M.

Example 9.1.3. If R is an integral domain, and M is a torsion R-module, then $\operatorname{rank}_R M = 0$. Indeed, for any singleton $\{m\} \subseteq M$ there exists a non-zero $r \in R$ such that $rm = 0$, which shows that every singleton is R-linearly dependent.

9.2 Noetherian Rings & Modules

Chain Conditions

Recalling the definition of Noetherian rings (see Definition 8.3.35), we can extend this notion to the evironment of modules:

Definition 9.2.1. Let R be a commutative ring. We say that an R-module M is *Noetherian* if every submodule of M is finitely generated.

Proposition 9.2.2. Let R be a commutative ring, and M be an R-module. The following are equivalent propositions:

(a) M is a *Noetherian* module.

(b) Every *ascending chain* of submodules of M *stabilizes*. In other words, if $(N_j)_{j \in \mathbf{N}}$ is a collection of submodules of M such that $N_j \subseteq N_{j+1}$, then there exists an index $j_0 \in \mathbf{N}$ such that $N_j = N_{j+1}$ for all $j \geqslant j_0$.

(c) Every non-empty collection of submodules of M has a *maximal* element with respect to inclusion.

Proof.
- (a) $\Rightarrow$ (b). Let M be Noetherian and define the module $N := \bigcup_{j\in\mathbf{N}} N_j$, which is a submodule of M. Since submodules of Noetherian rings are finitely generated, it follows that N is finitely generated. Let $N = \langle n_1, \dots, n_k \rangle$ be its generating set. For all $1 \leqslant i \leqslant k$, there must exist $j_i \in \mathbf{N}$ such that $n_i \in N_j$ for all $j \geqslant j_i$. Taking the maximum $j_0 := \max(j_1, \dots, j_k)$, one finds that $n_i \in N_j$ for each $1 \leqslant i \leqslant k$ and every $j \geqslant j_0$. Therefore $N \subseteq N_j$ for all $j \geqslant j_0$, which implies that $N_j = N$ for each of those indexes — therefore the chain stabilizes.
- (b) $\Rightarrow$ (c). We prove the contrapositive. Suppose there exists a non-empty collection $\mathcal{N}$ of submodules of M admitting no maximal element. Let $N_0 \in \mathcal{N}$ be any element. Inductively, for all $j \geqslant 1$, define $N_j \in \mathcal{N}$ such that $N_{j-1} \subsetneq N_j$, that is, N_{j-1} is a *proper* subset of N_j — this is possible since N_{j-1} isn't maximal. The collection $(N_j)_{j\in\mathbf{N}}$ forms an ascending chain of submodules, but by construction does not stabilize.
- (c) $\Rightarrow$ (a). Let $N \subseteq M$ be any submodule. Since $(0) \subseteq N$ is a finitely generated submodule of N, one can define a non-empty collection $\mathcal{N}$ of finitely generated submodules of N. From (c) one has that $\mathcal{N}$ admits a maximum element, say $W := \langle n_1, \dots, n_k \rangle$. Let $n \in N$ be any element and consider the finitely generated submodule $\langle n_1, \dots, n_k, n \rangle \in \mathcal{N}$. Since W is maximal, we have $\langle n_1, \dots, n_k, n \rangle \subseteq W$ — therefore $n \in W$ and $N \subseteq W$. Therefore $N = W$ is finitely generated, which proves that M is Noetherian.

$\natural$

Corollary 9.2.3. Every principal ideal domain R is a Noetherian module over itself, and thus every non-empty collection of ideals of R admits a maximal element.

Lemma 9.2.4 (Quotient of Noetherian rings is Noetherian)**.** Let R be a Noetherian ring, and $\mathfrak{a} \subseteq R$ be an ideal. Then the quotient ring $R/\mathfrak{a}$ is Noetherian.

Proof. From Proposition 8.3.36 we find that the canonical projection $R \twoheadrightarrow R/\mathfrak{a}$ implies that $R/\mathfrak{a}$ is Noetherian. $\natural$

Theorem 9.2.5 (Generalized Hilbert's basis theorem)**.** Let R be a ring. Then R is Noetherian if and only if the polynomial ring $R[x_1, \dots, x_n]$.

Prove generalized Hilbert's basis theorem

Corollary 9.2.6. Let R be a Noetherian ring, and $\mathfrak{a} \subseteq R[x_1, \dots, x_n]$ be an ideal of the polynomial ring. Then the quotient ring $R[x_1, \dots, x_n]/\mathfrak{a}$ is Noetherian.

Proof. Since $R[x_1, \dots, x_n]$ is Noetherian by Theorem 9.2.5, applying Lemma 9.2.4 we find that $R[x_1, \dots, x_n]/\mathfrak{a}$ is Noetherian. $\natural$

Corollary 9.2.7. Every *finite-type* algebra over a Noetherian ring is *Noetherian*.

Existence of Maximal Ideals

Proposition 9.2.8. Let R be a commutative ring. If $\mathfrak{a}$ is any *proper* ideal of R, then there exists a *maximal ideal* $\mathfrak{m}$ of R containing $\mathfrak{a}$.

Proof. Consider the collection I of proper ideals of R containing $\mathfrak{a}$—which is ordered by inclusion. Using this ordering, let $(\mathfrak{a}_j)_{j \in J}$ be the chain of all proper ideals with $\mathfrak{a}_j \in I$ and $\mathfrak{a}_j \subseteq \mathfrak{a}_{j+1}$. Define the set $\mathfrak{m} := \bigcup_{j \in J} \mathfrak{a}_j$, which is again an ideal of R. Since every $\mathfrak{a}_j$ contains $\mathfrak{a}$ and does *not* contain 1, it follows that $\mathfrak{m}$ contains $\mathfrak{a}$ and also doesn't contain 1—hence $\mathfrak{m}$ is a proper ideal. Therefore $\mathfrak{m}$ is a *maximal* ideal of R containing $\mathfrak{a}$, which proves the statement. ♮

9.3 Localization

Definition 9.3.1 (Multiplicative subset)**.** Given a commutative ring R, a subset $S \subseteq R$ is said to be *multiplicative* if $1_R \in S$ and S is closed under multiplication—that is, given $s, s' \in S$, we have $ss' \in S$.

Definition 9.3.2 (Localization)**.** Let R be a commutative ring, and $S \subseteq R$ be a *multiplicative subset*. A *localization* of R over the set S is a morphism of *commutative rings* $L_S \colon R \to R[S^{-1}]$ such that:

(a) For all $s \in S$, the image $L_S(s) \in R[S^{-1}]$ is a *unit*.

(b) For every morphism of rings $f \colon R \to K$ satisfying property (a), there exists a *unique* ring morphism $\phi \colon R[S^{-1}] \to K$ such that the diagram

$$\begin{tikzcd} R \arrow[r, "L_S"] \arrow[dr, bend right, "f"'] & R[S^{-1}] \arrow[d, dashed, "\phi"] \\ & K \end{tikzcd}$$

Equivalently, consider the minimal equivalence relation $\sim_{\text{frac}}$ on the set $R \times S$ defined by $(r, s) \sim_{\text{frac}} (r', s')$ if and only if $rs' - sr' = 0$. Then the localization of R under the multiplicative subset S is simply

$$R[S^{-1}] = (R \times S)/\sim_{\text{frac}}.$$

Example 9.3.3. Let $p \in \mathbf{Z}$ be a prime number, and consider the ring of fractions $\mathbf{Z}[p^{-1}]$. Then every proper submodule of the $\mathbf{Z}$-module given by the quotient $\mathbf{Z}[p^{-1}]/\mathbf{Z}$ is *finite*.

How to solve this?

Definition 9.3.4. Let R be an integral domain. We define the *field of fractions* of R to be the localization

$$\operatorname{Frac}(R) := R[(R \setminus \{0\})^{-1}].$$

It is immediate that $\operatorname{Frac}(R)$ is a field, since a field is simply an integral domain whose non-zero elements are units—and this exactly what we did with the above localization.

9.4 Principal Ideal Domains

Modules Over PID's

Theorem 9.4.1. Let R be a principal ideal domain, and M be a free R-module with $\operatorname{rank}_R M = m$ finite, and $N \subseteq M$ a submodule. Then:

(a) The submodule N is free with $\operatorname{rank}_R N = n$ satisfying $n \leqslant m$.

(b) There exists a basis $(y_1, \dots, y_m)$ of M, and a collection $(a_1, \dots, a_m)$ of R such that $(a_1 y_1, \dots, a_n y_n)$ is a basis for N and $a_1 \mid a_2 \mid \cdots \mid a_n$.

Proof. Fix a basis $(x_1, \dots, x_m)$ for M and define a collection of R-linear morphisms $(\pi_j \colon M \twoheadrightarrow R)_{j=1}^n$ with $\pi_j x_i = \delta_{ij}$.

(a) Since R is a domain, M is torsion-free and so is N. Since M has a finite rank, it must also be the case that N has a finite rank n satisfying $n \leqslant m$—which follows from Proposition 10.2.14. We proceed via induction on n: for the base case $n = 0$ we have $N = 0$ and therefore N is free. We assume for the induction hypothesis that the proposition is true for some $0 < n - 1 < m$. Now we consider the case $0 < n \leqslant m$: since n is non-zero, then N is a non-zero module and therefore there exists a non-zero $x = \sum_{j=1}^n b_j x_j \neq 0$ element of N. If $1 \leqslant j_0 \leqslant n$ is an index such that $b_{j_0} \neq 0$, then in particular $\pi_{j_0}|_N$ is a non-zero map of the form $N \to R$. Since R is a PID and $\pi_{j_0} N \subseteq R$ is a non-zero submodule, then there must exist $b_0 \in R \smallsetminus 0$ for which $\pi_{j_0} N = R b_0$. Since R is a domain, then $\{b_0\}$ is a basis for the free R-module $\pi_{j_0} N$. Then the short exact sequence

$$0 \longrightarrow \ker(\pi_{j_0}|_N) \hookrightarrow N \xrightarrow{\pi_{j_0}|_N} Rb_0 \longrightarrow 0$$

ends with a free module and thus splits, hence

$$N \simeq \ker(\pi_{j_0}|_N) \oplus Rb_0.$$

Since rank is additive over direct sums, then $n = \operatorname{rank}_R(\ker(\pi_{j_0}|_N)) + 1$ implies that $\ker(\pi_{j_0}|_N) \subseteq M$ is a submodule of rank $n - 1$. From the inductive hypothesis it follows that $\ker(\pi_{j_0}|_N)$ is a free module. Therefore N is the direct sum of two free modules, hence N is itself free.

(b) Notice that the case $M = 0$ is trivial, thus we shall consider only $m \geqslant 1$ and do induction on m—moreover, we'll assume that N is anon-zero submodule of M, since the zero case is trivially satisfied. For the base case $m = 1$, there exists an R-module isomorphism $\phi \colon M \xrightarrow{\simeq} R$. Then from the fact that R is a PID it follows that there exists $a \in R$ such that $\phi N = Ra = aR$, and if $u := \phi^{-1}(1)$ then $M = Ru$—so that $\{u\}$ is a basis for M. Moreover,

$$N = a(Ru) = (aR)u = (Ra)u = R(au),$$

therefore $\{au\}$ forms a basis for N.

Now let $m > 1$ and assume as a hypothesis for induction that the proposition is true for all modules of rank less than m. For each $\phi \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, R)$, let $a_\phi \in R$ denote a ring element such that $\phi N = Ra_\phi$, and define a collection of principal ideals

$$\Sigma := (Ra_\phi)_{\phi \in \mathrm{Mor}(M,R)}.$$

From our previous considerations in item (a), we know that there must exist a non-zero $\pi_{j_0} \in \mathrm{Mor}(M, R)$—thus Σ is non-empty and its elements are not all zero. Since R is Noetherian, it follows that Σ admits a maximal element $Ra_{\phi_1} \in \Sigma$. Define $a_1 := a_{\phi_1}$ and let $y \in N$ be such that $\phi_1 y = a_1$, which exists because $\phi_1 N = Ra_1$. We'll show that $a_1 \mid \varphi y$ for any $\varphi \in \mathrm{Mor}(M, R)$. Let $\varphi\colon M \to R$ be any such map, and let $\mathfrak{p} := \langle a_1, \varphi y\rangle$ be an ideal of R, so that there exists $d \in R$ such that $\mathfrak{p} = Rd$ since $\mathfrak{p}$ is principal. Therefore there exists a pair $r_1, r_2 \in R$ such that $d = r_1 a_1 + r_2 \varphi y$. Define a morphism $\psi\colon M \to R$ given by $\psi := r_1\phi_1 + r_2\varphi$ so that $\psi y = d$ and hence $d \in \psi N$ from the fact that $y \in N$. Since $a_1 \in Rd$, then $Ra_1 \subseteq Rd$ and from maximality we obtain the equality $Ra_1 = Rd$, which proves that $a_1 \mid d$—while $d \mid \varphi y$, thus $a_1 \mid \varphi y$ as we wanted.

In particular, the last paragraph shows that $a_1 \mid \pi_j y$ for each projection $1 \leqslant j \leqslant m$, thus we may find $b_j \in R$ such that $\pi_j y = a_1 b_j$. We shall define

$$y_1 := \sum_{j=1}^{m} b_j x_j.$$

Notice then that y can be rewritten as

$$y = \sum_{j=1}^{m} \pi_j(y) x_j = \sum_{j=1}^{m} (a_1 b_j) x_j = a_1 \sum_{j=1}^{m} b_j x_j = a_1 y_1. \tag{9.1}$$

Furthermore, one knows that $a_1 = \psi y = \psi(a_1 y_1) = a_1 \psi y_1$, therefore since R is a domain it follows that $\psi y_1 = 1$.

We'll now show that $M = Ry_1 \oplus \ker\psi$. Let $x \in M$ be any element and write $x = \psi(x) y_1 + (x - \psi(x) y_1)$. Notice that

$$\psi(x - \psi(x)y_1) = \psi x - \psi(x)\psi(y_1) = 0 \tag{9.2}$$

so $x - \psi(x) y_1 \in \ker\psi$. Therefore $M = Ry_1 + \ker\psi$. Now if $ry_1 \in \ker\psi$ then from the fact that $\psi y_1 = 1$ we have $r = 0$, hence $ry_1 = 0$ and thus $Ry_1 \cap \ker\psi = 0$. We thus conclude that $M = Ry_1 \oplus \ker\psi$.

Further, we'll prove that $N = R(a_1 y_1) \oplus (N \cap \ker\psi)$. Let $x \in N$ be any element and as before rewrite it as $x = \psi(x) y_1 + (x - \psi(x) y_1)$. Using the fact that $\psi N = Ra_1$, there exists $b \in R$ such that $\psi x = ba_1$—then one has

$$x = (ba_1) y_1 + (x - (ba_1) y_1)$$

where $ba_1y_1 \in R(a_1y_1)$ and $x - (ba_1)y_1 \in N \cap \ker\psi$ from Eq. (9.2). Therefore $N = R(a_1y_1) + (N \cap \ker\psi)$, and since $R(a_1y_1) \subseteq Ry_1$ while $N \cap \ker\psi \subseteq \ker\psi$ then $R(a_1y_1) \cap (N \cap \ker\psi) = 0$. Thus

$$N = R(a_1y_1) \oplus (N \cap \ker\psi).$$

Since $\ker\psi \subseteq M$ then it is free and the rank of the direct sum yields $\operatorname{rank}_R \ker\psi = m - 1$. Thus from the inductive hypothesis we conclude that there exists a basis $(y_2, \dots, y_m)$ of $\ker\psi$ and a collection of ring elements $(a_2, \dots, a_n)$ such that $(a_2y_2, \dots, a_ny_n)$ forms a basis for $N \cap \ker\psi$, and $a_2 \mid a_3 \mid \cdots \mid a_m$. Since $M = Ry_1 \oplus \ker\psi$ then $(y_1, y_2, \dots, y_m)$ is a basis for M, and since $N = R(a_1y_1) \oplus (N \cap \ker\psi)$ we have that $(a_1y_1, a_2y_2, \dots, a_ny_n)$ is a basis for N.

To conclude the proof, we must show that $a_1 \mid a_2$. Since $(y_1, \dots, y_m)$ is a basis of M, let $f : M \to R$ be an R-linear morphism such that $fy_1 = fy_2 = 1$ while $fy_j = 0$ for each $3 \leqslant j \leqslant m$. From Eq. (9.1) we have $a_1 = a_1 f y_1 = f(a_1y_1) = fy$, therefore since $y \in N$ then $a_1 \in fN$. This implies in $Ra_1 \subseteq fN$ but using maximality this yields the equality $fN = Ra_1$. We also know that $a_2 = f(a_2y_2)$ in a similar fashion, and since $a_2y_2 \in N$ then $a_2 \in fN$. This concludes the proof that $a_1 \mid a_2$, which was the last step for the end of the proof at large.

$\natural$

Decomposition Theorems

Theorem 9.4.2 (Fundamental theorem of invariant factors)**.** Let R be a principal ideal domain and M a finitely generated R-module. Then the following holds:

(a) There exists a number $r \in \mathbf{N}$ and a collection of non-zero and non-invertible elements $(a_1, \dots, a_m)$ of R such that

$$M \simeq R^r \oplus \Big(\bigoplus_{j=1}^{m} R/Ra_j \Big),$$

and $a_1 \mid a_2 \mid \cdots \mid a_m$. That is, M is isomorphic to the direct sum of a finite collection of cyclic R-modules. The elements a_j are said to be invariant factors of M.

(b) The module M is torsion-free if and only if M is free.

(c) There exists an isomorphism

$$\mathtt{tor}\, M \simeq \bigoplus_{j=1}^{m} R/Ra_j.$$

In particular, M is a torsion module if and only if $r = 0$, and in such a case we have $\operatorname{Ann} M = Ra_m$[1].

[1]The annihilator of a module M is defined to be the collection

$$\operatorname{Ann} M := \{r \in R : rm = 0 \text{ for all } m \in M\}.$$

Proof. Let $M = \langle x_1, \dots, x_n \rangle$ and let $e_j := (\delta_{ij})_{i=1}^n \in R^n$ be a base element of R^n. Define a morphism $\phi \colon R^n \to M$ as the map that sends $e_j \mapsto x_j$ for each j—which is unique by the free module universal property. Since $\{x_1, \dots, x_n\}$ generates M, then $\operatorname{im} \phi = M$ and ϕ is thus surjective. Via the first isomorphism theorem we obtain $R^n / \ker \phi \simeq M$. Since $\ker \phi$ is a submodule of R^n we can use Theorem 9.4.1 to obtain that $\ker \phi$ is free and has $\operatorname{rank}(\ker \phi) := m \leqslant n$. Furthermore, there exists a basis $(t_1, \dots, t_n)$ of R^n and a list of ring elements $(a_1, \dots, a_m)$ such that $(a_1 t_1, \dots, a_m t_m)$ is a basis of $\ker \phi$ satisfying $a_1 \mid a_2 \mid \cdots \mid a_m$. From the decomposition of free modules we obtain

$$M \simeq R^n / \ker \phi = \frac{R t_1 \oplus \cdots \oplus R t_n}{R a_1 t_1 \oplus \cdots \oplus R a_m t_m}$$

Define a morphism of R-modules $\psi \colon \bigoplus_{j=1}^n R t_j \to \left(\bigoplus_{j=1}^m R / R a_j \right) \oplus R^{n-m}$ by mapping

$$\psi(r_1 t_1, \dots, r_n t_n) := (r_1 + R a_1, \dots, r_m + R a_m, r_{m+1}, \dots, r_n).$$

From the definition, this map is certainly surjective. Moreover, we can see that

$$\ker \psi = \left(\bigoplus_{j=1}^m R a_j \right) \oplus \{0\}^{\oplus (n-m)} \simeq \bigoplus_{j=1}^m R a_j,$$

therefore by the first isomorphism theorem we find

$$\frac{R t_1 \oplus \cdots \oplus R t_n}{R a_1 t_1 \oplus \cdots \oplus R a_m t_m} \simeq \left(\bigoplus_{j=1}^m R / R a_j \right) \oplus R^{n-m}.$$

Notice that if $a \in R$ is an invertible element, then $Ra = R$ and hence $R/Ra = 0$. Therefore the proposition of item (a) is stablished since we can cut off of the direct sum every summand R/Ra_j where a_j is invertible. Let $b \in R$ be any element, then given any $r + Rb \in R/Rb$ we find that

$$b(r + Rb) = br + Rb = rb + Rb = Rb$$

therefore R/Rb is a torsion R-module. This shows that M is torsion-free if and only if $M \simeq R^r$, which proves item (b). As we just noted, one has $\mathtt{tor}\, M \simeq \bigoplus_{j=1}^m R/Ra_j$, and M is a torsion module we have Ann

Notice that this isn't finished: it is part of the lost work—this chapter was the most affected by the losses.

♮

Chinese Remainder Theorem

Theorem 9.4.3 (Chinese remainder)**.** Let R be a commutative ring, and $\mathfrak{a}_1, \dots, \mathfrak{a}_k$ be ideals of R such that $\mathfrak{a}_i + \mathfrak{a}_j = R$ for all $i \neq j$. Then:

- We have the equality $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_k = \mathfrak{a}_1 \cdot \ldots \cdot \mathfrak{a}_k$.

- The natural projection $R \twoheadrightarrow R/\mathfrak{a}_1 \times \cdots \times R/\mathfrak{a}_k$ is *surjective*, and induces a natural *isomorphism* of rings

$$\frac{R}{\mathfrak{a}_1 \ldots \mathfrak{a}_k} \simeq (R/\mathfrak{a}_1) \times \cdots \times (R/\mathfrak{a}_k)$$

Proof. In particular, we have $\mathfrak{a}_j + \mathfrak{a}_k = R$ for all $1 \leqslant j \leqslant k-1$. Therefore, for all such indices there exists $a_j \in \mathfrak{a}_k$ such that $1 - a_j \in \mathfrak{a}_j$, hence

$$(1-a_1)\ldots(1-a_{k-1}) \in \mathfrak{a}_1 \cdot \ldots \cdot \mathfrak{a}_{k-1},$$

and since $a_j \in \mathfrak{a}_k$, then $1 - \prod_{j=1}^{k-1}(1-a_j) \in \mathfrak{a}_k$. This shows that

$$(\mathfrak{a}_1 \cdot \ldots \cdot \mathfrak{a}_{k-1}) + \mathfrak{a}_k = R. \tag{9.3}$$

Notice that $\mathfrak{a}_1 \cdot \ldots \cdot \mathfrak{a}_k \subseteq \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_k$, thus it remains to prove the other side of the inclusion. From our last paragraph, we know that $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_k \subseteq \mathfrak{a}_1 \cdot \ldots \cdot \mathfrak{a}_k$ for all $k \geqslant 3$. Since $k = 1$ is trivial, we just need to prove the case for $k = 2$. Let $\mathfrak{b}, \mathfrak{c} \subseteq R$ be ideals such that $\mathfrak{b} + \mathfrak{c} = R$—then there exists $b_0 \in \mathfrak{b}$ and $c_0 \in \mathfrak{c}$ such that $b_0 + c_0 = 1$. If $x \in \mathfrak{b} \cap \mathfrak{c}$, then $x = b_0 x + c_0 x$, implying in $x \in \mathfrak{b} \cdot \mathfrak{c}$. Thus $\mathfrak{b} \cap \mathfrak{c} \subseteq \mathfrak{b} \cdot \mathfrak{c}$.

We now prove the second assertion via induction on k. For the base case $k = 1$, the statement follows trivially from the first isomorphism theorem. Assume as the hypothesis of induction that the statement is true $k - 1 > 1$—that is, we have an isomorphism $R/(\mathfrak{a}_1 \cdot \ldots \cdot \mathfrak{a}_{k-1}) \simeq (R/\mathfrak{a}_1) \times \cdots \times (R/\mathfrak{a}_{k-1})$. Consider the natural map

$$\pi \colon R \longrightarrow \frac{R}{\mathfrak{a}_1 \cdot \ldots \cdot \mathfrak{a}_{k-1}} \times (R/\mathfrak{a}_k). \tag{9.4}$$

From Eq. (9.3) the statement is reduced for the case of two ideals $\mathfrak{b} := \mathfrak{a}_1 \cdot \ldots \cdot a_{k-1}$ and $\mathfrak{c} := \mathfrak{a}_k$ such that $\mathfrak{b} + \mathfrak{c} = R$. Let $b, c \in R$ be any two elements—we shall show that there exists $r \in R$ such that $r - b \in \mathfrak{b}$ and $r - c \in \mathfrak{c}$. Since $\mathfrak{b}$ and $\mathfrak{c}$ are relatively prime, let as before $b_0 + c_0 = 1$ and define $r := b_0 c + c_0 b$. Then, one has

$$\begin{aligned} r &= b_0 c + (1 - b_0) b = b + b_0(c - b) \equiv b \pmod{\mathfrak{b}} \\ r &= (1 - c_0) c + c_0 b = c + c_0(b - c) \equiv c \pmod{\mathfrak{c}} \end{aligned}$$

since $b_0 \in \mathfrak{b}$ and $c_0 \in \mathfrak{c}$. Therefore $r - c \in \mathfrak{b}$ and $r - b \in \mathfrak{c}$ a wanted. This shows that the natural morphism of rings π (see Eq. (9.4)) is surjective. From our first considerations, we know that $\ker \pi = \mathfrak{b} \cap \mathfrak{c} = \mathfrak{b} \cdot \mathfrak{c}$, therefore the first isomorphism theorem establishes that $R/(\mathfrak{b} \cdot \mathfrak{c}) \simeq (R/\mathfrak{b}) \times (R/\mathfrak{c})$. ♮

Corollary 9.4.4 (Chinese remainder for PIDs)**.** Let R be a principal ideal domain, and $a_1, \ldots, a_k \in R$ be elements such that $\gcd(a_i, a_j) = 1$ for all pairs $i \neq j$. Then the natural map $r + (a_1 \cdots a_k) \mapsto (r + (a_1), \ldots, r + (a_k))$ establishes an *isomorphism* of rings

$$R/(a_1 \cdots a_k) \simeq \frac{R}{(a_1)} \times \cdots \times \frac{R}{(a_k)}.$$

9.5 Unique Factorisation Domains

Chapter 10

Linear Algebra Over Rings

Remark 10.0.1. For the remainder of the chapter, if M is said to be an R-module, then M can be *either a left or right R-module*—and R need *not be commutative*—we use this to *generalize* propositions for *both* left and right modules simultaneously, whenever possible. On the other hand, if we *specify* the side of the multiplicative structure of M, then it may well be the case that the proposition does *only work* for this strict case *or* that the *distiction* between right and left modules is *important* in some way.

10.1 Free Modules

Construction

As always, if S is a set and M is an R-module, for some ring R, we define $M^{\oplus S}$ to be the collection of *finitely supported set-functions $S \to M$*. We define on $M^{\oplus S}$ an R-module structure as follows: for every $\alpha \in M^{\oplus S}$ and $r \in R$ we define

$$(r\alpha)(s) := r(\alpha(s)).$$

If we consider the module of R over itself, we can define a set-function $\iota\colon S \to R^{\oplus S}$ by mapping $s \mapsto \mathbf{s}$, where

$$\mathbf{s}(x) := \begin{cases} 1, & \text{if } x = s, \\ 0, & \text{if } x \neq s. \end{cases}$$

For every set S, we define a corresponding module module $F_R S$ composed of of formal sums $\sum_{s\in S} a_s s$ such that $a_s \in R$ is non-zero only for finitely many $s \in S$.

Proposition 10.1.1. There exists a natural isomorphism of R-modules $F_R S \simeq R^{\oplus S}$.

Proof. Let $\Phi\colon F_R S \to R^{\oplus S}$ be a map given by $\sum_{s\in S} a_s s \mapsto \sum_{s\in S} a_s \mathbf{s}$. Clearly, such map is injective and preserves both multiplicative and additive structures, thus Φ is an R-module morphism. Moreover, if $\alpha \in R^{\oplus S}$ is any function, since α has finite support, the formal sum $\sum_{s\in S} \alpha(s)s$ is a well defined element of $F_R S$. Also, mapping this element under Φ yields $\sum_{s\in S} \alpha(s)\mathbf{s}$ and, for every $x \in S$, we have $\sum_{s\in S} \alpha(s)\mathbf{s}(x) = \alpha(s)\mathbf{x}(x) = \alpha(x)$. Therefore the map is surjective, hence an isomorphism. ♮

Proof. Let A and B be two sets such that $F_R A = F_R B$. In particular, it follows that for all $a \in A$, there exists an element $a = \sum_{j=1}^{n} r_j b_j \in F_R B$, moreover, each $b_i \in B$ can be written as $\sum_{j=1}^{m} r_j(b_i)a_j$ ♮

Proposition 10.1.2 (Free module universal property)**.** Let R be a ring and S be any set. Given any R-module M and a *set-function* $f\colon S \to M$, there exists a *unique R-module morphism* $\phi\colon F_R S \to M$ such that the following diagram commutes in `Set`[1]:

$$\begin{tikzcd} F_R S \arrow[r, "\phi"] & M \\ S \arrow[u, "\iota"] \arrow[ur, bend right, "f"'] & \end{tikzcd}$$

where $\iota\colon S \to F_R S$ is the mapping $s \mapsto s$.

Proof. Let $\phi\colon F_R S \to M$ be such that $\phi(\sum_{s\in S} a_s s) \mapsto \sum_{s \in S} a_s f(s)$ for any $\sum_{s\in S} a_s s \in F_R S$ so that clearly $\phi\iota = f$ since $\phi\iota(s) = \phi(s) = f(s)$. Moreover, for any two $\sum_{s\in S} a_s s, \sum_{s\in S} b_s s \in F_R S$ we have

$$\begin{aligned} \phi\left(\sum_{s\in S} a_s s + \sum_{s\in S} b_s s\right) &= \phi\left(\sum_{s\in S}(a_s + b_s)s\right) = \sum_{s\in S}(a_s+b_s)f(s) = \sum_{s\in S} a_s f(s) + \sum_{s\in S} b_s f(s) \\ &= \phi\left(\sum_{s\in S} a_s s\right) + \phi\left(\sum_{s\in S} b_s s\right). \end{aligned}$$

Furthermore, ϕ certainly preserves the multiplicative structure by R elements. Therefore ϕ is an R-module morphism. Since f completely determines the image of ϕ, it is the unique morphism such that $\phi\iota = f$. ♮

Corollary 10.1.3. The mapping $\iota\colon S \hookrightarrow F_R S$ given by $s \mapsto s$ is injective.

Proof. For any pair $s, s' \in S$ of distinct elements, consider the module $M := F_R\{s, s'\}$ and a set-function $f\colon S \to M$ with $f(s) = s$ and $f(s') = s'$. Then, by the universal of free modules, there exists a unique morphism of R-modules $\phi\colon F_R S \to M$ such that $\phi\iota = f$. If $\iota(s)$ was equal to $\iota(s')$, then $f(s)\ f(s')$, which cannot be the case—thus $\iota(s) \neq \iota(s')$ for all $s, s' \in S$ distinct, hence ι is injective. ♮

Proposition 10.1.4. Let $A := \{1, \dots, n\}$ be a set of n elements and define a map $\iota\colon A \to R[x_1, \dots, x_n]$ by $j \mapsto x_j$. Then $R[x_1, \dots, x_n]$ is a *free commutative R-algebra on A*.

Proof. We prove that $R[x_1, \dots, x_n]$ satisfies the universal property for A. Let $\alpha\colon R \to M$ be any R-algebra and $f\colon A \to M$ be a set-function. From Proposition 8.2.9 we find a unique extension $\overline{\alpha}\colon R[x_1, \dots, x_n] \to S$ such that $\overline{\alpha}|_R = \alpha$ and $\overline{\alpha}(x_j) := f(j)$. Therefore, $\overline{\alpha}$ is the uniquely determined R-algebra morphism such that $\overline{\alpha}\iota = f$. ♮

[1]It is to be noted the subtlety of not adding a dashed arrow (denoting uniqueness) for ϕ in the diagram—this is done purposefully since the diagram is commutative in the category of sets, so there may well be a distinct *set-function* $F_R S \to M$ also making the diagram commute in `Set`.

Free Modules from Subsets

Given any R-module M and a subset $S \subseteq M$ of its elements, one can define a free module $R^{\oplus S}$ out of S and, from the universal property of free modules, there exist a unique R-map $\phi\colon R^{\oplus S} \to M$ such that $\phi\iota = i$, where $\iota\colon S \hookrightarrow R^{\oplus S}$ and $i\colon S \hookrightarrow M$ is the canonical inclusion.

It is to be noted that the image $\phi(R^{\oplus S}) \subseteq M$ is a submodule of M and its elements are of the form $\sum_{s\in S} a_s s$ for $a_s \in R$ non-zero only for finitely many $s \in S$. We'll denote the submodule *generated* by S on M by the notation $\langle S \rangle$.

Definition 10.1.5 (Finitely generated module)**.** An R-module M is said to be *finitely generated* exactly when there exists a *finite subset* $S \subseteq M$ such that $M = \langle S \rangle$. Equivalently, M is finitely generated if there exists a *surjective* R-map $R^{\oplus n} \twoheadrightarrow M$ for a positive integer $n \in \mathbf{Z}_{>0}$.

Remark 10.1.6 (Submodules)**.** The reader should be cautious when working with finitely generated modules, for instance, *it is not true* that a finitely generated module has finitely generated modules.

For instance, if we consider the ring of polynomials $P := \mathbf{Z}[x_1, x_2, \dots]$ on infinitely many variables, then $P = \langle 1 \rangle$ is finitely generated *as a P-module*. However, if we consider the ideal $\mathfrak{a} := (x_1, x_2, \dots) \subseteq P$, one cannot take a finite collection of polynomials of P and generate all of $\mathfrak{a}$. Indeed, if $S \subseteq P$ is any finite collection, since polynomials have finitely many terms, there must exist, for all $p(x_1, x_2, \dots) \in S$, a maximum index $j \in \mathbf{Z}_{>0}$ such that x_j appears as a variable in p with a non-zero coefficient. Taking the maximum index over all polynomials of S, we are still left with only a finite index, say $n \in \mathbf{Z}_{>0}$, such that x_n is the highest-index variable appearing any of the polynomials of S—hence x_{n+1} is not contemplated by any of the polynomials of S, therefore this set cannot generate $\mathfrak{a}$.

Definition 10.1.7 (Noetherian module)**.** An R-module M is said to be a *Noetherian module* if every submodule of M is *finitely generated as an R-module*.

We can state the definition for *Noetherian rings* (see Definition 8.3.35) equivalently as follows: a ring R is said to be a Noetherian if R is a Noetherian R-module.

Lemma 10.1.8. Let M be an R-module, and $N \subseteq M$ be a submodule. Then if both N and M/N are finitely generated, then M is finitely generated.

Proof. Let $N = \langle A \rangle$ and $M/N = \langle B \rangle$ for finite sets $A \subseteq N$ and $B \subseteq M/N$. If $m \in M$ is any element, then the class $m + N \in M/N$ can be written as $m + N = \sum_{b\in B} r_b b + N$ for $r_b \in R$. Moreover, since $m - \sum_{b\in B} r_b b \in N$, we can write $m - \sum_{b\in B} r_b b = \sum_{a\in A} r_a a$ for $r_a \in R$. Therefore

$$m = \sum_{a\in A} r_a a + \sum_{b\in B} r_b b.$$

This shows that $A \cup B$ generates M, thus M is finitely generated. ♮

Proposition 10.1.9. Let M be an R-module, and $N \subseteq M$ be a submodule. Then M is Noetherian if and only if *both* N and M/N are Noetherian.

Proof. Analogous to the proof of Proposition 8.3.36, if M is a Noetherian module, then the natural projection $\pi\colon M \twoheadrightarrow M/N$ shows that M/N is Noetherian. Since N is a submodule of M, then N is finitely generated and every submodule $P \subseteq N$ is also a submodule of M, thus P must be finitely generated.

For the converse, suppose both M/N and N are Noetherian. Fix any submodule $P \subseteq M$. Consider the submodule $P \cap N$ of both P and N—from the hypothesis that N is Noetherian, $P \cap N$ is finitely generated. By Proposition 8.4.51 we find that there exists a canonical isomorphism $P/(P \cap N) \simeq (N + P)/N$. Since $(N + P)/N$ is a submodule of M/N, by the hypothesis that M/N is Noetherian it follows that $P/(P \cap N)$ is finitely generated. From Lemma 10.1.8 we find that P itself is finitely generated, making M Noetherian. ♮

Corollary 10.1.10. Given a *Noetherian ring* R, any *finitely generated* R-module M is *Noetherian*.

Proof. Since M is finitely generated, there exists a surjective R-module morphism $p\colon R^{\oplus n} \twoheadrightarrow M$, for some $n \in \mathbf{Z}_{>0}$—therefore $M \simeq R^{\oplus n}/\ker p$. Now, if $R^{\oplus n}$ is Noetherian, by Proposition 10.1.9, then $\ker p$ and $R^{\oplus n}/\ker p$ are both Noetherian, and therefore M is Noetherian.

We prove that $R^{\oplus n}$ is Noetherian by induction on n. For $n = 1$ we have that $R^{\oplus 1} \simeq R$ is Noetherian. Assume that $R^{\oplus(n-1)}$ is Noetherian for some $n > 1$. Notice that the inclusion $\iota\colon R^{\oplus(n-1)} \hookrightarrow R^{\oplus n}$ mapping $(r_1, \dots, r_{n-1}) \mapsto (r_1, \dots, r_{n-1}, 0)$ is an R-map and makes $R^{\oplus(n-1)}$ into a submodule of $R^{\oplus n}$. Considering the canonical projection of the n-th coordinate $\pi_n\colon R^{\oplus n} \twoheadrightarrow R$, we have $\ker \pi_n = R^{\oplus(n-1)}$. By the first isomorphism theorem we obtain an isomorphism

$$R^{\oplus n}/R^{\oplus(n-1)} \simeq R,$$

therefore $R^{\oplus n}/R^{\oplus(n-1)}$ is Noetherian. Applying Proposition 10.1.9 we find that $R^{\oplus n}$ is Noetherian—thus M is Noetherian. ♮

Finite Generation of R-Algebras

Definition 10.1.11. Let A be an R-algebra. We define the following two distinct concepts regarding the finiteness of the generation of A:

(a) The R-algebra A is said to be *finite* if there exists a *surjective R-module morphism*

$$\phi\colon R^{\oplus n} \twoheadrightarrow A$$

for some $n \in \mathbf{Z}_{>0}$, so that

$$A \simeq R^{\oplus n}/\ker \phi,$$

where $\ker \phi$ is a *submodule* of $R^{\oplus n}$. In other words, A is finitely generated as a *module over* R. This nomenclature is unfortunately misleading: even though we say that A is finite, A may well be an *infinite set*.

(a) The R-algebra A is said to be of *finite type* if there exists a *surjective R-algebra morphism*

$$\alpha\colon R[x_1, \dots, x_n] \twoheadrightarrow A$$

for some $n \in \mathbf{Z}_{>0}$, so that

$$A \simeq R[x_1, \dots, x_n]/\ker \alpha,$$

where $\ker \alpha$ is an *ideal* of the ring $R[x_1, \dots, x_n]$. In other words, A is finitely generated as an *algebra over R*.

Remark 10.1.12 (Finite type but not finite)**.** Let A be an R-algebra. If A is finite, then A is also of finite type. On the other hand, if we consider the R-algebra $R[x]$, we find that $R[x]$ is clearly of finite type, but there exists no surjective R-map from $R^{\oplus n}$ to $R[x]$, therefore $R[x]$ isn't finite.

10.2 Linear Independence & Bases

Linear Independence

Definition 10.2.1 (Linear independence)**.** Let J be a set and M be an R-module. An indexed set $j\colon J \to M$ is said to be *linearly independent* if the unique R-module morphism $\phi\colon F_R J \to M$, making the diagram

$$\begin{tikzcd} F_R J \arrow[r, dashed, "\phi"] & M \\ J \arrow[u, hook, "\iota"] \arrow[ur, bend right, "j"'] & \end{tikzcd}$$

commute in `Set`, is *injective*. The indexed set j is said to *generate* M if ϕ is *surjective*. Finally, if ϕ is an *isomorphism*, then j is a *basis* of M—in this case, since $F_R J \simeq R^{\oplus J}$, then $R^{\oplus J} \simeq M$.

Corollary 10.2.2. An R-module is *free* if and only if it admits a *basis*.

Proof. If M is free, then there exists a set S such that $M \simeq F_R S$, then S is a basis for M. For the converse, if M admits a basis S, then $F_R S \simeq M$. ♮

As with vector spaces we'll intentionally identify an indexed set $j\colon J \to M$ simply by J itself and say that J is a *subset* of M. The good old abuse of notation.

Remark 10.2.3 (Singletons)**.** Singletons do *not* need to be linearly independent. A simple example is $\{3\} \subseteq \mathbf{Z}/9\mathbf{Z}$.

The following lemma regarding maximality of linearly independent sets is equivalent to the Axiom of Choice.

Lemma 10.2.4 (Maximality)**.** Let M be an R-module and $S \subseteq M$ be a linearly independent set. There exists a *maximal linearly independent* subset of M containing S.

Proof. Let $\mathcal{S}$ be the non-empty collection of linearly independent sets of M containing S. Notice that a the union of a chain of elements of $\mathcal{S}$ is again a linearly independent set containing S—thus $\mathcal{S}$ is closed under arbitrary unions. In other words, every chain of elements has an upper bound in $\mathcal{S}$. By Zorn's lemma, it follows that the collection $\mathcal{S}$ has a maximal element. ♮

Remark 10.2.5 (Maximality, generation, and bases)**.** It should be noted right away that being a maximal linear independent set does *not* imply that the set generated the module. For instance, $\{2\} \subseteq \mathbf{Z}$ is a maximal linearly independent subset of $\mathbf{Z}$ containing $\{2\}$, but obviously it doesn't generate $\mathbf{Z}$. This, however is true for the case of vector spaces.

On the other hand, every base of a module is necessarily a maximal linearly independent set.

Definition 10.2.6 (Invariant basis number)**.** A ring R is said to satisfy the *invariant basis number* (IBN) property if $R^m \simeq R^n$ as R-modules if and only if $m = n$.

Corollary 10.2.7. A ring R does *not* satisfy the invariant basis number property if and only if there exists distinct natural numbers $n, m \in \mathbf{N}$, and matrices $A \in M_{m\times n}(R)$ and $B \in M_{n\times m}(R)$ such that

$$AB = \mathrm{id}_m \quad \text{and} \quad BA = \mathrm{id}_n .$$

Corollary 10.2.8. Every field satisfies the invariant basis number property.

Proposition 10.2.9. Let $f: R \to S$ be a ring morphism, for *any* two rings R and S. If S satisfies the invariant basis number property, then so does R.

Proof. Let $A := [a_{ij}] \in M_{m\times n}(R)$ and $B := [b_{ij}] \in M_{n\times m}(R)$ be matrices and define both $fA := [f(a_{ij})] \in M_{m\times n}(S)$ and $fB := [f(b_{ij})] \in M_{n\times n}(S)$. Notice that since f is a ring morphism we have $(fA)(fB) = f(AB)$. Since we cannot have $(fA)(fB)$ and $(fB)(fA)$ equal to their respective identities, it follows that AB and BA are also not equal to the identity matrices—therefore R is satisfies the IBN property. ♮

Theorem 10.2.10. All commutative rings satisfy the invariant basis number property.

Proof. Let R be a commutative ring. By Proposition 9.2.8 we find that R contains a proper maximal ideal $\mathfrak{m}$—therefore the quotient ring $R/\mathfrak{m}$ is a field. Considering the canonical projection $R \twoheadrightarrow R/\mathfrak{m}$, we obtain, by Proposition 10.2.9, that R is IBN since fields are IBN ♮

Example 10.2.11. Let V be an *infinite dimensional* k-vector space, for some field k. The ring $R := \mathrm{End}_{\mathtt{Vect}_k}(V)$ of endomorphisms of V does *not* satisfy the IBN property.

First we prove that $\mathrm{End}_{\mathtt{Vect}_k}(V \oplus V) \simeq R^4$. Notice that any k-linear map $f: V \oplus V \to V \oplus V$ can be decomposed into its projections $f_1, f_2: V \oplus V \to V$. Further, each projection f_j can be again decomposed into unique k-linear morphisms $p_j, q_j: V \to V$ such that $f_j = p_j + q_j$—for $j \in \{1, 2\}$. Therefore, the map f is uniquely determined by the quadruple (p_1, q_1, p_2, q_2), where $f = (p_1 + q_1, p_2 + q_2)$. Hence we may map bijectively $\mathrm{End}_{\mathtt{Vect}_k}(V \oplus V) \to R^4$ via $f \mapsto (p_1, q_1, p_2, q_2)$.

Notice that, since V is infinite dimensional, we have $\dim_k V = \dim_k(V \oplus V)$—thus $V \simeq V \oplus V$, and hence $\mathrm{End}_{\mathtt{Vect}_k}(V \oplus V) \simeq \mathrm{End}_{\mathtt{Vect}_k}(V)$ (see Proposition 8.4.19). Notice that from our earlier result, we just concluded that $R \simeq R^4$, which proves that R does not satisfy the IBN property.

Example 10.2.12. Let $M := \prod_{j \in \mathbf{N}} \mathbf{Z}$ be a $\mathbf{Z}$-module and a ring $R := \mathrm{End}_{\mathbf{Z}\text{-}\mathtt{Mod}}(M)$. We'll show that R does not satisfy the IBN property.

Define parallel $\mathbf{Z}$-module morphisms $\phi, \psi \colon M \rightrightarrows M$ given by

$$\phi(a_1, a_2, \dots) := (a_1, a_3, a_5, \dots),$$
$$\psi(a_1, a_2, \dots) := (a_2, a_4, a_6, \dots).$$

Then, if $f \in R$ is any endomorphism, let $f_j \colon M \to \mathbf{Z}$ be its j-th projection. Define endomorphisms $g, h \colon M \rightrightarrows M$, whose projections are given by

$g_{2j-1} := f_j$ for *odd* $j \in \mathbf{Z}_{>0}$ and *zero* for the remaining projections,
$h_{2j} := f_j$ for *even* $j \in \mathbf{Z}_{>0}$ and *zero* for the remaining projections.

Then we obtain the equality

$$\begin{aligned}\phi g + \psi h &= (g_1, g_3, g_5, \dots) + (h_2, h_4, h_6, \dots)\\ &= (f_1, 0, f_3, 0, f_5, \dots) + (0, f_2, 0, f_4, 0 f_6, \dots)\\ &= f\end{aligned}$$

which is uniquely defined for f—hence $\{\phi, \psi\}$ is a basis for the *right*-R-module R (over itself). From this we conclude that $R \simeq R^2$ by mapping $f \mapsto (g, h)$ and with an inverse $(g, h) \mapsto \phi g + \psi h = f$. This shows that R does not satisfy the invariant basis number property.

Free Modules Basis

Lemma 10.2.13. Let R be an integral domain, and $M := R^{\oplus A}$ be a free R-module. Considering the inclusion $M \hookrightarrow \mathrm{Frac}(R)^{\oplus A}$, a subset $S \subseteq M$ is linearly independent in M if and only if it is linearly independent in $\mathrm{Frac}(R)^{\oplus A}$.

Proof. We do the proof of both statements via the contrapositive. Define the notation $V := \mathrm{Frac}(R)^{\oplus A}$. Suppose S is linearly independent in V, then there exists a collection of elements $(a_s/b_s)_{s \in S}$ for $a_s/b_s \in \mathrm{Frac}(R)$ non-zero only for finitely many $s \in S$—but not all zero—such that

$$\sum_{s \in S} \frac{a_s}{b_s} \cdot s = 0 \tag{10.1}$$

Consider the *finite* set $S' := \{s \in S : a_s \neq 0\}$. Since $b_s \neq 0$ for all $s \in S'$, one can consider the *non-zero* finite product $b := \prod_{s \in S'} b_s$. Notice that, for all $s_0 \in S$ we have

$$b \cdot \frac{a_{s_0}}{b_{s_0}} = \Bigg(\prod_{s \in S \setminus \{s_0\}} b_s\Bigg) a_{s_0},$$

which is simply an element of R since the denominator of the fraction is 1. We then define a collection $(c_s)_{s\in S}$ as $c_{s'} := \prod_{s\in S\setminus\{s'\}} b_s$ for all $s' \in S'$, and $c_s := 1$ for all $s \in S \setminus S'$. Notice that Eq. (10.1) is equivalent to

$$\sum_{s\in S}(c_s a_s)s = 0, \tag{10.2}$$

with coefficients $c_s a_s \in R$ for which finitely many are non-zero—but not all are zero—therefore Eq. (10.2) lies in M. From this we conclude that S is linearly dependent on M.

For the converse, suppose that S is linearly dependent on M, then there exists a collection $(r_s)_{s\in S}$—of elements $r_s \in R$ that are non-zero for only finitely many $s \in S$, and not all zero—such that

$$\sum_{s\in S} r_s s = 0. \tag{10.3}$$

From the inclusion $R \hookrightarrow \mathrm{Frac}(R)$, we see that $r_s \in \mathrm{Frac}(R)$ for all $s \in S$, and therefore Eq. (10.3) lies in both M and V—which proves that S is also linearly dependent in V. ♮

Proposition 10.2.14. Let R be an integral domain, and M be a *free* R-module. If B is a maximal linearly independent subset of M, and S is a linearly independent subset of M, then

$$|S| \leqslant |B|.$$

Moreover, if C is another maximal linearly independent subset of M, then

$$|C| = |B|.$$

Proof. By Lemma 10.2.13 it is equivalent to prove the proposition for a field $R = k$ and a k-vector space $M = V$.

We construct a map $\iota\colon S \to B$ inductively. Assume a well-ordering on the set S. We use transfinite induction—that is, assume ι is defined, injectively, for all $w < v$ for some $v \in S$, and define B' to consist of the elements of B but for all $w < v$ we replace $\iota(w) \in B$ by w. Our inductive hypothesis will be that B' is still a maximal linearly independent set of V.

We now define $\iota(v)$. Since by hypothesis the set B' is maximal, then $B' \cup \{v\}$ must be linearly dependent, so that there exists a collection $(a_j)_{j=0}^n$ of elements $a_j \in k$, not all zero, such that

$$a_0 v + a_1 b_1 + \cdots + a_n b_n = 0 \tag{10.4}$$

for some finite subset $(b_j)_{j=1}^n \subseteq B'$. Since B' is linearly independent, it must be the case that $a_0 \neq 0$. Moreover, since S is also linearly independent, it follows that not all b_j can be members of S. With a possible change of indexing, we may thus assume that $a_1 \neq 0$ and $b_1 \in B' \setminus S$—this implies that $b_1 \neq j(w)$ for all $w < v$, hence we may set $\iota(v) := b_1$ and not loose injectivity of ι.

Consider now the set B'', whose elements are those of B' but with $b_1 = \iota(v)$ replaced by v. Since B' is linearly independent, then $B'' \setminus v$ is linearly independent therefore, by

Eq. (10.4) we find $v = -\sum_{j=1}^{n} \frac{a_j}{a_0} b_j$, hence no subset of B''—containing or not v—can be linearly dependent, since that would imply the linear dependence of B'. We conclude that B'' is linearly independent and maximal, which finishes the transfinite induction and proves that the injective set-function $\iota\colon S \rightarrowtail B$ can be constructed. This shows that $|S| \leqslant |B|$. For the equality, we can simply do the construction of injective maps $C \rightarrowtail B$ and $B \rightarrowtail C$ and conclude that $|C| = |B|$. ♮

Corollary 10.2.15. Let R be an integral domain, and A and B be sets. Then there exists an isomorphism of R-modules $F_R A \simeq F_R B$ if and only if there exists a bijection $A \simeq B$.

Proof. If $F_R A \simeq F_R B$, then A and B are maximal linearly independent sets of the "same" module, which by Proposition 10.2.14 implies in $|A| = |B|$.

If $|A| = |B|$, let $f\colon A \xrightarrow{\simeq} B$ be a bijective set function, and consider the canonical inclusions $i_A\colon A \hookrightarrow F_R A$ and $i_B\colon B \hookrightarrow F_R B$. Consider the indexed sets $j_A := i_B f\colon A \to F_R B$ and $j_B := i_A f^{-1}\colon B \to F_R A$—both of which are injective, since are a composition of an injection with a bijection. The are unique induced R-module morphisms $\phi\colon F_R A \to F_R B$ and $\psi\colon F_R B \to F_R A$ are, by construction, such that $\phi i_A = j_A$ and $\psi i_B = j_B$. Consider the following commutative diagram in $\mathtt{Set}$:

$$\begin{tikzcd}
F_R A \arrow[r, dashed, "\phi"] & F_R B \arrow[r, dashed, "\psi"] & F_R A \arrow[r, dashed, "\phi"] & F_R B \\
A \arrow[u, hook, "i_A"] \arrow[r, "f"'] & B \arrow[u, hook, "i_B"] \arrow[r, "f^{-1}"'] & A \arrow[u, hook, "i_A"] \arrow[r, "f"'] & B \arrow[u, hook, "i_B"]
\end{tikzcd}$$

Using the universal property on the first two squares we find that $\psi\phi$ is the unique R-module morphism induced by i_A—but since $\mathrm{id}_{F_R A}$ also satisfies the property, it follows that $\psi\phi = \mathrm{id}_{F_R A}$. Analogously, applying the universal property on the last two squares we obtain $\phi\psi = \mathrm{id}_{F_R B}$. Therefore ϕ and ψ are inverses of each other and $F_R A \simeq F_R B$. ♮

Lemma 10.2.16. Let R be a commutative ring, and F be a free R-module with a basis B. If $\mathfrak{m} \subseteq R$ is a maximal ideal, the collection $(b + \mathfrak{m}F)_{b \in B}$ forms a basis of the vector space $F/(\mathfrak{m}F)$ over the field $R/\mathfrak{m}$[2].

Proof. We shall prove that $B/(\mathfrak{m}F) := (b + \mathfrak{m}F)_{b \in B}$ forms a minimal generating set of $F/(\mathfrak{m}F)$—so that by Lemma 5.6.17 we obtain that $B/(\mathfrak{m}F)$ is a basis. If $x + \mathfrak{m}F \in F/(\mathfrak{m}F)$ is any element, consider a representative $x \in F$ and let $(x_b)_{b \in B}$ be a collection of elements $x_b \in R$—which are non-zero only for finitely many $b \in B$—such that $x = \sum_{b \in B} x_b b$. Taking the natural projection $F \twoheadrightarrow F/(\mathfrak{m}F)$ of *sets*, we find that $x + \mathfrak{m}F = \big(\sum_{b \in B} x_b b\big) + \mathfrak{m}F$.

[2] Let R be a commutative ring. If $\mathfrak{a} \subseteq R$ is an ideal, and M is an R-module, we there exists a natural structure of $(R/\mathfrak{a})$-module that can be endowed on the abelian group $M/(\mathfrak{a}M)$ via the multiplication

$$(r + \mathfrak{a}) \cdot (m + \mathfrak{a}M) := rm + \mathfrak{a}M.$$

By the $(R/\mathfrak{m})$-module structure of $F/(\mathfrak{m}F)$, we find that

$$x + \mathfrak{m}F = \left(\sum_{b \in B} x_b b \right) + \mathfrak{m}F = \sum_{b \in B} x_b b + \mathfrak{m}F = \sum_{b \in B} (x_b + \mathfrak{m})(b + \mathfrak{m}F).$$

From construction, the collection $(x_b + \mathfrak{m})_{b \in B}$ is composed of finitely many non-zero elements $x_b + \mathfrak{m} \in R/\mathfrak{m}$. From this we conclude that $B/(\mathfrak{m}F)$ is indeed a generating set for $F/(\mathfrak{m}F)$.

How can I finish this proof? One can prove the linear independence or minimality of $B/(\mathfrak{m}F)$, but what is the easiest way out?

♮

Proposition 10.2.17. Let R be a non-zero commutative ring. If F is a free R-module and both A and B are basis of F, then $|A| = |B|$.

Proof. From Proposition 9.2.8 we know that R admits a maximal ideal $\mathfrak{m} \subsetneq R$. We know from Lemma 10.2.16 that the collections $(a + \mathfrak{m}F)_{a \in A}$ and $(b + \mathfrak{m}F)_{b \in B}$ are both basis for the $(R/\mathfrak{m})$-vector space $F/(\mathfrak{m}F)$. Since basis of vector spaces have the same cardinality, it follows that $|A| = |B|$. ♮

Definition 10.2.18 (Rank)**.** Let R be an integral domain. We define the *rank* of a free R-module M to be the cardinality of a maximal linearly independent subset of M—this cardinal shall be denoted $\operatorname{rank}_R M$. If $R = k$ is a field, then M is a k-vector space and thus $\operatorname{rank}_k M = \dim_k M$.

Proposition 10.2.19. Let R be an integral domain and M be an R-module. If M is generated by a subset $S \subseteq M$, then S *contains* a maximal linearly independent subset of M—hence $|S| \geqslant \operatorname{rank}_R M$.

Proof. By Lemma 10.2.13, we can treat the case for the field $k := \operatorname{Frac}(R)$ and a vector space $V := M$. Consider the power set 2^S of subsets of S, ordered by inclusion—in particular, let $\mathcal{S} \subseteq 2^S$ contain all linearly independent subsets of S. By Zorn's lemma, there exists a maximal linearly independent set $B \in \mathcal{S}$ of 2^S. Since maximal linearly independent sets form a basis (mind you, this is true only for vector spaces), then in particular S is contained in the span of B—therefore B is a maximal linearly independent set generating V, that is, B is a basis for V. ♮

Theorem 10.2.20 (Every module is the quotient of a free module)**.** Given a ring R, every right-R-module M (or left-R-module) is a *quotient* of a *free* right-R-module F (or left-R-module). Moreover, M is finitely generated if and only if one can choose a finitely generated free module F.

Proof. Define a free right-R-module $F := R^{\oplus |M|}$ and let $(x_m)_{m \in M}$ be a basis of F. By the universal property of free modules, the indexing of the basis of F induces a unique morphism of right-R-modules $p \colon F \to M$ such that $p(x_m) = m$ for all $m \in M$—which by construction is surjective. From the first isomorphism theorem we find that

$$M \simeq F/\ker g,$$

therefore M is indeed the quotient of a free module.

For the last proposition, if F can be chosen to be finitely generated, then it is immediate that M is finitely generated. For the converse, suppose that M is finitely generated and $M = \langle m_1, \dots, m_n \rangle$. Define the free R-module $F := R^{\oplus n}$ with a basis $\{x_1, \dots, x_n\}$ and consider the R-map $p\colon F \to M$ defined by $p(x_j) := m_j$ for each $1 \leqslant j \leqslant n$. Since $\langle p(x_1), \dots, p(x_n) \rangle = \langle m_1, \dots, m_n \rangle$, then p is surjective, therefore the statement holds true. ♮

Corollary 10.2.21. For any R-module M there exists an exact sequence of R-modules

$$G \longrightarrow F \twoheadrightarrow M \longrightarrow 0$$

for *free* R-modules G and F.

Direct Sums & Products of Free R-Modules

Proposition 10.2.22. Let $(M_j)_{j \in J}$ be a family of free R-modules. Then the direct sum $\bigoplus_{j \in J} M_j$ is free.

Proof. Let $(S_j)_{j \in J}$ be set of basis for each respective M_j, and define $S := \bigcup_{j \in J} S_j$. Let N be any R-module and $f\colon S \to N$ a set-function. Since M_j is free for all $j \in J$, there exists a unique morphism $\phi_j\colon M_j \to N$ such that the diagram

$$\begin{tikzcd} M_j \arrow[r, "\phi_j"] & N \\ S_j \arrow[u, hook] \arrow[ur, bend right, "f|_{S_j}"'] & \end{tikzcd}$$

commutes in `Set`. From the universal property of coproduct, the collection $(\phi_j)_{j \in J}$ defines a unique morphism $\phi\colon \bigoplus_{j \in J} M_j \to N$ such that the diagram

$$\begin{tikzcd} M_j \arrow[d, hook] \arrow[dr, bend left, "\phi_j"] & \\ \bigoplus_{j \in J} M_j \arrow[r, dashed, "\phi"] & N \end{tikzcd}$$

commutes in R-`Mod` for all $j \in J$. Therefore in particular

$$\begin{tikzcd} \bigoplus_{j \in J} M_j \arrow[r, "\phi"] & N \\ S \arrow[u, hook] \arrow[ur, bend right, "f"'] & \end{tikzcd}$$

commutes in `Set`—which implies that $\bigoplus_{j \in J} M_j$ is free. ♮

Remark 10.2.23. If M is a free R-module and N is a free submodule of M, it *does not follow* that the quotient M/N is free. For instance, take $M := \mathbf{Z}$ and the free submodule $N := 2\mathbf{Z}$, then M/N is not free since $\{1\}$ is not linearly independent in $\mathbf{Z}/2\mathbf{Z}$.

Interesting Examples

Example 10.2.24. Let R be an integral domain, and M be a free R-module. If $a \in R$ and $m \in M$ are such that $am = 0$, than either $a = 0$ or $m = 0$.

Let S be a basis for M and consider R as a module over itself. Let $f\colon S \rightarrowtail R$ be any injective *set-function*. Since $\iota\colon S \hookrightarrow M$ is injective, then the unique R-module morphism $\phi\colon M \to R$—making the diagram

$$\begin{tikzcd}
M \arrow[r, "\phi"] & R \\
S \arrow[u, tail, "\iota"] \arrow[ur, tail, bend right, "f"'] &
\end{tikzcd}$$

commute in `Set`—must be injective, therefore $\ker\phi = 0$. Notice that if $am = 0$ then since $\phi(0) = 0$ we have $\phi(am) = a\phi(m) = 0$ but since R is an integral domain, then either $a = 0$ or $\phi(m) = 0$—that is, $m \in \ker\phi$, which implies in $m = 0$.

Example 10.2.25. Let M be a free $\mathbf{Z}$-module, and $v \in M$ a non-zero element. Then there exists only finitely many $n \in \mathbf{Z}$ such that the equation $v = nx$ has a solution $x \in M$.

Let S be a basis for M, and let $v = \sum_{s \in S} a_s s$—where $a_s \in \mathbf{Z}$ is non-zero for only finitely many $s \in S$, but not all zero. Let $n \in \mathbf{Z}$ be an integer such that there exists a solution $x = \sum_{s \in S} b_s s \in M$ for $v = nx$, then since S is a basis it follows that

$$\sum_{s \in S} (a_s - nb_s)s = 0$$

implies in $a_s = nb_s$ for all $s \in S$. Hence n must be a divisor of all a_s—thus be necessarily have $n \leqslant \gcd(a_s)_{s \in S}$. Since the greatest common divisor is finite, then n can only assume a finite number of values for there to be a solution of $v = nx$ in M.

Remark 10.2.26. The direct product of a family of free modules *need not be free*. To see this, we construct the following example.

Let $M := \prod_{j \in \mathbf{N}} \mathbf{Z}$, we'll show that M is not a free $\mathbf{Z}$-module. Define the submodule $N := \bigoplus_{j \in \mathbf{N}} \mathbf{Z}$ of M. Suppose, for the sake of *contradiction*, that M admits a basis B.

Notice that since each element of N has only finitely many non-zero entries, it follows that for each $n \in \mathbf{N}$ the of elements with n non-zero entries can be represented by a sequence of pairs $(a_j, i_j)_{j=1}^n$ with $a_j \in \mathbf{Z}$ and a corresponding index $i_j \in \mathbf{N}$. Therefore, the number of elements with n non-zero entries is given by the finite product of countable sets $(\mathbf{Z} \times \mathbf{N})^n$—which is itself countable. Then the number of elements of N is given by the countable union $\bigcup_{n \in \mathbf{N}} (\bigcup_{j \in \mathbf{N}} \mathbf{Z}^n)$, which is again countable.

Let $B_0 \subseteq B$ be the collection of all basis elements that appear when expanding the elements of N in terms of B—such collection is necessarily countable since N itself is countable. The free module $N_0 := F_R B_0 \subseteq M$ will then contain N. Since B_0 is countable and the elements of N_0 are finite linear combinations of the elements of B_0, then N_0 is countable.

Consider the module $\overline{M} \coloneqq M/N_0$, has a basis $B \setminus B_0$. From Example 10.2.25, for every non-zero element $\overline{x} \in \overline{M}$ there exists only finitely many $n \in \mathbf{N}$ such that the equation $\overline{x} = nm$ admits a solution $m \in \overline{M}$.

Consider the subset

$$S \coloneqq \{(b_j)_{j \in \mathbf{N}} \in M : b_j = (\pm j!) a_j, \text{ with } a_j \in \mathbf{Z}\},$$

which has cardinality of the continuum, $2^{\aleph_0}$, hence S is uncountable. Therefore there exists $s \in S$ such that $s \notin N_0$. Notice however that, from the construction of S, every $n \in \mathbf{N}$ is such that the equation $s = nm$ has a solution $m \in M$—but since $s \notin N_0$, this also implies that $s = n\overline{m}$ has a solution $\overline{m} \in M/N_0$ for all $n \in \mathbf{N}$. This contradicts our last paragraph since M/N_0 is supposedly free! From this we conclude that M/N_0 must *not* be free, and M should not admit a basis. Thus M *is not a free* $\mathbf{Z}$*-module*.

10.3 Exact Sequences & Decompositions of Modules

Direct Summands

Definition 10.3.1 (Direct summand)**.** Let P be an R-module and M be a submodule of P. We say that M is a *direct summand* of P if there exists a submodule N of P satisfying $M \cap N = \{0\}$, called a *complement* of M, for which

$$P = M \oplus N.$$

Definition 10.3.2 (Retract)**.** Let M be an R-module. A submodule N of M is said to be a *retract* of M if there exists an R-module morphism $\rho \colon M \to N$ such that $\rho(n) = n$ for all $n \in N$—such morphism is called a *retraction*.

Proposition 10.3.3 (Direct summands & retractions)**.** Given R-modules P and M, the module M is a *direct summand* of P if and only if there exists a *retraction* $\rho \colon P \to M$, in this case $P \simeq M \oplus \ker \rho$.

Proof. Suppose the retract ρ exists. Let $p \in P$ be any element and consider the element $\rho(p) \in M$. Notice that $\rho(p - \rho(p)) = \rho(p) - \rho(\rho(p)) = \rho(p) - \rho(p) = 0$, thus $p - \rho(p) \in \ker \rho$. Since $p = \rho(p) + (p - \rho(p))$, then $P = M + \ker \rho$. Moreover, if $s \in S \cap \ker \phi$ it follows that $\rho(s) = s$ and $\rho(s) = 0$, therefore $s = 0$—thus $S \cap \ker \phi = \{0\}$, and $P = M \oplus \ker \rho$.

For the converse, suppose that M is a direct summand of P and let N be its complement. Since any element $p \in P$ can be written uniquely as a sum $p = m + n$ for $m \in M$ and $n \in N$, the map $\rho \colon P \to M$ given by $m + n \mapsto m$ is well defined and unique—also being an R-module morphism. Notice also that $\rho|_M = \mathrm{id}_M$. ♮

Corollary 10.3.4. If $P = M \oplus N$ and $M \subseteq A \subseteq P$, then

$$A = M \oplus (A \cap N).$$

Proof. Let $\rho \colon P \to M$ be a retraction, so that $\ker \rho = N$. Since $M \subseteq A$, then $\rho|_A$ is a retraction between A and M, and $\ker(\rho|_A) = A \cap N$. This shows that $A \cap N$ is a complement of M in A and hence $A = M \oplus (A \cap N)$. ♮

Corollary 10.3.5. Let M be an R-module, and $p\colon M \to M$ be an idempotent endomorphism (that is, $p^2 = p$). Then there exists a canonical isomorphisms

$$M \simeq \operatorname{im} p \oplus \ker p.$$

Proof. Notice that the induced map $\overline{p}\colon M \to \operatorname{im} p$ given by $\overline{p}(m) = p(m)$ is a morphism of R-modules such that $\overline{p}(\ell) = p(\ell) = \ell$ for all $\ell \in \operatorname{im} p$—since $p^2 = p$. This implies that $\overline{p}$ is a retraction and from Proposition 10.3.3 it follows that $M \simeq \operatorname{im} p \oplus \ker p$. ♮

Proposition 10.3.6. Let M be an R-module and $(M_j)_{j=1}^n$ be a family of R-modules. Then

$$M \simeq M_1 \oplus \cdots \oplus M_n$$

if and only if there exists a collection of R-module morphisms $(\phi_j\colon M \to M)_{j=1}^n$ such that $\operatorname{im} \phi_j \simeq M_j$ for all j, we have $\phi_i \phi_j = 0$ for each pair $i \neq j$, and $\sum_{j=1}^n \phi_j = \mathrm{id}_M$.

Proof. Suppose $\psi\colon M_1 \oplus \cdots \oplus M_n \xrightarrow{\simeq} M$ is an isomorphism, then every element of $m \in M$ can be uniquely written as a sum $m = \sum_{j=1}^n a_j \psi((\delta_{ij} m_j)_{i=1}^n)$ for $a_j \in R$ and $m_j \in M_j$. Define, for $1 \leqslant j \leqslant n$, a morphism $\phi_k\colon M \to M$ given by

$$\sum_{j=1}^n a_j \psi((\delta_{ij} m_j)_{i=1}^n) \overset{\phi_k}{\longmapsto} a_k \psi((\delta_{ik} m_k)_{i=1}^n).$$

Clearly one has $\phi_i \phi_j = 0$ for indices $i \neq j$, and $\phi_1 + \cdots + \phi_n = \mathrm{id}_M$. Also, the map $\Phi_j\colon \operatorname{im} \phi_j \to M_j$ given by $a_j \psi((\delta_{ij} m_j)_{i=1}^n) \mapsto a_j m_j$ establishes an isomorphism $\operatorname{im} \phi_j \simeq M_j$.

For the converse, suppose that the collection of morphisms $(\phi_j)_{j=1}^n$ exist and satisfy the required properties. Let N be any R-module, together with a family of morphisms $(\psi_j\colon N \to \operatorname{im} \phi_j)$. There exists a uniquely defined morphism $\eta\colon N \to M$ given by $n \mapsto \sum_{j=1}^n \psi_j(n)$ such that the following diagram

$$\begin{tikzcd}
N \arrow[dr, bend left, "\psi_j"] \arrow[d, dashed, "\eta"'] & \\
M \arrow[r, "\phi_j"'] & \operatorname{im} \phi_j \simeq M_j
\end{tikzcd}$$

for all $1 \leqslant j \leqslant n$. This shows that M satisfies the universal property for the product of the family $(M_j)_{j=1}^n$, hence $M \simeq M_1 \oplus \cdots \oplus M_n$. ♮

Exact Sequences

Just like in the category of groups or vector spaces, we define an exact sequence of modules as follows.

Definition 10.3.7 (Exact sequence)**.** A sequence $(d_n\colon M_n \to M_{n-1})_{n \in \mathbf{Z}}$ of R-module morphisms is said to form an *exact sequence*

$$\dots \longrightarrow M_{n+1} \xrightarrow{d_{n+1}} M_n \xrightarrow{d_n} M_{n-1} \longrightarrow \dots$$

if for all $n \in \mathbf{Z}$ we have $\operatorname{im} d_{n+1} = \ker d_n$. In particular, a sequence is said to be exact *in* M_m if $\operatorname{im} d_{m+1} = \ker d_m$—thus an exact sequence is exact in each of its modules.

Proposition 10.3.8. Let A, B and C be R-modules. The following are properties concerning exact sequences:

(a) A sequence

$$0 \longrightarrow A \xrightarrow{\ \phi\ } B$$

is exact if and only if ϕ is *injective*.

(b) A sequence

$$B \xrightarrow{\ \psi\ } C \longrightarrow 0$$

is exact if and only if ψ is *surjective*.

(c) A sequence

$$0 \longrightarrow A \xrightarrow{\ \kappa\ } B \longrightarrow 0$$

is exact if and only if κ is an *isomorphism*.

Proof. (a) If the sequence is exact then $\ker \phi = 0$ and ϕ is injective. For the converse, if ϕ is injective, then $\ker \phi = 0$ and since the image of the morphism $0 \to A$ must be zero, the exactness condition is satisfied.

(b) If the sequence is exact then $\operatorname{im} \psi = \ker(B \to 0) = B$ and ψ is surjective. If on the contrary we have ψ surjective, then again $\operatorname{im} \psi = B$, and since the kernel of the morphism $B \to 0$ is the whole module B, it follows that the sequence is exact.

(c) From the previous two items we have that κ is both injective and surjective, thus a bijection—and since bijective morphisms of R-modules are isomorphisms, the statement follows.

♮

Definition 10.3.9. An exact sequence of R-modules of the form

$$0 \longrightarrow A \rightarrowtail B \twoheadrightarrow C \longrightarrow 0$$

is said to be a *short exact sequence*. Moreover, we shall call such a sequence an *extension of A by C*—we may sometimes name B as the "extension".

Proposition 10.3.10 (Isomorphism theorems)**.** We now restate the isomorphism theorems of R-modules in terms of exact sequences:

(a) If $0 \to A \overset{f}{\rightarrowtail} B \overset{g}{\twoheadrightarrow} C \to 0$ is a *short exact sequence*, then there exists natural isomorphisms

$$A \simeq \operatorname{im} f \quad \text{and} \quad B/\operatorname{im} f \simeq C.$$

(b) Let M be an R-module, and S and T be both submodules of M. Then the following commutative diagram has short exact rows and the third vertical morphism is an isomorphism:

$$\begin{tikzcd}
0 \arrow[r] & S \cap T \arrow[r] \arrow[d] & S \arrow[r] \arrow[d] & \frac{S}{S\cap T} \arrow[r] \arrow[d, "\simeq"] & 0 \\
0 \arrow[r] & T \arrow[r] & S + T \arrow[r] & \frac{S+T}{T} \arrow[r] & 0
\end{tikzcd}$$

(c) Let M be an R-module, and consider submodules S and T of M such that $S \subseteq T$. Then there exists a short exact sequence

$$0 \longrightarrow S/T \overset{\iota}{\rightarrowtail} M/T \overset{\pi}{\twoheadrightarrow} M/S \longrightarrow 0$$

Proof. (a) Since f is an injection, the induced morphism $f\colon A \xrightarrow{\simeq} \operatorname{im} f$ is an isomorphism. Moreover, since $\ker g = \operatorname{im} f$ and g is a surjection, by the first isomorphism theorem we have $B/\operatorname{im} f \simeq C$.

(b) By the second isomorphism theorem, the mapping $S/(S \cap T) \to (S+T)/T$ given by $s + S \cap T \mapsto s + T$ is an isomorphism.

(c) Simply define ι as the inclusion $S/T \hookrightarrow M/T$—so that $\operatorname{im}\iota = S/T$—and π as the natural projection $M/T \twoheadrightarrow M/S$ mapping $m + T \mapsto m + S$. Since π is surjective and, as argued in Proposition 8.4.52, $\ker \pi = S/T$. Therefore $\ker\pi = \operatorname{im}\iota$—the sequence is exact.

♮

Proposition 10.3.11. Let $0 \to M_j \overset{\alpha_j}{\rightarrowtail} N_j \overset{\beta_j}{\twoheadrightarrow} L_j \to 0$ be a short exact sequence of R-modules for all $1 \leqslant j \leqslant n$. Then the induced sequence

$$0 \longrightarrow \bigoplus_{j=1}^n M_j \overset{\alpha}{\rightarrowtail} \bigoplus_{j=1}^n N_j \overset{\beta}{\twoheadrightarrow} \bigoplus_{j=1}^n L_j \longrightarrow 0$$

Proof.
- We first prove the injectivity of α. Notice that $(m_j)_j \in \ker\alpha$ if and only if $\alpha_j(m_j) = 0$—which implies in $m_j = 0$ since α_j is injective—therefore $(m_j)_j = 0$. This implies in $\ker\alpha = 0$, hence α is injective.
- For the surjectivity of β, let $(\ell_j)_j \in \bigoplus_{j=1}^n L_j$ be any element. Since each β_j is surjective, there exists $n_j \in N_j$ such that $\beta_j(n_j) = \ell_j$. Now, if we consider the element $(n_j)_j \in \bigoplus_{j=1}^n N_j$, we obtain $\beta(n_j)_j = (\beta_j(n_j))_j = (\ell_j)_j$.
- For the exactness in $\bigoplus_{j=1}^n N_j$ we prove that $\operatorname{im}\alpha = \ker\beta$. Let $(m_j)_j \in \bigoplus_{j=1}^n$ be any element, then by definition $\alpha(m_j)_j = (\alpha_j(m_j))_j := (n_j)_j$. Since each $n_j \in \operatorname{im}\alpha_j$ and $\operatorname{im}\alpha_j \subseteq \ker\beta_j$, it follows that $\beta_j(n_j) = 0$. Then $\beta(n_j)_j = (\beta_j(n_j))_j = 0$, which implies in $\operatorname{im}\alpha \subseteq \ker\beta$.

  For the converse inclusion, let $(n_j)_j \in \ker\beta$ be any element— then in particular we must have $n_j \in \ker\beta_j$. Since $\ker\beta_j \subseteq \operatorname{im}\alpha_j$ for all $1 \leqslant j \leqslant n$, then there exists

$m_j \in M_j$ such that $\alpha_j(m_j) = n_j$. Therefore one obtains $\alpha(m_j)_j = (\alpha_j(m_j))_j = (n_j)_j$, which shows that $\ker \beta \subseteq \operatorname{im} \alpha$. Thus $\operatorname{im} \alpha = \ker \beta$.

♮

Lemma 10.3.12. Consider the diagram

$$\begin{tikzcd}
0 \arrow[dr] & & & & 0 \\
& A \arrow[dr, tail, "\alpha"] & & K \arrow[ur] & \\
& & X \arrow[ur, "\kappa"] \arrow[dr, two heads, "\beta"] & & \\
& D \arrow[ur, tail, "\delta"] & & B \arrow[dr] & \\
0 \arrow[ur] & & & & 0
\end{tikzcd}$$

composed of short exact sequences of R-modules. Then $\kappa\alpha \colon A \to K$ is an isomorphism if and only if $\beta\delta \colon D \to B$ is an isomorphism.

Proof. From symmetry of the sequences we simply prove the forward implication. That is, suppose $\kappa\alpha$ is an isomorphism. For injectivity, let $d \in \ker \beta\delta$ be any element then $\delta(d) \in \ker \beta$. By exactness, there exists $a \in A$ such that $\alpha(a) = \delta(d)$. However, since $\delta(d) \in \ker \kappa$, then $\kappa\alpha(a) = 0$—and by the injectivity of $\kappa\alpha$ we find that $a = 0$, implying in $\delta(d) = 0$. Since δ is also injective, then $d = 0$, therefore $\ker \beta\delta = 0$, proving that $\beta\delta$ is injective.

For surjectivity, take any $b \in B$ and, since β is surjective, let $x \in X$ be such that $\beta(x) = b$. Consider the element $\kappa(x) \in K$—since $\kappa\alpha$ is surjective, there exists $a \in A$ for which $\kappa\alpha(a) = \kappa(x)$, thus $x - \alpha(a) \in \ker \kappa$. From exactness there must exist $d \in D$ for which $\delta(d) = x - \alpha(a)$. With this in hands we obtain

$$\beta\delta(d) = \beta(x - \alpha(a)) = \beta(x) - \beta\alpha(a) = \beta(x) = b,$$

proving that $\beta\delta$ is surjective. ♮

Split Exact Sequences

Definition 10.3.13 (Split sequence)**.** A short exact sequence

$$0 \longrightarrow A \overset{\iota}{\rightarrowtail} B \overset{p}{\twoheadrightarrow} C \longrightarrow 0$$

is said to be *split* if p is a split epimorphism—that is, there exists a morphism of R-modules (a section of p) $s \colon C \rightarrowtail B$ such that $ps = \mathrm{id}_C$.

Proposition 10.3.14 (Split sequence extension)**.** If the short exact sequence

$$0 \longrightarrow A \overset{\iota}{\rightarrowtail} B \overset{p}{\twoheadrightarrow} C \longrightarrow 0$$

is split, then there exists a natural isomorphism of R-modules

$$B \simeq A \oplus C.$$

Proof. Since the sequence is split, let $s\colon C \to B$ be the section of p. We'll show that $B = \operatorname{im} \iota \oplus \operatorname{im} s$. Let $b \in B$ be any element, then

$$\begin{aligned} p(b - sp(b)) &= p(b) - p(sp(b)) \\ &= p(b) - (ps)p(b) \\ &= p(b) - \mathrm{id}_C\, p(b) \\ &= 0, \end{aligned}$$

therefore $b - sp(b) \in \ker p$. Since the sequence is exact, $\operatorname{im} \iota = \ker p$, there exists $a \in A$ such that $\iota(a) = b - sp(b)$—thus $b = \iota(a) + sp(b)$. From this we conclude that $B = \operatorname{im} \iota + \operatorname{im} s$. For this to be a direct summand, it remains to prove that the intersection of the images is empty. Suppose that $x \in B$ is common to both images, so that there exists $a \in A$ and $c \in C$ such that $\iota(a) = x = s(c)$. Post-composing with p we get $p(x) = p\iota(a) = 0$, thus $x \in \ker p$. Moreover, $ps(c) = 0$ but since $ps = \mathrm{id}_C$, then $c = 0$—thus $s(c) = 0$ and $x = 0$. We conclude that

$$B = \operatorname{im} \iota \oplus \operatorname{im} s,$$

but $\operatorname{im} \iota \simeq A$ and $\operatorname{im} s \simeq C$—since both are injective maps—then $B \simeq A \oplus C$. Notice that this could also be extracted as a consequence of Corollary 10.3.5. ♮

Remark 10.3.15. It should be emphasized that the converse of Proposition 10.3.14 *does not hold* in general. For instance, consider cyclic free groups $A := \langle a \rangle$, $B := \langle b \rangle$ and $C := \langle c \rangle$—where both A and C have order 2, while B has order 4. Endowing such groups with the structure of $\mathbf{Z}$-modules, we can define morphisms $\iota\colon A \to B$ mapping $a \mapsto 2b$, and $p\colon B \to C$ with $b \mapsto c$, then the sequence $0 \to A \xrightarrow{\iota} B \xrightarrow{p} C \to 0$ is exact

Continue this, I can't continue today

Proposition 10.3.16 (Equivalent definition for split sequences)**.** A short exact sequence $0 \to A \xrightarrow{\iota} B \xrightarrow{p} C \to 0$ is split if and only if ι is a split monomorphism—that is, there exists a morphism $r\colon B \to A$ such that $r\iota = \mathrm{id}_A$. Moreover, if $s\colon C \to B$ is the section of p, then

$$sp + \iota r = \mathrm{id}_B .$$

Proof. Suppose the sequence splits and let $s\colon C \rightarrowtail B$ be a section of p, then by Proposition 10.3.14 we have $B \simeq \operatorname{im} \iota \oplus \operatorname{im} s$. Since ι and s are both injective maps, for every $b \in B$ there exists a unique pair $a \in A$ and $c \in C$ such that $b = \iota(a) + s(c)$. We define a map $r\colon B \to A$ given by $r(b) = r(\iota(a) + s(c)) := a$. Therefore, for all $a \in A$, we have $r\iota(a) = a$—thus r is a retract of ι.

Suppose, for the converse, that ι is a split monomorphism with a retract $r\colon B \twoheadrightarrow A$. Notice that

$$(\iota r)^2 = (\iota r)(\iota r) = \iota(r\iota)r = \iota \operatorname{id}_B r = \iota r,$$

therefore $\iota r\colon B \to B$ is an idempotent endomorphism. From Corollary 10.3.5 we obtain that $B = \ker(\iota r) \oplus \operatorname{im}(\iota r)$. Since r is surjective, then $\operatorname{im}(\iota r) = \operatorname{im} \iota$. Moreover, since p is surjective, given any $c \in C$, there exists $b \in B$ such that $p(b) = c$. From the decomposition of B, there exists $k \in \ker(\iota r)$ and $a \in A$ such that $b = k + \iota(a)$, therefore

$$c = p(b) = p(k + \iota(a)) = p(k) + p\iota(a) = p(k),$$

since $\operatorname{im} \iota = \ker p$. Thus we may define a map $s\colon C \to B$ as $s(c) = s(p(k)) := k$, so that $sp = \operatorname{id}_B$. We conclude that p is a split epimorphism, which implies that the sequence is split.

For the second statement, notice that

$$\begin{aligned}
(sp + \iota r)(b) &= (sp + \iota r)(\iota(a) + s(c)) \\
&= (sp)f(a) + (sp)s(c) + (\iota r)\iota(a) + (\iota r)s(c) \\
&= s(pf(a)) + s(ps(c)) + \iota(r\iota(a)) + \iota(rs(c)) \\
&= \iota(a) + s(c),
\end{aligned}$$

therefore $sp + \iota r = \operatorname{id}_B$. ♮

Proposition 10.3.17. Let $0 \to A \overset{f}{\rightarrowtail} B \overset{g}{\twoheadrightarrow} F \to 0$ be a short exact sequence of R-modules. If F is a *free* R-module, then the sequence is *split*. That is, every short exact sequence ending with a free module is split.

Proof. Let $(e_j)_{j \in J}$ be a basis for F. Since g is surjective, let $(b_j)_{j \in J}$ be a collection such that $g(b_j) = e_j$. By the free module universal property, define the unique morphism $\rho\colon F \to B$ mapping $\rho(e_j) \mapsto b_j$ for each $j \in J$. Then notice that $g\rho(e_j) = g(b_j) = e_j$, thus we can again use the the uniqueness of the morphism of the universal property of free modules to obtain that $g\rho = \operatorname{id}_F$. Therefore we conclude that the sequence is split. ♮

Example 10.3.18. Every short exact sequence of vector spaces is split. Indeed, any vector space admits a basis, which implies that a vector space is free—hence by Proposition 10.3.17 we obtain the proposition.

Proposition 10.3.19. If $0 \to A \overset{f}{\rightarrowtail} B \overset{g}{\twoheadrightarrow} C \to 0$ is a short exact sequence of R-modules, then the following are properties concerning finite generation:

(a) If both A and C are finitely generated modules, then B is finitely generated.

(b) If B is finitely generated, then C is finitely generated.

Proof. (a) Let $A := \langle a_1, \dots, a_n \rangle$ and $C := \langle c_1, \dots, c_m \rangle$. Define $(b'_j)_{j=1}^m$ be a collection of elements $b_j \in g^{-1}(c_j)$—which is ensured to exist by the surjectivity of g—and $(b_j)_{j=1}^n$ to be the collection $b_j := f(a_j)$. We'll show that $X := (b_1, \dots, b_n, b'_1, \dots, b'_m)$ is a generating set for B. Let $b \in B$ be any element, then since C is finitely generated, we can write

$$g(b) = \sum_{j=1}^m c_j r'_j = \sum_{j=1}^m g(b'_j) r'_j = \sum_{j=1}^m g(b'_j r'_j) = g\Big(\sum_{j=1}^m b'_j r'_j\Big)$$

for some collection of elements $r'_j \in R$. Hence $x := b - \sum_{j=1}^m b'_j r'_j \in \ker g$. Notice however that, since A is generated by $(a_j)_{j=1}^n$, then in particular $\operatorname{im} f = \ker g$ is generated by $(b_j)_{j=1}^n$. Therefore there exists a collection $(r_j)_{j=1}^n$ of elements $r_j \in R$ such that

$$b = \sum_{j=1}^m b'_j r'_j + \sum_{j=1}^n b_j r_j.$$

This shows that the finite set X—whose cardinality is $m + n$—generates the R-module B, proving the proposition.

(b) For item (b), one may notice that since the sequence is exact, then in particular $g \colon B \twoheadrightarrow C$ is surjective. Moreover, since B is finitely generated, we let $B = \langle b_1, \dots, b_k \rangle$. Given any $c \in C$ one has $b \in B$ such that $g(b) = c$. By the generating property, we have $b = \sum_{j=1}^k b_j r_j$ for a collection $(r_j)_{j=1}^k$ of ring elements $r_j \in R$. Therefore, by the morphism property of g we obtain

$$c = g(b) = g\Big(\sum_{j=1}^k b_j r_j\Big) = \sum_{j=1}^k g(b_j) r_j.$$

This shows that C is finitely generated by the collection $(g(b_j))_{j=1}^k$.

♮

Morphisms of Exact Sequences

Definition 10.3.20 (Morphisms of short exact sequences)**.** A morphism between short exact sequences of R-modules $0 \to A \to B \to C \to 0$ and $0 \to X \to Y \to Z \to 0$ is a triple (α, β, γ), of morphisms of R-modules making the following diagram commute

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r, tail] \arrow[d, "\alpha"'] & B \arrow[r, two heads] \arrow[d, "\beta"] & C \arrow[r] \arrow[d, "\gamma"] & 0 \\
0 \arrow[r] & X \arrow[r, tail] & Y \arrow[r, two heads] & Z \arrow[r] & 0
\end{tikzcd}$$

Definition 10.3.21 (Equivalent sequences)**.** Two short exact sequences of R-modules $0 \to A \rightarrowtail B \twoheadrightarrow C \to 0$ and $0 \to X \rightarrowtail Y \twoheadrightarrow Z \to 0$ are said to be *equivalent* if $A = X$,

$C = Z$ and there exists an isomorphism of R-modules $BisoY$ such that the diagram

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r, tail] \arrow[d, equal] & B \arrow[r, two heads] \arrow[d, "\simeq"'] & C \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & X \arrow[r, tail] & Y \arrow[r, two heads] & Z \arrow[r] & 0 \end{tikzcd}$$

commutes in R-$\texttt{Mod}$.

Proposition 10.3.22. Let $A \xrightarrow{\alpha} B \overset{\beta}{\twoheadrightarrow} C \to 0$ and $X \xrightarrow{\chi} Y \overset{\gamma}{\twoheadrightarrow} Z \to 0$ be exact sequences. If there exists a surjective morphism $f\colon A \twoheadrightarrow X$ and an isomorphism $g\colon B \xrightarrow{\simeq} Y$ such that $g\alpha = \chi f$, then there exists a *unique isomorphism* $h\colon C \to Z$ making the diagram

$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[d, two heads, "f"'] & B \arrow[r, two heads, "\beta"] \arrow[d, "g"', "\simeq"] & C \arrow[r] \arrow[d, dashed, "h"', "\simeq"] & 0 \\ X \arrow[r, "\chi"'] & Y \arrow[r, two heads, "\gamma"'] & Z \arrow[r] & 0 \end{tikzcd}$$

commutative in R-$\texttt{Mod}$.

Proof. Let $c \in C$ be any element. Since β is surjective, there exists $b \in B$ such that $\beta(b) = c$, thus we may define $h\colon C \to Z$ as the mapping $c \mapsto \gamma g(a)$. To see that h is well defined, consider $b' \in \beta^{-1}(c)$, then $\beta(b - b') = 0$ and since $\ker\beta = \operatorname{im}\alpha$, there exists $a \in A$ such that $\alpha(a) = b - b'$. Therefore

$$\gamma g(b) - \gamma g(b') = \gamma g(b - b') = \gamma g\alpha(a) = \gamma\chi f(a) = 0$$

since $\operatorname{im}\chi = \ker\gamma$—thus h is indeed well defined, is a R-module morphism, and makes the diagram commute. For the uniqueness of h, suppose $h'\colon C \to Z$ is another morphism making the diagram commute—that is, $h'\beta = \gamma g$. For any $c \in C$, let $b \in \beta^{-1}(c)$ and notice that

$$h'(c) = h'\beta(b) = \gamma g(b) = h\beta(b) = h(b),$$

therefore $h' = h$.

We now show that h is an isomorphism. Let $c \in \ker h$ be any element, and $b \in \beta^{-1}(c)$, then $0 = h\beta(b) = \gamma g(b)$—therefore $g(b) \in \ker\gamma = \operatorname{im}\chi$, thus there exists $x \in X$ such that $\chi(x) = g(b)$. From the surjectivity of f, there exists $a \in A$ such that $f(a) = x$. Since the first square is commutative,

$$g\alpha(a) = \chi f(a) = \chi(x) = g(b),$$

but since g is injective then $\alpha(a) = b$. Therefore we conclude that

$$c = \beta(b) = \beta\alpha(a) = 0,$$

since $\ker\beta = \operatorname{im}\alpha$. Hence $\ker h = 0$ and h is injective.

For the surjectivity of h, let $z \in Z$ be any element. Since γ is surjective, let $y \in Y$ be such that $\gamma(y) = z$, then from the surjectivity of g, we let $b \in B$ be an element such that $g(b) = y$, then from the commutativity of the second square we obtain

$$h\beta(b) = \gamma g(b) = \gamma(y) = z$$

thus $\beta(b) \in C$ has image z under h, and h is therefore surjective. $\natural$

Proposition 10.3.23. Let $0 \to A \overset{\alpha}{\rightarrowtail} B \overset{\beta}{\to} C$ and $0 \to X \overset{\chi}{\rightarrowtail} Y \overset{\gamma}{\to} Z$ be exact sequences. If there exists morphisms $g\colon B \to Y$ and $h\colon C \to Z$ such that $\gamma g = h\beta$, then there exists a *unique* morphism $f\colon A \to X$ such that

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r, tail, "\alpha"] \arrow[d, dashed, "f"'] & B \arrow[r, "\beta"] \arrow[d, "g"'] & C \arrow[d, "h"'] \\
0 \arrow[r] & X \arrow[r, tail, "\chi"'] & Y \arrow[r, "\gamma"'] & Z
\end{tikzcd}$$

commutes in R-$\texttt{Mod}$. Moreover, if both g and h are isomorphisms, then f is an *isomorphism*.

Proof. Let $a \in A$ be any element and define $\alpha(a) := b$. Since the top row is exact we have $b \in \ker\beta$. From the commutativity of the second square, we find

$$\gamma g(b) = h\beta(b) = h(0) = 0$$

therefore $g(b) \in \ker\gamma$. Since the bottom row is exact, there must exist $x \in X$ such that $\chi(x) = g(b)$. Define a map $f\colon A \to X$ by sending $a \mapsto x$ as described above. From the injectivity of α and χ, one has that $\alpha^{-1}(b) = \{a\}$ and $\chi^{-1}(g(b)) = \{x\}$, therefore f is well defined. If $a, a' \in A$, and $r \in R$ are any elements, defining $\alpha(a) := b$ and $\alpha(a') := b'$ $\alpha(ar + a') = \alpha(a)r + \alpha(a') = br + b'$. Moreover, $g(b), g(b') \in \ker\gamma$ thus there exists $x, x' \in X$ such that $\chi(x) = g(b)$ and $\chi(x') = g(b')$. Therefore

$$f(ar + a') = xr + x' = f(a)r + f(a'),$$

which shows that f is a morphism of R-modules. For the commutativity of the first square we have, for any $a \in A$,

$$g\alpha(a) = g(b) = \chi(x) = \chi f(a).$$

For the second statement, suppose that both g and h are isomorphisms. Let $a \in \ker f$ be any element, then if $\alpha(a) = b$ we obtain by commutativity of the first square that

$$g\alpha(a) = g(b) = \chi f(a) = \chi(0) = 0,$$

but since g is injective, then $b = 0$. Therefore $\alpha(a) = 0$, which implies in $a = 0$, since α is injective. This shows that $\ker f = 0$ and thus f is injective. For surjectivity, let $x \in X$ be any element, and let $y := \chi(x)$ so that $y \in \ker\gamma$ by the exactness of the bottom row. Since g is surjective, let $b \in B$ such that $g(b) = y$. From the commutativity of the second square we have

$$h\beta(b) = \gamma g(b) = \gamma(y) = 0,$$

therefore $\beta(b) \in \ker h$. Since h is injective, then $\beta(b) = 0$ and thus $\beta \in \operatorname{im}\alpha$ from the exactness of the top row. Let $a \in A$ be such that $\alpha(a) = b$, then from the commutativity of the first square we get

$$\chi f(a) = g\alpha(a) = g(b) = y.$$

Since χ is injective and $\chi(x) = y$, then $f(a) = x$. Therefore f is an isomorphism. ♮

Proposition 10.3.24 (Five lemma)**.** Consider the following commutative diagram in R-`Mod`, whose *rows are exact*:

$$\begin{tikzcd}
A_1 \arrow[r] \arrow[d, "\phi_1"] & A_2 \arrow[r] \arrow[d, "\phi_2"] & A_3 \arrow[r] \arrow[d, "\phi_3"] & A_4 \arrow[r] \arrow[d, "\phi_4"] & A_5 \arrow[d, "\phi_5"] \\
B_1 \arrow[r] & B_2 \arrow[r] & B_3 \arrow[r] & B_4 \arrow[r] & B_5
\end{tikzcd}$$

The following properties hold:

(a) If ϕ_2 and ϕ_4 are *surjective* and ϕ_5 is *injective*, then ϕ_3 is *surjective*.

(b) If ϕ_2 and ϕ_4 are *injective* and ϕ_1 is *surjective*, then ϕ_3 is *injective*.

(c) If ϕ_1, ϕ_2, ϕ_4 and ϕ_5 are *isomorphisms*, then ϕ_3 is an *isomorphism*.

Proof. Let $\alpha_j \colon A_j \to A_{j+1}$ and $\beta_j \colon B_j \to B_{j+1}$ for $1 \leqslant j \leqslant 4$ be the morphisms shown in the diagram. We prove each item:

(a) Let $b_3 \in B_3$ be any element, and define $b_4 := \beta_3(b_3)$. Since ϕ_4 is surjective, there exists $a_4 \in A_4$ such that $\phi_4(a_4) = b_4$. From the exactness of the bottom row we have $b_4 \in \ker\beta_4$, since $b_4 \in \operatorname{im}\beta_3$. By the commutativity of the forth square we have

$$\phi_5\alpha_4(a_4) = \beta_4\phi_4(a_4) = \beta_4(b_4) = 0$$

therefore $\alpha_4(a_4) \in \ker\phi_5$. Since ϕ_5 is injective, then $a_4 \in \ker\alpha_4$. By the exactness of the top row, there exists $a_3 \in A_3$ such that $\alpha_3(a_3) = a_4$. From the commutativity of the third square we get

$$\beta_3\phi_3(a_3) = \phi_4\alpha_3(a_3) = \phi_4(a_4) = b_4.$$

Since $\beta_3(b_3) = b_4 = \beta_3\phi_3(a_3)$, then $\phi_3(a_3) - b_3 \in \ker\beta_3$. Using again the exactness of the bottom row, there exists $b_2 \in B_2$ such that $\beta_2(b_2) = \phi_3(a_3) - b_3$. Since ϕ_2 is surjective, there exists $a_2 \in A_2$ such that $\phi_2(a_2) = b_2$. From the commutativity of the second square one has

$$\phi_3\alpha_2(a_2) = \beta_2\phi_2(a_2) = \beta_2(b_2) = \phi_3(a_3) - b_3,$$

which implies in $\phi_3(a_3 - \alpha_2(a_2)) = b_3$. This proves the surjectivity of ϕ_3.

(b) Let $a_3 \in \ker \phi_3$ be any element. By the commutativity of the third square we obtain

$$\phi_4 \alpha_3(a_3) = \beta_3 \phi_3(a_3) = \beta_3(0),$$

thus $\alpha_3(a_3) \in \ker \phi_4$—but since ϕ_4 is injective, then $a_3 \in \ker \alpha_3$. From the exactness of the top row, there exists $a_2 \in A_2$ such that $\alpha_2(a_2) = a_3$. By the commutativity of the second square we have

$$\beta_2 \phi_2(a_2) = \phi_3 \alpha_2(a_2) = \phi_3(a_3) = 0,$$

hence $\phi_2(a_2) \in \ker \beta_2$. From the exactness of the bottom row there must exist $b_1 \in B_1$ such that $\beta_1(b_1) = \phi_2(a_2)$. Since ϕ_1 is surjective, let $a_1 \in A_1$ be such that $\phi_1(a_1) = b_1$. Using the commutativity of the first square we obtain

$$\phi_2 \alpha_1(a_1) = \beta_1 \phi_1(a_1) = \beta_1(b_1) = \phi_2(a_2),$$

therefore $\phi_2(\alpha_1(a_1) - a_2) = 0$. Since ϕ_2 is injective, then $\alpha_1(a_1) = a_2$. From exactness of the top row, we have $a_2 \in \ker \alpha_2$, but since $\alpha_2(a_2) = a_3$, then $a_3 = 0$. This shows that $\ker \phi_3 = 0$, thus ϕ_3 is injective.

(c) This last item is a direct consequence of the above items (a) and (b).

$\natural$

Proposition 10.3.25 (3 × 3 lemma)**.** Consider the following commutative diagram in R-`Mod`, whose *columns are exact*:

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & A \arrow[r, "\alpha"] \arrow[d, tail, "f_1"'] & B \arrow[r, "\beta"] \arrow[d, tail, "g_1"'] & C \arrow[r] \arrow[d, tail, "h_1"] & 0 \\
0 \arrow[r] & X \arrow[r, "\chi"] \arrow[d, two heads, "f_2"'] & Y \arrow[r, "\gamma"] \arrow[d, two heads, "g_2"'] & Z \arrow[r] \arrow[d, two heads, "h_2"] & 0 \\
0 \arrow[r] & L \arrow[r, "\lambda"'] \arrow[d] & S \arrow[r, "\sigma"'] \arrow[d] & D \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

Then, the following properties hold:

(a) If the *bottom two rows are exact*, then the *top row is exact*.

(b) If the *top two rows are exact*, then the *bottom row is exact*.

Proof. Let's diagram chase!

(a) First we show that $\operatorname{im}\alpha \subseteq \ker\beta$. Let $a \in A$ be any element and define $b := \alpha(a)$, $x := f_1(a)$ and $y := \chi(x)$. From the commutativity of the top left square:

$$g_1\alpha(a) = \chi f_1(a) = \chi(x) = y,$$

thus $g_1(b) = y$. Since $y \in \operatorname{im}\chi$, by the exactness of the middle row we obtain $y \in \ker\gamma$. From the commutativity of the top right square:

$$h_1\beta(b) = \gamma g_1(b) = \gamma(y) = 0,$$

therefore $\beta(b) \in \ker h_1$, and since h_1 is injective, we conclude that $\beta(b) = 0$. Therefore $\operatorname{im}\alpha \subseteq \ker\beta$ as wanted.

For the final part, we show that $\ker\beta \subseteq \operatorname{im}\alpha$. Let $b \in \ker\beta$ be any element. From the commutativity of the top right square:

$$\gamma g_1(b) = h_1\beta(b) = h_1(0) = 0$$

therefore $g_1(b) \in \ker\gamma$. From the exactness of the middle row there exists $x \in X$ such that $\chi(x) = g_1(b)$. From both the commutativity of the bottom left square and the exactness of the middle row:

$$\gamma f_2(x) = g_2\chi(x) = g_2g_1(b) = 0,$$

hence $f_2(x) \in \ker\gamma$, but since γ is injective we conclude that $x \in \ker f_2$. From the exactness of the left column there exists $a \in A$ such that $f_1(a) = x$. By the commutativity of the top left square:

$$g_1\alpha(a) = \chi f_1(a) = \chi(x) = y.$$

Since g_1 is injective and $g_1(b) = y$, then $\alpha(a) = b$. Thus $\ker\beta \subseteq \operatorname{im}\alpha$.

(b) We show that $\operatorname{im}\lambda \subseteq \ker\sigma$. Let $\ell \in L$ be any element and define $s := \gamma(\ell)$. Since f_2 is surjective, let $x \in X$ be such that $f_2(x) = \ell$. From the exactness of the middle row we have $y := \chi(x) \in \ker\gamma$. By the commutativity of the bottom left square:

$$g_2\chi(x) = \lambda f_2(x) = \lambda(\ell) = s,$$

therefore $g_2(y) = s$. Now using the commutativity of the bottom right square:

$$\sigma g_2(y) = h_2\gamma(y) = h_2(0) = 0,$$

then $\sigma(s) = 0$, hence $\operatorname{im}\lambda \subseteq \ker\sigma$.

Finally, we show that $\ker\sigma \subseteq \operatorname{im}\lambda$. Let $s \in \ker\sigma$ be any element. Since g_2 is surjective, let $y \in Y$ be such that $g_2(y) = s$. From the commutativity of the bottom right square:

$$h_2\gamma(y) = \sigma g_2(y) = \sigma(s) = 0,$$

therefore $\gamma(y) \in \ker h_2$. From the exactness of the right column, there exists $c \in C$ such that $h_1(c) = \gamma(y)$. Since β is surjective, there exists $b \in B$ for which $\beta(b) = c$. Using the commutativity of the top right square:

$$\gamma g_1(b) = h_1\beta(b) = h_1(c) = \gamma(y),$$

hence $\gamma(y - g_1(b)) = 0$. By the exactness of the middle row, we can find $x \in X$ such that $\chi(x) = y - g_1(b)$. Applying the commutativity of the bottom left square:

$$\lambda f_2(x) = g_2\chi(x) = g_2(y - g_1(b)) = g_2(y) - g_2g_1(b) = g_2(y) = s,$$

therefore $s \in \operatorname{im}\lambda$. Thus indeed $\ker\sigma \subseteq \operatorname{im}\lambda$.

♮

Proposition 10.3.26. Consider the following commutative diagram in R-`Mod`:

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r, "\alpha"] \arrow[d, "\phi_1", "\simeq"'] & B \arrow[r, "\beta"] \arrow[d, "\phi_2", "\simeq"'] & C \arrow[r] \arrow[d, "\phi_3"', "\simeq"] & 0 \\
0 \arrow[r] & X \arrow[r, "\chi"'] & Y \arrow[r, "\gamma"'] & Z \arrow[r] & 0
\end{tikzcd}$$

Then the top row is exact if and only if the bottom row is exact.

Proof. ($\Rightarrow$) Suppose the top row is exact. We prove that the bottom row is exact in two parts:

- We show that $\operatorname{im}\chi \subseteq \ker\gamma$. Let $x \in X$ be any element, and define $y := \chi(x)$. Since ϕ_1 is surjective, let $a \in A$ be such that $\phi_1(a) = x$. By the commutativity of the first square:
$$\phi_2\alpha(a) = \chi\phi_1(a) = \chi(x) = y.$$
From the exactness of the top row we have $b := \alpha(a) \in \ker\beta$. Hence, by the commutativity of the second square:
$$\gamma\phi_2(b) = \gamma(y) = \phi_3\beta(b) = \phi_3(0) = 0,$$
thus $y \in \ker\gamma$—which implies in $\operatorname{im}\chi \subseteq \ker\gamma$.
- We now show that $\ker\gamma \subseteq \operatorname{im}\chi$. Let $y \in \ker\gamma$ be any element. Since ϕ_2 is surjective, let $b \in B$ be such that $\phi_2(b) = y$. From the commutativity of the second square:
$$\phi_3\beta(b) = \gamma\phi_2(b) = \gamma(y) = 0,$$
thus $\beta(b) \in \ker\phi_3$—but since ϕ_3 is injective, then $b \in \ker\beta$. From the exactness of the top row we find $a \in A$ such that $\alpha(a) = b$. Using the commutativity of the first square:
$$\chi\phi_1(a) = \phi_2\alpha(a) = \phi_2(b) = y,$$
therefore $y \in \operatorname{im}\chi$ and $\ker\gamma \subseteq \operatorname{im}\chi$.

($\Leftarrow$) If on the contrary we assume that the bottom row is exact, since ϕ_1, ϕ_2 and ϕ_3 are invertible, one can simply consider the following commutative diagram:

$$\begin{tikzcd}
0 \arrow[r] & X \arrow[r, tail, "\chi"] \arrow[d, "\phi_1^{-1}", "\simeq"'] & Y \arrow[r, two heads, "\gamma"] \arrow[d, "\phi_2^{-1}", "\simeq"'] & Z \arrow[r] \arrow[d, "\phi_3^{-1}"', "\simeq"] & 0 \\
0 \arrow[r] & A \arrow[r, "\alpha"'] & B \arrow[r, "\beta"'] & C \arrow[r] & 0
\end{tikzcd}$$

We can now apply the first part of the proof and conclude that the sequence of modules $0 \to A \to B \to C \to 0$ is exact. ♮

Exact Functors

Definition 10.3.27. Let $F\colon R\text{-}\mathtt{Mod} \to \mathtt{Ab}$ be a covariant functor, and $G\colon R\text{-}\mathtt{Mod}^{\mathrm{op}} \to \mathtt{Ab}$ be a contravariant functor. Consider a sequence of R-modules

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0\ ,$$

We define the following notions:

(a) If $0 \to A \to B \to C$ is *exact*, the functor F is said to be *left exact* if the induced sequence of abelian groups

$$0 \longrightarrow FA \xrightarrow{Ff} FB \xrightarrow{Fg} FC$$

is exact. Dually, G is said to be *right exact* if

$$GC \xrightarrow{Gg} GB \xrightarrow{Gf} GA \longrightarrow 0$$

is an exact sequence.

(b) If $A \to B \to C \to 0$ is *exact*, the functor F is said to be *right exact* if the induced sequence of abelian groups

$$FA \xrightarrow{Ff} FB \xrightarrow{Fg} FC \longrightarrow 0$$

is exact. Dually, G is said to be *left exact* if

$$0 \longrightarrow GC \xrightarrow{Gg} GB \xrightarrow{Gf} GA$$

is an exact sequence.

(c) If $0 \to A \to B \to C \to 0$ is *short exact*, the functor F is said to be *exact* if the induced sequence of abelian groups

$$0 \longrightarrow FA \xrightarrow{Ff} FB \xrightarrow{Fg} FC \longrightarrow 0$$

is exact. Analogously, G is said to be *exact* if

$$0 \longrightarrow GC \xrightarrow{Gg} GB \xrightarrow{Gf} GA \longrightarrow 0$$

Proposition 10.3.28. Given any R-module M, the functors

$$\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M,-)\colon R\text{-}\mathtt{Mod} \longrightarrow \mathtt{Ab} \quad \text{and} \quad \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(-,M)\colon R\text{-}\mathtt{Mod}^{\mathrm{op}} \longrightarrow \mathtt{Ab}$$

are both left-exact.

Proof. (a) We first show that $\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M,-)$ is exact. Let $0 \to A \overset{f}{\rightarrowtail} B \overset{g}{\to} C$ be an exact sequence of R-modules, and consider the induced sequence of abelian groups

$$0 \longrightarrow \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M,A) \overset{f_*}{\rightarrowtail} \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M,B) \xrightarrow{g_*} \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M,C).$$

First we show that f_* is injective. Let $\ell \in \ker f_*$ be any morphism, then by definition we have $f_*(\ell) = f\ell = 0$. Since f is injective, we have $\ker f = 0$, which implies in $\operatorname{im}\ell = 0$—hence $\ell = 0$. This shows that $\ker f_* = 0$, therefore f_* is an injective map.

To show that the induced sequence is exact in $\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M,B)$, we prove that $\operatorname{im} f_* = \ker g_*$. Let $\ell \in \operatorname{im} f_*$ be any element and let $k \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M,A)$ be such that $f_*(k) = fk = \ell$. Since the original sequence is exact, we have

$$g_*(\ell) = g\ell = g(fk) = 0$$

since $\operatorname{im} f = \ker g$. Therefore $\ell \in \ker g_*$ and $\operatorname{im} f_* \subseteq \ker g_*$.

For the converse of this inclusion, let $h \in \ker g_*$ be any map so that $gh(m) = 0$ for all $m \in M$—therefore $h(m) \in \ker g$ and by exactness this means that there must exist $a \in A$ such that $f(a) = m$, which is unique since f is an injective map. Define $p\colon M \to A$ by mapping $m \mapsto a$ if $h(m) = f(a)$. To check that p is a morphism of modules, let $m, m' \in M$ be elements and let $a, a' \in A$ be such that $h(m) = f(a)$ and $h(m') = f(a')$. Then since

$$h(m+m') = h(m) + h(m') = f(a) + f(a') = f(a+a'),$$

then $p(m+m') = a + a' = p(m) + p(m')$. If $r \in R$ is any ring element, then one also has

$$h(mr) = h(m)r = f(a)r = f(ar),$$

therefore $p(mr) = ar = p(m)r$. Then we conclude that $p \in \mathrm{Mor}_{\mathrm{Mor}\,R}(M,A)$ and therefore $f_*(p) = h$, showing that $h \in \operatorname{im} f_*$ and that $\ker g_* \subseteq \operatorname{im} f_*$. This finishes the proof that $\operatorname{im} f_* = \ker g_*$.

(b) We now prove the left exactness of the contravariant functor $\mathrm{Mor}_{\mathrm{Mor}\,R}(-,M)$. Let $A \overset{i}{\to} B \overset{p}{\twoheadrightarrow} C \to 0$ be an exact sequence of R-modules, and consider the induced sequence of abelian groups

$$0 \longrightarrow \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(C,M) \overset{p^*}{\rightarrowtail} \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(B,M) \xrightarrow{i^*} \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(A,M).$$

We first show the injectivity of p^*. Let $h \in \ker p^*$ be any map, then for every $b \in B$ one has $hp(b) = 0$, therefore $\operatorname{im} p \subseteq \ker h$. Since p is surjective, $\operatorname{im} p = C$ and since $\ker h \subseteq C$, then we conclude that $\ker h = \operatorname{im} p = C$ and therefore $h = 0$.

Let's now show that $\operatorname{im} p^* = \ker i^*$. If $g \in \operatorname{Mor}_{R\text{-}\mathtt{Mod}}(C, M)$ is any map, then one has

$$i^* p^*(g) = i^*(gp) = (gp)i = g(pi) = 0,$$

from the fact that $\operatorname{im} i = \ker p$. Therefore $g \in \ker i^*$ and $\operatorname{im} p^* \subseteq \ker i^*$. For the converse, take $g \in \ker i^*$. Define a map $f: C \to Y$ by $c \mapsto g(b)$ if $c = p(b)$—an element $b \in B$ with such property is ensured to exist by the surjectivity of p. To check that f is well defined, let $b, b' \in B$ be any two elements with $p(b) = p(b')$—then $p(b - b') = 0$, which implies that there exists $a \in A$ such that $i(a) = b - b'$. Moreover, since $i^*(g) = gi = 0$, then

$$gi(a) = g(b - b') = g(b) - g(b') = 0,$$

therefore $f(b) = f(b')$. To show that f is a morphism of R-modules, let $r \in R$ and $c, c' \in C$ be any elements with corresponding $b, b' \in B$ for which $p(b) = c$ and $p(b') = c'$. Then $p(b + b') = c + c'$ and then

$$f(c + c') = g(b + b') = g(b) + g(b') = f(c) + f(c').$$

Moreover, we have $p(br) = p(b)r = cr$, hence

$$f(cr) = g(br) = g(b)r = f(c)r.$$

This shows that f is indeed R-linear. Therefore we can finally note that $p^*(f) = fp = g$ so that $g \in \operatorname{im} p^*$, hence $\ker i^* \subseteq \operatorname{im} p^*$.

♮

Remark 10.3.29. The functors $\operatorname{Mor}_{R\text{-}\mathtt{Mod}}(M, -)$ and $\operatorname{Mor}_{R\text{-}\mathtt{Mod}}(-, M)$ are *not right-exact*. In other words, if $0 \to A \xrightarrow{f} B$ is an exact sequence of R-modules, *it is not always true* that the sequence of abelian groups $\operatorname{Mor}_{R\text{-}\mathtt{Mod}}(A, M) \xrightarrow{f^*} \operatorname{Mor}_{R\text{-}\mathtt{Mod}}(B, M) \to 0$ is exact.

For instance, one can take the exact sequence $0 \to \mathbf{Z} \overset{f}{\rightarrowtail} \mathbf{Z}$ where $f(x) := 2x$, and the module $M := \mathbf{Z}/2\mathbf{Z}$, so that $\operatorname{Mor}_{\mathbf{Z}\text{-}\mathtt{Mod}}(\mathbf{Z}, M) = \{0, \pi\}$—where $\pi: \mathbf{Z} \twoheadrightarrow \mathbf{Z}/2\mathbf{Z}$ is the natural projection. Therefore the induced map $f^*: \operatorname{Mor}_{\mathbf{Z}\text{-}\mathtt{Mod}}(\mathbf{Z}, M) \to \operatorname{Mor}_{\mathbf{Z}\text{-}\mathtt{Mod}}(\mathbf{Z}, M)$ is *not* surjective, since $\pi \mapsto \pi f = 0$.

Proposition 10.3.30 (Converse of Proposition 10.3.28)**.** Consider R-modules A, B and C, and morphisms of R-modules $f: A \to B$ and $g: B \to C$. We have the following two properties concerning the functors $\operatorname{Mor}_{R\text{-}\mathtt{Mod}}(M, -)$ and $\operatorname{Mor}_{R\text{-}\mathtt{Mod}}(-, M)$:

(a) If the sequence of abelian groups

$$0 \longrightarrow \operatorname{Mor}_{R\text{-}\mathtt{Mod}}(M, A) \overset{f_*}{\rightarrowtail} \operatorname{Mor}_{R\text{-}\mathtt{Mod}}(M, B) \xrightarrow{g_*} \operatorname{Mor}_{R\text{-}\mathtt{Mod}}(M, C)$$

is *exact for every* $M \in R\text{-}\mathtt{Mod}$, then the corresponding sequence of R-modules

$$0 \longrightarrow A \overset{f}{\rightarrowtail} B \xrightarrow{g} C$$

is also *exact*.

(b) If the sequence of abelian groups

$$0 \longrightarrow \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(C, M) \overset{g^*}{\rightarrowtail} \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(B, M) \xrightarrow{f^*} \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(A, M).$$

is *exact for every* $M \in R\text{-}\mathtt{Mod}$, then the corresponding sequence of R-modules

$$A \xrightarrow{f} B \overset{g}{\twoheadrightarrow} C \longrightarrow 0$$

is also *exact*.

Proof. (a) We first show the injectivity of f. Let $M := \ker f$ and consider the natural inclusion $\iota \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, A)$, then $f_*(\iota) = f\iota = 0$, but since f_* is injective then $\iota = 0$—which implies in $\ker f = 0$.

For the proof that $\operatorname{im} f = \ker g$, take $M := \operatorname{im} f$ and consider the natural inclusion $i \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, B)$. Since f is injective, the induced map $\overline{f}\colon A \to \operatorname{im} f$ is a bijection, therefore one may define $\ell\colon \operatorname{im} f \to A$ by $\ell = (\overline{f})^{-1}$. Notice that $f_*(\ell) = f\ell = i$, therefore $i \in \operatorname{im} f_*$. Since $\operatorname{im} f_* \subseteq \ker g_*$, then $g_*(i) = gi = 0$, which implies in $\operatorname{im} f \subseteq \ker g$.

For the converse inclusion, define $M := \ker g$ and consider the canonical inclusion $j \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, B)$, so that $g_*(j) = gj = 0$. From the exactness of the sequence of abelian groups, there must exist $h \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, A)$ such that $f_*(\ell) = f\ell = j$—therefore, for all $b \in \ker g$ there exists $\ell(b) \in A$ such that $f(\ell(b)) = b$, thus $\ker g \subseteq \operatorname{im} f$.

(b) First we prove that g is surjective. Since the sequence is exact for all modules M, take $M := C/\operatorname{im} g$ and let $\pi\colon C \twoheadrightarrow M$ be the natural projection morphism. Then we have $g^*(\pi) = \pi g = 0$, but since g^* is injective we find $\pi = 0$—which can only be the case if $\operatorname{im} g = C$, therefore g is surjective.

We shall now show that $\operatorname{im} f = \ker g$. Notice that since $\operatorname{im} g^* \subseteq \ker f^*$ then $f^*g^* = 0$ but then $(gf)^* = 0$ since $\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(-, M)$ is contravariant. Therefore, if we take $M := C$ and consider the element $\mathrm{id}_C \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(C, M)$, we obtain

$$0 = (gf)^*(\mathrm{id}_C) = \mathrm{id}_C(gf) = gf,$$

therefore $\operatorname{im} f \subseteq \ker g$.

For the other side of the inclusion, consider the module $M := B/\operatorname{im} f$ and let $p\colon B \to M$, then denote the natural projection morphism—so that $f^*(p) = pf = 0$. Since the sequence of abelian groups is exact, there must exist $h \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(C, M)$ such that $g^*(h) = hg = p$. If, for the sake of contradiction, $\ker g$ is *not* contained in $\operatorname{im} f$, there must exist $b \in B$ such that $g(b) = 0$ but $b \notin \operatorname{im} f$. This however implies that $p(b) \neq 0$, while $hg(b) = p(b)$—and since $g(b) = 0$ then $hg(b) = 0$, which is a contradiction. This shows that $\ker g \subseteq \operatorname{im} f$.

♮

10.4 Bimodules

Definition 10.4.1 (Bimodule)**.** Let R and S be any two rings, and M be an abelian group. We say that M is a *(R, S)-bimodule* if M is both a left-R-module and a right-S-module, and M must satisfy the associative law:

$$r(ms) = (rm)s,$$

for all $r \in R$, $m \in M$, and $s \in S$.

Example 10.4.2. Let R be a commutative ring and M be an R-module, then we can define the structure of (R, R)-bimodule on M by making the identification $rm = mr$ for any $r \in R$ and M. Such identification respects the module structure since, given any other $r' \in R$ one has

$$(rr')m = r(r'm) = r'mr = (mr)r' = m(rr') = m(r'r).$$

Proposition 10.4.3. Let R and S be rings, and M be a *(R, S)-bimodule*. The following properties concern the relations between functors $\mathrm{Mor}(M, -)$ and $\mathrm{Mor}(-, M)$, and bimodules:

(a) For any *left-R-module B* we have that $\mathrm{Mor}_{{}_R\mathtt{Mod}}(M, B)$ is a *left-S-module* with a product $s \cdot f \colon M \to B$ mapping $m \mapsto f(ms)$ for any $f \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(M, B)$ and $s \in S$. This can be summarized by the fact that

$$\mathrm{Mor}_{{}_R\mathtt{Mod}}(M, -) \colon {}_R\mathtt{Mod} \longrightarrow {}_S\mathtt{Mod}$$

is a covariant functor.

(b) For any *right-S-module B* we have that $\mathrm{Mor}_{\mathtt{Mod}_S}(M, B)$ is a *right-R-module* with a product $f \cdot r \colon M \to B$ mapping $m \mapsto f(rm)$ for any $f \in \mathrm{Mor}_{\mathtt{Mod}_S}(M, B)$ and $s \in S$. This can be summarized by the fact that

$$\mathrm{Mor}_{\mathtt{Mod}_S}(M, -) \colon \mathtt{Mod}_S \longrightarrow \mathtt{Mod}_R$$

is a covariant functor.

(c) For any *right-S-module A* we have that $\mathrm{Mor}_{\mathtt{Mod}_S}(A, M)$ is a *left-R-module* with a product $r \cdot f \colon A \to M$ mapping $m \mapsto rf(m)$. This is summarized by the fact that

$$\mathrm{Mor}_{\mathtt{Mod}_S}(-, M) \colon {\mathtt{Mod}_S}^{\mathrm{op}} \longrightarrow {}_R\mathtt{Mod}$$

is a contravariant functor.

(d) For any *left-R-module A* we have that $\mathrm{Mor}_{{}_R\mathtt{Mod}}(A, M)$ is a *right-S-module* with a product $f \cdot s \colon A \to M$ mapping $m \mapsto f(m)s$. This is summarized by the fact that

$$\mathrm{Mor}_{{}_R\mathtt{Mod}}(-, M) \colon {}_R\mathtt{Mod}^{\mathrm{op}} \longrightarrow \mathtt{Mod}_S$$

is a contravariant functor.

Proof. The proof of the above statements are rather repetitive, but we shall lay all of them down:

(a) Given a morphism $f \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(M, B)$, elements $s, s' \in S$ and any $m \in M$, one has

$$((ss') \cdot f)(m) = f(m(ss')) = f((ms)s') = (s' \cdot f)(ms) = (s \cdot (s' \cdot f))(m),$$

therefore $(ss') \cdot f = s \cdot (s' \cdot f)$. This shows that $\mathrm{Mor}_{{}_R\mathtt{Mod}}(M, B)$ is a left-S-module. Now, if $g \colon B \to C$ is any morphism of left-R-modules, consider the map $g_* \colon \mathrm{Mor}_{{}_R\mathtt{Mod}}(M, B) \to \mathrm{Mor}_{{}_R\mathtt{Mod}}(M, C)$. Take a morphism $f \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(M, B)$ and consider any two elements $s \in S$ and $m \in M$, then

$$g_*(sf)(m) = g(sf)(m) = gf(ms) = s(gf(m)) = sg_*(f)(m),$$

which shows that g_* is indeed a morphism between left-S-modules.

(b) Let $f \in \mathrm{Mor}_{\mathtt{Mod}_S}(M, B)$, and $r, r' \in R$ be any elements, then given any $m \in M$ one has

$$(f \cdot (rr'))(m) = f((rr')m) = f(r(r'm)) = (f \cdot r)(r'm) = ((f \cdot r) \cdot r')(m).$$

Therefore $\mathrm{Mor}_{\mathtt{Mod}_S}(M, B)$ has a structure of right-R-module. Let $g \colon B \to C$ be any morphism of right-S-modules and consider its induced map $g_* \colon \mathrm{Mor}_{\mathtt{Mod}_S}(M, B) \to \mathrm{Mor}_{\mathtt{Mod}_S}(M, C)$. If $f \in \mathrm{Mor}_{\mathtt{Mod}_S}(M, B)$ is any morphism, consider elements $r \in R$ and $m \in M$, then

$$g_*((f \cdot r))(m) = g(f \cdot r)(m) = gf(rm) = r(gf(m)) = rg_*(f)(m).$$

Therefore g_* is a morphism between right-R-modules.

(c) Let $f \in \mathrm{Mor}_{\mathtt{Mod}_S}(A, M)$ be any morphism, and consider two elements $r, r' \in R$. From the product definition we obtain, for any $m \in M$:

$$((rr') \cdot f)(m) = (rr')f(m) = r(r'f(m)) = r(r' \cdot f)(m) = (r \cdot (r' \cdot f))(m),$$

which shows that $\mathrm{Mor}_{\mathtt{Mod}_S}(A, M)$ has a structure of left-R-module. For the second part of the proposition, consider a morphism $g \colon A \to B$ between right-S-modules, and its corresponding map $g^* \colon \mathrm{Mor}_{\mathtt{Mod}_S}(B, M) \to \mathrm{Mor}_{\mathtt{Mod}_S}(A, M)$. Let $f \in \mathrm{Mor}_{\mathtt{Mod}_S}(B, M)$, $r \in R$, and $b \in B$ be any elements, then one has

$$g^*(r \cdot f)(m) = (r \cdot f) \circ g(m) = rf(g(m)) = rg^*(f)(m).$$

Therefore g^* is a morphism between left-R-modules.

(d) Let $f \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(A, M)$ be any morphism and take elements $s, s' \in S$, then for any $m \in M$ we have

$$(f \cdot (ss'))(m) = f(m)(ss') = (f(m)s)s' = (f \cdot s)(m)s' = ((f \cdot s) \cdot s')(m),$$

therefore $\mathrm{Mor}_{{}_R\mathtt{Mod}}(A, M)$ has the structure of a right-S-module. For the last part, take any morphism $g \colon A \to B$ of left-R-modules and consider the induced map

$g^*\colon \mathrm{Mor}_{{}_R\mathtt{Mod}}(B,M) \to \mathrm{Mor}_{{}_R\mathtt{Mod}}(A,M)$. If $f \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(B,M)$, $s \in S$, and $b \in B$ are any elements, then one has

$$g^*(f\cdot s)(m) = (f\cdot s)\circ g(m) = fg(m)sg^*(f)(m)s.$$

This shows that g^* is a morphism between right-S-modules.

$\natural$

Example 10.4.4. If R is a commutative ring, and A and B are R-modules, then both modules also have the structure of (R,R)-bimodules (see Example 10.4.2). From Proposition 10.4.3 we find that

$$\begin{gathered}\mathrm{Mor}_{(R,R)\text{-}\mathtt{Mod}}(A,-)\colon (R,R)\text{-}\mathtt{Mod} \longrightarrow (R,R)\text{-}\mathtt{Mod},\\ \mathrm{Mor}_{(R,R)\text{-}\mathtt{Mod}}(-,B)\colon (R,R)\text{-}\mathtt{Mod}^{\mathrm{op}} \longrightarrow (R,R)\text{-}\mathtt{Mod}\end{gathered}$$

are the Mor functors.

Corollary 10.4.5. Let R be a ring and M be a left-R-module. Then $\mathrm{Mor}_{{}_R\mathtt{Mod}}(R,M)$ is a left-R-module and there exists a natural isomorphism of left-R-modules

$$\mathrm{Mor}_{{}_R\mathtt{Mod}}(R,M) \simeq M,$$

mapping $f \mapsto f(1)$.

Proof. Since R has a natural structure of (R,R)-bimodule, by means of item (a) of Proposition 10.4.3 we obtain that $\mathrm{Mor}_{{}_R\mathtt{Mod}}(R,M)$ has a structure of left-R-module via the product $r\cdot f \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(R,M)$ mapping $a \mapsto f(ar)$—where $r \in R$ and $f \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(R,M)$ are any two elements.

We first shows that the map $\phi\colon \mathrm{Mor}_{{}_R\mathtt{Mod}}(R,M) \to M$ given by $f \mapsto f(1)$ is a morphism of left-R-modules. Let $f,g \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(R,M)$ be any two elements, then $\phi(f+g) = (f+g)(1) = f(1)+g(1) = \phi(f)+\phi(g)$. Moreover, if $r \in R$ then

$$\phi(r\cdot f) = (r\cdot f)(1) = f(1\cdot r) = f(r) = f(r\cdot 1) = rf(1) = r\phi(f).$$

Therefore ϕ is indeed a morphism as wanted. For injectivity it is simple to realize that $\phi(f) = f(1) = 0$ if and only if $f = 0$, since $f(r) = f(r\cdot 1) = rf(1)$—therefore $\ker\phi = 0$. Moreover, given any $m \in M$, there exists a morphism $g\colon R \to M$ uniquely determined by $g(1) := m$, so that $\phi(g) = m$. $\natural$

Theorem 10.4.6. Let R and S be any two rings, and M be an (R,S)-bimodule. Then for any collection $(B_j)_{j\in J}$ of left-R-modules, the natural isomorphism of abelian groups

$$\mathrm{Mor}_{{}_R\mathtt{Mod}}\Big(A, \prod_{j\in J} B_j\Big) \simeq \prod_{j\in J} \mathrm{Mor}_{{}_R\mathtt{Mod}}(A,B_j)$$

is also an *isomorphism of left-S-modules*.

Proof. Define $B \coloneqq \prod_{j \in J} B_j$, and name the isomorphism by ϕ—which maps $f \mapsto (\pi_j f)_{j \in J}$, where π_j is the j-th canonical projection. Notice that, given any morphism $f \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(A, B)$, $s \in S$ and $a \in A$, we have

$$\pi_j \circ (s \cdot f)(a) = \pi_j(f(as)) = \pi_j f(as) = (s \cdot (\pi_j f))(a)$$

for any $j \in J$. Therefore, we conclude that

$$\phi(s \cdot f) = (\pi_j(s \cdot f))_{j \in J} = (s \cdot (\pi_j f))_{j \in J} = s(\pi_j f)_{j \in J} = s\phi(f),$$

which proves that ϕ is a morphism of left-S-modules. Since ϕ is also an isomorphism of abelian groups, then ϕ is bijective, thus ϕ is indeed an isomorphism of left-S-modules. $\natural$

Definition 10.4.7 (Dual module)**.** Let M be a right-R-module (or left). We define the *dual module* of M to be the *left-R-module* (or right):

$$M^* \coloneqq \mathrm{Mor}_{\mathtt{Mod}_R}(M, R).$$

10.5 Projective Modules

Lifting Property For Projective Modules

Theorem 10.5.1 (Free modules have the lifting property)**.** Let R be a ring and F be a free R-module. For every surjective morphism of R-modules $p \colon M \to N$ and morphism $h \colon F \to N$, there *exists a unique morphism* $\ell \colon F \to M$, called *lifting of h*, such that the diagram

$$\begin{tikzcd} & F \arrow[dl, dashed, bend right, "\ell"'] \arrow[d, "h"] & \\ M \arrow[r, two heads, "p"'] & N \arrow[r] & 0 \end{tikzcd}$$

commutes in R-$\mathtt{Mod}$.

Proof. Since F is free, let $B \coloneqq (b_j)_{j \in J}$ be a basis for F. From the surjectivity of p, for every $j \in J$ there exists $m_j \in M$ such that $p(m_j) = h(b_j)$. From the free module universal property there exists a unique $\ell \colon F \to M$ such that $\ell(b_j) = m_j$ for each $j \in J$. From construction we have $p\ell(b_j) = p(m_j) = h(b_j)$, therefore $p\ell = h$, since B generates F, and the diagram commutes. $\natural$

Remark 10.5.2 (Uniqueness of the lift)**.** The lift ℓ of h *need not be unique* in case F isn't free!

Definition 10.5.3 (Projective module)**.** Let R be a ring and P be an R-module. We say that P is a *projective R-module* if

$$\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(P, -) \colon R\text{-}\mathtt{Mod} \longrightarrow \mathtt{Ab}$$

is an *exact covariant functor*.

Proposition 10.5.4 (Equivalences for projective modules)**.** Let R be a ring and P be an R-module. The following properties are equivalent:

(a) The module P is *projective*.

(b) For every exact sequence of R-modules $M \overset{g}{\twoheadrightarrow} N \to 0$ the sequence of abelian groups

$$\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(P, M) \xrightarrow{g_*} \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(P, N) \longrightarrow 0$$

is *exact*. Equivalently, for every $h \colon P \to N$, there *exists a lifting* $\ell \colon P \to N$ of h—*not necessarily unique*—such that the diagram

$$\begin{tikzcd} & P \arrow[dl, bend right, "\ell"'] \arrow[d, "h"] & \\ M \arrow[r, two heads, "p"'] & N \arrow[r] & 0 \end{tikzcd}$$

is commutative in R-$\mathtt{Mod}$.

(c) Every short exact sequence of R-modules of the form

$$0 \longrightarrow L \rightarrowtail M \twoheadrightarrow P \longrightarrow 0$$

is a *split* sequence.

(d) The module P is a *direct summand* of a *free* R-module.

(e) There exists a collection $(x_j)_{j \in J}$ of elements $x_j \in P$, and a collection of associated morphisms of R-modules $(\phi_j \colon P \to R)_{j \in J}$[3] such that, for all $x \in P$ we have:

- The elements $\phi_j(x) \in R$ are *non-zero for only finitely many* $j \in J$.
- The element x can be written as $x = \sum_{j \in J} x_j \phi_j(x)$.

Proof. From the definition, the equivalence of (a) and (b) is immediate. We prove the following:

- (b) $\Rightarrow$ (c). Let $g \colon M \twoheadrightarrow P$ be the epimorphism depicted in the sequence. Consider the identity morphism $\mathrm{id}_P \colon P \to P$ and apply (b) to obtain $\rho \colon P \to M$ such that

$$\begin{tikzcd} & P \arrow[dl, bend right, "\rho"'] \arrow[d, "\mathrm{id}_P"] & \\ M \arrow[r, two heads, "g"'] & P \arrow[r] & 0 \end{tikzcd}$$

is a commutative diagram of R-modules. Notice that $g\rho = \mathrm{id}_P$, therefore ρ is a section of g—thus the sequence splits.

[3]The collection of pairs $(x_j, \phi_j \colon P \to R)_{j \in J}$ is sometimes referred to as the *dual "basis"*, but it should be noted right away that such family *may not form a basis* for the module P—*not every projective module is free*!

- (c) $\Rightarrow$ (d). Via Theorem 10.2.20 let $p\colon F \twoheadrightarrow M$ be a surjective morphism of R-modules, where F is free. Thus we have a short exact sequence

$$0 \longrightarrow \ker p \hookrightarrow F \xrightarrow{p} P \longrightarrow 0.$$

By item (c) we find that the above sequence is split, therefore if $\iota\colon P \rightarrowtail F$ is a section of p, then
$$F = \ker p \oplus \operatorname{im} \iota \simeq \ker p \oplus P.$$

- (d) $\Rightarrow$ (b). Let $f\colon M \twoheadrightarrow N$ be a surjective morphism of R-modules, and $\psi\colon P \to N$ be any morphism. By item (d), let F be a free module with $F \simeq P \oplus P'$, where P' is the complement R-module of P with respect to F—also, let B be a basis of F. Considering the natural projection $\pi_P\colon F \twoheadrightarrow P$, define a map $\phi'\colon F \to M$ as follows: given $b \in B$, by the surjectivity of f, there exists $m \in M$ such that $f(m) = \psi\pi_P(b)$—we shall define $\phi'(b) := m$. It is easily seen that ϕ' is R-linear, and that $f\phi' = \psi\pi_P$. Moreover, the surjectivity of f implies that ϕ' is the unique morphism of modules with such property. Considering the natural inclusion $\iota_P\colon P \hookrightarrow F$—which is a section of π_P—define $\phi := \phi'\iota_P\colon P \to M$ and notice that

$$f\phi = f(\phi'\iota_P) = (f\phi')\iota_P = (\psi\pi_P)\iota_P = \psi(\pi_P\iota_P) = \psi.$$

This finishes the proof of the equivalence of the items (a), (b), (c), and (d). For item (e), we shall prove its equivalence with (d).

- (d) $\Rightarrow$ (e). Via item (d), there exists an indexing set J and an isomorphism $\psi\colon \bigoplus_{j\in J} R \xrightarrow{\simeq} P \oplus P'$, where P' is the complement of P. If $\pi_P\colon P \oplus P' \twoheadrightarrow P$ denotes the canonical projection, define a collection $(x_j)_{j\in J}$ by $x_j := \pi_P\psi(e_j)$—where $e_j := (\delta_{ij})_{i\in J}$. Let $\iota_P\colon P \hookrightarrow P \oplus P'$ be the canonical inclusion of P, and $\pi_j\colon \bigoplus_{j\in J} R \twoheadrightarrow R$ be the canonical projection of the j-th coordinate. Define a collection of morphisms $(\phi_j)_{j\in J}$ by $\phi_j := \pi_j\psi^{-1}\iota_P\colon P \to R$, so that—since $\psi^{-1}\iota_P(x) \in \bigoplus_{j\in J} R$ has finitely many non-zero components—there are finitely many $j \in J$ such that $\pi_j\psi^{-1}\iota_P(x) \in R$ is non-zero. For the last condition of item (e), if $x \in P$ is any element, we have

$$\begin{aligned} x &= \pi_P\iota_P(x) = \pi_P(\psi\psi^{-1})\iota_P(x) = \pi_P\psi(\psi^{-1}\iota_P)(x) = \pi_P\psi(\phi_j(x))_{j\in J} \\ &= \pi_P\psi\left(\sum_{j\in J} e_j\phi_j(x)\right) = \sum_{j\in J} \pi_P\psi(e_j\phi_j(x)) = \sum_{j\in J} \pi_P\psi(e_j)\phi_j(x) \\ &= \sum_{j\in J} x_j\phi_j(x). \end{aligned}$$

- (e) $\Rightarrow$ (d). The collection $(\phi_j)_{j\in J}$ induces, by the universal property of the product, a unique morphism of R-modules $\phi\colon P \to \prod_{j\in J} R$ mapping $x \mapsto (\phi_j(x))_{j\in J}$. For any $x \in P$ we know from hypothesis that the collection $(\phi_j(x))_{j\in J}$ has finitely many non-zero elements, therefore $\operatorname{im}\phi \subseteq \bigoplus_{j\in J} R$. Then we may naturally restrict the codomain of ϕ, obtaining a morphism $\phi\colon P \to \bigoplus_{j\in J} R$. Define a collection $(e_j)_{j\in J}$

where $e_j := (\delta_{ij})_{i\in J}$, and consider an R-linear map $\lambda\colon \bigoplus_{j\in J} R \to P$ defined by sending $e_j \mapsto x_j$. By hypothesis, one has

$$\lambda\phi(x) = \lambda(\phi_j(x))_{j\in J} = \sum_{j\in J} x_j \phi_j(x) = x,$$

therefore ϕ is a section of λ, showing that λ is a split epimorphism. Thus P is a direct summand of the free module $\bigoplus_{j\in J} R$.

$\natural$

Example 10.5.5. From Theorem 10.5.1 we find that every free module is a projective module.

Example 10.5.6. Let R be a ring. If there exists an idempotent element $e \in R$ (that is, $e^2 = e$), then we have a decomposition $R = eR \oplus (1-e)R$. Therefore eR is projective.

Indeed, given $r \in R$ we can write it as $r = er + (1-e)r$—so that $R = eR + (1-e)R$. Moreover, if $a \in eR \cap (1-e)R$, let $r, r' \in R$ be such that $a = er = (1-e)r'$, however, notice that

$$a = er = e(er) = e((1-e)r') = (e - e^2)r' = (e-e)r' = 0.$$

Therefore $eR \cap (1-e)R = 0$, thus $R = eR \oplus (1-e)R$.

As an example of a *projective module that isn't free*: if e is a central in R, then given any $er \in eR$ we have

$$(er)(1-e) = er - (er)e = er - e(re) = er - e(er) = er - er = 0.$$

That is, the singleton $\{er\}$ is *R-linearly dependent*—therefore eR does *not* admit a basis.

Proposition 10.5.7. Let $(P_j)_{j\in J}$ be a family of R-modules. Then the module $P := \bigoplus_{j\in J} P_j$ is projective if and only if P_j is projective for each $j \in J$.

Proof. Suppose that P is projective, then we let F be a free module of which P is a direct summand, say $F = P \oplus P'$. If $\sigma\colon J \to J$ is any permutation, we know that $\bigoplus_{j\in J} P_j \simeq \bigoplus_{j\in J} P_{\sigma(j)}$, therefore for all $i \in J$ we have

$$F = P \oplus P' = \Big(\bigoplus_{j\in J} P_j\Big) \oplus P' \simeq \Big(P_i \oplus \bigoplus_{j\in J\setminus i} P_j\Big) \oplus P' = P_i \oplus \Big(\Big(\bigoplus_{j\in J\setminus i} P_j\Big) \oplus P'\Big)$$

Since P_i is a direct summand of a free module, it is a projective module.

For the converse, suppose that P_j is projective for all $j \in J$. Let $g\colon M \twoheadrightarrow N$ be a surjective morphism of R-modules, and $\phi\colon P \to N$ be any morphism. Since P_j is projective, if $\iota_j\colon P_j \hookrightarrow P$ is the canonical inclusion, there exists a morphism $\psi_j\colon P_j \to M$ making the diagram

$$\begin{tikzcd}
 & P_j \arrow[d, "\phi\iota_j"] \arrow[dl, bend right, "\psi_j"'] & \\
M \arrow[r, two heads, "g"'] & N \arrow[r] & 0
\end{tikzcd}$$

commute in R-$\texttt{Mod}$. By the universal property of the coproduct P, the collection of morphisms $(\psi_j)_{j\in J}$ induce a unique morphism $\psi\colon P \to M$ such that $\psi\iota_j = \psi_j$ for all $j \in J$. To show that $g\psi = \phi$, notice that, given any $(x_j)_{j\in J} \in P$, one has

$$g\psi(x_j)_{j\in J} = g\Big(\sum_{j\in J} \psi_j(x_j)\Big) = \sum_{j\in J} g\psi_j(x_j) = \sum_{j\in J} \phi\iota_j(x_j) = \phi(x_j)_{j\in J}.$$

Therefore, the following diagram commutes in R-$\texttt{Mod}$:

$$\begin{tikzcd}
 & P \arrow[dl, bend right, "\psi"'] \arrow[d, "\phi"] & \\
M \arrow[r, two heads, "g"'] & N \arrow[r] & 0
\end{tikzcd}$$

which shows that P is a projective module. ♮

Proposition 10.5.8 (Eilenberg's trick)**.** Let P be a *projective* R-module. Then there exists a *free R-module F* such that the direct sum $P \oplus F$ is a *free mocule*.

Proof. Since P is projective, there exists a free R-module F' and an R-module Q such that $F' = P \oplus Q$. Notice that

$$(P \oplus Q) \oplus F' = P \oplus (Q \oplus (P \oplus Q)) = P \oplus ((Q \oplus P) \oplus Q)$$

is a direct sum of free modules, thus also free—however, $(Q \oplus P) \oplus Q$ may not be free, since Q is only ensured to be projective. However, if we define $F \coloneqq \bigoplus_{j\in\mathbf{N}}(Q \oplus P)$, then F is a direct sum of projective modules and hence projective itself. Using the fact that $P \oplus (Q \oplus P) = (P \oplus Q) \oplus P$, we obtain

$$P \oplus F = P \oplus \bigoplus_{j\in\mathbf{N}}(Q \oplus P) = \bigoplus_{j\in\mathbf{N}}(P \oplus Q),$$

which is a direct sum of free modules—hence a projective module. Therefore F satisfies the requirement of the statement. ♮

Proposition 10.5.9 (Schanuel)**.** Let P and P' be projective R-modules. If there exists short exact sequences

$$\begin{tikzcd}
0 \arrow[r] & K \arrow[r, tail, "f"] & P \arrow[r, two heads, "g"] & M \arrow[r] & 0
\end{tikzcd}$$

$$\begin{tikzcd}
0 \arrow[r] & K' \arrow[r, tail, "\phi"] & P' \arrow[r, two heads, "\psi"] & M \arrow[r] & 0
\end{tikzcd}$$

of R-modules, then there exists an isomorphism of R-modules

$$K \oplus P' \simeq K' \oplus P.$$

Proof. Since P is projective, there exists a morphism of R-modules $\varepsilon\colon P \to P'$ such that the diagram

$$\begin{tikzcd}
 & P \arrow[d, two heads, "g"] \arrow[dl, bend right, "\varepsilon"'] & \\
P' \arrow[r, two heads, "\psi"'] & M \arrow[r] & 0
\end{tikzcd}$$

commutes in R-`Mod`. Notice that for any $k \in K$ one has

$$\psi \varepsilon f(k) = g f(k) = 0,$$

therefore $\varepsilon f(k) \in \ker \psi$. Since $\ker \psi \subseteq \operatorname{im} \phi$, there must exist $k' \in K'$ such that $\phi(k') = \varepsilon f(k)$—which needs to be unique with such image since ϕ is injective. Define $\delta\colon K \to K'$ to be the map $k \mapsto k'$, where $k' \in K'$ is as described above. So far, we have the following commutative diagram in R-`Mod`

$$\begin{tikzcd}
0 \arrow[r] & K \arrow[r, tail, "f"] \arrow[d, "\delta"'] & P \arrow[r, two heads, "g"] \arrow[d, "\varepsilon"] & M \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & K' \arrow[r, tail, "\phi"'] & P' \arrow[r, two heads, "\psi"'] & M \arrow[r] & 0
\end{tikzcd}$$

We now define a map $\omega\colon K \to K' \oplus P$ given by $k \mapsto (\delta(k), f(k))$, and $\gamma\colon K' \oplus P \to P'$ by $(k', p) \mapsto \varepsilon(p) - \phi(k')$. Notice that both maps are clearly R-linear and from definition $(k', p) \in \ker \gamma$, that is, $\varepsilon(p) = \phi(k')$ if and only if there exists a common $k \in K$ for which $f(k) = p$ and $\delta(k) = k'$—therefore $\ker \gamma = \operatorname{im} \omega$. To show that ω is injective, let $k \in \ker \omega$ be any element, then from definition $(\delta(k), f(k)) = (0, 0)$, but since f is injective, then $k = 0$—thus $\ker \omega = 0$ and ω is injective. For the surjectivity of γ, if $p' \in P'$, let $p \in P$ be such that $g(p) = \psi(p')$—which exists because g is surjective. Then by the commutativity of the right square one has

$$\psi \varepsilon(p) = g(p) = \psi(p'),$$

that is, $\varepsilon(p) - p' \in \ker \psi$ and thus exists $k' \in K'$ for which $\phi(k') = \varepsilon(p) - p'$. This shows that

$$\gamma(k', p') = \varepsilon(p) - \phi(k') = p',$$

so that γ is surjective.

We've just shown that the sequence

$$\begin{tikzcd}
0 \arrow[r] & K \arrow[r, tail, "\omega"] & K' \oplus P \arrow[r, two heads, "\gamma"] & P' \arrow[r] & 0
\end{tikzcd}$$

is short exact. Since P' is projective, the sequence is also split and therefore there exists an isomorphism of R-modules $K' \oplus P \simeq K \oplus P'$. $\natural$

Finitely Generated Projective Modules

Proposition 10.5.10. Let P be a finitely generated projective right-R-module (or left). Then the dual module P^* is a projective left-R-module (or right).

Proof. If P is finitely generated and projective we can choose a finitely generated free right-R-module F of which P is a direct summand—say, $F = P \oplus P'$. Since F is finitely generated, there exists $n \in \mathbf{Z}_{>0}$ such that $F \simeq \bigoplus_{j=1}^{n} R$. We know that there exists a natural isomorphism of *left*-R-modules

$$F^* = \operatorname{Mor}_{\mathtt{Mod}_R}(P \oplus P', R) \simeq \operatorname{Mor}_{\mathtt{Mod}_R}(P, R) \oplus \operatorname{Mor}_{\mathtt{Mod}_R}(P', R) = P^* \oplus P'^*.$$

Moreover, we can rewrite the morphism set using the fact that F is finitely generated:

$$F^* \simeq \operatorname{Mor}_{\mathtt{Mod}_R}\Big(\bigoplus_{j=1}^{n} R, R\Big) \simeq \bigoplus_{j=1}^{n} \operatorname{Mor}_{\mathtt{Mod}_R}(R, R) \simeq \bigoplus_{j=1}^{n} R,$$

therefore F^* is a finitely generated free left-R-module. Since P^* is a direct summand of F^*, then P^* is a finitely generated *projective* left-R-module. ♮

Proposition 10.5.11. Let P be a right-R-module (or left). The *double dual* P^{**} is a *right-R-module* (or left), and the natural evaluation map

$$\operatorname{eval}\colon P \longrightarrow P^{**}$$

sending $x \mapsto \operatorname{eval}_x\colon P^* \to R$—where $\operatorname{eval}_x(f) = f(x)$—is an *injective morphism of right-R-modules*.

Proof. Let $(x_j, \phi_j)_{j\in J}$ be the dual "basis" of P, where $x_j \in P$ and $\phi_j \in P^*$. If $x \in \ker \operatorname{eval}$, then $f(x) = 0$ for all $f \in P^*$—thus in particular $x = \sum_{j\in J} x_j \phi_j(x) = 0$. This shows that $\ker \operatorname{eval} = 0$, thus the evaluation morphism is injective. ♮

Proposition 10.5.12. Let P be a finitely generated projective right-R-module (or left), with a dual "basis" $(x_j, \phi_j)_{j=1}^{n}$. The following properties hold:

(a) The finite collection $(\phi_j, \operatorname{eval}_{x_j})_{j=1}^{n}$ forms a *dual "basis"* for the dual module P^*.

(b) The dual P^* is a *finitely generated projective left-R-module* (or right), with a generating set $(\phi_j)_{j=1}^{n}$.

(c) The double dual P^{**} is a *finitely generated projective right-R-module* (or left), with a generating set $(\operatorname{eval}_{x_j})_{j=1}^{n}$.

(d) The natural evaluation morphism $\operatorname{eval}\colon P \rightarrowtail P^{**}$ is an *isomorphism* of right-R-modules (or left).

Proof. (a) Let $f \in P^*$ be any functional, then given any $x \in P$ one has

$$\sum_{j=1}^{n} \operatorname{eval}_{x_j}(f)\phi_j(x) = \sum_{j=1}^{n} f(x_j\phi_j(x)) = f\Big(\sum_{j=1}^{n} x_j\phi_j(x)\Big) = f(x),$$

therefore $f = \sum_{j=1}^{n} \operatorname{eval}_{x_j}(f)\phi_j$—which proves that $(\phi_j, \operatorname{eval}_{x_j})_j$ is a dual basis for P^*.

(b) It is immediate from the last item's proof that $(\phi_j)_{j=1}^n$ is a generating set for P^*. Moreover, since P^* admits a dual basis it follows that P^* is projective.

(c) Define a map $\mathrm{eval}^*\colon P^* \longrightarrow (P^{**})^*$ given by $f \mapsto \mathrm{eval}_f^* \in \mathrm{Mor}_{\mathtt{Mod}_R}(P^{**}, R)$, where $\mathrm{eval}_f^*(\Phi) := \Phi(f)$ for any $\Phi \in \mathrm{Mor}_{{}_R\mathtt{Mod}}(P^*, R)$. Notice that, for any $f \in P^*$ one has

$$\sum_{j=1}^n \mathrm{eval}_{x_j}(f)\,\mathrm{eval}_{\phi_j}^*(\Phi) = \sum_{j=1}^n f(x_j)\Phi(\phi_j) = \sum_{j=1}^n \Phi(f(x_j)\phi_j) = \Phi\Big(\sum_{j=1}^n f(x_j)\phi_j\Big) = \Phi(f),$$

that is, $\Phi = \sum_{j=1}^n \mathrm{eval}_{x_j} \cdot \mathrm{eval}_{\phi_j}^*(\Phi)$. Therefore $(\mathrm{eval}_{x_j})_{j=1}^n$ is a generating set for P^{**}, and $(\mathrm{eval}_{\phi_j}^*, \mathrm{eval}_{x_j})_{j=1}^n$ is a dual "basis" for P^{**}. Therefore P^{**} is a finitely generated projective right-R-module.

(d) Given an element $\Phi \in P^{**}$, we can rewrite it as $\Phi = \sum_{j=1}^n \mathrm{eval}_{x_j} \cdot \mathrm{eval}_{\phi_j}^*(\Phi)$, therefore, if we take $\sum_{j=1}^n x_j\Phi(\phi_j) \in P$, one obtains

$$\begin{aligned}
\mathrm{eval}\Big(\sum_{j=1}^n x_j\Phi(\phi_j)\Big) &= \sum_{j=1}^n \mathrm{eval}(x_j\Phi(\phi_j)) = \sum_{j=1}^n \mathrm{eval}(x_j)\Phi(\phi_j)\\
&= \sum_{j=1}^n \mathrm{eval}_{x_j} \cdot\Phi(\phi_j) = \sum_{j=1}^n \mathrm{eval}_{x_j} \cdot \mathrm{eval}_{\phi_j}^*(\Phi)\\
&= \Phi.
\end{aligned}$$

This shows that eval is also a surjective morphism. Therefore eval is an isomorphism of right-R-modules

$$P \simeq P^{**}.$$

♮

10.6 Injective Modules

Lifting Property for Injective Modules

Definition 10.6.1 (Injective module)**.** Let R be a ring, and E be an R-module. We say that E is an *injective R-module* if

$$\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(-, E)\colon R\text{-}\mathtt{Mod}^{\mathrm{op}} \longrightarrow \mathtt{Ab}$$

is an *exact contravariant functor*.

Proposition 10.6.2. Let E be an R-module. The following are equivalent properties:

(a) The module E is *injective*.

(b) Given any exact sequence of R-modules $0 \to M \overset{f}{\rightarrowtail} N$, the sequence

$$\mathrm{Mor}_{R\text{-}\mathtt{Mod}}(N, E) \overset{f^*}{\twoheadrightarrow} \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, E) \longrightarrow 0$$

is an *exact sequence of abelian groups*

(c) For any injective morphism of R-modules $f\colon M \rightarrowtail N$, and morphism $\phi\colon M \to E$, there *exists* a morphism $\psi\colon N \to E$—we do not require uniqueness—for which the diagram

$$\begin{tikzcd} & E & \\ & & \\ 0 \arrow[r] & M \arrow[uu, "\phi"] \arrow[r, tail, "f"'] & N \arrow[uul, bend right, "\psi"'] \end{tikzcd}$$

is *commutative* in $R\text{-}\mathtt{Mod}$.

(d) Any short exact sequence of R-modules of the form

$$0 \longrightarrow E \rightarrowtail A \twoheadrightarrow B \longrightarrow 0$$

is *split*.

Proof. The equivalence between (a) and (b) comes exactly from the definition of what an injective module is. We first prove the equivalence between (b) and (c), then the equivalence between (c) and (d):

- (b) $\Rightarrow$ (c). If we let $0 \to M \overset{f}{\rightarrowtail} N$ be any exact sequence of R-modules, then by (b) we know that $f^*\colon \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(N, E) \twoheadrightarrow \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, E)$ is a surjective morphism. Therefore, if we are given any $\phi \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, E)$, one can choose a $\psi \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(N, E)$ such that $f^*(\psi) = \psi f = \phi$.

- (c) $\Rightarrow$ (b). Let $0 \to M \overset{f}{\rightarrowtail} N$ be any exact sequence of R-modules, then if we are given any morphism $\phi\colon M \to E$, we can find $\psi\colon N \to E$ such that $\phi = f\psi$, but $f\psi = f^*(\psi)$, therefore we've just shown that the map f^* is surjective.

- (c) $\Rightarrow$ (d). Let $0 \to E \overset{f}{\rightarrowtail} A \twoheadrightarrow B \to 0$ be a short exact sequence of R-modules. If we consider the identity morphism id_E, by item (c) we are able to find a morphism $\psi\colon A \to E$ such that $\psi f = \mathrm{id}_E$—therefore f is a split monomorphism and the sequence splits.

- (d) $\Rightarrow$ (c). Let $0 \to L \overset{f}{\rightarrowtail} M$ be an exact sequence of R-modules and let $\phi\colon L \to E$ be a morphism. We may define an R-module

$$X := \frac{E \oplus M}{\{(\phi(x), -f(x)) : x \in L\}},$$

and—given the natural projection $\pi\colon E \oplus M \twoheadrightarrow X$ and the canonical inclusion morphisms $\iota_E\colon E \hookrightarrow E \oplus M$ and $\iota_M\colon M \hookrightarrow E \oplus M$—we define $\ell_E\colon E \to X$ and $\ell_M\colon M \to X$ by $\ell_E := \pi\iota_E$ and $\ell_M := \pi\iota_M$. By Proposition 8.4.64, we know that (X, ℓ_E, ℓ_M) is the pushout of the pair (f, ϕ):

$$\begin{tikzcd} L \arrow[r, tail, "f"] \arrow[d, "\phi"'] & M \arrow[d, "\ell_M"] \\ E \arrow[r, "\ell_E"'] & X \arrow[ul, phantom, "\ulcorner", very near start] \end{tikzcd}$$

Let $e \in \ker \ell_E$ be any element, then

$$\ell_E(e) = [e, 0] \in \{[\phi(x), -f(x)] : x \in L\}.$$

Let $x_e \in L$ be such that $(e, 0) = (\phi(x_e), -f(x_e))$, then $x_e \in \ker f$—but since f is injective, it follows that $x_e = 0$. Therefore $e = \phi(0) = 0$.

From this we can consider the following short exact sequence

$$0 \longrightarrow E \xrightarrow{\ell_E} X \longrightarrow\!\!\!\!\rightarrow \operatorname{coker} \ell_E \longrightarrow 0$$

which by item (d) is split—therefore there exists a retract $r\colon X \to E$ of ℓ_E, that is, $r\ell_E = \mathrm{id}_E$. From this we can create a morphism $\psi\colon M \to E$ given by $\psi := r\ell_M$—we'll show that this map satisfies the condition for item (c). Let $x \in L$ be any element and notice that

$$\psi f(x) = (r\ell_M) f(x) = r[0, f(x)] = r[\phi(x), 0] = r(\ell_E(\phi(x))) = (r\ell_E)\phi(x) = \phi(x),$$

therefore $\psi f = \phi$ as needed, which proves that E is injective.

♮

Corollary 10.6.3. Let M be an R-module, and E be a submodule of M. If E is *injective*, then it is a *direct summand* of M.

Proof. Indeed, since the short exact sequence

$$\begin{tikzcd} 0 \arrow[r] & E \arrow[r, hook, "\iota"] & M \arrow[r, two heads, "\pi"] & M/E \arrow[r] \arrow[l, bend left, "\rho"] & 0 \end{tikzcd}$$

is split, where ρ is a section of π. Therefore

$$M = \operatorname{im} \iota \oplus \operatorname{im} \rho = E \oplus \operatorname{im} \rho.$$

♮

Proposition 10.6.4. Any *direct summand of an injective* R-module is an *injective* R-module

Proof. Let $E = X \oplus Y$, and consider a morphism $\phi\colon L \to X$, and an injective morphism $f\colon L \rightarrowtail M$. If $\iota_X\colon X \hookrightarrow E$, we can define a morphism $\phi'\colon L \to E$ given by $\phi' := \iota_X \phi$. Since E is injective, there exists $\psi'\colon M \to E$ such that $\psi' f = \phi'$. However, taking the canonical projection $\pi_X\colon E \twoheadrightarrow X$ we can define $\psi\colon M \to X$ to be given by $\psi := \pi_X \psi'$—therefore the diagram

$$\begin{tikzcd} & X & \\ 0 \arrow[r] & L \arrow[u, "\phi"] \arrow[r, tail, "f"'] & M \arrow[ul, bend right, "\psi"'] \end{tikzcd}$$

is commutative in R-$\mathtt{Mod}$, hence X is injective. ♮

Proposition 10.6.5. Let $(E_j)_{j\in J}$ be a collection of R-modules. Then the product $\prod_{j\in J} E_j$ is an injective R-module if and only if E_j is injective for each $j \in J$.

Proof. ($\Rightarrow$) We know that any permutation $\sigma\colon J \xrightarrow{\simeq} J$ is such that $\prod_{j\in J} E_j \simeq \prod_{j\in J} E_{\sigma(j)}$ in R-`Mod`. Therefore, for any $i \in J$ we have a natural isomorphism

$$\prod_{j\in J} E_j \simeq E_i \times \Big(\prod_{j\in J\setminus i} E_j \Big),$$

which by Proposition 10.6.4 implies in the injectivity of E_i.

($\Leftarrow$) Let E_j be injective for each $j \in J$. Let $f\colon L \rightarrowtail M$ be any injective morphism of R-modules and consider a morphism $\phi\colon L \to E$. Since each E_j is injective, define $(\psi_j\colon M \to E_j)_{j\in J}$ to be the collection of morphisms such that the diagram

$$\begin{tikzcd}
 & \prod_{j\in J} E_j \arrow[r, two heads, "\pi_j"] & E_j \\
0 \arrow[r] & L \arrow[u, "\phi"] \arrow[r, tail, "f"] & M \arrow[u, "\psi_j"']
\end{tikzcd}$$

commutes for all $j \in J$—where π_j is the natural j-th projection. From the universal property of products, the collection $(\psi_j)_{j\in J}$ defines a unique morphism $\psi\colon M \to E$ such that $\pi_j\psi = \psi_j$ for each $j \in J$. Note that ψf and ϕ are equal if and only if each of their projections match, but since

$$\pi_j(\psi f) = (\pi_j\psi) f = \psi_j f = \pi_j\phi,$$

then $\psi f = \phi$ and the diagram

$$\begin{tikzcd}
 & \prod_{j\in J} E_j & \\
0 \arrow[r] & L \arrow[u, "\phi"] \arrow[r, tail, "f"'] & M \arrow[ul, bend right, "\psi"']
\end{tikzcd}$$

commutes in R-`Mod`—therefore $\prod_{j\in J} E_j$ is an injective module. ♮

Remark 10.6.6. Contrary to the behaviour of projective modules, *the direct sum of a collection of injective modules need not be injective.*

Lemma 10.6.7. Let R be an integral domain and consider its field of fractions $\operatorname{Frac}(R)$ as an *R-module*. The following holds:

(a) Let $I \subseteq \operatorname{Frac}(R)$ be an R-submodule of the R-module $\operatorname{Frac}(R)$, and let $\phi\colon I \to \operatorname{Frac}(R)$ be a *morphism of R-modules*. Then there exists $q \in \operatorname{Frac}(R)$ such that $\phi(y) = qy$ for all $y \in I$.

(b) Let $\mathfrak{a}$ be a submodule (ideal) of R. Then any morphism of R-modules $\phi\colon \mathfrak{a} \to \operatorname{Frac}(R)$ can be extended to a morphism $\overline{\phi}\colon R \to \operatorname{Frac}(R)$.

Proof. (a) If $I = 0$, then $\phi = 0$ and any $q \in \operatorname{Frac}(R)$ satisfies the requirement. On the contrary, suppose I is non-zero, and fix any non-zero $b := w/z \in I$. If $a := u/v \in \operatorname{Frac}(R)$ is any other element, let $r := zv$ so that

$$br = b(zv) = (bz)v = wv \in R \quad \text{and} \quad ar = a(zv) = a(vz) = (av)z = uz \in R,$$

where we used the commutativity of R in order to have $ar \in R$. Therefore one has

$$\phi((ba)r) = \phi(b(ar)) = \phi(b) \cdot (ar),$$
$$\phi((ba)r) = \phi((ab)r) = \phi(a(br)) = \phi(a) \cdot (br).$$

Thus $\phi(a)br = \phi(b)ar$—multiplying this equality by $(br)^{-1} \in \operatorname{Frac}(R)$ we obtain

$$\phi(a) = \phi(b) \cdot (ar)(br)^{-1} = \phi(b) \cdot (ab^{-1}) = \phi(b) \cdot (b^{-1}a) = (\phi(b) \cdot b^{-1})a.$$

Therefore we may define $q := \frac{\phi(b)}{b} \in \operatorname{Frac}(R)$, so that $\phi(y) = qy$.

(b) Given any morphism of R-modules $\phi\colon \mathfrak{a} \to \operatorname{Frac}(R)$, let $q \in \operatorname{Frac}(R)$ be such that $\phi(y) = qy$—which exists because of the last items's result. Define $\overline{\phi}\colon R \to \operatorname{Frac}(R)$ to be an R-module morphism given by $\overline{\phi}(1) := q$—this completely defines ψ since $\overline{\phi}(r) = \overline{\phi}(1 \cdot r) = \overline{\phi}(1)r = qr$. Therefore the diagram

$$\begin{tikzcd} & \operatorname{Frac}(R) & \\ 0 \arrow[r] & \mathfrak{a} \arrow[u, "\phi"] \arrow[r, hook] & R \arrow[ul, bend right, "\overline{\phi}"'] \end{tikzcd}$$

♮

Theorem 10.6.8 (Baer's criterion)**.** Let E be a right-R-module (or left). Then E is *injective* if and only if for every *right ideal* (or left) $\mathfrak{a}$ of R, and morphism of right-R-modules (or left) $\phi\colon \mathfrak{a} \to E$, there exists an extension $\overline{\phi}\colon R \to E$ of ϕ—that is, the diagram

$$\begin{tikzcd} & E & \\ 0 \arrow[r] & \mathfrak{a} \arrow[u, "\phi"] \arrow[r, hook] & R \arrow[ul, bend right, "\overline{\phi}"'] \end{tikzcd}$$

is commutative in $\mathtt{Mod}_R$ (or ${}_R\mathtt{Mod}$).

Proof. ($\Rightarrow$) If E is injective, then the existence of the extension follows immediatly.

($\Leftarrow$) Suppose the latter condition is satisfied. Consider morphisms $\phi\colon L \to E$, and $f\colon L \rightarrowtail M$ injective, in $\mathtt{Mod}_R$. Define $\mathcal{L}$ to be the collection of all pairs $(L', \psi'\colon L' \to E)$ such that $f(L) \subseteq L' \subseteq M$, where L' is a submodule of M, and $\psi' f = \phi$. Notice that since f is injective we may define $\psi' := \phi \circ (f|_{f(L)})^{-1}$ and take $L' := f(L)$, proving that $\mathcal{L}$ is a non-empty set.

We define a partial order $\preccurlyeq$ on $\mathcal{L}$ as follows: $(L', \psi') \preccurlyeq (L'', \psi'')$ if and only if $L' \subseteq L''$ and $\psi''|_{L'} = \psi'$. Define an ascending chain (with respect to $\preccurlyeq$) of pairs $(L'_j, \psi'_j)_{j \in J}$ of elements of $\mathcal{L}$, and let $L'_{\mathrm{M}} \coloneqq \bigcup_{j\in J} L'_j$. From construction, one has that $f(L) \subseteq L'_{\mathrm{M}} \subseteq M$, where L'_{M} is a submodule of M. Define a morphism $\psi'_{\mathrm{M}} \colon L'_{\mathrm{M}} \to E$ as follows: if $x \in L'_j$, define $\psi'_{\mathrm{M}}(x) \coloneqq \psi'_j(x)$—which is well defined because given $(L'_j, \psi'_j) \preccurlyeq (L'_i, \psi'_i)$, one has $\psi_i|_{L_j} = \psi_j$. From its construction the pair $(L'_{\mathrm{M}}, \psi'_{\mathrm{M}})$ is a maximal element of the chain (L'_j, ψ'_j).

Since $\mathcal{L}$ is non-empty and every chain of elements has a maximal element, we can use Zorn's lemma to conclude that $\mathcal{L}$ admits a maximal element (L_0, ψ_0). If, for the sake of contradiction, there exists an element $x \in M \smallsetminus L_0$, define a right-ideal

$$\mathfrak{a} \coloneqq \{r \in R : xr \in L_0\} \subseteq R,$$

and a morphism $\lambda \colon \mathfrak{a} \to E$ by $r \mapsto \psi_0(xr) = \psi_0(x)r$. From hypothesis, λ admits an extension $\overline{\lambda} \colon R \to E$ making the diagram

$$\begin{tikzcd}
 & E & \\
0 \arrow[r] & \mathfrak{a} \arrow[u, "\lambda"] \arrow[r, hook] & R \arrow[ul, bend right, "\overline{\lambda}"']
\end{tikzcd}$$

commute in $\mathtt{Mod}_R$. Define a map $\psi_1 \colon L_0 + xR \to E$ by $\psi_1(y + xr) \coloneqq \psi_0(y) + \overline{\lambda}(r)$, which is certainly a morphism of right-R-modules. To prove that ψ_1 is well defined, if $y + xr = y' + xr' \in L_0 + xR$ then $y - y' = x(r' - r) \in xR$, therefore $y - y' \in L_0$, while $x(r' - r) \in xR$. This implies in $r' - r \in \mathfrak{a}$, therefore both $\psi_0(y - y')$ and $\lambda(r' - r)$. Finally, we see that

$$\psi_0(y - y') = \psi_0(x(r' - r)) = \lambda(r' - r) = \overline{\lambda}(r' - r) = \overline{\lambda}(r') - \overline{\lambda}(r),$$

therefore $\psi_1(y + xr) = \psi_1(y' + xr')$, thus ψ_1 is indeed well defined morphism. From construction, we have $L_0 \subseteq L_0 + xR$ and $\psi_1|_{L_0} = \psi_0$, which implies in

$$(L_0, \psi_1) \preccurlyeq (L_0 + xR, \psi_1),$$

contradicting the maximality of (L_0, ψ_1). Therefore it must be the case that $L_0 = M$, which implies in $\psi_0 \colon M \to E$ being such that $\psi_0 f = \phi$. We conclude that the morphism of right-R-modules ψ_0 makes the diagram

$$\begin{tikzcd}
 & E & \\
0 \arrow[r] & L \arrow[u, "\phi"] \arrow[r, tail, "f"'] & M \arrow[ul, bend right, "\psi_0"']
\end{tikzcd}$$

commutative in $\mathtt{Mod}_R$, showing that E is an injective module. ♮

Corollary 10.6.9. Let R be an integral domain. Then the following holds:

(a) The field of fractions $\mathrm{Frac}(R)$ is an *injective R-module*

(b) Every $\mathrm{Frac}(R)$*-module* is an *injective R-module*.

(c) Every *vector space* is *injective*.

Proof. (a) Notice that by Lemma 10.6.7 item (b) we find that $\mathrm{Frac}(R)$ satisfies the lifting property of injective modules described by Baer's criterion, therefore $\mathrm{Frac}(R)$ is an injective R-module.

(b) Let M be a $\mathrm{Frac}(R)$-module, and $\mathfrak{a} \subseteq R$ be a non-zero ideal of R—if $\mathfrak{a}$ where zero, then the extension property would be satisfies right away. Since M is a module over a field, it's a free module and we can assume that $M \simeq \bigoplus_{j \in J} \mathrm{Frac}(R)$ for some set J.

Let $\phi\colon \mathfrak{a} \to M$ be any morphism of R-modules. If $\pi_j \in \mathrm{Mor}_{R\text{-}\mathtt{Mod}}(M, \mathrm{Frac}(R))$ denotes the canonical projection of the j-th coordinate, we define a collection $(f_j\colon \mathfrak{a} \to \mathrm{Frac}(R))_{j \in J}$ of R-module morphisms by $f_j := \pi_j \phi$ for each $j \in J$. Notice that since

$$\phi(a) = (f_j(a))_{j \in J} = (f_j(1)a)_{j \in J} \in \bigoplus_{j \in J} \mathrm{Frac}(R) \simeq M,$$

it must be the case that $f_j(1)$ is non-zero only for finitely many $j \in J$. Therefore $u := (f_j(1))_{j \in J}$ is an element of M and we can define completely define a morphism of R-modules $\overline{\phi}\colon R \to M$ by mapping $\overline{\phi}(1) := u$. We conclude that such map extends ϕ, that is

$$\begin{tikzcd} & M & \\ 0 \arrow[r] & \mathfrak{a} \arrow[u, "\phi"] \arrow[r, hook] & R \arrow[ul, bend right, "\overline{\phi}"'] \end{tikzcd}$$

is a commutative diagram in R-$\mathtt{Mod}$—which, by Baer's criterion, proves that M is injective.

(c) Given a k-vector space V, where k is any field, we know that $k = \mathrm{Frac}(k)$. Therefore V is an injective k-module by item (b).

♮

Proposition 10.6.10 (Schanuel's dual (see Proposition 10.5.9))**.** Let E and E' be injective R-modules. If there exists short exact sequences

$$0 \longrightarrow M \xrightarrow{f} E \xrightarrow{g} L \longrightarrow 0$$

$$0 \longrightarrow M \xrightarrow{\phi} E' \xrightarrow{\psi} L' \longrightarrow 0$$

of R-modules, then there exists an isomorphism of R-modules

$$E \oplus L' \simeq E' \oplus L.$$

Proof. From the injectivity of E', there exists a morphism of R-modules $\varepsilon\colon E \to E'$ such that the diagram

$$\begin{tikzcd} & E' & \\ 0 \arrow[r] & M \arrow[u, tail, "\phi"] \arrow[r, tail, "f"'] & E \arrow[ul, bend right, "\varepsilon"'] \end{tikzcd}$$

commutes in R-$\texttt{Mod}$. Define a collection $(e_\ell)_{\ell\in L}$ where $g(e_\ell)=\ell$—which is possible because g is surjective—and define the map $\lambda\colon L\to L'$ to be given by $\ell\mapsto\psi\varepsilon(\ell)$, then λ is certainly a morphism of R-modules and $\lambda g=\psi\varepsilon$. We've constructed the following commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & M \arrow[r, tail, "f"] \arrow[d, equal] & E \arrow[r, two heads, "g"] \arrow[d, "\varepsilon"'] & L \arrow[r] \arrow[d, "\lambda"] & 0 \\
0 \arrow[r] & M \arrow[r, tail, "\phi"'] & E' \arrow[r, two heads, "\psi"'] & L' \arrow[r] & 0
\end{tikzcd}$$

in R-$\texttt{Mod}$. Define mappings $\alpha\colon E\to E'\oplus L$ given by $e\mapsto(\varepsilon(e),g(e))$, and $\beta\colon E'\oplus L\to L'$ by $(e',\ell)\mapsto\lambda(\ell)-\psi(e')$. From construction, both are R-module morphisms. Moreover, notice that $(e',\ell)\in\ker\beta$ if and only if $\lambda(\ell)=\psi(e')$, hence if we take $e\in E$ such that $g(e)=\ell$, we find by the commutativity of the right square that

$$\psi\varepsilon(e)=\lambda g(e)=\lambda(\ell)=\psi(e').$$

Therefore $\varepsilon(e)-e'\in\ker\psi$ and by exactness there exists $m\in M$ for which $\phi(m)=\varepsilon(e)-e'$. From the commutativity of the left square one has

$$\varepsilon f(m)=\varepsilon(e)-e',$$

thus $\varepsilon(e-f(m))=e'$. Notice that

$$\alpha(e-f(m))=(\varepsilon(e-f(m)),g(e-f(m)))=(e',g(e)-gf(m))=(e',g(e))=(e',\ell),$$

which proves that $(e',\ell)\in\operatorname{im}\alpha$. Therefore $\ker\beta\subseteq\operatorname{im}\alpha$. For the converse, given any $e\in E$ we have

$$\beta\alpha(e)=\beta(\varepsilon(e),g(e))=\lambda g(e)-\psi\varepsilon(e)=0$$

since $\lambda g=\psi\varepsilon$—thus $\operatorname{im}\alpha\subseteq\ker\beta$.

We now show that α is injective, while β is surjective. If $e\in\ker\alpha$ then by definition $(\varepsilon(e),g(e))=0$. Since $\operatorname{im}f=\ker g$, let $m\in M$ be such that $f(m)=e$. By the commutativity of the left square we know that

$$\phi(m)=\varepsilon f(m)=\varepsilon(e)=0,$$

but ϕ is injective, thus $m=0$. This proves that $e=f(m)=0$ and therefore $\ker\alpha=0$. For surjectivity, let $\ell'\in L'$ be any element. Since ψ is surjective, choose $e'\in E'$ with image $\psi(e')=\ell'$. Taking the pair $(-e',0)\in E'\oplus L$ we get

$$\beta(-e',0)=\lambda(0)-\psi(-e')=\psi(e')=\ell,$$

therefore β is surjective.

In the last two paragraphs we've shown that the sequence of R-modules

$$0\longrightarrow E\overset{\alpha}{\rightarrowtail}E'\oplus L\overset{\beta}{\twoheadrightarrow}L'\longrightarrow 0$$

is short exact. Since E is injective, then the sequence splits, proving the existence of an isomorphism $E'\oplus L\simeq E\oplus L'$. ♮

Divisible Modules

Definition 10.6.11 (Divisible module)**.** Let R be a *domain*, and D be an R-module. We say that D is a *divisible R-module* if for all $x \in D$ and $r \in R \setminus \{0\}$, there exists $y \in D$ for which $yr = x$.

Proposition 10.6.12. Let R be an integral domain. The following holds:

(a) The field of fractions $\mathrm{Frac}(R)$ is a divisible R-module.

(b) The sum (either direct or not) of divisible R-modules is a divisible R-module.

(c) The direct product of divisible R-modules is a divisible R-module.

(d) If D is a divisible R-module and $\phi\colon D \to M$ is a morphism of R-modules, then the image $\phi(D) \subseteq M$ is a divisible R-module.

(e) The quotient of a divisible R-module is a divisible R-module.

(f) Every direct summand of a divisible R-module is itself a divisible R-module.

Proof. (a) Given any element $x := u/v \in \mathrm{Frac}(R)$ and a non-zero $r \in R$, one has $y := u/(vr) \in \mathrm{Frac}(R)$ such that

$$yr = \frac{u}{vr} \cdot r = \frac{u}{v} = x.$$

(b) Let $(D_j)_{j \in J}$ be a family of divisible R-modules. First we prove that the sum of the family is divisible. Let $\sum_{j \in F} x_j \in \sum_{j \in J} D_j$ be any element, where $F \subseteq J$ is a finite subset. If $r \in R$ is any non-zero element, since D_j is divisible, let $y_j \in D_j$ be such that $y_j r = x_j$. Since F is finite, we have $\sum_{j \in F} y_j \in \sum_{j \in J} D_j$, therefore

$$\Big(\sum_{j \in F} y_j\Big) r = \sum_{j \in F} y_j r = \sum_{j \in F} x_j.$$

This proves that $\sum_{j \in J} D_j$ is a divisible R-module.

We now consider the direct sum $\bigoplus_{j \in J} D_j$ and any element $(x_j)_{j \in J} \in \bigoplus_{j \in J} D_j$. If $r \in R$ is non-zero, define a collection $(y_j)_{j \in J}$ as follows: if $x_j = 0$, let $y_j = 0 \in D_j$, otherwise we use the divisibility of D_j and let $y_j \in D_j$ be an element such that $y_j r = x_j$. Since finitely many $j \in J$ have a non-zero x_j, it follows that $(y_j)_{j \in J} \in \bigoplus_{j \in J} D_j$. Therefore we have

$$(y_j)_{j \in J} r = (y_j r)_{j \in J} = (x_j)_{j \in J},$$

proving that $\bigoplus_{j \in J} D_j$ is a divisible R-module.

(c) Let $(D_j)_{j \in J}$ be a collection of divisible R-modules, and let $(x_j)_{j \in J} \in \prod_{j \in J} D_j$ be any element. If $r \in R$ is a non-zero element, we define a collection $(y_j)_{j \in J} \in \prod_{j \in J} D_j$ such that $y_j r = x_j$—which is possible since each D_j is divisible. Therefore $(y_j)_{j \in J} r = (y_j r)_{j \in J} = (x_j)_{j \in J}$, proving that the direct product $\prod_{j \in J} D_j$ is a divisible R-module.

(d) Let $m \in \phi(D)$ be any element and take $d \in \phi^{-1}(m)$. Given any non-zero $r \in R$, since D is divisible, let $y \in D$ be such that $yr = d$. Applying ϕ to such element,

$$m = \phi(d) = \phi(yr) = \phi(y)r.$$

Since $\phi(y) \in \phi(D)$, this shows that m is divisible by r—hence $\phi(D)$ is a divisible R-module.

(e) Given a divisible R-module D and a submodule $Q \subseteq D$, the natural projection $\pi\colon D \twoheadrightarrow D/Q$ shows that $\pi(D) = D/Q$ is a divisible R-module via last item's result.

(f) Let D be a divisible R-module and suppose that $D = X \oplus Y$. Then the natural projection $\pi_X\colon D \twoheadrightarrow X$ has an image $\pi_X(D) = X$—therefore X is divisible by the result of item (d).

$\natural$

Proposition 10.6.13 (Injective module is divisible)**.** Let R be an *integral domain*. Then every *injective* R-module *is* a *divisible* R-module.

Proof. Let E be an injective R-module, and $x \in E$ be any element. If $r \in R$ is any non-zero element, consider the submodule (ideal) $rR \subseteq R$—which is free with a basis $\{r\}$. By the injectivity of E, if $\phi\colon rR \to E$ is defined by $\phi(r) := x$, there exists an extension $\psi\colon R \to E$ such that $\psi|_{rR} = \phi$. Therefore one has

$$x = \psi(r) = \psi(1 \cdot r) = \psi(1)r,$$

which proves that x is divisible by r, and hence E itself is a divisible R-module. $\natural$

Proposition 10.6.14 (Divisible modules in PIDs)**.** Let R be a *principal ideal domain*. The following holds:

(a) An R-module is *injective* if and only if it is *divisible*.

(b) Every *quotient* of an *injective* R-module is itself *injective*.

Proof. (a) If R is a principal ideal domain, then in particular it's an integral domain, therefore any injective R-module is divisible. We prove the converse. Let D be a divisible R-module and take any non-zero submodule (ideal) $\mathfrak{a}$ of R, together with a morphism of R-modules $\phi\colon \mathfrak{a} \to E$. Since R is a PID, assume $\mathfrak{a} = aR$ for some non-zero $a \in R$. By the divisibility of D, we know that there exists $d \in D$ such that $\phi(a) = da$—that is, $\phi(a) \in D$ is divisible by d. We define a map $\psi\colon R \to D$ by mapping $r \mapsto dr$, which is clearly a morphism of R-modules. Notice that if $x := ar \in \mathfrak{a}$ is any element, then

$$\psi(x) = \psi(ar) = \psi(a)r = (da)r = \phi(a)r = \phi(ar) = \phi(x),$$

that is, ψ extends ϕ, making the diagram

$$\begin{tikzcd} & D & \\ 0 \arrow[r] & \mathfrak{a} \arrow[u, "\phi"] \arrow[r, hook] & R \arrow[ul, bend right, "\psi"'] \end{tikzcd}$$

commutative in R-$\mathtt{Mod}$ and showing that D is injective.

(b) Let E be an injective module, hence divisible since R is in particular an integral domain. Therefore, if $M \subseteq E$ is any submodule, we know that E/M is a divisible R-module by Proposition 10.6.12 item (e). Therefore, by the fact that R is a PID, we use last item's result to obtain that E/M is injective.

♮

Corollary 10.6.15 (Embedding abelian groups into injective ones)**.** Every *abelian group* can be *embedded* as a subgroup of an *injective abelian group*.

Proof. Let M be an abelian group ($\mathbf{Z}$-module) and by Theorem 10.2.20 we know that there exists a free abelian group $F := \bigoplus_{j \in J} \mathbf{Z}$ such that $M = F/K$ for some subgroup $K \subseteq F$. Considering the natural inclusion morphism $\iota \colon \mathbf{Z} \hookrightarrow \mathbf{Q}$ of $\mathbf{Z}$-modules, we find that F is a submodule of the $\mathbf{Z}$-module $\bigoplus_{j \in J} \mathbf{Q}$. Since $\mathbf{Q}$ is a divisible $\mathbf{Z}$-module, by Proposition 10.6.12 we find that $E := (\bigoplus_{j \in J} \mathbf{Q})/K$ is again a divisible $\mathbf{Z}$-module. Since $\mathbf{Z}$ is a PID, by Proposition 10.6.14 we find that E is injective. Therefore the proposition follows by the natural inclusion $M = F/K \hookrightarrow E$. ♮

Theorem 10.6.16 (Embedding modules into injective ones)**.** Every right-R-module (or left) can be *embedded as a submodule* of some *injective right-R-module* (or left).

Proof. Let M be any right-R-module. By Corollary 10.6.15, let D be an injective abelian group containing the abelian group M. Considering the isomorphism $M \simeq \mathrm{Mor}_{\mathtt{Mod}_R}(R, M)$ given by $m \mapsto f_m$—where $f_m(r) := mr$—we have an injective right-R-module morphism $M \rightarrowtail \mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, D)$ given by the composition:

$$M \xrightarrow{\simeq} \mathrm{Mor}_{\mathtt{Mod}_R}(R, M) \hookrightarrow \mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, M) \hookrightarrow \mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, D)$$

This proves that M can be embedded as a subgroup of the right-R-module $\mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, D)$.

♮

Proposition 10.6.17. Let D be a *divisible abelian group* and R be any ring. Then the abelian group $\mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, D)$ is an *injective right-R-module* (also, $\mathrm{Mor}_{{}_{\mathbf{Z}}\mathtt{Mod}}(R, D)$ is an injective left-R-module), with a product

$$(f \cdot r)(x) := f(rx)$$

for every $f \in \mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, D)$, and $r, x \in R$.

Proof. We take D to be a right-$\mathbf{Z}$-module—which is isomorphic to the left module since $\mathbf{Z}$ is commutative. Any ring R has the structure of $(R, \mathbf{Z})$-bimodule, thus the product is well defined and produces a structure of right-R-module in the abelian group $\mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, D)$.

We now prove that $\mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, D)$ is injective using Baer's criterion. Since $\mathbf{Z}$ is a PID, the divisibility of D implies that it is an injective $\mathbf{Z}$-module. Let $\mathfrak{a} \subseteq R$ be any right-ideal of R and $\phi \colon \mathfrak{a} \to \mathrm{Mor}_{\mathtt{Mod}_{\mathbf{Z}}}(R, D)$ be a morphism of R-modules. Define a map

$\phi' : \mathfrak{a} \to D$ by $a \mapsto \phi(a)(1)$, which is a morphism of right-$\mathbf{Z}$-modules. By the injectivity of D, there exists a morphism of right-$\mathbf{Z}$-modules $\psi' : R \to D$ such that the diagram

$$\begin{tikzcd} & D & \\ 0 \arrow[r] & \mathfrak{a} \arrow[u, "\phi'"] \arrow[r, hook] & R \arrow[ul, bend right, "\psi'"'] \end{tikzcd}$$

commutes in $\texttt{Mod}_{\mathbf{Z}}$. Define a morphism of right-R-modules $\psi : R \to \mathrm{Mor}_{\texttt{Mod}_{\mathbf{Z}}}(R, D)$ by assigning $\psi(1) := \psi'$, which completely determines ψ since

$$\psi(s)(r) = (\psi(1)s)(r) = (\psi' s)(r) = \psi'(sr)$$

for any $s, r \in R$. Moreover, if $a \in \mathfrak{a}$ is any element, then for any $r \in R$ one has

$$\psi(a)(r) = \psi'(ar) = \phi'(ar) = \phi(ar)(1) = (\phi(a)r)(1) = \phi(a)(r),$$

since $ar \in \mathfrak{a}$. Therefore $\psi|_{\mathfrak{a}} = \phi$, making the diagram

$$\begin{tikzcd} & \mathrm{Mor}_{\texttt{Mod}_{\mathbf{Z}}}(R, D) & \\ 0 \arrow[r] & \mathfrak{a} \arrow[u, "\phi"] \arrow[r, hook] & R \arrow[ul, bend right, "\psi"'] \end{tikzcd}$$

commute in $\texttt{Mod}_R$—which shows that $\mathrm{Mor}_{\texttt{Mod}_{\mathbf{Z}}}(R, D)$ is an injective right-R-module. ♮

Lemma 10.6.18. Let E be an injective R-module, and $0 \to A \overset{\alpha}{\rightarrowtail} B \overset{\beta}{\twoheadrightarrow} C \to 0$ be a short exact sequence of R-modules. If there exists a morphism $\gamma : A \to E$, then one can complete the given exact sequence to a commutative diagram in R-$\texttt{Mod}$:

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r, tail, "\alpha"] \arrow[d, "\gamma"'] & B \arrow[r, two heads, "\beta"] \arrow[d, "\gamma'"] & C \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & E \arrow[r, tail, "\alpha'"] & P \arrow[r, two heads, "\beta'"] & C \arrow[r] & 0 \end{tikzcd}$$

whose rows are short exact sequences.

Proof. Let P be the pushout of the pair (γ, α), that is:

$$P := \frac{E \oplus B}{\{(\gamma(a), -\alpha(a)) \in E \oplus B : a \in A\}}$$

together with natural inclusion maps $\alpha' : E \to P$ given by $e \mapsto [e, 0]$ and $\gamma' : B \to P$ sending $b \mapsto [0, b]$. Define $\beta' : P \to C$ to be the mapping $[e, b] \mapsto \beta(b)$. To see that β' is well defined, let $[e, b] = [e', b']$ and notice this means that there exists a common $a \in A$ such that $\alpha(a) = b - b'$ and $\gamma(a) = e - e'$. Since the top row is exact, then

$$0 = \beta\alpha(a) = \beta(b - b') = \beta(b) - \beta(b'),$$

that is, $\beta'[e, b] = \beta'[e', b']$. Also it is clear that β' is a morphism of R-modules and that $\beta'\gamma' = \beta$.

Given any $e \in \ker \alpha'$ one has $\alpha'(e) = [e, 0] = [0, 0]$, that is, there exists $a \in A$ for which $(\gamma(a), -\alpha(a)) = (e, 0)$ but since α is injective, then $a = 0$—proving that $\ker \alpha' = 0$ and that α' is injective. For the surjectivity of β', we use the fact that β is surjective: given any $c \in C$, there exists $b \in B$ for which $\beta(b) = c$—therefore $\beta'\gamma'(b) = \beta'[0, b] = \beta(b) = c$, showing that β' is surjective.

We now show the exactness of the bottom row. Since $\beta'\alpha'(e) = \beta'[e, 0] = \beta(0) = 0$, then $\operatorname{im} \alpha' \subseteq \ker \beta'$. Given any $[e, b] \in \ker \beta'$ by definition we have $\beta(b) = 0$, therefore by exactness of the top row there must exist $a \in A$ such that $\alpha(a) = b$. Therefore

$$[e, b] = [e, \alpha(a)] = [e, \alpha(a)] + [\gamma(a), -\alpha(a)] = [e + \gamma(a), 0]$$

since $[\gamma(a), -\alpha(a)] = [0, 0]$. From this we obtain that $\alpha'(e + \gamma(a)) = [e, b]$ and hence $\ker \beta' \subseteq \operatorname{im} \alpha'$. ♮

Proposition 10.6.19. An R-module E is *injective* if and only if *every* short exact sequence of R-modules

$$0 \longrightarrow E \overset{f}{\rightarrowtail} B \overset{g}{\twoheadrightarrow} C \longrightarrow 0$$

ending with a *cyclic* module C is *split*.

Proof. The forward implication is immediate since any short exact sequence of right-R-modules starting with an injective module is split. For the converse, assume that E has the described property, and let $\mathfrak{b} \subseteq R$ be any right-ideal together with a morphism of right-R-modules $\phi\colon \mathfrak{b} \to E$. Considering the short exact sequence $0 \to \mathfrak{b} \hookrightarrow R \twoheadrightarrow R/\mathfrak{b}$, using Lemma 10.6.18 there exists a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathfrak{b} \arrow[r, hook, "\iota"] \arrow[d, "\phi"'] & R \arrow[r, two heads] \arrow[d, "\phi'"] & R/\mathfrak{b} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & E \arrow[r, tail, "\alpha"] & P \arrow[r, two heads] & R/\mathfrak{b} \arrow[r] & 0
\end{tikzcd}$$

Since $R/\mathfrak{b}$ is a cyclic right-R-module, by hypothesis the bottom sequence splits, proving the existence of a retract $q\colon P \to E$ such that $q\alpha = \mathrm{id}_E$. Therefore the composition $q\phi'\colon R \to E$ is a morphism of right-R-modules such that $(q\phi)\iota = \phi$, making the diagram

$$\begin{tikzcd}
 & E & \\
0 \arrow[r] & \mathfrak{b} \arrow[u, "\phi"] \arrow[r, hook, "\iota"'] & R \arrow[ul, bend right, "q\phi"']
\end{tikzcd}$$

commute in $\mathtt{Mod}_R$. Thus by Baer's criterion E is an injective module. ♮

Part III

Combinatorics

Chapter 11

A Short Toy-Manual of Graphs

11.1 General Definition Of a Graph

Notation 11.1.1 (k-subsets)**.** Given a set A, we define a k-subset—where $0 \leqslant k \leqslant |A|$—to be the a subset of A with exactly k elements. We denote the collection of all k-subsets of A as $[A]^k$.

Notation 11.1.2 (Range)**.** For any $n \in \mathbf{N}$, we denote by $[n]$ the set of natural numbers $\{0, 1, \dots, n\}$.

Definition 11.1.3 (General Graph)**.** We define a *general graph* $G = (V, E, d)$ as a collection of disjoint sets V and E—called, respectively, the vertices and edges of G—together with a map d that defines the relations of incidence (see Definition 11.1.4) between vertices and edges—that is, it defines the formation of edges between vertices.

Definition 11.1.4 (Incidence)**.** We say that a vertex x is *incident* with an edge e if $x \in e$. Moreover if $x \in e$, we say that e is an edge *at* x.

In order to ease the the way on which we talk about graphs, I'll introduce some notation that I judge will be quite appropriate to avoid confusion.

Notation 11.1.5 (Vertices and edges of a graph)**.** Given a graph G, its collection of vertices and edges are denoted by, respectively, $\mathrm{Vert}(G)$ and $\mathrm{Edge}(G)$.

Notation 11.1.6 (Existence of an edge)**.** Let G be any graph and consider two vertices $x, y \in \mathrm{Vert}(G)$. We denote the relation of existence of an edge joining x and y by $\mathrm{Edge}(x, y)$—that is, if $\mathrm{Edge}(x, y)$ is `true`, there exists an edge $e \in \mathrm{Edge}(G)$ such that x and y are its end-vertices, otherwise, if $\mathrm{Edge}(x, y)$ is `false`, then there is no edge joining the two vertices.

Definition 11.1.7 (Adjacency)**.** Given vertices $x, y \in V$, we say that x and y are neighbours if $\mathrm{Edge}(x, y)$—in such case, x and y are said to be end-vertices of the existent edge joining them.

We say that two edges e and g are *adjacent* or *neighbours* if there exists a unique vertex x common to both of them—$x \in e$ and $x \in g$.

A set of vertices, or edges, is said to be independent if no pair of elements is adjacent. An independent set of vertices V is called stable.

Definition 11.1.8 (Diagonal set of vertices)**.** Let $G = (V, E)$ be a graph. We define the diagonal subset of V to be the collection $\Delta_V = \{(x, x) : x \in V\}$. So that we have $V^2 \smallsetminus \Delta_V = \{(x, y) : x \neq y\}$—the collection of ordered pairs of distinct vertices.

The following definition will be useful when we are talking about graphs in which there is no concept of direction, that is, an edge joining x and y is exactly the same as the edge joining y and x—that is, edges have no intrinsic orientation. Graph theorists like to think about graphs as being entities that satisfy such invariance on the order of the vertices—but, when it's possible, we shall take the most general approach.

Definition 11.1.9 (Edge invariance)**.** We define the set $\langle V\rangle^2$ to be the quotient

$$\langle V\rangle^2 = (V^2 \smallsetminus \Delta_V)/_{(x,y)\sim(y,x)}.$$

Notation 11.1.10 (Collection of edges)**.** Given sets X and Y, we denote the collection of all edges of the form xy, where $x \in X$ and $y \in Y$, by $\mathrm{Edge}(X, Y)$. The collection of all edges that a given vertex x is incident with is denoted $\mathrm{Edge}(x)$.

Definition 11.1.11 (Order and size)**.** Let G be a graph. We define the *order* of G as $|G| = |\mathrm{Vert}(G)|$, moreover, the *size* of G is defined as $\|G\| = |\mathrm{Edge}(G)|$.

Definition 11.1.12 (Trivial graph)**.** A graph G is said to be trivial if $|G| \leqslant 1$. In the case of the 0 order, we denote the graph as $\varnothing$.

11.2 The Simple Graph — Looping All Over

Definition 11.2.1 (Simple Graph)**.** A simple graph is defined to be a graph $G = (V, E, d)$ (in the sense of Definition 11.1.3) where

$$d\colon E \to V^2/_{(x,y)\sim(y,x)}, \text{ mapping } e \overset{d}{\longmapsto} (x, y) = (y, x)$$

That is, d defines a reflexive and symmetric relation — that is, $\mathrm{Edge}(x, x)$ is true for all $x \in V$, and $\mathrm{Edge}(x, y) = \mathrm{Edge}(y, x)$ for all pairs $x, y \in V$.

We'll now construct a category for simple graphs using presheaves on the the following category.

Definition 11.2.2 (Walking simple graph)**.** A walking simple graph is defined as a category S consisting of

- An object V, called the set of vertices — containg elements x.
- An object E, called the set of edges — containing ordered pairs $e = (x, y)$, where $x, y \in V$.

- Morphisms $s, t\colon E \to V$ called source and target, respectively — given an edge $e = (x, y)$ we have $e \overset{s}{\mapsto} x$ and $e \overset{t}{\mapsto} y$. We also impose that the map defined by $e \mapsto (s(e), t(e))$ in $\mathrm{End}_{\mathtt{Set}}(E)$ is injective.

- Identity morphisms $\mathrm{id}_V\colon V \to V$ and $\mathrm{id}_E\colon E \to E$ (reflexivity of vertices and edges).

- An automorphism $\sigma\colon E \xrightarrow{\simeq} E$ called symmetry — mapping $e = (x, y) \overset{\sigma}{\mapsto} (y, x)$.

Proposition 11.2.3 (Simple graph is a presheaf)**.** Let S be a walking simple graph. A simple graph is a presheaf

$$G\colon \mathsf{S}^{\mathrm{op}} \to \mathtt{Set}.$$

Proof. The source and target morphisms correspond to $G(s), G(t)\colon G(E) \to G(V)$ in such a way that, given any $e = (x, y) \in G(E)$, $e \xmapsto{G(s)} x$ and $e \xmapsto{G(t)} y$ (edges have sources and targets). Moreover, since the map $e \mapsto (s(e), t(e))$ is injective, if $e, e' \in G(E)$ are edges such that $(s(e), t(e)) = (s(e'), t(e'))$, then $e = e'$ (edges are well defined). The automorphism σ corresponds to $G(\sigma)\colon G(E) \to G(E)$ which is again an automorphism and hence defines an equivalence between edges (x, y) and (y, x) — which implies that the source and target of an edge are indiscernible from each other. $\natural$

Definition 11.2.4 (Morphism of simple graphs)**.** We define morphisms of simple graphs in two equivalent ways — respectively following Definition 11.2.1 and Proposition 11.2.3:

- A morphism $G \to H$ of simple graphs $G = (V, E)$ and $H = (V', E')$ is a map $f\colon V \to V'$ such that $\mathrm{Edge}(x, y)$ implies $\mathrm{Edge}(f(x), f(y))$.

- Let $G = (V, E, s, t)$ and $H = (V', E', s', t')$ be simple graphs. A morphism $G \to H$ is a natural transformation $\alpha\colon G \Rightarrow H$

$$\begin{tikzcd} \mathsf{S}^{\mathrm{op}} \arrow[r, bend left, "G"{name=U}] \arrow[r, bend right, "H"'{name=D}] & \mathtt{Set} \arrow[Rightarrow, from=U, to=D, "\alpha"] \end{tikzcd}$$

 Where α is defined by a pair of morphisms $\alpha_V\colon V \to V'$ and $\alpha_E\colon E \to E'$ such that the following diagrams commute

$$\begin{tikzcd} E \arrow[r, "s"] \arrow[d, "\alpha_E"'] & V \arrow[d, "\alpha_V"] \\ E' \arrow[r, "s'"'] & V' \end{tikzcd} \qquad \begin{tikzcd} E \arrow[r, "t"] \arrow[d, "\alpha_E"'] & V \arrow[d, "\alpha_V"] \\ E' \arrow[r, "t'"'] & V' \end{tikzcd}$$

from the symmetric and reflective relations in Mor(S) we see that this is equivalent to the commutativity of the diagram[1]

$$\begin{tikzcd} E \arrow[r, "s\times t"] \arrow[d, "\alpha_E"'] & V^2/_{(x,y)\sim(y,x)} \arrow[d, "\alpha_V\times\alpha_V"] \\ E' \arrow[r, "s'\times t'"'] & V'^2/_{(x',y')\sim(y',x')} \end{tikzcd}$$

Definition 11.2.5 (Simple graphs category)**.** The category of simple graphs, denoted by `sGraph`, consists of simple graphs and morphisms between them. That is, $\texttt{sGraph} = \texttt{Set}^{\texttt{S}^{\mathrm{op}}}$ — the category of presheaves on the walking simple graphs category S.

11.3 The Simple Loopless Graph

Sometimes (and by that I mean almost always) graph theorists like the incidence relation to be irreflexive—loops are therefore forbidden, that is, Edge(x, x) is always false for any $x \in V$—but this construction does not get us a good category. We are going to reserve the word "simple graph" to mean a simple graph with loops and the word "graph" for simple graphs with no loops. The following is a formal definition of a loopless simple graph—a "graph".

Definition 11.3.1 (Graph)**.** $G = (V, E, d)$ is said to be a graph if

$$d\colon E \longrightarrow \langle V\rangle^2, \text{ mapping } e \overset{d}{\longmapsto} (x, y) = (y, x) \text{ and } x \neq y.$$

That is, Edge(x, y) if and only if Edge(y, x), and Edge(x, x) is always false.

From now on we are mostly going to assume the existence of the incidence relation map d (symmetric and irreflexive) and simply denote a graph G by (V, E)—where we assume that G is a simple loopless graph.

Definition 11.3.2 (Morphism of graphs)**.** Let $G = (V, E)$ and $H = (V', E')$ be graphs. A morphism $G \to H$ is a map $\varphi\colon V \to V'$ such that Edge(x, y) implies Edge$(\varphi(x), \varphi(y))$—that is, adjacency of vertices need to be preserved.

Definition 11.3.3 (Graph category)**.** We define the category of simple loopless graphs, denoted by `Graph`, as the category with graph objects and morphisms of graphs.

Lemma 11.3.4 (Stable preimage)**.** Let $\varphi\colon G \to H$ be a morphism of graphs, then, for any $x' \in \mathrm{Vert}(H)$, the collection $\varphi^{-1}(x') \subseteq \mathrm{Vert}(G)$ is stable in G.

[1] Where, for any $e \in E$, $e' \in E'$, and $(x, y) \in \langle V\rangle^2$, we define the maps $s \times t$, $s' \times t'$ and $\alpha_V \times \alpha_V$ as

- $e \xmapsto{s\times t} (s(e), t(e)) = (t(e), s(e))$.
- $e' \xmapsto{s'\times t'} (s'(e'), t'(e')) = (t'(e'), s'(e'))$.
- $(x, y) \xmapsto{\alpha_V\times\alpha_V} (\alpha_V(x), \alpha_V(y)) = (\alpha_V(y), \alpha_V(x))$.

Proof. Suppose, for the sake of contradiction, that there exists vertices $x, y \in \varphi^{-1}(x)$ such that $\mathrm{Edge}(x, y)$ is true in G. In particular, this implies that $\mathrm{Edge}(\varphi(x), \varphi(y)) = \mathrm{Edge}(x', x')$ is true in H, which is forbidden—hence there can be no such pair of vertices in $\varphi^{-1}(x')$. ♮

Definition 11.3.5 (Graph property)**.** A class of graphs is said to be a property of graphs if it is closed up to isomorphism.

Definition 11.3.6 (Graph invariant)**.** A map $\phi\colon \mathtt{Graph} \to S$—where $S \in \mathtt{Set}$, possibly $\mathbf{R}$ for instance—is said to be a graph invariant if for all $G, G' \in \mathtt{Graph}$ such that $G \simeq G'$ then

$$\phi(G) = \phi(G').$$

Definition 11.3.7 (Union and intersections)**.** Let $G = (V, E), G' = (V', E') \in \mathtt{Graph}$. We define their union as the graph

$$G \cup G' = (V \cup V', E \cup E').$$

Analogously, their intersection is defined as

$$G \cap G' = (V \cap V', E \cap E').$$

Definition 11.3.8 (Complete Graph)**.** A graph G is said to be complete if for all pairs of vertices $x, y \in G$, $\mathrm{Edge}(x, y)$ is true.

Notation 11.3.9. We denote by K^n the complete graph on n vertices.

Definition 11.3.10 (Subgraph)**.** Let $G = (V, E)$ and $G' = (V', E')$ be graphs. G' is said to be a *subgraph* of G—denoted $G' \subseteq G$—if $V' \subseteq V$ and $E' \subseteq E$. On the other hand, G is said to be a *supergraph* of G'.

Definition 11.3.11 (Induced subgraph)**.** Let $G' = (V', E')$ be a subgraph of G. G' is said to be an *induced subgraph* of G if for all $x, y \in V$, then $\mathrm{Edge}_{G'}(x, y)$ if $\mathrm{Edge}_G(x, y)$. We denote $G' = G[V']$. If $G[V'] = G$, then we say that G' is a *spanning subgraph* of G.

Definition 11.3.12 (Embedding of graphs)**.** Let G and H be graphs. A morphism of graphs $\iota\colon H \hookrightarrow G$ is said to be an embedding of H into G if its underlying map $\mathrm{Vert}(H) \rightarrowtail \mathrm{Vert}(G)$ is injective.

Definition 11.3.13 (Deletion vertices or edges)**.** Let $G = (V, E)$ be a graph and $V' \subseteq V$ a subset of the vertices of G. The deletion of the parts of G associated with the vertices V' is defined as the graph $G[V \setminus V']$—which we'll shortly identify as $G \setminus V'$. Equivalently we define the deletion of edges $E' \subseteq E$ as the graph $G - E' := (V, E \setminus E')$.

Notice that the non-connected vertices are not deleted from the resulting graph $G \setminus E'$, contrary to the case of the deletion of vertices $G \setminus V'$, where edges with endpoints out of V' where removed.

Definition 11.3.14 (Edge-maximal)**.** Given a graph property P, a graph $G = (V, E)$ is said to be *edge-maximal* with respect to P if $G \in P$ and for all $E' \supsetneq E$ the graph (V, E') does not have the property P.

Definition 11.3.15 (Clique)**.** Let G be a graph and $S \subseteq \operatorname{Vert}(G)$ be a subset with k vertices. If the graph $G[S]$ is complete, we call it a k-clique. Trivially, $G[S] \simeq K_k$.

Definition 11.3.16 (Complement graph)**.** Given a graph $G = (V, E)$, the complement of G is defined as the graph $\overline{G} = (V, \langle V \rangle^2 \setminus E)$.

Definition 11.3.17 (Line graph)**.** Given a graph G, the line graph of G, denoted by $L(G)$, is the graph such that, if $e = (x, y), e' = (y, z) \in \operatorname{Edge}(G)$ are adjacent edges then $\operatorname{Edge}_{L(G)}(x, z)$—that is $L(G) = (\operatorname{Vert}(G), E)$ where

$$E = \{(x, z) \in \langle V \rangle^2 : \operatorname{Edge}_G(x, y) \text{ and } \operatorname{Edge}_G(y, z) \text{ for some } y \in \operatorname{Vert}(G)\}.$$

Definition 11.3.18 (Join)**.** Let $G = (V, E), G' = (V', E') \in \texttt{Graph}$ be disjoint—$G \cap G' = \varnothing$. We define their join as the graph $G * G'$ with vertices $V \cup V'$ and edges $E \cup E' \cup E^*$, where E^* is defined as $E^* = \{(x, y) : x \in G, y \in G'\}$—that is, the collection of all edges connecting the vertices of G and G'.

The join definition fails to be a coproduct in the category of graphs exactly because of the addition of the E^* edges—the edges between elements of G with G' in $G * G'$ cannot be ensured to be preserved by a morphism $G * G' \to H$ as described in Proposition 11.5.4.

Vertex Degree

Notation 11.3.19 (Neighbourhood)**.** Given a graph G, we denote by $N(v)$ the collection of all neighbours of $v \in G$.

Definition 11.3.20 (Degree)**.** Let G be a graph, we define the degree (or valency) of a vertex $v \in G$ as $\deg v = |N(v)|$. If G is a finite graph,

- The minimum degree of G is defined as $\delta(G) = \min_{v \in \operatorname{Vert}(G)} \deg v$.
- The maximum degree of G is defined as $\Delta(G) = \max_{v \in \operatorname{Vert}(G)} \deg v$.
- The average degree of G is defined as

$$\deg G = \frac{1}{|G|} \sum_{v \in \operatorname{Vert}(G)} \deg v.$$

- The edge-vertex ratio of G is $\epsilon(G) = \frac{\|G\|}{|G|}$. Since an edge is composed of two vertices, the sum $\sum_{v \in \operatorname{Vert}(G)} \deg v$ counts every edge exactly twice, so that $\epsilon(G) = \frac{1}{2} \sum_{v \in \operatorname{Vert}(G)} \deg v$. Thus we arrive in the relation $\epsilon(G) = \frac{1}{2} \deg G$.

Lemma 11.3.21. The number of vertices of odd degree in a graph is even.

Proof. Let G be a graph. Since $\|G\| \in \mathbf{N}$ and $\|G\| = \frac{1}{2}\sum_{v \in \mathrm{Vert}(G)} \deg v$, then $\sum_{v \in \mathrm{Vert}(G)} \deg v$ is even. If the number of odd degree vertices where odd, then the sum of their degrees would also be odd. Since the sum of even degree vertices is always even, it is necessary for the number of odd degree vertices to be even. ♮

Proposition 11.3.22 (Edge-dense subgraph)**.** Let G be a finite graph with size $\|G\| \geqslant 1$. Then there exists a subgraph $H \subseteq G$ such that

$$\delta(H) > \epsilon(H) \geqslant \epsilon(G).$$

Proof. We'll construct the subgraph H by means of a chain of consequent single vertex deletion—as in Definition 11.3.13. Let $H_0 = G$ Notice that, if we want to have the edge-vertex ratio unaltered (or increased) by the deletion, we got to choose a vertex $v \in H_0$—if existent—such that the graph $H_1 = H_0 \setminus v$ has $\epsilon(H_1) \geqslant \epsilon(H_0)$, that is, $\deg v$ has to satisfy

$$\epsilon(H_1) = \frac{\|H_0\| - \deg v}{|H_0| - 1} \geqslant \frac{\|H_0\|}{|H_0|} = \epsilon(H_0).$$

Solving the above equation for the degree of v, we find that $\deg v \leqslant \epsilon(H_0)$. Since $\delta(H_0) \leqslant \deg H_0 \leqslant \Delta(H_0)$ and $\epsilon(H_0) \leqslant \deg H_0$, then $\deg v \leqslant \Delta(H_0)$—that is, $\delta(H_1) \leqslant \delta(H_0)$. If such vertex doesn't exists in $\mathrm{Vert}(H_0)$, we terminate our algorithm and find $H = H_0 = G$. Otherwise, we continue recursively, generating a finite chain of subgraphs—the finiteness of the chain comes from the fact that G itself is finite

$$G = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_n = H$$

for some $n \in \mathbf{N}$. Where $\epsilon(H_{j+1}) \geqslant \epsilon(H_j)$ and $\delta(H_{j+1}) \geqslant \delta(H_j)$ for all $0 \leqslant j < n$. At some point the chain will terminate into the subgraph H. Moreover, since the recursion terminates in $H \neq \varnothing$, then necessarily $\deg v > \epsilon(H)$ for all $v \in H$—in particular, this implies in $\delta(H) > \epsilon(H)$. ♮

Definition 11.3.23 (Regular graph)**.** A graph G is said to be k-regular—for some $k \in \mathbf{N}$—if $\deg v = k$ for all $v \in G$.

11.4 Path to Glory

Definition 11.4.1 (Walk)**.** A walk on a graph G is defined to be a subgraph $W \subseteq G$—assume $|W| = n + 1$ for some $n \in \mathbf{N}$—such that there exists a *surjective map* $\ell \colon [n] \twoheadrightarrow \mathrm{Vert}(W)$ satisfying the condition of $\mathrm{Edge}_G(\ell(k), \ell(k+1))$ for all $0 \leqslant k < n$. We'll generally assume this underlying surjective vertex-labelling map and denote $W = (\ell(0), \dots, \ell(n)) = (v_0, \dots, v_n)$. The walk W is said to be closed if $v_0 = v_n$.

Definition 11.4.2 (Length)**.** We define the length of a walk W to be its size $\|W\|$.

Definition 11.4.3 (Walk operations)**.** Given a path $W = (v_0, \dots, v_n)$ we can define the following:

- (Subwalk) Let $v_j \in \mathrm{Vert}(W)$ be any vertex. We define a subwalk of W restricted to v_j to be one of the following subgraphs of W: $v_j W = (v_j, v_{j+1}, \dots, v_n)$ and $W v_j = (v_0, \dots, v_{j-1}, v_j)$. We can also define excluding subwalk induced by v_j to be $\breve{v}_j W = (v_{j+1}, \dots, v_n)$ and $W \breve{v}_j = (v_0, \dots, v_{j-1})$.
- (Inner walk) The inner walk of W is given by $\breve{W} = (v_1, \dots, v_{n-1})$.
- (Walk concatenation) Let $Q = (w_0, \dots, w_m)$ be a walk. If P and Q have coinciding endings, say $v_n = w_0 = x$, we can define the concatenation of P with Q to be the walk $PQ = (v_0, \dots, v_{n-1}, x, w_1, \dots, w_m)$.

Definition 11.4.4 (Path)**.** We define a path to be walk with no repeating vertices. A path graph P_n is a graph of order $n+1$ and size n such that there exists a *bijective map* $\ell_n \colon [n] \xrightarrow{\simeq} \mathrm{Vert}(P_n)$ for which $\mathrm{Edge}_{P_n}(\ell_n(k), \ell_n(k+1))$ for all $0 \leqslant k < n$. We'll usually assume the existence of the underlying vertex-labelling bijection ℓ_n. This way we can naturally write the path P_n as a unique sequence of vertices $(v_0, v_1, \dots, v_n)$ *up to isomorphism* of graphs.

Given a graph G, we define a path of size $n+1$ *on* G to be the *induced subgraph* of G given by $G[\iota(P_n)]$, where $\iota \colon P_n \hookrightarrow G$ is an *embedding of graphs*—see Definition 11.3.12. The underlying labelling is now given by the composition $\iota \ell_n \colon [n] \to \mathrm{Vert}(G)$.

Since a path is just a special type of walk, every operation described on Definition 11.4.3 is carried over to the path graphs.

Proposition 11.4.5 (Transforming paths into walks)**.** A map $\phi \colon P_n \to G$ is a morphism of graphs if and only if the sequence $(\phi(v_0), \dots, \phi(v_n))$ is a walk on G.

Proof. Throughout, assume $0 \leqslant j < n$. If ϕ is a morphism of graphs then $\mathrm{Edge}_{P_n}(v_j, v_{j+1})$ implies $\mathrm{Edge}_G(\phi(v_j), \phi(v_{j+1}))$ thus inducing a walk on G. On the other hand, if $(\phi(v_0), \dots, \phi(v_n))$ is a walk on G, then since the only edges on P_n occur between consequent vertices v_j and v_{j+1}, we have that ϕ indeed preserves the adjacency structure of the path P_n, that is, $\mathrm{Edge}_{P_n}(v_j, v_{j+1})$ implies $\mathrm{Edge}_G(\phi(v_j), \phi(v_{j+1}))$—hence ϕ is a morphism. ♮

Lemma 11.4.6 (Paths on walks)**.** Let W be a walk. If $x, y \in \mathrm{Vert}(W)$, then there exists a path $P \subseteq W$ such that $x, y \in \mathrm{Vert}(P)$.

Proof. Assume W is a k-walk and contains x and y at its end-vertices—if not, we would be analogously analysing the subwalk xWy. Let $\ell \colon [k] \twoheadrightarrow \mathrm{Vert}(W)$ be a surjective labelling on W—since x and y are end-vertices, we have

$$\ell(0) = x \quad \text{and} \quad \ell(k) = y. \tag{11.1}$$

Let $\mathcal{C}$ be the collection of all cycles on W—refer to Definition 11.4.10. Define the equivalence relation $\sim$ on the vertices $\mathrm{Vert}(W)$ to be so that $v \sim u$ if and only if $v, u \in \mathrm{Vert}(C)$ for some $C \in \mathcal{C}$.

The quotient graph $W/{\sim}$ (see Definition 11.5.1) is a path joining the vertex classes $\overline{x}$ and $\overline{y}$. To see that, we can first construct an equivalence relation $\sim_\ell$ on the indexing set $[k]$, defined as follows: $i \sim_\ell j$ if and only if $\ell(i) \sim \ell(j)$—clearly $[k]/{\sim_\ell}$ is an indexing set for the vertices of the quotient graph. Lets assume that $[k]/{\sim_\ell} = \{\overline{0}, \dots, \overline{m}\}$. Let $\ell' \colon [k]/{\sim_\ell} \to \operatorname{Vert}(W/{\sim})$ be a labelling on $W/{\sim}$ defined as $\ell'(\overline{j}) = \overline{\ell(j)}$. Using Eq. (11.1), the sequence of vertices $(\ell'(\overline{0}), \dots, \ell'(\overline{m}))$ is such that $\ell'(\overline{0}) = \overline{x}$ and $\ell'(\overline{m}) = \overline{y}$. Also, $\operatorname{Edge}_{W/\sim}(\ell'(\overline{j}), \ell'(\overline{j+1})) = \operatorname{Edge}_W(\ell(j), \ell(j+1))$ is always true (because W is a walk). So far, we've shown that $W/{\sim}$ is a walk between $\overline{x}$ and $\overline{y}$—we now need to show that no vertex is repeated in the given labelled sequence, which is actually trivial. Suppose that $\ell'(\overline{i}) = \ell'(\overline{j})$, then from definition of ℓ' we have $\overline{\ell(i)} = \overline{\ell(j)}$, which implies that $\ell(i) \sim \ell(j)$ and hence $i \sim_\ell j$—thus $\overline{i} = \overline{j}$. This shows that

$$W/{\sim} = (\overline{x}, \ell'(\overline{1}), \dots, \ell'(\overline{m-1}), \overline{y})$$

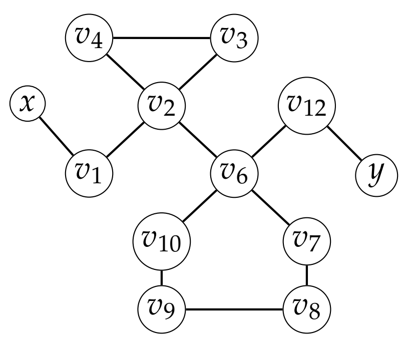
is indeed a path joining $\overline{x}$ and $\overline{y}$.

For the last part we just need to consider the embedding of graphs $\iota \colon W/{\sim} \hookrightarrow W$ such that $\operatorname{Edge}_{W/\sim}(\overline{v}, \overline{u})$ implies $\operatorname{Edge}_W(\iota(\overline{v}), \iota(\overline{u}))$. The sequence of vertices

$$\iota(W/{\sim}) \coloneqq (\iota(\overline{x}), \iota\ell'(\overline{1}), \dots, \iota\ell'(\overline{m-1}), \iota(\overline{y})) \subseteq W$$

is clearly a path on W joining the vertices $\iota(\overline{x}) = x$ and $\iota(\overline{y}) = y$. Hence, $P = W[\operatorname{im} \iota]$ gives us the wanted path—finally the proposition is proved. ♮

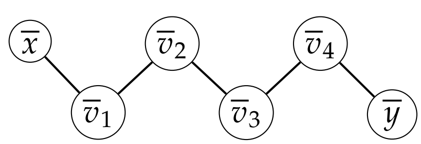
The last proof may got really clumsy at some points, so here goes an example of the basic operations we've developed and used throughout the proof. Let $W = (x, v_1, \dots, v_{12}, y)$ be a walk, visually given by—where $v_5 = v_2$ and $v_{11} = v_6$



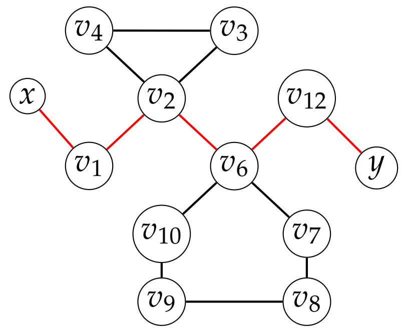
Now, the process of taking the quotient of the graph amounts to the identification of the cycles

$$C = \{W[\{v_2, v_3, v_4, v_5\}], W[\{v_6, v_7, v_8, v_9, v_{10}, v_{11}\}]\}$$

The cycles are respectively reduced to classes $\overline{v}_2$ and $\overline{v}_3$, both elements of $\operatorname{Vert}(W)/{\sim}$. The quotient graph $W/{\sim}$ can be visually depicted as follows



Now for the embedding of graphs $\iota \colon W/{\sim} \hookrightarrow W$, we can view $P = W[\operatorname{im} \iota]$ as the subgraph consisting of the red edges and their end-vertices—that is, the collection of vertices $\{x, v_1, v_2, v_6, v_{12}, y\}$—which can be visualised as follows



Corollary 11.4.7 (Distances & morphisms of graphs)**.** Consider graphs G and H. If $\phi\colon G \to H$ is a morphism of graphs then, for all $x, y \in G$, we have

$$d_H(\phi(x), \phi(y)) \leqslant d_G(x, y).$$

Proof. Assume ϕ is a morphism. Suppose $x, y \in G$ are separated by a finite distance—if not, the proposition follows trivially. Let P_k be a minimal path on G joining x and y. Since ϕ is a morphism, the image $\phi(P_k) \subseteq H$ is a walk where $\phi(x), \phi(y) \in P_k$—this follows directly from Proposition 11.4.5. Now, using Lemma 11.4.6 on $\phi(P_k)$ and the vertices $\phi(x)$ and $\phi(y)$, it follows that there exists a path $P \subseteq \phi(P_k)$ joining $\phi(x)$ and $\phi(y)$. Moreover, clearly $\|P\| \leqslant \|\phi(P_k)\| \leqslant k$, thus $d_H(\phi(x), \phi(y)) \leqslant d_G(x, y)$. ♮

Definition 11.4.8 (Path between sets)**.** Let A and B be disjoint sets of vertices. $P = (v_0, \dots, v_n)$ is said to be an A-B path if $v_0 \in A$ and $v_n \in B$.

Definition 11.4.9 (Independent paths)**.** Two paths are said to be independent if their inner path share no vertex.

Cycles

Definition 11.4.10. A cycle is a closed walk with size at least 3. A k-cycle is a cycle C^k whose length is k.

A k-cycle *on* a graph G is the *induced subgraph* of G given by $G[\iota(C^k)]$ such that $\iota\colon C^k \to G$ is an embedding of graphs.

Definition 11.4.11 (Chords)**.** Let G be a graph and C be a cycle on G. Given $x, y \in C$, a chord is an edge $(x, y) \in \mathrm{Edge}(G)$ such that $\mathrm{Edge}_C(x, y)$ is false.

Definition 11.4.12 (Induced cycle)**.** An induced cycle on a graph is a cycle with no chords.

Definition 11.4.13 (Distance)**.** Let G be a graph and $x, y \in \mathrm{Vert}(G)$ be any vertices. Let $\mathcal{P}$ be the collection of all paths in G containing the vertices x and y. If $\mathcal{P}$ is non-empty, we define the distance of x and y on G to be the minimum length of the subpaths linking x to y, in other words

$$d_G(x, y) = \min_{P \in \mathcal{P}} \|xPy\|.$$

On the other hand, if $\mathcal{P}$ is empty, then $d_G(x, y) = \infty$.

Definition 11.4.14 (Miscellaneous definitions)**.** We define the following:

(a). (Girth) The girth of a graph G is the minimum length of a cycle on G. In other words, let $\mathcal{C}$ be the collection of all cycles on G, the girth of G is

$$g(G) = \min_{C \in \mathcal{C}} \|C\|.$$

(b). (Circumference) The circumference of a graph G is the maximal length of a cycle on G. In other words, let $\mathcal{C}$ be the collection of all cycles on G. If $\mathcal{C}$ is non-empty, we set

$$\operatorname{circ}(G) = \max_{C \in \mathcal{C}} \|C\|.$$

(c). (Diameter) The diameter of a graph G is the maximal distance between two vertices of G, that is

$$\operatorname{diam}(G) = \max_{x, y \in \operatorname{Vert}(G)} d_G(x, y).$$

Proposition 11.4.15. Let G be a graph such that $\delta(G) \geqslant 2$. Then there exists a path on G with length $\delta(G)$ and a cycle on G with length at least $\delta(G) + 1$.

11.5 Universal Properties of `Graph`

Quotients

Definition 11.5.1 (Quotient)**.** Let G be a graph and $\sim$ be an equivalence relation on the collection of vertices $\operatorname{Vert}(G)$. We define the quotient graph $G/\sim$ to be the graph whose vertices are the vertex classes $\operatorname{Vert}(G)/\sim$ and whose edges are such that $\operatorname{Edge}_{G/\sim}([x], [y])$ if and only if $\operatorname{Edge}_G(x, y)$.

Proposition 11.5.2 (Universal property of quotients)**.** The quotient graph is a quotient in the category `Graph`. In other words, let G be a graph and $\sim$ be an equivalence relation on $\operatorname{Vert}(G)$. Consider any graph H together with a morphism of graphs $\psi \colon G \to H$. Then there exists a unique morphism of graphs $\phi \colon G/\sim \to H$ such that the following diagram commutes

$$\begin{tikzcd}
G \arrow[r, "\psi"] \arrow[d, two heads, "\pi"'] & H \\
G/\sim \arrow[ur, dashed, bend right, "\phi"'] &
\end{tikzcd}$$

where π is the naturally defined projection.

Proof. First we show that ϕ is indeed unique. Suppose ϕ_1 and ϕ_2 are both morphisms that satisfy the commutativity of the diagram. Given any $x \in \operatorname{Vert}(G)$ we have $\psi(x) = \phi_1([x]) = \phi_2([x])$ and since $\pi(\operatorname{Vert}(G)) = \operatorname{Vert}(G/\sim)$ (surjective property) then ϕ_1 and ϕ_2 have equal images throughout their whole domain — implying $\phi_1 = \phi_2$. To show that ϕ is a morphism of graphs, it is sufficient to consider any $x, y \in G$ such that $\operatorname{Edge}_G(x, y)$: ψ being a morphism implies $\operatorname{Edge}_G(\psi(x), \psi(y))$, then — since $\phi([x]) = \psi(x)$ and $\phi([y]) = \psi(y)$ — we get that $\operatorname{Edge}_G(\phi([x]), \phi([y]))$ is true. $\natural$

Coproducts

Definition 11.5.3 (Coproduct)**.** Let $G = (V, E)$ and $H = (V', E')$ be graphs, we define their coproduct $G \oplus H$ to be the disjoint union of vertices and edges of the original graphs — that is, $G \oplus H = (V \amalg V', E \amalg E')$[2].

Proposition 11.5.4 (Universal property of coproducts)**.** Given graphs $G, H \in \mathtt{Graph}$, the graph $G \oplus H$ is a coproduct of G and H in the category of simple loopless graphs. That is, given any graph $W \in \mathtt{Graph}$ and graph morphisms $f\colon G \to W$ and $g\colon H \to W$, there exists a unique graph morphism $\phi\colon G \oplus H \to W$ such that the following diagram commutes

$$\begin{tikzcd}
G \arrow[dr, hook, "\iota_G"] \arrow[ddr, bend right, "f"'] & & H \arrow[dl, hook', "\iota_H"'] \arrow[ddl, bend left, "g"] \\
& G \oplus H \arrow[d, dashed, "\phi"] & \\
& W &
\end{tikzcd}$$

Where the inclusion morphisms ι_G and ι_H are naturally defined.

Proof. Notice that ϕ is defined by a map of vertices $\mathrm{Vert}(G) \cup \mathrm{Vert}(H) \to \mathrm{Vert}(W)$ with the restriction that $\mathrm{Edge}_{G\oplus H}(x, y)$ implies $\mathrm{Edge}_W(\phi(x), \phi(y))$, we'll first prove that f together with g completely identify ϕ. Let $v \in G \oplus H$ be any vertex. If $v \in \mathrm{Vert}(G)$, suppose $f(v) = h$ then necessarily $\phi\iota_G(v) = h$. On the other hand, if $v \in \mathrm{Vert}(H)$ and if $g(v) = h'$ then $\phi\iota_H(v) = h'$. Since $\operatorname{cod}\phi = \mathrm{Vert}(G) \amalg \mathrm{Vert}(H)$ and we have $\iota_G(G) = \mathrm{Vert}(G)$ and $\iota_H(H) = \mathrm{Vert}(H)$, this shows that f and g completely determine the image of ϕ — hence ϕ is uniquely defined.

Finally we show that ϕ is indeed a morphism of graphs. Given $x, y \in \mathrm{Vert}(G)$, suppose $\mathrm{Edge}_G(x, y)$, then $\mathrm{Edge}_W(f(x), f(y)) = \mathrm{Edge}_W(\phi(x), \phi(y))$ — since $\phi(x) = f(x)$ and $\phi(y) = f(y)$. The case $x, y \in \mathrm{Vert}(H)$ is completely analogous and we'll therefore omit for the sake of brevity. We don't need to inspect the case where $x \in \mathrm{Vert}(G)$ and $y \in \mathrm{Vert}(H)$ since $\mathrm{Edge}_{G\oplus H}(x, y)$ is always false in such instance. Thus ϕ is a morphism of graphs. ♮

Products

Definition 11.5.5 (Kronecker product)**.** Let $G, H \in \mathtt{Graph}$. We define the Kronecker product of G and H to be the graph $G \otimes H = (V, E)$ whose vertices are $V = \mathrm{Vert}(G) \times \mathrm{Vert}(H)$ and edges defined by $\mathrm{Edge}_{G\otimes H}(v \otimes h, v' \otimes h')$ if and only if $\mathrm{Edge}_G(v, v')$ and $\mathrm{Edge}_H(h, h')$.

Proposition 11.5.6 (Products)**.** The Kronecker product as defined above is a product in the category of simple loopless graphs $\mathtt{Graph}$. That is, given graphs $G, H, W \in \mathtt{Graph}$

[2] $\amalg$ denotes the standard disjoint union in $\mathtt{Set}$.

and graph morphisms $f \colon W \to G$ and $g \colon W \to H$, there exists a unique morphism of graphs $\phi \colon W \to G \otimes H$ such that the following diagram commutes

$$\begin{tikzcd} & W \arrow[d, dashed, "\phi"] \arrow[ddl, bend right, "f"'] \arrow[ddr, bend left, "g"] & \\ & G \otimes H \arrow[dl, "\pi_G"] \arrow[dr, "\pi_H"'] & \\ G & & H \end{tikzcd}$$

Where π_G and π_H are the naturally defined projection morphisms.

Proof. Consider the morphism $\phi \colon W \to G \otimes H$ defined by the mapping $w \overset{\phi}{\mapsto} f(w) \otimes g(w)$. Let $w, w' \in \mathrm{Vert}(W)$ be any vertices such that $\mathrm{Edge}_W(w, w')$, since f and g are graph morphisms, then $\mathrm{Edge}_G(f(w), f(w'))$ and $\mathrm{Edge}_H(g(w), g(w'))$. Since $\phi(w) = f(w) \otimes g(w)$ and $\phi(w') = f(w') \otimes g(w')$ it follows from the construction of the Kronecker product that $\mathrm{Edge}_{G \otimes H}(\phi(w), \phi(w'))$. This shows that ϕ is a graph morphism.

We now inspect its uniqueness. Let $\phi, \phi' \in \mathrm{Mor}_{\mathsf{Graph}}(W, G \otimes H)$ be morphisms satisfying the commutativity of the diagram. Then $f = \pi_G \phi = \pi_G \phi'$ and $g = \pi_H \phi = \pi_H \phi'$, thus, given any $w \in W$ we have $\pi_G \phi(w) = \pi_G \phi'(w)$ and $\pi_H \phi(w) = \pi_H \phi'(w)$, which implies in $\phi(w) = \phi'(w)$. Therefore ϕ is unique. ♮

11.6 Ramsey and Schur theorem

Theorem 11.6.1 (Ramsey theorem on graphs)**.** Given $r \in \mathbf{N}_{\geqslant 1}$ we have that the colouring function $C \colon \binom{\mathbf{N}}{2} \to [r]$ is such that there exists an infinite monochromatic subset $A \subseteq \mathbf{N}$, that is, for all $a, b \in A$, we have $C(ab) = i \in [r]$.

Proof. Firstly, define the sets $\Gamma_k(x) := \{y \in \mathbf{N} : C(xy) = k \in [r]\}$. We know that for any element $x \in \mathbf{N}$ the colouring function C provides at most r of such sets. Notice that for every vertex $v_i \in \mathbf{N}$ there are infinitely many edges connecting to other vertices of $\mathbf{N}$ this way we conclude that since $[r]$ partitions $\mathbf{N}$ only into a finite number of subsets, by the pigeonhole principle, we can conclude that there is one set, which we'll denote by Γ_{k_i}, that is infinite.

Let now the sequence $(v_n)_{n \in \mathbf{N}}$ such that for all $n \in \mathbf{N}$, $v_n \in \Gamma_{k_{n-1}}(v_{n-1})$ for some colour $k_{n-1} \in [r]$. A trivial result of such construction is that given any $n \in N$ we have $C(v_n v_{n-1}) = k_{n-1}$. In fact we can extend such result by simply recalling that since we are always taking the infinite sets, surely we'll get the sequence

$$(\Gamma_{k_n}(v_n))_{n \in \mathbf{N}} \text{ such that } \mathbf{N} \supseteq \Gamma_{k_1}(v_1) \supseteq \Gamma_{k_2}(v_2) \supseteq \dots$$

What this means is that given any pair $\{v_i, v_j\}$ of elements of such sequence, we have that $C(v_i, v_j) = k_{\min\{i,j\}}$. Define now the sequence $\left(k_{\min\{i,j\}}\right)_{\{i,j\} \in \binom{\mathbf{N}}{2}}$, since this is a sequence with infinitely many elements, we can say that the finite colouring $[r]$ induces a finite partitioning of such sequence. By the pigeonhole principle we conclude finally that there exists an infinite monochromatic subsequence. With this we conclude the proof since this construction is obtained directly from $(v_n)_{n\in\mathbf{N}}$. ♮

We actually didn't stated the general version of Ramsey theorem, which extends for the colouring of the set $\binom{X}{k}$ where X is a countably infinite poset and $k \in \mathbf{N}$.

Definition 11.6.2 (Hypergraph)**.** An ordered pair $H = (V, E)$ of sets, where V are the vertices and E the hyperedges, is called an **hypergraph** if $E \subseteq \binom{V}{k}$, thus an edge connect two or more vertices.

Theorem 11.6.3 (Ramsey theorem on Hypergraphs)**.** Let X a countably infinite poset. Then, for all $k, r \in \mathbf{N}$, the colouring $C\colon \binom{X}{k} \to [r]$ is such that there exists an infinite subset $A \subseteq X$ such that, for all $a, b \in A$, we have $C(ab) = i \in [r]$.

Proof. We proceed via induction on k. For the case $k = 1$ we have that $\binom{X}{1} = X$ and thus the colouring C induces a partition of the infinite set X into a finite number of subsets. By the pigeonhole principle we conclude that there is one of such subsets that is infinite and also monochromatic.

For the inductive step, let $\theta \in X$ be the smallest element of X (we can do such a thing since we said X was a poset), let the colouring $C\colon \binom{X}{k+1} \to [r]$ and another colouring

$$C_0\colon \binom{X \setminus \{\theta\}}{k} \longrightarrow [r] \text{ such that } C_0(E) := C(E \cup \{\theta\}), \text{ for all } E \in \binom{X}{k}.$$

By the inductive hypothesis we say that there exists an infinite subset $X_0 \subseteq X$ such that, for all $E_0 \in \binom{X_0}{k}$, we have $C_0(E_0) = c_0 \in [r]$.

Now let $X_0' := \{x \in X_0 : \theta < x\}$ and define $x_1 := \min(X_0)$. As before, set the colouring $C_1\colon \binom{X_0'}{k} \to [r]$ such that $C_1(E) := C(E \cup \{x_1\})$. Since X_0' is an infinite set by construction, we know from the pigeonhole principle that there exists an infinite subset $X_1 \subseteq X_0'$ such that, for all $E_1 \in \binom{X_1}{k}$, we have $C_1(E_1) = c_1 \in [r]$.

We continue constructing a sequence of elements $(\theta, x_1, x_2, \dots)$ in such a way that we have a corresponding sequence $(c_0, c_1, c_2, \dots)$ of colours. Note now that we can construct a new and final colouring function

$$C'\colon \{\theta, x_1, x_2, \dots\} \to [r] \text{ such that } C'(x_j) := c_j \in [r].$$

There are countably infinite elements on the domain and only finite colours, the partitioning is such that the pigeonhole principle is applicable and thus there is a infinite subset $A \subseteq \{\theta, x_1, x_2, \dots\} \subseteq X$ such that for all $x \in A$ we have $C'(x) = c \in [r]$. This concludes the proof. ♮

Part IV

Classical Topology

Chapter 12

Topological Spaces

12.1 Topology

Definition 12.1.1 (Topology)**.** Let X be a set and $\tau \subseteq 2^X$. We say that τ is a topology for X if the following properties are satisfied

(T1) $X, \varnothing \in \tau$.

(T2) The arbitrary union of elements of τ is an element of τ.

(T3) The finite intersection of elements of τ is an element of τ.

The elements of the topology τ are called open sets of X.

Example 12.1.2. We proceed by listing some examples of topologies that are somewhat interesting, they are included here in order to familiarize the reader with the possible constructions for the topology on a given set X:

- $\tau_1 = \{U \subseteq X : U = \varnothing \text{ or } X \setminus U \text{ is finite}\}$ is the *cofinite topology* on X.
- $\tau_2 = \{U \subseteq X : U = \varnothing \text{ or } X \setminus U \text{ is countable}\}$ is the *cocountable topology* on X.
- Let $p \in X$, then $\tau_3 = \{U \subseteq X : U = \varnothing \text{ or } p \in U\}$ is the *particular point topology* on X.
- Let $p \in X$, then $\tau_4 = \{U \subseteq X : U = X \text{ or } p \notin U\}$ is the *excluded point topology* on X.
- The collection 2^X forms what is called the *discrete topology* on X.

Definition 12.1.3 (Comparing topologies)**.** Let X be a set, and τ and τ' be topologies on X. If $\tau' \supseteq \tau$, then we say that

- τ' is finer than τ. Moreover, if τ is strictly contained in τ', we say that τ' is strictly finer than τ
- τ is coarser than τ'. Moreover, if τ is strictly contained in τ', we say that τ' is strictly coarser than τ

In general, we say that two topologies are comparable if either of them contains the other.

Topological Basis

Definition 12.1.4 (Basis)**.** Let X be a topological space. A collection $\mathcal{B} \subseteq 2^X$ is said to be a basis for the topology of X if it satisfies the following

(B1) Every element of $\mathcal{B}$ is an open set of X.

(B2) Every open subset $U \subseteq X$ can be written as a union of elements of B, that is, exists $\{B_i\}_{i\in I} \subseteq \mathcal{B}$ for which $U = \bigcup_{i\in I} B_i$.

Proposition 12.1.5 (Necessary and sufficient condition for a basis)**.** Let X be a set and $\mathcal{B} \subseteq 2^X$. Then $\mathcal{B}$ is a basis for some topology of X if and only if it satisfies

B1 $X = \bigcup_{B\in\mathcal{B}} B$.

B2 If $x \in A \cap B$, where $A, B \in \mathcal{B}$, then there exists $C \in \mathcal{B}$ such that $x \in C \subseteq A \cap B$.

Proof. Let $\mathcal{B}$ be a basis for the space X. Let $p \in X$ be any point and let $U \subseteq X$ be a neighbourhood of x. From the definition of a basis, there exists a subcollection of sets such that their union equals U, which implies in the existence of $B \in \mathcal{B}$ such that $x \in B$. Moreover, since $B \subseteq X$ for all $B \in \mathcal{B}$, it follows that $X = \bigcup_{B\in\mathcal{B}} B$. Let $A, B \subseteq \mathcal{B}$ be any intersecting sets of $\mathcal{B}$ and take any point $p \in A \cap B$. Notice that $A \subseteq B \subseteq X$ is open, hence there exists a subcollection of open sets of the basis $\mathcal{B}$ whose union equals $A \cap B$. It is immediate that there exist $C \in \mathcal{B}$ such that $p \in C$ and necessarily $C \subseteq A \cap B$.

Let $\mathcal{B} \subseteq 2^X$ satisfying both conditions specified above. We first show that $\mathcal{B}$ is a collection of open sets. Let τ be the collection of all possible unions of sets of $\mathcal{B}$. From the first property, $X \in \tau$ and clearly $\varnothing \in \tau$ — satisfying Item 1. From construction, unions of sets in τ are unions of unions of sets of $\mathcal{B}$, which is certainly contained in τ — hence the collection satisfies Item 2. Let $T, T' \in \tau$ be any intersecting sets and, for every $p \in T \cap T'$, choose any $B, B' \subseteq \mathcal{B}$ such that $p \in B \subseteq T$ and $p \in B' \subseteq T'$ — which are ensured to exist. From the second property of $\mathcal{B}$, there exists $C \in \mathcal{B}$ such that $p \in C \subseteq B \cap B'$, thus $C \subseteq B \cap B'$ and, in particular $C \subseteq T \cap T'$. We can see that $T \cap T'$ is again the union of a collection of elements of $\mathcal{B}$, hence τ is closed under finite intersections, satisfying Item 1. We can now finally conclude that τ is a topology on X and hence $\mathcal{B}$ is composed of open sets of X — and, even better than that, τ is the unique topology generated by $\mathcal{B}$. ♮

Definition 12.1.6 (Subbase)**.** Let (X, τ) be a topological space. A collection $\mathcal{S} \subseteq \tau$ is called a subbase for (X, τ) if the collection of all finite intersections $U_1 \cap \cdots \cap U_n$, where $U_i \in \mathcal{S}$, is a base for (X, τ).

Definition 12.1.7 (Weight)**.** Let X be a topological space and $\mathfrak{B}$ be the collection of all bases for the topology of X. We define the weight of X as

$$w(X) = \min_{\mathcal{B}\in\mathfrak{B}} |\mathcal{B}|.$$

Definition 12.1.8 (Basis at a point)**.** Let X be a topological space and $p \in X$ be any fixed point. We define the collection $\mathcal{B}_p \subseteq 2^X$ of neighbourhoods of p to be the *neighbourhood basis for the topology of X at p* if for any neighbourhood $U_p \subseteq X$, there exists $B \in \mathcal{B}_p$ such that $B \subseteq U_p$.

Definition 12.1.9. Let (X, τ) be a topological space. Let $x \in X$ be any point and consider $\mathfrak{B}_x$ the collection of all bases at x. Then we define the character of X at the point x as

$$\chi(x, (X, \tau)) = \min_{\mathcal{B}_x \in \mathfrak{B}_x} |\mathcal{B}_x|$$

Closed and Open Sets

The notion of a closed set and an open set are closely related — pardon for the pun. They have a dual relationship, allowing for us to define topologies via either of them. Lets first define what we mean by a closed set.

Definition 12.1.10 (Closed set)**.** Let X be a topological space. We define a set $A \subseteq X$ to be closed if $X \setminus A$ is open.

The following proposition realizes the idea that the duality of open and closed sets allow us to work with topological spaces by analysing both open and closed elements of the space of interest.

Proposition 12.1.11. If X is a topological space, then

1. The sets X and $\varnothing$ are closed.
2. The finite union of closed sets is closed.
3. The arbitrary intersection of closed sets is closed.

Proof. Notice that $X \setminus X = \varnothing$ and $X \setminus \varnothing = X$ are both open sets from Definition 12.1.1, hence X and $\varnothing$ are closed. Let $\{C_j\}_{j=1}^n$ be a finite collection of closed sets, then $X \setminus \bigcup_{j=1}^n C_j = \bigcap_{j=1}^n X \setminus C_j$ but since $X \setminus C_j$ is open for all j, then their finite intersection is open and hence $X \setminus \bigcup_{j=1}^n C_j$ is also open, which implies by definition that $\bigcup_{j=1}^n C_j$ is closed. Consider now any collection of closed sets $\{C_j\}_{j \in J}$ — where J is possibly infinite. Then $X \setminus \bigcap_{j \in J} C_j = \bigcup_{j \in J} X \setminus C_j$ is open by the arbitrary union of open sets being open, hence $\bigcap_{j \in J} C_j$ is closed. ♮

We now define four important operations on sets of a topological space, which will accompany us for the rest of these notes on general point-set topology.

Definition 12.1.12. Let X be a topological space and $A \subseteq X$ be a set. We define

(a) The closure of A is the least closed set, $\operatorname{Cl} A$, that contains A. This can be equivalently described as

$$\operatorname{Cl} A = \bigcap \{F \subseteq X : A \subseteq F \text{ and } F \text{ is closed}\}.$$

(b) The interior of A is the biggest open set $\operatorname{Int} A$ contained in A. That is

$$\operatorname{Int} A = \bigcup \{U \subseteq X : U \subseteq A, U \text{ is open}\}.$$

(c) The exterior of A in X is defined as

$$\operatorname{Ext} A = X \setminus \operatorname{Cl} A.$$

(d) The boundary of A in X is defined as

$$\partial A = X \setminus (\operatorname{Int} A \cup \operatorname{Ext} A).$$

Remark 12.1.13. For the remainder of this section, unless specified on the contrary, we let X be any topological space and $A \subseteq X$ be any subset of X.

Proposition 12.1.14. $\operatorname{Int} A$ and $\operatorname{Ext} A$ are open sets, while on the other hand $\operatorname{Cl} A$ and ∂A are closed sets.

Proof. Since $\operatorname{Int} A$ is the union of a collection of open sets, $\operatorname{Int} A$ is open. $\operatorname{Cl} A$ is the intersection of a collection of closed sets, hence $\operatorname{Cl} A$ is closed. The exterior set $\operatorname{Ext} A$ is simply the complement of a closed set, hence it's open. The boundary of A is the complement of the union of open sets (which is open), hence ∂A is closed. ♮

Proposition 12.1.15. The following are equivalences on the definitions of open and closed sets.

(a) Open set equivalences:

- A is open.
- $A = \operatorname{Int} A$.
- A contains none of its boudary points, i.e. $A \cap \partial A = \varnothing$.
- For all $x \in A$ there exists a neighbourhood $U \subseteq A$ containing x.

(b) Closed set equivalences:

- A is closed.
- $A = \operatorname{Cl} A$.
- For all $x \in X \setminus A$, there exists a neighbourhood $U \subseteq X \setminus A$ of x.

Proof. (a) Let A be open, then, in particular, $A \in \{U \subseteq X : U \subseteq A, U \text{ is open}\}$, thus $A = \operatorname{Int} A$. Moreover, from the definition of ∂A it follows that $A \cap \partial A = \operatorname{Int} A \cap A \setminus (\operatorname{Int} A \cup \operatorname{Ext} A) = \varnothing$. On the other hand, if $p \in A$ is any point, then A itself is a neighbourhood of p.

In order to finish the equivalence chain, let A be such that all of its points have a neighbourhood contained in A. We can then define the collection of neighbourhoods $\mathcal{U} = \{U_p \subseteq A : p \in A, p \in U\}$. Notice that $A \subseteq \bigcup_{U \in \mathcal{U}} U$ and the opposite inclusion is clearly true — thus A is the union of a collection of open sets of X, hence A is open.

(b) Let A be closed, then A is clearly the least closed set containing itself, implying in $A = \operatorname{Cl} A$. Let $x \in X \setminus A$ be any point. Since $X \setminus A$ is open, we use the previous equivalence for open sets to conclude that there exists $U \subseteq X \setminus A$ neighbourhood of x. To conclude the equivalence chain, suppose the last property is true for $X \setminus A$. Then $X \setminus A$ is open by the previous item, which, in turn, this implies that A closed. ♮

Proposition 12.1.16 (Basis criterion for open sets)**.** Let X a topological space and $\mathcal{B}$ a base for the topology of X. A set $A \subseteq X$ is open if and only if for all points $p \in A$ there exists a neighbourhood of p, $B_p \in \mathcal{B}$, such that $B_p \subseteq A$.

Proof. Let A be open, then there exists a collection of elements of $\mathcal{B}$ whose union is A — which implies that there exists, for all $p \in A$, a set $B \subseteq \mathcal{B}$ such that $p \in B$ and $B \subseteq A$. For the contrary, if $p \in A$ is any point and $B \in \mathcal{B}$ is the corresponding neighbourhood $p \in B \subseteq A$, then from Proposition 12.1.15 we find that A is open. ♮

Proposition 12.1.17. The following propositions are equivalent, regarding points on the closure of a set A

(a) $x \in \operatorname{Cl} A$.

(b) Every neighbourhood $U \subseteq X$ of x is such that $U \cap A \neq \varnothing$.

(c) There exists a basis $\mathcal{B}_x$ at the point x such that for all $U \in \mathcal{B}_x$ we have $U \cap A = \varnothing$.

Proof. (a) implies (b): Let $x \in X$ and suppose that (b) is false for x, so that there exists $U \subseteq X$ neighbourhood of x such that $U \cap A = \varnothing$. Then $A \subseteq X \setminus U$, that is, A is a subset of the complement of U. Since U is open, then $X \setminus U \in C_A := \{F \subseteq X : A \subseteq F,\ F \text{ is closed}\}$. From the definition of closure, we have that $\operatorname{Cl} A \subseteq X \setminus U$ and therefore $x \notin \operatorname{Cl} A$. (b) implies (c): From the definition of a basis at the point x, we know that every $U \in \mathcal{B}_x$ is a neighbourhood of x, hence if (b) is true for x, proposition (c) follows immediately. (c) implies (a): Let $x \in X$ such that $x \notin \operatorname{Cl} A$, so that proposition (a) is false for x. From the definition of closure, there exists $F \subseteq C_A$ such that $x \notin F$. Consider the open complement $V = X \setminus F$ so that $x \in V$ and $V \cap A = \varnothing$. Hence, given any basis at x there exists a neighbourhood of x, say U, such that $U \subseteq V$ and hence $U \cap A = \varnothing$, which implies that proposition (c) is false for x. ♮

Corollary 12.1.18. If U is an open set and $U \cap A = \varnothing$, then $U \cap \operatorname{Cl} A = \varnothing$. Also, if U and V are disjoint sets, then $U \cap \operatorname{Cl} V = \operatorname{Cl} U \cap V = \varnothing$.

Proof. Suppose $U \cap A = \varnothing$ and that there exists $x \in U \cap \operatorname{Cl} A$, so that $x \in \operatorname{Cl} A$. Since U is open, it is a neighbourhood of x, hence from Proposition 12.1.17 we find that $U \cap A \neq \varnothing$, which is false, thus $U \cap \operatorname{Cl} A = \varnothing$. ♮

In order to classify points as being interior, exterior or on the boundary of a set, we may use the following important proposition.

Proposition 12.1.19 (Interior, exterior and boundary points)**.** Classification of points:

(a) $x \in \operatorname{Int} A$ if and only if there exists $U \subseteq A$ neighbourhood of x.

(b) $x \in \operatorname{Ext} A$ if and only if there exists $U \subseteq X \setminus A$ neighbourhood of x.

(c) $x \in \partial A$ if and only if all neighbourhoods $U \subseteq X$ of x are such that $U \cap A \neq \varnothing$ and $U \cap (X \setminus A) \neq \varnothing$.

Proof. (a) Let $x \in A$ be any point. Suppose $U \subseteq X$ is neighbourhood of x such that $U \subseteq A$, then from the fact that U is open we conclude that $U \subseteq \operatorname{Int} A$ from the definition of the interior operator. Suppose $x \in \operatorname{Int} A$ then from definition there is a neighbourhood of x contained in A.

(b) Suppose exists $U \subseteq X \setminus A$ neighbourhood of x. From Proposition 12.1.17 we find $x \notin \operatorname{Cl} A$ and therefore $x \in \operatorname{Ext} A$. The other side of the implication is trivial.

(c) Since all neighbourhoods of x contain a point of $X \setminus A$, from the first item of this proposition we find that x cannot belong to the interior of A. Moreover, every neighbourhood contains a point of A, then x cannot be an element of $\operatorname{Ext} A$. Therefore $x \in \partial A$.

$\natural$

Proposition 12.1.20 (Decomposition of the closure)**.** $\operatorname{Cl} A = A \cup \partial A = \operatorname{Int} A \cup \partial A$.

Proof. For the first equality we have

$$\begin{aligned} A \cup \partial A &= A \cup [X \setminus (\operatorname{Int} A \cup \operatorname{Ext} A)] \\ &= A \cup [(X \setminus \operatorname{Int} A) \cap (X \setminus \operatorname{Ext} A)] \\ &= A \cup [(X \setminus \operatorname{Int} A) \cap \operatorname{Cl} A] \\ &= [A \cup (X \setminus \operatorname{Int} A)] \cap [A \cup \operatorname{Cl} A] \\ &= X \cap \operatorname{Cl} A = \operatorname{Cl} A. \end{aligned}$$

Analogously, for the second equality

$$\begin{aligned} \operatorname{Int} A \cup \partial A &= \operatorname{Int} A \cup [(X \setminus \operatorname{Int} A) \cap \operatorname{Cl} A] \\ &= [\operatorname{Int} A \cup (X \setminus \operatorname{Int} A)] \cap [\operatorname{Int} A \cup \operatorname{Cl} A] \\ &= X \cap \operatorname{Cl} A = \operatorname{Cl} A. \end{aligned}$$

This proves the proposition. $\natural$

Proposition 12.1.21. Let X be a topological space and $A \subseteq X$. Then $\operatorname{Cl}(X \setminus A) = X \setminus \operatorname{Int} A$ and also $\operatorname{Int}(X \setminus A) = X \setminus \operatorname{Cl} A$.

Proof. We prove the first equality. Let $\mathcal{A} := \{U \subseteq A : U \text{ is open}\}$, from definition we have that $\operatorname{Int} A = \bigcup_{U \in \mathcal{A}} U$, moreover, $X \setminus \operatorname{Int} A = X \setminus \bigcup_{U \in \mathcal{A}} U = \bigcap_{U \in \mathcal{A}} X \setminus U$ from de Morgan's Laws. Notice that obviously $X \setminus U \subseteq X$ and moreover since U is open, then the complement $X \setminus U$ is closed. This makes $X \setminus \operatorname{Int}(A) = \bigcap_{U \in \mathcal{A}} X \setminus U = \operatorname{Cl}(X \setminus A)$. Now we show the second equality. Define $\widetilde{\mathcal{A}} := \{U \subseteq A : U \text{ is closed}\}$, then $\operatorname{Cl} A = \bigcap_{U \in \widetilde{\mathcal{A}}} U$ and hence $X \setminus \operatorname{Cl} A = X \setminus \bigcap_{U \in \widetilde{\mathcal{A}}} U = \bigcup_{U \in \widetilde{\mathcal{A}}} X \setminus U$. Notice that $X \setminus U \subseteq X \setminus A$ and since U is closed, then $X \setminus U$ is open. From this we conclude that $X \setminus \operatorname{Cl} A = \bigcup_{U \in \widetilde{\mathcal{A}}} X \setminus U = \operatorname{Int}(X \setminus A)$. $\natural$

Proposition 12.1.22. Let a finite collection of subsets $\{A_i\}_{i=1}^n \subseteq 2^X$, where X is a topological space, then we have that

$$\mathrm{Cl}\left(\bigcup_{i=0}^{n} A_i\right) = \bigcup_{i=0}^{n} \mathrm{Cl}\, A_i$$

Proof. It suffices to prove that for $A, B \subseteq X$ we have $\mathrm{Cl}(A \cup B) = \mathrm{Cl}\, A \cup \mathrm{Cl}\, B$. Notice that $\mathrm{Cl}\, A, \mathrm{Cl}\, B \subseteq \mathrm{Cl}(A \cup B)$, hence $\mathrm{Cl}\, A \cup \mathrm{Cl}\, B \subseteq \mathrm{Cl}(A \cup B)$. Now, since $A \subseteq \mathrm{Cl}\, A$ and $B \subseteq \mathrm{Cl}\, B$, we find $A \cup B \subseteq \mathrm{Cl}\, A \cup \mathrm{Cl}\, B$, then $\mathrm{Cl}(A \cup B) \subseteq \mathrm{Cl}(\mathrm{Cl}\, A \cup \mathrm{Cl}\, B) = \mathrm{Cl}\, A \cup \mathrm{Cl}\, B$. ♮

Definition 12.1.23 (Locally finite family)**.** A collection of subsets $\{A_i\}_{i \in I} \subseteq 2^X$ of a topological space X is said to be a locally finite family if for every point x there exists a neighbourhood $U \subseteq X$ for which the collection $\{i \in I : U \cap A_i \neq \varnothing\}$ is finite.

Definition 12.1.24 (Discrete family)**.** A collection of subsets $\{A_i\}_{i \in I} \subseteq 2^X$ of a topological space X is said to be a discrete family if for all $x \in X$ there exists a neighbourhood $U \subseteq X$ such that there exists at most one A_i in the family such that $U \cap A \neq \varnothing$. A discrete family is also a locally finite family.

Proposition 12.1.25. Given a locally finite family of sets $\{A_i\}_{i \in I} \subseteq 2^X$, where X is a topological space, we have that

$$\mathrm{Cl}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} \mathrm{Cl}\, A_i.$$

Proof. Notice that clearly $A_i \subseteq \mathrm{Cl}\left(\bigcup_{i \in I} A_i\right)$ for all $i \in I$, hence $\bigcup_{i \in I} \mathrm{Cl}\, A_i \subseteq \mathrm{Cl}\left(\bigcup_{i \in I} A_i\right)$. Moreover, we have from hypothesis that $\{A_i\}_{i \in I}$ is a locally finite family, thus given $x \in \mathrm{Cl}\left(\bigcup_{i \in I} A_i\right)$ we can find a neighbourhood of x, say $U \subset X$, such that $I_0 := \{i \in I : U \cap A_i \neq \varnothing\}$ is finite. Notice that from Proposition 12.1.17 we have that x cannot be a limit point of any of the sets with index $i \in I \setminus I_0$, hence $x \notin \mathrm{Cl}\left(\bigcup_{i \in I \setminus I_0} A_i\right)$. On the other hand, we have $x \in \mathrm{Cl}\left(\bigcup_{i \in I} A_i\right) = \mathrm{Cl}\left(\bigcup_{i \in I_0} A_i\right) \cup \mathrm{Cl}\left(\bigcup_{i \in I \setminus I_0} A_i\right)$, which implies that $x \in \mathrm{Cl}\left(\bigcup_{i \in I_0} A_i\right) = \bigcup_{i \in I_0} \mathrm{Cl}\, A_i \subseteq \bigcup_{i \in I} \mathrm{Cl}\, A_i$ (the equality comes from the fact that I_0 is finite and hence Proposition 12.1.22 hold). ♮

Proposition 12.1.26. If $\{A_i\}_{i \in I}$ is a locally finite (resp. discrete) family, then the family $\{\mathrm{Cl}\, A_i\}_{i \in I}$ is locally finite (resp. discrete). Conversely, if $\{\mathrm{Cl}\, A_i\}_{i \in I}$ is locally finite, then $\{A_i\}_{i \in I}$ is locally finite.

Proof. Let the locally finite (resp. discrete) family $\{A_i\}_{i \in I}$. Given any element $x \in X$ and a neighbourhood U of x such that the indexing set $I_0 := \{i \in I : U \cap A_i \neq \varnothing\}$ is finite (resp. is either empty or a singleton). Notice that since $U \cap A_i \subseteq U \cap \mathrm{Cl}\, A_i$, we find that for all $i \in I_0, U \cap \mathrm{Cl}\, A_i \neq \varnothing$, so that $I_0 = I_0' := \{i \in I : U \cap \mathrm{Cl}\, A_i \neq \varnothing\}$. Since U is an open set, from Corollary 12.1.18 we find that $U \cap \mathrm{Cl}\, A_i = U \cap A_i = \varnothing$ for all $i \in I \setminus I_0$. For the converse, if the collection of closures is locally finite, then clearly $\{A_i\}_{i \in I}$ is locally finite. ♮

Derived and Dense Sets

Definition 12.1.27 (Limit point and derived set)**.** A point $p \in A$ is called a limit point of A if every neighbourhood of p has a point of $A \setminus \{p\}$. We define the set A' as the collection of all limit points of A, which we'll call derived set.

Proposition 12.1.28. A point $p \in A$ is a limit point if and only if $p \in \mathrm{Cl}(A \setminus \{p\})$.

Proof. If p is a limit point, then from Proposition 12.1.17 we see that $p \in \mathrm{Cl}(A \setminus \{p\})$. Moreover, if $p \in \mathrm{Cl}(A \setminus \{p\})$, then again from Proposition 12.1.17 we see that there is a basis $\mathcal{B}_p$ at p such that for all $U \in \mathcal{B}_p$ we have $U \cap (A \setminus \{p\}) \neq \varnothing$ — thus p is a limit point of A. ♮

Definition 12.1.29 (Isolated points)**.** If $p \in A \setminus A'$, we say that p is an isolated point of A.

Proposition 12.1.30. A point $p \in A$ is isolated if and only if there exists a neighbourhood of p for which the only point of intersection with A is p, i.e. $U \subseteq X$ neighbourhood of p with $U \cap A = \{p\}$.

Proof. If p is isolated then there exists a neighbourhood $U \subseteq X$ of p such that $U \cap (A \setminus \{p\}) = \varnothing$ but since $p \in U$, then it follows that $U \cap A = \{p\}$. Now, suppose there exists such U, then in particular p does not satisfy the condition to be a limit point of A, thus $p \notin A$, hence $p \in A \setminus A'$ and therefore p is a limit point. ♮

Proposition 12.1.31. The closure of a set is the union of the set with its limit points, i.e. $\mathrm{Cl}\, A = A \cup A'$.

Proof. Let $p \in A'$, then from definition any neighbourhood of p intersects $A \setminus \{p\}$ — and, in particular, intersects A — therefore, by means of Proposition 12.1.17 we see that $p \in \mathrm{Cl}\, A$, thus $A' \subseteq \mathrm{Cl}\, A$. It's obvious that $A \subseteq \mathrm{Cl}\, A$, thus $A \cup A' \subseteq \mathrm{Cl}\, A$. On the other hand, if $q \in \mathrm{Cl}\, A$, then we assume that $q \notin A$ — since, on the contrary, it's clear that $q \in A \cup A'$. Again using Proposition 12.1.17 we get that every neighbourhood of q intersects A, and since $q \notin A$, then the intersection contains a point other than q — that is, q is a limit point of A. Hence $\mathrm{Cl}\, A \subseteq A \cup A'$, which finishes the proof. ♮

Corollary 12.1.32. A set A is closed if and only if A contains all of its limit points, that is, $A = A'$.

Proof. If A is closed, then $A = \mathrm{Cl}\, A$ and hence $A' \subseteq A$ from Proposition 12.1.31. Otherwise, if A contains all of its limit points, then $A \cup A' = A = \mathrm{Cl}\, A$ — that is, A is closed. ♮

Definition 12.1.33 (Dense set)**.** A set $A \subseteq X$ is said to be dense if $\mathrm{Cl}\, A = X$.

Proposition 12.1.34. A set A is dense in X if and only if for all non-empty open subsets of X contains a point of A.

Proof. Let A be dense in X and let $U \subseteq X$ be any non-empty open subset of the space. Suppose, for the sake of contradiction, that $U \cap A = \varnothing$. Since U is non-empty, take any $x \in U$, then from Proposition 12.1.17 we find that $x \notin \operatorname{Cl} A$ but since $\operatorname{Cl} A = X$, then $x \notin X$, which is a contradiction — thus $U \cap A$ is non-empty. On the other hand, if we suppose that every non-empty open set of the space intersects A, then given any $x \in X$ we see that x is a limit point of A, hence $x \in \operatorname{Cl} A$, thus $X \subseteq \operatorname{Cl} A$, which proves the proposition since $\operatorname{Cl} A \subseteq X$. ♮

12.2 Morphisms of Topological Spaces

Definition 12.2.1 (Continuous map)**.** Let X and Y be topological spaces and consider the map $f\colon X \to Y$. We say that f is continuous if for all $U \subseteq Y$ open, the preimage $f^{-1}(U)$ is open in X.

Proposition 12.2.2. A map $f\colon X \to Y$ is continuous if and only if the preimage of every closed subset is closed.

Proof. Suppose that f satisfies the latter, the given any $U \subseteq Y$ there exists a closed set $C \subseteq Y$ such that $U = Y \setminus C$, thus $f^{-1}(C)$ is closed, then $f^{-1}(X \setminus C) = f^{-1}(U)$ is open. For the contrary, the analogous argument is used. ♮

Proposition 12.2.3. Let X, Y, Z be topological spaces. The following properties of continuous maps between topological spaces hold

(CM1) Every constant map $f\colon X \to Y$ is continuous.

(CM2) The identity map is continuous.

(CM3) If $f\colon X \to Y$ is continuous, then for all open set $U \subseteq X$ we have $f|_U$ continuous.

(CM4) If $f\colon X \to Y$ and $g\colon Y \to Z$ are continuous maps, then $gf\colon X \to Z$.

Proof. (1) Let any point $a \in Y$ and consider the constant map $x \overset{f}{\mapsto} a$ for all $x \in X$. Then, for all $U \subseteq Y \setminus \{a\}$ open, we have that $f^{-1}(U) = \varnothing$, thus have open preimage. On the other hand, the fibre $f^{-1}(a) = X$, hence also open, thus f is continuous.

(2) Notice that if $U \subseteq Y$ is any open set, then $\mathrm{id}_X^{-1}(U) = U$ and thus open.

(3) Let $g\colon X \to Y$ be a continuous map, and $U \subseteq X$ be any open set, then we can take any open $V \subseteq g(U)$ and conclude that $g^{-1}(V)$ is open (from the hypothesis that g is continuous); let now an open $H \subseteq Y \setminus g(U)$, then certainly $g|_U^{-1}(H) = \varnothing$, thus open. Hence $g|_U\colon U \to Y$ is indeed continuous.

(4) Let $U \subseteq Z$ be open, then $(gf)^{-1}(U) = f^{-1}(g^{-1}(U))$. Moreover, since from hypothesis g is continuous, then $g^{-1}(U)$ is open, now, from the continuity of f we conclude that $f^{-1}(g^{-1}(U))$ is open, hence gf is continuous.

♮

Definition 12.2.4 (Category of topological spaces)**.** We define the category of topological spaces to be composed of objects named topological spaces and morphisms being continuous maps between them — we'll denote such category by `Top`.

Proposition 12.2.5 (Local criterion for continuity)**.** Let $f\colon X \to Y$ be a map between topological spaces. Then f is continuous if and only if for all points $x \in X$ there exists a neighbourhood $U_x \subseteq X$ of x such that $f|_{U_x}$ is continuous.

Proof. If f is a morphism, then for any $x \in X$ we can take the whole space X as a neighbourhood of x and the proposition follows. On the other hand If f is locally continuous for every point of X, let $V \subseteq Y$ be any open set and consider, for every $x \in f^{-1}(V)$ a neighbourhood U_x such that $f|_{U_x}$ is continuous — and hence $f|_{U_x}^{-1}(V)$ is open for all $x \in f^{-1}(V)$. Notice that $f^{-1}(V) \cap U_x = f|_{U_x}^{-1}(V)$, which has to be open on X. Moreover, from construction $f^{-1}(V) = \bigcup_{x \in f^{-1}(V)} f^{-1}(V) \cap U_x$ is the union of open sets, thus $f^{-1}(V)$ is open — and therefore f is continuous. ♮

Proposition 12.2.6 (Basis criterion for continuity)**.** Let $f\colon X \to Y$ be a map between topological spaces, and $\mathcal{B}$ be a basis for the topology of Y. Then, f is continuous (hence a morphism) if and only if for all $B \in \mathcal{B}$ we have $f^{-1}(B) \subseteq X$ open.

Proof. ($\Rightarrow$) If f is continuous, then certainly $f^{-1}(B)$ is open. ($\Leftarrow$) On the other hand, if $A \subseteq X$ is any open set, then $A = \bigcup_{p \in A} B_p$ for $B_p \in \mathcal{B}$ neighbourhood of p. Hence $f^{-1}(A) = f^{-1}(\bigcup_{p \in A} B_p) = \bigcup_{p \in A} f^{-1}(B_p)$ is the union of open sets, thus $f^{-1}(A)$ is open and therefore f is continuous. ♮

Proposition 12.2.7. Let $f\colon X \to Y$ be a morphism of topological spaces and $\mathcal{B}$ be a basis for the space X. Then $f(\mathcal{B}) = \{f(B) : B \in \mathcal{B}\}$ is a basis for the space Y if and only if f is surjective and open.

Proof. ($\Rightarrow$) Suppose $f(\mathcal{B})$ is a basis for Y, then for all $V \subseteq Y$ open, there exists an indexing set I such that $V = \bigcup_{i \in I} U_i$, where $U_i \in f(\mathcal{B})$. This implies in the existence of $B \in \mathcal{B}$ such that $U_i = f(B)$ hence f is open. Consider now any point $y \in Y$ and any neighbourhood V_y of y. From the same argument as above we have $V_y = \bigcup_{i \in I_y} U_i$, where there exists some $i \in I_y$ such that $y \in U_i$ and hence $y \in U_i = f(B)$ for some $B \in \mathcal{B}$. ($\Leftarrow$) Suppose f is surjective and open. Let $V \subseteq Y$ be any open set. Since f is continuous and surjective, we have that $U = f^{-1}(V)$ is a non-empty open set. Since $\mathcal{B}$ is a basis, we can write $U = \bigcup_{i \in I} B_i$ where $B_i \in B$, and then $f(U) = f\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f(B_i)$, where $f(B_i) \in f(\mathcal{B})$ and $f(U) = V$ from the fact that f is surjective. Since f is open, $f(B)$ is open for all $B \in \mathcal{B}$. Hence $f(\mathcal{B})$ is a basis for the space Y. ♮

Definition 12.2.8 (Isomorphism)**.** Let f be a morphism of topological spaces. If f is bijective and has a continuous inverse, then we say that f is an isomorphism of topological spaces — which can also be called an homeomorphism.

Definition 12.2.9 (Open & Closed maps)**.** Let $f\colon X \to Y$ be any set-function between topological spaces.

- We say that f is an open map if for all $U \subseteq X$ open the image $f(U) \subseteq Y$ is open.
- We say that f is a closed map if for all $C \subseteq X$ closed, the image $f(C) \subseteq Y$ is closed.

Proposition 12.2.10. Let $f\colon X \to Y$ be a map of topological spaces and consider that f is an isomorphism, then f is an open and a closed map.

Proof. Let $U \subseteq X$ be an open (resp. closed) set and consider $V := f(U) \subseteq Y$. Since f is a bijection, we find that $f(U) = V$ is open (resp. closed) by the continuity of f^{-1}. Hence f is an open (resp. closed) map. ♮

Corollary 12.2.11. Let $f\colon X \to Y$ be a bijective set-function, where X and Y are topological spaces. The following propositions are equivalent

(a) f is an isomorphism of topological spaces.

(b) f is open.

(c) f is closed.

Proof. (a) $\Rightarrow$ (b): Suppose f is an isomorphism, then since f^{-1} is also an isomorphism, it follows that, given any $U \subseteq X$ the image $f(U)$ is open — hence f is open. (b) $\Rightarrow$ (c): Let f be open, then given any closed set $C \subseteq X$, we have that $f(X \setminus C) \subseteq Y$ is open, moreover, since f is injective, $f(X \setminus C) = f(X) \setminus f(C)$ and since $f(X) = Y$ from the surjectivity of f, we conclude that $f(C)$ is closed. (c) $\Rightarrow$ (a) Let f be closed, then given any $V \subseteq Y$ open, we know that $Y \setminus V$ is closed and since f is a bijection, $f^{-1}(Y \setminus V) = f^{-1}(Y) \setminus f^{-1}(V) = X \setminus f^{-1}(V) \subseteq X$ is closed, thus $f^{-1}(V)$ is open — which implies in the continuity of f. Moreover, since f is a bijection, f^{-1} is also closed, and therefore continuous by the same analogous proof, thus f is an isomorphism. ♮

Proposition 12.2.12. Let X and Y be topological spaces and $f\colon X \to Y$ be a set-function. We can classify the behaviour of f by the following conditions

(a) f is a morphism of topological spaces if and only if $f(\operatorname{Cl} A) \subseteq \operatorname{Cl}(f(A))$ for every set $A \subseteq X$.

(b) f is a morphism of topological spaces if and only if $f^{-1}(\operatorname{Int} B) \subseteq \operatorname{Int} f^{-1}(B)$ for every set $B \subseteq Y$.

(c) f is closed if and only if $f(\operatorname{Cl} A) \supseteq \operatorname{Cl}(f(A))$ for every set $A \subseteq X$.

(d) f is open if and only if $f^{-1}(\operatorname{Int} B) \supseteq \operatorname{Int} f^{-1}(B)$ for every set $B \subseteq Y$.

Proof.

prove

♮

Definition 12.2.13 (Local isomorphism)**.** Let $f\colon X \to Y$ be a set-function between topological spaces X and Y. We say that f is a local isomorphism of topological spaces if, for all $x \in X$, there exists a neighbourhood $U \subseteq X$ such that $f(U) \subseteq Y$ is open and the induced map $f\colon U \xrightarrow{\simeq} f(U)$ is an isomorphism of topological spaces.

Proposition 12.2.14. The following are properties pertaining to local isomorphisms

(a) Every isomorphism is a local isomorphism.

(b) Every local isomorphism is continuous and open.

(c) Every bijective local isomorphism is an isomorphism.

Proof. (a) Let $f\colon X \xrightarrow{\simeq} Y$ be an isomorphism, then for all $x \in X$ we can choose the neighbourhood X and the restriction $f|_X = f\colon X \xrightarrow{\simeq} f(X) = Y$ is an isomorphism.

(b) Let $g\colon X \to Y$ be a local isomorphism. Let $x \in X$ be any element and consider $U \subseteq X$ such that $g|_U\colon U \xrightarrow{\simeq} g(U)$ is an isomorphism, then in particular $g|_U$ is continuous — by Proposition 12.2.5 we find that g is continuous. Now let $U \subseteq X$ be any open set, for each $x \in U$ take $U_x \subseteq X$ neighbourhood of x such that $g|_{U_x}\colon U_x \to g(U_x)$ is an isomorphism. Notice that the restriction $g|_{U \cap U_x}\colon U \cap U_x \xrightarrow{\simeq} f(U \cap U_x)$ is also an isomorphism which implies in $f(U \cap U_x) \subseteq Y$ being open — moreover, $V = \bigcup_{x \in U} f(U \cap U_x)$ thus V is open.

(c) Let $f\colon X \to V$ be a bijective local isomorphism, then by the last item, f is open. Using Corollary 12.2.11 we see that f is an isomorphism.

♮

Definition 12.2.15 (Embedding)**.** Let $f\colon X \rightarrowtail Y$ be an injective morphism of topological spaces. If the induced morphism $f\colon X \to f(X)$ is an isomorphism, then we say that f is a topological embedding of X in Y.

Topology Generated by Mappings

Proposition 12.2.16 (Topology generated by a collection of mappings)**.** Let X be a topological space and $\{Y_i\}_{i \in I}$ be a collection of topological spaces. Let $\{f_i\colon X \to Y_i\}_{i \in I}$ be the collection of mappings between such topological spaces. Then there exists a initial topology on X such that f_i is continuous, for all $i \in I$. Such topology is generated by the base

$$\mathcal{B} = \left\{ \bigcap_{j \in J} f_j^{-1}(U_j)\colon U_i \subseteq Y_i \text{ is open}, J \subseteq I \text{ is finite} \right\}.$$

We call such topology as the topology generated by the collection of mappings $\{f_i\}_{i \in I}$.

Proof. First we show that $\mathcal{B}$ is indeed a basis for X. Let $x \in X$ be any point, then clearly $x \in f_i^{-1}(Y_i)$ for all $i \in I$, hence $x \in \bigcap_{j \in J} f_j^{-1}(Y_j) \in \mathcal{B}$ for a finite indexing set $J \subseteq I$. Let $J, S \subseteq I$ be finite indexing sets, then consider the sets $A := \bigcap_{j \in J} f_j^{-1}(U_j)$, $B := \bigcap_{s \in S} f_s^{-1}(V_s) \in \mathcal{B}$ and let any point $x \in A \cap B$. Then in particular we can let an non-empty indexing set $T = J \cap S$ so that $x \in f_t^{-1}(U_t)$ for all $t \in T$ and therefore $C := \bigcap_{t \in T} f_t^{-1}(U_t)$ is such that $x \in C \subseteq A \cap B$. This shows that $\mathcal{B}$ indeed satisfies the basis properties.

Now we show the initial topology property. Let τ be the topology generated by $\mathcal{B}$ and consider τ' to be any other topology on X for which the functions f_i are continuous for all $i \in I$. Trivially we must have $\mathcal{B} \subseteq \tau'$ and therefore $\tau \subseteq \tau'$. This says that τ is coarser than τ', which proves the proposition. ♮

Proposition 12.2.17. Let X and Y be topological spaces, and the topology of Y be generated by the collection of maps $\{f_i : Y \to Y_i\}_{i \in I}$, where $\{Y_i\}_{i \in I}$ is a collection of topological spaces. Then a map $f : X \to Y$ is continuous if and only if the composition $f_i f : X \to Y_i$ is continuous for every $i \in I$.

Proof. ($\Leftarrow$) Suppose $f_i f$ is continuous for all $i \in I$, then since f_i is continuous on the topology of Y it follows that for any given open set $U \subseteq Y_i$ we have $f_i^{-1}(U) = V \subseteq Y$ open. In particular, notice that $(f_i f)^{-1}(U) = f^{-1}(f_i^{-1}(U)) = f^{-1}(V) \subseteq X$ which must be open from the hypothesis of the continuity of $f_i f$. ($\Rightarrow$) Moreover, if f is continuous, then clearly the composition of continuous functions is continuous. ♮

12.3 Metric Spaces

Write on metric spaces

Definition 12.3.1. Let M be a set. We say that a map $d : M \times M \to \mathbf{R}_{\geqslant 0}$ is a *pre-metric* if for all points $x, y, z \in M$ it satisfies

(a) Symmetry: $d(x, y) = d(y, x)$.

(b) Triangle inequality: $d(x, z) \leqslant d(x, y) + d(y, z)$.

Moreover, if d happens to satisfy the condition that

(c) $d(x, y) = 0$ implies in $x = y$.

then we say that d establishes a *metric* in M.

The set M together with the (pre)metric d is called a (pre)metric space.

12.4 Hausdorff Spaces

Definition 12.4.1. Let X be a topological space. We define

(1) X is T_0 (or Kolmogorov) if and only if for every pair of distinct points $x, y \in X$ there exists an open set containing one, but not both, of them.

(2) X is T_1 (or Fréchet) if and only if for every pair of distinct points $x, y \in X$ there exists open sets $U, V \subseteq X$ such that $x \in U$ and $y \in V$, but $x \notin V$ and $y \notin U$

Proposition 12.4.2. A space X is T_1 if and only if the constant sequence $(x)_{i \in \mathbf{N}}$ converges to x and only to x.

Proof. ($\Rightarrow$) Suppose X is T_1, then certainly $x \to x$. Let $y \in X$ be distinct of x, then by the T_1 property, there exists a neighbourhood $U \ni y$ such that $x \notin U$, hence $(x)_{i \in \mathbf{N}}$ does not converge to y. ($\Leftarrow$) Suppose X is not T_1, then choose $y \in X$ for which every neighbourhood $U \ni y$ contains x, then $x \to y$. ♮

Definition 12.4.3 (Hausdorff space)**.** We say that a topological space X is Hausdorff (or T_2) if for all pair of distinct points $x, y \in X$, there exists neighbourhoods U_x and U_y, subsets of X, such that $U_x \cap U_y = \varnothing$.

Corollary 12.4.4. Every open subset of a Hausdorff space is Hausdorff.

Proof. Let X be Hausdorff and $U \subseteq X$ be an open. Choose any distinct pair $x, y \in U$ and let $U_x, U_y \subseteq X$ be neighbourhoods of x, y such that $U_x \cap U_y = \varnothing$. Since $U_x \cap U, U_y \cap U \subseteq U$ are neighbourhoods of x and y, respectively, then $(U_x \cap U) \cap (U_y \cap U) = \varnothing$, hence U is Hausdorff. ♮

Proposition 12.4.5 (Hausdorff properties)**.** Let X be a Hausdorff space. Then

(a) Every finite subset of X is closed.

(b) If a sequence $(x_i) \subseteq X$ converges to $x \in X$, then the limit is unique.

Proof. (a) Let $p_0 \in X$ be any point and $p \in X$ be a distinct point. Since X is Hausdorff, choose $U_p, U_{p_0} \subseteq X$ such that $U_p \cap U_{p_0} = \varnothing$, hence $U_p \subseteq X \smallsetminus \{p_0\}$. Since for all $x \in X \smallsetminus \{p_0\}$ there exists a neighbourhood $U_x \subseteq X \smallsetminus \{p\}$, hence $\{p\}$ must be a closed set (every point of $X \smallsetminus A$ has a neighbourhood in $X \smallsetminus A$ is equivalent of saying that A is closed).

(b) Suppose that $x_i \to x$ and $x_i \to y$ are two limits of the sequence. Since X is Hausdorff, if $x \neq y$ implies that there exists neighbourhoods $U_x, U_y \subseteq X$ for which $U_x \cap U_y = \varnothing$. From the definition of convergence, there exists $N \in \mathbf{N}$ such that for all $i \geqslant N, x_i \in U_x$ and exists $M \in \mathbf{N}$ for which every $i \geqslant M$, we have $x_i \in U_y$. Then, for $i \geqslant \max(N, M), x_i \in U_x \cap U_y$, which cannot happen if $x \neq y$, hence $x = y$. ♮

Proposition 12.4.6 (Neighbourhoods of limit points)**.** Let X be a Hausdorff space and $A \subseteq X$. If $p \in X$ is a limit point of A, then every neighbourhood of p contains infinitely many points of A.

Proof. Let p be a limit point of A, then for any neighbourhood $U_p \subseteq X$ we must have $(U_p \smallsetminus \{p\}) \cap A \neq \varnothing$. For the sake of contradiction, suppose that there exists a finite number of points, $n > 1$, in the set $U_p \cap A = \{x_i\}_{i=1}^n$, which implies that $U_p \cap A$ is closed. Consider now the non-empty closed set $\{x_i\}_{i=1}^n \smallsetminus \{p\}$, then $X \smallsetminus (\{x_i\}_{i=1}^n \smallsetminus \{p\})$ is open and contains p, thus is a neighbourhood of p. Notice that since the intersection of finitely many open sets is open, then $U_p \cap \left(X \smallsetminus \left(\{x_i\}_{i=1}^n \smallsetminus \{p\}\right)\right)$ is also open and is a neighbourhood of p. However, notice that

$$A \cap U_p \cap (X \smallsetminus (\{x_i\}_{i=1}^n \smallsetminus \{p\})) = \{x_i\}_{i=1}^n \cap ((X \smallsetminus \{x_i\}_{i=1}^n) \cup \{p\}) = \{p\}$$

which is a contradiction to the fact that p is a limit point of A (because any neighbourhood of p should also contain a point of A other than p). Hence we conclude that $A \cap U_p$ must be infinite. ♮

Proposition 12.4.7. Let X be a Hausdorff space and $A \subseteq X$. Then the set of limit points of A, denoted by A', is closed in X.

Proof. Consider the complement set $X \setminus A'$, we must show that it is open. Consider $x \in X \setminus A'$, so that x is not a limit point of A and hence there exists $U \subseteq X$ neighbourhood of x such that $(U \setminus \{x\}) \cap A = \varnothing$. We now show that $U \subseteq X \setminus A'$. Let $p \in U$ be any point, then the set $U \setminus \{x\}$ is a neighbourhood of p and is disjoint with A, hence p is not a limit point of A, that is, $p \notin A'$, which proves that $U \subseteq X \setminus A'$. Moreover, since X is Hausdorff, the singleton $\{x\}$ is closed, hence $U \setminus \{x\}$ is open ♮

Proposition 12.4.8. Let $f, g\colon X \to Y$ be morphisms of topological spaces, and Y be a Hausdorff space. Then the set $\{x \in X : f(x) = g(x)\}$ is closed in X.

Proof. Consider the complement set $A := X \setminus \{x \in X : f(x) = g(x)\}$. Let any point $x \in A$, since $f(x) \neq g(x)$ and Y is Hausdorff, there exists neighbourhoods $U_1, U_2 \subseteq Y$ of $f(x), g(x)$, respectively, such that $U_1 \cap U_2 = \varnothing$. Since U_1, U_2 are open, then $f^{-1}(U_1), g^{-1}(U_2)$ are both open, hence $f^{-1}(U_1) \cap g^{-1}(U_2)$ is a neighbourhood of x such that $f^{-1}(U_1) \cap g^{-1}(U_2) \subseteq A$. From Proposition 12.1.19 we find that A is open. ♮

Proposition 12.4.9. Every metric space is Hausdorff.

Proof. Let (M, d) be a metric space and $x, y \in M$ any distinct points. Let $r := d(x, y)$. The open balls $B_{r/2}(x)$ and $B_{r/2}(y)$ are disjoint, thus M is Hausdorff. ♮

Proposition 12.4.10. Every totally ordered set endowed with the order topology is a Hausdorff space.

Proof. Let X be a space endowed with the order topology. Let $x, y \in X$ be distinct points. Suppose that $x < y$. If there exists a point $z \in X$ such that $x < z < y$ then the neighbourhoods of x and y, respectively, $U = \{u \in X : u < z\}$ and $V = \{u \in X : u > z\}$ are disjoint, that is $U \cap V = \varnothing$. Suppose there is no such middle element, then $G = \{u \in X : u < y\} = \{u \in X : u \leqslant x\}$ and $H = \{u \in X : u > x\} = \{u \in X : u \geqslant y\}$ are neighbourhoods of x and y, respectively, and moreover $G \cap V = \varnothing$. This shows that X is Hausdorff. ♮

12.5 Countability

Definition 12.5.1 (First countable)**.** Let X be a topological space. We say that X is first countable if for all points $p \in X$ there exists a countable basis of X composed of neighbourhoods of p. Equivalently $\chi(p, X) \leqslant \aleph_0$.

Convergence on First Countable Spaces

Proposition 12.5.2 (Sufficient condition for Hausdorff)**.** Let X be a first countable space. Then, X is Hausdorff if and only if every sequence has at most one limit

Proof. Let $x, y \in X$ be distinct points, and $\mathcal{B}$ be a neighbourhood basis at x and $\mathcal{A}$ be a neighbourhood basis at y. If X is not Hausdorff, then for all $n \in \mathbf{N}$, choose a point $x_n \in B_n \cap A_n$, where $B_n \in \mathcal{B}$ and $A_n \in \mathcal{A}$ and consider the sequence $\{x_n\}_{n\in\mathbf{N}}$. Then there exists sub-sequences $\{x_k\}$ and $\{x_j\}$ for which $x_k \to x$ and $x_j \to y$. ♮

Definition 12.5.3 (Nested neighbourhood basis)**.** Let X be a topological space and a point $p \in X$. The infinite sequence of neighbourhoods of p, namely $(U_i)_{i\in\mathbf{N}}$ is said to be a nested neighbourhood basis at p if for all $i \in \mathbf{N}$, $U_{i+1} \subseteq U_i$ and for all neighbourhood V of p, there exists $i \in \mathbf{N}$ for which $U_i \subseteq V$.

Lemma 12.5.4. Let X be a first countable topological space. Then, for all points $p \in X$ there exists a nested neighbourhood basis at p.

Proof. Since X is first countable, let $\mathcal{V}$ be a countable basis for the topology of X at p. If $|\mathcal{V}| < \infty$ then define $U_i := V_1 \cap \cdots \cap V_{|\mathcal{V}|}$ for all $i \in \mathbf{N}$. If $|\mathcal{V}|$ is infinite, then define for all $i \in \mathbf{N}$ the set $U_i := V_1 \cap \cdots \cap V_i$. For both cases, the sequence $(U_i)_{i\in\mathbf{N}}$ is a nested neighbourhood basis. ♮

Definition 12.5.5 (Eventually in)**.** Let X be a topological space, and a sequence $(x_i)_{i\in\mathbf{N}} \subseteq X$, and a set $A \subseteq X$. We say that $(x_i)_{i\in\mathbf{N}}$ is eventually in A if $x_i \in A$ for all but finitely many $i \in \mathbf{N}$.

Lemma 12.5.6 (Sequence lemma)**.** Let X be a first countable space, and a set $A \subseteq X$, and a point $p \in X$. Then

(a) $p \in \operatorname{Cl} A$ if and only if p is a limit of points of A.

(b) A is closed in X if and only if A contains every limit point of sequences in A.

(c) $p \in \operatorname{Int}(A)$ if and only if all sequences that converge to p are eventually in A.

(d) A is open in X if and only if every sequence in X converging to a point of A is eventually in A.

Proof. (a) ($\Rightarrow$) Let $p \in \operatorname{Cl} A$, then for all neighbourhoods $V \subseteq A$ of p, the set $V \cap (A \setminus \{p\})$ is non-empty. Since X is first countable, consider $(U_i)_{i\in\mathbf{N}}$ a nested neighbourhood basis at p and construct the sequence $x\colon \mathbf{N} \to \bigcup_{i\in\mathbf{N}} U_i \cap (A \setminus \{p\})$ defined as $i \mapsto x_i \in U_i \cap (A \setminus \{p\})$. We'll show that $x_i \to p$. Consider $V \subseteq X$ any neighbourhood of p, since $(U_i)_{i\in\mathbf{N}}$ is a basis then there exists an indexing set $I_V \subseteq \mathbf{N}$ for which $V = \bigcup_{i\in I_V} U_i$. Consider the index $n := \min(I_V)$, then, from the definition of the nested basis, we have for all $i \geqslant n$ the elements $x_i \in V \cap A$, hence $x_i \to p$. ($\Leftarrow$) Suppose that $(x_i)_{i\in\mathbf{N}}$ is a sequence of points in A such that $x_i \to p$. From definition, for all neighbourhood $V \subseteq X$ of p, there exists $n \in \mathbf{N}$ such that $\forall i \geqslant N, x_i \in V$, moreover, since $x_i \in A$ then $x_i \in V \cap A$ which implies that $p \notin \operatorname{Ext}(A)$, hence $p \in \operatorname{Cl} A$.

(b) ($\Rightarrow$) Suppose that A is closed, then $A = \operatorname{Cl} A$. Consider any sequence $(x_i)_{i\in\mathbf{N}}$ of points in A and let $x_i \to p$. From item (a) we have that $p \in A$. ($\Leftarrow$) Suppose the contrary, then given any $p \in \operatorname{Cl} A$ we have $p \in A$, which implies that A is closed.

(c) ($\Rightarrow$) Suppose $p \in \mathrm{Int}(A)$. Let $(x_i)_{i\in\mathbf{N}} \subseteq X$ be a sequence such that $x_i \to p$. Consider any neighbourhood $V \subseteq A$ of p, then from the definition of convergence, exists $N \in \mathbf{N}$ such that $\forall i \geqslant N, x_i \in V_p$, then there exists at most $N-1$ points of $(x_i)_{i\in\mathbf{N}}$ outside A, hence the sequence is eventually in A. ($\Leftarrow$) Suppose $(x_i)_{i\in\mathbf{N}}$ is not-eventually in A and $x_i \to p$. For the sake of contradiction, suppose that $p \in \mathrm{Int}(A)$, then given a neighbourhood $V \subseteq A$ of p, there must exist $N \in \mathbf{N}$ such that $\forall i \geqslant N, x_i \in V \subseteq A$, which is a contradiction to the fact that $(x_i)_{i\in\mathbf{N}}$ is not-eventually in A, hence $p \notin \mathrm{Int}(A)$.

(d) ($\Rightarrow$) Let A be open, then $A = \mathrm{Int}(A)$. Given any $p \in A$ we have from item (c) that p is the limit of all of its converging sequences are eventually in A. ($\Leftarrow$) Suppose the contrary, and let $p \in A$ be any point. From item (c) we see that $p \in \mathrm{Int}(A)$, which implies that $A \subseteq \mathrm{Int}(A)$ and hence A is open. ♮

Second Countable Spaces and Covers

Definition 12.5.7 (Second countable)**.** A topological space X is second countable if it admits a countable basis for its topology. Equivalently $w(X) \leqslant \aleph_0$.

Definition 12.5.8 (Cover)**.** Let X be a topological space and $\mathcal{U} \subseteq 2^X$. We say that $\mathcal{U}$ is a cover of X if for all points $x \in X$ there exists $U \in \mathcal{U}$ such that $x \in U$.

Definition 12.5.9 (Open and closed covers)**.** Let $\mathcal{U}$ be a cover of a space X. If all elements $U \in \mathcal{U}$ are opens in X, then $\mathcal{U}$ is said to be open. On the other hand, if all $U \in \mathcal{U}$ are closed in X, then $\mathcal{U}$ is said to be closed.

Definition 12.5.10 (Subcover)**.** Let $\mathcal{U}$ be a cover of a space X. If $\mathcal{U}' \subseteq \mathcal{U}$ is a cover for X, then we call it a subcover of $\mathcal{U}$.

Although formally we should always write down if the cover is either open or closed, I'll slip here and there and every time that I use the term "cover" and do not specify its type, it should be understood that I'm talking about open covers — otherwise I'll specifically write "closed cover" in order to avoid any confusion.

Proposition 12.5.11 (Basis out of covers)**.** Let $\mathcal{U}$ be an open cover of the space X. For all $U \in \mathcal{U}$, define $\mathcal{B}_U$ as the basis for the subspace U. Then union $\bigcup_{U\in\mathcal{U}} \mathcal{B}_U$ is a basis for X.

Proof. Define $\mathcal{B} := \bigcup_{U\in\mathcal{U}} \mathcal{B}_U$. Clearly $\bigcup_{B\in\mathcal{B}} B = \bigcup_{U\in\mathcal{U}} \bigcup_{B\in\mathcal{B}_U} B = X$ — since every element of X can be found in $\mathcal{U}$ and $\mathcal{B}_U \subseteq 2^U \subseteq 2^X$. Let $x \in X$ be any point and let $U \in \mathcal{U}$ be a neighbourhood of x. Notice that on U the basis $\mathcal{B}_U$ satisfies the local intersecting condition (see Item 2). Since $x \in X$ is any point, then the condition is true globally for $\mathcal{B}$, thus $\mathcal{B}$ is a basis for X. ♮

Corollary 12.5.12 (Second countable out of a cover)**.** Let $\mathcal{U}$ be a countable open cover of the space X. If every $U \in \mathcal{U}$ is second countable, then X is second countable.

Proof. We consider $\mathcal{B}$ as in Proposition 12.5.11, the union of the basis $\mathcal{B}_U$ for the cover elements $U \in \mathcal{U}$ — and since U is second countable, we choose $\mathcal{B}_U$ to be a countable cover. Since $\mathcal{U}$ is countable and so is every basis contained in the basis $\mathcal{B}$, we conclude that $\mathcal{B}$ itself is countable and by Proposition 12.5.11 we find that $\mathcal{B}$ is a basis for X. ♮

Definition 12.5.13 (Lindelöf space)**.** A space X is said to be Lindelöf if every open cover of X has a countable subcover.

Definition 12.5.14 (Separable space)**.** A space X is said to be separable if it contains a countable dense subset.

Proposition 12.5.15 (Properties of second countable spaces)**.** The following properties hold

(SC1) A second countable space is first countable.

(SC2) A second countable space is separable.

(SC3) A second countable space is Lindelöf.

Proof. Let $\mathcal{B}$ be a countable basis for X. (SC1) Let a point $p \in X$, then the set $\mathcal{B}_p \subseteq \mathcal{B}$ of neighbourhoods of p is a countable basis for X at p.

(SC2) Let $f \colon I \to \bigcup_{B \in \mathcal{B}} B$ with $n \mapsto x_n$, where $|I| = |B|$ is an indexing set, and $x \in X$ be any point, and any neighbourhood of $V_x \subseteq X$ of x. Since B is a basis, there exists an indexing set I_{V_x} such that $V_x = \bigcup_{i \in I_{V_x}} B_i$, hence $V_x \cap A$ is non-empty. Moreover, define a sequence $(x_i)_{i \in I_{V_x}}$ such that $x_i \in B_i$. Notice that clearly $x_i \to x$, from the fact that V_x is a neighbourhood of x, and $x_i \in A$, from Lemma 12.5.6 we have that $x \in \operatorname{Cl}(\operatorname{im}(f))$. Hence we conclude that $X = \operatorname{Cl}(\operatorname{im}(f))$ and thus $\operatorname{im}(f)$ is a countable dense subset of X.

(SC3) Let $\mathcal{U}$ be a cover for X and define $\mathcal{B}' := \{B \in \mathcal{B} : B \subseteq U (U \in \mathcal{U})\}$, and $\mathcal{U}' := \{U \in \mathcal{U} : B \subseteq U (B \in \mathcal{B}')\}$. We'll show that $\mathcal{U}'$ is a subcover of $\mathcal{U}$. Let $x \in X$ be any point. since $\mathcal{U}$ covers X then there exists $U_x \in \mathcal{U}$ such that $x \in U_x$. Moreover, since $\mathcal{B}$ is a basis for the topology of X, then there exists $B_x \in \mathcal{B}$ such that $B_x \subseteq U_x$, so that $B_x \in \mathcal{B}'$ and $U_x \in \mathcal{U}'$ from the construction of the collections. Hence $\mathcal{U}'$ covers X. ♮

Proposition 12.5.16. Given a metric space M, the following properties are equivalent

(MS1) M is second countable.

(MS2) M is separable.

(MS3) M is Lindelöf.

Proof. (Separable $\Rightarrow$ Lindelöf) Suppose M is separable and let A be a countable dense set in M. Let $\mathcal{U}$ be a cover of M and define the collection $\mathcal{U}' := \{U \in \mathcal{U} : U \supseteq B(a, r),\ (a, r) \in A \times \mathbf{Q}\}$, where $B(a, r)$ is the open ball of radius r around a. We show that $\mathcal{U}'$ is a subcover of M. Let $x \in M$ be any point. Since $\mathcal{U}$ covers M then there exists $U_x \in \mathcal{U}$ such that $x \in U_x$. Since M is a metric space, there exists an open ball $B(x, \ell) \subseteq U_x$ for some $\ell \in \mathbf{R}$. Since A is dense in M, there exists a point $a \in A$ such that $a \in B(x, \ell/2)$. Now, from the fact that $\mathbf{Q}$ is dense in $\mathbf{R}$ we conclude that there exists an

$r \in \mathbf{Q}$ such that $d(x,a) < r < \ell/2$. Notice that $B(a,r) \subseteq B(x,r) \subseteq U_x$, hence $U_x \in \mathcal{U}'$. Therefore $\mathcal{U}'$ is a subcover of $\mathcal{U}$, and since $|\mathcal{U}'| = |A \times \mathbf{Q}|$, we find that $\mathcal{U}'$ is countable.

(Lindelöf $\Rightarrow$ second countable) Suppose M is Lindelöf, then given a cover $\mathcal{U} := \{B(x, \frac{1}{n}) \subseteq M : x \in M, n \in \mathbf{N}\}$ of M there exists a countable subcover $\mathcal{U}'$. Let $U \subseteq M$ be any open set and let $x \in U$ be any point. Since M is a metric space, there exists $r > 0$ such that $B(x,r) \subseteq U$. Define now $n \in N$ such that $\frac{1}{n} < \frac{r}{2}$. Since $\mathcal{U}'$ is a cover of M, there exists a $B(p, \frac{1}{n}) \in \mathcal{U}'$ such that $x \in B(p, \frac{1}{n})$. We now show that $B(p, \frac{1}{n}) \subseteq B(x,r)$. Let $y \in B(p, \frac{1}{n})$ be any point, then $d(p,y) < \frac{1}{n}$, moreover $d(x,p) < \frac{1}{n}$, hence we conclude that $d(x,y) \leqslant d(x,p) + d(p,y) < \frac{2}{n} < r$. This implies in particular that $y \in B(x,r)$ and in general that $B(p, \frac{1}{n}) \subseteq B(x,r) \subseteq U$. Hence we conclude that U can be written as a union of elements of $\mathcal{U}'$, which implies that $\mathcal{U}'$ is a countable basis for M.

By means of Proposition 12.5.15 we conclude the equivalence chain. ♮

Corollary 12.5.17. Euclidean spaces are countable.

Proof. Lets consider $\mathbf{R}^n$, notice that $\mathbf{Q}^n$ is dense in $\mathbf{R}^n$ and countable, thus $\mathbf{R}^n$ is separable, which by Proposition 12.5.16 implies that $\mathbf{R}^n$ is second countable. ♮

Weights and Cardinality

Proposition 12.5.18. Let X be a topological space and $w(X) \leqslant \mathbf{m}$. Then for every collection $\{U_i\}_{i \in I} \subseteq 2^X$ of open sets, there exists $I_0 \subseteq I$ such that $|I_0| \leqslant \mathbf{m}$ and $\bigcup_{i \in I_0} U_i = \bigcup_{i \in I} U_i$.

Proof. Since $I_0 \subseteq I$ then clearly $\bigcup_{i \in I_0} U_i \subseteq \bigcup_{i \in I} U_i$. Let $\mathcal{B}$ be a base for X such that $|\mathcal{B}| \leqslant \mathbf{m}$ and define the collection $\mathcal{B}_0 := \{U \in \mathcal{B} : U \subseteq U_i (i \in I)\}$. Define a function $f : \mathcal{B}_0 \to I$ such that for all $U \in \mathcal{B}_0$ we have $U \subseteq U_{f(U)} \in \{U_i\}_{i \in I}$. Define the indexing set $I_0 := f(\mathcal{B}_0) \subseteq I$. Notice that $|I_0| \leqslant |\mathcal{B}| \leqslant \mathbf{m}$. For any point $x \in \bigcup_{i \in I} U_i$ there exists $i \in S$ such that $x \in U_i$, and hence exists $U \in \mathcal{B}$ such that $x \in U \subseteq U_i$, from the fact that $\mathcal{B}$ is a basis. From the definition of $\mathcal{B}_0$ it follows that $U \in \mathcal{B}_0$ and hence $f(U) \in I_0$. Therefore, from the construction of f it follows that $x \in U \subseteq U_{f(U)} \subseteq \bigcup_{i \in I_0} U_i$. From this we conclude that $\bigcup_{i \in I} U_i \subseteq \bigcup_{i \in I_0} U_i$. ♮

Proposition 12.5.19. Let X be a topological space. If $w(X) \leqslant \mathbf{m}$, then for every base $\mathcal{B}$ for the topology of X there exists a base $\mathcal{B}_0$ such that $|\mathcal{B}_0| \leqslant \mathbf{m}$ and $\mathcal{B}_0 \subseteq \mathcal{B}$.

Proof. Suppose $\mathbf{m} \geqslant \aleph_0$. Let $\mathcal{B} := \{U_i\}_{i \in I}$ be any base for the space X. Define a base $\mathcal{B}_1 := \{W_t\}_{t \in T}$ for X such that $|T| \leqslant \mathbf{m}$.

For all $t \in T$ define the set $I(t) := \{i \in I : U_i \subseteq W_t\}$. From the fact that W_t is open and $\mathcal{B}$ is a base for X, we have that $W_t = \bigcup_{i \in I(t)} U_i$. From Proposition 12.5.18 we find that there exists $I_0(t) \subseteq I(t)$ for which $|I_0(t)| \leqslant \mathbf{m}$ and

$$W_t = \bigcup_{i \in I(t)} U_i = \bigcup_{i \in I_0(t)} U_i \tag{12.1}$$

Now, define the set $\mathcal{B}_0 := \{U_i\}_{i \in I(t), t \in T} \subseteq \mathcal{B}$. Since $|T|, |S_0(t)| \leqslant \mathbf{m}$ and the fact that $\mathbf{m}^2 = \mathbf{m}$, we find that $|\mathcal{B}_0| \leqslant \mathbf{m}$. We now show that $\mathcal{B}_0$ is a base for the space X. Let a any point $x \in X$ and a neighbourhood $U \subseteq X$ of x. From the hypothesis that $\mathcal{B}_1$ is a basis, there exists $t \in T$ such that $W_t \subseteq U$. On the other hand Eq. (12.1) ensures the existence of an $i \in I_0(t)$ such that $U_i \subseteq W_t \subseteq U$. Since $U_i \in \mathcal{B}_0$, we find that $\mathcal{B}_0$ is indeed a basis for X. ♮

12.6 Filters and Nets

Definition 12.6.1 (Filter on a set)**.** A filter on a set X is a collection $\mathcal{F} \subseteq 2^X$ that satisfies the following properties

(F1) (Downward directed) Given $A, B \in \mathcal{F}$, then there exists $C \in \mathcal{F}$ such that $C \subseteq A \cap B$.

(F2) (Upward closed) If $A \in \mathcal{F}$ and $A \subseteq B$ then $B \in \mathcal{F}$.

(F3) $\mathcal{F}$ is non-empty.

Definition 12.6.2 (Proper filter)**.** If $\mathcal{F}$ is a filter of X such that there exists $A \subseteq X$ for which $A \notin \mathcal{F}$, then we say that $\mathcal{F}$ is a proper filter. This is equivalent of saying that $\varnothing \notin \mathcal{F}$.

Proposition 12.6.3 (Eventuality filter)**.** Let X be a topological space and $(x_i)_{i \in \mathbf{N}}$ be a sequence of points of X. The collection of sets such that (x_i) is eventually in, explicitly

$$\mathcal{E}_{(x_i)} = \{U \subseteq X : \forall i \geqslant N, x_i \in U (N \in \mathbf{N})\},$$

is a filter on X.

Definition 12.6.4 (Filter base)**.** A non-empty downward directed set is called a filter base.

Proposition 12.6.5. Given a filter base $\mathcal{G}$ on X, we define the filter generated by $\mathcal{G}$ as the collection

$$\mathcal{G}^{\uparrow} = \{A \subseteq X : G \subseteq A (G \in \mathcal{G})\}.$$

Proof. Since $\mathcal{G}$ is non-empty, $\mathcal{G}^{\uparrow}$ is non-empty. Let $A, B \in \mathcal{G}^{\uparrow}$, then there are $G, H \in \mathcal{G}$ such that $G \subseteq A$ and $H \subseteq B$. Since $G \cap H \in \mathcal{G}$, then $G \cap H \subseteq A \cap B$ and thus $A \cap B \in \mathcal{G}^{\uparrow}$. Suppose that $A \subseteq C$ where $C \subseteq X$, then since $G \subseteq A$, it follows that $G \subseteq C$ and hence $C \in \mathcal{G}^{\uparrow}$. Therefore $\mathcal{G}^{\uparrow}$ satisfies all of the three requirements. ♮

Proposition 12.6.6. Let X be a topological space. The base at a point $p \in X$, namely $\mathcal{B}_p$, is a filter base.

Proof. Let $U, V \in \mathcal{B}_p$ be neighbourhoods of p. Since $p \in U \cap V$ and the finite intersection of open sets is open, then $U \cap V$ is a neighbourhood of p and hence $U \cap V \in \mathcal{B}_p$. ♮

Definition 12.6.7 (Convergence of filters)**.** Let $\mathcal{F}$ be a filter on a topological space (X, τ). We say that $\mathcal{F}$ converges to x, $\mathcal{F} \to x$, if and only if $\mathcal{B}_x \subseteq \mathcal{F}$.

Filters determine Hausdorff, closure and continuity

Proposition 12.6.8 (Hausdorff)**.** A topological space is Hausdorff if and only if limits of convergent proper filters are unique.

Proof. ($\Rightarrow$) Let X be a Hausdorff space and $\mathcal{F}$ be a proper filter on X. Suppose that for distinct $x, y \in X$ we have $\mathcal{F} \to x$ and also $\mathcal{F} \to y$. Since X is Hausdorff, let $U, V \subseteq X$ neighbourhoods of x and y, respectively, and such that $U \cap V$ is empty. Notice that since U is a neighbourhood of x then $U \in \mathcal{F}$, and analogously, $V \in \mathcal{F}$, but from the downward directness, it follows that $U \cap V = \varnothing \in \mathcal{F}$, which can't be the case since $\mathcal{F}$ is supposed to be proper. Hence $\mathcal{F}$ cannot converge to distinct points of X. ($\Leftarrow$) Suppose that X is not Hausdorff and choose distinct points x, y that are not separable by open sets. Consider the collection $\mathcal{B} = \mathcal{B}_x \cap \mathcal{B}_y$, then given two sets $A, B \in \mathcal{B}$ we have that $A \cap B \in \mathcal{B}$ and therefore $\mathcal{B}$ is a filter base. Notice that the filter $\mathcal{B}^{\uparrow}$ converges both to x and y. ♮

Proposition 12.6.9 (Closed)**.** Let X be a topological space and $A \subseteq X$. A point $p \in \operatorname{Cl} A$ if and only if there exists a proper filter $\mathcal{F}$ with $A \in \mathcal{F}$ such that $\mathcal{F} \to p$.

Proof. ($\Rightarrow$) Let $p \in \operatorname{Cl} A$, then for all neighbourhoods $U \in \mathcal{B}_p$ the set $U \cap (A \setminus \{p\})$ is non-empty, in particular we have that $\mathcal{B} := \{U \cap A : U \in \mathcal{B}_p\}$ does not contain the empty set. Hence the proper filter $\mathcal{B}^{\uparrow}$converges to p. ($\Leftarrow$) Let $\mathcal{F}$ be a proper filter with $\mathcal{F} \to p$ and $A \in \mathcal{F}$, then in particular we have that the downward directness implies $\mathcal{B} \subseteq \mathcal{F}$ and thus $\varnothing \notin \mathcal{B}$, hence $x \in \operatorname{Cl} A$. ♮

Definition 12.6.10 (Pushforward of filters)**.** Let $f: X \to Y$ be a map of sets. The collection of images $\{f(A) : A \in \mathcal{F}\}$ form a filter base whose generated filter is defined to the be the pushforward of $\mathcal{F}$ with respect to f, namely $f_*(\mathcal{F})$. Hence

$$f_*(\mathcal{F}) = \{B \subseteq Y : f(A) \subseteq B(A \in \mathcal{F})\}.$$

Proposition 12.6.11 (Continuity)**.** A map $f: X \to Y$ is continuous if and only if for every given filter $\mathcal{F}$ on X such that $\mathcal{F} \to x$, then $f_*(\mathcal{F}) \to f(x)$, where $x \in X$.

Proof. ($\Rightarrow$) Suppose f is a continuous map and $\mathcal{F} \to x$. From continuity of f, for any given neighbourhood of $f(x)$, $B \in \mathcal{B}_{f(x)}$, there exists a corresponding neighbourhood of x, $V \in \mathcal{B}_x$, such that $f^{-1}(B) \subseteq V$. Since $\mathcal{F} \to x$ if and only if $\mathcal{B}_x \subseteq \mathcal{F}$ (from Definition 12.6.7), then in particular $f^{-1}(B) \in \mathcal{F}$, since $f^{-1}(B) \in \mathcal{B}_x$, is a neighbourhood of x. Hence, if $B \in \mathcal{B}_{f(x)}$ is any element and $A = f^{-1}(B)$, then $f(A) \subseteq B$, which implies that $\mathcal{B}_{f(x)} \subseteq f_*(\mathcal{F})$ from the definition of the pushforward, therefore $f_*(\mathcal{F}) \to f(x)$.

($\Leftarrow$) Suppose now that for any filter $\mathcal{F}$ on X such that $\mathcal{F} \to x$, implies $f_*(\mathcal{F}) \to f(x)$. Given any open set $U \subseteq Y$, if $U \cap \operatorname{im}(f) = \varnothing$ then $f^{-1}(U) = \varnothing$ and hence is open, otherwise there exists $x \in X$ for which $f(x) \in U$. Given such a point $x \in X$, take $\mathcal{F} = \mathcal{B}_x^{\uparrow}$ so that from hypothesis $f_*(\mathcal{F}) \to f(x)$ and hence $\mathcal{B}_{f(x)} \subseteq f_*(\mathcal{F})$. Since U is open, then $U \in \mathcal{B}_{f(x)}$, which in turn implies that $U \in f_*(\mathcal{F})$ and from definition there must exist a set $V \in \mathcal{F}$, which happens to be a neighbourhood of x, such that $f(V) \subseteq B$, thus f is continuous. ♮

Write on nets later

12.7 Topological Manifolds

Locally Euclidean

Definition 12.7.1 (Locally euclidean)**.** Let M be a topological space. We say that M is locally euclidean of dimension n if for all points $x \in X$ there exists a neighbourhood $U_x \subseteq M$ of x such that exists an open set $V \subseteq \mathbf{R}^n$ for which $U_x \simeq V$, that is, U_x is isomorphic to V.

Proposition 12.7.2. Let $B^n \subseteq \mathbf{R}^n$ be any open ball, then the morphism

$$\varphi\colon B^n \to \mathbf{R}^n,\ x \longmapsto \frac{x}{1-|x|}$$

is an isomorphism $B^n \simeq \mathbf{R}^n$.

Lemma 12.7.3. A topological space M is locally euclidean of dimension n if and only if either of the following is true:

(a) Every point of M has a neighbourhood in M that is isomorphic to an open ball in $\mathbf{R}^n$.

(b) Every point of M has a neighbourhood in M that is isomorphic to $\mathbf{R}^n$.

Proof. From Proposition 12.7.2, we find that the proof of one of the two statements proves the other. Lets prove the first one of them. ($\Leftarrow$) Since $B^n \subseteq \mathbf{R}^n$ is an open of $\mathbf{R}^n$ we conclude from the definition Definition 12.7.1 that M is indeed locally euclidean. ($\Rightarrow$) Suppose M is a locally euclidean n-dimensional topological space, then let any point $p \in M$ and define $U_p \subseteq M$ be a neighbourhood of p such that exists $V \subseteq \mathbf{R}^n$ for which we can define an isomorphism $\phi\colon U_p \to V$. From the fact that the collection of open balls in $\mathbf{R}^n$ form a basis for $\mathbf{R}^n$, we conclude that there exists an open ball $B^n \subseteq V$ for which $p \in B^n$. Thus the open set $\psi^{-1}(B) \subseteq M$ is a neighbourhood of p. Since ϕ is an isomorphism, then the morphism $\psi|_B^{-1}\colon B \to \phi^{-1}(B)$ is an isomorphism, thus $B \simeq \phi^{-1}(B)$ as wanted. ♮

Proposition 12.7.4. Every locally euclidean space is first countable.

prove

Definition 12.7.5 (Miscellaneous)**.** Let M be a locally euclidean n-dimensional topological space. We define:

(C1) (Coordinate domain) An open set $U \subseteq M$ is called a coordinate domain of M if it is isomorphic to an open set of $\mathbf{R}^n$.

(C2) (Coordinate map) An isomorphism ϕ from a coordinate domain to an open set of $\mathbf{R}^n$ is called a coordinate map.

(C3) (Coordinate chart) The pair (U, ϕ) of a coordinate domain and one of its coordinate maps is called a coordinate chart for M. Given $x \in M$, we say that (U, ϕ) is a chart *at* x if $x \in U$.

(C4) (Coordinate ball) A coordinate domain isomorphic to an open ball of $\mathbf{R}^n$ is called a coordinate ball.

(C5) (Coordinate neighbourhood or euclidean neighbourhood) Given a point $p \in M$, if $U \subseteq M$ is a coordinate domain of M such that $p \in U$, then we say that U is a coordinate neighbourhood of p.

Proposition 12.7.6. Every coordinate ball is second countable.

Proof. Let M be a locally euclidean n-dimensional topological space and let $U \subseteq M$ be a coordinate ball of M and $\phi\colon U \xrightarrow{\simeq} B^n$ be an isomorphism. From Corollary 12.5.17 we find that there exists a countable basis $\mathcal{B} \subseteq 2^{\mathbf{R}^n}$ for B^n. Since ϕ is an isomorphism, then the collection of preimages $\{\phi^{-1}(B) \subseteq U : B \in \mathcal{B}\}$ is also a basis for U and is clearly countable — hence U is second countable. ♮

Proposition 12.7.7 (Locally euclidean from surjective morphism)**.** Let X be locally euclidean of dimension n, and $f\colon X \to Y$ be a surjective local isomorphism. Then Y is locally euclidean of dimension n.

Proof. Let $\mathcal{B}$ be a basis for X. Since f is continuous, surjective and a open (see Proposition 12.2.14), we can use Proposition 12.2.7 to conclude that $f(\mathcal{B})$ is a basis for Y. Since f is a local isomorphism, given any point $y \in Y$, choose a neighbourhood $V_y \in f(\mathcal{B})$, so that there exists a $B \in \mathcal{B}$ such that $f(B) = V_y$. Consider the isomorphism $f|_B\colon B \to V_y$. Since X is locally euclidean of dimension n, there exists an open set $W \subseteq \mathbf{R}^n$ such that $B \simeq W$. Moreover, since $B \simeq V_y$, then $V_y \simeq W$. We conclude that Y is locally euclidean of dimension n. ♮

Topological Manifold

Definition 12.7.8 (Topological manifold)**.** An n-dimensional topological manifold is a second countable Hausdorff space that is locally euclidean n-dimensional.

Proposition 12.7.9. Every topological manifold admits a basis of coordinate balls.

Proof. Let M be a n-dimensional topological manifold. From Lemma 12.7.3 we can take, for every point $p \in M$, a neighbourhood $U \subseteq M$ such that $U \simeq B^n$ (an open ball of $\mathbf{R}^n$). Define $\mathcal{U}$ to be the collection of all coordinate balls on M, from our last argument it follows that $\mathcal{U}$ covers M. Let $U, U' \in \mathcal{U}$ be intersecting coordinate balls and ϕ be a coordinate isomorphism of either U or U'. Let $p \in U \cap U'$ be any point and define a ball $B^n_p \subseteq \phi(U \cap U') \subseteq \mathbf{R}^n$ that is a neighbourhood of $\phi(p) \in \mathbf{R}^n$. Let $V \subseteq M$ be defined as $V := \phi^{-1}(B^n)$ so that $V \subseteq U \cap U'$ and also $p \in V$. Notice that the induced map $\psi\colon V \to B^n$, defined by $\psi(x) := \phi(x)$, is an isomorphism — that is, $V \in \mathcal{U}$ and hence $\mathcal{U}$ is a basis for M (see Proposition 12.1.5). ♮

Proposition 12.7.10. Every open subset of an n-manifold is itself an n-manifold.

Proof. Consider $U \subseteq M$ an open set and let $p \in U$. Then consider $V \subseteq M$ to be a coordinate neighbourhood of x such that $V \simeq B \subseteq \mathbf{R}^n$, then the set $U \cap V$ is open and isomorphic to a subset of B and hence V is n-dimensional locally euclidean. Moreover, since a subset of a Hausdorff space is Hausdorff and a subset of a second countable space is second countable, then U is indeed a n-manifold. ♮

Definition 12.7.11. The empty topological space is an n-manifold for all $n > 0$.

Theorem 12.7.12 (Dimension invariance)**.** If $m \neq n$, then a non-empty topological space cannot be both n-manifold and m-manifold.

Proposition 12.7.13. A separable metric space that is locally euclidean of dimension n is an n-manifold.

Proof. From Proposition 12.5.16, since M is separable then it is also second countable. Moreover, from Proposition 12.4.9 we find that M is Hausdorff. Hence M is an n-manifold. ♮

Proposition 12.7.14. Every topological manifold is separable and metrizable.

Proof. Notice that since a manifold is second countable, then by Proposition 12.5.15 we find that it is separable.

After proving Urysohn metrization theorem, prove the metrizability property

♮

Manifolds with Boundary

Definition 12.7.15 (Upper half-space)**.** We define the closed n-dimensional upper half-space $\mathbf{H}^n \subseteq \mathbf{R}^n$ as

$$\mathbf{H}^n := \{x \in \mathbf{R}^n : \pi_n(x) \geqslant 0\}$$

where π_n is the projection of the n-th coordinate. We define the boundary of $\mathbf{H}^n$ as $\partial \mathbf{H}^n := \{x : \pi_n(x) = 0\}$, and the interior as $\operatorname{Int}(\mathbf{H}^n) := \{x : \pi_n(x) > 0\}$.

Definition 12.7.16 (Manifold with boundary)**.** We define an n-dimensional topological manifold with boundary to be a second countable Hausdorff space such that each point has a neighbourhood isomorphic to an open set of $\mathbf{R}^n$ or $\mathbf{H}^n$.

Definition 12.7.17 (Miscellaneous)**.** Let M be an n-manifold with boundary. We define

(MB1) A coordinate chart for M is a pair (U, ϕ), where $U \subseteq M$ is an open set and $\phi\colon U \xrightarrow{\simeq} V$ is an isomorphism, where $V \subseteq \mathbf{R}^n$ or $V \subseteq \mathbf{H}^n$. We say that the chart is an interior chart if V is an open subset of $\mathbf{R}^n$. A chart is said to be a boundary chart if V is an open subset of $\mathbf{H}^n$ with $\operatorname{im}(\phi) \cap \partial \mathbf{H}^n \neq \varnothing$.

(MB2) A point $p \in M$ is called an interior point of M if it is contained in the domain of an interior chart. The collection of such points is called the interior of M, and is denoted $\text{Int}(M)$.

(MB3) A point $p \in M$ is called a boundary point of M if it is in the domain of a boundary chart that maps p to a point of $\mathbf{H}^n$. The boundary of M, is defined as the collection of such points and is denoted by ∂M.

Proposition 12.7.18. If M is an n-dimensional manifold with boundary, then $\text{Int}(M)$ is an open subset of M, which is itself an n-dimensional manifold (without boundary).

Proof. Let $p \in \text{Int}(M)$ be any point, then by definition we have that $p \in U$ where (U, ϕ) is an interior chart for M and therefore $U \simeq V$ where V is some open subset of $\mathbf{R}^n$, hence locally euclidean, which makes $\text{Int}(M)$ an n-dimensional manifold (since a subset of a Hausdorff space is Hausdorff and a subset of a second countable space is again second countable). To prove that $\text{Int}(M)$ is open, we can use the fact that M is first countable and lemma Lemma 12.5.6. First, suppose the converse, so that $(x_i)_{i\in\mathbf{N}}$ is a sequence of points in M that is not-eventually in $\text{Int}(M)$ and, for the sake of contradiction, $x_i \to p$. Consider the open neighbourhood U of p (which happens to be the chart domain), if $x_i \to p$ then there exists $N \in \mathbf{N}$ such that $x_i \in U$ for all $i \geqslant N$, but notice that every point of U is an interior point of M, therefore $x_i \in \text{Int}(M)$, which is a contradiction to the hypothesis that the sequence is not-eventually in $\text{Int}(M)$. Hence such sequences cannot converge to p, and $\text{Int}(M)$ is open. ♮

Theorem 12.7.19 (Boundary invariance)**.** Let M be a manifold with boundary, then

$$\partial M \cap \text{Int}(M) = \varnothing$$

Proof.

Proof to come far ahead

♮

Corollary 12.7.20. If M is a non-empty n-manifold with boundary, then the collection ∂M is closed in M, and M is an n-manifold if and only if $\partial M = \varnothing$.

Proof. Since $\partial M = M \setminus \text{Int}(M)$ from theorem Theorem 12.7.19 and $\text{Int}(M)$ is open from Proposition 12.7.18, then it follows that ∂M is closed. ($\Rightarrow$) Moreover, suppose that M is a manifold, then given any point $p \in M$ there exists an interior chart (U, ϕ) such that $p \in U \subseteq \text{Int}(M)$, hence $\text{Int}(M) = M$, and from Theorem 12.7.19 we conclude that $\partial M = \varnothing$. ($\Leftarrow$) Suppose that $\partial M = \varnothing$, then from Theorem 12.7.19 we conclude that $M = \text{Int}(M)$, which is a manifold by Proposition 12.7.18. ♮

Chapter 13

Top — Universal Properties

13.1 Prelude

Before diving into the construction of new spaces from old using the classical categorical notions of limits over diagrams, lets first stablish the definition of initial and final topologies, that will classify the kind of topology each of these new constructions receive.

Definition 13.1.1 (Initial & final topology)**.** Let S and J be sets, and consider a collection of topological spaces $(X_j)_{j\in J}$. We define the following notions:

(a) Given a collection of *set-functions* $(f_j\colon S \to X_j)_{j\in J}$, we define the *initial topology* τ_{initial} on S — induced by $(f_j)_{j\in J}$ — to be the *minimum* collection of open sets such that $f_j\colon (S, \tau_{\text{initial}}) \to X_j$ are *continuous maps*.

(b) Given a collection of *set-functions* $(g_j\colon X_j \to S)_{j\in J}$, we define the *final topology* τ_{final} on S — induced by $(g_j)_{j\in J}$ — to be the *maximum* collection of open sets such that $g_j\colon X_j \to (S, \tau_{\text{final}})$ are *continuous maps*.

13.2 Subspace

Construction

Definition 13.2.1 (Subspace topology)**.** Given a topological space (X, τ) and a subset $S \subseteq X$, we define the space (S, τ_S) as a subspace of X if

$$\tau_S = \{U \subseteq S : U = S \cap V,\ V \in \tau\}.$$

If so, we call τ_S the relative topology or subspace topology.

Proposition 13.2.2. Let X be a topological space and S be a subspace of X. A set $A \subseteq S$ is closed in S if and only if $A = S \cap F$ for some closed set $F \subseteq X$ with respect to X. The closure of A with respect to S, denoted $\widetilde{A}$, is such that $\widetilde{A} = S \cap \operatorname{Cl} A$.

Proof. ($\Rightarrow$) Suppose A is a closed set in S, then $S \setminus S = S \cap U$ is open and hence U is some open set of X. Then

$$A = S \setminus (S \setminus A) = S \setminus (S \cap U) = S \cap (X \setminus U)$$

where $X \setminus U$ is closed, hence A equals to the intersection of S with a closed set of X. ($\Leftarrow$) Suppose that $A = S \cap F$ for a closed set F in X. Then we have

$$S \setminus A = S \setminus (S \cap F) = S \cap (X \setminus F)$$

Since $X \setminus F$ is open, then $S \setminus A$ is open. We conclude that A is closed in S.

For the last proposition, notice that $\widetilde{A} = \bigcap\{F \subseteq S : F \supseteq A, F \text{ is closed in } S\}$, that is, the intersection of sets $S \cap C$ such that C is closed in X and $C \supseteq A$, so that $\widetilde{A} = S \cap \operatorname{Cl} A$. ♮

Proposition 13.2.3. Let X be a topological space and S be a subspace. Then

(a) If $U \subseteq S$ is open (resp. closed) and S is an open (resp. closed) subset of X, then U is open (resp. closed) in X.

(b) If $U \subseteq S$ and U is open (resp. closed) in X, then it is open (resp. closed) in S.

Proof. (a) Suppose that U is open in S and S is open in X. From the definition of the relative topology, there exists an open $V \subseteq X$ in X such that $U = S \cap V$. Since S and V are both open in X, then U is open in X. From Proposition 13.2.2 we find that if U is closed in S then $B = S \cap F$ for some closed set $F \subseteq X$. Moreover, if S is closed in X, then clearly B is closed in X.

(b) If $B \subseteq S$ then clearly $B \cap S = B$, hence if B is open in X, so it is in S. Moreover, if B is closed in X, from Proposition 13.2.2 B is closed in S. ♮

Definition 13.2.4 (Second definition of the subspace topology)**.** Let X be a topological space and $S \subseteq X$ be any set. The subspace topology on S is the initial topology on the set such that the inclusion $\iota_S \colon S \to X$ is continuous.

Corollary 13.2.5. Definitions 13.2.1 and 13.2.4 are equivalent.

Proof. Suppose τ_{ι_S} is the initial topology such that ι_S is continuous and τ_S be the subspace topology. Our goal is to show that they are, in fact, equal. Given any $U = S \cap V \in \tau_S$ we have $\iota_S^{-1}(V) = S \cap V = U$, since V is open, we find that U is open in τ_{ι_S}. This implies in $\tau_S \subseteq \tau_{\iota_S}$. Moreover, let $O \in \tau_{\iota_S}$, then there must exist some $A \subseteq X$ open set for which $\iota_S^{-1}(A) = A \cap S = O$ (this comes directly from the fact that τ_{ι_S} was solely constructed for the purpose of making ι_S continuous), since A is open, then $O = A \cap S \in \tau_S$. This implies that $\tau_{\iota_S} \subseteq \tau_S$. We conclude finally that both definitions are indeed equivalent. ♮

Theorem 13.2.6 (Universal property of the subspace topology)**.** Let X be a topological space and S be a subspace. Given any topological space Y, a map $f \colon Y \to S$ is continuous if and only if $\iota_S f : Y \to X$ is continuous, where $\iota_S : S \hookrightarrow X$ is the inclusion map. Hence the following diagram commutes

$$\begin{tikzcd} Y \arrow[r, "f"] \arrow[dr, "\iota_S f"'] & S \arrow[d, hook, "\iota_S"] \\ & X \end{tikzcd}$$

Moreover, on the converse, if $S \subseteq X$ is a topological space such that the above property holds, then it is equipped with the subspace topology.

Proof. First we show that if (S, τ_S) is a subspace of (X, τ) then it satisfies the universal property. Suppose then that S a subspace of X. ($\Rightarrow$) Let f be continuous. If $V \subseteq X$ is any open subset, we have that

$$(\iota_S f)^{-1}(V) = f^{-1}(\iota^{-1}(V)) = f^{-1}(S \cap V)$$

since V is said to be open, then $S \cap V$ is open in S, which implies that $f^{-1}(S \cap V) = (\iota_S f)^{-1}(V)$ is open. This shows that the map $\iota_S f$ is open. ($\Leftarrow$) Let $\iota_S f$ be continuous. Consider $U = S \cap V = \iota_S^{-1}(V)$ to be any open set of the subspace S (that is V is an open of the space X). Then we have

$$f^{-1}(U) = f^{-1}(\iota_S^{-1}(V)) = (\iota_S f)^{-1}(V)$$

since $\iota_S f$ is continuous, then $f^{-1}(U)$ is open, therefore f is continuous.

We now show that if an object satisfies such the property, then it is the subspace. Let (S, τ') be a space satisfying the universal property. In particular we can take the subspace (S, τ_S) of (X, τ) and the identity map $\mathrm{id}_S : (S, \tau_S) \to (S, \tau')$. Since (S, τ') satisfies the universal property, we have that id_S is continuous if and only if $\iota_S \,\mathrm{id}_S = \iota_S$ is continuous. That is

$$\begin{tikzcd} (S, \tau_S) \arrow[r, "\mathrm{id}_S"] \arrow[dr, "\iota_S\,\mathrm{id}_S = \iota_S"'] & (S, \tau') \arrow[d, "\iota_S"] \\ & (X, \tau) \end{tikzcd}$$

We know from Definition 13.2.4 that $\iota_S \,\mathrm{id}_S = \iota_S$ is continuous for the subspace topology τ_S, hence the universal property implies that id_S is continuous. In particular, this says that $\tau' \subseteq \tau_S$. In order to show the other side of the equality, consider now the space (S, τ') and the identity map $\mathrm{id}'_S : (S, \tau') \to (S, \tau')$ so that from the universal property of (S, τ') the map id'_S is continuous if and only if $\iota_S \,\mathrm{id}'_S = \iota_S$ is continuous. That is, the following diagram commutes

$$\begin{tikzcd} (S, \tau') \arrow[r, "\mathrm{id}'_S"] \arrow[dr, "\iota_S\,\mathrm{id}'_S = \iota_S"'] & (S, \tau') \arrow[d, "\iota_S"] \\ & (X, \tau) \end{tikzcd}$$

Notice that since id'_S is continuous (see Proposition 12.2.3) then from the universal property the map ι_S is continuous on τ'. Since τ_S is the initial topology such that ι_S is continuous (see Definition 13.2.4), then clearly $\tau_S \subseteq \tau'$. This finishes the proof that $\tau' = \tau_S$ and hence the space $(S, \tau') = (S, \tau_S)$ is the subspace of (X, τ). $\natural$

Corollary 13.2.7. Let X, Y be topological spaces and $f\colon X \to Y$ be a continuous map. Then the following hold

(a) (Domain restriction) Let S be a subspace of X. Then $f|_S$ is continuous.

(b) (Codomain restriction) Let T be a subspace of Y such that $f(X) \subseteq T$. Then $f\colon X \to T$ is continuous.

(c) (Codomain expansion) Let Y be a subspace of Z. Then the map $f\colon X \to Z$ is continuous.

Proof. (a) Notice that $f|_S = f\iota_S$. Applying the universal property, we find that since f is continuous, so is $f|_S$. (b) From the universal property we have that $\iota_T f = f$ is continuous, so is $f\colon X \to T$. (c) Notice that from the universal property we have $\iota_Y f\colon X \to Z$ continuous, since f is continuous. The following are the universal property diagrams for items (b) and (c):

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "\iota_T f = f"'] & T \arrow[d, hook, "\iota_T"] \\ & Y \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "\iota_Y f"'] & Y \arrow[d, hook, "\iota_Y"] \\ & Z \end{tikzcd}$$

$\natural$

Proposition 13.2.8 (Subspace properties)**.** Let (X, τ) be a topological space and (S, τ_S) be a subspace of X. The following are properties concerning the subspace topology

(SP1) Let T be a subspace of S. Then T is a subspace of X.

(SP2) Let $\mathcal{B}$ be a basis for X. Then the collection $\mathcal{B}_S = \{B \cap S : B \in \mathcal{B}\}$ is a basis for S.

(SP3) Let $(p_i)_{i \in \mathbf{N}} \subseteq S$ be a sequence and $p \in S$. Then $p_i \to p$ in S if and only if $p_i \to p$ in X.

(SP4) A subspace of a Hausdorff space is Hausdorff.

(SP5) A subspace of a first countable space is first countable.

(SP6) A subspace of a second countable space is second countable.

Proof. (SP1) Let Z be some topological space, we can choose maps $f\colon Z \to T$ and $g = \iota_T f\colon Z \to S$ and apply the universal property on both T and S in order to get

$$\begin{tikzcd} Z \arrow[r, "f"] \arrow[dr, "g"'] \arrow[ddr, bend right, "\iota_S g = \iota'_T f"'] & T \arrow[d, hook, "\iota_T"] \arrow[dd, bend left=60, "\iota_S \iota_T := \iota'_T"] \\ & S \arrow[d, hook, "\iota_S"] \\ & X \end{tikzcd} \quad \Leftrightarrow \quad \begin{tikzcd} Z \arrow[r, "f"] \arrow[dr, "\iota'_T f"'] & T \arrow[d, hook, "\iota'_T"] \\ & X \end{tikzcd}$$

Moreover, if f is continuous, from the universal property of T we find that g is continuous, but using the universal property of S that tells us that $\iota_S g = \iota'_T f$ is continuous. The converse is true by using the same argumentation. Hence T satisfies the universal property and therefore is a subspace of X.

(SP2) First notice that since $B \in \mathcal{B}$ is open in X, then $\mathcal{B}_S \subseteq \tau_S$. Let $s \in S$ be any element, then in particular we have $s \in B_s \in \mathcal{B}$ for some element of the basis. Then $s \in S \cap B_s \in \mathcal{B}_S$. Let $A = S \cap B_1, B = S \cap B_2 \in \mathcal{B}_S$ and $x \in A \cap B$. Then in particular $x \in B_1 \cap B_2$. Since $\mathcal{B}$ is a basis for X, then exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$. Hence the corresponding set $C = S \cap B_3 \in \mathcal{B}_S$ is such that $x \in C \subseteq A \cap B$. This proves that $\mathcal{B}_S$ is a basis.

(SP3) ($\Rightarrow$) Suppose $p_i \to p$ in S, that is, for all $U_p = S \cap V_p \in \tau_S$, where $V_p \in \tau$, we have some $N \in \mathbf{N}$ such that for all $n \geqslant N$ we have $p_n \in U_p$. In particular, this implies that for each $n \geqslant N$ we get $p_n \in V_p$, hence $p_i \to p$ in X. ($\Leftarrow$) Suppose that $p_i \to p$ in X. Let $V_p \subseteq X$ be a neighbourhood of p, then since $p \in S$ we have that $V_p \cap S \neq \varnothing$ is an element of τ_S. Let $U_p \in \tau_S$ be any neighbourhood of p, then there exists $V \in \tau$ such that $U_p = S \cap V$ (from the definition of the subspace topology) and also $p \in V$, so that V is a neighbourhood of p. Therefore there exists $M \in \mathbf{N}$ such that, for all $n \geqslant M$ we have $p_n \in S \cap V = U_p \subseteq V$. This implies that $p_i \to p$ in S.

(SP4) Let X be Hausdorff. Let $x, y \in S$ be distinct points. In particular, there exits $A, B \in X$ neighbourhoods of x and y, respectively, such that $A \cap B = \varnothing$. Hence the sets $U = S \cap A, V = S \cap B \in \tau_S$ are neighbourhoods of x and y respectively and since $U \subseteq A$ and $V \subseteq B$ we find that $U \cap V = \varnothing$. Hence S is Hausdorff.

(SP5) Let X be a first countable space. Let $p \in S$ be any point. Define a countable base $\mathcal{B}_p$ of neighbourhoods of p for X. Define the, clearly countable, collection $\mathcal{B}'_p = \{S \cap B_p : B_p \in \mathcal{B}_p\} \subseteq 2^S$. From Item 2 we have that $\mathcal{B}'_p$ is a basis for S. Moreover, every element of $\mathcal{B}'_p$ is clearly a neighbourhood of p. This proves the property.

(SP6) Let X be second countable and $\mathcal{B}$ be a countable basis for the topology of X. Define the, clearly countable, collection $\mathcal{B}' := \{S \cap B : B \in \mathcal{B}\}$. Since $\mathcal{B}$ is a basis for X, we use Item 2 to conclude that $\mathcal{B}'$ is a basis for S. ♮

Proposition 13.2.9. Let X be a topological space and $A, B \subseteq X$ be subspaces such that

- Either A and B are *both* open subsets of X or are *both* closed.
- They *cover* X, that is, $X = A \cup B$.

Then the following commutative diagram is a *pushout* in `Top`

$$\begin{tikzcd} A \cap B \arrow[r] \arrow[d] & A \arrow[d, hook] \\ B \arrow[r, hook] & X \end{tikzcd} \tag{13.1}$$

In other words, if $f : X \to Y$ is any *set-function* between topological spaces, then f is *continuous* if the restrictions $f|_A$ and $f|_B$ are continuous.

Proof. Since diagram Eq. (13.1) is a pushout in `Set`, one just has to prove that X has the final topology induced by the natural inclusions $\iota_A : A \hookrightarrow X$ and $\iota_B : B \hookrightarrow X$ — that is,

$U \subseteq X$ is open if and only if $\iota_A^{-1}(U) = U \cap A$ is open in A and $\iota_B^{-1}(B) = U \cap B$ is open in B.

Indeed, if U is an open set of X, then by definition of the subspace topology we conclude that $\iota_A^{-1}(U)$ is open in A and $\iota_B^{-1}(U)$ is open in B.

On the other hand, assume that $U \cap A$ is open in A, while $U \cap B$ is open in B. We separate the proof in two cases:

- Suppose that A and B are open subspaces of X. Since $A \cap U$ is open, there must exist an open set $V_A \subseteq X$ such that $\iota_A^{-1}(V_A) = V_A \cap A = U \cap A$ — furthermore, since A and V_A are open in X, it follows that $U \cap A$ is open in X. The same argument can be made for the subspace B, concluding that $U \cap B$ is also open in X. Finally, one notices that
$$U = U \cap X = U \cap (A \cup B) = (U \cap A) \cup (U \cap B),$$
which implies that U is open in X — since it is a union of open sets.
- Suppose that A and B are closed subspaces of X. Since A and B are subspaces, just as in the open subspace case, there exists open sets $V_A, V_B \subseteq X$ such that $U \cap A = V_A \cap A$ and $U \cap B = V_B \cap B$. Notice that since A and B are closed in X, then both $A \setminus U$ and $B \setminus U$ are closed in X — therefore by definition both $X \setminus (A \setminus U)$ and $X \setminus (B \setminus U)$ are open in X. Furthermore, one can write
$$\begin{aligned} U &= X \setminus (X \setminus U) \\ &= X \setminus ((A \cup B) \setminus U) \\ &= X \setminus ((A \setminus U) \cup (B \setminus U)) \\ &= (X \setminus (A \setminus U)) \cap (X \setminus (B \setminus U)) \end{aligned}$$
that is, U is the intersection of open sets, thus U is an open set of X.

♮

Topological Embeddings

Definition 13.2.10 (Embedding)**.** Let $f\colon Y \rightarrowtail X$ be a continuous injective map of topological spaces. We call f an embedding when $f'\colon Y \xrightarrow{\simeq} f(Y)$, for $f'(y) := f(y)$ for all $y \in Y$, is an isomorphism.

Example 13.2.11. Let X be a topological space and S be a subspace of X. We show that $\iota_S\colon S \hookrightarrow X$ is an embedding. Notice that $\iota_S(S) = S$ and hence $\iota'_S\colon S \xrightarrow{\simeq} \iota(S) = S$ is equal to the identity map id_S, which clearly establishes an isomorphism.

Proposition 13.2.12. Let f be a continuous injective map between topological spaces. If f is either open or closed, then f is an embedding.

Proof. Let $f\colon X \to Y$. Since f is injective, the codomain restricted map $f'\colon X \xrightarrow{f} f(X)$ is a bijection and continuous by Corollary 13.2.7. Suppose f is open (resp. closed) and consider any open (resp. closed) set $U \subseteq X$, then $f(U) \subseteq f(X)$ is open (resp. closed) in $f(X)$ this shows that f' is an isomorphism. ♮

Proposition 13.2.13. A surjective embedding is an isomorphism.

Proof. Let $f\colon X \to Y$ be a surjective embedding. Notice that in a surjective map we have $f(X) = Y$, hence $f' = f\colon X \xrightarrow{\simeq} f(X) = Y$. ♮

Example 13.2.14 (2-torus surface)**.** Consider first a circle $(x-d)^2 + z^2 = r^2$, with centre at $(d, 0)$. Let ϕ be the angle going up from the x to the z axis, then we can parametrize such circle as $x = r\cos(\phi) + d,\ z = r\sin(\phi)$. Lets now consider the revolution of such circle around the z axis. If θ is the angle going up from the x axis to the y axis, we find the new equation $(\sqrt{x^2+y^2} - d)^2 + z^2 = r^2$ has a parametrization given by

$$(x, y, z) = ((r\cos(\phi) + d)\cos(\theta), (r\cos(\phi) + d)\sin(\theta), r\sin(\phi)).$$

This mapping is clearly continuous. Although not injective, we can choose any point $p = (x, y)$ and find a neighbourhood of p such that the mapping is injective.

Lemma 13.2.15 (Gluing)**.** Let X and Y be topological spaces and consider $\mathcal{U} = \{U_i\}_i$ an open cover (or finite closed cover) of X. Let $\{f_i\colon U_i \to Y\}_i$ be a collection of continuous maps such that $f_i|_{U_i \cap U_j} = f_j|_{U_i \cap U_j}$ for all indices i and j. Then there exists a unique continuous map $f\colon X \to Y$ such that $f|_{U_i} = f_i$ for all i.

Proof. Let $\mathcal{U}$ be open, then given any point $p \in X$, $f|_{U_p} = f_p$ is continuous for some neighbourhood $U_p \in \mathcal{U}$ of p, which implies that f is continuous. On the other hand, if $\mathcal{U}$ is a finite closed cover of X, then consider $C \subseteq Y$ to be any closed set, then we have $f_i^{-1}(C) = f^{-1}(V) \cap U_i$ is closed in U_i (since f_i is continuous). Since U_i is closed in X, we find that $f_i^{-1}(C)$ is closed in X (see Proposition 13.2.3). Notice that, if $|\mathcal{U}| = n$, then $f^{-1}(C) = \bigcup_{i=1}^{n} f_i^{-1}(C)$ is the finite union of closed sets in X, which implies that $f^{-1}(C)$ itself is closed. This shows us that the inverse of f maps closed sets to closed sets, which implies that f is continuous. Since $\mathcal{U}$ is a cover of X, it is clear that f is unique. ♮

13.3 Product Space

Definition 13.3.1 (Product topology)**.** Let $\{X_i\}_{i\in I}$ be a collection of sets. The product topology on the set $\prod_{i\in I} X_i$ is defined to be the initial topology such that for all $i \in I$ the projection $\pi_i\colon \prod_{i\in I} X_i \twoheadrightarrow X_i$ is continuous.

Definition 13.3.2 (Second definition of the product topology)**.** Let $\{X_i\}_{i\in I}$ be a collection of topological spaces. We define the product topology on the set $\prod_{i\in I} X_i$ to be the topology generated by the bases

$$\mathcal{B} = \left\{ \prod_{i\in I} U_i : U_i \subseteq X_i \text{ is open, and only finitely many } U_i \neq X_i \right\}.$$

Corollary 13.3.3. The definitions 13.3.1 and 13.3.2 are equivalent.

Proof. To prove the equivalence, we first show that if $\prod_{i\in I} X_i$ is endowed with the product topology, then the collection $\mathcal{B}$ is a basis for $\prod_{i\in I} X_i$. ($\Rightarrow$) Let $\prod_{i\in I} U_i \in \mathcal{B}$, then for all $i_j \in I$ we have

$$\pi_{i_j}^{-1}(U_j) = \prod_{i\in I} V_i, \text{ where } V_i = \begin{cases} X_i, & i \neq i_j \\ U_j, & \text{otherwise} \end{cases}$$

where $\pi_{i_j}^{-1}(U_j)$ is open from the hypothesis that $\prod_{i\in I} X_i$ is endowed with the product topology. Moreover, we also have $\prod_{i\in I} W_i \cap \prod_{i\in I} W_i' = \prod_{i\in I} W_i \cap W_i'$. Notice that clearly

$$\mathcal{B} = \left\{ \bigcap_{j\in J} \pi_j^{-1}(U_j) : U_j \subseteq X_j \text{ is open, and } J \subseteq I \text{ is finite} \right\}$$

and from Proposition 12.2.16 this shows that $\mathcal{B}$ is a basis for the product topology. ($\Leftarrow$) The converse is clear. ♮

Theorem 13.3.4 (Product topology universal product)**.** Let $\{X_i\}_{i\in I}$ be a collection of topological spaces and $\pi_i \colon \prod_{i\in I} X_i \twoheadrightarrow X_i$ be the projection map. For all given topological space Z and a map $f \colon Z \to \prod_{i\in I} X_i$. If $\prod_{i\in I} X_i$ is endowed with the product topology, then the map f is continuous if and only if for every $i \in I$ the map $\pi_i f \colon Z \to \prod_{i\in I} X_i$. That is, the following diagram commutes in `Top`

$$\begin{tikzcd} Z \arrow[d, "f"] \arrow[dr, bend left, "\pi_i f"] & \\ \prod_{i\in I} X_i \arrow[r, two heads, "\pi_i"] & X_i \end{tikzcd}$$

Moreover, if the space $\prod_{i\in I} X_i$ satisfies such universal property, then its topology is the product topology.

Proof. Let $\prod_{i\in I} X_i$ be endowed with the product topology. ($\Rightarrow$) Let f be a continuous map and $U \subseteq Z$ be any open set. Since π_i is continuous for all $i \in I$ from hypothesis then $\pi_i^{-1}(U) = V \subseteq \prod_{i\in I} X_i$ is open, we conclude that $(\pi_i f)^{-1}(U) = f^{-1}(\pi_i^{-1}(U)) = f^{-1}(V)$ is open, hence $\pi_i f$ is continuous. ($\Leftarrow$) Suppose $\pi_i f$ is continuous, then for all given open set $U \subseteq \prod_{i\in I} X_i$ we have $f^{-1}(\pi_i^{-1}(U)) \subseteq Z$ open. Notice that since $\pi_i^{-1}(U)$ is open in the product topology of $\prod_{i\in I} X_i$ for all $i \in I$, then $f^{-1}(V) \subseteq Z$ is open for all open set $V \subseteq \prod_{i\in I} X_i$.

For the second part of the theorem, suppose that $(\prod_{i\in I} X_i, \tau')$ be a space satisfying the property. In particular, consider a space $Z = (\prod_{i\in I} X_i, \tau')$ and a map $f = \mathrm{id}$. Then the following diagram commutes

$$\begin{tikzcd} (\prod_{i\in I} X_i, \tau') \arrow[d, "\mathrm{id}"] \arrow[dr, bend left, "\pi_i\,\mathrm{id} = \pi_i"] & \\ (\prod_{i\in I} X_i, \tau') \arrow[r, two heads, "\pi_i"] & X_i \end{tikzcd}$$

We can now assert that id is continuous since both domain and codomain have the same topology τ', hence π_i is continuous for all $i \in I$. Since π_i is continuous for all $i \in I$ for the topology τ' then we can use the Definition 13.3.1 to conclude that $\tau' \subseteq \tau$, where τ is the product topology. For the second inclusion, consider the commutative diagram

$$\begin{tikzcd} (\prod_{i\in I} X_i, \tau') \arrow[d, "\mathrm{id}'"] \arrow[dr, bend left, "\pi_i \, \mathrm{id}' = \pi_i'"] & \\ (\prod_{i\in I} X_i, \tau) \arrow[r, two heads, "\pi_i"] & X_i \end{tikzcd}$$

where $\pi_i' \colon (\prod_{i\in I} X_i, \tau') \twoheadrightarrow X_i$. We know from the previous discussion that since $(\prod_{i\in I} X_i, \tau')$ satisfies the universal property, then π_i' is continuous and hence id' is continuous. In particular, this implies that if $U \subseteq (\prod_{i\in I} X_i, \tau)$ is open, then $\mathrm{id}'^{-1}(U) = U \subseteq (\prod_{i\in I} X_i, \tau')$ is open. This implies in $\tau \subseteq \tau'$. Hence we conclude that if an object satisfies the product topology universal property, then it is endowed with the product topology. $\natural$

Lemma 13.3.5 (Projections are open maps under the product topology)**.** Let $\{X_i\}_{i\in I}$ be a collection of topological spaces. Then the projections

$$\pi_j \colon \prod_{i\in I} X_i \twoheadrightarrow X_j$$

are open maps (see Definition 12.2.9) under the product topology. Moreover, such projections are not in general closed maps.

Proof. Consider the basis $\mathcal{B}$ from Definition 13.3.2, then let $U_i \subseteq X_i$ be an open set. Then notice that (see the proof of Corollary 13.3.3)

$$\pi_j(\pi_i^{-1}(U_i)) = \begin{cases} U_i, & i = j \\ X_j, & i \neq j \end{cases}$$

which are both open sets. Moreover, we have that

$$\pi_j\left(\pi_{i_1}^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_k}^{-1}(U_{i_k})\right) = \begin{cases} X_j, & j \neq i \\ U_{i_\ell}, & j = i_\ell \text{ for some } 1 \leqslant \ell \leqslant k \end{cases}$$

is open. In particular, this implies in

$$\pi_j\left(\bigcup_{i_1,\ldots,i_k \in I} \pi^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_k}^{-1}(U_{i_k})\right) = \bigcup_{i_1,\ldots,i_k\in I} \pi_j\left(\pi^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_k}^{-1}(U_{i_k})\right)$$

which from the last assertion is the union of open sets, hence open. This implies that π_j is open (from the fact that $\mathcal{B}$ is a basis for $\prod_{i\in I} X_i$).

On the other hand, we can build a counterexample to show why the projections are not necessarily closed. Consider the product $\mathbf{R}^2$ and let the closed set $C := \{(x, y) \in \mathbf{R}^2 : xy = 1\}$, which is the hyperbola on the plane. Notice that the first projection $(x, y) \overset{\pi_1}{\longmapsto} x$ is not closed since $\pi_1(C) = \{x \in \mathbf{R} : x \neq 0\}$. $\natural$

Definition 13.3.6 (Box topology)**.** We define yet another natural topology on the set $\prod_{i\in I} X_i$: the box topology, which is generated by the basis

$$\mathcal{B}_{\text{box}} = \left\{ \prod_{i\in I} U_i : U_i \subseteq X_i \text{ is open} \right\}.$$

For the case where I is an infinite indexing set, we find that $\tau_{\text{box}} \supsetneq \tau_{\text{prod}}$ and from this we find that certainly the projections are still continuous on τ_{box}, although the box topology fails to satisfy the universal product Theorem 13.3.4.

Proposition 13.3.7 (Product of morphisms)**.** Let $\phi\colon X \to Z$ and $\psi\colon Y \to W$ be morphisms in $\texttt{Top}$. Then the product morphism $\phi \times \psi\colon X \times Y \to Z \times W$ is continuous with respect to the product topology.

Proof. Consider the morphisms $\phi\pi_X\colon X \times Y \to Z$ and $\psi \times \pi_Y\colon X \times Y \to W$, so that

$$\begin{tikzcd}
& X \times Y \arrow[d, "\phi\times\psi"] \arrow[ddl, bend right, "\phi\pi_X"'] \arrow[ddr, bend left, "\psi\pi_Y"] & \\
& Z \times W \arrow[dl, "\pi_Z"] \arrow[dr, "\pi_W"'] & \\
Z & & W
\end{tikzcd}$$

commutes. Therefore by the universal property of the product topology we conclude that $\phi \times \psi$ is continuous. ♮

13.4 Coproduct Space

Definition 13.4.1 (First definition)**.** Let $\{X_i\}_{i\in I}$ be any collection of topological spaces and consider the disjoint union $\coprod_{i\in I} X_i = \bigcup_{i\in I} X_i \times \{i\}$. We define the coproduct topology on $\coprod_{i\in I} X_i$ via the following property. A set $U \subseteq \coprod_{i\in I} X_i$ is open if and only if it is of the form $U = \coprod_{i\in I} U_i$, where each $U_i \subseteq X_i$ is open.

Definition 13.4.2 (Coproduct topology)**.** Let $\{X_i\}_{i\in I}$ be a collection of topological spaces. The coproduct topology on the set $\coprod_{i\in I} X_i$ is defined to be the final topology such that for all $j \in I$ we have that the inclusions $\iota_j\colon X_j \rightarrowtail \coprod_{i\in I} X_i$ are continuous.

Theorem 13.4.3 (Coproduct topology universal property)**.** Let $\{X_i\}_{i\in I}$ be a collection of topological spaces and let $\iota_j\colon X_j \to \coprod_{i\in I} X_i$ be the jth inclusion. The coproduct topology on $\coprod_{i\in I} X_i$ satisfies the following property. Let Z be a topological space and a collection of continuous maps $\{f_i\colon X_i \to Z\}_{i\in I}$. Then there exists a unique map $f\colon \coprod_{i\in I} X_i \to Z$ such that the following diagram commutes in the category $\texttt{Top}$ for all

$j \in I$:

$$\begin{tikzcd} X_j \arrow[r, "\iota_j"] \arrow[dr, bend right, "f_j"'] & \coprod_{i\in I} X_i \arrow[d, dashed, "f"] \\ & Z \end{tikzcd}$$

On the other hand, if $(\coprod_{i\in I} X_i, \tau')$ satisfies such property, then τ' is the coproduct topology.

Proof. (Uniqueness) Suppose that $f\iota_j = f_j$, then from definition $f_j(x) = f(\iota(x)) = f(x, j)$ for each $j \in I$, which is clearly unique. (Existence) Suppose now that f_j is continuous for all $j \in I$. Let $U \subseteq Z$ be any open set. Notice that since $f_j^{-1} = \iota_j^{-1} f^{-1}$ then $f_j^{-1}(U) = \iota_j^{-1}(f^{-1}(U))$. Now, if $f^{-1}(U) \subseteq \coprod_{i\in I} X_i$ is closed then its preimage under ι_j would be closed (from the continuity of ι_j), hence $f^{-1}(U)$ is open, which implies in f continuous.

Suppose that $(\coprod_{i\in I} X_i, \tau')$ satisfies the universal property and denote by τ the coproduct topology. Then in particular we have

$$\begin{tikzcd} X_j \arrow[r, "\iota_j"] \arrow[dr, bend right, "\iota_j"'] & (\coprod_{i\in I} X_i, \tau') \arrow[d, dashed, "g"] \\ & (\coprod_{i\in I} X_i, \tau) \end{tikzcd} \qquad \begin{tikzcd} X_j \arrow[r, "\iota_j"] \arrow[dr, bend right, "\iota_j"'] & (\coprod_{i\in I} X_i, \tau) \arrow[d, dashed, "f"] \\ & (\coprod_{i\in I} X_i, \tau') \end{tikzcd}$$

and hence f and g are both identities on $\mathrm{End}(\coprod_{i\in I} X_i)$. If $U \subseteq (\coprod_{i\in I} X_i, \tau)$ is open, then $g^{-1}(U) = U \subseteq (\coprod_{i\in I} X_i, \tau')$ is also open. On the other hand, from the second diagram, if $V \subseteq (\coprod_{i\in I} X_i, \tau')$ is open, then $f^{-1}(V) = V \subseteq (\coprod_{i\in I} X_i, \tau)$ is open. This concludes that $\tau' = \tau$ and hence the coproduct topology is unique. ♮

Proposition 13.4.4. Let $\{X_i\}_{i\in I}$ be a collection of topological spaces. The following are properties of the coproduct space.

(a) A subset $C \subseteq \coprod_{i\in I} X_i$ is closed if and only if for all X_i, we have that $C \cap X_i$ is closed.

(b) The canonical injection $\iota_j\colon X_j \to \coprod_{i\in I} X_i$ is a topological embedding and an open and closed map.

(c) If X_i is Hausdorff for all $i \in I$, then $\coprod_{i\in I} X_i$ is Hausdorff.

(d) If X_i is first countable for all $i \in I$, then $\coprod_{i\in I} X_i$ is first countable.

(e) If X_i is second countable for all $i \in I$ and I is a countable indexing set, then $\coprod_{i\in I} X_i$ is second countable.

Properties for the coproduct top

13.5 Quotient Space

Quotient Topology

The motivation for the construction of the quotient topology is the study of surjective set-functions $\pi\colon X \twoheadrightarrow S$ between topological spaces X and sets S, which induce an equivalence relation on the initial topological space by means of arranging the points of X in classes where $x \sim y$ for $x, y \in X$ if and only if $\pi(x) = \pi(y)$, that is, x and y are common points of a fibre $\pi^{-1}(s)$ for some $s \in S$.

The geometrically appealing version of such construction would be the idea of gluing every element $x \in \pi^{-1}(s)$ into a unique point — this way we loose some initial information about the topological space X. However, we would like to make this gluing process compatible with the theory of topology so far constructed. To achieve that, one may make the point that we should make S into a topological space for which its topology makes π a continuous map. The first idea that may come to mind is that we can simply force the continuity of π by defining a topology τ on S so that $U \subseteq S$ is an open set if and only if $\pi^{-1}(U)$ is open — which is exactly the definition of a continuous map. Lets give this a formal definition.

Definition 13.5.1. Let X be a topological space and S be any set. Let also $\pi\colon X \twoheadrightarrow S$ be a set-function. The quotient topology on S induced by the map π is the final topology such that π is a continuous map.

Proposition 13.5.2. The quotient topology is a topology.

Proof. Let $X/{\sim}$ be a space with the topology induced by the map $\pi\colon X \twoheadrightarrow X/{\sim}$. Consider an arbitrary collection $\{U_j\}_j$ of open sets of $X/{\sim}$. Notice that since $\pi^{-1}(\bigcup_j U_j) = \bigcup_j \pi^{-1}(U_j)$ then since $\pi^{-1}(U_j)$ is open for all index j, we conclude that $\bigcup_j \pi^{-1}(U_j) \subseteq X$ is open, hence the set $\bigcup_j U_j \subseteq X/{\sim}$ is necessarily open — since π is continuous. Let now $A, B \subseteq X/{\sim}$ be any open sets, then $\pi^{-1}(A \cap B) = \pi^{-1}(A) \cap \pi^{-1}(B) \subseteq X$, which is open and therefore $A \cap B$ is open in $X/{\sim}$. We conclude that the quotient topology indeed satisfies the properties of a topology. ♮

In other terms, let T_S be the collection of all topologies on S such that π is a continuous map. The above definition simply says that the quotient topology τ on the set S is the intersection $\tau = \bigcap_{T \in T_S} T$. Even better than that definition is the fact that we can determine the quotient topology by the following universal property.

Theorem 13.5.3 (Universal property of the quotient topology)**.** Let X be a topological space and S be any set, together with a surjective set-function $\pi\colon X \twoheadrightarrow S$. The quotient topology τ on S induced by π is such that, for all topological spaces Z and every morphism $g\colon X \to Z$ for which all $x, y \in X$ such that $\pi(x) = \pi(y)$ then $g(x) = g(y)$ — i.e. g is constant on the fibres of π — there exists a unique morphism $f\colon (S, \tau) \to Z$

such that the diagram

$$
\begin{tikzcd}
X \arrow[r, "g"] \arrow[d, two heads, "\pi"'] & Z \\
S \arrow[ur, dashed, bend right, "f"'] &
\end{tikzcd}
$$

commutes in $\mathtt{Top}$. Moreover, if τ' is a topology on S such that the diagram commutes, then necessarily $\tau' = \tau$.

Proof. Since π is surjective, we can completely define a map $f\colon S \to Z$ sending $s \mapsto g(x)$ such that $x \in \pi^{-1}(s)$, which is well defined because, for all $x, y \in \pi^{-1}(s)$, we have $g(x) = g(y)$, so that the image of each $s \in S$ under the map f is uniquely defined. We now show that f is, in fact, continuous — and hence a morphism. Let $U \subseteq Z$ be any open set of Z, then, since $g^{-1}(U)$ is open and $(f\pi)^{-1}(U) = \pi^{-1}f^{-1}(U) = g^{-1}(U)$, $f^{-1}(U)$ cannot be a closed set — in fact, it needs to be an open set, because $\pi^{-1}(f^{-1}(U)) \subseteq X$ must be open, since π is continuous. It follows that f is continuous and thus the said morphism indeed exists.

For the uniqueness, let f and f' be two morphisms such that $f\pi = g$ and $f'\pi = g$. Let $s \in S$ be any point. Since π is surjective, there exists $x \in X$ for which $x \in \pi^{-1}(s)$, therefore, $f\pi(x) = f(s) = f'\pi(x) = f'(s)$ for every element of their domain — hence $f = f'$.

Suppose now that both (S, τ) and (S, τ') satisfy the universal property, that is, the following diagrams commutes for unique morphisms f and h

$$
\begin{tikzcd}
 & Z & \\
(S,\tau) \arrow[ur, dashed, bend left, "f"] & X \arrow[l, two heads, "\pi"] \arrow[r, two heads, "\pi"'] \arrow[u, "g"] & (S,\tau') \arrow[ul, dashed, bend right, "f'"']
\end{tikzcd}
$$

Since the diagram commutes for all Z, let $Z = (S, \tau)$, and consider the map $f' = \mathrm{id}'\colon (S, \tau') \to (S, \tau)$. Then, given any $U \in \tau$ we find that since id' is continuous that $\mathrm{id}'^{-1}(U) = U \subseteq (S, \tau')$ is open, hence $\tau \subseteq \tau'$. Analogously, let $Z = (S, \tau')$ and consider $f = \mathrm{id}\colon (S, \tau) \to (S, \tau')$. Let $U' \in \tau'$ then from the continuity of id we find that $\mathrm{id}^{-1}(U') = U' \subseteq (S, \tau)$ is open, therefore $\tau' \subseteq \tau$. Thus indeed $\tau = \tau'$ as wanted. ♮

Proposition 13.5.4 (Quotient topology as a coequalizer)**.** Let X be a topological space and $f\colon X \twoheadrightarrow S$ be a surjective map, where S is some set. Define the equivalence relation set

$$R \coloneqq \{(x, y) \in X \times X : f(x) = f(y)\},$$

together with two maps $r_1, r_2\colon R \rightrightarrows X$ given by the following commutative diagram

$$
\begin{tikzcd}
R \arrow[r, hook] \arrow[rr, bend left, "r_1"] \arrow[rr, bend right, "r_2"'] & X \times X \arrow[r, shift left, "\pi_1"] \arrow[r, shift right, "\pi_2"'] & X
\end{tikzcd}
$$

The *quotient topology* is exactly the topology that makes S into the *coequalizer* of r_1 and r_2.

Proof. Let Y be any space and $h\colon X \to Y$ be a continuous map such that $hr_1 = hr_2$. Let $x, x' \in X$ be any two points such that $f(x) = f(y)$, then from construction $x, x' \in R$. Then $h(x) = hr_1(x, x') = hr_2(x, x') = h(x')$. Now by means of the universal property Theorem 13.5.3, if τ is the quotient topology on S induced by f, we conclude that there exists a unique continuous map $g\colon (S, \tau) \to Y$ such that the following diagram commutes

$$\begin{tikzcd} Y & & \\ (S,\tau) \arrow[u, dashed, "g"] & X \arrow[l, two heads, "f"] \arrow[ul, bend right, "h"'] & R \arrow[l, shift right, "r_1"'] \arrow[l, shift left, "r_2"] \end{tikzcd}$$

Therefore $(S, \tau) = \operatorname{coeq}(r_1, r_2)$, with associated morphism f. ♮

To ease the way in which we refer to quotients and surjective morphisms that induce quotients between topological spaces, we define the following terminology.

Definition 13.5.5 (Quotient morphism)**.** A surjective morphism $\pi\colon X \to Y$ of topological spaces X and Y is said to be a *quotient morphism* (or quotient map) if π induces the universal property of quotients — in other words, open sets of Y are exactly those that have open preimage on π, that is, $V \subseteq Y$ is open if and only if $\pi^{-1}(V) \subseteq X$ is open.

Theorem 13.5.6 (Quotient descent)**.** Let $q\colon X \to Y$ be a quotient map between topological spaces. For any space Z together with a morphism $f\colon X \to Z$ that is constant on the fibres of q — that is, $q(x_1) = q(x_2)$ implies $f(x_1) = f(x_2)$ — there exists a *unique* morphism $f_*\colon Y \to Z$ such that the following diagram commutes

$$\begin{tikzcd} X \arrow[d, two heads, "q"'] \arrow[dr, bend left, "f"] & \\ Y \arrow[r, dashed, "f_*"'] & Z \end{tikzcd}$$

Proof. Since q is surjective, for every $y \in Y$ there exists $x \in X$ such that $q(x) = y$, hence we define $f_*(y) \coloneqq f_*(x)$ for every $x \in q^{-1}(y)$. For the uniqueness, since f is constant on the fibres of q then f_* is fully defined by f — thus unique. From the universal property we obtain that f_* is continuous. ♮

Some Examples And Applications

Many important spaces can be obtained with the inclusion of the quotient topology to our toolkit, I'll now briefly discuss some of those, which will most probably come up further into this text.

Example 13.5.7 (Projective space)**.** Let $\sim$ be the equivalence relation for which $x \sim y$ if and only if $x = \gamma y$, where $x, y \in \mathbf{R}^{n+1}$ $\{0\}$ and $\gamma \in \mathbf{R}$. We define the n-dimensional real projective space as the quotient $(\mathbf{R}^{n+1} \setminus \{0\})/\sim$, which is denoted by $\mathbf{RP}^n$.

Definition 13.5.8 (Cone)**.** Let X be any topological space and I be the standard interval. The topological space $X \times I$ is known as the cylinder on X. Via a quotient operation, we can collapse regions of this cylinder. For instance, we can create a cone by collapsing one of the sides of the cylinder, such as $(X \times I)/(X \times \{1\})$. Such object is denoted $\operatorname{Cone} X$, the cone on X.

Definition 13.5.9 (Suspension)**.** The *suspension* of a topological space is an endofunctor $\mathrm{S}\colon \mathtt{Top} \to \mathtt{Top}$ mapping each space X to the quotient space

$$\mathrm{S}\,X \coloneqq (X \times I)/(X \times \{0, 1\})$$

and for each morphism $f\colon X \to Y$ we have the naturally induced morphism

$$\mathrm{S}\,f\colon \mathrm{S}\,X \to \mathrm{S}\,Y \ \text{ mapping } \ [x, t] \mapsto [f(x), t].$$

Example 13.5.10. Let J be an indexing set and $\{X_j\}_{j\in J}$ be a collection of non-empty topological spaces. For each $j \in J$, choose any $p_j \in X_j$ as a base point. We define the wedge sum of the collection $\{X_j\}_{j\in J}$ with respect to the base points $\{p_j\}_{j\in J}$ as the topological space

$$\bigvee_{j\in J} X_j = \coprod_{j\in J} X_j/\{p_j\}_{j\in J}$$

An interesting fact about wedge sums of topological spaces preserve the Hausdorff property if every component is Hausdorff.

Proposition 13.5.11. Let $\{X_j\}_{j\in J}$ be an indexed collection of Hausdorff topological spaces. Then the wedge sum $\bigvee_{j\in J} X_j$ with respect to any choice of base points is Hausdorff.

Proof. Let $\{p_j \in X_j\}_{j\in J}$ be any choice of base points for the given collection. Let $x, y \in \bigvee_{j\in J} X_j$ be any distinct points in the wedge sum space. If $x, y \in X_j$ for some $j \in J$, then it is clear that there exists non-intersecting neighbourhoods of x and y on X_j — of which we can take their intersection with the disjoint union $\coprod_{j\in J} X_j$ and the proposition will hold for $\bigvee_{j\in J} X_j$. On the other hand, if $i, j \in J$ are distinct indices and $x \in X_i$ while $y \in X_j$, then there exists $U_x \subseteq X_i$ and $U_y \subseteq X_j$ neighbours of x and y, respectively, such that $U_x \cap X_j = \varnothing$ and $U_y \cap X_i = \varnothing$, which in particular imply in $U_x \cap U_y = \varnothing$. For our end, we just need to consider the neighbourhoods $U'_x = U_x \cap \coprod_{j\in J} X_j$ and $U'_y = U_y \cap \coprod_{j\in J} X_j$ so that $U'_x = U'_y$. ♮

Proposition 13.5.12. Let X be a second countable space and $M = X/{\sim}$ be a quotient. If M is locally Euclidean, then M is second countable.

Proof. Let $\pi\colon X \twoheadrightarrow M$ be the quotient map that induces the equivalence relation $\sim$ in X. If we assume that M is locally euclidean, we can let $\mathcal{C}$ be a cover of M composed of coordinate balls. Since π is surjective, $\mathcal{H} \coloneqq \{\pi^{-1}(U) : U \in \mathcal{C}\}$ is a cover for the space X. Moreover, since X is second countable, any cover of X contains a countable subcover — in particular, let $\mathcal{U} \subseteq \mathcal{H}$ be a countable subcover. Define the countable set

$C' = \{U \in C : \pi^{-1}(U) \in \mathcal{U}\}$. Since $\mathcal{U}$ covers X and π is surjective, it follows that $C' \subseteq C$ is a countable subcover of M composed of coordinate balls — that is, M is Lindelöf. Better than that, since coordinate balls are second countable (see Proposition 12.7.6) we can apply Corollary 12.5.12 to see that M is second countable. ♮

Corollary 13.5.13 (Manifold from a quotient)**.** In the context of the preceding proposition, if M is both locally Euclidean and Hausdorff, then M is a topological manifold.

Hausdorffness & Quotient Spaces

Remark 13.5.14. The quotient topology *doesn't preserve* Hausdorffness.

As an example, take the space $\mathbf{R} \times \{-1, 1\}$ with the usual topology and define the equivalence relation $(x, 1) \sim (x, -1)$ if and only if $x \neq 0$. The resulting space $(\mathbf{R} \times \{-1, 1\})/{\sim}$ is not Hausdorff: consider the sequence $(1/n, 1)_{n \in \mathbf{Z}_{>0}}$, we have that $1/n \to 0$ but since $1/n$ does never assume the value of zero, for all $n \in \mathbf{Z}_{>0}$ we have $(1/n, 1) \sim (1/n, -1)$ — thus the has two distinct limits $(0, 1)$ from one side and $(0, -1)$ from the other, thus $(\mathbf{R} \times \{-1, 1\})/{\sim}$ isn't Hausdorff.

Proposition 13.5.15 (Hausdorff from open quotients)**.** Let $\pi\colon X \twoheadrightarrow Y$ be a surjective morphism of topological spaces X and Y. Then Y is Hausdorff if and only if the collection of pairs of points with common fibre, $C := \{(p, q) \in X \times X : \pi(p) = \pi(q)\}$, is closed in $X \times X$.

Proof. Let Y be Hausdorff, then, given any $(p, q) \in X \setminus C$, there are neighbourhoods $V_p, V_q \subseteq Y$ of $\pi(p)$ and $\pi(q)$, respectively, such that $V_p \cap V_q = \varnothing$. Since these neighbourhoods are disjoint, then in particular the collection fibres $\pi^{-1}(V_p) \times \pi^{-1}(V_q)$ is contained in $X \times X \setminus C$, that is, $X \times X \setminus C$ is open — hence C is closed.

Let C be closed, then given any distinct points $a, b \in Y$, the surjectivity of π implies that there exists $p, q \in X$ such that $\pi(p) = a$ and $\pi(q) = b$ — in particular $(p, q) \in X \times X \setminus C$ and since C is closed, there exists a neighbourhood $U_p \times U_q \subseteq X \times X$ of (p, q) such that $U_p \times U_q \subseteq X \times X \setminus C$, that is, $\pi(U_p), \pi(U_q) \subseteq Y$ are non intersecting open sets (from the fact that π is open) that are neighbourhoods of a and b, respectively — thus Y is Hausdorff. ♮

Proposition 13.5.16. Every topological space is the quotient of a Hausdorff space.

Proof. Let X be any topological space. Consider the product of real lines $P := \prod_{x \in X} \mathbf{R}$ under the product topology and let Y be the subspace of P given by all points with one, and only one, rational coordinate — such space Y is therefore Hausdorff. Furthermore, define a collection $\{Y_x\}_{x \in X}$ to consist of subsets $Y_x \subseteq Y$ given by the set of points whose rational coordinate has index x — hence Y_x is dense in Y. It should be noted that the above construction is not necessarily unique, so the reader may try to construct a different Hausdorff space Y and a collection of dense sets $\{Y_x\}_{x \in X}$ in Y.

Define now the subspace of $X \times Y$ given by $Z := \bigcup_{x \in X} x \times Y_x$ — endowed with the subspace topology. Let $\sim$ be the equivalence relation on Z given by $(x_1, y_1) \sim (x_2, y_2)$ if and only if $x_1 = x_2 = x$ for some $x \in X$ and $y_1, y_2 \in Y_x$. Define a map $\phi\colon Z/{\sim} \to X$

sending $x \times Y_x \mapsto x$ — which is clearly both surjective and injective. We now show that ϕ is a topological isomorphism.

Let $V \subseteq X$ be any open set, then $\phi^{-1}(V) = \{x \times Y_x : x \in V\}$ — which in turn is open in $Z/\sim$. On the other hand, let $U := \{x \times Y_x : x \in U'\}$ be any open set in $Z/\sim$ given by some indexing set of points $U' \subseteq X$ — our goal will be to prove that U' is open, so that the inverse of ϕ is continuous. Consider $x_0 \in U'$ to be any point and likewise $(x_0, y_0) \in x_0 \times Y_{x_0}$. By the fact that $\bigcup_{x \in U'} x \times Y_x \subseteq Z$ is open, we are able to find neighbourhoods $O_X(x_0) \subseteq X$ and $O_Y(y_0) \subseteq Y$ — of x_0 and y_0, respectively — such that $N := Z \cap (O_X(x_0) \times O_Y(y_0))$ is a neighbourhood of (x_0, y_0) in $Z/\sim$. If $x' \in O_X(x_0)$ is any point, then by the fact that $Y_{x'}$ is dense in Y, by Proposition 12.1.34, the sets $x' \times Y_{x'}$ and N have a non-empty intersection and thus $x' \times Y_{x'}$ also intersects $\bigcup_{x \in U'} x \times Y_x$. Therefore $x' \times Y_{x'} \in U$ and hence $x' \in U'$ — which implies that U' is open. ♮

Quotient Morphisms In More Depth

So far we've been studying the construction of quotients out of surjective set-functions, but what about being able to classifying a surjective morphisms between given topological spaces as inducing the universal property of the quotient space? This will be our goal with this subsection — identifying quotient morphisms. For that end, we shall profit from the main idea behind quotients: fibres. For that, we define a set given by fibres of f as being saturated.

Definition 13.5.17 (Saturated set)**.** Let $f\colon X \to Y$ be a set-function. We say that a set $U \subseteq X$ is saturated with respect to f if there exists $V \subseteq Y$ such that $U = f^{-1}(V)$.

Proposition 13.5.18 (Equivalences for saturated sets)**.** Let $f\colon X \to Y$ be a set-function and $U \subseteq X$ be any subset. The following propositions are equivalent

(a) The set U is saturated with respect to f.

(b) $U = f^{-1}(f(U))$.

(c) Let $p \in U$ be any point, U contains every element $x \in X$ with common fibre to p — that is, $f(x) = f(p)$ implies $x \in U$.

Proof. (c) ⇒ (b): Suppose U satisfies proposition (c), it is clear that $U \subseteq f^{-1}(f(U))$, on the other hand, given $x \in f^{-1}(f(U))$, it follows that x has a common fibre with some point of U, which implies that $x \in U$. (b) ⇒ (a): Trivial from the definition. (a) ⇒ (c): Let $V \subseteq Y$ be such that $U = f^{-1}(V)$, then, given any $p \in f^{-1}(V)$, it is clear that $f(p) \in V$, hence every point $x \in X$ such that $f(x) = f(p) \in V$ then $x \in U$, which finishes the equivalence chain. ♮

Proposition 13.5.19 (Classification of surjective morphisms)**.** Let $\pi\colon X \twoheadrightarrow Y$ be a surjective morphism of topological spaces. The map π is a quotient morphism — that is, induces the universal property of quotients for X and Y — if and only if every saturated open (or closed) set of X has an open (or closed) image in Y.

Proof. Let π be any surjective morphism taking saturated open sets to open images. Let $V \subseteq Y$ be an open set. Since π is surjective and continuous, then $\pi^{-1}(V) \subseteq X$ is open. On the other hand, let $V \subseteq Y$ be any set of Y (not necessarily open), such that $\pi^{-1}(V) := U \subseteq X$ is open. This implies directly that U is saturated with respect to π and from our initial hypothesis, $\pi(U) = V \subseteq Y$ is open. Thus π is a quotient morphism.

For the contrary, let π be a quotient morphism. Then, given any $U \subseteq X$ open set, saturated with respect to π, define $V \subseteq Y$ such that $U = \pi^{-1}(V)$. Since π is a quotient morphism, it follows that V is necessarily open in Y, thus $\pi(U) = V$ is open.

The proof for the closed set case is completely analogous. ♮

Proposition 13.5.20 (Properties of quotient morphisms)**.** The following properties pertain to quotient morphisms between topological spaces.

(a) The composition of quotient morphisms is a quotient morphism.

(b) Injective quotient morphisms are isomorphisms.

(c) Let $\pi \colon X \twoheadrightarrow Y$ be a quotient morphism. Then, $C \subseteq Y$ is closed if and only if $\pi^{-1}(C) \subseteq X$ is closed.

(d) Let $\pi \colon X \twoheadrightarrow Y$ be a quotient morphism and $U \subseteq X$ be any saturated set (open or closed) with respect to π. Then, the restriction $\pi|_U \colon U \twoheadrightarrow \pi(U)$ is a quotient map.

(e) Let J be an indexing set and $\{\pi_j \colon X_j \twoheadrightarrow Y_j\}_{j\in J}$ be an indexed collection of quotient morphisms. The map $\pi \colon \coprod_{j\in J} X_j \twoheadrightarrow \coprod_{j\in J} Y_j$ defined by the restrictions $\pi(x_j) = \pi_j(x_j)$ for every $x_j \in X_j \cap \coprod_{j\in J} X_j$ is a quotient map.

Proof. (a) Let $\pi \colon X \twoheadrightarrow Y$ and $\pi' \colon Y \twoheadrightarrow Z$ be quotient morphisms, and consider the map $\pi'\pi \colon X \twoheadrightarrow Z$, which is clearly surjective. From hypothesis a subset $U \subseteq Z$ is open if and only if $\pi'^{-1}(U)$ is open, moreover, $\pi'^{-1}(U) \subseteq Y$ is open if and only if $\pi^{-1}(\pi'^{-1}(U)) \subseteq X$ is open — the proposition follows.

(b) If the quotient morphism $\pi \colon X \twoheadrightarrow Y$ is injective, then π is a bijection. Let $U \subseteq X$ be any open set, since π is bijective, there exists a set $V \subseteq Y$ such that $\pi^{-1}(V) = U$ — moreover, such set must be open on Y from the quotient topology. This shows that $\pi(U) = V$ is open and hence π is a topological isomorphism.

(c) Let $C \subseteq Y$ be any set, notice that $Y \smallsetminus C$ is open if and only if $\pi^{-1}(Y \smallsetminus C) = \pi^{-1}(Y) \smallsetminus \pi^{-1}(C) \subseteq X$ is open — thus the proposition follows.

(d) Let $V \subseteq \pi(U)$ be any set. Since π is a quotient morphism, V is open if and only if $\pi^{-1}(V) \subseteq U \subseteq X$ — moreover, since U is saturated, the whole set can have its subsets classified by $\pi|_U$ into open or closed sets, thus $\pi|_U$ is a quotient map.

(e) Let $U \subseteq \coprod_{j\in J} Y_j$ be any set and let $j_0 \in J$ be such that $U \subseteq Y_{j_0}$. From the mappings, U is open if and only if $\pi_{j_0}^{-1}(U) \subseteq X_{j_0}$ is open — but since $\pi^{-1}(U) = \pi_{j_0}^{-1}(U)$, the proposition follows.

♮

Example 13.5.21 (Cones)**.** An application of the last proposition takes us back to Definition 13.5.8, where we defined the cone $\operatorname{Cone} X$ of a topological space X as

$(X \times I)/(X \times \{1\})$ — the collapse the top of the cylinder. Notice that, given any point $(x,t) \in (X \times I) \setminus (X \times \{1\})$, there exists a neighbourhood $U \subseteq (X \times I) \setminus (X \times \{1\})$ of (x,t), therefore $X \times \{1\}$ is closed in $X \times I$. If we consider the quotient morphism of the cone $\pi\colon X \times I \twoheadrightarrow \operatorname{Cone} X$, the restriction $\pi|_{X \times \{0\}}\colon X \times \{0\} \twoheadrightarrow \pi(X \times \{0\})$ is also a quotient map — moreover, such quotient map is injective, thus $\pi|_{X \times \{0\}}$ is an isomorphism. Therefore we have a sequence of isomorphisms

$$X \xrightarrow{\simeq} X \times \{0\} \xrightarrow{\simeq} \pi(X \times \{0\}) \subseteq \operatorname{Cone} X,$$

thus we can identify X as a subspace of $\operatorname{Cone} X$.

The following is a *sufficient*, but *not* necessary condition for a quotient morphism.

Proposition 13.5.22. Let $\pi\colon X \twoheadrightarrow Y$ be a surjective topological morphism. If π is either open or closed, then π is a quotient morphism.

Proof. If π is open (respectively, closed), in particular we have that saturated sets open (respectively, closed) of X are mapped to open (respectively, closed) sets of Y, hence π is a quotient morphism. ♮

Proposition 13.5.23. Let $f\colon X \to Y$ be a topological morphism that is either open or closed. The following are properties hold:

(a) If f is injective, it is a topological *embedding*.

(b) If f is surjective, it is a *quotient map*.

(c) If f is bijective, it is an *isomorphism*.

Proof. We work out the proof for the case where f is open, the closed case is equivalent:

1. Consider the restriction of the codomain $f'\colon X \to f(X)$ — which is certainly surjective, thus bijective. Let $g\colon f(X) \to X$ denote the inverse of f. Notice that since f is open then any $U \subseteq X$ open implies in $f(U) \subseteq Y$ also open, therefore $g(U) = f(U)$ is open and hence g is continuous.
2. Let $U \subseteq Y$ be any set. Since f is surjective, there exists $V \subseteq X$ such that $f(V) = U$. Moreover, since f is open, U can only be open in Y if its preimage $f^{-1}(U) = V$ is open in X — which implies that f is a quotient map.
3. If f is bijective, then by the fist item f is an embedding, that is, it yields an isomorphism $X \simeq f(X)$. Moreover since f is also surjective then $f(X) = Y$ and thus f is an isomorphism $X \xrightarrow{\simeq} Y$.

♮

13.6 Attaching Space

Definition 13.6.1 (Attaching space)**.** Let X and Y be topological spaces, and consider a subspace $A \subseteq X$ together with a continuous map $f\colon A \to X$. The *attaching space of X and Y along f* is defined to be the pushout of the canonical inclusion $\iota\colon A \hookrightarrow X$ and f, that is

$$\begin{tikzcd} A \arrow[r, hook, "\iota"] \arrow[d, "f"'] & X \arrow[d] \\ Y \arrow[r] & X \cup_f Y \arrow[ul, phantom, "\ulcorner", very near start] \end{tikzcd}$$

Corollary 13.6.2. In the notation of Definition 13.6.1, the attaching space of X and Y along f is given by the quotient space

$$X \cup_f Y \simeq (X \amalg Y)/\sim,$$

where $\sim$ is the smallest equivalence relation on $X \amalg Y$ such that $x \sim f(x)$ for all $x \in A$.

Proof. Let Z be any space together with two continuous maps $p\colon X \to Z$ and $q\colon Y \to Z$ such that $p\iota = qf$ — that is, for every $x \in A$ we have $p(x) = q(f(x))$. We define a map $\phi\colon (X \amalg Y)/\sim \to Z$ by $[(x, X)] \mapsto p(x)$ and $[(y, Y)] \mapsto q(y)$. Indeed, the image of a class point under ϕ does not depend on the representative since, for any $x \in A$, we have $[(x, X)] \sim [(f(x), Y)]$, but $p(x) = q(f(x))$. Also, since $[X] := \{[(x, X)]\}_{x \in X}$ and $[Y] := \{[(y, Y)]\}_{y \in Y}$ are open subspaces covering $(X \amalg Y)/\sim$, since $\phi|_{[X]} = p$ and $\phi|_{[Y]} = q$ are continuous maps, by Proposition 13.2.9 we conclude that ϕ is continuous.

If we consider the inclusions $\iota_X\colon X \hookrightarrow (X \amalg Y)/\sim$ and $\iota_Y\colon Y \hookrightarrow (X \amalg Y)/\sim$, one has that $\phi\iota_X = p$ and $\phi\iota_Y = q$. On the other hand, since these inclusions are monomorphisms in `Top`, we conclude that ϕ is the only continuous map making the following diagram commutative in `Top`

$$\begin{tikzcd} A \arrow[r, hook, "\iota"] \arrow[d, "f"'] & X \arrow[d, hook, "\iota_X"] \arrow[ddr, bend left, "p"] & \\ Y \arrow[r, hook, "\iota_Y"] \arrow[drr, bend right, "q"'] & (X \amalg Y)/\sim \arrow[dr, dashed, "\phi"] & \\ & & Z \end{tikzcd}$$

We conclude that $(X \amalg Y)/\sim$ is the pushout of f with ι and since pushouts are unique up to isomorphism, then $(X \amalg Y)/\sim \simeq X \cup_f Y$. ♮

Proposition 13.6.3 (Attaching space properties)**.** Let $A \subseteq Y$ be a *closed* subspace, $f\colon A \to X$ be a continuous map, and $\pi\colon X \amalg Y \twoheadrightarrow X \cup_f Y$ be the canonical projection to the attaching space along f. The following properties hold:

(a) The restriction $\pi|_X$ is a *topological embedding*, and its image $\pi(X)$ is a *closed subspace* of $X \cup_f Y$.

(b) The restriction $\pi|_{Y \setminus A}$ is a *topological embedding*, and its image $\pi(Y \setminus A)$ is an *open subspace* of $X \cup_f Y$.

(c) The attaching space $X \cup_f Y$ is isomorphic to the disjoint union $\pi(X) \amalg \pi(Y \setminus A)$.

Proof. (a)

Prove attaching space properties and continue on attaching manifolds along the boundary.

$\natural$

Chapter 14

Connectivity and Compactness

14.1 Connected Spaces

Connectedness

Definition 14.1.1 (Connected space)**.** Let X be a topological space. We say that X is *connected* if and only if one of the following conditions hold:

(a) The space X cannot be expressed as the union of two disjoint non-empty open sets.

(b) Every morphism $f: X \to \{0, 1\}$ is constant—where $\{0, 1\}$ is a space endowed with the discrete topology.

Corollary 14.1.2. The two conditions in Definition 14.1.1 are equivalent.

Proof. Assume X satisfies the condition (a), and let $x \in X$ be any point. Suppose $f(x) = 0$ (or $f(x) = 1$, we do not loose generality by choosing a point of the domain), and consider $f^{-1}(0) \subseteq X$, which must be open from the continuity of f, thus $f^{-1}(0)$ is a neighbourhood of x in X. If we suppose, for the sake of contradiction, that there exists $y \in X$ such that $f(y) = 1$, then $f^{-1}(1)$ is also open and is a neighbourhood of y—notice however that $f^{-1}(0) \cap f^{-1}(1) = \varnothing$, thus we arrive at a contradiction, there must be no point y with image different than zero.

We prove the counter positive: not (b) implies not (a). Suppose $f: X \to \{0, 1\}$ is not constant, so that there exists two points $x, y \in X$ such that $f(x) = 0$ and $f(y) = 1$ (and therefore f is surjective), notice however that, since f is continous, ♮

Definition 14.1.3 (Connected components)**.** Let X be a topological space. Define an equivalence relation $\sim$ by, given $x, y \in X$, we have $x \sim y$ if and only if there exists a connected subspace of X containing both x and y. The collection of equivalence classes $X/\sim$ is called *connected components* of X.

Definition 14.1.4 (Totally disconnected)**.** A space is said to be totally disconnected if the only connected subsets are singletons.

Example 14.1.5. The set of rational numbers $\mathbf{Q}$ is totally disconnected.

Path Connectedness

Notation 14.1.6 (Standard interval)**.** From now on, when talking about *paths* and *homotopies*, we shall reserve the symbol I to denote the *standard topological interval*, which is defined by
$$I := [0,1] \hookrightarrow \mathbf{R}.$$

Definition 14.1.7 (Paths & loops)**.** A *path* in a topological space X is any continuous map $\gamma\colon I \to X$. A *loop* in X is a continuous map $\ell\colon I \to X$ such that $\ell(0) = \ell(1)$.

Definition 14.1.8 (Path connected space)**.** A topological space X is said to be *path connected* if and only if for all $x, y \in X$ there exists a *path* connecting x and y.

Proposition 14.1.9 (Path connected equivalence relation)**.** There exists an equivalence relation $\sim$ on the topological X defined by: given $x, y \in X$, we have $x \sim y$ if and only if there exists a path in X connecting x and y.

Proof. The constant path $x\colon I \to X$ given by $t \mapsto x$ is a path on X, thus $x \sim x$. Let $x \sim y$ and γ be a path from $\gamma(0) = x$ to $\gamma(1) = y$, then we can define a map $\lambda\colon [0,1] \to X$ given by $\lambda(t) := \gamma(1-t)$, so that λ is both continuous, and $\lambda(0) = y$ while $\lambda(1) = x$—thus λ is a path between y and x, therefore $y \sim x$. Suppose now $y \sim z$ and let η be a path connecting y to z. We define a map $\phi\colon I \to X$ given by
$$\phi(t) := \begin{cases} \lambda(2t), & \text{for } t \in [0, 1/2] \\ \eta(2t-1), & \text{for } t \in [1/2, 1] \end{cases}$$
which is surely continuous and connects both x and z—thus $x \sim z$. ♮

Definition 14.1.10 (Path connected components)**.** Let X be a path connected topological space, and $\sim$ be the equivalence relation described in Proposition 14.1.9. The collection $X/{\sim}$ is called the *path connected components* of the space X. We denote the collection of all path components of X by $\pi_0(X)$—the collection of homotopy classes between maps $* \to X$.

Definition 14.1.11 (π_0 functor)**.** The concept of connected components of a space induce a covariant functor $\pi_0\colon \mathtt{Top} \to \mathtt{Set}$ defined by mapping objects $X \mapsto \pi_0 X$, and morphisms $f\colon X \to Y$ to $\pi_0 f\colon \pi_0 X \to \pi_0 Y$—which is a well defined map since, given a path component P of X, the set $f(X) \subseteq Y$ is connected and therefore contained in a unique path component of Y.

Properties of Connectivity

Theorem 14.1.12 (Morphisms preserve connectivity)**.** Let X be a (path) connected space and $f\colon X \to Y$ be a topological morphism. Then $f(X) \subseteq Y$ is (path) connected.

Proof. Suppose that $f(X)$ is not connected, and let $g\colon f(X) \to \{0,1\}$ be a non-constant morphism—in particular, this is equivalent to the condition of $gf\colon X \to \{0,1\}$ being a non-constant morphism, thus implying in the non-connectedness of X.

On the other hand, assume now that X is path connected and let $u, v \in f(X)$—define $x, y \in X$ so that $f(x) = u$ and $f(y) = v$. Let $\gamma: I \to X$ be a path connecting x and y—then $f\gamma: I \to f(X)$ is a path connecting u and v, which proves the proposition. ♮

Corollary 14.1.13. Connectedness and path connectedness are both topological properties.

Corollary 14.1.14. The quotient of a (path) connected topological space is (path) connected.

Proposition 14.1.15. Let $f: X \twoheadrightarrow Y$ be a morphism of topological spaces X and Y. If Y is connected and, for all $y \in Y$, the fibre $f^{-1}(y)$ is connected, then X is connected.

Proof. Let $g: X \to \{0, 1\}$ be a morphism. From the connectedness condition on the fibres of f, it follows that g must be constant throughout the fibres of f—this implies in the existence of a morphism $g^*: Y \to \{0, 1\}$ such that $g = g^* f$. Since Y is connected, g^* must be constant, therefore the composition $g^* f$ is constant and so is g—that is, X is connected. ♮

Proposition 14.1.16. Let X be a topological space. The following are properties concerning connectivity:

(a) Let U and V be disjoint open subsets of X. If $A \subseteq X$ is connected and is contained in $U \cup V$, then either $A \subseteq U$ or $A \subseteq V$.

(b) If X contains a dense connected set, then X is connected.

(c) Let $A \subseteq X$ be a connected set. Then $\mathrm{Cl}(A)$ is connected and any subset $B \subseteq X$ with $A \subseteq B \subseteq \mathrm{Cl}(A)$ is connected.

Proof. (a) Suppose A has points in both sets, then $A \cap U$ and $A \cap V$ would be two disjoint non-empty sets whose union is A—which would imply that A is not connected.

(b) Let $D \subseteq X$ be a dense connected set and suppose there exists non-empty disjoint open sets U and V whose union is X. Therefore $D \subseteq U \cup V$ and by item (a) we have $D \subseteq U$ or $D \subseteq V$. If $D \subseteq U$, then $X = \mathrm{Cl}(D) \subseteq \mathrm{Cl}(U)$ but since $\mathrm{Cl}(U) \subseteq X$ it follows that $V = \varnothing$.

(c) Notice that since $A \subseteq B \subseteq \mathrm{Cl}(A)$ then $\mathrm{Cl}(A) = B$ and from item (b) we conclude that B is connected.

♮

Proposition 14.1.17. Let J be a set, and consider a collection $(X_j)_{j \in J}$ of (path) connected topological spaces. Define the space $X := \bigcup_{j \in J} X_j$. If the intersection $\bigcap_{j \in J} X_j$ is non-empty, then X is (path) connected.

Proof. Suppose the intersection is non empty and let $x \in \bigcup_{j \in J} X_j$ be any point. We split the proposition into the two given cases.

Suppose X_j is connected for all $j \in J$. Let $f: X \to \{0, 1\}$ be a morphism—we want to show that it has to be constant. Since X_j is connected, it follows that the restriction

$f|_{X_j}$ has to be constant, and, since $x \in X_j$, then $f(x) = f(y)$ for all $y \in X_j$. The fact that this must be true for all X_j shows that f must be constant throughout the whole set X.

Suppose X_j is path connected for all $j \in J$—that is, from hypothesis, X_j is a path connected component of X. We now prove that the intersection is connected to each of the path connected components. For each $j \in J$, choose any point $y \in X_j$. Since $x \in X_j$, then $x \sim y$ and therefore every point of X is connected by a path. ♮

Theorem 14.1.18. In the real space $\mathbf{R}$, the only connected sets are intervals and singletons.

Proof. If A is not an interval, there must exist a pair $x, y \in A$ for which there is $z \notin A$ with $x < z < y$. This way, we can write A as the union of two disjoint non-empty sets: $A = [A \cap (-\infty, z)] \cup [A \cap (z, \infty)]$—that is, A is disconnected.

We now show that intervals are connected. For the sake of contradiction, assume there exists an interval I such that there are disjoint and non-empty sets A and B for which $I = A \cup B$—that is, we assume I is disconnected. Let $x, y \in I$ be any two elements with $x < y$ and $x \in A$, while $y \in B$. By the archimedian principle, valid in $\mathbf{R}$, since the set $C := [x, y) \cap A$ is both non-empty and bounded above, thus C has a supremum, let $s := \sup C$—where $x < s \leqslant y$. Notice that $s \notin A$, since if so then s would not be the supremum of C, therefore $s \in B$—which also cannot be the case, since s would not, again, be the supremum of C. We conclude that the sets A and B cannot be constructed at all, hence there is no disconnected interval in $\mathbf{R}$. ♮

Proposition 14.1.19 (Partition by components)**.** The (path) connected components of any topological space form a *partition* of the space.

Proof. Let X be a topological space. Let U and V be (path) connected components of X and suppose that $U \cap V$ is non-empty. Since the union of (path) connected sets is (path) connected, it follows that $U \cup V$ is (path) connected from Proposition 14.1.17. Since (path) connected components are maximal, it follows that $U \cup V = U = V$. Therefore the intersection of (path) connected components is always empty.

Let $p \in X$ be any point. Considering the (path) connected set $\{p\}$, one finds a maximal (path) connected component U containing $\{p\}$. Therefore the collection of (path) connected components covers X. ♮

Corollary 14.1.20. Let X be a non-empty topological space. The following are properties concerning the components of X:

(a) Any non-empty (*path*) *connected* subset of X is contained in a *single connected component*.

(b) Any non-empty *path* connected subset of X is contained in a *single path connected component*.

(c) Each *connected* component of X is the disjoint union of *path* components.

Proof. For the proof of (a), let $A \subseteq X$ be a non-empty (path) connected set. Let $p \in A$ be any point. Since connected components cover X, there exists a connected component

$C \subseteq X$ such that $p \in C$. Since the union of connected sets is connected, then $A \cup C$ is connected (if A is path connected, then it's also connected). Since $A \cup C$ contains C, by the maximality of the components it must be the case that $A \cup C = C$. Therefore $A \subseteq C$. For item (b) we have a completely analogous proof.

For the proof of item (c), let C be a connected component of X. For any point $p \in C$, since path components cover X, let P be a path component containing p. Since path components are connected sets, the union $P \cup C$ is connected and contains C. Again, by the maximality of C this implies in $P \subseteq C$. Since path components are also maximal and any point of C can be found in a path component contained in C, it follows that C is the disjoint union of path components. ♮

Lemma 14.1.21. The connected components of a non-empty topological space are closed.

Proof. Let C be a connected component of a topological space X. Since C is dense in $\mathrm{Cl}(C)$ it follows that $\mathrm{Cl}(C)$ is connected (see Proposition 14.1.16 item (b)). Since connected components are maximal, it follows that $\mathrm{Cl}(C) = C$, thus C is closed. ♮

Proposition 14.1.22. Any path connected space is connected.

Proof. Let X be a path connected space and $f\colon X \to \{0,1\}$ be any continuous map. Let $x, y \in X$ be any two points. Since X is path connected, there exists a path $\gamma\colon I \to X$ between x and y. From the hypothesis that f is continuous, f must be continuous on $\gamma(I)$, therefore constant. Notice that, since this is true for all points of X we find that $f(x) = f(y)$ for all $x, y \in X$. ♮

Proposition 14.1.23. Connectedness and path connectedness are *homotopy invariants*.

Proof. Let $f\colon X \xrightarrow{\simeq} Y$ be a homotopy equivalence, and $g\colon Y \xrightarrow{\simeq} X$ be its homotopy inverse—furthermore, consider a homotopy $h\colon fg \to \mathrm{id}_Y$.

Suppose X is connected, we shall prove that Y is connected. Let $k\colon Y \to \{0,1\}$ be any continuous map and $y, y' \in Y$ be any two points. Since X is connected, in particular the map $kf\colon X \to \{0,1\}$ is constant—therefore, $kfg(y) = kfg(y')$. Since h is a homotopy, then the induced maps $h(y,-), h(y',-)\colon I \rightrightarrows Y$ are such that $h(y,0) = fg(y)$ and $h(y,1) = y$, while $h(y',0) = fg(y')$ and $h(y',1) = y'$. We conclude that $h(y,-)$ is a path from $fg(y)$ to y and $h(y',-)$ is a path from $fg(y')$ to y'. Since k is continuous, it follows that $kfg(y) = k(y)$ and $kfg(y') = k(y')$. Finally we obtain $k(y) = k(y')$, showing that k is constant.

Suppose X is path connected. Then $f(X)$ is path connected and therefore k must be constant in every point of $f(X)$. Now, if $y \in Y \setminus f(X)$, we consider the induced map $h(y,-)\colon I \rightrightarrows Y$—which is a path from $fg(y) \in f(X)$ to y. Therefore, every point of $Y \setminus f(X)$ can be connected by a path to a point of $f(X)$, proving that Y itself is path connected. ♮

Proposition 14.1.24 (Products preserve connectedness)**.** Let $(X_j)_{j \in J}$ be a collection of (path) connected topological spaces. Then $\prod_{j \in J} X_j$ is (path) connected.

Proof. Lets define the notation $X := \prod_{j\in J} X_j$.

Suppose $(X_j)_{j\in J}$ is composed of connected spaces. Let $k\colon X \to \{0,1\}$ be any continuous map. For every $j_0 \in J$, let $p \in \prod_{j\in J\setminus j_0} X_j$ be any point—which, from construction, excludes the j_0-th coordinate from the original product space X. Consider the continuous map $\iota_{j_0}\colon X_{j_0} \to X$ for which $\pi_j \iota_{j_0}(x) := \pi_j(p)$ for $j \neq j_0$ and $\pi_{j_0}\iota_{j_0}(x) := x$—that is, ι_{j_0} embeds X_{j_0} in X where every coordinate, but j_0, is fixed using the pre-chosen point p. Since k is continuous and X_j is connected, the continuous map $k\iota_{j_0}\colon X_j \to \{0,1\}$ has to be constant. Now, if we take any pair of points $x, y \in X$, one sees that $k(x) = k(y)$ from the fact that $k\iota_j$ has to be constant for all $j \in J$—for the suitable choice of initial fixed point.

Suppose $(X_j)_{j\in J}$ is a collection of path connected spaces. Let $x, x' \in X$ be any two elements. For each $j \in J$, let $\gamma_j\colon I \to X$ be a path connecting $\gamma_j(x)$ to $\gamma_j(x')$. By the universal property of the product topology, there exists a unique continuous map $\gamma\colon I \to X$ such that $\pi_j\gamma = \gamma_j$ for every $j \in J$. Since the product of the paths, γ, is a path from x to x', we conclude that X is path connected. ♮

Theorem 14.1.25. A space X is *connected* if and only if the covariant functor $\mathrm{Mor}_{\mathsf{Top}}(X, -)$ *preserves coproducts*.

Prove connectedness iff preserves coproducts.

Local Connectedness

Definition 14.1.26 (Locally (path) connected space)**.** A space X is said to be *locally (path) connected* if it admits a basis of (path) connected open subsets. Equivalently, for every $p \in X$ and neighbourhood $U \subseteq X$ of p, there exists a (path) connected neighbourhood $V \subseteq U$ of p.

Proposition 14.1.27 (Properties for locally (path) connected spaces)**.** Let X be a locally (path) connected space. Then the following properties hold:

(a) Every open subset of X is locally (path) connected.

(b) Every (path) connected component of X is open.

Proof. Since any open subset A can be written as the union of elements of the basis, then the elements of the basis of X contained in A form a basis of (path) connected sets for A. Therefore A is locally (path) connected.

Let C be a (path) connected component of X, and $p \in C$ be any point. Since X is locally (path) connected, there exists a (path) connected neighbourhood $U \subseteq X$ of p. Therefore by Corollary 14.1.20 we find that $U \subseteq C$. Therefore C is open. ♮

Proposition 14.1.28 (Properties of locally path connected spaces)**.** Let X be a locally *path* connected space. Then the following properties hold:

(a) X is locally connected.

(b) The path components of X are equal to its connected components.

(c) X is connected if and only if it is path connected.

Proof. (a) From Proposition 14.1.22 we find that the path connected basis of X is also a connected basis, therefore X is locally connected.

(b) Let $p \in X$ be any point and consider P and C to be, respectively, the path component and the connected component containing p. Since P contains a point of C, then $P \subseteq C$—since $P \cup C$ is connected and contains C. Moreover, by Corollary 14.1.20 we know that C can be written as the disjoint union of path components.

Moreover, it needs to be the case that P is the only path component in C. On the contrary, then P and $C \setminus P$ would be non-empty disjoint open sets whose union is C—making C disconnected. Therefore $C = P$.

(c) X is connected if and only if it has a unique connected component. Since path components equal connected components, it follows that uniqueness of connected components is equivalent to uniqueness of path components. Since X is path connected if and only if it has a unique path component, the proposition follows. ♮

Proposition 14.1.29 (Connectivity of manifolds)**.** Let M be an n-manifold with or without boundary. The following properties hold:

(a) M is *locally path connected*.

(b) M has countably many connected components, each being an open subset of M and a connected topological space.

Proof. (a) Since M has a basis of coordinate balls and those are isomorphic to an open subset of $\mathbf{R}^n$, which are path connected, it follows that the coordinate balls are path connected.

(b) By item (a) and Proposition 14.1.27, the collection of connected components of M forms an *open* cover of M. Since M is second countable, there exists a countable subcover. Since connected components are disjoint, the subcover must be the cover itself, showing that the collection of connected components of M is countable.

Since connected components are open, together with the initial topology generated by the canonical inclusion—that is, the subspace topology—we find that connected components are topological manifolds. ♮

Corollary 14.1.30. A connected manifold is path connected.

Proof. Follows directly from Propositions 14.1.28 and 14.1.29. ♮

14.2 Compact Spaces

Definition 14.2.1 (Compact space)**.** A topological space X is said to be *compact* if for *every* open cover there exists a *finite subcover*.

Proposition 14.2.2 (Image of compact space)**.** If $f\colon X \to Y$ is a topological morphism and X is compact, then $f(X)$ is compact in Y.

Proof. Let C be an open cover of $f(X)$. Consider the preimage collection $\mathcal{U} := \{f^{-1}(V)\}_{V \in C}$, which by construction covers X. Since X is compact, let $\mathcal{U}'$ be the finite subcover given by $\mathcal{U}$. If we now consider the image collection $C' := \{f(U)\}_{U \in C'}$, we find that C' covers $f(X)$ and is contained in C, therefore a finite subcover. ♮

Corollary 14.2.3. Compactness is an *invariant* property of topological spaces.

Corollary 14.2.4. The quotient of a compact space is compact.

Proof. Since any quotient space is the image of a continuous projection, the proposition follows from Proposition 14.2.2. ♮

Proposition 14.2.5 (Closed subset is compact)**.** In a compact space any closed subset is compact.

Proof. Let X be a space and $A \subseteq X$ any closed subset. If $\mathcal{U}$ is an open cover of A, then $\mathcal{U} \cup \{X \setminus A\}$ is a cover of X. Moreover, since X is compact, there exists a finite subcover $\mathcal{U}' \subseteq \mathcal{U} \cup \{X \setminus A\}$. In particular, since $\mathcal{U}'$ covers X, it also covers A. ♮

Checking if a Space is Compact

Definition 14.2.6 (Finite intersection property)**.** A collection of sets $\mathcal{A}$ is said to satisfy the *finite intersection property* if and only if for every finite collection $\{A_1, \dots, A_n\} \subseteq \mathcal{A}$, we have $\bigcap_{j=1}^{n} A_j \neq \varnothing$, that is, the intersection is non-empty

Proposition 14.2.7. A space is compact if and only if every collection of closed subsets satisfying the finite intersection property has non-empty intersection.

Proof. Let X be a compact space and C be any collection of closed subsets satisfying the finite intersection property. Suppose, for the sake of contradiction, that C has an empty intersection—thus the complement of the intersection covers X. Hence there must exist a finite collection $\{A_1, \dots, A_n\} \subseteq C$ whose corresponding complement covers X and therefore $\bigcup_{j=1}^{n}(X \setminus A_j) = X \setminus \big(\bigcap_{j=1}^{n} A_j\big) = X$—then $\bigcap_{j=1}^{n} A_j = \varnothing$, which contradicts the hypothesis that C satisfies the finite intersection property.

Suppose that the latter condition is true. Suppose there exists an open cover $\mathcal{U}$ of X that has no finite subcover. In particular, the collection $\mathcal{U}' := \{X \setminus U \colon U \in \mathcal{U}\}$ is composed of closed sets and satisfies the finite intersection property—thus $\bigcup_{U \in \mathcal{U}}(X \setminus U) = X \setminus \big(\bigcap_{U \in \mathcal{U}} U\big) = X$, which immediately implies that the intersection of $\mathcal{U}$ is empty, yielding a contradiction. Hence $\mathcal{U}$ must have a finite subcover. ♮

Lemma 14.2.8 (Non-empty countable intersection)**.** Let X be a compact space and $\{C_j\}_{j \in \mathbf{N}}$ be a countable collection of non-empty subsets of X for which $C_{j+1} \subseteq C_j$ for every $j \in \mathbf{N}$—that is, the sets are nested. Then the countable intersection $\bigcap_{j \in \mathbf{N}} C_j$ is non-empty.

Proof.

Prove, this is an exercise

♮

Definition 14.2.9 (Proper map)**.** A continuous map $f: X \to Y$ is said to be *proper* if for all compact sets $C \subseteq Y$ the preimage $f^{-1}(C) \subseteq X$ is compact.

Theorem 14.2.10. Let X be a Hausdorff space and $x \in X$ any point. For every compact set $K \subseteq X \setminus \{x\}$ there exists two disjoint open sets U and V such that $x \in U$ and $K \subseteq V$.

Proof. Since X is Hausdorff and $x \notin K$, for any $k \in K$ there exists disjoint neighbourhoods U_x and V_k of x and k, respectively. Let $\{V_k\}_{k \in K}$ be a collection of such neighbourhoods—which is also an open cover of K. Since K is compact, there exists a finite collection of points such that $\{V_{k_1}, \dots, V_{k_n}\}$ is a finite subcover. Defining $V := \bigcup_{j=1}^n V_{k_j}$ and $U := \bigcup_{j=1}^n U_{k_j}$, we find that $K \subseteq V$ and $x \in U$. ♮

Corollary 14.2.11 (Hausdorff compact subsets)**.** Any *compact* subset of a Hausdorff space is *closed*.

Proof. If X is Hausdorff and $C \subseteq X$ is a compact set, let $x \in X \setminus C$ be any point. From Theorem 14.2.10 we know the existence of a neighbourhood U of x that is disjoint from K—thus $U \subseteq X \setminus C$ and hence C is closed. ♮

Corollary 14.2.12 (Maps from compact to Hausdorff spaces)**.** Let X be compact and Y Hausdorff. Every continuous map $f: X \to Y$ is *closed* and *proper*. Moreover, the following are consequential properties:

(a) If f is injective, then it is an *embedding*.

(b) If f is surjective, then it is a *quotient map*.

(c) If f is bijective, then it is an *isomorphism*.

Proof. We first prove that f is closed. Let $C \subseteq X$ be any closed set of X, which is therefore compact. From Proposition 14.2.2 we find that $f(C)$ is compact and from Corollary 14.2.11 we have that $f(C)$ is closed.

Now we prove that f is proper. Let $K \subseteq Y$ be any compact set and consider the preimage $f^{-1}(K)$. Since Y is Hausdorff, as pointed before, K is closed. Now since f is continuous, the preimage of closed sets is closed—hence $f^{-1}(K)$ is closed. From the fact that X is compact, we conclude that $f^{-1}(K)$ is compact. For the last three consequences, they come from Proposition 13.5.23. ♮

Lemma 14.2.13 (Tube lemma)**.** Let X be any space and Y be compact. For every $x \in X$ and open set $U \subseteq X \times Y$ containing $\{x\} \times Y$, there exists a neighbourhood $V \subseteq X$ of x such that $V \times Y \subseteq U$.

Proof. Since the product of open sets form a basis for the product topology, for every $y \in Y$ there exists a neighbourhood $V \times W \subseteq U$ of (x, y). Since $\{x\} \times Y \simeq Y$ and Y is compact, then $\{x\} \times Y$ is compact and therefore must exist a finite collection $\{V_j \times W_j\}_{j=1}^n$ of open sets of $X \times Y$ covering $\{x\} \times Y$. Then if $V := \bigcap_{j=1}^n V_j$ we find that $V \times Y \subseteq U$. ♮

Theorem 14.2.14 (Closed projection)**.** Given topological spaces X and Y, the space X is compact if and only if the canonical projection $\pi\colon X \times Y \twoheadrightarrow Y$ is closed.

Proof. Suppose X is compact, then if $C \subseteq X \times Y$ is any closed set, let $y \in Y \smallsetminus \pi(C)$ be any point—we shall show that there exists a neighbourhood of y outside of $\pi(C)$. Consider the open set $U := X \times (Y \smallsetminus \pi(C))$, which certainly contains $X \times \{y\}$. From Lemma 14.2.13, we find a neighbourhood $V \subseteq Y$ of y such that $X \times V \subseteq U$, that is, $V \subseteq Y \smallsetminus \pi(C)$, which settles that $\pi(C) \subseteq Y$ is closed.

Suppose now that X is a space such that $\pi\colon X \times Y \twoheadrightarrow Y$ is closed for any space Y. Let $\mathcal{C}$ be any collection of closed subsets of X satisfying the finite intersection property—we'll show that $\bigcap_{C\in\mathcal{C}} C$ is non-empty. Define Y to be the space to consisting of the underlying set $X \cup \{*\}$ for some point $*$ and the topology given by 2^X and $\{C \cup \{*\} : C \in \mathcal{C}\}$. Let $K \subseteq X \times Y$ be the closure of the diagonal of X—that is, $K := \operatorname{Cl} \Delta_X$. From hypothesis, $\pi(K) \subseteq Y$ is closed and from construction $X \subseteq \pi(K)$.

We now show that $* \in \pi(K)$. Suppose on the contrary that $* \notin \pi(K)$. Since $\pi(K)$ is closed, we can find a neighbourhood $V \subseteq Y \smallsetminus \pi(K)$ of $*$—and therefore $V \cap X = \varnothing$. From the construction of the topology of Y, one could only hope to write V as the intersection of finitely many sets of the form $C \cup \{*\}$ for $C \in \mathcal{C}$—on the other hand, since $\mathcal{C}$ satisfies the finite intersection property, for any finite collection of sets of $\mathcal{C}$ one can find $x \in X$ such that $x \in C_1 \cap \dots \cap C_n$, therefore $(C_1 \cup \{*\}) \cap \dots \cap (C_n \cup \{*\})$ still contains a point of X. This shows that it is impossible to build V out of such sets—hence we obtain a contradiction and thus $* \in \pi(K)$ and there exists $(x_0, *) \in K$ for some $x_0 \in X$.

For every $C \in \mathcal{C}$, any neighbourhood $U \times (C \cup \{*\})$ of $(x_0, *)$ has a non-empty intersection with the diagonal Δ_X—thus $C \cap U$ is non-empty. Moreover, $x_0 \in C$ for all $C \in \mathcal{C}$, otherwise, since C is closed, one can find a neighbourhood $U \subseteq X \smallsetminus C$ of x_0, which should not be possible. Therefore $x_0 \in \bigcap_{C\in\mathcal{C}} C$ and by Proposition 14.2.7 we conclude that X is compact. ♮

Tychonoff Theorem

Lemma 14.2.15. Let J be a set, and $(X_j)_{j\in J}$ be a collection of topological spaces. For any point $x \in \prod_{j\in J} X_j$ and subset $A \subseteq \prod_{j\in J} X$ of the product space, we have $x \in \operatorname{Cl} A$ if, for every finite $F \subseteq J$, we have $\pi_F(x) \in \operatorname{Cl}(\pi_F(A))$—where $\pi_F\colon \prod_{j\in J} X_j \twoheadrightarrow \prod_{j\in F} X_j$ is the canonical projection map.

Proof. Suppose, on the contrary, that $x \notin \operatorname{Cl} A$—then there exists a neighbourhood $U \subseteq \prod_{j\in J} X_j$ of x which is disjoint from A. From the definition of the product topology, the collection of preimages $\pi_j(U_j)$ for open sets $U_j \subseteq X_j$ form a sub-basis of the product space. In particular, from the basis properties, this allows for the existence of an open set $V := U_{j_1} \times \dots \times U_{j_n}$ such that, denoting $F := \{j_1, \dots, j_n\} \subseteq J$, the basis element $\pi_F^{-1}(V)$ of the product space is a neighbourhood of x and $\pi_F^{-1}(V) \subseteq U$. Therefore π_F^{-1} and A are disjoint, which is equivalent to

$$A \subseteq \left(\prod_{j\in J} X_j\right) \smallsetminus \pi_F^{-1}(V) = \pi_F^{-1}\left(\prod_{j\in F} X_j \smallsetminus V\right)$$

that is, $\pi_F(A) \subseteq \left(\prod_{j\in F} X_j\right) \setminus V$. Since V is a neighbourhood of $\pi_F(x)$ which is disjoint from $\pi_F(A)$ we conclude, by Proposition 12.1.17, that $\pi_F(x) \notin \mathrm{Cl}(\pi_F(A))$—which proves the lemma. ♮

Theorem 14.2.16 (Tychonoff)**.** The Cartesian product of a *set* of compact topological spaces, endowed with the product topology, is compact.

Proof. Denote by $\{X_\alpha\}_{\alpha<\kappa}$ a collection of compact spaces indexed by an ordinal κ. We now show via induction on κ that, for any space Y, the canonical projection $Y \times \prod_{\alpha<\kappa} X_\alpha \twoheadrightarrow Y$ is closed.

For the ease of notation we define $X^\gamma := Y \times \prod_{\alpha<\gamma} X_\alpha$ for all $\gamma \leqslant \kappa$—moreover for $\lambda \leqslant \gamma$ we denote by $\pi^\gamma_\lambda \colon X^\gamma \twoheadrightarrow X^\lambda$ the canonical projection between such spaces. We also define that if $C \subseteq X^\kappa$ is a *closed* set, then $C_\lambda := \mathrm{Cl}(\pi^\kappa_\lambda(C))$—thus our goal is equivalent of showing that $\pi^\kappa_0(C) = C_0$.

Assume as inductive hypothesis that for all $x_0 \in C_0$ there exists $x_\lambda \in C_\lambda$ for every $\lambda < \kappa$ such that, if $\lambda < \gamma < \kappa$, then

$$\pi^\gamma_\lambda(x_\gamma) = x_\lambda,$$

and in particular $\pi^\lambda_0(x_\lambda) = x_0$. Again, equivalent to our goal is to show that $\pi^\kappa_0(x_\kappa) = x_0$.

If $\kappa = \lambda + 1$ is a successor ordinal, then the projection $\pi^\kappa_\lambda \colon X^\lambda \times X_\lambda \twoheadrightarrow X^\lambda$ is closed from the fact that X_λ is compact, by Theorem 14.2.14. In particular, $\pi^\kappa_\lambda(C) \subseteq X^\lambda$ is closed and, by the inductive hypothesis, $\pi^\kappa_\lambda(C) = C_\lambda$—thus there exists $x_\kappa \in K$ for which $\pi^\kappa_\lambda(x_\kappa) = x_\lambda$, hence

$$\pi^\kappa_0(x_\kappa) = \pi^\lambda_0 \pi^\kappa_\lambda(x_k) = \pi^\lambda_0(x_\lambda) = x_0,$$

which was out goal.

Now, if κ is a limit ordinal, then $X^\kappa = \lim_{\lambda<\kappa} X^\lambda$ together with transitions maps π^γ_λ. The limit of the tuple $(x_\lambda)_{\lambda<\kappa}$ defines a point $x_\kappa \in X^\kappa$, we wish to show that $x_\kappa \in C$. For every finite set of ordinals F below κ there exists $\lambda < \kappa$ above all ordinals of F, therefore

$$\pi_F(x_\kappa) = \pi^\lambda_F \pi^\kappa_\lambda(x_\kappa) = \pi^\lambda_F(x_\lambda).$$

Moreover $\pi^\lambda_F(x_\lambda) \in \pi^\lambda_F(C_\lambda)$, where from definition $C_\lambda = \mathrm{Cl}(\pi^\kappa_\lambda(C))$. Since π^λ_F is a continuous map,

$$\pi^\lambda_F(C_\lambda) = \pi^\lambda_F(\mathrm{Cl}(\pi^\kappa_\lambda(C))) \subseteq \mathrm{Cl}(\pi^\lambda_F \pi^\kappa_\lambda(C)) = \mathrm{Cl}(\pi_F(C)).$$

This shows that $\pi_F(x_\kappa) \in \mathrm{Cl}(\pi_F(C))$, which by Lemma 14.2.15 shows that $\pi(C)$ is closed. ♮

Applications to Real & Metric spaces

Lemma 14.2.17 (Closed intervals are compact)**.** Every closed and bounded interval in **R** is compact.

Proof. Consider a closed interval $[a, b]$ in $\mathbf{R}$ and let $\mathcal{U}$ be a cover for $[a, b]$. Define S to be the set of all points $x \in [a, b]$ for which the interval $[a, x]$ is covered by finitely many sets of $\mathcal{U}$—since there must exist a set $U \in \mathcal{U}$ for which $a \in U$, then $a \in S$. Since S is non-empty, by the least upper bound property one can define a point $x_0 := \sup S$.

Let $U_0 \in \mathcal{U}$ be a set containing x_0 and let $\varepsilon > 0$ be such that $(x_0 - \varepsilon, x_0] \subseteq U_0$. Since x_0 is the supremum of S, there must also exist $x \in S$ such that $x \in (x_0 - \varepsilon, x_0]$—that is, the interval $[a, x_0]$ can be covered by finitely many sets of $\mathcal{U}$, say $[a, x_0] \subseteq U_0 \cup U_1 \cup \cdots \cup U_n$, therefore $x_0 \in S$. If, for the sake of contradiction, $x_0 < b$, then by the fact that U_0 is an open set, there must exist $x \in U_0$ with $x > x_0$ and yet $x \in [a, b]$—which implies that $[a, x] \subseteq \bigcup_{j=0}^{n} U_j$ and therefore x_0 isn't the supremum of S, leading to a contradiction. Thus $x_0 = b$ and hence $[a, b] \subseteq \bigcup_{j=0}^{n} U_j$—which proves the proposition. ♮

Corollary 14.2.18 (Heine-Borel)**.** A subset of $\mathbf{R}^n$ is compact if and only if it is both closed and bounded.

Proof. Let $K \subseteq \mathbf{R}^n$ be a compact set. If we let $\mathcal{U}$ be the cover of $\mathbf{R}^n$ by open balls centred at the origin—with any real radius—in particular $\mathcal{U}$ will be a cover for K and since K is compact, there exists a finite subcover $\mathcal{U}$ covering K. From this we conclude that there are only finitely many balls of real radius that cover K—hence K is necessarily bounded. From the Hausdorffness of $\mathbf{R}^n$, we obtain from Corollary 14.2.11 that K is closed.

On the other hand, let $C \subseteq \mathbf{R}^n$ be a closed and bounded set. Since C is a bounded set, its projection into each coordinate must also be bounded, therefore one can obtain a collection of intervals $([a_j, b_j])_{j=1}^{n}$ in $\mathbf{R}$ for which

$$C \subseteq \prod_{j=1}^{n} [a_j, b_j].$$

From Lemma 14.2.17 we know that each $[a_j, b_j]$ is compact—thus by Theorem 14.2.16 we conclude that $\prod_{j=1}^{n} [a_j, b_j]$ is compact, but since C is a closed subset of a compact set, by Proposition 14.2.5 we find that C is compact. ♮

Corollary 14.2.19 (Extreme values on compact sets)**.** Let X be a compact space and $f \colon X \to \mathbf{R}$ be a continuous map. Then f is *bounded* and attains its *maximum* and *minimum* values on X.

Proof. Since $f(X) \subseteq \mathbf{R}$ is compact, then by Corollary 14.2.18 we obtain that $f(X)$ is both closed and bounded—thus in particular $f(X)$ contains its supremum and infimum. ♮

We can even extend Corollary 14.2.18 to a more general context, encompassing all metric spaces—this is what the following proposition does.

Proposition 14.2.20. In any metric space X, if $A \subseteq X$ is *compact* then A is both *bounded* and *closed* in X.

Proof. Since every metric space is Hausdorff, A is closed by Corollary 14.2.11. Moreover, if $x \in A$ is any point, we find that the collection of open balls $\mathcal{B}_x := \{B_x(n)\}_{n \in \mathbf{N}}$ forms an open cover of A and, since A is compact, there exists a finite subcover of $\mathcal{B}_x$. Therefore there exists a maximal $m \in \mathbf{N}$ for which $A \subseteq B_x(m)$—thus A is bounded. ♮

Remark 14.2.21. Notice that a bounded and closed set in a metric space does *not* need to be compact, for instance, consider the space $\ell^2(\mathbf{N})$ (see Example 23.1.15) and the subset $A := \{f_n\}_{n \in \mathbf{N}}$ composed of sequences $f_n \in \ell^2(\mathbf{N})$ such that $f_n(j) := \delta_{nj}$—that is, sequences where the only non-zero term is the n-th one, which equals 1. Clearly A is bounded since $\|f_n\|_2 = 1$, moreover, A is closed because it has no limit points. However, since no subset of A contains a limit point of A, we find that A is not limit point compact—thus not compact (see Theorem 14.3.8).

Tychonoff & The Axiom of Choice

Theorem 14.2.22 (Tychonoff & Choice)**.** Tychonoff's theorem is *equivalent* to the axiom of choice.

Prove

Examples of Explicit Isomorphisms

Example 14.2.23 (Torus T^2)**.** The 2-torus T^2 is defined as the quotient $T^2 := (I \times I)/\sim$ where $\sim$ is the smallest equivalence relation on $I \times I$ for which $(0, t) \sim (1, t)$ and $(s, 0) \sim (s, 1)$ for all $t, s \in I$. Let's prove that there exists a topological isomorphism

$$T^2 \simeq S^1 \times S^1.$$

First, lets consider the interval I and the equivalence relation $0 \sim_1 1$ so that $I/\sim_1$ is just the interval with its ends glued to a common point. Consider the morphism $f\colon I \twoheadrightarrow S^1$ given by $t \mapsto (\cos(2\pi t), \sin(2\pi t))$ and notice that $f(1) = f(0) = 1$—that is, f is constant on the fibres of the canonical projection $\pi\colon I \twoheadrightarrow I/\sim_1$—moreover, f is certainly surjective since cosine and sine are both maps of period 2π. By Theorem 13.5.3 we find a unique morphism $g\colon I/\sim_1 \to S^1$ such that the following diagram commutes

$$\begin{tikzcd} I \arrow[d, two heads] \arrow[dr, bend left, "f"] & \\ I/\sim_1 \arrow[r, dashed, "\simeq", "g"'] & S^1 \end{tikzcd}$$

Notice that besides g being surjective, given $[t], [s] \in I/\sim_1$ such that $g([t]) = g([s])$ then $t - s \in \mathbf{Z}$ since cosine and sine have a period of 2π—but since $s, t \in I$, this can only be the case if $t = s$, thus g is injective. We therefore conclude that g is a bijective map. Since I is compact, by Corollary 14.2.4 we know that $I/\sim_1$ is compact. Moreover, the inclusion map $S^1 \hookrightarrow \mathbf{R}^2$ is continuous and induces the subspace topology on S^1—since $\mathbf{R}^2$ is Hausdorff, then S^1 is Hausdorff. Therefore g is a bijective map from a compact space $I/\sim_1$ to a Hausdorff space S^1, which by Corollary 14.2.12 is an isomorphism.

Therefore, we find a unique isomorphism $g \times g \colon T^2 \xrightarrow{\simeq} S^1 \times S^1$ such that the following diagram commutes

$$\begin{tikzcd} I \times I \arrow[d, two heads] \arrow[dr, bend left, "f \times f"] & \\ T^2 \arrow[r, dashed, "\simeq", "g \times g"'] & S^1 \times S^1 \end{tikzcd}$$

which proves the isomorphism.

Let $\mathbf{R}^2 \times \mathbf{Z}^2 \to \mathbf{R}^2$ be a group action of $\mathbf{Z}^2$ on $\mathbf{R}^2$ given by

$$((x, y), (n, m)) \mapsto (x + n, y + m),$$

then the quotient of $\mathbf{R}^2$ by such group action, which we'll denote by $\mathbf{R}^2/\mathbf{Z}^2$ is such that there exists a topological isomorphism

$$T^2 \simeq \mathbf{R}^2/\mathbf{Z}^2.$$

Indeed, one sees right away that $\mathbf{R}^2/\mathbf{Z} \simeq (I \times I)/\sim$, where we map $(x, y) + \mathbf{Z}^2 \mapsto (x - \lfloor x \rfloor, y - \lfloor y \rfloor)$ is the explicit isomorphism. For free we then obtain an isomorphism

$$S^1 \times S^1 \simeq \mathbf{R}^2/\mathbf{Z}^2.$$

Example 14.2.24 (Projective space $\mathbf{RP}^2$)**.** Let $\sim_1$ and $\sim_2$ be the smallest equivalence relations on, respectively, S^2 and D^2 such that $p \sim_1 -p$ for all $p \in S^2$, and $v \sim_2 u$ if $v = \lambda u$ for some $\lambda \in \mathbf{R}$, for all $u, v \in D^2$. We'll show that there exists topological isomorphisms

$$\mathbf{RP}^2 \simeq S^2/\sim_1 \simeq D^2/\sim_2.$$

For the first isomorphism, consider the continuous mapping $f_1 \colon \mathbf{RP}^2 \to S^2/\sim_1$ given by $\mathbf{R}\ell \mapsto \left[\frac{\ell}{\|\ell\|}\right]$—which collapses the line $\mathbf{R}\ell$ to the class of the unitary vector $\frac{\ell}{\|\ell\|}$ generating the line, which ensures that f_1 is surjective. Notice that f_1 is necessarily injective since, if $f_1(\mathbf{R}\ell) = f_1(\mathbf{R}\ell')$, then the lines $\mathbf{R}\ell$ and $\mathbf{R}\ell'$ intersect both at 0 and at the common point of the sphere—this implies in $\mathbf{R}\ell = \mathbf{R}\ell'$. Notice that the inclusion $i_1 \colon S^2/\sim_1 \hookrightarrow \mathbf{RP}^2$ is continuous and is the inverse of f_1, sending each unitary vector $v \in S^2/\sim_1$ of the sphere quotient to the respective line generated by v, namely $\mathbf{R}v \in \mathbf{RP}^2$.

For the second isomorphism we make an analogous construction, take the map $f_2 \colon D^2/\sim_2 \to S^2/\sim_1$ given by the identification $[v]_{\sim_2} \mapsto \left[\frac{v}{\|v\|}\right]_{\sim_1}$, which is both continuous and bijective. Moreover, the inclusion $i_2 \colon S^2/\sim_1 \hookrightarrow D^2/\sim_2$ is a continuous map which is inverse to f_2, thus f_2 establishes the required isomorphism.

14.3 Sequential & Limit Point Compactness

Definition 14.3.1 (Limit point compactness)**.** A space X is said to be *limit point compact* if every infinite subset of X has a limit point in X.

The following important theorem establishes that every compact space is limit point compact.

Theorem 14.3.2 (Bolzano-Weierstraß)**.** Let X be a compact space. Any *infinite subset* $S \subseteq X$ has a limit point.

Proof. Suppose, for the sake of contradiction, that there exists $S \subseteq X$, infinite, with no limit points—from this hypothesis, any point $x \in X$ is not a limit point of S and is either in or out of S. In the former case $x \in S$ there exists a neighbourhood of x, say U_x, for which $U_x \cap S = \{x\}$. In the latter case $x \notin S$, there must exist a neighbourhood U_x of x such that S and U_x are disjoint. From the law of excluded middle, the collection $\mathcal{U} := \{U_x\}_{x \in X}$ is an open cover of X—on the other hand, there exists no finite subcover of $\mathcal{U}$ since each one must only intersect S at a unique point, but S is infinite, hence the contradiction. The infinite set S must therefore contain at least one limit point. ♮

Remark 14.3.3. One should beware that the converse of Theorem 14.3.2 does not hold at all. A simple counterexample goes as follows: endow $\mathbf{R}$ with the topology given by $\{\varnothing, \mathbf{R}\}$ and the open intervals $\{(x, \infty) : x \in \mathbf{R}\}$—in such a space *any* set has a limit point, although the space itself is not compact.

Definition 14.3.4 (Sequential compactness)**.** A space X is said to be *sequentially compact* if every sequence of points $(x_j)_j$ in X, there exists a subsequence $(x'_j)_j \subseteq (x_j)_j$ that converges in X.

Lemma 14.3.5 (In first countable Hausdorff spaces)**.** Let X be a first countable Hausdorff space. If X is limit point compact, then X is sequentially compact.

Proof. Suppose X is indeed limit point compact and let $(x_j)_{j \in \mathbf{N}}$ be any sequence of points in X and let $S := \{x_j\}_{j \in \mathbf{N}}$ be the set of values that the sequence attains. If S is finite, then the sequence contains a constant subsequence—which is therefore convergent in X.

On the other hand, if S is infinite, then by the limit point compactness property of X we find that there exists a limit point $x \in X$ of S. If it is the case that $x_j = x$ for infinitely many $j \in \mathbf{N}$, then the collection of such points form a constant sequence—which converges to x. If the former is not the case, then at most a finite amount of points x_j are equal to x—therefore, one can discard these points from the sequence and only consider the subsequence $(x'_j)_j \subseteq (x_j)_{j \in \mathbf{N}}$ such that $x'_j \neq x$ for all indices j. Since X is assumed to be first countable, there must exist a neighbourhood basis $(B_j)_{j \in \mathbf{N}}$ at x.

We'll now construct a sequence $(x_{j_i})_{i \in \mathbf{N}} \subseteq (x'_j)_j$ such that $x_{j_i} \in B_i$ and thus $x_{j_i} \to x$. Since x is a limit point, one can choose $x_{j_0} \in B_0$. For the hypothesis of induction, suppose we have chosen $j_0 < j_1 < \cdots < j_n$ indices so that $x_{j_i} \in B_i$. By Proposition 12.4.6 we find that, since B_{n+1} is a neighbourhood of x, then B_{x+1} has infinitely many points of S—therefore, one can certainly choose $j_{n+1} > j_n$ for which $x_{j_{n+1}} \in B_{n+1}$. Thus the sequence $(x_{j_i})_{i \in \mathbf{N}}$ was successfully constructed so that it converges to x. ♮

Lemma 14.3.6. Every sequentially compact metric space is second countable.

Proof. Evoking Proposition 12.5.16 it's sufficient to show that a sequentially compact metric space M is separable. We first show that, for every $\varepsilon > 0$, the open cover composed of ε-balls has a finite subcover. For the sake of contradiction, assume there exists some $\varepsilon_0 > 0$ such that the statement is false.

We now construct a sequence $(x_j)_{j\in\mathbf{N}}$: choose any $x_0 \in M$, now since $B_{\varepsilon_0}(x_0)$ does not cover M, one can choose $x_1 \in M \smallsetminus B_{\varepsilon_0}(x_0)$—since no finite collection of ε_0-balls covers M, we may proceed indefinitely for each $j \in \mathbf{N}$, always choosing $x_j \in M \smallsetminus \bigcup_{i<j} B_{\varepsilon_0}(x_i)$.

Since M is sequentially compact, there must exist a convergent subsequence $(x_{j_n})_{n\in\mathbf{N}}$ of $(x_j)_{j\in\mathbf{N}}$—with $x_{j_n} \to x$ for some $x \in M$. Since convergent sequences are Cauchy in metric spaces, for large enough $j \in \mathbf{N}$, we have $d(x_{j+1}, x_j) < \varepsilon_0$—which implies in $x_{j+1} \in B_{\varepsilon_0}(x_j)$, a contradiction by the construction of the sequence. Therefore there must indeed exist a finite collection of ε_0-balls that cover M.

For the separability of M, for each $n \in \mathbf{N}$, define F_n to be a finite set of points of M for which the $1/n$-balls centred at each point of F_n covers M. Since each F_n is finite, the union $F := \bigcup_{n\in\mathbf{N}} F_n$ is countable. Now, if $U \subseteq X$ is a non-empty open subset of X and $p \in U$ is any point and let $B_\varepsilon(p) \subseteq U$ be any neighbourhood of p. Since $\mathbf{R}$ is archimedian, there must exist $n \in \mathbf{N}$ for which $1/n < \varepsilon$—hence there exists $q \in F_n$ for which $p \in B_{1/n}(q)$ and $q \in B_\varepsilon(p)$. We conclude that F is dense in M—thus M is separable. ♮

Proposition 14.3.7 (Metric & second countable spaces)**.** For second countable spaces and metric spaces, sequential compactness implies compactness.

Proof. Let X be a second countable and sequentially compact space. Let $\mathcal{U}$ be any open cover of X. Since every second countable space is Lindelöf (see Proposition 12.5.15), there exists a countable subcover $\mathcal{U}' := \{U_j\}_{j\in\mathbf{N}} \subseteq \mathcal{U}$. For the sake of contradiction, suppose that there is no finite subcover of $\mathcal{U}'$—that is, for every $j \in \mathbf{N}$, there exists $x_j \in X$ such that $x_j \notin \bigcup_{j=0}^{n} U_j$. If $(x_j)_{j\in\mathbf{N}}$ is a sequence formed by points with such a property, since X is sequentially compact, there must exist a convergent subsequence $(x_{j_n})_{n\in\mathbf{N}}$ with $x_{j_n} \to x$ for some point $x \in X$. Let $m \in \mathbf{N}$ be such that $x \in U_m$—from the definition of limit point, there must exist only finitely many x_{j_n} not contained in U_m. However, that cannot be the case since, for all $j_n \geqslant m$, we must have $x_{j_n} \notin \bigcup_{n=0}^{m} U_n$ thus in particular $x_{j_n} \notin U_m$—this implies that $(x_{j_n})_{n\in\mathbf{N}}$ does not converge to x, which is a contradiction.

By Lemma 14.3.6 if M is sequentially compact then it's second countable—therefore, by what was shown above we conclude that M is compact. ♮

Theorem 14.3.8. For metric spaces and second countable Hausdorff spaces, limit point compactness, sequential compactness and compactness are *equivalent* properties.

Proof. This is the result of Theorem 14.3.2, Lemma 14.3.5 and Proposition 14.3.7. ♮

Corollary 14.3.9 (Metric space completeness)**.** Every compact metric space is complete.

Proof. Let M be compact and $(x_j)_{j\in\mathbf{N}}$ be any Cauchy sequence. Since M is compact, then it's also sequentially compact, which implies in the existence of a convergent subsequence $(x'_j)_j \subseteq (x_j)_{j\in\mathbf{N}}$ for which $x'_j \to x$ for some $x \in M$. We'll show that, in fact, $x_j \to x$. Let $B_\varepsilon(x)$ be any open ball centred at x. Since the sequence is Cauchy, there exists $N \in \mathbf{N}$ such that $d(x_m, x_n) < \varepsilon/2$ for every $m, n > N$. On the other hand, there exists $M \in \mathbf{N}$ such that $d(x'_\ell, x) < \varepsilon/2$ for every $\ell > M$—therefore, for all $n, \ell > \max(N, M)$ we have $d(x_n, x) < d(x_n, x'_\ell) + d(x'_\ell, x) < \varepsilon$, thus $x_n \in B_\varepsilon(x)$ for all $n > \max(N, M)$. We conclude that all but finitely many points of $(x_j)_{j\in\mathbf{N}}$ are contained in $B_\varepsilon(x)$ for an arbitrary $\varepsilon > 0$—therefore $x_j \to x$. ♮

14.4 Local Compactness

Definition 14.4.1 (Locally compact spaces)**.** A space X is said to be *locally compact* if and only if, for every point $x \in X$, there exists a compact set $K \subseteq X$ and a neighbourhood $U \subseteq X$ of x for which $U \subseteq K$.

Definition 14.4.2 (Relatively compact)**.** Given a space X, a subset $A \subseteq X$ is said to be *relatively compact in* X if its closure $\operatorname{Cl} A$ is compact.

Proposition 14.4.3 (Local & relative compactness in Hausdorff spaces)**.** Let X be a *Hausdorff* space. The following are equivalent properties:

(a) X is locally compact.

(b) Every point of X has a relatively compact neighbourhood in X.

(c) X has a basis of relatively compact open sets.

Proof. From the definition of local and relative compactness, it is clear that (c) implies (b) and that (b) implies (a). We therefore only prove that (a) implies (c). Let X be a locally compact Hausdorff space. Take $x_0 \in X$ to be any point and let $K \subseteq X$ be a compact set containing a neighbourhood $U \subseteq X$ of x_0. If we let $\mathcal{B}$ be the collection of all neighbourhoods of x_0 which are *contained* in U, then $\mathcal{B}$ forms a neighbourhood basis at x_0. Moreover, since X is Hausdorff, K is closed and every $V \in \mathcal{B}$ is also contained in K—therefore $\operatorname{Cl} V \subseteq K$ and since a closed subset of a compact set is compact then $\operatorname{Cl} V$ is compact. We conclude that V is relatively compact in X and thus $\mathcal{B}$ is a basis composed of relatively compact open sets. ♮

Lemma 14.4.4. Let X be a locally compact Hausdorff space. For every point $x \in X$ and neighbourhood $U \subseteq X$ of x, there exists a relatively compact neighbourhood V of x for which $\operatorname{Cl} V \subseteq U$.

Proof. Since X is locally compact Hausdorff space, by Proposition 14.4.3 there exists a relatively compact neighbourhood W of x. Since $\operatorname{Cl}(W) \setminus U$ is closed in $\operatorname{Cl} W$, it follows that it's compact. Evoking Theorem 14.2.10 one can find disjoint open sets T and Q such that $x \in T$ and $\operatorname{Cl}(W) \setminus U \subseteq Q$. Define a set $V := T \cap W$, then $\operatorname{Cl} V \subseteq \operatorname{Cl} W$ implies in $\operatorname{Cl} V$ compact—thus V is a relatively compact neighbourhood of x, we just need to show that $\operatorname{Cl} V \subseteq U$.

Since T and Q are disjoint, then $T \subseteq X \setminus Q$ and therefore $V \subseteq X \setminus Q$—since $X \setminus Q$ is closed, this implies in $\operatorname{Cl} V \subseteq X \setminus Q$. Therefore, since $\operatorname{Cl} V \subseteq \operatorname{Cl} W$, we find that $\operatorname{Cl} V \subseteq \operatorname{Cl}(W) \setminus Q$. From construction, we have $\operatorname{Cl}(W) \setminus U \subseteq Q$, which implies in $\operatorname{Cl}(W) \setminus Q \subseteq U$—hence in particular $\operatorname{Cl} V \subseteq U$. ♮

Proposition 14.4.5. Any open or closed subset of a locally compact Hausdorff space is itself locally compact Hausdorff space.

Proof. Let X be a locally compact Hausdorff space. If $U \subseteq X$ is an open set, then by Proposition 14.4.3, since U is also Hausdorff, any point $x \in U$ has a relatively compact neighbourhood of x contained in U—thus U is locally compact Hausdorff space.

Let $C \subseteq X$ be a closed set and let $x \in C$ be any point. Since C is also Hausdorff, we argue analogously that x has a relatively compact neighbourhood K of x contained in C. Since $\operatorname{Cl} K$ is compact, the closed subset $\operatorname{Cl}(K \cap C) \subseteq \operatorname{Cl} K$ is compact. Moreover, since C is closed, $\operatorname{Cl}(K \cap C) \subseteq \operatorname{Cl} C = C$, therefore$K \cap C$ is a relatively compact neighbourhood of x in C—hence C is a locally compact Hausdorff space. ♮

Lemma 14.4.6 (Product of locally compact spaces)**.** Any finite product of locally compact spaces is locally compact.

Proof. Let $\{X_1, \dots, X_n\}$ be a finite collection of locally compact spaces and consider the product space $X := \prod_{j=1}^n X_j$. Let $x \in X$ be any point and consider, for every index $1 \leqslant j \leqslant n$, the projection $\pi_j(x) \in X_j$. Since X_j is locally compact, there exists a compact set $K_j \subseteq X_j$ containing a neighbourhood $U_j \subseteq X_j$ of $\pi_j(x)$. From this we consider the product sets $K := \prod_{j=1}^n K_j$ and $U := \prod_{j=1}^n$. By Theorem 14.2.16 we know that K is compact and by the product topology the set U is open and also a neighbourhood of x. Since $U_j \subseteq K_j$ for all $1 \leqslant j \leqslant n$, it follows that $U \subseteq K$—which proves that X is locally compact. ♮

Theorem 14.4.7 (Baire category theorem)**.** In every locally compact Hausdorff space or complete metric space, each *countable* collection of *dense* open subsets has a *dense intersection*.

Proof. Let X be either a locally compact Hausdorff space or a complete metric space and $\{D_n\}_{n \in \mathbf{N}}$ be a countable collection of dense open subsets of X. Evoking Proposition 12.1.34, it suffices to prove that every non-empty subset $U \subseteq X$ contains a point of $D := \bigcap_{n \in \mathbf{N}} D_n$. We split the cases in two:

1. If X is a locally compact Hausdorff space, we construct a nested sequence of compact sets inductively as follows. From hypothesis, D_0 is dense, thus in particular $U \cap D_0$ contains a point of X, therefore evoking Proposition 14.4.3 we can find a non-empty relatively compact open set $C_0 \subseteq X$ for which $\operatorname{Cl}(C_0) \subseteq U \cap D_0$. We proceed analogously, finding a relatively compact set $C_1 \subseteq X$ such that $\operatorname{Cl}(C_1) \subseteq C_0 \cap D_1 \subseteq U \cap D_0 \cap D_1$. By induction we find a sequence of compact open sets $(\operatorname{Cl}(C_n))_{n \in \mathbf{N}}$ such that $\operatorname{Cl}(C_{n+1}) \subseteq \operatorname{Cl}(C_n)$ and $\operatorname{Cl}(C_n) \subseteq U \cap \bigcap_{j=1}^n D_j$ for every $n \in \mathbf{N}$. By Lemma 14.2.8, there exists a point $x \in \bigcap_{n \in \mathbf{N}} \operatorname{Cl}(C_n) \subseteq U$ and from construction $x \in \bigcap_{n \in \mathbf{N}} D_n$.
2. If X is a complete metric space, we use the inductive argument made above to construct a Cauchy sequence. For every $n \in \mathbf{N}$, we have a non-empty open set $C_{n-1} \cap D_n$—hence we may choose a point x_n and a neighbourhood $B_{\varepsilon_n}(x_n) \subseteq C_{n-1} \cap D_n$ for some $\varepsilon_n > 0$. By choosing for each $n \in \mathbf{N}$ a radius $r_n < \min(\varepsilon_n, 1/n)$,

we construct a sequence of closed balls $(\mathrm{Cl}(B_{r_n}(x_n)))_{n\in\mathbf{N}}$ for which $\mathrm{Cl}(B_{r_n}(x_n)) \subseteq U \cap \bigcap_{j=1}^n D_j$. In the limit $n \to \infty$ the radius r_n of the closed balls goes to zero and $(x_n)_{n\in\mathbf{N}}$ form a Cauchy sequence. Since X is complete, such sequence converges to a point of $U \cap \bigcap_{n\in\mathbf{N}} D_n$.

♮

Manifolds

Definition 14.4.8 (Regular coordinate ball)**.** Let M be an n-manifold. We say that a coordinate ball $B \subseteq M$ is *regular* if there exists a neighbourhood U of $\mathrm{Cl}(B)$ and a topological isomorphism $\phi\colon U \xrightarrow{\simeq} B_{r'}(x)$—for some $r' > 0$ and $x \in \mathbf{R}^n$—such that $\phi(B) = B_r(x)$ and $\phi(\mathrm{Cl}\, B) = \mathrm{Cl}(B_r(x))$ for some $0 < r < r'$.

Corollary 14.4.9. Regular coordinate balls are relatively compact.

Proof. Given a regular coordinate ball B, let $U \supseteq B$ be an open set such that there exists an isomorphism $\phi\colon U \to B_{r'}(x)$. Let $0 < r < r'$ be such that $\phi(B) = B_r(x)$ and $\phi(\mathrm{Cl}\, B) = \mathrm{Cl}(B_r(x))$. In particular, the restriction $\phi|_{\mathrm{Cl}\, B}$ is a topological isomorphism, and since $\mathrm{Cl}(B_r(x))$ is a compact set, so is $\mathrm{Cl}\, B$. ♮

Lemma 14.4.10. Let M be an n-manifold and $B \subseteq M$ be a coordinate ball. If $\phi\colon B \xrightarrow{\simeq} B_{r'}(x)$ is a topological isomorphism—for some $r' > 0$ and $x \in \mathbf{R}^n$—then $\phi^{-1}(B_r(x)) \subseteq M$ is a *regular coordinate ball for all* $0 < r < r'$.

Proof. Let $\phi\colon B \to B_{r'}(x)$ be an isomorphism, and define $U := \phi^{-1}(B_r(x))$ for any $0 < r < r'$. The induced restriction $\phi\colon U \to B_r(x)$ is clearly an isomorphism.

Consider the restriction $\phi^{-1}\colon \mathrm{Cl}(B_r(x)) \to M$. From Corollary 14.2.12 we obtain that ϕ^{-1} is a closed map. From Proposition 12.2.12 we find that

$$\phi^{-1}(\mathrm{Cl}(B_r(x))) = \mathrm{Cl}(\phi^{-1}(B_r(x))) = \mathrm{Cl}(U).$$

Therefore $\phi(\mathrm{Cl}\, U) = \mathrm{Cl}(B_r(x))$, proving that U is a regular coordinate ball. ♮

Proposition 14.4.11. Every manifold without boundary has a *countable basis* of *regular* coordinate balls.

Proof. Let M be any n-manifold. Since M is second countable, let $\{U_j\}_{j\in\mathbf{N}}$ be a countable open cover of M consisting of coordinate neighbourhoods. For each $j \in \mathbf{N}$, let $\phi_j\colon U_j \xrightarrow{\simeq} V_j$ be a topological isomorphism, where $V_j \subseteq \mathbf{R}^n$ is open. For every $x \in V_j$, let $r(x) > 0$ be such that $B_{r(x)}(x) \subseteq V_j$.

Define $\mathcal{B}$ as the collection of all $\phi_j^{-1}(B_r(x)) \subseteq M$ such that $x \in \mathbf{R}^n$ is a point consisting of *rational coordinates*, and $0 < r < r(x)$ is also rational. From Lemma 14.4.10 we find that $\phi_j^{-1}(B_r(x))$ is a regular coordinate ball of M. Since the rationals are countable, for each $j \in \mathbf{N}$, the collection of regular coordinate balls of the form $\phi_j^{-1}(B_r(x))$ described above is countable. Therefore we conclude that $\mathcal{B}$ is a countable set.

Since each ϕ_j is an isomorphism, for every $p \in U_j$, if $\phi(p) = x$, then $\phi_j^{-1}(B_r(x))$ is a neighbourhood of p. Therefore $M = \bigcup_{B \in \mathcal{B}} B$. Moreover, if $B, B' \in \mathcal{B}$ are any intersecting sets, let $p \in B \cap B'$ be a point. Take U_j neighbourhood of p and consider the point $y := \phi_j(p)$. Let $U \subseteq (B \cap B') \cap U_j$ be any neighbourhood of p. Let $0 < r < r(y)$ be any rational number such that $\phi_j(U) \subseteq B_r(y)$, then $\phi_j^{-1}(B_r(y)) \subseteq U \subseteq B \cap B'$ and $\phi_j^{-1}(B_r(y)) \in \mathcal{B}$—which proves that $\mathcal{B}$ is a basis of M. ♮

Definition 14.4.12 (Regular coordinate half-ball)**.** Let M be an n-manifold with boundary. A subset $B \subseteq M$ is a *regular coordinate half-ball* if there exists an open set $U \supseteq \operatorname{Cl}(B)$ and a topological isomorphism $\phi\colon U \to B_{r'}(0) \cap \mathbf{H}^n$, for some $r' > 0$, such that $\phi(B) \subseteq B_r(0) \cap \mathbf{H}^n$ and $\phi(\operatorname{Cl} B) \subseteq \operatorname{Cl}(B_r(0)) \cap \mathbf{H}^n$ for some $0 < r < r'$.

Corollary 14.4.13. Regular coordinate half-balls are relatively compact.

Proof. Follows from the same reasoning of the proof of Corollary 14.4.9. ♮

Proposition 14.4.14. Every manifold with boundary has a countable basis composed of regular coordinate balls and regular coordinate half-balls.

Proof. The proof is analogous to Proposition 14.4.11, one just needs to beware that V_j may be either an open subset of $\mathbf{R}^n$ or an open subset of $\mathbf{H}^n$. ♮

Proposition 14.4.15. Every manifold with or without boundary is locally compact.

Proof. From Corollary 14.4.9 and Proposition 14.4.11 we find that manifolds without boundary are indeed locally compact.

For manifolds with boundary we analogously use Corollary 14.4.9 and Proposition 14.4.14. ♮

14.5 Paracompactness

Definition 14.5.1 (Cover refinement)**.** Let X be a topological space and $\mathcal{U}$ be a cover of X. We say that a cover $\mathcal{A}$ of X is a *refinement* of $\mathcal{U}$ if for each $A \in \mathcal{A}$ there exists $U \in \mathcal{U}$ such that $A \subseteq U$—moreover, $\mathcal{A}$ is said to be an *open* refinement if each $A \in \mathcal{A}$ is open in X.

Corollary 14.5.2. Let $\mathcal{U}$ be an *open cover* of X. If each element of $\mathcal{U}$ intersects only finitely many others, then $\mathcal{U}$ is *locally finite*.

Proof. Let $p \in X$ be any element. Since $\mathcal{U}$ covers X, let $U \in \mathcal{U}$ be a neighbourhood of p. From hypothesis, U intersects only finitely many elements of $\mathcal{U}$, therefore $\mathcal{U}$ is locally finite. ♮

Definition 14.5.3 (Paracompactness)**.** A space X is said to be *paracompact* if for *every* open cover of X there exists al *locally finite open refinement*.

Corollary 14.5.4. Every compact space is paracompact.

Proof. Let X be compact and $\mathcal{U}$ be an open cover. If $C \subseteq \mathcal{U}$ is a finite subcover, then in particular C is a locally finite open refinement of $\mathcal{U}$. ♮

Definition 14.5.5 (Compact exhaustion)**.** Let X be a space. We define a sequence $(K_j)_{j\in\mathbf{N}}$ to be a *exhaustion of X by compact sets* if $X = \bigcup_{j\in\mathbf{N}} K_j$ and, for all $j \in \mathbf{N}$, each set K_j is compact in X, and the sequence satisfies $K_j \subseteq \operatorname{Int} K_{j+1}$.

Lemma 14.5.6. Every second countable, locally compact Hausdorff topological space admits an exhaustion by compact sets.

Proof. Let X be a second countable, locally compact Hausdorff space. Since X is locally compact Hausdorff, there exists a basis consisting of relatively compact open subsets of X. Since X is second countable, one can consider a countable cover of X by relatively compact sets $C := \{C_j\}_{j\in\mathbf{N}}$. We proceed by induction for the construction of each K_j. Define $K_0 := \overline{U}_0$, and assume, for the inductive hypothesis, that we've constructed a sequence $(K_0, \dots, K_{n-1})$ of compact sets satisfying $K_j \subseteq \operatorname{Int} K_{j+1}$ for $j < n-1$. Since K_{n-1} is compact and C covers K_{n-1}, it follows that there exist a finite $k_{n-1} \in \mathbf{N}$ such that $K_{n-1} \subseteq \bigcup_{j=0}^{k_{n-1}} C_j$—in particular, choose k_{n-1} so that $k_{n-1} > n$. We thus define $K_n := \bigcup_{j=0}^{k_{n-1}} \operatorname{Cl}(C_j)$ so that $K_n \subseteq \operatorname{Int} K_{n-1}$. Moreover, since $k_{n-1} > n$ then $U_n \subseteq K_n$. Therefore the sequence $(K_j)_{j\in\mathbf{N}}$ is an exhaustion of X by compact sets. ♮

Theorem 14.5.7. Every second countable, locally compact Hausdorff space is paracompact.

Proof. Let X satisfy the above-mentioned hypothesis, and $\mathcal{U}$ be an open cover of X. By Lemma 14.5.6, let $(K_j)_{j\in\mathbf{N}}$ be an exhaustion of X by compact sets. Define the following collections:

- For all $j \in \mathbf{N}$, let $A_j := K_{j+1} \setminus \operatorname{Int}(K_j)$—which is a closed compact set.
- Define $W_0 := \operatorname{Int} K_2$ and, for each $j > 0$, let $W_j := \operatorname{Int}(K_{j+2}) \setminus K_{j-1}$—forming a collection of open sets.

Notice that $A_j \subseteq W_j$ for all $j \in \mathbf{N}$.

For each point $p \in A_j$, let $U_p \in \mathcal{U}$ be a neighbourhood of p, and define $V_p := U_p \cap W_j$—which is again a neighbourhood of p. The collection $\{V_p\}_{p\in A_j}$ forms an open cover for the compact set A_j, and from compactness there exists a finite subcover, say $\mathcal{V}_j$. The union $\mathcal{V} := \bigcup_{j\in\mathbf{N}} \mathcal{V}_j$ is an open refinement of $\mathcal{U}$.

Consider the open set W_j for each $j \in \mathbf{N}$, notice that since $W_{j+3} = \operatorname{Int}(K_{j+5}) \setminus K_{j+2}$, then W_k for $k > j+3$ does not intersect W_j. On the other hand $W_{j-3} = \operatorname{Int}(K_{j-1}) \setminus K_{j-4}$ also does not intersect W_j and neither do W_k for $k < j-3$. Therefore W_j intersects W_k only for $j-2 \leqslant k \leqslant j+2$—therefore every $V \in \mathcal{V}$ intersects only finitely many members, showing that $\mathcal{V}$ is locally finite. ♮

Corollary 14.5.8. Every manifold with or without boundary is paracompact.

14.6 Normal Spaces

Definition 14.6.1 (Normal space)**.** A topological space X is said to be *normal* if it is *Hausdorff*, and for every pair of *disjoint closed sets* $C, F \subseteq X$ there exists a pair of *disjoint open sets* $U, V \subseteq X$ such that $C \subseteq U$ and $F \subseteq V$.

Proposition 14.6.2. Every compact Hausdorff space is normal.

Proof. Since X is Hausdorff, for each $c \in C$ and $f \in F$ let $C_f, F_f \subseteq X$ be *disjoint* neighbourhoods of, respectively, c and f. Now consider the collections $(C_f)_{f\in F}$ and $(F_f)_{f\in F}$, where the latter forms an open cover of F. Since F is closed, thus compact, let $\mathcal{F}_c := (F_{f_j})_{j=1}^n \subseteq (F_f)_{f\in F}$ be a *finite subcover* of F—and consider the finite intersection $C_c := \bigcap_{j=1}^n C_{f_j}$, which is a neighbourhood of c and is disjoint from every member of $\mathcal{F}_c$.

By the compactness of C, let $\mathcal{C} := (C_{c_j})_{j=1}^m$ be a finite subcover of the open cover $(C_c)_{c\in C}$ of the set C. If we consider the corresponding family of finite covers $(\mathcal{F}_{c_j})_{j=1}^n$ of the set F, we can define $\mathcal{F} := \bigcap_{j=1}^m \mathcal{F}_{c_j}$—which is again a finite open cover of F, since each $\mathcal{F}_{c_j}$ is finite. Now, defining the open sets $U := \bigcup \mathcal{C}$ and $V := \bigcup \mathcal{F}$, we find that $C \subseteq U$ and $F \subseteq V$—moreover, from the construction of the open covers, the sets U and V are also disjoint. ♮

Corollary 14.6.3. Every closed subspace of a normal space is normal.

Definition 14.6.4 (Regular space)**.** A topological space X is said to be *regular* if it is *Hausdorff* and for every *closed set* $C \subseteq X$ and point $a \in X \setminus C$, there exists a pair of *open sets* $U, V \subseteq X$ for which $a \in U$ and $C \subseteq V$.

Lemma 14.6.5. Let X be a Hausdorff space. Then X is *normal* if and only if, for any closed subset $C \subseteq X$ and open set $U \subseteq X$ containing C, there exists a neighbourhood V of C such that $\mathrm{Cl}(V) \subseteq U$.

Proof. If X is normal, we can consider a closed set $C \subseteq X$ and, given an open set $U \subseteq X$ containing C, the closed set $F := X \setminus U$. From definition, there must exist disjoint open sets $V, U' \subseteq X$ for which $C \subseteq V$ and $F \subseteq U'$. Then in particular $\mathrm{Cl}(V)$ and F are disjoint—otherwise, if $y \in F \cap \mathrm{Cl}(V)$ then $y \in U'$, but U' and V are disjoint, yielding a contradiction.

For the converse, let A and B be disjoint closed subsets of X and assume the latter hypothesis. Define the open set $U' := X \setminus B$ containing A. Then one can choose a neighbourhood V of A such that $\mathrm{Cl}(V) \subseteq U'$, therefore $B \subseteq U := X \setminus \mathrm{Cl}(V)$. Therefore U and V form the desired pair of disjoint open sets for normality. ♮

Theorem 14.6.6. Every paracompact Hausdorff space is normal.

Proof. Let X be a paracompact Hausdorff space, and consider a closed subset $A \subseteq X$ and a point $x \in X \setminus A$—we shall prove that X is regular first. Since X is Hausdorff, for each $p \in A$, let U_p and V_p be disjoint neighbourhoods of p and x, respectively. Therefore $(U_p)_{p\in A}$ forms an open cover of A—thus $\mathcal{C}' := (U_p)_{p\in P} \cup \{X \setminus A\}$ is an open cover of X. Since X is paracompact, let $\mathcal{C}$ be a locally finite refinement of $\mathcal{C}'$, and consider the subset $\mathcal{U} \subseteq \mathcal{C}$ of members $U \subseteq \mathcal{C}$ such that $U \subseteq U_p$ for some $p \in A$. The collection $\mathcal{U}$ forms a locally finite open cover of A. Also, given any $U \in \mathcal{U}$, if $U \subseteq U_p$, then V_p is a neighbourhood of x that is disjoint from U—therefore $\mathrm{Cl}(U)$ cannot contain the point x.

Define $L := \bigcup_{U \in \mathcal{U}} U$ and $W := X \setminus \mathrm{Cl}(L)$. From the fact that $\mathcal{U}$ is locally finite, we know that $\mathrm{Cl}(L) = \bigcup_{U \in \mathcal{U}} U$. This shows that x is not contained in L, while W is a neighbourhood of x. Since L and W are disjoint open sets such that $A \subseteq L$ and $x \in W$, this shows that X is regular. For the proof that X is normal, one simply needs to interchange the Hausdorff condition used to build the initial sets by the regularity of X. ♮

Corollary 14.6.7. Every topological manifold is normal.

Theorem 14.6.8 (Urysohn)**.** Let X be a *normal* topological space, and $C, F \subseteq X$ be a pair of *disjoint closed* subsets. Then there exists a *continuous* map $f : X \to [0, 1]$ such that $f|_C = 0$ and $f|_F = 1$.

Chapter 15

Topological Homotopies & The Fundamental Group

15.1 Homotopy

Definition 15.1.1 (Left homotopy in `Top`)**.** Let $f, g\colon X \rightrightarrows Y$ be parallel topological morphisms. We define a *left homotopy* $\eta\colon f \Rightarrow g$ between f and g to be a morphism $\eta\colon X \times I \to Y$ such that the following diagram commutes

$$\begin{tikzcd} X \arrow[r, "i_0"] \arrow[dr, "f"'] & X \times I \arrow[d, "\eta"] & X \arrow[l, "i_1"'] \arrow[dl, "g"] \\ & Y & \end{tikzcd}$$

where $i_0, i_1\colon X \rightrightarrows X \times I$ are parallel morphisms with $x \overset{i_0}{\mapsto} (x, 0)$ and $x \overset{i_1}{\mapsto} (x, 1)$.

If topological spaces are viewed as categories whose objects are its points and morphisms are the identities, then topological morphisms $f, g\colon X \rightrightarrows Y$ are functors between the categories X and Y and a homotopy $\eta\colon f \Rightarrow g$ is a natural transformation between f and g, where the following diagram commutes in the category Y for any two points $x, x' \in X$:

$$\begin{tikzcd} f(x) \arrow[r, "\eta(x,-)"] \arrow[d, "f"'] & g(x) \arrow[d, "g"] \\ f(x') \arrow[r, "\eta(x',-)"'] & g(x') \end{tikzcd}$$

To put concretely, left homotopy between f and g is a continuous map η such that $\eta(-, 0) = f$ and $\eta(-, 1) = g$—which can be visually thought of as a deformation of the morphism f to g. If f and g are indeed homotopic, we denote this by $f \sim_{\mathrm{h}} g$.

Corollary 15.1.2 (Equivalence relation)**.** Left homotopy induces an *equivalence relation* $\sim_{\mathrm{h}}$ on the collection of topological morphisms. Moreover, homotopy respects composition of morphisms.

Proof. First we prove that $\sim_{\mathrm{h}}$ is an equivalence relation:

- (Reflexive) Given a continuous map $f\colon X \to Y$, we explicitly define a homotopy $f \Rightarrow f$ by mapping $(x,t) \mapsto x$ for all $(x,t) \in X \times I$.
- (Symmetric) Let $g\colon X \to Y$ be another continuous map and suppose that there exists a homotopy $\eta\colon f \Rightarrow g$. We can define a homotopy $\eta'\colon g \Rightarrow f$ by mapping $(x,t) \mapsto \eta(x, 1-t)$—so that $\eta'(-,0) = \eta(-,1) = g$ and $\eta'(-,1) = \eta(-,0) = f$.
- (Transitive) Consider yet another morphism $h\colon X \to Y$ and suppose that there exists a homotopy $\sigma\colon g \Rightarrow h$. In order to construct a homotopy $\delta\colon f \Rightarrow h$, we define a X-parametrized path by concatenating η and σ appropriately

  $$\delta(x,t) \coloneqq \begin{cases} \eta(x, 2t), & t \in [0, 1/2], \\ \sigma(x, 2t-1), & t \in [1/2, 1]. \end{cases}$$

  Indeed, $\delta(-,0) = \eta(-,0) = f$ while $\delta(-,1) = \sigma(-,1) = h$. It remains to prove that δ is continuous. Notice that the sets $A \coloneqq X \times [0,1/2]$ and $B \coloneqq X \times [1/2, 1]$ are both closed in the product topology and cover the whole space $X \times I$. By Proposition 13.2.9, since δ is continuous on both A and B then δ is continuous on $A \cup B = X \times I$. We say that δ is the *product* of the homotopies η and σ.

To prove that homotopy preserves composition, consider the following commutative diagram in $\mathtt{Top}$

$$\begin{tikzcd} X \arrow[r, "f"] & Y \arrow[r, shift left, "g"] \arrow[r, shift right, "g'"'] & Z \arrow[r, "h"] & W \end{tikzcd}$$

If there exists a homotopy $\eta\colon g \Rightarrow g'$, we want to show that hgf is homotopy equivalent to $hg'f$. To do that, consider the following commutative diagram

$$\begin{tikzcd} X \arrow[r, "i_0"] \arrow[d, "f"'] & X \times I \arrow[d, "f \times \mathrm{id}_I"] & X \arrow[l, "i_1"'] \arrow[d, "f"] \\ Y \arrow[r, "i_0"] \arrow[dr, "g"'] & Y \times I \arrow[d, "\eta"] & Y \arrow[l, "i_1"'] \arrow[dl, "g'"] \\ & Z \arrow[d, "h"] & \\ & W & \end{tikzcd}$$

From the diagram we see that a natural choice of homotopy $\sigma\colon hgf \Rightarrow hg'f$ is given by $\sigma = h\eta(f \times \mathrm{id}_I)$ thus indeed $hgf \sim_{\mathrm{h}} hg'f$. $\natural$

Given any two topological spaces X and Y, we denote the collection of continuous maps $X \to Y$ *up to homotopy equivalence* by

$$[X, Y] \coloneqq \mathrm{Mor}_{\mathtt{Top}}(X, Y)/\sim_{\mathrm{h}}.$$

We call a morphism $[f] \in [X, Y]$ a *homotopy class*. Moreover, for any three spaces $X, Y, Z \in \mathtt{Top}$, there exists a *unique* compositional map

$$[X, Y] \times [Y, Z] \longrightarrow [X, Z]$$

such that the following diagram commutes

$$\begin{tikzcd}
\mathrm{Mor}_{\mathtt{Top}}(X, Y) \times \mathrm{Mor}_{\mathtt{Top}}(Y, Z) \arrow[r] \arrow[d, two heads] & \mathrm{Mor}_{\mathtt{Top}}(X, Z) \arrow[d, two heads] \\
{[X, Y] \times [Y, Z]} \arrow[r, dashed] & {[X, Z]}
\end{tikzcd}$$

That is, given morphisms $X \xrightarrow{f} Y \xrightarrow{g} Z$, we have a composition of $[f]$ with $[g]$ uniquely defined as

$$[g] \circ [f] \coloneqq [gf].$$

Definition 15.1.3 (Homotopy category)**.** We therefore define a category $\mathtt{Ho}(\mathtt{Top})$ composed of topological spaces—which, viewed as an object of $\mathtt{Ho}(\mathtt{Top})$, is called a *homotopy type*—and classes of continuous morphisms between them up to homotopy.

This quotient operation on the category of topological spaces induces a canonically defined projective functor

$$\kappa \colon \mathtt{Top} \longrightarrow \mathtt{Ho}(\mathtt{Top}).$$

Definition 15.1.4 (Homotopy equivalence)**.** Let X and Y be topological spaces. We say that a continuous map $f \colon X \to Y$ is a *homotopy equivalence* of X and Y if there exists a continuous map $g \colon Y \to X$ and homotopies $fg \Rightarrow \mathrm{id}_Y$ and $gf \Rightarrow \mathrm{id}_X$.

If there exists such homotopy equivalence, we write that $X \simeq_{\mathrm{h}} Y$—it is to be noted that homotopy equivalences are exactly the isomorphisms in the homotopy category $\mathtt{Ho}(\mathtt{Top})$.

Corollary 15.1.5. Every topological isomorphism is a homotopy equivalence.

Proof. Let $f \colon X \xrightarrow{\simeq} Y$ be a topological isomorphism. We consider its image under the functor $\mathtt{Top} \to \mathtt{Ho}(\mathtt{Top})$, which we'll name $[f] \colon X \to Y$. Then since $f^{-1} \colon Y \xrightarrow{\simeq} X$ is also a continuous morphism, we can consider its class $[f^{-1}]$ and notice that $[f][f^{-1}] = [ff^{-1}] = [\mathrm{id}_Y]$ and $[f^{-1}][f] = [f^{-1}f] = [\mathrm{id}_X]$. Therefore, there exists two homotopies $ff^{-1} \Rightarrow \mathrm{id}_Y$ and $f^{-1}f \Rightarrow \mathrm{id}_X$. ♮

Corollary 15.1.6. Homotopy equivalence is an equivalence relation on the class of topological spaces.

Definition 15.1.7 (Relative homotopy)**.** Let X be a space and $A \subseteq X$ be a subspace. If $f, g \colon X \rightrightarrows Y$ are parallel continuous maps such that $f|_A = g|_A$, a *homotopy between f and g relative to A*—if existent—is a homotopy $\eta \colon f \Rightarrow_{\mathrm{rel}\, A} g$ such that

$$\eta(a, t) = f(a) = g(a) \quad \text{for all } a \in A \text{ and } t \in I.$$

If f and g are homotopic relative to A, we shall also denote this by $f \sim_{\mathrm{rel}\, A} g$.

Contractibility

Definition 15.1.8 (Null-homotopy)**.** A morphism of topological spaces $X \to Y$ is said to be *null-homotopic* if it is homotopic to a *constant map*. A homotopy between a morphism and a constant morphism is called a *null-homotopy*. In particular, a null-homotopy of the identity map $\mathrm{id}\colon X \to X$ is a *contraction* of X.

Proposition 15.1.9. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be continuous maps. If either f or g is null-homotopic, then $gf\colon X \to Z$ is null-homotopic.

Proof. Let f be null-homotopic and $\eta\colon f \Rightarrow c_p$ be a homotopy between f and the constant map $c_p\colon X \to Y$ on $p \in Y$. The composition $g\eta\colon X \times I \to Z$ is a continuous map such that $g\eta(-,0) = gf$ and $g\eta(-,1) = gc_p$, where $gc_p = \overline{c}_{g(p)}$ is the constant map on $g(p) \in Z$. Therefore $g\eta$ is a null-homotopy of gf.

If g is null-homotopic with a homotopy $\varepsilon\colon g \Rightarrow c_z$ for some constant map $c_z\colon Y \to Z$ on $z \in Z$. Notice that the map $\varepsilon \circ (f \times \mathrm{id}_I)\colon X \times I \to Z$ is continuous and for all $x \in X$ we have

$$\begin{aligned}
\varepsilon \circ (f \times \mathrm{id}_I)(x,0) &= \varepsilon(f(x),0) = gf(x), \\
\varepsilon \circ (f \times \mathrm{id}_I)(x,1) &= \varepsilon(f(x),1) = c_z f(x) = z.
\end{aligned}$$

Therefore establishing a null-homotopy $gf \sim_{\mathrm{h}} \overline{c}_z$—where $\overline{c}_z\colon X \to Z$ is constant on z. ♮

Proposition 15.1.10. A continuous map $f\colon S^n \to Y$ is null-homotopic if and only if there exists a continuous map $F\colon D^{n+1} \to Y$ such that $F|_{S^n} = f$.

Proof. First we prove that $\mathrm{Cone}(S^n) \simeq D^{n+1}$. Let $\phi\colon \mathrm{Cone}(S^n) \to D^{n+1}$ to be the map given by $[x,t] \mapsto tx$, then $\|tx\| = t\|x\| \leqslant 1$ and $\operatorname{im}\phi \subseteq D^{n+1}$. Moreover, the map is well defined, since $\phi[x,0] = 0$—which is the only uncertain region of the cone. Since the quotient of a compact space is compact, the space $\mathrm{Cone}(S^n)$ is compact—moreover, D^{n+1} is Hausdorff. Notice that ϕ establishes a bijection between a compact space and a Hausdorff space, thus ϕ is an isomorphism.

Suppose f is null-homotopic to a continuous map $c_p\colon S_n \to Y$, for some point $p \in Y$. Let $\eta\colon c_p \Rightarrow f$ be an inverse null-homotopy for f. Define F via the isomorphism $\mathrm{Cone}(S^n) \simeq D^{n+1}$ as $F(tx) := \eta(x,t)$, where $x \in S^n$ and $t \in I$. From this, for every $s \in S^n$ we get $F(s) = \eta(s,1) = f(s)$, and $F(0) = p$.

For the converse, let $F\colon D^{n+1} \to Y$ be an extension of f. Define a map $\eta\colon X \times I \to Y$ given by a collection of maps $(\eta_t\colon S^n \to Y)_{t \in I}$, where $\eta_t(x) := F(tx)$—which is a continuous map—so that the following diagram commutes in `Top` for all $t \in I$:

$$\begin{tikzcd}
S^n \arrow[r, "\pi_t"] \arrow[dr, "\eta_t"'] & \mathrm{Cone}(S^n) \arrow[r, "\simeq"] & D^{n+1} \arrow[dl, "F"] \\
& Y &
\end{tikzcd}$$

where π_t is the continuous map $x \mapsto [x, t]$. From this we find that $\eta_0(x) = F(0)$ is constant and $\eta_1(x) = F(x) = f(x)$ since $x \in S^n$.

It just remains to be shown that η is continuous. If $U \subseteq Y$ is any open set, let $(x_0, t_0) \in \eta^{-1}(U)$ be any pair. Since F is continuous, then there exists $V \subseteq F^{-1}(U)$ neighbourhood of the point $t_0 x_0 \in D^{n+1}$. Since ϕ is an isomorphism, there exists an open set $P' \times T \subseteq \mathrm{Cone}(S^n)$ that is a neighbourhood of the class $[x_0, t_0]$, and $P' \times T \subseteq \phi^{-1}(V)$. In particular, we know that η_{t_0} is continuous, therefore there exists a neighbourhood $P \subseteq \eta_{t_0}^{-1}(P' \times T)$ of x_0 in S^n.

If $(x, t) \in P \times T$ is any point, then from construction we have $F(tx) \in U$—thus $\eta(x, t) \in U$, which shows that $P \times T \subseteq \eta^{-1}(U)$ is a neighbourhood of (x_0, t_0), hence $\eta^{-1}(U)$ is open. ♮

Definition 15.1.11 (Contractible space)**.** A space X is said to be *contractible* if the unique continuous map $X \to *$ is a homotopy equivalence. From its construction, contractible spaces are the terminal objects of $\mathtt{Ho(Top)}$. A null-homotopy of id_X is said to be a *contraction* of X.

Example 15.1.12 (A ball is contractible)**.** Consider the open (or closed) ball $B^n \subseteq \mathbf{R}^n$. Let $f\colon B^n \to *$ be the unique continuous map from the ball to the point space, and $\iota\colon * \to B^n$ be the map $* \mapsto 0$. We define a homotopy $\eta\colon \iota f \Rightarrow \mathrm{id}_{B^n}$ given by $(p, t) \mapsto tp$, and a homotopy $\sigma\colon f\iota \Rightarrow \mathrm{id}_*$ mapping $(*, t) \mapsto *$, since $f\iota = \mathrm{id}_*$. Therefore B^n is contractible.

Furthermore, since $B^n \simeq \mathbf{R}^n$ in $\mathtt{Top}$, it follows that $\mathbf{R}^n \simeq_{\mathrm{h}} B^n$ and thus $\mathbf{R}^n \simeq_{\mathrm{h}} *$—the euclidean space is contractible.

Proposition 15.1.13. A space X is contractible if and only if id_X is null-homotopic.

Proof. Suppose X is contractible, then $f\colon X \to *$ is a homotopy equivalence. Let $g\colon * \to X$ be a homotopy inverse of f, and let $x_0 \in X$ be the image of g—that is $g(*) = x_0$ and hence $gf\colon X \to X$ is a constant map $gf(x) = x_0$. Therefore id_X is homotopic to the constant map gf—and the choice of x_0 was arbitrary.

For the converse, let $x_0 \in X$ be any point and let $\eta\colon \mathrm{id}_X \Rightarrow c_{x_0}$ be a homotopy—where $c_{x_0}\colon X \to X$ is the constant map $x \mapsto x_0$. Consider the unique continuous map $f\colon X \to *$ and let $g\colon * \to X$ be given by $g(*) = x_0$, then $gf = c_{x_0} \sim_{\mathrm{h}} \mathrm{id}_X$ via η and $fg = * = \mathrm{id}_*$. Therefore f is a homotopy equivalence. ♮

Proposition 15.1.14. If X is contractible, then it is path connected.

Proof. Let $x, y \in X$ be any two points. Since X is contractible, then id_X is null-homotopic—hence there exists a constant map $c_p\colon X \to X$ and a homotopy $\eta\colon \mathrm{id}_X \Rightarrow c_p$. Notice that by definition $\eta(x, 0) = x$ and $\eta(x, 1) = p$, therefore $\eta(x, -)\colon I \to X$ is a path from x to p—and analogously, $\eta(y, -)\colon I \to X$ is a path from y to p. To construct a path from x to y we can simply concatenate the path $\eta(x, -)$ with the inverse path $\eta^{-1}(y, -)$, which is given by $\eta^{-1}(y, t) = \eta(y, 1 - t)$. Concretely, let $\gamma\colon I \to X$ be defined by

$$\gamma(t) := \begin{cases} \eta(x, 2t), & \text{if } t \in [0, 1/2] \\ \eta(y, 2 - 2t), & \text{if } t \in [1/2, 1] \end{cases}$$

which is continuous by Proposition 13.2.9. Then γ is a path from x to y. ♮

Remark 15.1.15. The converse of Proposition 15.1.14 is not true in general, for instance, the sphere S^2 is path connected, but not contractible.

Corollary 15.1.16. If X is contractible and Y is path connected, then every pair of parallel morphisms $X \rightrightarrows Y$ is homotopic, in short, $[X, Y]$ has a single point.

Proof. Let $f, g\colon X \rightrightarrows Y$ be topological morphisms. Since X is contractible, let $\eta\colon \mathrm{id}_X \Rightarrow c_p$ be a homotopy from the identity to the constant function $c_p\colon X \to X$ at the point $p \in X$. Notice that $f\eta(-,0) = f\,\mathrm{id}_X = f$ and $f\eta(-,1) = f c_p = c_{f(p)}$, where $\overline{c}_{f(p)}\colon X \to Y$ is the map constant on $f(p)$. Given any point $y \in Y$, the collection of paths $\mathrm{Path}_Y(f(p), y)$ is non-empty, thus one can define a collection $(\gamma_x)_{x\in X}$ of paths $\gamma_x \in \mathrm{Path}_Y(f(p), g(x))$. Define a map $\varepsilon\colon X \times I \to Y$ given by

$$\varepsilon(x,t) := \begin{cases} f\eta(x,t), & \text{if } t \in [0, 1/2] \\ \gamma_x(2t-1), & \text{if } t \in [1/2, 1] \end{cases}$$

then ε is continuous by Proposition 13.2.9 and defines a homotopy $\varepsilon\colon f \Rightarrow g$. ♮

Corollary 15.1.17. Let Y be a contractible space. Then for every topological space X any pair of parallel morphisms $X \rightrightarrows Y$ are homotopic, that is, $[X, Y]$ has a single point.

Proof. Let $f, g\colon X \rightrightarrows Y$ be any two parallel morphisms. Since Y is contractible, there exists $p \in Y$ such that $\eta\colon \mathrm{id}_Y \Rightarrow c_p$, where $c_p\colon Y \to Y$ is the constant map on p. Notice that the composition $\eta \circ (f \times \mathrm{id}_I)$ is a continuous map such that, for all $x \in X$ we have

$$\begin{aligned} \eta \circ (f \times \mathrm{id}_I)(x, 0) &= \eta(f(x), 0) = \mathrm{id}_Y(f(x)) = f(x), \\ \eta \circ (f \times \mathrm{id}_I)(x, 1) &= \eta(f(x), 1) = c_p(f(x)) = p. \end{aligned}$$

Therefore $\eta \circ (f \times \mathrm{id}_I)\colon X \times I \to Y$ defines a homotopy $f \sim_{\mathrm{h}} c_p f$ where $c_p f\colon X \to Y$ is merely the constant map $x \mapsto p$, which we'll call $\overline{c}_p\colon X \to Y$. Now, taking the inverse homotopy $(\eta)^{-1}(y,t) = \eta(y, 1-t)$ we obtain a homotopy $c_p \sim_{\mathrm{h}} \mathrm{id}_Y$. We can then consider the continuous map $\eta^{-1} \circ (g \times \mathrm{id}_I)\colon X \times I \to Y$

$$\begin{aligned} \eta^{-1} \circ (g \times \mathrm{id}_I)(x, 0) &= \eta^{-1}(g(x), 0) = c_p(g(x)) = p, \\ \eta^{-1} \circ (g \times \mathrm{id}_I)(x, 1) &= \eta^{-1}(g(x), 1) = \mathrm{id}_Y(g(x)) = g(x). \end{aligned}$$

therefore $\eta^{-1} \circ (g \times \mathrm{id}_I)$ is a homotopy $c_p g = \overline{c}_p \sim_{\mathrm{h}} \mathrm{id}_Y$. Therefore the concatenation of homotopies $\varepsilon\colon X \times I \to Y$ given by

$$\varepsilon(x,t) := \begin{cases} \eta \circ (f \times \mathrm{id}_I)(x, 2t), & \text{if } t \in [0, 1/2] \\ \eta^{-1} \circ (g \times \mathrm{id}_I)(x, 2t-1), & \text{if } t \in [1/2, 1] \end{cases}$$

defines a homotopy $f \sim_{\mathrm{h}} \overline{c}_p \sim_{\mathrm{h}} g$ as wanted. ♮

15.2 Mapping Spaces

Compact-Open Topology

Definition 15.2.1 (Evaluation map)**.** Let X and Y be topological spaces. We define the *evaluation map* on X and Y to be the set-function:

$$\mathrm{eval}\colon \mathrm{Mor}_{\mathsf{Top}}(X,Y)\times X \longrightarrow Y \quad \text{mapping} \quad (f,x)\longmapsto f(x).$$

Definition 15.2.2 (Admissible topology)**.** A topology on the set of continuous maps $\mathrm{Mor}_{\mathsf{Top}}(X,Y)$ is said to be *admissible* if the evaluation map $\mathrm{eval}\colon \mathrm{Mor}_{\mathsf{Top}}(X,Y)\times X\to Y$ is *continuous*.

Definition 15.2.3 (Compact-open topology)**.** Let X and Y be topological spaces. A pair (K,U)—composed of a compact set $K\subseteq X$ and an open set $U\subseteq Y$—defines a a collection of continuos maps

$$\mathrm{co}(K,U) := \{f\in \mathrm{Mor}_{\mathsf{Top}}(X,Y) : f(K)\subseteq U\}$$

on the space $\mathrm{Mor}_{\mathsf{Top}}(X,Y)$. We define the *compact-open topology* on $\mathrm{Mor}_{\mathsf{Top}}(X,Y)$ to be the topology whose subbasis is the collection of sets $\mathrm{co}(K,U)$ for each $K\subseteq X$ compact and $U\subseteq Y$ open. We'll denote the compact-open topology of $\mathrm{Mor}_{\mathsf{Top}}(X,Y)$ by $\mathrm{co}(X,Y)$.

Proposition 15.2.4. The compact-open topology in $\mathrm{Mor}_{\mathsf{Top}}(X,Y)$ is the coarsest between the admissible topologies in the mapping space.

Proof. Suppose τ is an admissible topology for $\mathrm{Mor}_{\mathsf{Top}}(X,Y)$. We show that for any $\mathrm{co}(K,U)\in\mathrm{co}(X,Y)$, we have $\mathrm{co}(K,U)\in\tau$. Let $k\in K$ be any point and $f\in\mathrm{co}(K,U)$ be any map with $f(K)\subseteq U$. From definition $\mathrm{eval}(f,k)=f(k)\in U$—moreover, since eval is continuous, then $\mathrm{eval}^{-1}(U)\subseteq \mathrm{Mor}_{\mathsf{Top}}(X,Y)\times X$ is open—there must exist a neighbourhood $V_k\subseteq(\mathrm{Mor}_{\mathsf{Top}}(X,Y),\tau)$ of f and a neighbourhood $W_k\subseteq K$ of k such that

$$\mathrm{eval}(V_k\times W_k)\subseteq U.$$

The collection $(W_k)_{k\in K}$ is a cover for K—and since K is compact, let $\{W_{k_1},\dots,W_{k_n}\}$ be a finite subcover. Therefore, for all $1\leqslant j\leqslant n$ we have $\mathrm{eval}(V_{k_j}\times W_{k_j})\subseteq U$. Since open sets are closed under finite intersections, consider the neighbourhood $V := V_{k_1}\cap\cdots\cap V_{k_n}$ of f. Let $k\in K$ be any point, then there must exist $1\leqslant j\leqslant n$ such that $k\in W_{k_j}$. Therefore if $g\in V$ is any map, we have

$$g(k)=\mathrm{eval}(g,k)\in\mathrm{eval}(V\times W_{k_j})\subseteq \mathrm{eval}(V_{k_j}\times W_{k_j})\subseteq U,$$

hence $g(K)\subseteq U$—thus we conclude that $g\in\mathrm{co}(K,U)$ and $V\subseteq\mathrm{co}(K,U)$ is an open subset containing f. This shows that $\mathrm{co}(K,U)$ is open in $(\mathrm{Mor}_{\mathsf{Top}}(X,Y),\tau)$, therefore $\mathrm{co}(K,U)\in\tau$, which proves the proposition. ♮

Proposition 15.2.5. If X is a *locally compact Hausdorff* space, then the compact-open topology on $\mathrm{Mor}_{\mathsf{Top}}(X,Y)$ is *admissible*.

Proof. Let $U \subseteq Y$ be any open set, if $\mathrm{eval}^{-1}(U)$ is non-empty, let $(f, x) \in \mathrm{eval}^{-1}(U)$ be any pair. Since f is continuous, then $f^{-1}(U)$ is an open set and there must exist a neighbourhood $W \subseteq X$ of x such that $W \subseteq f^{-1}(U)$—thus $f(W) \subseteq U$. Since X is locally compact Hausdorff, then W is also locally compact Hausdorff and therefore there must exist a relatively compact $V \subseteq X$ neighbourhood of x such that $V \subseteq \mathrm{Cl}(V) \subseteq W$—so that $L := \mathrm{co}(\mathrm{Cl}(V), U) \times V$ is a neighbourhood of the pair (f, x). In fact, if $(g, y) \in L$ is any pair, then $g(y) \in U$, therefore $(g, y) \in \mathrm{eval}^{-1}(U)$. This implies in $L \subseteq \mathrm{eval}^{-1}(U)$ and proves that $\mathrm{eval}^{-1}(U)$ is open in $\mathrm{Mor}_{\mathsf{Top}}(X, Y) \times X$. ♮

Corollary 15.2.6. When X is locally compact Hausdorff, the compact-open topology is the coarsest admissible topology on $\mathrm{Mor}_{\mathsf{Top}}(X, Y)$.

Exponential Objects

Definition 15.2.7 (Adjoint map)**.** Let $f : X \times Y \to Z$ be a topological morphism. We define the *adjoint map* of f to be the continuous map $f^\wedge : X \to \mathrm{Mor}_{\mathsf{Top}}(Y, Z)$ given by $f^\wedge(x)(y) := f(x, y)$—this adjoint map is the *currying of f*.

Proposition 15.2.8. Given a continuous map $f : X \times Y \to Z$, the adjoint map $f^\wedge : X \to \mathrm{Mor}_{\mathsf{Top}}(Y, Z)$ is continuous and, for every $x \in X$, the map $f^\wedge(x) : Y \to Z$ is continuous.

Proof. Let $\mathrm{co}(K, U) \subseteq \mathrm{Mor}_{\mathsf{Top}}(Y, Z)$ be an open set in the compact-open topology. Let $f^\wedge(x) \in \mathrm{co}(K, U)$ for some $x \in X$—hence $f(\{x\} \times K) \subseteq U$. From tube lemma (see Lemma 14.2.13) there exists $V \subseteq X$ neighbourhood of x such that $V \times K \subseteq f^{-1}(U)$. Therefore $f^\wedge(V) \subseteq \mathrm{co}(K, U)$, which shows that $f^\wedge(\mathrm{co}(K, U)) \subseteq X$ is an open set. Since this is the case for any element of the subbasis of $\mathrm{Mor}_{\mathsf{Top}}(Y, Z)$, it follows that $f^\wedge$ is continuous.

The last assertion is trivial since $f^\wedge(x) = f\iota_x$, where $\iota_x : Y \hookrightarrow X \times Y$ is the injective map $y \mapsto (x, y)$, which is continuous—therefore $f^\wedge(x)$ is a continuous map. ♮

The notion of an adjoint map associated with the space $\mathrm{Mor}_{\mathsf{Top}}(X \times Y, Z)$ gives rise to a set-function

$$\begin{aligned} \mathrm{curry} : \mathrm{Mor}_{\mathsf{Top}}(X \times Y, Z) &\longrightarrow \mathrm{Mor}_{\mathsf{Top}}(X, \mathrm{Mor}_{\mathsf{Top}}(Y, Z)) \\ f &\longmapsto f^\wedge. \end{aligned}$$

Dual to *currying* is the notion of *uncurrying*, which is a set-function

$$\begin{aligned} \mathrm{uncurry} : \mathrm{Mor}_{\mathsf{Top}}(X, \mathrm{Mor}_{\mathsf{Top}}(Y, Z)) &\longrightarrow \mathrm{Mor}_{\mathsf{Top}}(X \times Y, Z) \\ \phi &\longmapsto \phi^\vee := \mathrm{eval}_{Y,Z} \circ (\phi \times \mathrm{id}_Y), \end{aligned}$$

that is, $\phi^\vee(x, y) := \phi(x)(y)$.

Proposition 15.2.9. Let X, and Z be any two spaces and Y be a locally compact Hausdorff space. Given a continuous map $g : X \to \mathrm{Mor}_{\mathsf{Top}}(Y, Z)$, the uncurried map $g^\vee : X \times Y \to Z$ is continuous.

Proof. Let $U \subseteq Z$ be an open set, and let $(x,y) \in (g^\vee)^{-1}(U)$ be any point. Since $g(x) \in \mathrm{Mor}_{\mathsf{Top}}(Y,Z)$ and $y \in (g(x))^{-1}(U)$, there must exist a neighbourhood $W \subseteq Y$ of y such that $W \subseteq (g(x))^{-1}(U)$. Since Y is locally compact Hausdorff, the open set W is also locally compact Hausdorff and therefore there exists a relatively compact neighbourhood $V \subseteq W$ of y with $V \subseteq \mathrm{Cl}(V) \subseteq W$. Thus $g(x)(\mathrm{Cl}(V)) \subseteq U$, therefore $g(x) \in \mathrm{co}(\mathrm{Cl}(V), U)$—that is $g(x)$ is *open* in $\mathrm{Mor}_{\mathsf{Top}}(Y,Z)$.

Notice that since g is continuous, there exists a neighbourhood $Q \subseteq X$ of x such that $g(Q) \subseteq \mathrm{co}(\mathrm{Cl}(V), U)$. Moreover the neighbourhood $T \times V \subseteq X \times Y$ of (x,y) is contained in $(g^\vee)^{-1}(U)$, since for any $(a,b) \in T \times V$ we have $g^\vee(a,b) = g(a)(b) \in U$. Therefore we've shown that $(g^\vee)^{-1}(U) \subseteq X \times Y$ is an open set, which proves that $g^\vee$ is continuous. ♮

Using Proposition 15.2.8 and Proposition 15.2.9, we have proven the following corollary.

Corollary 15.2.10. Let X, Y and Z be topological spaces. If Y is a locally compact Hausdorff space, then currying is a *set bijection* $\mathrm{Mor}_{\mathsf{Top}}(X \times Y, Z) \simeq \mathrm{Mor}_{\mathsf{Top}}(X, \mathrm{Mor}_{\mathsf{Top}}(Y,Z))$.

Theorem 15.2.11. Let X, Y and Z be topological spaces. If both X and Y are Hausdorff and Y is also locally compact, then

$$\mathrm{Mor}_{\mathsf{Top}}(X \times Y, Z) \xrightarrow[\text{curry}]{\simeq} \mathrm{Mor}_{\mathsf{Top}}(X, \mathrm{Mor}_{\mathsf{Top}}(Y,Z)).$$

is a *topological isomorphism*.

Proof. By Corollary 15.2.10 it is sufficient to prove that curry and uncurry are continuous maps. Under the compact-open topology we know that an open set of the *subbasis* of $\mathrm{Mor}_{\mathsf{Top}}(X, \mathrm{Mor}_{\mathsf{Top}}(Y,Z))$ is of the form $\mathrm{co}(K,V)$, for $K \subseteq X$ compact and $V \subseteq \mathrm{Mor}_{\mathsf{Top}}(Y,Z)$ open—that is, there exists an open set $U \subseteq Z$ and compact subset $L \subseteq Y$ such that $V = \mathrm{co}(L,U)$. Therefore $\mathrm{co}(K,V) = \mathrm{co}(K, \mathrm{co}(L,U))$.

Consider any such open set $W := \mathrm{co}(K, \mathrm{co}(L,U)) \subseteq \mathrm{Mor}_{\mathsf{Top}}(X, \mathrm{Mor}_{\mathsf{Top}}(Y,Z))$. Since K and L are compact, then $K \times L \subseteq X \times Y$ is a compact set. If $f \in \mathrm{curry}^{-1}(W)$, then $f(K \times L) = f^\wedge(K)(L) \subseteq U$. This implies that $\mathrm{co}(K \times L, U) \subseteq \mathrm{curry}^{-1}(W)$ is a neighbourhood of f—therefore $\mathrm{curry}^{-1}(W)$ is open in $\mathrm{Mor}_{\mathsf{Top}}(X \times Y, Z)$. From this we conclude that curry is a continuous map.

Consider now any open set $\mathrm{co}(Q,U)$ of the subbasis of $\mathrm{Mor}_{\mathsf{Top}}(X \times Y, Z)$. Define $K := \pi_X(Q)$ and $L := \pi_Y(Q)$, with $Q \subseteq K \times L$—which are compact sets since the image of a compact set under a continuous map is compact. Let $g \in \mathrm{uncurry}^{-1}(\mathrm{co}(Q,U))$ be a curried map. If $(x,y) \in Q$, then $g(x)(y) = g^\vee(x,y) \in U$—which implies that $\mathrm{co}(K, \mathrm{co}(L,U)) \subseteq \mathrm{uncurry}^{-1}(\mathrm{co}(Q,U))$ is a neighbourhood of g. ♮

Proposition 15.2.12. Let Z be a locally compact topological space, and $p \colon X \to Y$ be a quotient map. Then the product morphism

$$p \times \mathrm{id}_Z \colon X \times Z \to Y \times Z$$

is a quotient map.

Proof. Let $h: Y \times Z \to W$ be a set-function—where W is a topological space—such that the composition $h \circ (p \times \mathrm{id}_Z)$ is a continuous map. From Proposition 15.2.8 we find that the adjoint $(h \circ (p \times \mathrm{id}_Z))^\wedge = h^\wedge p$ is continuous. Since p is a quotient map, then from the universal property $h^\wedge$ is continuous. Now, since Z is locally compact, the uncurried h is continuous by Proposition 15.2.9. ♮

Linear Homotopy

Definition 15.2.13 (Star-shaped spaces)**.** A subspace $A \subseteq \mathbf{R}^n$ is said to be *star-shaped* with respect to a point $x \in A$ if, for all $y \in A$, the line path $(1-t)x + ty$ is contained in A for all $t \in [0,1]$. A set $C \subseteq \mathbf{R}^n$ is said to be *convex* if and only if it is star-shaped with respect to each of its points.

Proposition 15.2.14. Let $p: X \to Y$ be a quotient morphism of topological spaces. If $\eta: Y \times I \to Z$ is a set-function between topological spaces such that the map $\varepsilon: X \times I \to Z$ given by $(x,t) \mapsto \eta(p(x),t)$ is a homotopy, then η is a homotopy.

Proof. Since p is a quotient map and I is compact (hence locally compact), we know from Proposition 15.2.12 that $p \times \mathrm{id}_I: X \times I \to Y \times I$ is a quotient map. Therefore, since $\eta \circ (p \times \mathrm{id})$ is continuous—since ε is a homotopy— we find that η is a continuous map, thus a homotopy. ♮

Definition 15.2.15. Let X be a space and Y be a star-shaped space. We say that a homotopy $\ell: X \times I \to Y$ is *linear* if there exists continuous maps $f, g: X \rightrightarrows Y$ such that

$$\ell(x,t) = (1-t)f(x) + tg(x).$$

15.3 Retractions & Deformations

Retractions & Cofibrations

Definition 15.3.1 (Retract)**.** Let X be a topological space. A subspace $A \subseteq X$ is said to be a *retract* of X if the inclusion map $\iota: A \hookrightarrow X$ is a split monomorphism in $\mathtt{Top}$—that is, there exists a continuous map $r: X \to A$ such that $r\iota = \mathrm{id}_A$. We call r a *retraction of X to A*.

A subspace $B \subseteq X$ is said to be a *weak retract* of X if $\iota: B \hookrightarrow X$ is a split monomorphism in $\mathtt{Ho}(\mathtt{Top})$—that is, there exists a continuous map $r: X \to A$ such that $r\iota \sim_{\mathrm{h}} \mathrm{id}_A$.

Definition 15.3.2 (Homotopy extension property)**.** Let X and Y be topological spaces, and $A \subseteq X$ be a subspace. The pair (X, A) is said to have the *homotopy extension property with respect to Y* if, given continuous maps $g: X \to Y$ and $\varepsilon: A \times I \to Y$—such that $g(a) = \varepsilon(a, 0)$ for all $a \in A$—there exists a continuous map $\eta: X \times I \to Y$ such that $\eta(x,0) = g(x)$ for all $x \in X$, and $\eta|_{A \times I} = \varepsilon$—that is, η is an extension of ε. This can all be summarized by the following commutative diagram in Top:

$$\begin{tikzcd}
A \times 0 \arrow[rr, hook] \arrow[dd, hook] & & A \times I \arrow[dl, "\varepsilon"'] \arrow[dd, hook] \\
& Y & \\
X \times 0 \arrow[ur, "g"] \arrow[rr, hook] & & X \times I \arrow[ul, "\eta"']
\end{tikzcd}$$

Corollary 15.3.3 (Extending morphisms)**.** Let (X, A) have the homotopy extension property with respect to Y, and consider parallel morphisms $f, f' \colon A \rightrightarrows Y$. If f is homotopic to f', and f has an *extension* to X, then f' also admits an extension to X.

Proof. Let $g \colon X \to Y$ be an extension of f—that is, $g|_A = f$— and $\varepsilon \colon f \Rightarrow f'$ be a homotopy then, from the homotopy extension property of the pair (X, A), there exists a continuous map $\eta \colon X \times I \to Y$ which is an *extension* of ε. From this we conclude that $\eta(-, 1) \colon X \to Y$ is an extension of f'. ♮

Definition 15.3.4 (Cofibration)**.** Let $f \colon Z \to X$ be a morphism between topological spaces. We say that f is a *cofibration* if, for any topological space Y, and continuous maps $g \colon X \to Y$ and $\varepsilon \colon Z \times I \to Y$ such that $g(f(z)) = \varepsilon(z, 0)$, for all $z \in Z$, there exists a continuous map $\eta \colon X \times I \to Y$ such that $\eta(x, 0) = g(x)$ for all $x \in X$, and $\eta(f(z), t) = \varepsilon(z, t)$ for all $z \in Z$ and $t \in I$. This is summarized by the following commutative diagram in Top:

$$\begin{tikzcd}
Z \times 0 \arrow[rr, hook] \arrow[dd, "f \times \mathrm{id}_0"'] & & Z \times I \arrow[dl, "\varepsilon"'] \arrow[dd, "f \times \mathrm{id}_I"] \\
& Y & \\
X \times 0 \arrow[ur, "g"] \arrow[rr, hook] & & X \times I \arrow[ul, "\eta"']
\end{tikzcd}$$

Example 15.3.5. The inclusion map $A \hookrightarrow X$ is a cofibration if and only if the pair (X, A) has the homotopy extension property every space.

Theorem 15.3.6 (Weakness of retracts)**.** Let X be a topological space and $A \subseteq X$ be a subspace. If the pair (X, A) has the *homotopy extension property with respect to* A, then A is a *weak retract* of X if and only if A is a *retract* of X.

Proof. We know that if A is a retract of X then in particular A is a weak retract. We now prove the converse. Let $r \colon X \to A$ be a weak retraction of X to A. Since $r\iota \sim_{\mathrm{h}} \mathrm{id}_A$, let $\varepsilon \colon r\iota \Rightarrow \mathrm{id}_A$ be a homotopy. Since (X, A) has the homotopy extension property with respect to A, it follows that there exists a map $\eta \colon X \times I \to A$ extending ε—that is, since $\varepsilon(a, 0) = r(a)$ for all $a \in A$, then $\eta(x, 0) = r(x)$ for all $x \in X$. If we define $r' \colon X \to A$ to be the map $r' := \eta(-, 1)$, then r' is a retraction of X to A—since $r'|_A = \mathrm{id}_A$— and η establishes a homotopy $r \sim_{\mathrm{h}} r'$. ♮

Deformations

Definition 15.3.7 (Deformation)**.** Let X be a space and $A \subseteq X$ be a subspace. We define a *deformation of A in X* to be a homotopy

$$\delta \colon A \times I \longrightarrow X \quad \text{such that} \quad \delta(-,0) = \mathrm{id}_A .$$

If $\delta(A \times 1) \subseteq B$, for some subspace B of X, then δ is said to be a *deformation of A into B*—and A is said to be *deformable* in X into B.

In particular, the space X is called *deformable* if there exists a subspace $A \subseteq X$ such that X is deformable in itself into A—hence, contractibility of X is equivalent to X being deformable into a point.

Lemma 15.3.8. A space X is *deformable* into a subspace A if and only if the inclusion map $A \hookrightarrow X$ admits a *right homotopy inverse*.

Proof. Suppose that $\iota \colon A \hookrightarrow X$ has a right homotopy inverse $f \colon X \to A$. Let $\eta \colon \mathrm{id}_X \Rightarrow \iota f$ be a homotopy, then $\eta(-,0) = \mathrm{id}_X$ and $\eta(X,1) = \iota f(X) \subseteq A$—therefore η is a deformation of X into A.

For the converse, suppose X is deformable into A, and let $\delta \colon X \times I \to X$ be such a deformation. Define a map $f \colon X \to A$ to be such that $\iota f = \delta(-,1)$, then f is continuous and δ establishes a homotopy $\mathrm{id}_X \sim_{\mathrm{h}} \iota f$—thus f is a right homotopy inverse of ι. ♮

Definition 15.3.9 (Deformation retract)**.** Let X be a space, and $A \subseteq X$ a subspace. We define the following concepts:

(a) The subspace A is said to be a *weak deformation retract* of X if the canonical inclusion $A \hookrightarrow X$ is a *homotopy equivalence*.

(b) The subspace A is said to be a *deformation retract* of X if there exists a retraction $r \colon X \to A$ of ι—that is, $r\iota = \mathrm{id}_X$—such that there exists a homotopy $\iota r \sim_{\mathrm{h}} \mathrm{id}_X$. A homotopy $\eta \colon \mathrm{id}_X \Rightarrow \iota r$ is called a *deformation retraction of X to A*.

(c) The subspace A is said to be a *strong deformation retract* of X if there is a retraction $r \colon X \to A$ of ι—that is, $r\iota = \mathrm{id}_X$—such that there exists a *relative* homotopy $\iota r \sim_{\mathrm{rel}\, A} \mathrm{id}_X$. A homotopy $\eta \colon \mathrm{id}_X \Rightarrow_{\mathrm{rel}\, A} \iota r$ is called a *strong deformation retraction of X to A*.

Lemma 15.3.10. Let $A \subseteq X$ be a subspace of a topological space X. Then A is a *weak deformation retract* of X if and only if A is a *weak retract* of X and X is deformable into A.

Proof. From Lemma 15.3.8, there exists a left homotopy inverse of the inclusion $\iota \colon A \hookrightarrow X$ if and only if X is deformable into A. On the other hand, A is a weak retract of X if and only if ι admits a right homotopy inverse. Therefore the proposition follows. ♮

Example 15.3.11. The n-th sphere S^n is a strong deformation retraction of the punctured euclidean space $\mathbf{R}^{n+1} \smallsetminus 0$. This is realized by the linear homotopy $\delta \colon (\mathbf{R}^{n+1} \smallsetminus 0) \times I \to \mathbf{R}^{n+1} \smallsetminus 0$ given by

$$\delta(x,t) := (1-t)x + t\frac{x}{\|x\|}.$$

Lemma 15.3.12. Let X be a space, and $A \subseteq X$ be a *retract* of X. If X is *deformable* into the retract A, then A is a *deformation retract* of X.

Proof. Since A is a retract, there exists a left homotopy inverse $r \colon X \to A$ of the inclusion $\iota \colon A \hookrightarrow X$. From hypothesis, X is deformable into A, then there exists a right homotopy inverse $f \colon X \to A$ of ι (by Lemma 15.3.8), notice however that

$$f = \mathrm{id}_A f = (r\iota)f = r(\iota f) = r\,\mathrm{id}_X = r.$$

Therefore we have $\mathrm{id}_X \sim_{\mathrm{h}} \iota r$—thus A is a deformation retract of X. ♮

Corollary 15.3.13 (Removing weaknesses via homotopy extensions)**.** If the pair (X, A) has the homotopy extension property with respect to the subspace A, then A is a *weak deformation retract* of X if and only if A is a *deformation retract* of X.

Proof. If A is a weak deformation retract of X, then by Lemma 15.3.10 we know that A is a weak retract of X and X is deformable into A. From Theorem 15.3.6, since (X, A) has the homotopy extension property, then A is a weak retract if and only if it is a retract. Therefore, since A is a retract of X and X is deformable into A, by Lemma 15.3.12 we conclude that A is a deformation retract of X. The converse is clearly true since a deformation retract is always a weak deformation retract. ♮

Theorem 15.3.14 (Strengthening deformation retracts)**.** Let X be a space and $A \subseteq X$ be a subspace. Define a subspace

$$Y := (X \times 0) \cup (A \times I) \cup (X \times 1)$$

of the cylinder $X \times I$. Then, if $(X \times I, Y)$ has the *homotopy extension property with respect to* X and A is *closed* in X, then A is a *deformation retract* of X if and only if A is a *strong deformation retract* of X.

Proof. Since strong deformation retracts are always deformation retracts, we prove the other side of the equivalence. Assume the hypothesis, and let A be a deformation retract of X. Let $\delta \colon \mathrm{id}_X \Rightarrow \iota r$ be a deformation retraction of X to A—where, as usual, $\iota \colon A \hookrightarrow X$ is the inclusion and $r \colon X \to A$ is the retraction. Define a map $\varepsilon \colon Y \times I \to X$ as follows

$$\varepsilon((x,t),s) := \begin{cases} x, & \text{if } x \in X \text{ and } t = 0, \\ \delta(x,(1-s)t), & \text{if } x \in A \text{ and } t \in I, \\ \delta(r(x), 1-s), & \text{if } x \in X \text{ and } t = 1. \end{cases}$$

The map is indeed well defined since, given any point $a \in A$, one has $\varepsilon((a,0),s) = a = \delta(a,0)$—since δ is a deformation retraction. Moreover, the compatibility of the last two equalities is met because

$$\varepsilon((a,1),1-s) = \delta(a,1-s) = \delta(r(a),1-s),$$

since $r\iota = \mathrm{id}_A$. For the continuity of ε, we use the fact that the sets $(X \times 0) \times I$, $(A \times I) \times I$ and $(X \times 1) \times I$ form a *closed* cover of $Y \times I$—moreover, since ε is continuous in each of the sets of such closed cover—by Proposition 13.2.9 we conclude that ε is continuous.

If $(x,t) \in Y$ is any pair, since $\delta(x,0) = x$ and

$$\delta(r(x),1) = \iota r(r(x)) = r(x) = \delta(x,1),$$

then in general $\varepsilon((x,t),0) = \delta(x,t)$. Using the homotopy extension property of the pair $(X \times I, Y)$ with respect to X, there exists an *extension*

$$\chi\colon (X \times I) \times I \to X$$

for which $\chi((x,t),0) = \delta(x,t)$ for all $(x,t) \in X \times I$, and $\chi|_{Y \times I} = \varepsilon$. Now define a map $\eta\colon X \times I \to X$ given by $\eta(x,t) \coloneqq \chi((x,t),1)$, then we obtain

$$\eta(x,t) = \begin{cases} \varepsilon((x,0),1) = x, & \text{if } x \in X \text{ and } t = 0, \\ \varepsilon((x,t),1) = \delta(x,0) = x, & \text{if } x \in A \text{ and } t \in I, \\ \chi((x,t),1), & \text{if } x \in X \text{ and } t \in I, \\ \varepsilon((x,1),1) = \delta(r(x),0) = \iota r(x), & \text{if } x \in X \text{ and } t = 1, \end{cases}$$

therefore $\eta(-,0) = \mathrm{id}_X$ and $\eta(-,1) = \iota r$, while $\eta(a,t) = a$ for all $(a,t) \in A \times I$. Therefore η establishes a homotopy $\mathrm{id}_X \sim_{\mathrm{rel}\, A} \iota r$, which shows that A is a strong deformation retract of X. ♮

Mapping Cylinder & Cofibrations

Definition 15.3.15 (Mapping cylinder)**.** Let $f\colon X \to Y$ be a topological morphism. We define the *mapping cylinder* of f to be the pushout

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "i_0"'] & Y \arrow[d] \\ X \times I \arrow[r] & \mathrm{Cyl}(f) \arrow[ul, phantom, "\ulcorner", very near start] \end{tikzcd}$$

in $\mathtt{Top}$, where $i_0\colon X \to X \times I$ is the morphism $x \mapsto (x,0)$—the definition can be "isomorphically" given by replacing i_0 by $i_1\colon X \to X \times I$ mapping $x \mapsto (x,1)$, the resulting pushouts are isomorphic topological spaces.

Put more concretely, the cylinder of f is the topological space

$$\mathrm{Cyl}(f) = Y \cup_f (X \times I),$$

where one identifies $(x,0) \sim f(x)$ for all $x \in X$.

Together with the mapping cylinder, we have two distinguished *embedding* morphisms $\iota_X\colon X \hookrightarrow \mathrm{Cyl}(f)$ with $x \mapsto [x,0]$, and $\iota_Y\colon Y \hookrightarrow \mathrm{Cyl}(f)$ mapping $y \mapsto [y]$. Moreover, one has a *retraction* $r\colon \mathrm{Cyl}(f) \to Y$ given by $[x,t] \mapsto f(x)$ for $(x,t) \in X \times I$, while $[y] \mapsto y$ for $y \in Y$.

Theorem 15.3.16. Given a topological morphism $f\colon X \to Y$, there exists morphisms $\iota_X\colon X \to \mathrm{Cyl}(f)$ and $r\colon \mathrm{Cyl}(f) \to Y$ such that the following holds:

(a) The diagram

$$\begin{tikzcd} X \arrow[rr, "\iota_X"] \arrow[dr, "f"'] & & \mathrm{Cyl}(f) \arrow[dl, "r"] \\ & Y & \end{tikzcd}$$

commutes in `Top`. That is, every continuous map can be factored through an embedding and a retraction of its mapping cylinder.

(b) There exists a *relative homotopy*

$$\mathrm{id}_{\mathrm{Cyl}(f)} \sim_{\mathrm{rel}\, Y} \iota_Y r.$$

Thus ι_Y and r are homotopy equivalences.

(c) The embedding ι_X of X into the cylinder of f is a *cofibration*.

Proof. Item (a) follows directly from definition: $r\iota_X(x) = r[x,0] = f(x)$ for any $x \in X$. For item (b), consider the map $\eta\colon \mathrm{Cyl}(f) \times I \to \mathrm{Cyl}(f)$ given by

$$\eta(p,s) := \begin{cases} [x,(1-s)t], & \text{if } p = [x,t] \text{ for } (x,t) \in X \times I, \\ [y], & \text{if } p = [y] \text{ for } y \in Y. \end{cases}$$

Then η is continuous in both $(X \times I) \times I$ and $Y \times I$, which forms a cover of $\mathrm{Cyl}(f) \times I$—therefore η is continuous. Moreover, for any $y \in Y$ we have $\eta([y],-) = [y]$, hence η is a homotopy relative to Y. Since $\eta([x,t],0) = [x,t]$ and $\eta([y],0) = [y]$ then $\eta(-,0) = \mathrm{id}_{\mathrm{Cyl}(f)}$. On the other hand, $\eta([x,t],1) = [x,0] = [f(x)]$ and $\eta([y],1) = [y]$, that is, $\eta(-,0) = \iota_Y r$. Therefore we can conclude that η establishes a relative homotopy $\mathrm{id}_{\mathrm{Cyl}(f)} \sim_{\mathrm{rel}\, Y} \iota_Y r$.

In order to prove item (c), let W be any topological space and consider continuous maps $g\colon \mathrm{Cyl}(f) \to W$ and $\varepsilon\colon X \times I \to W$ such that $g\iota_X(x) = \varepsilon(x,0)$ for all $x \in X$. Define a map $\delta\colon \mathrm{Cyl}(f) \times I \to W$ given by

$$\delta([y],s) := g[y],$$
$$\delta([x,t],s) := \begin{cases} g\left[x, \frac{2t-s}{2-s}\right], & \text{if } s \leqslant 2t, \\ \varepsilon\left(x, \frac{s-2t}{1-t}\right), & \text{if } 2t \leqslant s, \end{cases}$$

where $y \in Y$ and $x \in X$. Therefore, one has $\delta([x,t],0) = g[x,t]$ while $\delta([y],0) = g[y]$. Moreover, δ is an extension of ε since $\delta|_{X \times I} = \varepsilon$. This shows that ι_X is a cofibration. ♮

Lemma 15.3.17. A continuous map $f\colon X \to Y$ is a *homotopy equivalence* if and only if X is a *weak deformation retract* of the mapping cylinder $\mathrm{Cyl}(f)$.

Proof. From the factorization of Theorem 15.3.16 we have $f = r\iota_X$. Since $\iota_Y r \sim_{\mathrm{h}} \mathrm{id}_{\mathrm{Cyl}(f)}$, then $\iota_X \sim_{\mathrm{h}} \iota_Y f$. If f is a homotopy equivalence, then let $g\colon Y \to X$ be its homotopy inverse. It follows that the map $gr\colon \mathrm{Cyl}(f) \to X$ is a homotopy inverse of ι_X. Therefore X is a weak deformation retract of $\mathrm{Cyl}(f)$

Conversely, if X is a weak deformation retract of $\mathrm{Cyl}(f)$, let $k\colon \mathrm{Cyl}(f) \to X$ be an homotopy inverse of ι_X. Then the map $k\iota_Y\colon Y \to X$ is a homotopy inverse of f. ♮

Theorem 15.3.18. Two topological spaces X and Y have the *same homotopy type* if and only if both can be embedded as a *weak deformation retract* of a common space Z.

Proof. If $f: X \to Y$ is a homotopy equivalence, then by Lemma 15.3.17 we find that X is a weak deformation retract of $\mathrm{Cyl}(f)$. Since Y is also a weak deformation retract of $\mathrm{Cyl}(f)$, the statement follows. For the converse, if $i_X: X \to Z$ and $i_Y: Y \to Z$ are embeddings, that are homotopy equivalences, then let $r: Z \to Y$ be the retract of i_Y and define $f := r i_X: X \to Y$, then f is a homotopy equivalence $X \simeq_{\mathrm{h}} Y$. ♮

15.4 Fundamental Groupoid & The Fundamental Group

Paths

Notation 15.4.1 (Family of paths)**.** Let X be a space and $x, y \in X$ be any two points. We'll denote by $\mathrm{Path}_X(x, y)$ the family of paths $\gamma: I \to X$ with $\gamma(0) = x$ and $\gamma(1) = y$.

Definition 15.4.2 (Operations on paths)**.** Let X be a topological space. We define the following operations on the space of paths of X:

- Given a path $\gamma: I \to X$ we define the *reverse* path of γ to be a path $\gamma^{-1}: I \to X$ given by $\gamma^{-1}(t) := \gamma(1 - t)$.
- If $p \in \mathrm{Path}_X(x, y)$ and $q \in \mathrm{Path}_X(y, z)$ are paths in X, we define the *concatenation* of p with q to be a path $q \cdot p: I \to X$ given by

$$(q \cdot p)(t) := \begin{cases} p(2t), & t \in [0, 1/2], \\ q(2t - 1), & t \in [1/2, 1]. \end{cases}$$

  Yielding a path $q \cdot p \in \mathrm{Path}_X(x, z)$.
- Moreover, we define $\mathrm{cons}_x: I \to X$ to be the unique constant path on x—that is, $\mathrm{cons}_x(t) = x$ for all $t \in I$.

Proposition 15.4.3. If X is a path connected topological space, then any two paths on X are homotopic.

Proof. Let $f, g: I \rightrightarrows X$ be any two paths. Define, for each $t \in I$, a path $\gamma_t \in \mathrm{Path}_X(f(t), g(t))$—which exists since X is path connected. Define a map $\eta: I \times I \to X$ given by $\eta(t, s) := \gamma_t(s)$, then it is clear that η is a homotopy between f and g. ♮

Definition 15.4.4 (Homotopy relative boundary)**.** Let X be a space, and $\gamma, \gamma': I \rightrightarrows X$ be paths with common endpoints:

$$\gamma(0) = \gamma'(0) =: p_0 \quad \text{and} \quad \gamma(1) = \gamma'(1) =: p_1.$$

We define a *homotopy relative boundary* between γ and γ' to be a homotopy $\eta: \gamma \Rightarrow \gamma'$ such that η is constant on the endpoints p_0 and p_1—that is,

$$\eta(0, -) = \mathrm{cons}_{p_0} \quad \text{and} \quad \eta(1, -) = \mathrm{cons}_{p_1}.$$

Proposition 15.4.5. Homotopy relative boundary is an equivalence relation on the family of paths of a topological space.

Proof. Let X be a topological space and $x, y \in X$ be any two points—we'll consider the family $\mathrm{Path}_X(x, y)$. The constant homotopy $\gamma \Rightarrow \gamma$ is clearly a homotopy relative boundary, thus the relation is reflexive.

If $\delta \in \mathrm{Path}_X(x, y)$ is another path, and $\eta\colon \gamma \Rightarrow \delta$ is a homotopy relative boundary, then we can construct the a reverse homotopy $\sigma\colon \delta \Rightarrow \gamma$ as $\sigma(-, t) := \eta(-, 1-t)$. Notice that $\sigma(0, t) = \eta(0, 1-t) = \mathrm{cons}_x(1-t)$ is constant on x, while $\sigma(1, t) = \eta(1, 1-t) = \mathrm{cons}_y(1-t)$ is constant on y. Therefore the relation is symmetric.

Consider yet another path $p \in \mathrm{Path}_X(x, y)$ and a homotopy relative boundary $\eta\colon \delta \Rightarrow p$. We define a map $\varepsilon\colon I \times I \to X$ given by

$$\varepsilon(-, t) := \begin{cases} \sigma(-, 2t), & t \in [0, 1/2], \\ \eta(-, 2t-1), & t \in [1/2, 1]. \end{cases}$$

This map is continuous by the same argument used in Corollary 15.1.2 and thus establishes a homotopy relative boundary $\varepsilon\colon \gamma \Rightarrow p$. ♮

The Fundamental Groupoid

Definition 15.4.6 (Fundamental groupoid)**.** Given a topological space X, we define the *fundamental groupoid* of X to be the category $\Pi_1(X)$ whose objects are the points of X, and whose morphisms are paths between those points up to homotopy relative boundary. That is, given $x, y \in X$ we have

$$\mathrm{Mor}_{\Pi_1(X)}(x, y) = \mathrm{Path}_X(x, y)/\sim_{\mathrm{hrb}},$$

where $\sim_{\mathrm{hrb}}$ is the equivalence relation on the family of paths $x \to y$ given by homotopy relative boundary.

Composition of morphisms in $\Pi_1(X)$ is naturally defined by the concatenation of paths—in other words, given points $x, y, z \in X$ we have

$$\begin{tikzcd} \mathrm{Path}_X(x, y) \times \mathrm{Path}_X(y, z) \arrow[r] \arrow[d, two heads] & \mathrm{Path}_X(x, z) \arrow[d, two heads] \\ \mathrm{Mor}_{\Pi_1(X)}(x, y) \times \mathrm{Mor}_{\Pi_1(X)}(y, z) \arrow[r, dashed] & \mathrm{Mor}_{\Pi_1(X)}(x, z) \end{tikzcd}$$

where, for any paths $\gamma \in \mathrm{Path}_X(x, y)$ and $\delta \in \mathrm{Path}_X(y, z)$, the concatenation of paths $(\gamma, \delta) \mapsto \delta \cdot \gamma$ induces a concatenation operation on the respective class paths

$$([\gamma], [\delta]) \longmapsto [\delta] \cdot [\gamma] := [\delta \cdot \gamma].$$

Corollary 15.4.7. $\Pi_1(X)$ is a groupoid.

Proof. We show that $\Pi_1(X)$ is a category whose morphisms are isomorphisms.

- Given any point $x \in X$ one has an identity map $[\mathrm{cons}_x] \in \mathrm{Mor}_{\Pi_1(X)}(x, x)$.
- Concatenation of class paths is unital with respect to constant paths: given any path class $[\gamma] \in \mathrm{Mor}_{\Pi_1(X)}(x, y)$, we have

$$[\mathrm{cons}_y] \cdot [\gamma] = [\mathrm{cons}_y \cdot \gamma] = [\gamma] = [\gamma \cdot \mathrm{cons}_x] = [\gamma] \cdot [\mathrm{cons}_x].$$

- Concatenation is associative. Let $x, y, z, w \in X$ be any four points and consider class paths $[\alpha]: x \to y$, $[\beta]: y \to z$, and $[\gamma]: z \to w$.

  First we have to show that $\gamma \cdot (\beta \cdot \alpha)$ and $(\gamma \cdot \beta) \cdot \alpha$ are relative boundary homotopic. To that end, define a continuous map $\tau: I \to I$ by

$$\tau(t) := \begin{cases} 2t, & t \in [0, 1/4], \\ t + \frac{1}{4}, & t \in [1/4, 1/2], \\ \frac{t}{2} + \frac{1}{2}, & t \in [1/2, 1]. \end{cases}$$

  Then, one sees right away that

$$(\gamma \cdot (\beta \cdot \alpha))(\tau(t)) = ((\gamma \cdot \beta) \cdot \alpha)(t)$$

  for all $t \in I$—hence $\gamma \cdot (\beta \cdot \alpha) \sim_{\mathrm{hrb}} (\gamma \cdot \beta) \cdot \alpha$. From this, we finally obtain the associativity of the path classes,

$$[\gamma] \cdot ([\beta] \cdot [\alpha]) = [\gamma \cdot (\beta \cdot \alpha)] = [(\gamma \cdot \beta) \cdot \alpha] = ([\gamma] \cdot [\beta]) \cdot [\alpha].$$

- Every path class $[\gamma] \in \mathrm{Mor}_{\Pi_1(X)}(x, y)$ is an isomorphism, since it has an inverse $[\gamma^{-1}] \in \mathrm{Mor}_{\Pi_1(X)}(y, x)$. Indeed, one has

$$[\gamma] \cdot [\gamma^{-1}] = [\gamma \cdot \gamma^{-1}] = [\mathrm{cons}_y] \quad \text{and} \quad [\gamma^{-1}] \cdot [\gamma] = [\gamma^{-1} \cdot \gamma] = [\mathrm{cons}_x].$$

♮

Definition 15.4.8 (The category $\mathtt{Top}^{*/}$)**.** A *pointed topological space* is a pair (X, x) consisting of a topological space X together with a base-point $x \in X$. We define a category $\mathtt{Top}^{*/}$ whose objects are pointed topological spaces, and whose morphisms $f: (X, x) \to (Y, y)$, for any $(X, x), (Y, y) \in \mathtt{Top}^{*/}$, are continuous maps $f: X \to Y$ such that $f(x) = y$. We say that the morphisms of $\mathtt{Top}^{*/}$ preserve base-points.

In $\mathtt{Top}^{*/}$, we define a *pointed homotopy* $\eta: f \Rightarrow g$ between parallel morphisms $f, g: (X, x) \rightrightarrows (Y, y)$ to be a homotopy preserving base-points, that is,

$$\eta(x, -) = \mathrm{cons}_y .$$

Analogously to homotopies in $\mathtt{Top}$, pointed homotopies define an equivalence relation $\sim_{\mathrm{ph}}$ in $\mathtt{Top}^{*/}$. We define the homotopy category of $\mathtt{Top}^{*/}$ to be the category $\mathtt{Ho}(\mathtt{Top}^{*/})$ composed of pointed topological spaces and morphisms

$$\mathrm{Mor}_{\mathtt{Ho}(\mathtt{Top}^{*/})}((X, x), (Y, y)) := \mathrm{Mor}_{\mathtt{Top}^{*/}}((X, x), (Y, y))/{\sim_{\mathrm{ph}}}.$$

This quotient induces a natural projective functor

$$\kappa^{*/}: \mathtt{Top}^{*/} \longrightarrow \mathtt{Ho}(\mathtt{Top}^{*/}).$$

The Fundamental Group

Definition 15.4.9 (Fundamental group)**.** Let X be a topological space and $x \in X$ be any point. We define the *fundamental group* of X at the base-point x as the family of *loops*

$$\pi_1(X, x) := \mathrm{Aut}_{\Pi_1(X)}(x),$$

endowed with the operation of concatenation of paths. Therefore, we see that the fundamental group is a functor

$$\pi_1 \colon \mathtt{Top}^{*/} \longrightarrow \mathtt{Grp}.$$

Proposition 15.4.10. Let X be a path connected space. For every point $x \in X$ the inclusion functor $\pi_1(X, x) \to \Pi_1 X$ is an *equivalence of categories*.

Proof. Indeed, notice that $\pi_1(X, x) = \mathrm{Aut}_{\Pi_1 X}(x)$ and since X is a connected space, $\Pi_1 X$ is a connected groupoid—that is, for every pair of points $x, y \in \Pi_1 X$, there exists an isomorphism $x \simeq y$ in $\Pi_1 X$. It follows that $\mathrm{Aut}_{\Pi_1 X}(x) = \mathrm{sk}(\Pi_1 X)$, and therefore the inclusion functor is an equivalence of categories (see Example 1.4.13). ♮

Definition 15.4.11 (Pushforwards in π_1)**.** Let $f \colon (X, x) \to (Y, y)$ be a morphism of pointed topological spaces. There exists an induced pushforward

$$f_* \colon \pi_1(X, x) \longrightarrow \pi_1(Y, y)$$

mapping $[\gamma] \mapsto [f\gamma]$, which establishes a group morphism between the fundamental groups of the initial pointed topological spaces.

Proposition 15.4.12. Let $f, g \colon (X, x) \rightrightarrows (Y, y)$ be morphisms in $\mathtt{Top}^{*/}$. If there exists a *pointed homotopy* $\eta \colon f \Rightarrow g$, then $f_* = g_*$.

Proof. Let γ be a loop at x representing some class of $\pi_1(X, x)$. The pointed homotopy η naturally induces a homotopy $f\gamma \Rightarrow g\gamma$—thus $[f\gamma] = [g\gamma]$ in $\pi_1(Y, y)$—therefore $f_* = g_*$. ♮

By Proposition 15.4.12 we obtain a factorization of the fundamental group functor through the homotopy category of pointed topological spaces. To put briefly, the following diagram is quasi-commutative

$$\begin{tikzcd}
\mathtt{Top}^{*/} \arrow[r, "\pi_1"] \arrow[d, "\kappa^{*/}"'] & \mathtt{Grp} \\
\mathtt{Ho}(\mathtt{Top}^{*/}) \arrow[ur, bend right] &
\end{tikzcd}$$

Corollary 15.4.13 (Preserving isomorphisms)**.** If $f \colon (X, x) \xrightarrow{\simeq} (Y, y)$ is a *homotopy equivalence* in $\mathtt{Top}^{*/}$, then $f_* \colon \pi_1(X, x) \xrightarrow{\simeq} \pi_1(Y, y)$ is an *isomorphism* in $\mathtt{Grp}$.

Proof. If g is the homotopy inverse of f, then for any $[\gamma] \in \pi_1(X, x)$ we have $g_* f_* [\gamma] = g_*[f\gamma] = [gf\gamma] = [\gamma]$, therefore $g_* f_* = \mathrm{id}_{\pi_1(X,x)}$. Analogously we have $f_* g_* = \mathrm{id}_{\pi_1(Y,y)}$. Therefore f_* is an isomorphism of groups. ♮

Simply Connected Spaces

Definition 15.4.14 (Simply connected space)**.** A topological space X is said to be *simply connected* if it is path connected and its fundamental group is trivial, $\pi_1(X,x) \simeq 1$ for any base-point $x \in X$.

Continue on semi-locally simply connected spaces

Examples of Fundamental Groups

Example 15.4.15 (π_1 of the euclidean space)**.** Let $x \in \mathbf{R}^n$ be any point. Given any loop $\ell\colon x \to x$, one can define a homotopy relative boundary $\eta\colon \ell \Rightarrow \mathrm{cons}_x$ given by $\eta(s,t) := (1-t)\ell(s) + tx$. Therefore the fundamental group of the n-dimensional euclidean space is trivial,

$$\pi_1(X,x) = *.$$

Proposition 15.4.16. The fundamental group of a topological manifold is countable.

Proof. Let $\mathcal{B}$ be a countable open cover of M consisting of coordinate balls. Since M has countably many connected components (see Proposition 14.1.29), those connected components are also path components. This implies that any two $B, B' \in \mathcal{B}$ are such that the intersection $B \cap B'$ are composed of at most countably many path components (since M is locally path connected). Let $\mathcal{X}$ be a collection containing a single point from each $B \cap B'$ for any pair $(B, B') \in \mathcal{B} \times \mathcal{B}$. For every $B \in \mathcal{B}$ and points $x, y \in \mathcal{X}$ with $x, y \in B$, fix a path $\gamma^B_{(x,y)} \in \mathrm{Path}_B(x,y)$.

Clearly, the fundamental groups of a path connected component of M based at any two points are always isomorphic. Since $\mathcal{X}$ contains at least one point of each path component of M, we may choose a base-point $p \in \mathcal{X}$. Let Γ_p denote the collection of loops based at p that are equal to a finite concatenation of paths of the form $\gamma^B_{(x,y)}$, for some $B \in \mathcal{B}$. Since $\mathcal{B}$ is countable, so is Γ_p. We'll settle to prove that every element of $\pi_1(M,p)$ can be represented by a loop in Γ_p.

Let f be any loop based at p. Consider the collection $(f^{-1}(B))_{B \in \mathcal{B}}$, which is an open cover of I. Since I is compact, there exists a finite subcover out of such collection. The finite subcover gives rise to a finite collection of numbers

$$0 =: a_0 < a_1 < \cdots < a_k := 1$$

for which the closed interval $[a_{j-1}, a_j]$ is contained in some $f^{-1}(B)$.

For each $0 < j \leqslant k$, define $f_j\colon I \to M$ to be the path given by

$$f_j(t) := f((1-t)a_{j-1} + t a_j),$$

that is, f_j is a path $f(a_{j-1}) \to f(a_j)$—and let $B_j \in \mathcal{B}$ be such that $\operatorname{im} f_j \subseteq B_j$. From construction we have that $f(a_j) \in B_j \cap B_{j+1}$ for each $0 \leqslant j < k$.

For each $0 \leqslant j < k$ let $g_j \in \mathrm{Path}_{B_j \cap B_{j+1}}(x_j, f(a_j))$, where $x_j \in \mathcal{X}$ as in the first paragraph—and $x_0, x_k := p$, with constant paths $g_0, g_k := \mathrm{cons}_p$. Notice that

$$\begin{aligned} f &\sim_{\mathrm{h}} f_k \cdots f_1 \\ &\sim_{\mathrm{h}} g_k^{-1} f_k (g_{k-1} g_{k-1}^{-1}) \cdots (g_2 g_2^{-1}) f_2 (g_1 g_1^{-1}) f_1 g_0 \\ &\sim_{\mathrm{h}} (g_k^{-1} f_k g_{k-1}^{-1}) \cdots (g_1^{-1} f_1 g_0), \end{aligned}$$

where $g_j^{-1} f_j g_{j-1} \in \mathrm{Path}_{B_j}(x_{j-1}, x_j)$ for all $0 < j \leqslant k$. Since each B_j is simply connected, then $g_j^{-1} f_j g_{j-1}$ is relative boundary homotopic to chosen path $\gamma^B_{(x_{j-1}, x_j)}$. This shows that f is homotopic to a path in Γ_p—therefore $\pi_1(M, p)$ is countable. ♮

15.5 Mappings $S^1 \longrightarrow S^1$

We shall study the circle via the inclusion $S^1 \hookrightarrow \mathbf{C}$ mapping $(\cos(2\pi t), \sin(2\pi t)) \mapsto e^{2\pi \mathrm{i} t}$ and the identification $q\colon I \to S^1$ given by $t \mapsto e^{2\pi \mathrm{i} t}$. In general, we shall also make use of the map $\mathbf{R} \to S^1$ given by $x \mapsto e^{2\pi \mathrm{i} x}$.

Each path $\phi\colon (I, 0) \to (\mathbf{R}, 0)$ between pointed spaces, can be *lifted* to a unique path $\widehat{\phi}\colon (S^1, 1) \to (S^1, 1)$ such that the diagram

$$\begin{tikzcd} I \arrow[r, "\phi_n"] \arrow[d] & \mathbf{R} \arrow[d] \\ S^1 \arrow[r, dashed, "\widehat{\phi}_n"'] & S^1 \end{tikzcd}$$

commutes in $\mathtt{Top}$. Explicitly, we have

$$\widehat{\phi}_n(e^{2\pi \mathrm{i} t}) = e^{2\pi \mathrm{i} \phi_n(t)}.$$

Proposition 15.5.1 (Unwinding pointed maps). Let $f\colon (S^1, 1) \to (S^1, 1)$ be a *pointed* continuous map. Then there exists a *unique pointed morphism* $\phi\colon (I, 0) \to (\mathbf{R}, 0)$ such that $f = \widehat{\phi}$.

Proof. Define $h\colon I \to S^1$ to be the morphism $h := f q$—which is uniformly continuous since I is a compact space. From the latter properties, one can find a finite partition

$$0 =: t_0 < t_1 < \cdots < t_k := 1$$

of I, such that $|h(t) - h(t_j)| < 2$ for all $t \in [t_j, t_{j+1}]$, for each $0 \leqslant j < k$. This is done in order to ensure that $h(t) \neq -h(t_j)$—so that the complex logarithm[1] of the quotient

[1] If $z = r e^{\mathrm{i}\theta} \in \mathbf{C}$ is a complex number, in polar form, satisfying $r > 0$ and $\theta \in (-\pi, \pi)$, then the *complex logarithm* of z is $\mathrm{Log}(z) = \log(r) + \mathrm{i}\theta$.

$h(t)/h(t_j)$ is well defined. With these conditions being satisfied, we can define a map $\phi\colon I \to \mathbf{R}$ given by

$$\phi(t) := \frac{1}{2\pi\mathrm{i}}\left(\operatorname{Log}\left(\frac{h(t_1)}{h(t_0)}\right) + \cdots + \operatorname{Log}\left(\frac{h(t_k)}{h(t_{k-1})}\right) + \operatorname{Log}\left(\frac{h(t)}{h(t_j)}\right)\right)$$

for all $t \in [t_j, t_{j+1}]$. Therefore one has

$$\begin{aligned} q\phi(t) &= e^{2\pi\mathrm{i}\phi(t)} \\ &= e^{\operatorname{Log}\left(\frac{h(t_1)}{h(t_0)}\right)+\cdots+\operatorname{Log}\left(\frac{h(t_j)}{h(t_{j-1})}\right)+\operatorname{Log}\left(\frac{h(t)}{h(t_j)}\right)} \\ &= e^{\operatorname{Log}\left(\frac{h(t_1)}{h(t_0)}\right)}\cdots e^{\operatorname{Log}\left(\frac{h(t_j)}{h(t_{j-1})}\right)} e^{\operatorname{Log}\left(\frac{h(t)}{h(t_k)}\right)} \\ &= \frac{h(t_1)}{h(t_0)}\cdots\frac{h(t_j)}{h(t_{j-1})}\cdot\frac{h(t)}{h(t_j)} \\ &= \frac{h(t)}{h(t_0)}. \end{aligned}$$

Now, from construction we know that $h(t_0) = h(0) = 1$, thus $q\phi(t) = h(t)$. Moreover $\phi(0) = 0$, which implies that ϕ is a continuous pointed map of the form $(I, 0) \to (\mathbf{R}, 0)$ such that

$$\widehat{\phi}(t) = f(e^{2\pi\mathrm{i}t}).$$

For the uniqueness of ϕ, suppose that $\psi\colon (I, 0) \to (\mathbf{R}, 0)$ is another pointed morphism such that $f = \widehat{\psi}$. Then in particular $\widehat{\phi} = \widehat{\psi}$, which implies in $e^{2\pi\mathrm{i}\phi(t)} = e^{2\pi\mathrm{i}\psi(t)}$ for each $t \in I$. This can only be the case if $\phi(t) - \psi(t) \in \mathbf{Z}$. Therefore, since the map $\phi - \psi$ is a continuous map of the form $(I, 0) \to (\mathbf{Z}, 0)$, it follows that $\phi = \psi$—since I is connected and $\mathbf{Z}$ is discrete. ♮

Theorem 15.5.2 (Unwinding maps)**.** Let $f\colon S^1 \to S^1$ be any morphism, then there exists a *unique pointed morphism* $\phi\colon (I, 0) \to (\mathbf{R}, 0)$ such that $f = f(1)\widehat{\phi}$.

Proof. From f we can define a *pointed* morphism $g\colon (S^1, 1) \to (S^1, 1)$ by $g := f(1)^{-1}f$, so that indeed $g(1) = 1$. From Proposition 15.5.1 there exists a unique pointed morphism $\phi\colon (I, 0) \to (\mathbf{R}, 0)$ such that $g = \widehat{\phi}$. Therefore, from the definition of g we find that $f = f(1)\widehat{\phi}$. ♮

Lemma 15.5.3. Let $n \in \mathbf{Z}$ and $\phi \in \operatorname{Path}_{\mathbf{R}}(0, n)$. If we consider the linear path $\phi_n\colon I \to \mathbf{R}$ given by $\phi_n(t) := tn$, then there exists a relative homotopy $\widehat{\phi} \sim_{\mathrm{rel}\,1} \widehat{\phi}_n$ between the lifted paths in the circle.

Proof. There exists a relative linear homotopy $\ell\colon \phi \Rightarrow_{\mathrm{rel}\,\partial I} \phi_n$ given by

$$\ell(s, t) := (1 - t)\phi(s) + t\phi_n(s).$$

Notice that this homotopy can be lifted to $\widehat{\ell}\colon S^1 \times I \to S^1$ mapping

$$\widehat{\ell}(e^{2\pi\mathrm{i}s}, t) = e^{2\pi\mathrm{i}\left((1-t)\phi(s)+t\phi_n(s)\right)} = e^{2\pi\mathrm{i}(1-t)\phi(s)}e^{2\pi\mathrm{i}t\phi_n(s)}$$

which is a continuous map such that $\widehat{\ell}(e^{2\pi \mathrm{i}s}, 0) = e^{2\pi \mathrm{i}\phi(s)} = \widehat{\phi}(e^{2\pi \mathrm{i}s})$ while on the other end $\widehat{\ell}(e^{2\pi \mathrm{i}s}, 1) = e^{2\pi \mathrm{i}\phi_n(s)} = \widehat{\phi}_n(e^{2\pi \mathrm{i}s})$. That is, $\widehat{\ell}$ establishes a relative homotopy $\widehat{\phi} \sim_{\mathrm{rel}\,1} \widehat{\phi}_n$ as we desired. ♮

Proposition 15.5.4. Let $f: S^1 \to S^1$ be a morphism. Then there exists a unique integer $n \in \mathbf{Z}$ for which $f \sim_{\mathrm{h}} \widehat{\phi}_n$, where $\phi_n: I \to \mathbf{R}$ is the linear path $\phi_n(t) := tn$.

Proof. From Theorem 15.5.2 we find that $f = f(1)\widehat{\phi}$ for a unique pointed map $\phi: (I, 0) \to (\mathbf{R}, 0)$. Since $\widehat{\phi}: (S^1, 1) \to (S^1, 1)$ is a pointed continuous map, it follows that

$$\widehat{\phi}(1) = \widehat{\phi}(e^{2\pi \mathrm{i}}) = e^{2\pi \mathrm{i}\phi(1)} = 1,$$

which implies in $\phi(1) := n \in \mathbf{Z}$. From Lemma 15.5.3 we get $\widehat{\phi} \sim_{\mathrm{rel}\,1} \widehat{\phi}_n$.

Notice that if $\zeta_0 := e^{2\pi \mathrm{i}s_0} \in S^1$ is any fixed point, the multiplication map $\mathrm{mul}: S^1 \to S^1$ given by $\mathrm{mul}(\zeta) := \zeta_0\zeta$ is a rotation of the point ζ on the circle—we'll show that $\mathrm{mul} \sim_{\mathrm{h}} \mathrm{id}_{S^1}$. Let $\delta: S^1 \times I \to S^1$ be the map $\delta(e^{2\pi \mathrm{i}s}, t) := e^{2\pi \mathrm{i}((1-t)s_0 + s)}$, so that $\delta(-, 0) = \mathrm{mul}$ and $\delta(-, 1) = \mathrm{id}_{S^1}$.

From the last paragraph we find that

$$f = f(1) \cdot \widehat{\phi} \sim_{\mathrm{h}} \mathrm{id}_{S_1} \widehat{\phi} = \widehat{\phi} \sim_{\mathrm{h}} \widehat{\phi}_n,$$

therefore $f \sim_{\mathrm{h}} \widehat{\phi}_n$ as wanted. ♮

Definition 15.5.5 (Degree)**.** Given an endomorphism $f: S^1 \to S^1$, let $\phi: I \to \mathbf{R}$ be the unique path such that $f = f(1)\widehat{\phi}$—then we define the *degree* of f to be $\deg f := \phi(1) \in \mathbf{Z}$.

Lemma 15.5.6. Let $f, g: S^1 \rightrightarrows S^1$ be endomorphisms of the circle. Then $f \sim_{\mathrm{h}} g$ if and only if $\deg f = \deg g$.

Proof. ($\Rightarrow$) Let $\eta: f \Rightarrow g$ be a homotopy. For each $s \in I$ there exists an endomorphism $\eta_s := \eta(-, s): S^1 \to S^1$ and, from Theorem 15.5.2, there is a unique pointed continuous map $\phi_s: (I, 0) \to (\mathbf{R}, 0)$ with $\phi_s(1) \in \mathbf{Z}$ for which

$$\eta_s = \eta_s(1)\widehat{\phi}_s. \tag{15.1}$$

We shall construct an explicit equation for ϕ_s and prove that the mapping $\Phi: I \times I \to \mathbf{R}$ given by $\Phi(t, s) := \phi_s(t)$ is a homotopy.

We proceed as in the proof of Proposition 15.5.1: define a map $\varepsilon: I \times I \to S^1$ by

$$\varepsilon(t, s) := \eta_s(e^{2\pi \mathrm{i}t}) = \eta_s q(t).$$

Since $I \times I$ is compact, ε is uniformly continuous—therefore one can choose a partition

$$0 =: t_0 < t_1 < \cdots < t_k := 1$$

of I such that $|\varepsilon(t, s) - \varepsilon(t_j, s)| < 2$ for all $s \in I$ and $t \in [t_j, t_{j+1}]$—this ensures that the complex logarithm of $\varepsilon(t, s)/\varepsilon(t_j, s) \neq -1$ is well defined. For each $s \in I$, we construct a map $\psi_s: I \to S^1$ by

$$\psi_s(t) := \frac{1}{2\pi \mathrm{i}}\left(\mathrm{Log}\left(\frac{\varepsilon(t_1, s)}{\varepsilon(t_0, s)}\right) + \cdots + \mathrm{Log}\left(\frac{\varepsilon(t_j, s)}{\varepsilon(t_{j-1}, s)}\right) + \mathrm{Log}\left(\frac{\varepsilon(t, s)}{\varepsilon(t_j, s)}\right)\right) \tag{15.2}$$

for all $t \in [t_j, t_{j+1}]$. Then for every $s \in I$ one has

$$\begin{aligned}
q\psi_s(t) &= e^{2\pi i \psi_s(t)} \\
&= e^{\operatorname{Log}\left(\frac{\varepsilon(t_1,s)}{\varepsilon(t_0,s)}\right) + \cdots + \operatorname{Log}\left(\frac{\varepsilon(t_j,s)}{\varepsilon(t_{j-1},s)}\right) + \operatorname{Log}\left(\frac{\varepsilon(t,s)}{\varepsilon(t_j,s)}\right)} \\
&= e^{\operatorname{Log}\left(\frac{\varepsilon(t_1,s)}{\varepsilon(t_0,s)}\right)} \dots e^{\operatorname{Log}\left(\frac{\varepsilon(t_j,s)}{\varepsilon(t_{j-1},s)}\right)} e^{\operatorname{Log}\left(\frac{\varepsilon(t,s)}{\varepsilon(t_j,s)}\right)} \\
&= \frac{\varepsilon(t_1,s)}{\varepsilon(t_0,s)} \cdots \frac{\varepsilon(t_j,s)}{\varepsilon(t_{j-1},s)} \cdot \frac{\varepsilon(t,s)}{\varepsilon(t_j,s)} \\
&= \frac{\varepsilon(t,s)}{\varepsilon(t_0,s)} \\
&= \frac{\eta_s(e^{2\pi i t})}{\eta_s(1)}.
\end{aligned}$$

Therefore $\eta_s = \eta_s(1)\widehat{\psi}_s$, which shows that $\psi_s = \phi_s$ (from Eq. (15.1)). Looking at Eq. (15.2) we see that the mapping Φ above-mentioned is continuous, hence a homotopy.

Considering the continuous map $\Phi(1,-)\colon I \to \mathbf{R}$ given by $s \mapsto \phi_s(1)$, we know that $\phi_s(1) \in \mathbf{Z}$ from earlier considerations—therefore using the fact that I is connected and $\mathbf{Z}$ is discrete, it must bae the case that the induced map $\Phi(1,-)\colon I \to \mathbf{Z}$ is *constant*. In particular, we shall have $\phi_0(1) = \phi_1(1)$—but from definition we have $f = f(1)\widehat{\phi}_0$ and $g = g(1)\widehat{\phi}_1$, so that $\deg f = \phi_0(1)$ and $\deg g = \phi_1(1)$. From these considerations we can finally conclude that $\deg f = \deg g$.

($\Leftarrow$) For the converse, suppose that $\deg f = \deg g := n$. Therefore there exists two unique pointed morphisms $\phi, \psi \in \mathrm{Path}_{\mathbf{R}}(0, n)$ such that $f = f(1)\widehat{\phi}$ and $g = g(1)\widehat{\psi}$. Recall that since $f(1), g(1) \in S^1$ are unitary complex numbers, multiplying a circle point by them amounts to a rotation of the initial point through the circle—which was already shown to be homotopic to the identity map of the circle. Therefore we conclude that $\phi \sim_{\mathrm{h}} \phi_n \psi$, where $\phi_n \in \mathrm{Path}_{\mathbf{R}}(0, n)$ is the linear path $\phi_n(t) := tn$. Thus

$$f \sim_{\mathrm{h}} \widehat{\phi} \sim_{\mathrm{h}} \widehat{\phi}_n \sim_{\mathrm{h}} \widehat{\psi} \sim_{\mathrm{h}} g,$$

proving that f and g are homotopic. ♮

Lemma 15.5.7. For each $n \in \mathbf{Z}$ the map $e_n\colon S^1 \to S^1$ given by $\zeta \mapsto \zeta^n$ is a continuous map of degree

$$\deg e_n = n.$$

Proof. Notice that from definition we have $e_n(e^{2\pi i t}) = e^{2\pi i (nt)} = q\phi_n(t)$, where $\phi_n \in \mathrm{Path}_{\mathbf{R}}(0, n)$ is the linear path $\phi_n(t) := tn$. Therefore $\deg e_n = \phi_n(1) = n$. ♮

Theorem 15.5.8 (deg is a ring isomorphism)**.** The map $\deg\colon [S^1, S^1] \to \mathbf{Z}$ given by $[f] \mapsto \deg f$ is a *ring isomorphism*. In other words, the degree map is an isomorphism of the *first cohomology group of* S^1 with $\mathbf{Z}$:

$$H^1(S^1) \simeq \mathbf{Z}.$$

Proof. From Lemma 15.5.6 we know that deg is a bijective map, we need to show that it's also a ring morphism. Consider any two endomorphism classes $[f], [g] \in [S^1, S^1]$ with degrees $\deg[f] := n$ and $\deg[g] := m$, then:

- (Additive structure). From Lemma 15.5.7 we know that $f \sim_{\mathrm{h}} e_n$ and $g \sim_{\mathrm{h}} e_m$, moreover given any $\zeta \in S^1$ one has

  $$(e_n \cdot e_m)(\zeta) = e_n(\zeta)e_m(\zeta) = \zeta^{n+m}.$$

  Therefore $e_n \cdot e_m = e_{n+m}$ and, consequently

  $$\deg([f]\cdot[g]) = \deg([e_n]\cdot[e_m]) = \deg[e_n \cdot e_m] = \deg[e_{n+m}] = n+m = \deg[f]+\deg[g].$$

- (Multiplicative structure). Given any $\zeta \in S^1$ we have

  $$e_n e_m(\zeta) = e_n(\zeta^m) = \zeta^{mn} = \zeta^{nm},$$

  therefore $e_n e_m = e_{nm}$. From this we find that

  $$\deg([f]\circ[g]) = \deg([e_n]\circ[e_m]) = \deg[e_n \circ e_m] = \deg[e_{nm}] = nm = \deg[f]\cdot\deg[g].$$

This finishes the proof that deg is an isomorphism of rings

$$[S^1, S^1] \simeq \mathbf{Z}.$$

♮

Corollary 15.5.9 (Automorphisms of S^1)**.** If $f \in \mathrm{Aut}_{\mathrm{Top}}(S^1)$ is an automorphism, then $\deg f = \pm 1$. As a consequence, either $f \sim_{\mathrm{h}} \mathrm{id}_{S^1}$ or $f \sim_{\mathrm{h}} \rho$, where ρ represents the reflection by complex conjugation.

Proof. If f is an automorphism, let $f^{-1}\colon S^1 \to S^1$ be its inverse. Therefore $ff^{-1} = \mathrm{id}_{S^1}$, implying in

$$1 = \deg \mathrm{id}_{S^1} = \deg(ff^{-1}) = \deg f \cdot \deg f^{-1},$$

which can only be the case if $\deg f = \deg f^{-1} = \pm 1$. Notice that the identity morphism is $\mathrm{id}_{S^1} = e_1$, therefore $\deg \mathrm{id}_{S^1} = 1$. On the other hand, the reflection $\rho\colon S^1 \to S^1$ maps $e^{2\pi i t} \mapsto e^{-2\pi i t}$—thus $\rho = e_{-1}$, implying in $\deg \rho = -1$. From the latter two considerations, the last proposition follows. ♮

Example 15.5.10. A null-homotopic map $f\colon S^1 \to S^1$ has $\deg f = 0$, since it's homotopic to a constant map.

Corollary 15.5.11. The circle S^1 is not contractible.

Proof. Since $\deg \mathrm{id}_{S^1} = 1$, from our considerations from Example 15.5.10 we conclude that id_{S^1} isn't null-homotopic and therefore there exists no contraction of the circle. ♮

Corollary 15.5.12. There exists no retraction $D^2 \to S^1$.

Proof. Suppose, for the sake of contradiction, that there exists a retraction $r\colon D^2 \longrightarrow S^1$ such that $r\iota = \mathrm{id}_{S^1}$, where $\iota\colon S^1 \hookrightarrow D^2$ is the canonical inclusion. Since D^2 is a contractible space, then r is null-homotopic. This would imply that $r\iota$ was null-homotopic and thus would be id_{S^1}, which contradicts Corollary 15.5.11. Therefore r cannot exist. ♮

Theorem 15.5.13 (Brouwer's fixed point)**.** Every continuous map $D^2 \longrightarrow D^2$ has a fixed point.

Proof. Let $f\colon D^2 \longrightarrow D^2$ be any continuous map. Suppose, for the sake of contradiction, that f admits no fixed point. We construct a map $r\colon D^2 \longrightarrow S^1$ as follows: for each $x \in D^2$, since $fx \neq x$, we can define a point $rx \in S^1$ given by $rx := \frac{fx-x}{\|fx-x\|}$. Notice r defines a retraction of the disk to the sphere, contradicting the result of Corollary 15.5.12, therefore there must exist at least one point of f. ♮

Chapter 16

Covering Spaces

16.1 The van Kampen Theorem

For the Fundamental Groupoid

Theorem 16.1.1 (van Kampen theorem for the fundamental groupoid)**.** Let X be a space, and $\mathcal{O}$ be a connected open cover[1] of X that is closed under finite intersections—which is a subcategory of `Top` whose morphisms are inclusions. Considering the functor $\Pi_1\colon \mathtt{Top} \to \mathtt{Grpd}$, the fundamental groupoid $\Pi_1 X$ is the colimit of the functor $\Pi_1|_{\mathcal{O}}$, that is:

$$\Pi_1 X \simeq \operatorname{colim}_{U \in \mathcal{O}} \Pi_1 U$$

in the category `Grpd`.

Proof. We'll show that $\Pi_1|$ Let $\mathcal{G}$ be a groupoid and consider the constant functor $C\colon \mathcal{O} \to \mathtt{Grpd}$ mapping $U \mapsto \mathcal{G}$ and $\iota \mapsto \mathrm{id}_{\mathcal{G}}$ for any object $U \in \mathcal{O}$ and morphism ι of $\mathcal{O}$. Let $\eta\colon \Pi_1|_{\mathcal{O}} \Rightarrow C$ be a natural transformation. The pair $(\mathcal{G}, \eta)$ forms a *cocone* over the functor $\Pi_1|_{\mathcal{O}}$: indeed, given an object $U \in \mathcal{O}$ there exists a morphism of groupoids $\eta_U\colon \Pi_1 U \to CU = \mathcal{G}$ and from naturality, given any inclusion $\iota\colon U \hookrightarrow V$ in $\mathcal{O}$ one has that

$$\begin{tikzcd}
\Pi_1 U \arrow[r, "\eta_U"] \arrow[d, "\Pi_1 \iota"'] & CU = \mathcal{G} \arrow[d, "C\iota = \mathrm{id}_{\mathcal{G}}"] \\
\Pi_1 V \arrow[r, "\eta_V"'] & CV = \mathcal{G}
\end{tikzcd}$$

commutes in `Grpd`—showing that $\eta_U = \mathrm{id}_{\mathcal{G}}\, \eta_U = \eta_V \circ \Pi_1 \iota$, hence compatibility Π_1 is satisfied, making $(\mathcal{G}, \eta)$ a cocone.

Consider the collection $(i_U\colon \Pi_1 U \to \Pi_1 X)_{U \in \mathcal{O}}$ of canonical inclusions of groupoids. To show the universal property of $\Pi_1 X$ we must construct a *unique morphism* of

[1]That is, composed of connected open subsets of X.

groupoids (that is, a functor between groupoids) $\chi\colon \Pi_1 X \to \mathcal{G}$ such that

$$\begin{tikzcd}
\mathcal{G} & \Pi_1 U \arrow[l, "\eta_U"'] \arrow[dl, bend left, "i_U"] \\
\Pi_1 X \arrow[u, dashed, "\chi"] &
\end{tikzcd} \tag{16.1}$$

commutes in Grpd for all $U \in \mathcal{O}$.

To that end, for every $x \in X$, if $x \in U$ let $\chi x \coloneqq \eta_U x$. For the morphisms of $\Pi_1 X$, consider any path $f \in \mathrm{Path}_X(x, y)$ on X. If $\operatorname{im} f$ lies interely in some object $U \in \mathcal{O}$, simply define $\chi[f] \coloneqq \eta_U[f]$—since $\mathcal{O}$ is closed under finite intersections, if $\operatorname{im} f \subseteq U \cap V$ for some $U, V \in \mathcal{O}$, then the image $\eta_\bullet[f]$ is independent of the choice of U or V. Consider now the case where f has an image not entirely contained in a single element of $\mathcal{O}$, but multiple ones, say $\operatorname{im} f \subseteq \bigcup_{j=1}^n U_j$ for some finite collection of sets $U_j \in \mathcal{O}$—since $\mathcal{O}$ is closed under finite intersections, this, we shall define a corresponding collection of paths $(f_j\colon I \to U_j)_{j=1}^n$ such that

$$f = f_n f_{n-1} \cdots f_2 f_1.$$

With this collection in hands we may define $\chi[f] \coloneqq \eta_{U_n}[f_n] \cdots \eta_{U_1}[f_1]$.

We must ensure that this is well defined: let $g \in \mathrm{Path}_X(x, y)$ be another path and suppose there exists a homotopy $\varepsilon\colon f \Rightarrow g$. Take a decomposition $(g_j\colon I \to V_j)_{j=1}^m$ for sets $V_j \in \mathcal{O}$. Consider a partition $(J_j \times S_j)_{j=1}^\ell$ of the square $I \times I$ such that $\operatorname{im} \varepsilon|_{J_j \times S_j} \subseteq U$ for some $U \in \mathcal{O}$, and such that $(J_j)_{j=1}^\ell$ is a refinement for the decompositions of f and g—that is, $f|_{J_j}$ and $g|_{J_j}$ are paths entirely contained in some set of $\mathcal{O}$. In this way we see that ε induces a collection of homotopies $(\varepsilon_j\colon f_j \Rightarrow g_j)_{j=1}^\ell$ proving that $[f_j] = [g_j]$ in some $\Pi_1 U$, therefore $\chi[f] = \chi[g]$. Hence χ is a uniquely defined functor satisfying Eq. (16.1). ♮

For the Fundamental Group

Theorem 16.1.2 (van Kampen theorem for the fundamental group)**.** Let $(X, x) \in \mathsf{Top}^{*/}$ be a path-connected space, and $\mathcal{O}$ be a path-connected open cover of X closed under finite intersections—and such that $x \in U$ for every $U \in \mathcal{O}$, therefore $\mathcal{O}$ is a subcategory of $\mathsf{Top}^{*/}$ whose morphisms are inclusions. Then the fundamental groupoid of X is the colimit of the functor $\pi_1|_{\mathcal{O}}\colon \mathcal{O} \to \mathsf{Grp}$, that is:

$$\pi_1(X, x) \simeq \operatorname{colim}_{U \in \mathcal{O}} \pi_1(U, x).$$

We first prove a particular case of the classical van Kampen theorem and after generalize.

Lemma 16.1.3. The van Kampen theorem for the fundamental group holds when $\mathcal{O}$ is finite.

Proof. Let G be a group and $C\colon \mathcal{O} \to \mathtt{Grp}$ be the constant functor on G and consider the cocone $(G, \eta\colon \pi_1|_{\mathcal{O}} \Rightarrow C)$ over the functor $\pi_1|_{\mathcal{O}}$. We'll construct a morphism of groups $\chi\colon \pi_1(X,x) \to G$. Recall that the inclusion functor $J\colon \pi_1(X,x) \to \Pi_1 X$ is an equivalence of categories, since $\pi_1(X,x)$ is a skeleton of $\Pi_1 X$ by Proposition 15.4.10. Define a quasi-inverse of J as follows: consider a collection $(\gamma_y)_{y\in X}$ of paths $\gamma_y \in \mathrm{Path}_X(x,y)$ where $\operatorname{im}\gamma_y \subseteq U$ when $y \in U$ and $\gamma_x := \mathrm{cons}_x$—this is possible because $\mathcal{O}$ is closed under finite intersections—then define $F\colon \Pi_1 X \to \pi_1(X,x)$ by mapping $f\colon a \to b$ to $Ff := \gamma_b f \gamma_a^{-1}\colon x \to x$.

Notice that the quasi-inverse functors J and F induce, for each $U \in \mathcal{O}$, a corresponding pair of quasi-inverse functors

$$F_U : \Pi_1 U \rightleftarrows \pi_1(U,x) : J_U.$$

Then we can construct a cocone $(G, \delta\colon \Pi_1|_{\mathcal{O}} \Rightarrow C)$ over the functor $\Pi_1|_{\mathcal{O}}$, where $\delta_U := \eta_U F_U\colon \Pi_1 U \to G$. By means of Theorem 16.1.1 there exists a *unique* morphism of groupoids $\xi\colon \Pi_1 X \to G$ (where G is interpreted as a groupoid with a single object) such that

$$\begin{tikzcd}
\Pi_1 U \arrow[r, "F_U"] \arrow[drr, bend right, "i_U"'] & \pi_1(U,x) \arrow[r, "\eta_U"] & G \\
 & & \Pi_1 X \arrow[u, dashed, "\xi"']
\end{tikzcd}$$

commutes in $\mathtt{Grpd}$ for every $U \in \mathcal{O}$. Define $\chi := \xi J\colon \pi_1(X,x) \to G$, and notice that since $\eta_U F_U = \xi i_U$ we can precompose with $J_U\colon \pi_1(U,x) \to \Pi_1 U$ and use that $F_U J_U = \mathrm{id}_{\pi_1(U,x)}$ to obtain that $\eta_U = \xi i_U J_U$. Notice however that given any $[g] \in \pi_1(U,x)$ one has $i_U J_U[g] = [g] \in \Pi_1 X$ while $J j_U[g] = [g] \in \Pi_1 X$ again—for the canonical inclusion $j_U\colon \pi_1(U,x) \hookrightarrow \pi_1(X,x)$—therefore $i_U J_U = J j_U$. This proves that $\eta_U = \xi J j_U = \chi j_U$ for every $U \in \mathcal{O}$, that is

$$\begin{tikzcd}
\pi_1(U,x) \arrow[r, "\eta_U"] \arrow[dd, bend right, "j_U"'] & G \\
 & \Pi_1 X \arrow[u, dashed, "\xi"'] \\
 & \pi_1(X,x) \arrow[u, "J"] \arrow[uu, dashed, bend right, "\chi"']
\end{tikzcd}$$

commutes in $\mathtt{Grp}$, which proves the universal property for the colimit $\pi_1(X,x)$. ♮

Proof of the Classical van Kampen Theorem

Let $\mathcal{O}$ be a path-connected open cover of X closed under intersections and composed of neighbourhoods of the chosen base-point x. Let $\mathfrak{F} \subseteq 2^{\mathcal{O}}$ be the category whose objects are the *finite* subsets of $\mathcal{O}$ of the cover that is *closed under finite intersections*, and morphisms are *inclusions*. Given any such subset $C \in \mathfrak{F}$, we know from Lemma 16.1.3 that the space $U_C := \bigcup_{U\in C} U$ satisfies

$$\pi_1(U_C, x) \simeq \operatorname{colim}_{U\in C} \pi_1(U,x). \tag{16.2}$$

- Let's prove that the colimit of the functor $\pi_1|_{\mathfrak{F}} \colon \mathfrak{F} \to \texttt{Grp}$—which maps each $C \in \mathfrak{F}$ to the group $\pi_1(U_C, x)$—is the fundamental group $\pi_1(X, x)$. Given any group G and a its corresponding constant functor $C_G \colon \mathfrak{F} \to \texttt{Grp}$ with $C_G C := G$, let $\eta \colon \pi_1|_{\mathfrak{F}} \Rightarrow C_G$ be a natural transformation. The pair (G, η) is then a cocone over the functor $\pi_1|_{\mathfrak{F}}$.

  We'll construct a unique morphism of groups $\chi \colon \pi_1(X, x) \to G$ satisfying the coherence of the cocones using the same technique from Theorem 16.1.1. If $f \colon x \to x$ is a loop contained entirely in a set $U_C \subseteq X$ for some $C \in \mathfrak{F}F$, we simply map $\chi[f] := \eta_C[f]$. If on the other hand f is not entirely contained in a single set, say that f is contained in the union $\bigcup_{j=1}^n U_j$ for sets $U_j \in C$ and define a collection of decompositions of f, namely $(f_j \colon I \to U_j)_{j=1}^n$, for which f is the result of the concatenation of paths. From the same argument as before, merely map $\chi[f] := \eta_{U_n}[f_n] \cdots \eta_{U_1}[f_1]$, which is well defined and unique[2]. Therefore one has

$$\operatorname{colim}_{C \in \mathfrak{F}} \pi_1(U_C, x) \simeq \pi_1(X, x). \tag{16.3}$$

- For the final part of the proof, we shall prove that the colimits of the functors $\pi_1|_{\mathcal{O}}$ and $\pi_1|_{\mathfrak{F}}$ agree so that the van Kampen theorem is true. Recalling Eq. (16.2), one has

$$\begin{aligned}\operatorname{colim}_{C \in \mathfrak{F}} \pi_1(U_C, x) &\simeq \operatorname{colim}_{C \in \mathfrak{F}}(\operatorname{colim}_{U \in C} \pi_1(U, x)) \\ &\simeq \operatorname{colim}_{(\mathcal{O}, \mathfrak{F})} \pi_1(U, x),\end{aligned}$$

  where $(\mathcal{O}, \mathfrak{F})$ is the category whose objects are pairs $(U, C) \in C \times \mathfrak{F}$, and morphisms are paired inclusions—also $\pi_1(-, x)|_{(\mathcal{O}, \mathfrak{F})} \colon (\mathcal{O}, \mathfrak{F}) \to \texttt{Grp}$ is defined to map (U, C) to $\pi_1(U, x)$. Notice that the functors $\pi_1(-, x)|_{\mathcal{O}}$ and $\pi_1(-, x)|_{(\mathcal{O}, \mathfrak{F})}$ factor as

$$\begin{tikzcd}
\mathcal{O} \arrow[rr, "\pi_1(-,x)|_{\mathcal{O}}"] \arrow[dr, "\iota"] & & \texttt{Grp} \\
& (\mathcal{O}, \mathfrak{F}) \arrow[ur, "\pi_1(-,x)|_{(\mathcal{O},\mathfrak{F})}"'] \arrow[ul, bend left, "p"] &
\end{tikzcd}$$

  Where $\iota U := (U, \{U\})$ and $p(U, C) := U$. Therefore one has an isomorphism

$$\operatorname{colim}_{U \in \mathcal{O}} \pi_1(U, x) \simeq \operatorname{colim}_{(U,C) \in (\mathcal{O}, \mathfrak{F})} \pi_1(U, x). \tag{16.4}$$

Therefore, by Eqs. (16.3) and (16.4) we have

$$\operatorname{colim}_{U \in \mathcal{O}} \pi_1(U, x) \simeq \pi_1(X, x)$$

as wanted.

[2]Simply refer to the proof of Theorem 16.1.1, now it should be clear that the proof follows exacly the same steps

16.2 Covering Spaces

Initial Constructions

Definition 16.2.1 (Covering space)**.** A surjective continuous map $p\colon E \to B$ is said to be a *covering space over* B if for every point $x \in B$, there exists a neigbourhood $U \subseteq B$ of x that is *evenly covered* by p, that is: there exists a bundle isomorphism over B between the pullback[3] of p over U and a product bundle $\pi\colon U \times E_x \to U$ with *discrete*[4] *fibre* $E_x = p^{-1}x$. Diagrammatically, one has that the diagram

$$\begin{tikzcd} U \times E_x \arrow[r, "\simeq"] \arrow[dr, "\pi"'] & p^{-1}U \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & E \arrow[d, "p"] \\ & U \arrow[r, hook] & B \end{tikzcd}$$

commutes in `Top`, where the topological isomorphism $U \times E_x \xrightarrow{\simeq} p^{-1}U$ is explicitly given by $(u, e) \mapsto e$.

Corollary 16.2.2. Given a covering $p\colon E \twoheadrightarrow B$, the induced morphism of groups

$$p_*\colon \pi_1(E, e) \rightarrowtail \pi_1(B, pe)$$

is a monomorphism for any $e \in E$.

Corollary 16.2.3. A covering space is a local topological isomorphism.

Proof. Given a covering space $p\colon E \to B$ and $x \in B$, we can consider the neighbourhood $U \subseteq B$ of x such that p admits a trivialization $\phi\colon p^{-1}U \xrightarrow{\simeq} U \times E_x$. Since E_x is discrete, then $U \times E_x \simeq \coprod_{e \in E_x} U \times \{e\}$ and certainly $U \times \{e\} \simeq U$. Therefore p yields an isomorphism

$$U_e := \phi^{-1}(U \times \{e\}) \simeq U, \tag{16.5}$$

proving that p is a local isomorphism. ♮

Notation 16.2.4 (Sheets)**.** The open sets $U_e \simeq U \times \{e\}$ of E from Eq. (16.5) are called *sheets*.

Proposition 16.2.5 (Covering projections are open)**.** Given a covering space $p\colon E \to B$, the projection p is open.

Proof. Consider the open cover $(U_x)_{x \in B}$ of B—where U_x is a neighbourhood of x such that there exists a topological isomorphism $p^{-1}U_x \simeq U_x \times \operatorname{Disc} E_x$, for each $x \in B$. Under the product topology, projections are open maps (see Lemma 13.3.5), therefore

[3] It should be noted that the pullback of p over U is nothing more than the induced bundle $\iota^* p$ given by the inclusion $\iota\colon U \hookrightarrow B$.

[4] Here, discrete means a space together with the discrete topology. We shall make use of the functor $\operatorname{Disc}\colon$ `Set` $\to$ `Top` mapping a bare set S to the topological space $\operatorname{Disc} S$ with underlying set S and endowed with the discrete topology.

$p|_{p^{-1}U_x}$ is an open map for each $x \in B$. Given any open set $V \subseteq E$, notice that $V = \bigcup_{x \in B}(W \cap p^{-1}U_x)$, therefore

$$pV = p\Big(\bigcup_{x \in B}(W \cap p^{-1}U_x)\Big) = \bigcup_{x \in B} p(W \cap p^{-1}U_x)$$

is the union of open sets, thus open—which shows that p is an open map. ♮

Lemma 16.2.6 (Fibre-wise diagonal of covering space is open and closed)**.** Let $p\colon E \to B$ be a covering space, and $E \times_B E$ be the pullback of p with itself. Then the diagonal of E with respect to the fibre product over B, namely

$$\Delta_B E = \{(e, e) \in E \times_B E : e \in E\},$$

is an *open and closed set* in $E \times_B E$.

Proof. First we prove that $\Delta_B E$ is open. Let $e \in E$ be any point and $U_{pe} \subseteq B$ be a neighbourhood of pe together with an isomorphism

$$p^{-1}U_{pe} \simeq U_{pe} \times E_{pe} = U_{pe} \times \{e\}.$$

Since $p^{-1}U_{pe}$ is open, it follows that $U_{pe} \times \{e\} \hookrightarrow E$ is open. Then the set $(E \times_B E) \cap (U_{pe} \times U_{pe})$ is an open neighbourhood of (e, e) in $E \times_B E$. This shows that $\Delta_B E$ is open.

To show that $\Delta_B E$ is closed, let $(x, y) \in E \times_B E$ with $x \neq y$. Let $U \subseteq B$ be a neighbourhood of $px = py$ that is evenly covered, and consider the induced sheets $U_x, U_y \subseteq E$ for x and y, respectively. From the assumption that x and y are distinct points, we have an empty intersection $U_x \cap U_y = \varnothing$—which shows that $U_x \times U_y$ is disjoint from $\Delta_B E$. Therefore $(E \times_B E) \cap (U_x \times U_y)$ is a neighbourhood for $(x, y) \in E \times_B E$ outside of $\Delta_B E$, showing that $\Delta_B E$ is closed. ♮

Theorem 16.2.7 (Lifting out of connected space)**.** Let $p\colon E \to B$ be a covering space, and $f\colon Y \to X$ be a continuous map, where Y is a connected space. Consider two *lifts of f along p*: continuous maps $\widehat{f}_1, \widehat{f}_2\colon Y \rightrightarrows E$ such that the triangle

$$\begin{tikzcd}
 & E \arrow[d, "p"] \\
Y \arrow[ur, shift left, "\widehat{f}_1"] \arrow[ur, shift right, "\widehat{f}_2"'] \arrow[r, "f"'] & X
\end{tikzcd}$$

commutes in `Top`. If there exists $y \in Y$ such that $\widehat{f}_1 y = \widehat{f}_2 y$, then the lifts agree everywhere $\widehat{f}_1 = \widehat{f}_2$.

Proof. Consider the pullback $E \times_B E$ of p with itself and consider the uniquely defined morphism $(\widehat{f_1}, \widehat{f_2})\colon Y \to E \times_B E$ making the diagram

$$\begin{tikzcd}
Y \arrow[dr, dashed] \arrow[drr, bend left, "\widehat{f_1}"] \arrow[ddr, bend right, "\widehat{f_2}"'] & & \\
& E \times_B E \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & E \arrow[d, "p"] \\
& E \arrow[r, "p"'] & B
\end{tikzcd}$$

commute. Define $\Delta_B E \subseteq E \times_B E$ for the diagonal of E with respect to the fibre product. Using Lemma 16.2.6 we know that $\Delta_B E$ is both open and closed, thus $V \coloneqq (\widehat{f_1}, \widehat{f_2})^{-1}(\Delta_B E) \subseteq Y$ is both open and closed in Y, which is a non-empty set since by hypothesis $\widehat{f_1}$ and $\widehat{f_2}$ agree at least in one point of Y. Since V is closed, then $Y \setminus V$ is open in Y and certainly disjoint from V. Since their union is the whole space Y, by the hypothesis that Y is connected, it must be the case that $V = Y$. Therefore $\widehat{f_1} = \widehat{f_2}$ as wanted. ♮

Proposition 16.2.8. Let $q\colon E \to B \times I$ be a locally trivial covering space with fibre F. Then B admits an open cover $\mathcal{U}$ for which q is trivial over $U \times I$ for each $U \in \mathcal{U}$.

Prove when needed

Coverings and G-Actions

Definition 16.2.9 (Properly discontinuous action)**.** Given a discrete topological group G, a left action $G \times E \to E$ is said to be *properly discontinuous* if for each pair $(g, x) \in G \times E$, where $g \neq e$, there exists a neighbourhood $U \subseteq E$ of x such that

$$U \cap gU = \varnothing.$$

In particular, every properly discontinuous action is free.

Definition 16.2.10 (G-principal covering space)**.** Let G be a discrete topological group. A *left G-principal covering space* is a covering $p\colon E \to B$ together with a properly discontinuous left action $G \circlearrowleft E$ for which $p(gx) = px$ for every pair $(g, x) \in G \times E$, and such that the induced action on the fibres is transitive.

Proposition 16.2.11. If $G \circlearrowleft X$ is a properly discontinuous action of a discrete group G on a space X, the canonical projection $q\colon X \twoheadrightarrow X/G$ is a G-principal covering space.

Proof. First of all, it is clear that $q(gx) = qx \in X/G$ for any pair $(g, x) \in G \times X$. Let $[x] \in X/G$ be any point, and $V \subseteq X$ be a neighbourhood of x such that $V \cap gV \neq \varnothing$ implies $g = e$, and define $U \coloneqq qV$. Notice that since G acts by topological isomorphisms, one has $q^{-1}U = \bigcup_{g \in G} gV$—now, since each gV is open, it follows that $q^{-1}U \subseteq X$ is open. Since X/G has the quotient topology, then $U \subseteq X/G$ is open.

Consider $q^{-1}U = \coprod_{g \in G} V_g$, where $V_g := gV$. Given any $g \in G$, suppose there exists $h \in G$ such that $V_g \cap V_h \neq \varnothing$—then for any $x' \in V_g \cap V_h$, one has $h^{-1}x' \in V_g \cap V$, and from construction this implies in $h^{-1}g = e$, thus $h = g$.

It remains for us to show a trivialization for q—we shall prove that $q|_{V_g} \colon V_g \to U$ is an isomorphism. Since quotient maps are open, it suffices to show that $q|_{V_g}$ is bijective. Let $[x'] \in U$ be any point, and take $x'' \in V$ such that $qx'' = [x']$, then in particular $gx'' \in gV$ and $q(gx'') = [gx''] = [x']$—thus $q|_{V_g}$ is surjective. For injectivity, let $x', x'' \in V$ be a pair of points such that $q(gx') = q(gx'')$, then $[x'] = [x'']$ in X/G, which implies in the existence of a point $h \in G$ such that $x' = hx''$, then $x'' \in V \cap V_h$—therefore $h = e$ and hence $x' = x''$, thus in particular $gx' = gx''$ as wanted. ♮

Definition 16.2.12 (Deck transformations)**.** Given a covering space $p \colon E \to B$, we define a group of automorphisms $\mathrm{Aut}(p)$ of the cover p to be composed of topological isomorphisms $\alpha \colon E \xrightarrow{\simeq} E$ such that $p\alpha = p$—such maps are called *deck transformations* of the covering p.

Example 16.2.13 (Translations)**.** Given a left G-principal covering $p \colon E \to B$, the *left translation* of E by $g \in E$, the isomorphism $\ell_g \colon E \xrightarrow{\simeq} E$ mapping $e \mapsto ge$, is a deck transformation of p. The collection of such maps $(\ell_g)_{g \in G}$ define a morphism of groups

$$\ell \colon G \longrightarrow \mathrm{Aut}(p).$$

Assume that E is a connected space. Let $x \in B$ be any point and consider any deck transformation $\alpha \in \mathrm{Aut}(p)$. From the fact that α is bijective and $p\alpha = p$, it acts as a permutation on the fibre $p^{-1}x$. Since p is a G-principal covering, then G acts transitively on the fibres: hence given any two points $e, e' \in p^{-1}x$, since $\alpha e \in p^{-1}(pe)$, it follows that there exists $g \in G$ such that $\alpha e' = g(\alpha e)$. Therefore α, using the connectedness of E, is uniquely defined by its image under a single point. This shows that, under these assumptions, ℓ is an isomorphism $G \simeq \mathrm{Aut}(p)$.

Proposition 16.2.14. Let $p \colon E \to B$ be a covering space. The following are properties concerning the automorphism group of p:

(a) If E is a connected space, then $\mathrm{Aut}(p)$ has a properly discontinuous action on E.

(b) If B is locally path connected and $H \subseteq \mathrm{Aut}(p)$ is a subgroup, then the induced map $q \colon E/H \to B$ is a covering.

Proof. (a) Consider any point $e \in E$ and deck transformation $g \in \mathrm{Aut}(p)$. Since p is a covering, choose $U \subseteq B$ to be an evenly covered neighbourhood of pe and let U_e be a sheet over U with $e \in U_e$. Suppose there exists a point $e' \in U_e \cap gU_e$ and notice that, since $pg = p$, we have $pe' = p(g^{-1}e')$.

Continueeeee

♮

Proposition 16.2.15. Let E be a simply connected space. If $G \circlearrowright E$ is a properly discontinuous action of a discrete topological group G, then the action induces an isomorphism of groups

$$\pi_1(E/G) \simeq G.$$

Proof. Let $[x] \in E/G$ be any base point, and choose $x \in q^{-1}[x]$—where $q\colon E \twoheadrightarrow E/G$ is the canonical projection. Define a map $\psi_x\colon \pi_1(E/G, [x]) \to G$ such that $\psi_x[\alpha] = g$ if and only if $\widehat{\alpha}_x(1) = gx$—where $\widehat{\alpha}_x\colon I \to E$ is the lift of α over q. We now show that ψ_x is the required isomorphism of groups:

- (Well defined) Suppose that $[\alpha] = [\beta]$ and that $\psi_x[\alpha] = g$ while $\psi_x[\beta] = h$. This means that $\widehat{\alpha}_x(1) = gx$ and $\widehat{\beta}_x(1) = hx$, but since $\widehat{\alpha}_x = \widehat{\beta}_x$, then $gx = hx$ and hence $h^{-1}gx = x$. Since G has a properly discontinuous action, there exists a neighbourhood $U \subseteq E$ of x such that $U \cap (h^{-1}g)U$ is *non-empty*, thus it is necessarily the case that $h^{-1}g = e$, thus $h = g$. This proves that $\psi_x[\alpha] = \psi_x[\beta]$.
- (Injective) Suppose $\psi_x[\alpha] = \psi_x[\beta]$, then $\widehat{\alpha}_x(1) = gx = \widehat{\beta}_x(1)$ for some $g \in G$. This shows that $\widehat{\alpha}_x \sim_{\mathrm{rel}\,\partial I} \widehat{\beta}_x$, hence $\alpha \sim_{\mathrm{rel}\,\partial I} \beta$ thus $[\alpha] = [\beta]$.
- (Surjective) From the fact that $\pi_1(E/G, [x])$ acts transitively on the fibres of q, it folows that ψ_x is surjective.
- (Group morphism) Let $\psi_x[\alpha] = g$ and $\psi_x[\beta] = h$ and define $k := \psi_x[\beta \cdot \alpha]$, so that

$$kx = (\widehat{\beta \cdot \alpha})_x(1) = (\widehat{\beta}_{\widehat{\alpha}_x(1)} \cdot \widehat{\alpha}_x)(1) = \widehat{\beta}_{gx}(1) = g\widehat{\beta}_x(1) = ghx,$$

  therefore $k = gh$.

$\natural$

Fibre Transport

Definition 16.2.16 (Homotopy lifting property)**.** A continuous map $p\colon E \to B$ is said to heve the *homotopy lifting property* (HLP) for a given space X if: for any homotopy $\eta\colon X \times I \to B$ and continuous map $a\colon X \to E$ such that $pax = \eta(x, 0)$ *exists* a homotopy $\delta\colon X \times I \to E$ such that the diagram

$$\begin{tikzcd} X \arrow[r, "a"] \arrow[d, "i_0"'] & E \arrow[d, "p"] \\ X \times I \arrow[r, "\eta"'] \arrow[ur, dashed, "\delta"] & B \end{tikzcd}$$

commutes in `Top`—the map $i_0\colon X \hookrightarrow X \times I$ is defined as $x \mapsto (x, 0)$

Check proof of these last fibration stuff—tom Dieck's book

Definition 16.2.17 (Fibration)**.** A continuous map $p\colon E \to B$ is said to be a *fibration* if it satisfies the homotopy lifting property for every space.

Theorem 16.2.18. A covering space is a fibration.

Proposition 16.2.19 (Unique path lifting)**.** Let $p\colon E \to B$ be a covering space, and $\gamma\colon I \to B$ be a path with initial point $\gamma 0 := pe \in B$. Then there exists a *unique lifting* $I \to E$ of γ that begins in e. Furthermore, any two paths $u, v\colon I \rightrightarrows E$ with $u0 = v0$ are *homotopic* if and only if their images are homotopic in B.

Lemma 16.2.20. Let $p\colon E \to B$ be a covering space and define $pe_0 := b_0$. Let Y be a path connected and locally path connected space, and $f\colon Y \to B$ be a continuous map with $fy_0 = b_0$. There exists a lift $Y \to E$ of f over p such that $y_0 \mapsto e_0$ if and only if

$$f_*\pi_1(Y, y_0) \subseteq p_*\pi_1(E, e_0).$$

If such a lift exists, then it's unique.

See Munkres, page 478

Chapter 17

Cellular Structures

17.1 Compactly Generated Spaces

Weak-Hausdorff Spaces

Definition 17.1.1 (Weak-Hausdorff space)**.** A space X is said to be *weak-Hausdorff* if for every compact Hausdorff space K and continuous map $g: K \to X$ one has that $g(K)$ is *closed* in X.

Corollary 17.1.2. A weak-Hausdorff space is T_1.

Proof. Let X be a weak-Hausdorff space and $x \in X$, we shall take advantage of the fact that X is T_1 if and only if $\{x\}$ is closed in X. Let K be a compact set and $\mathrm{cons}_x: K \to X$ be the constant map $k \mapsto x$, which is continuous. Since $gK = \{x\}$, and gK is closed by hypothesis, the property follows. ♮

Lemma 17.1.3. If X is weak-Hausdorff, K is a compact Hausdorff space, and $g: K \to X$ is a morphism, then gK is a *compact Hausdorff subspace of* X.

Proof. We already know from Proposition 14.2.2 that gK is compact, thus we shall only prove that gK is Hausdorff. Let $x, y \in gK$ be any two distinct points. Using Corollary 17.1.2 and Proposition 14.6.2 we can find *disjoint open sets* $U, V \subseteq K$ such that $g^{-1}x \subseteq U$ and $g^{-1}y \subseteq V$. Then the sets $K \setminus U$ and $K \setminus V$ are *closed* in K, thus *compact* (see Proposition 14.2.5). Therefore the sets $g(K \setminus U)$ and $g(K \setminus V)$ are both *closed in* X. From this we know that $gK \setminus g(K \setminus U)$ and $gK \setminus g(K \setminus V)$ are both *open in* gK, both of which are *disjoint* and contain x and y, respectively—which shows that X is Hausdorff. ♮

Definition 17.1.4 (Compactly closed)**.** A subset A of X is said to be *compactly closed* if, for every compact space K and morphism $g: K \to X$, the preimage $g^{-1}A$ is closed in K.

Proposition 17.1.5. If X is a weak-Hausdorff space, a subset $A \subseteq X$ is compactly closed if and only if the intersection of A with each compact subset of X is closed.

Proof.

- (⇒) Suppose A is compactly closed, and let $K \subseteq X$ be any compact subspace of X. If we consider the inclusion morphism $\iota : K \hookrightarrow X$, we have that $\iota^{-1}A = K \cap A$ is closed.
- (⇐) Assume the latter property, and let K be a compact space together with a morphism $g : K \to X$. Since gK is compact in X it follows from assumption that $A \cap gK$ is closed in X. Since g is continuous, then $g^{-1}(A \cap gK) = g^{-1}A$ is closed in K.

♮

k-Spaces

Definition 17.1.6 (k-space)**.** A space X is said to be a *k-space* if every compactly closed subspace of X is closed. The full-subcategory of `Top` whose objects are k-spaces will be denoted by k-`Top`.

Lemma 17.1.7 (k-ification)**.** Given a topological space (X, τ), we can transform X into a k-space by creating a topology τ_k where C is closed in (X, τ_k) if and only if C is compactly closed in (X, τ). We shall shortly denote the k-space (X, τ_k) by kX.

Definition 17.1.8 (k-ification functor)**.** We define the k-ification functor

$$k : \mathtt{Top} \longrightarrow k\text{-}\mathtt{Top}$$

to be the functor mapping topological spaces X to its k-ified space kX, and $kf := f$ for every topological morphism f.

Lemma 17.1.9. If X is a weak-Hausdorff space, then kX is also a weak-Hausdorff k-space.

Notation 17.1.10 (Products)**.** In what follows, we shall denote by $X \times_{\mathrm{c}} Y$ the *cartesian product* of spaces X and Y, which shall be endowed with the usual product topology. Moreover, from now on we shall reserve the notation

$$X \times Y := k(X \times_{\mathrm{c}} Y),$$

which may seem odd, but is a convention used throughout the literature and we'll adopt here.

Proposition 17.1.11 (Quotients)**.** The quotient of a k-space is a k-space.

Proof. Let X be a k-space and $q : X \twoheadrightarrow Y$ be a quotient map—we want to show that Y is a k-space. Let $A \subseteq Y$ be a compactly closed subset. We'll prove that $q^{-1}A$ is compactly closed in X, thus closed, yielding the conclusion that A is closed in Y. Given any compact space K and a morphism $\phi : K \to X$. Then the map $q\phi : K \to Y$ is such that $(q\phi)^{-1}A = \phi^{-1}(q^{-1}A)$ is closed in K since A is compactly generated—the result follows. ♮

Proposition 17.1.12. Let X and Y be k-spaces, and consider quotient maps $q\colon X \twoheadrightarrow X'$ and $p\colon Y \twoheadrightarrow Y'$. Then the product

$$q \times p\colon X \times Y \twoheadrightarrow X' \times Y'$$

is also a quotient map.

Proposition 17.1.13. A k-space X is *weak-Hausdorff* if and only if the *diagonal* ΔX is *closed* in $X \times X$.

Proof.
- ($\Leftarrow$) Assume that ΔX is closed in $X \times X$. Let K be a compact space and $\phi\colon K \to X$ a continuous map. Since X is a k-space, we may simply show that ϕK is compactly closed. To that end, let C be another compact space together with a morphism $\psi\colon C \to X$, then

$$\psi^{-1}(\phi K) = \pi_2(\phi \times \psi)^{-1}(\Delta X)$$

is a closed subset of C—where $\pi_2\colon K \times C \twoheadrightarrow C$ is the canonical second projection. This proves that ϕK is compactly closed.

- ($\Rightarrow$) Suppose X is a weak-Hausdorff k-space. It suffices to show that ΔX is compactly closed in $X \times_{\mathrm{c}} X$, so that ΔX is closed in the k-space $X \times X$. Let K be any compact set and $\phi\colon K \to X \times_{\mathrm{c}} X$ be a morphism. Considering the canonical projections $\pi_1, \pi_2\colon X \times_{\mathrm{c}} X \rightrightarrows X$, define the set

$$A := \pi_1(\phi K) \cup \pi_2(\phi K)$$

of X. The set A is constructed so that one has $\phi K \subseteq A \times_{\mathrm{c}} A$ and hence $\phi^{-1}(\Delta X) = \phi^{-1}(\Delta A)$. Since X is weak-Hausdorff, if we consider the maps $\pi_1\phi, \pi_2\phi\colon K \rightrightarrows X$, we find that $\pi_1\phi K$ and $\pi_2\phi K$ are both compact Hausdorff subspaces of X—hence A is a compact Hausdorff space. Since A is Hausdorff, it follows that ΔA is closed in $A \times_{\mathrm{c}} A$, therefore by continuity $\phi^{-1}(\Delta A) = \phi^{-1}(\Delta X)$ is closed in K.

♮

CG Spaces

Definition 17.1.14 (Compactly generated space)**.** A space X is said to be *compactly generated* (or, shortly, CG space) if X is a *weak-Hausdorff k-space*. The full-subcategory of `Top` composed of compactly generated spaces will be denoted by `cgTop`.

The k-ification functor $k\colon \mathtt{Top} \to k\text{-}\mathtt{Top}$ can also act on the category of weak-Hausdorff spaces `wH-Top`, producing compactly generated spaces:

$$k\colon \mathtt{wH\text{-}Top} \longrightarrow \mathtt{cgTop}.$$

Lemma 17.1.15. The canonical forgetful functor $j\colon \mathtt{cgTop} \to \mathtt{wH\text{-}Top}$ embedding CG spaces as weak-Hausdorff spaces is such that there exists a *bijection*

$$\mathrm{Mor}_{\mathtt{cgTop}}(X, kY) \simeq \mathrm{Mor}_{\mathtt{wH\text{-}Top}}(jX, Y)$$

for every $X \in \texttt{cgTop}$ and $Y \in \texttt{wH-Top}$. Then k is *right-adjoint* to j:

$$\texttt{cgTop} \underset{k}{\overset{j}{\rightleftarrows}} \texttt{wH-Top}$$

Example 17.1.16. As examples of compactly generated spaces we have:

(a) If X is *locally compact*, then it is a compactly generated space.

(b) If X is a *first-countable weak-Hausdorff* space, then it is compactly generated.

Lemma 17.1.17. Let X and Y be topological spaces, then

(a) If X is locally compact and Y is compactly generated, then

$$X \times Y = X \times_{c} Y.$$

(b) If both X and Y are weak-Hausdorff, then

$$X \times Y = kX \times kY.$$

(c) If both X and Y are compactly generated, then $X \times Y$ is a *product* in the category $\texttt{cgTop}$.

Lemma 17.1.18. Let X be a compactly generated space. A set-function $f\colon X \to Y$ is *continuous* if and only if the restriction $f|_K$ is continuous for every compact subspace $K \subseteq X$.

Proposition 17.1.19 (Quotients of CG spaces)**.** Let X be a compactly generated space, and $q\colon X \twoheadrightarrow Y$ be a quotient map. Then Y is *compactly generated* if and only if the preimage of the diagonal of Y,

$$(q \times q)^{-1}(\Delta Y),$$

is *closed* in $X \times X$.

Proof.
- ($\Rightarrow$) Suppose Y is compactly generated. From Proposition 17.1.13 we know that ΔY is a closed subspace of $Y \times Y$, therefore by continuity of $q \times q$ with respect to the product topology $X \times_{c} X$ (see Proposition 13.3.7) we obtain that $(q \times q)^{-1}(\Delta Y)$ is a closed subspace of $X \times_{c} X$, it follows[1] that $(q \times q)^{-1}(\Delta Y)$ is closed in $X \times X$.
- ($\Leftarrow$) Suppose that $(q \times q)^{-1}(\Delta Y)$ is closed in $X \times X$. Since q is a quotient map, then Y is a quotient space of a k-space X—therefore Y itself is a k-space.

  Since q is a quotient map and X is in particular a k-space, from Proposition 17.1.12 we know that $q \times q\colon X \times X \twoheadrightarrow Y \times Y$ is a quotient map. Since $(q \times q)^{-1}(\Delta Y)$ is closed in $X \times X$ then ΔY is closed in $Y \times Y$. Therefore by Proposition 17.1.13 we conclude that Y is compactly generated.

$\natural$

[1]Given any space Z, if $C \subseteq Z$ is a closed set then, given any compact set K and continuous map $g\colon K \to Z$, we are always ensured that $g^{-1}C$ is closed—this merely follows from the continuity of g. Therefore kZ preserves the closed sets of Z, in the sense that if C is closed in Z then C is also closed in kZ

Proposition 17.1.20. Let X and Y be compactly generated spaces, and $A \subseteq X$ be a *closed subspace*. Then for every morphism $f\colon A \to Y$ the attaching space $Y \cup_f X$ is *compactly generated*.

Definition 17.1.21 (Weak topology)**.** Let $(X_j)_{j\in J}$ be a collection of topological spaces together with inclusions $X_j \hookrightarrow X_{j+1}$. We define the *weak topology* on the union set $X := \bigcup_{j\in J} X_j$ as follows: a set $U \subseteq X$ is *open* if and only if $U \cap X_j$ is open in X_j for all $j \in J$.

Proposition 17.1.22 (Colimits)**.** Let $(X_j)_{j\in J}$ be a collection of compactly generated spaces together with inclusions $X_j \hookrightarrow X_{j+1}$ with closed images. Then the *colimit*

$$\operatorname{colim}_{j\in J} X_j = \bigcup_{j\in J} X_j$$

is *compactly generated*[2].

17.2 Construction of CW-Complexes

Definition 17.2.1 (Attaching cells to a space)**.** Let $f\colon \coprod_{j\in J} S_j^{n-1} \to X$ be a morphism of topological spaces, where J is a set and S_j^{n-1} is an indexed copy of the $(n-1)$-sphere. We shall consider the attaching space Y given by the *pushout*

$$\begin{tikzcd}
\coprod_{j\in J} S_j^{n-1} \arrow[r, "f"] \arrow[d, hook] \arrow[dr, phantom, "\ulcorner"] & X \arrow[d] \\
\coprod_{j\in J} D_j^{n} \arrow[r] & X \cup_f \left(\coprod_{j\in J} D_j^n\right) =: Y
\end{tikzcd}$$

where $D_j^n \hookleftarrow S^{n-1}$ denotes an indexed copy of the n-disk. We define the following notions concerning the triple (Y, X, f):

- An *n-cell* in Y is defined to be the image e_j^n of a disk D_j^n in Y, so that by construction we have $Y = X \cup \left(\bigcup_{j\in J} e_j^n\right)$.
- One can decompose f into a collection of maps $(f_j^n\colon D_j^n \to Y)_{j\in J}$, each of these is called a *characteristic map*.

Definition 17.2.2 (Filtered space)**.** A *filtered topological space* X is a space together with an increasing sequence $(X_n)_{n\in\mathbf{N}}$ of closed subspaces, with inclusions $X_n \hookrightarrow X_{n+1}$ for all $n \in \mathbf{N}$, and

$$X = \operatorname{colim}_{n\in\mathbf{N}} X_n = \bigcup_{n\in\mathbf{N}} X_n.$$

Where X is endowed with the *weak topology*.

[2] The colimit of the sequence is endowed with the weak topology (see Definition 17.1.21).

Definition 17.2.3 (Relative cell complex)**.** A *relative cell complex* is a pair (X, A) where A is a k-space, and X is a filtered space with:

- The initial space of the sequence $(X_n)_{n\in\mathbf{N}}$ associated to X is

$$X_0 = A \amalg \Big(\coprod_{j\in J_0} D^0_j\Big),$$

  where D^0_j is a copy of the 0-disk indexed by a set J_0.

- For each $n \in \mathbf{N}$ we have an associated attaching map $f_n\colon \coprod_{j\in J_{n+1}} S^n_j \to X_n$ such that

$$X_{n+1} = X_n \cup_{f_n} \Big(\coprod_{j\in J_{n+1}} D^{n+1}_j\Big)$$

  where S^n_j and D^{n+1}_j are copies of the n-sphere and $(n+1)$-disk indexed by a set J_{n+1}, respectively. In other words, the following diagram is a *pushout* in $\mathtt{Top}$:

$$\begin{tikzcd} \coprod_{j\in J_{n+1}} S^{n-1}_j \arrow[r, "f_n"] \arrow[d, hook] & X_n \arrow[d, hook] \\ \coprod_{j\in J_{n+1}} D^n_j \arrow[r] \arrow[ur, phantom, "\ulcorner"] & X_{n+1} \end{tikzcd}$$

  Each space X_n is called the *n-skeleton* of (X, A).

We say that (X, A) is a *finite* relative cell complex if X has finitely many n-cells for each $n \in \mathbf{N}$. Moreover, if $X = X_n$ we say that (X, A) is an *n-dimensional* relative cell complex.

Together with relative cell complexes we define a *cellular map* $\phi\colon (X, A) \to (Y, B)$ between relative cell complexes to be a continuous map $\phi(X_n) \subseteq Y_n$ for each $n \in \mathbf{N}$. We say that a pair (Y, A) is a *subcomplex* of (X, A) if Y is a subspace of X that is given by the union of A with cells of X.

A *CW-complex* is a relative cell complex $(X, \varnothing)$, that is, each skeleton X_n is formed exclusively by attaching n-cells on X_{n-1}.

Chapter 18

Bundles

18.1 Bundle

Initial Construction

Definition 18.1.1 (Bundle)**.** A *bundle* is defined to be a triple (E, p, B), where E and B are *spaces* and $p: E \to B$ is a *morphism*. We refer to B as the *base space*, while E is the *total space*, and p is the *projection* of the bundle. As usual, given any $b \in B$, we name the object $p^{-1}b$ the *fibre* of the bundle over b.

A *subbundle* of (E, p, B) is a bundle (E', p', B') such that E' and B' are *subspaces* of E and B, respectively, and $p' = p|_{E'}: E' \to B'$.

Definition 18.1.2 (Product bundle)**.** A *product bundle over B with fibre F* is a triple $(B \times F, p, B)$ where $p(x, y) := x$ is the first projection.

Definition 18.1.3 (Cross section)**.** Given a bundle (E, p, B), a *cross section* of the bundle is a *section* $s: B \to E$ of p—that is, $ps = \mathrm{id}_B$. As an immediate consequence of this definition, if (E', p', B') is a subbundle of (E, p, B), then s is a cross section of (E', p', B') if and only if $sB \subseteq E'$.

Lemma 18.1.4 (Product bundle cross section)**.** Given a product bundle $(B \times F, p, B)$, a cross section $s: B \to B \times F$ will always have the form $s = \mathrm{id}_B \times f$, where $f: B \to F$ is a uniquely defined morphism. Therefore the collection of cross sections of product bundles is in bijection with the collection of maps $B \to F$.

Proof. Let s be any cross section of $(B \times F, p, B)$, from the definition of s, there exists unique morphisms $s': B \to B$ and $f: F \to B$ such that $s = s' \times f$—it remains to be shown that s' is the identity on B. From definition of a product bundle, we know that $ps = s'$—since p is the projection of the first factor—moreover, from the definition of a cross section, $ps = \mathrm{id}_B$, therefore $s' = \mathrm{id}_B$ as wanted. ♮

Definition 18.1.5 (Stiefel variety)**.** We define the *Stiefel variety of orthonormal k-frames*[1] in $\mathbf{R}^n$ to be the *compact subspace* $\operatorname{Stie}_k \mathbf{R}^n \subseteq (S^{n-1})^k$ for which $(v_1, \dots, v_k) \in \operatorname{Stie}_k \mathbf{R}^n$ if and only if $\langle v_i, v_j \rangle = \delta_{ij}$—where $\langle -, - \rangle$ is the standard euclidean inner product.

Definition 18.1.6 (Grassmann variety)**.** The real *k-Grassmann variety* is defined to be the topological space $\operatorname{Grass}_k \mathbf{R}^n$ whose points are k-dimensional subspaces of $\mathbf{R}^n$, and endowed with the quotient topology generated by the map $\operatorname{Stie}_k \mathbf{R}^n \twoheadrightarrow \operatorname{Grass}_k \mathbf{R}^n$ given by $(v_1, \dots, v_k) \mapsto \langle v_1, \dots, v_k \rangle$. Since $\operatorname{Stie}_k \mathbf{R}^n$ is compact, it follows that $\operatorname{Grass}_k \mathbf{R}^n$ is also compact.

Example 18.1.7. Notice that $\operatorname{Stie}_1 \mathbf{R}^n = S^{n-1}$ and by the construction of the Grassmannian variety, we see that $\operatorname{Grass}_1 \mathbf{R}^n = \mathbf{RP}^{n-1}$.

Definition 18.1.8 (Bundle morphism)**.** Given two bundles (E, p, B) and (E', p', B'), a *morphism of bundles* $(E, p, B) \to (E', p', B')$ is a pair (u, f) of morphisms $u \colon E \to E'$ and $f \colon B \to B'$ such that the diagram

$$\begin{tikzcd} E \arrow[r, "u"] \arrow[d, "p"'] & E' \arrow[d, "p'"] \\ B \arrow[r, "f"'] & B' \end{tikzcd}$$

commutes in Top—equivalently, $u(p^{-1}B) \subseteq p'^{-1}(fB)$. The special case where the base spaces coincide, we define a *bundle morphism over B* (also refered to as B-morphism) $(E, p, B) \to (E', p', B)$ to be a morphism $u \colon E \to E'$ such that the triangle

$$\begin{tikzcd} E \arrow[rr, "u"] \arrow[dr, "p"'] & & E' \arrow[dl, "p'"] \\ & B & \end{tikzcd}$$

commutes in Top, which can equivalently be expressed as the condition $u(p^{-1}B) \subseteq p'^{-1}B$.

Given any two bundle morphisms $(u, f) \colon (E, p, B) \to (E', p', B')$ and $(v, g) \colon (E', p', B') \to (E'', p'', B'')$, we define the *composition* of those morphisms to be the pair

$$(v, g) \circ (u, f) \coloneqq (vu, gf) \colon (E, p, B) \longrightarrow (E'', p'', B''),$$

which is again a bundle morphism, since

$$\begin{tikzcd} E \arrow[r, "u"] \arrow[d, "p"'] & E' \arrow[r, "v"] \arrow[d, "p'"] & E'' \arrow[d, "p''"] \\ B \arrow[r, "f"'] & B' \arrow[r, "g"'] & B'' \end{tikzcd}$$

is a commutative diagram in Top.

[1] A k-frame in an n-dimensional vector space is an ordered collection of k linearly independent vectors.

Example 18.1.9 (Cross section as bundle morphism)**.** Notice that a cross section is nothing more than a bundle morphism over B of the form $s\colon (B, \mathrm{id}_B, B) \to (E, p, B)$.

Definition 18.1.10 (Category of bundles)**.** We denote by $\mathtt{Bun}$ the category composed of bundles and bundle morphisms. Given a space B, we can also define a full subcategory $\mathtt{Bun}_B$ of $\mathtt{Bun}$, whose objects are bundles with base space B and bundle morphisms over B.

Definition 18.1.11 (Fibre of a bundle)**.** We say that a space F is *the fibre* of a bundle (E, p, B) if there exists a topological isomorphism $p^{-1}b \simeq F$ for every $b \in B$. A bundle (E, p, B) is said to be *trivial with fibre F* if there exists a bundle B-isomorphism $(E, p, B) \simeq (B \times F, p, B)$.

Universal Properties

Proposition 18.1.12 (Products in $\mathtt{Bun}$)**.** Given a family of bundles $(E_j, p_j, B_j)_{j \in J}$, we define the *product* of this family of bundles to be the bundle

$$\Big(\prod_{j\in J} E_j, \prod_{j\in J} p_j, \prod_{j\in J} B_j\Big),$$

which is defines a product in the category $\mathtt{Bun}$.

Proposition 18.1.13 (Pullbacks in $\mathtt{Bun}_B$)**.** Given two bundles $\xi = (E, p, B)$ and $\xi' = (E', p', B)$, define

$$E \oplus E' := \{(x, x') \in E \times E' : px = p'x'\},$$

and $q\colon E \oplus E' \to B$ to be the morphism $q(x, x') := px = p'x'$. Then the triple

$$\xi \oplus \xi' := (E \oplus E', q, B),$$

called *fibre product over B* of ξ and ξ', is the *pullback* of the pair (ξ, ξ') in the category $\mathtt{Bun}_B$.

Induced Bundle

Definition 18.1.14 (Induced bundle)**.** Let $\xi = (E, p, B)$ be a bundle, and $f\colon B_0 \to B$ be a continuous map. There exists an *induced bundle* of ξ under f, denoted $f^*\xi$, with base space B_0 and total space E_0 defined as the pullback:

$$\begin{tikzcd} E_0 \arrow[r, "f_\xi"] \arrow[d, "p_0"'] \arrow[dr, phantom, "\lrcorner", very near start] & E \arrow[d, "p"] \\ B_0 \arrow[r, "f"'] & B \end{tikzcd}$$

Explicitly, E_0 consists of pairs $(b_0, x) \in B_0 \times E$ such that $fb_0 = px$. The projection of the bundle $f^*\xi$ is the map $p_0\colon E_0 \to B_0$ given by $(b_0, x) \mapsto b_0$.

The mapping $f_\xi\colon E_0 \to E$ given by $(b_0, x) \mapsto x$ induces a morphism of bundles $(f_\xi, f)\colon f^*\xi \to \xi$, the so called *canonical morphism of an induced bundle*.

Proposition 18.1.15. Let $\xi = (E, p, B)$ be a bundle and $f: B_0 \to B$ be a continuous map. Considering the canonical morphism $(f_\xi, f): f^*\xi \to \xi$, for each $b_0 \in B_0$ the restricted map

$$f_\xi : p_0^{-1} b_0 \xrightarrow{\simeq} p^{-1}(f b_0)$$

is a *topological isomorphism*. Furthermore, given a bundle $\eta = (E', p', B_0)$ and a morphism $(v, f): \eta \to \xi$, there *exists a unique B_0-morphism $w: \eta \to f^*\xi$* such that $f_\xi w = v$. In other words, the following diagram

$$\begin{tikzcd}
E' \arrow[r, "v"] \arrow[dr, dashed, "w"'] & E \\
 & E_0 \arrow[u, "f_\xi"']
\end{tikzcd}$$

commutes in Top.

Proof. For the first part of the proposition, recalling the definition one has that the fibre $p_0^{-1} b_0$ is composed of pairs $(b_0, x) \in b_0 \times E$ such that $px = f b_0$—that is, $p_0^{-1} b_0 = b_0 \times p^{-1}(f b_0)$. Since $f_\xi(b_0, x) = x$, then its restriction is a topological isomorphism with local inverse $x \mapsto (b_0, x)$.

To prove the second part, define a map $w: E' \to E_0$ by $w := (p', v)$, therefore

$$\begin{tikzcd}
E' \arrow[rr, "w"] \arrow[dr, "p'"'] & & E_0 \arrow[dl, "p_0"] \\
 & B_0 &
\end{tikzcd}$$

commutes, showing that w is a B_0-morphism. Moreover, we have for any $y \in E'$ that

$$f_\xi w y = f_\xi(p' y, v y) = v y,$$

therefore $f_\xi w = v$ as wanted. For the uniqueness of w, suppose $\ell: E' \to E_0$ satisfies both $p_0 \ell = p'$ and $f_\xi \ell = v$, then since $\ell = (p_0 \ell, f_\xi \ell) = (p', v)$, this shows that $\ell = w$. ♮

Proposition 18.1.16. Given any continuous map $f: B_0 \to B$, the induced bundle construction via f is a functor

$$f^*: \mathtt{Bun}_B \to \mathtt{Bun}_{B_0}.$$

Moreover, given any morphism $u: \xi \to \eta$ in $\mathtt{Bun}_B$, the diagram

$$\begin{tikzcd}
 & E(f^*\eta) \arrow[rr, "f_\eta"] \arrow[dd] & & E\eta \arrow[dd] \\
E(f^*\xi) \arrow[ur, "f^*u"] \arrow[rr, "f_\xi"] \arrow[dr] & & E\xi \arrow[ur, "u"] \arrow[dr] & \\
 & B_0 \arrow[rr, "f"'] & & B
\end{tikzcd}$$

Proof. For the functoriallity, given any B-bundle morphism $u\colon \xi \longrightarrow \eta$, the associated map $f^*u\colon f^*\xi \longrightarrow f^*\eta$ is canonically given by the mapping $(b_0, x) \mapsto (b_0, ux)$, which is a B_0-morphism of bundles. Moreover, if we consider the identity morphism $\mathrm{id}_\xi\colon \xi \longrightarrow \xi$ one has

$$f^*(\mathrm{id}_\xi)(b_0, x) = (b_0, x) = \mathrm{id}_{f^*\xi}(b_0, x),$$

therefore $f^* \mathrm{id}_\xi = \mathrm{id}_{f^*\xi}$. Also, if $v\colon \eta \longrightarrow \zeta$ is any other bundle morphism, we have

$$f^*(vu)(b_0, x) = (b_0, vux) = f^*(v)(b_0, ux) = f^*(v)(f^*(u)(b_0, x)),$$

that is, $f^*(vu) = f^*v \circ f^*u$. This finishes the proof that f^* is indeed a functor.

The only additional information the diagram brings is that uf_ξ should equal $f_\eta f^*u$, and this is what we'll show. Let $(b_0, x) \in E(f^*\xi)$ be any point, then

$$uf_\xi(b_0, x) = ux = f_\eta(b_0, ux) = f_\eta(f^*u)(b_0, x),$$

which proves the commutativity of the diagram. ♮

Proposition 18.1.17 (Functorial transitivity of the induced bundle)**.** Consider continuous maps $B_1 \xrightarrow{g} B_0 \xrightarrow{f} B$, and a bundle $\xi = (E, p, B)$. Then the following are properties conserning the induced bundles over ξ:

(a) Considering the identity map $\mathrm{id}\colon B \longrightarrow B$, there exists a bundle B-isomorphism

$$\mathrm{id}^* \xi \simeq \xi.$$

(b) There exists a B_1-isomorphism of bundles

$$g^*f^*\xi \simeq (fg)^*\xi.$$

Proof. Notice that the morphisms of bundles $\xi \longrightarrow \mathrm{id}^* \xi$ given by $x \mapsto (px, x)$ has an inverse $(b, x) \mapsto x$, proving the first isomorphism. For the second item, define $u\colon g^*f^*\xi \longrightarrow (fg)^*\xi$ by the mapping $u(b_1, (b_0, x)) := (b_1, x)$ then by the fact that $(b_1, (b_0, x)) \in E(g^*f^*\xi)$ if and only if $gb_1 = p_1(b_0, x) = b_0$, we can conclude that u is an isomorphism of B_1-bundles. ♮

18.2 Fibre Bundles

Definition 18.2.1 (Bundle projection)**.** Let X, B, and F be Hausdorff spaces. We say that a continuous map $p\colon X \longrightarrow B$ is a *bundle projection* with *fibre F* if for each $b \in B$ there exists a neighbourhood $U \subseteq B$ of b such that there is a *topological isomorphism*

$$\phi\colon U \times F \longrightarrow p^{-1}U, \quad \text{such that} \quad p\phi(x, y) = x$$

for all $x \in U$ and $y \in F$—the map ϕ is called a *trivialization of the bundle over U*. This means that on the set $p^{-1}U$, the map p is a *projection* of the type $U \times F \twoheadrightarrow U$.

Definition 18.2.2 (Fibre bundle)**.** Let G be a topological group acting *effectively* on a Hausdorff space F—seen as a group of topological isomorphisms. Let X and B be Hausdorff spaces. We define a *fibre bundle* (or simply *bundle*) over the *base space* B with *total space* X, *fibre* F, and *structure group* G, to be a pair (p, Φ) where $p: X \to B$ is a *bundle projection* and Φ is a collection of trivializations of p (as described in Definition 18.2.1)—the members of Φ will be called *charts* over U—such that:

- For each $b \in B$ there exists a neighbourhood $U \subseteq B$ of b and a chart $\phi \in \Phi$ of the form $\phi: U \times F \to p^{-1}U$.
- Given a chart $\phi: U \times F \to p^{-1}U$, member of Φ, then any subset $V \subseteq U$ is such that the restriction $\phi|_{V\times F}$ belongs to the family Φ.
- Given any pair of charts $\phi, \psi \in \Phi$ over a common open set U, there exists a continuous map $\theta: U \to G$ such that

$$\psi(u, y) = \phi(u, \theta(u)(y)).$$

- The family Φ is *maximal* among the collections satisfying the previous properties.

The fibre bundle is said to be *smooth* if each object above is a *smooth manifold* and all maps are *smooth morphisms*.

18.3 Vector Bundle

First definitions

Definition 18.3.1 (Vector bundle)**.** A (topological) *vector bundle* is a fibre bundle with a fibre $\mathbf{R}^n$ and structure group contained in $\mathrm{GL}_n(\mathbf{R})$. Given a vector bundle ξ, we denote its total space by $E\xi$ and base space by $B\xi$.

Notation 18.3.2. Given a vector bundle $\xi = (E, p, B)$, we denote by $\xi_b := p^{-1}b$ (which can also be denoted by E_b) the fibre of $b \in B$ over p.

Definition 18.3.3 (Morphism of vector bundles)**.** If $\xi = (E, p, B)$ and $\xi' = (E', p', B')$ are any two vector bundles, we define a *bundle morphism* $\xi \to \xi'$ is a pair (u, f) of continuous maps $u: E \to E'$ and $f: B \to B'$ such that the diagram

$$\begin{tikzcd} E \arrow[r, "u"] \arrow[d, "p"'] & E' \arrow[d, "p'"] \\ B \arrow[r, "f"'] & B' \end{tikzcd}$$

commutes and the restriction $u_b: \xi_b \to \xi'_{fb}$ is $\mathbf{R}$-*linear* for every $b \in B$.

Definition 18.3.4. We denote by `VecBun` the category of vector bundles and morphisms between them. Furthermore, given a base space B, we also define a full subcategory $\mathtt{VecBun}_B$ of vector bundles with base B and B-morphisms.

Charts and Atlases

Definition 18.3.5 (Vector bundle chart & atlas)**.** Given a vector bundle (E, p, B), we define an n-dimensional vector bundle chart (U, ϕ), for some open set $U \subseteq B$, to be a topological isomorphism

$$\phi\colon p^{-1}U \xrightarrow{\simeq} U \times \mathbf{R}^n$$

such that the diagram

$$\begin{tikzcd} p^{-1}U \arrow[r, "\phi"] \arrow[d, "p"'] & U \times \mathbf{R}^n \arrow[dl, bend left, "\pi_1"] \\ U & \end{tikzcd}$$

commutes in `Top`—and π_1 is the projection of the first factor. The isomorphism ϕ induces a collection $(\phi_x\colon p^{-1}x \xrightarrow{\simeq} \mathbf{R}^n)_{x\in U}$ of isomorphisms given by the composition

$$\begin{tikzcd} p^{-1}x \arrow[r, "\simeq", "\phi"'] \arrow[rr, bend left, "\phi_x"] & x \times \mathbf{R}^n \arrow[r, "\simeq", "\pi_2"'] & \mathbf{R}^n \end{tikzcd}$$

Therefore, given any $y \in p^{-1}U$, if $y \in p^{-1}x$, then $\phi y = (x, \phi_x y)$.

Definition 18.3.6. A family $\Phi \coloneqq (U_j, \phi_j)_{j\in J}$ of vector bundle charts on (E, p, B) with domain covering B and values in $\mathbf{R}^n$ is said to form a *vector bundle atlas* for (E, p, B) if for any two vector bundle charts (U, ϕ) and (V, ψ) of Φ we have:

(a) For every $x \in U \cap V$, the transition map

$$\psi_x\phi_x^{-1}\colon \mathbf{R}^n \xrightarrow{\simeq} \mathbf{R}^n$$

is an $\mathbf{R}$-*linear topological isomorphism*.

(b) The map $g\colon U \cap V \to \mathrm{GL}_n(\mathbf{R})$ sending $x \mapsto \psi_x\phi_x^{-1}$ is *continuous*.

If such conditions are satisfied, by the requirement of item (b), the atlas Φ induces a collection of continuous maps

$$(g_{ij}\colon U_i \cap U_j \longrightarrow \mathrm{GL}_n(\mathbf{R}))_{(i,j)\in J\times J},$$

called *cocycle* of Φ. For any three $i, j, k \in J$, and $x \in U_i \cap U_j \cap U_k$, one has that

$$g_{ij}(x)g_{jk}(x) = ((\phi_i)_x(\phi_j^{-1})_x)((\phi_j)_x(\phi_k^{-1})_x) = (\phi_i)_x(\phi_k^{-1})_x = g_{ik}x.$$

Moreover, for any $j \in J$ we find $g_{ii}x = \mathrm{id}_{\mathbf{R}^n}$. The tuple (E, p, B, Φ), is said to be a *vector bundle with n-dimensional fibre*. Furthermore an atlas *for* (E, p, B) is a subatlas of Φ.

For every $x \in B$, we can endow the fibre E_x with the structure of an $\mathbf{R}$-vector space for which $\phi_x\colon E_x \xrightarrow{\simeq} \mathbf{R}^n$ is an $\mathbf{R}$-linear isomorphism, independently of the choice of $(\phi, U) \in \Phi$. We'll thus call E an *n-plane bundle*

Definition 18.3.7 (Zero section)**.** Given a vector bundle ξ, we shall denote by Zero: $B\xi \to E\xi$ the *zero section* of ξ, that is, the mapping $x \mapsto 0 \in E_x$.

Definition 18.3.8 (Trivial vector bundle)**.** The *n-dimensional trivial vector bundle* is the vector bundle

$$\varepsilon_B^n := (B \times \mathbf{R}^n, p, B, \Phi),$$

where p is the projection of the first component, and Φ is the unique maximal vector bundle atlas on ε_B^n containing identity maps for each open set of $B \times \mathbf{R}^n$. A vector bundle over B is said to be trivial if it is isomorphic—such isomorphism is said to be a *trivialization*—to ε_B^n for some $n \in \mathbf{N}$.

Definition 18.3.9 (Smooth vector bundle)**.** A vector bundle ξ is said to be smoth if its associated spaces are smooth manifolds and the projection is a C^∞-morphism.

Example 18.3.10 (Tangent bundle)**.** Given a smooth n-manifold M, we define the *tangent bundle* (see the discussion at Section 21.4) of M to be the vector bundle (TM, π, M). For each chart $\phi: U \to \mathbf{R}^n$ we define a vector bundle chart $\pi^{-1}U \to U \times \mathbf{R}^n$ mapping tangent vectors $X \mapsto (x, \phi_{*x}X)$.

Given any C^∞-morphism $f: M \to N$ of manifolds, there is an induced vector bundle morphism $Tf: TM \to TN$.

Theorem 18.3.11 (Isomorphism of vector bundles)**.** Let $u: \xi \to \eta$ be a B-morphism of n-dimensional vector bundles. Then u is a B-isomorphism of vector bundles if and only if $u_b: \xi_b \to \eta_b$ is an $\mathbf{R}$-linear isomorphism for each $b \in B$.

Proof. ($\Rightarrow$) If u is a B-isomorphsim of vector bundles, then the restriction mapping $u^{-1}|_{\eta_b}: \eta_b \to \xi_b$ is an inverse for u_b.

($\Leftarrow$) For the converse, suppose that the restriction $u_b: \xi_b \xrightarrow{\simeq} \eta_b$ is a linear isomorphism for all $b \in B$. Construct a set-function $v: E\eta \to E\xi$ such that $v|_{\eta_b} := u_b^{-1}$—we must show that v is continuous. Let $U \subseteq B$ be any open set, and $\phi: \xi^{-1}U \xrightarrow{\simeq} U \times \mathbf{R}^n$ and $\psi: \eta^{-1}U \xrightarrow{\simeq} U \times \mathbf{R}^n$ be vector bundle charts for ξ and η, respectively. Then if we consider the map $\psi u \phi^{-1}$, we find that it has the form $(b, x) \mapsto (b, f_b x)$—where $b \mapsto f_b$ is a mapping $U \to \mathrm{GL}_n(\mathbf{R})$. On the other hand, the map $\phi v \psi^{-1}$ is of the form $(b, x) \mapsto (b, f_b^{-1} x)$—where $b \mapsto f$ is again a map $U \to \mathrm{GL}_n(\mathbf{R})$. ♮

Definition 18.3.12 (Whitney sum)**.** Given vector bundles $\xi, \eta \in \mathtt{VecBun}_B$, we define the *Whitney sum* of ξ and η to be the fibre product $\xi \oplus \eta \in \mathtt{VecBun}_B$, where $(\xi \oplus \eta)_b = \xi_b \oplus \eta_b$ has the structure of the direct sum of vector spaces. Given charts $\phi: U \times \mathbf{R}^n \to \xi^{-1}U$ and $\psi: U \times \mathbf{R}^m \to \eta^{-1}U$ for the respective vector bundles, we define an induced vector bundle chart for $\xi \oplus \eta$ to be

$$\phi \oplus \psi: U \times \mathbf{R}^{n+m} \to \xi^{-1}U \oplus \eta^{-1}U.$$

Induced Vector Bundles

Proposition 18.3.13 (Induced vector bundle)**.** Let $\xi = (E, p, B)$ be an n-dimensional vector bundle, and$f: B_0 \to B$ be a continuous map. Then the induced bundle $f^*\xi = (E_0, p_0, B_0)$ admits a unique vector bundle structure, and the canonical morphism $(f_\xi, f): f^*\xi \to \xi$ is a vector bundle morphism. Moreover, $f_\xi: p_0^{-1}b_0 \to \xi^{-1}b$ is a a linear isomorphism (refer to Definition 18.1.14).

Proof. From Definition 18.1.14 we know that $p_0^{-1}b_0 = b_0 \times p^{-1}(fb_0)$. Consider a pair of fibre points $(b_0, x), (b_0, y) \in p_0^{-1}b_0$ and define vector space structures $(b_0, x) + (b_0, y) := (b_0, x + y)$, and $\lambda(b_0, x) = (b_0, \lambda x)$ for any $\lambda \in \mathbf{R}$. We know that the restriction $f_\xi: p_0^{-1}b_0 \to p^{-1}(fb_0)$ is a topological isomorphism (see Proposition 18.1.15), but via the linear structure just defined in $p_0^{-1}b_0$, we see that f_ξ is also a linear map, thus it's an isomorphism of vector spaces. This shows that $p_0^{-1}b_0$ is restrained to this linear structure, which proves uniqueness.

For a local trivialization of $f^*\xi$, consider, for any open set $U \subseteq B$, a vector bundle chart $h: U \times \mathbf{R}^n \to p^{-1}U$ for ξ. We can construct a map $h': f^{-1}U \times \mathbf{R}^n \to p_0^{-1}(f^{-1}U)$ mapping $(b_0, x) \mapsto (b_0, h(fb_0, x))$, which forms a vector bundle chart for $f^*\xi$ over the open set $f^{-1}U$. ♮

Homotopy Properties of Vector Bundles

Lemma 18.3.14 (Local trivialization)**.** Let $\xi = (E, p, B \times I)$ be a C^r vector bundle for some $0 \leqslant r \leqslant \infty$. Then every point $b \in B$ admits a neighbourhood $V \subseteq B$ such that $\xi|_{V \times I}$ is *trivial*.

Proof. Since ξ is locally trivial and I is compact, then given $b \in B$ consider a neighbourhood $V_j \subseteq B$ of b and a partition $0 = t_0 < \cdots < t_m = 1$ of I such that ξ is trivial in a neihbourhood of $V_j \times I_j := V_j \times [t_{j-1}, t_j]$ for each $0 < j \leqslant m$. Define $V := \bigcap_{j=1}^m V_j$ and let $(U_j)_{j=1}^m$ be a collection where $U_j \subseteq I$ is a neighbourhood of I_j such that $\xi|_{V \times U_j}$ is trivial.

We do induction on m. For the base case $m = 1$ it follows by construction that $\xi|_{V \times U_1}$ is trivial and since $I_1 = I$ then $U_1 = I$. Now if $m > 1$ we can proceed assuming that the case is true for all $n < m$: for each $1 \leqslant j \leqslant m$ we find a neighbourhood $J \subseteq I$ of the interval $I_1 \cup \cdots \cup I_j = [0, t_j]$ for which $\xi|_{V \times J}$ is trivial. This hints at the fact that it is sufficient to prove this construction for the case $m = 2$, and this is what we'll set out to do.

Take two subintervals $U_1 := [0, b]$ and $U_2 := [a, 1]$ where $0 < a < b < 1$. Let $\phi_j: p^{-1}(V \times U_j) \to (V \times U_j) \times \mathbf{R}^n$, for $j \in \{1, 2\}$, be C^r-charts (where we assumed ξ to be n-dimensional). Associated with these maps is the cocycle mapping $g_{21}: V \times (U_1 \cap U_2) \to \mathrm{GL}_n(\mathbf{R})$ mapping $x \mapsto \phi_{1\,x}\phi_{2\,x}^{-1}$. Let $a < c < b$ and consider a C^r-map $\lambda: U_2 \to U_1 \cap U_2$ for which $\lambda|_{[a,c]} = \mathrm{id}_{[a,c]}$. Defining $\mu := \mathrm{id}_V \times \lambda: V \times [a, 1] \to V \times [a, b]$, we can construct a map

$$h := g\mu: V \times [a, 1] \longrightarrow \mathrm{GL}_n(\mathbf{R})$$

such that $h|_{V\times[a,c]} = g|_{V\times[a,c]}$. This allows us to construct a trivialization ψ where we define, for each $x \in V \times I$, a map $\psi_x \colon \xi_x \to \mathbf{R}^n$, and

$$\psi_x := \begin{cases} \phi_{1\,x}, & \text{if } x \in V \times [0, c] \\ h(x)\phi_{2\,x}, & \text{if } x \in V \times [a, 1] \end{cases}$$

so that by construction for any $x \in [a, c]$ one has $h(x)\phi_{2\,x} = (\phi_{1\,x}\phi_{2\,x}^{-1})\phi_{2\,x} = \phi_{1\,x}$, showing that ψ_x is well defined. This defines a trivialization $\psi \colon p^{-1}(V\times I) \to (V\times I)\times\mathbf{R}^n$ as wanted. ♮

Corollary 18.3.15. Any C^r vector bundle, with $0 \leqslant r \leqslant \infty$, over an *interval* is *trivial*

Oriented Vector Bundles

Definition 18.3.16. Let ξ be a vector bundle. We define an *orientation* for ξ to be a family $\omega = (\omega_x)_{x\in B\xi}$ where ω_x is an orientation for the vector space fibre ξ_x such that ξ has an atlas for which: every chart $\phi \colon \xi|_U \to \mathbf{R}^n$ in the atlas Φ of ξ has an *orientation preserving* map $\phi_x \colon (\xi_x, \omega_x) \to (\mathbf{R}^n, \omega_n)$[2]. If this is the case, ω is said to be a *coherent* family of orientations of the fibres and Φ is an *oriented atlas* of ξ.

Lemma 18.3.17. Let $f \colon \eta \xrightarrow{\cong} \xi$ be an isomorphism in `VecBun`, and let ω be an orientation for ξ. Then there exists a unique orientation θ for η for which f preserves the orientations of the fibres.

Proposition 18.3.18. Every vector bundle over a simply connected manifold admits an orientation.

Proposition 18.3.19. A vector bundle ξ over a manifold M is *orientable* if and only if every loop $\gamma \in \Omega M$ preserves the orientation of $\xi_{\gamma(0)}$.

Corollary 18.3.20. An orientable vector bundle over a connected manifold has only two orientations.

[2]The use of ω_n denotes the standard orientation for the euclidean space $\mathbf{R}^n$ (see Definition 6.6.13).

Part V

Homotopy Theory

Chapter 19

Simplicial Sets

19.1 Simplex Category

Construction

Definition 19.1.1 (Skeletal simplex category)**.** We denote by $\mathbf{\Delta}$ the category whose objects are natural numbers $[n] := \{0 \leqslant 1 \leqslant \ldots \leqslant n\}$, and whose morphisms $\phi \colon [n] \to [m]$ are order-preserving—that is, if $i \leqslant j$ then $\phi(i) \leqslant \phi(j)$. The category $\mathbf{\Delta}$ is called the *skeletal simplex category*.

Along with the simplex category comes distinguished morphisms:

- For each $0 \leqslant j \leqslant n$, we denote by $\delta^n_j \colon [n-1] \rightarrowtail [n]$ the order-preserving *injective* morphism skipping the j-th value:

$$\delta^n_j(i) := \begin{cases} i, & \text{if } i < j, \\ i+1, & i \geqslant j. \end{cases}$$

  This morphism is called *elementary faces*.

- For each $0 \leqslant j \leqslant n$, we denote by $\sigma^n_j \colon [n+1] \twoheadrightarrow [n]$ the *surjective* morphism repeating the j-th value twice and every other value only once:

$$\sigma^n_j(i) := \begin{cases} i, & \text{if } i \leqslant j, \\ i-1, & \text{if } i > j. \end{cases}$$

  This morphism is called *elementary degeneracies*.

When convenient, we may drop the superscript of these maps and simply refer to them as δ_i and σ_j.

Notation 19.1.2 (Morphisms in $\mathbf{\Delta}$)**.** When convenient, a morphism $f \colon [n] \to [m]$ in $\mathbf{\Delta}$ shall be explicitly denoted by

$$f := \langle f_0 f_1 \ldots f_n \rangle$$

where $f_j := f j \in [m]$ for each $0 \leqslant j \leqslant n$. Moreover these maps can also be refered to as *simplicial operators*.

Lemma 19.1.3 (Generating morphisms in $\mathbf{\Delta}$)**.** Every morphism of $\mathbf{\Delta}$ can be generated by a composition of elementary faces and degeneracies.

Proof. Let $f\colon [n] \to [m]$ be any morphism of $\mathbf{\Delta}$. We can factor f through an injection $\iota\colon [n] \rightarrowtail [k]$ and a surjection $s\colon [k] \twoheadrightarrow [m]$:

$$\begin{tikzcd}
{[n]} \arrow[rr, "f"] \arrow[dr, two heads, "s"'] & & {[m]} \\
& {[k]} \arrow[ur, tail, "\iota"'] &
\end{tikzcd}$$

where $k := |\operatorname{im} f|$. Since s and ι must be order-preserving maps, it follows that they can be written as a finite composition of elementary degeneracies and elementary faces. ♮

Corollary 19.1.4. Every monomorphism of $\mathbf{\Delta}$ is split (admits a retraction), while every epimorphism of $\mathbf{\Delta}$ is split (admits a section).

Proof. Given an elementary face $\delta_j\colon [n] \rightarrowtail [n+1]$ the elementary degeneracy map $\sigma_j\colon [n+1] \twoheadrightarrow [n]$ is a retract of δ_j, and δ_j is a section of σ_j. ♮

Corollary 19.1.5 (Cosimplicial identities)**.** Fix any $n \in \mathbf{N}$ and consider indices $0 \leqslant i, j \leqslant n$. The following identities correlate elementary faces and degeneracies:

(1) If $i < j$ then $\delta^n_j \delta^{n-1}_i = \delta^n_i \delta^{n-1}_{j-1}$.

(2) If $i < j$ then $\sigma^{n-1}_i \sigma^n_j = \sigma^{n-1}_{j-1} \sigma^n_i$.

(3) If $i < j$ then $\sigma^{n-1}_i \delta^n_j = \delta^{n-1}_{j-1} \sigma^{n-2}_i$.

(4) If $i = j-1$ or $i = j$, then $\sigma^n_i \delta^{n+1}_j = \mathrm{id}_n$.

(5) If $i > j$ then $\sigma^{n-1}_i \delta^n_j = \delta^{n-1}_j \sigma^{n-2}_{i-1}$.

These identities can be found in the following four commutative diagrams:

$$\begin{tikzcd}
{[n-2]} \arrow[r, "\delta^{n-1}_i"] \arrow[d, "\delta^{n-1}_{j-1}"'] & {[n-1]} \arrow[d, "\delta^n_j"] \\
{[n-1]} \arrow[r, "\delta^n_i"'] & {[n]}
\end{tikzcd}
\qquad
\begin{tikzcd}
{[n+1]} \arrow[r, "\sigma^n_j"] \arrow[d, "\sigma^n_i"'] & {[n]} \arrow[d, "\sigma^{n-1}_i"] \\
{[n]} \arrow[r, "\sigma^{n-1}_{j-1}"'] & {[n-1]}
\end{tikzcd}$$

$$\begin{tikzcd}
{[n-1]} \arrow[r, "\delta^n_j"] \arrow[d, "\sigma^{n-2}_i"'] & {[n]} \arrow[d, "\sigma^{n-1}_i"] \\
{[n-2]} \arrow[r, "\delta^{n-1}_{j-1}"'] & {[n-1]}
\end{tikzcd}
\qquad
\begin{tikzcd}
{[n-1]} \arrow[r, "\delta^n_j"] \arrow[d, "\sigma^{n-2}_{i-1}"'] & {[n]} \arrow[d, "\sigma^{n-1}_i"] \\
{[n-2]} \arrow[r, "\delta^{n-1}_j"'] & {[n-1]}
\end{tikzcd}$$

Definition 19.1.6 (Cosimplicial object)**.** Given a category C, we define a *cosimplicial object* in C to be a covariant functor

$$F\colon \mathbf{\Delta} \longrightarrow \mathsf{C}.$$

Any cosimplicial object F is completely determined by $F[n]$ for all $n \in \mathbf{N}$ and by the maps $F\delta_i$ and $F\sigma_j$. The collection of all cosimplicial objects of C will be denoted by $\mathsf{CoSimp}(\mathsf{C})$.

Limits & Colimits in $\mathbf{\Delta}$

Lemma 19.1.7 (Pushout of injections)**.** Let $\iota\colon [k] \hookrightarrow [n]$ and $\tau\colon [k] \hookrightarrow [m]$ be order-preserving inclusions where

$$\iota(j) := j \quad \text{and} \quad \tau(j) := j + (m-k),$$

that is, ι sends $[k]$ to the initial segment of $[n]$, while τ sends $[k]$ to the terminal segment of $[m]$. There exists a pushout

$$\begin{tikzcd}
[k] \arrow[r, hook, "\iota"] \arrow[d, hook, "\tau"'] \arrow[dr, phantom, "\ulcorner", very near end] & {[n]} \arrow[d] \\
{[m]} \arrow[r] & {[n] \cup_{[k]} [m]}
\end{tikzcd}$$

in the simplex category $\mathbf{\Delta}$.

Proof. Indeed, if we consider $[m+n-k]$ as a candidate for the pushout, notice that the following diagram commutes

$$\begin{tikzcd}
[k] \arrow[r, hook, "\iota"] \arrow[d, hook, "\tau"'] & {[n]} \arrow[d, "\tau"] \\
{[m]} \arrow[r, "\iota"'] & {[m+n-k]}
\end{tikzcd}$$

Now, consider any object $[\ell] \in \mathbf{\Delta}$ together with two morphisms $f\colon [n] \to [\ell]$ and $g\colon [m] \to [\ell]$ such that $f\iota = g\tau$. Define a map $\phi\colon [m+n-k] \to [\ell]$ as follows

$$\phi(j) := \begin{cases} g(j), & \text{if } j \leqslant m-k, \\ g(j) = f(j-(m-k)), & \text{if } m-k \leqslant j \leqslant m, \\ f(j-(m-k)), & \text{if } m \leqslant j \leqslant m+n-k. \end{cases}$$

It is easy to see that ϕ is an order-preserving map and is uniquely defined so that the following diagram commutes

$$\begin{tikzcd}
[k] \arrow[r, hook, "\iota"] \arrow[d, hook, "\tau"'] & {[n]} \arrow[d, "\tau"] \arrow[ddr, bend left, "f"] & \\
{[m]} \arrow[r, "\iota"'] \arrow[drr, bend right, "g"'] & {[m+n-k]} \arrow[dr, dashed, "\phi"] & \\
 & & {[\ell]}
\end{tikzcd}$$

Therefore $[m+n-k] = [n] \cup_{[k]} [m]$ since we are in a skeletal category and isomorphism classes contain a unique representative. ♮

These pushouts lead to an interesting construction, any object $[n] \in \mathbf{\Delta}$ is the colimit of a diagram consisting of $[0]$'s and $[1]$'s, since

$$\begin{tikzcd}
{[0]} \arrow[r, hook, "0"] \arrow[d, hook, "m"'] \arrow[dr, phantom, "\ulcorner", very near end] & {[n]} \arrow[d] \\
{[m]} \arrow[r] & {[m+n]}
\end{tikzcd}$$

Lemma 19.1.8 (Pushout of surjections)**.** The following properties concern the pushout of pairs of surjective morphisms in $\mathbf{\Delta}$:

(a) Considering the cosimplicial identity (2) (see Corollary 19.1.5), where $i < j$, there exists *sections* $\alpha\colon [n] \to [n+1]$ of σ_i^n, and $\beta\colon [n-1] \to [n]$ of σ_i^{n-1} such that the following diagram commutes

$$\begin{tikzcd}
{[n+1]} \arrow[r, "\sigma_j^n"] \arrow[d, bend left, "\sigma_i^n"] & {[n]} \arrow[d, bend right, "\sigma_i^{n-1}"'] \\
{[n]} \arrow[r, "\sigma_{j-1}^{n-1}"'] \arrow[u, bend left, "\alpha"] & {[n-1]} \arrow[u, bend right, "\beta"']
\end{tikzcd}$$

that is, $\sigma_j^n \alpha = \beta \sigma_{j-1}^{n-1}$ — these sections are said to be *compatible* with the square. Therefore, the square is an *absolute pushout*.

(b) Let $p\colon [n] \twoheadrightarrow [k]$ and $q\colon [n] \twoheadrightarrow [\ell]$ be *surjections* in $\mathbf{\Delta}$. Then the *pushout* of p and q exists and is *absolute*:

$$\begin{tikzcd}
{[n]} \arrow[r, two heads, "q"] \arrow[d, two heads, "p"'] \arrow[dr, phantom, "\ulcorner", very near end] & {[\ell]} \arrow[d] \\
{[k]} \arrow[r] & {[m]}
\end{tikzcd}$$

Proof. For the proof of item (a), proceed as follows. From the cosimplicial identity (3) we have $\delta_i^{n+1}\sigma_i^n = \sigma_i^{n+1}\delta_{i+1}^{n+2}$, moreover, from (4) we obtain $\sigma_i^{n+1}\delta_{i+1}^{n+2} = \mathrm{id}_{n+1}$ — thus δ_i^{n+1} is a section of σ_i^n. Analogously, we find that δ_i^n is a section of σ_i^{n-1}. For the compatibility condition, notice that since $i < j$, then from (5) we know that $\sigma_j^n \delta_i^{n+1} = \delta_i^n \sigma_{j-1}^{n-1}$. Therefore we may define $\alpha := \delta_i^{n+1}$ and $\beta := \delta_i^n$.

For item (b), since surjections are finite compositions of degeneracies, via item (a) and Proposition 2.5.27, we conclude that the square in (b) is an absolute pushout. ♮

Lemma 19.1.9. The following are properties of pullbacks along monomorphisms and pushouts along epimorphisms in the simplex category $\mathbf{\Delta}$:

(a) If $f\colon [m] \rightarrowtail [n]$ is a *monomorphism*, then if $g\colon [k] \to [n]$ is any morphism such that $\operatorname{im} g \cap \operatorname{im} f \neq \varnothing$, then the square

$$\begin{tikzcd} {[a]} \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & {[k]} \arrow[d, "g"] \\ {[m]} \arrow[r, tail, "f"'] & {[n]} \end{tikzcd}$$

is a *pullback* in $\mathbf{\Delta}$.

(b) If $f\colon [m] \twoheadrightarrow [n]$ is an *epimorphism*, then for any morphism $g\colon [m] \to [k]$, then square

$$\begin{tikzcd} {[m]} \arrow[r, two heads, "f"] \arrow[d, "g"'] & {[n]} \arrow[d] \\ {[k]} \arrow[r] \arrow[ur, phantom, "\ulcorner", very near end] & {[b]} \end{tikzcd}$$

is a *pushout* in $\mathbf{\Delta}$.

I did not understand how to prove this, the proof presented on the book is a bit mysterious

19.2 Simplicial Sets

Construction

Definition 19.2.1 (Simplicial object)**.** For any category C, a *simplicial object* in C is a contravariant functor $\mathbf{\Delta}^{\mathrm{op}} \to \mathsf{C}$. We define $\mathtt{Simp}(\mathsf{C})$ to be the category whose objects are simplicial objects and natural transformations between them—such category is commonly referred to as the *simplicial category of* C.

Notation 19.2.2 (Unwrapping the definition)**.** Less compactly, a simplicial object $X \in \mathtt{Simp}(\mathsf{C})$ consists of a collection of objects $(X_n)_{n\in\mathbf{N}}$, for $X_n := X[n] \in \mathsf{C}$, and arrows $\alpha^*\colon X_n \to X_m$ in C for each map $\alpha\colon [m] \to [n]$ in $\mathbf{\Delta}$. The nature of these arrows are as discussed in Remark 1.7.3.

Notation 19.2.3 (Simplicial sets, notations and nomenclatures)**.** In particular, the most important case we'll study for the time being is the *simplicial category of sets*, which we denote by $\mathtt{sSet}$. The objects of $\mathtt{sSet}$ shall be called *simplicial sets*.

Given a simplicial set $X \in \mathtt{sSet}$, the points of $X_n \in \mathtt{Set}$ will be referred to as the *n-cells* (or *n-simplices*) of $X \in \mathtt{sSet}$.

Since every map α in $\mathbf{\Delta}$ may be decomposed into elementary face and degeneracies, the the collection δ_i^* and σ_j^* also generate the morphisms α^* between the objects of

$(X_n)_{n\in\mathbf{N}}$ of C. In the context of the simplicial category, we denote them by

$$d^n_i := (\delta^n_i)^*\colon X_n \longrightarrow X_{n-1},$$
$$s^n_j := (\sigma^n_j)^*\colon X_n \longrightarrow X_{n+1},$$

for all $0 \leqslant i, j \leqslant n$. The arrows d_i are called *face maps* (or *cofaces*), while the arrows s_j are called *degeneracy maps* (or *codegeneracies*) of the simplicial object X.

A map $\eta\colon X \Rightarrow Y$ is a natural transformation between simplicial objects if and only if it is compatible with the face and degeneracy maps, that is, the following two diagrams commute in C for all $n \in \mathbf{N}$, $0 \leqslant i, j \leqslant n$:

$$\begin{tikzcd} X_n \arrow[r, "\eta_n"] \arrow[d, "d^n_i"'] & Y_n \arrow[d, "d^n_i"] \\ X_{n-1} \arrow[r, "\eta_{n-1}"'] & Y_{n-1} \end{tikzcd} \qquad \begin{tikzcd} X_n \arrow[r, "\eta_n"] \arrow[d, "s^n_j"'] & Y_n \arrow[d, "s^n_j"] \\ X_{n+1} \arrow[r, "\eta_{n+1}"'] & Y_{n+1} \end{tikzcd}$$

Corollary 19.2.4 (Simplicial identities)**.** Since the face and degeneracy maps d_i and s_j are *dual* to the elementary face and degeneracy maps δ_i and σ_j, they satisfy the following identities—which are dual to Corollary 19.1.5:

(1) $d_i d_j = d_{j-1} d_i$, for $i < j$.

(2) $s_j s_i = s_i s_{j-1}$, for $i < j$.

(3) $d_j s_i = s_i d_{j-1}$, for $i < j - 1$.

(4) $d_j s_i = \mathrm{id}$, if $i = j$ or $i = j - 1$.

(5) $d_j s_i = s_{i-1} d_j$, for $i > j$.

Subcomplexes

Definition 19.2.5 (Subcomplex)**.** By a *subcomplex* of a simplicial set we mean a subobject.

Lemma 19.2.6. Given a morphism of simplicial sets $f\colon X \to Y$, the image fX is a subcomplex of Y.

Discrete Simplicial Sets

Definition 19.2.7 (Discrete simplicial set)**.** We say that a simplicial set X is *discrete* if every simplicial operator $f\colon [m] \to [n]$ induces a *bijective set-function* $f^*\colon X_n \to X_m$. We denote by $\mathtt{sSet}^{\mathrm{disc}}$ the full subcategory of $\mathtt{sSet}$ consisting of discrete simplicial sets.

For every set S, there exists a discrete simplicial set S^{disc} where $S^{\mathrm{disc}}_n := S$ for each $n \in \mathbf{N}$, and for any $f\colon [n] \to [m]$ we have $S^{\mathrm{disc}} f := \mathrm{id}_S$.

Corollary 19.2.8. For any set S and simplicial set X, there exists a bijection

$$\mathrm{Mor}_{\mathtt{sSet}}(S^{\mathrm{disc}}, X) \simeq \mathrm{Mor}_{\mathtt{Set}}(S, X_0).$$

Proof. Define a set-function $\Phi\colon \mathrm{Mor}_{\mathtt{sSet}}(S^{\mathrm{disc}}, X) \to \mathrm{Mor}_{\mathtt{Set}}(S, X_0)$ by mapping each simplicial set morphism $f\colon S^{\mathrm{disc}} \to X$ to the corresponding set function $f_0\colon S \to X_0$. To prove that Φ is injective, consider simplicial morphisms $f, g\colon S^{\mathrm{disc}} \rightrightarrows X$ such that $f_0 = g_0$. Applying the naturallity of f and g to the simplicial operator $j\colon [n] \to [0]$ we obtain that $j^* f_0 = f_n$ and $j^* g_0 = g_n$, but since $f_0 = f_0$, then $f_n = g_n$—proving that $f = g$. Surjectivity is immediate, therefore Φ is a bijection. ♮

Proposition 19.2.9 ($\mathtt{sSet}^{\mathrm{disc}}$ & $\mathtt{Set}$)**.** The full subcategory of discrete simplicial sets is *equivalent* to the category of sets.

Proof. We start by constructing a functor $\mathrm{disc}\colon \mathtt{Set} \to \mathtt{sSet}^{\mathrm{disc}}$ where we define the simplicial set S^{disc} associated with a given set S to be the simplicial set with $S_n^{\mathrm{disc}} = S$ for every S. Each set-function $f\colon A \to B$ is mapped to their corresponding simplicial set morphism $f\colon A^{\mathrm{disc}} \to B^{\mathrm{disc}}$ given by $f_n := f\colon A \to B$ for each $n \in \mathbf{N}$. We now show that disc is an equivalence of categories:

- (Fully faithful) Given any two sets, S and T, define a set-function $\Phi\colon \mathrm{Mor}_{\mathtt{Set}}(S, T) \to \mathrm{Mor}_{\mathtt{sSet}^{\mathrm{disc}}}(S^{\mathrm{disc}}, T^{\mathrm{disc}})$ by mapping each set function $f\colon S \to T$ to $f\colon S^{\mathrm{disc}} \to T^{\mathrm{disc}}$ as above. Then it's clear that Φ is a *bijection*.
- (Essentially surjective) Given any $X \in \mathtt{sSet}^{\mathrm{disc}}$, consider the set $S := X_0$—we'll show that $X \simeq S^{\mathrm{disc}}$. Define a morphism of simplicial sets $f\colon S^{\mathrm{disc}} \to X$ where $f_0 := \mathrm{id}_{X_0}$. Let $\alpha\colon [n] \to [0]$ be any simplicial operator. Since $\alpha^*\colon X_0 \xrightarrow{\simeq} X_n$ is a bijection, then by naturality of f we find that

$$\begin{tikzcd} & X_0 \arrow[dd, "\alpha^*", "\simeq"'] \\ S \arrow[ur, "\mathrm{id}_{X_0}"] \arrow[dr, "f_n"'] & \\ & X_n \end{tikzcd}$$

  commutes in $\mathtt{Set}$, making f_n a bijection, which proves that f is a natural isomorphism of simplicial sets.

♮

Standard Simplices

Definition 19.2.10 (Standard simplex)**.** We define the *standard n-simplex* Δ^n to be the contravariant functor $\mathbf{\Delta}^{\mathrm{op}} \to \mathtt{Set}$ represented by $[n]$, that is, the *simplicial set*

$$\Delta^n := \mathrm{Mor}_{\mathbf{\Delta}}(-, [n]).$$

The *generator* of the standard n-simplex corresponds to the n-cell

$$\iota_n := \mathrm{id}_{[n]} \in \Delta^n[n].$$

Let $\phi\colon [\ell] \to [k]$ be any simplicial operator, then the action of ϕ on the cells of Δ^n is given by the map $\Delta^n\phi = \phi^*\colon \Delta^k \to \Delta^\ell$. Explicitly, given a k-cell $\alpha\colon [k] \to [n]$ of Δ^n, the action of the simplicial operator ϕ on α yields an ℓ-cell

$$\big(\phi^*\alpha = \alpha\phi\colon [\ell] \longrightarrow [n]\big) \in \Delta^n[\ell].$$

Lemma 19.2.11 (Yoneda consequences)**.** Given a simplicial set X, there exists a bijection

$$\mathrm{Mor}_{\mathtt{sSet}}(\Delta^n, X) \simeq X_n,$$

which is explicitly given by the map $\eta \mapsto \eta_n\iota_n$. Equivalently, for each n-cell $x \in X_n$, there *exists* a *unique* morphism of simplicial sets $\eta^x\colon \Delta^n \to X$ such that $\eta^x_n\iota_n = x$.

Proof. Consider the Yoneda functor $よ_{\mathbf{\Delta}}\colon \mathbf{\Delta} \to \mathtt{sSet}$, mapping each object $[n] \mapsto \Delta^n$ and each simplicial operator $f\colon [n] \to [m]$ to the morphism of simplicial sets $f_*\colon \Delta^n \to \Delta^m$. Explicitly, for each $k \in \mathbf{N}$ we have $(f_*)_k\colon \Delta^n[k] \to \Delta^m[k]$ mapping $\alpha \mapsto f\alpha$. By the Yoneda lemma we know that there exists, for each $n \in \mathbf{N}$, a bijection

$$\mathrm{Mor}_{\mathtt{sSet}}(\Delta^n, X) = \mathrm{Mor}_{\mathtt{sSet}}(よ_{\mathbf{\Delta}}[n], X) \simeq X_n.$$

♮

Definition 19.2.12 (Representing map of a cell)**.** Regarding Lemma 19.2.11, we shall refer to η^x as the *representing map* of the n-cell x. When convenient, we'll simply denote η^x by $x\colon \Delta^n \to X$.

Corollary 19.2.13. There exists a bijection

$$\mathrm{Mor}_{\mathtt{sSet}}(\Delta^n, \Delta^m) \simeq \Delta^m[n].$$

The *inverse* of this set-function is explicitly given by mapping each simplicial operator $\alpha\colon [n] \to [m]$ to the corresponding morphism of simplicial sets $\alpha_*\colon \Delta^n \to \Delta^m$—where α_* acts on any k-cell $x \in \Delta^n[k]$ as $\alpha_* x = \alpha x \in \Delta^m[k]$.

Proof. In the proof of Lemma 19.2.11, merely consider the case where $X := \Delta^m$. ♮

Corollary 19.2.14 (Standard simplices & $\mathbf{\Delta}$)**.** The full subcategory $\mathtt{sSet}_\Delta$ of $\mathtt{sSet}$ consisting of standard simplices is *equivalent* to the simplex category $\mathbf{\Delta}$.

Proof. Consider the functor $F\colon \mathbf{\Delta} \to \mathtt{sSet}_\Delta$ given by $[n] \mapsto \Delta^n$ and mapping each simplicial operator $f\colon [n] \to [m]$ to its corresponding morphism of simplicial sets $f_*\colon \Delta^n \to \Delta^m$. By Lemma 19.2.11 we see that F is fully faithful and essentially surjective—thus an equivalence of categories. ♮

Definition 19.2.15 (Standard simplices on totally ordered sets)**.** Given a totally ordered set S, we define the *standard S-simplicial set* Δ^S to be the functor[1]

$$\Delta^S := \mathrm{Mor}_{\mathtt{tOrd}}(-, S)|_{\mathbf{\Delta}},$$

that is, for each $n \in \mathbf{N}$ we define Δ^S_n to be the collection of order preserving maps $[n] \to S$.

[1]We denote by $\mathtt{tOrd}$ the category of totally ordered sets.

Definition 19.2.16 (Boundary of a standard simplex)**.** For every $n \in \mathbf{N}$, we define the *boundary* of the standard n-simplex Δ^n to be the subobject

$$\partial \Delta^n \coloneqq \operatorname{colim}_{k \in [n]} \Delta^{[n] \smallsetminus k} = \bigcup_{k \in [n]} \Delta^{[n] \smallsetminus k}$$

composed of all *codimension-one faces* of Δ^n. Explicitly, for each $k \in \mathbf{N}$ we have

$$\partial \Delta^n[k] = \{ f \in \operatorname{Mor}_{\boldsymbol{\Delta}}([k],[n]) : f[k] \neq [n] \} \subseteq \Delta^n[k],$$

composed of all *non-surjective k-cells* of Δ^n.

Proposition 19.2.17. The boundary $\partial \Delta^n$ is the *maximal proper subcomplex of* Δ^n.

Proof. Let X be a proper subcomplex of Δ^n.

Prove

$\natural$

Geometric Realization of the Simplex Category

We define, for each $n \in \mathbf{N}$, a corresponding *standard topological n-simplex* Δ^n_{top} given by

$$\Delta^n_{\text{top}} \coloneqq \left\{ (t_0, \dots, t_n) \in \mathbf{R}^{n+1} : \sum_{j=0}^{n} t_j = 1 \text{ and } t_j \geqslant 0 \text{ for all } j \right\}.$$

Each Δ^n_{top} is composed of $n+1$ vertices $v_j \coloneqq (\delta_{ij})_{i=0}^n$.

From a categorical point of view, standard topological simplices are nothing more than a functor

$$\Delta^\bullet_{\text{top}} \colon \boldsymbol{\Delta} \longrightarrow \mathtt{Top},$$

mapping objects $\Delta^\bullet_{\text{top}}[n] \coloneqq \Delta^n_{\text{top}}$—where Δ^n_{top} is endowed with the standard euclidean topology—and for each morphism $f \colon [n] \to [m]$ in $\boldsymbol{\Delta}$, we map $\Delta^\bullet_{\text{top}} f \coloneqq f_*$, where $f_* \colon \Delta^n_{\text{top}} \to \Delta^m_{\text{top}}$ is a uniquely determined continuous map such that $f_*(v_j) \coloneqq v_{f(j)}$. From this definition we obtain

$$f_*(t_0, \dots, t_n) = (s_0, \dots, s_m), \text{ where } s_j = \sum_{f(i)=j} t_i,$$

that is, s_j is the sum of the points that are collapsed to the j-th coordinate.

This functor gives us a geometric visualization of the action of the elementary face and degeneracies:

- Given an elementary *face* map $\delta_j \colon [n-1] \rightarrowtail [n]$, for any $0 \leqslant j \leqslant n$, the corresponding map $(\delta_j)_* \colon \Delta^{n-1}_{\text{top}} \hookrightarrow \Delta^n_{\text{top}}$ is given by

$$(\delta_j)_* v_i = \begin{cases} v_i, & \text{if } i < j, \\ v_{i+1}, & \text{if } i \geqslant j. \end{cases}$$

  That is, $(\delta_j)_*$ embeds $\Delta^{n-1}_{\text{top}}$ as a face of Δ^n_{top} opposite to the j-th vertex.

- An elementary *degeneracy* map $\sigma_j\colon [n+1] \twoheadrightarrow [n]$, for any $0 \leqslant j \leqslant n$, has a map $(\sigma_j)_*\colon \Delta^{n+1}_{\text{top}} \twoheadrightarrow \Delta^n_{\text{top}}$ mapping the vertices as follows

$$(\sigma_j)_* v_i = \begin{cases} v_i, & \text{if } i \leqslant j, \\ v_{i-1}, & \text{if } i > j. \end{cases}$$

  Geometrically, the degeneracy map makes $\Delta^{n+1}_{\text{top}}$ into Δ^n_{top} by removing a face of dimension 1—through the projection parallel to the line connecting v_j and v_{j+1}.

Geometric Realization of a Simplicial Set

Given a simplicial set $X\colon \mathbf{\Delta}^{\text{op}} \longrightarrow \mathtt{Set}$, we consider the topological space

$$\coprod_{n\in\mathbf{N}} X_n \times \Delta^n_{\text{top}}$$

and construct in this space a minimal equivalence relation $\sim_{\text{gr}}$ for which points $(x,t) \in X_n \times \Delta^n_{\text{top}}$ and $(x',t') \in X_m \times \Delta^m_{\text{top}}$ are *equivalent*—$(x,t) \sim_{\text{gr}} (x',t')$—if and only if there exists a morphism $\alpha\colon [m] \longrightarrow [n]$ in $\mathbf{\Delta}$ such that $x' = \alpha^* x$ and $t = \alpha_* t'$. In an equivalent manner, we may summarize this equivalence relation as gluing points of the form

$$(x, \alpha_* t) \sim_{\text{gr}} (\alpha^* x, t).$$

The points of the resulting quotient space $(\coprod_{n\in\mathbf{N}} X_n \times \Delta^n_{\text{top}})/{\sim_{\text{gr}}}$ are denoted by $x \otimes t$—corresponding to the class of a pair (x,t).

Study why this construction is related to a tensor product of the form $X \otimes_{\mathbf{\Delta}} \Delta^{\bullet}_{\text{top}}$.

Definition 19.2.18 (Geometric realization functor)**.** We define the *geometric realization* of the category of simplicial sets to be a functor

$$|-|\colon \mathtt{sSet} \longrightarrow \mathtt{Top},$$

mapping $|X| \coloneqq (\coprod_{n\in\mathbf{N}} X_n \times \Delta^n_{\text{top}})/{\sim_{\text{gr}}}$ and for each natural transformation $\eta\colon X \longrightarrow Y$ we have a topological morphism $|\eta|\colon |X| \longrightarrow |Y|$ given by $x \otimes t \mapsto \eta_n x \otimes t$, for $(x,t) \in X_n \times \Delta^n_{\text{top}}$.

For any simplicial set X, each n-cell $x \in X_n$ induces *topological* morphism

$$\widehat{x}\colon \Delta^n_{\text{top}} \longrightarrow |X|, \text{ mapping } t \longmapsto x \otimes t.$$

From construction, given a morphism $\alpha\colon [n] \longrightarrow [m]$ in $\mathbf{\Delta}$ and a point $y \in X_m$ such that $y = \alpha^* x$, the diagram

$$\begin{tikzcd} \Delta^m_{\text{top}} \arrow[rr, "\alpha_*"] \arrow[dr, "\widehat{y}"'] & & \Delta^n_{\text{top}} \arrow[dl, "\widehat{x}"] \\ & {|X|} & \end{tikzcd} \tag{19.1}$$

commutes in $\mathtt{Top}$.

This looks like there exists an induced slice over category $\mathtt{C}/|X|$—where $\mathtt{C}$ is a subcategory of $\mathtt{Top}$ consisting of standard topological simplices and morphisms between them—whose objects are $\widehat{x}$ and morphisms $\phi\colon \widehat{x} \to \widehat{y}$ are maps $\phi_*\colon \Delta^m_{\text{top}} \to \Delta^n_{\text{top}}$ for some $\phi\colon [n] \to [m]$ in $\mathbf{\Delta}$.

Given any category $\mathtt{C}$ and a functor $F\colon \mathbf{\Delta} \to \mathtt{C}$, we can define a *simplicial set* induced by any object $C \in \mathtt{C}$ given by

$$\mathrm{Mor}_{\mathtt{C}}(F(-), C)\colon \mathbf{\Delta}^{\mathrm{op}} \longrightarrow \mathtt{Set}.$$

This simplicial set maps each $[n] \in \mathbf{\Delta}$ to the set of morphisms $\mathrm{Mor}_{\mathtt{C}}(Fn, C)$, and each morphism $\alpha\colon [m] \to [n]$ of $\mathbf{\Delta}$ to the set-function $\alpha^*\colon \mathrm{Mor}_{\mathtt{C}}(Fn, C) \to \mathrm{Mor}_{\mathtt{C}}(Fm, C)$.

Definition 19.2.19 (Singular complex)**.** Let $\mathtt{C}$ be a category and $F\colon \mathbf{\Delta} \to \mathtt{C}$ be a functor. We define the *singular complex* of $\mathtt{C}$ to be the functor

$$\mathrm{Sing}_F\colon \mathtt{C} \longrightarrow \mathtt{sSet}$$

mapping each object $C \in \mathtt{C}$ to the simplicial set

$$\mathrm{Sing}_F(C) \coloneqq \mathrm{Mor}_{\mathtt{C}}(F(-), C)\colon \mathbf{\Delta}^{\mathrm{op}} \longrightarrow \mathtt{Set},$$

and each map $\alpha\colon [m] \to [n]$ to the set-function

$$\mathrm{Sing}_F(\alpha) \coloneqq \alpha^*\colon \mathrm{Mor}_{\mathtt{C}}(Fn, C) \longrightarrow \mathrm{Mor}_{\mathtt{C}}(Fm, C).$$

In particular, the standard topological simplices functor $\Delta^\bullet_{\text{top}}\colon \mathbf{\Delta} \to \mathtt{Top}$ induces a singular complex on each topological space. Since we shall be mostly interested in this particular case for the time being, we shall reserve the notation

$$\mathrm{Sing} \coloneqq \mathrm{Sing}_{\Delta^\bullet_{\text{top}}}\colon \mathtt{Top} \longrightarrow \mathtt{sSet},$$

with no subscripts, for the standard topological simplices functor. In this case, given a topological space T, we shall denote by $\mathrm{Sing}(T)_n$ the image of $[n] \in \mathbf{\Delta}$ under the simplicial set $\mathrm{Sing}(T)$.

Given a simplicial set X, the collection of simplexes $(\widehat{x}_n)_{n\in\mathbf{N}}$, where $x_n \in X_n$, covers the whole topological space $|X|$—in the sense that collection of images forms a cover of $|X|$. Therefore, given any topological morphism $\phi\colon |X| \to T$, this map is completely defined by the family of compositions $(\phi\widehat{x}_n\colon \Delta^n_{\text{top}} \to T)_{n\in\mathbf{N}}$. Given a morphism $\alpha\colon [n] \to [n]$ in $\mathbf{\Delta}$, by Eq. (19.1), the diagram

$$\begin{tikzcd}
\Delta^m_{\text{top}} \arrow[rr, "\alpha_*"] \arrow[dr, "\widehat{y}_m"'] \arrow[ddr, bend right, "\phi\widehat{y}_m"'] & & \Delta^n_{\text{top}} \arrow[dl, "\widehat{x}_n"] \arrow[ddl, bend left, "\phi\widehat{x}_n"] \\
& {|X|} \arrow[d, "\phi"] & \\
& T &
\end{tikzcd}$$

commutes in $\mathtt{Top}$. This construction induces unique a collection of maps

$$(\phi_n \colon X_n \longrightarrow \mathrm{Sing}(T)_n)_{n\in\mathbf{N}},$$

where $\phi_n(x) \mapsto \phi\widehat{x}$. Notice that this family of arrows is nothing more than a natural transformation $\phi\colon X \Rightarrow \mathrm{Sing}(T)$ between simplicial sets. From this we conclude that there exists a natural bijection

$$\mathrm{Mor}_{\mathtt{Top}}(|X|, T) \simeq \mathrm{Mor}_{\mathtt{sSet}}(X, \mathrm{Sing}(T)),$$

thus the singular complex functor is *right adjoint* to the geometric realization,

$$\begin{tikzcd} \mathtt{sSet} \arrow[r, shift left, "|-|"] & \mathtt{Top} \arrow[l, shift left, "\mathrm{Sing}"] \end{tikzcd}$$

Geometric Realization as a CW-complex

Definition 19.2.20 (Degenerate n-cell)**.** Given a simplicial set X, we say that an n-cell $x \in X_n$ is *degenerate* if $x \in s_j X_{n-1}$ for some *codegeneracy* map $s_j \colon X_{n-1} \to X_n$, where $0 \leqslant j \leqslant n-1$.

Equivalently, x is denenerate if there exists a *surjective* map $\alpha\colon [n] \twoheadrightarrow [m]$ and m-cell $y \in X_m$ for which $x = \alpha^* y$.

Lemma 19.2.21 (Eilenberg-Zilber)**.** Let x be a n-cell of a given simplicial set X. There *exists a unique* pair (α, y) such that $\alpha\colon [n] \twoheadrightarrow [k]$ is a *surjective* map and y is a *non-degenerate* k-cell of X satisfying $\alpha^* y = x$.

Proof. The existence of the pair (α, y) comes straight from definition. Now suppose (β, z) is another pair satisfying the said property, where z is a non-degenerate ℓ-cell of X. Since pushouts of a pair of surjections in $\mathbf{\Delta}$ are absolute (see Lemma 19.1.8), the pushout of the pair (α, β):

$$\begin{tikzcd} {[n]} \arrow[r, two heads, "\alpha"] \arrow[d, two heads, "\beta"'] & {[k]} \arrow[d, "\gamma"] \\ {[\ell]} \arrow[r, "\omega"'] \arrow[ur, phantom, "\ulcorner"] & {[s]} \end{tikzcd}$$

is turned into a pullback by the simplicial set $X\colon \mathbf{\Delta}^{\mathrm{op}} \to \mathtt{Set}$, that is:

$$\begin{tikzcd} X_n & X_k \arrow[l, "\alpha^*"'] \\ X_\ell \arrow[u, "\beta^*"] \arrow[ur, phantom, "\lrcorner"] & X_s \arrow[l, "\omega^*"] \arrow[u, "\gamma^*"'] \end{tikzcd}$$

From the pushout property, γ and ω are epimorphisms, thus split, X ensures that γ^* and ω^* are split-epimorphisms in $\mathtt{Set}$. Therefore there exists s-cells $a, b \in X_s$ such that $\gamma^* a = y$ and $\omega^* b = z$. Notice however that we assumed y and z to be both non-degenerate, hence it must be the case that γ and ω are *identities*. This implies in $\beta = \alpha$ and $y = z$. $\natural$

Notice that, given a simplicial set X, we can naturally construct a *filtration* for X (see Definition 17.2.2) by defining for each $n \in \mathbf{N}$ the subspace $|X|_n$ of $|X|$ given by

$$|X|_n := \{x \otimes t \in |X| : (x, t) \in X_k \times \Delta^k_{\text{top}} \text{ for some } k \leqslant n\},$$

so that the collection $(|X|_n)_{n\in\mathbf{N}}$ defines a filtration—that is, $|X|_n \leqslant |X|_{n+1}$ and $|X| = \bigcup_{n\in\mathbf{N}} |X|_n$ is endowed with the *weak topology* (see Definition 17.1.21).

Lemma 19.2.22. Given a simplicial set X, the subspace $|X|_0 \subseteq |X|$ is *discrete* and given by

$$|X|_0 = X_0 \times \Delta^0_{\text{top}}.$$

Proof. We shall create an isomorphism between both spaces. Define a surjective map $X_0 \times \Delta^0_{\text{top}} \twoheadrightarrow |X|_0$ by mapping $(x, 1) \mapsto x \otimes 1$. Now, suppose that for some $x, y \in X_0$ one has $x \otimes 1 = y \otimes 1$, then there must exist a pair $(z, t) \in X_n \times \Delta^n_{\text{top}}$, for some $n \in \mathbf{N}$, and parallel morphisms $\alpha, \beta \colon [0] \rightrightarrows [n]$ in $\mathbf{\Delta}$ for which

$$(x, \alpha_* 1) = (\alpha^* z, t) \quad \text{and} \quad (y, \beta_* 1) = (\beta^* z, t).$$

Since $\Delta^0_{\text{top}} = 1$ is a single point, then $\alpha_* = \beta_*$ and hence $\alpha = \beta$, which shows that $x = y$. ♮

Notation 19.2.23. Denote by X^{nd}_n the collection of all *non-degenerate n-cells* of a given simplicial set X.

Lemma 19.2.24. Let $\xi \in |X|$ be any point, and let $n \in \mathbf{N}$ be the *minimal* index such that there exists $(x, t) \in X_n \times \Delta^n_{\text{top}}$ for which $\xi = x \otimes t$. Then x is a non-degenerate simplex, and if $n \geqslant 1$ then $t \in \operatorname{Int} \Delta^n_{\text{top}}$—also, if that is the case, then the pair (x, t) is unique with such property.

Proof. We shall prove this lemma in three steps:

- (x is non-degenerate) Suppose for the sake of contradiction that x is degenerate, so that there exists a pair $(\alpha \colon [n] \twoheadrightarrow [m], y)$ with $y \in X_m$ and $m < n$ such that $\alpha^* y = x$. From this one obtains

  $$x \otimes t = \alpha^* y \otimes t = y \otimes \alpha_* t$$

  then $(y, \alpha_* t) \in X_m \times \Delta^m_{\text{top}}$ satisfies $\xi = y \otimes \alpha^* t$, which contradicts the assumption that n was the minimal element of $\mathbf{N}$ with this property.

- ($t \in \operatorname{Int} \Delta^n_{\text{top}}$) Assume that $n \geqslant 1$ and, for the sake of contradiction, suppose that $t \in \partial \Delta^n_{\text{top}}$—so that there exists at least one coordinate of t that is zero. From this last comment, it follows that there exists an injective map $\beta \colon [k] \rightarrowtail [n]$, for some $k < n$, such that $t \in \beta_* \Delta^k_{\text{top}}$. If $s \in \Delta^k_{\text{top}}$ is the point such that $\beta_* s = t$, then

  $$\xi = x \otimes t = x \otimes \beta_* s = \beta^* x \otimes s,$$

  but $\beta^* x \in \Delta^k_{\text{top}}$, which contradicts again the minimality of n.

- (Uniqueness of the pair (x,t)) Assume again that $n \geqslant 1$, and suppose that $\xi = x \otimes t = y \otimes s$ for some $x, y \in X_n^{\mathrm{nd}}$ and $t, s \in \operatorname{Int}\Delta^n_{\mathrm{top}}$—which can be assumed using the result of the last item. From the definition of $\otimes$ we may consider a finite zig-zag $(\alpha_j, \beta_j)_{j=1}^N$ in $\mathbf{\Delta}$:

$$\begin{tikzcd} {[n_{j-1}]} & & {[m_j]} \\ & {[k_j]} \arrow[ul, "\alpha_j"] \arrow[ur, "\beta_j"'] & \end{tikzcd}$$

where n_0 and m_N are both defined to be n and the intersection $\operatorname{im}\beta_j \cap \operatorname{im}\alpha_{j+1}$ is *non-empty* for every $1 \leqslant j < N$—together with a collection of pairs

$$\big((x_j, t_j), (y_j, s_j)\big)_{j=0}^N \in \prod_{j=0}^N (X_{k_j} \times \Delta^{k_j}_{\mathrm{top}}) \times (X_{m_j} \times \Delta^{m_j}_{\mathrm{top}})$$

for which one has the following relations:

$$(x_j, (\alpha_j)_* t_j) = (\alpha_j^* y_{j-1}, s_{j-1}) \quad \text{and} \quad (x_j, (\beta_j)_* t_j) = (\beta_j^* y_j, s_j),$$

and $(x_0, t_0) \coloneqq (x, t)$, while $(y_N, s_N) \coloneqq (y, s)$.

By hypothesis, since t and s are interior points of Δ^n_{top}, then it must be the case that both α_1 and β_N are *surjective*, so that $k_1, k_N \geqslant n$. We proceed by induction on the zig-zag length N. If $N = 1$, then the zig-zag is composed by two *surjective* maps $\alpha, \beta \colon [k] \rightrightarrows [n]$. Since pushouts of surjective maps exist in $\mathbf{\Delta}$ and are *absolute*, if we consider the pushout of the pair (α, β):

$$\begin{tikzcd} {[k]} \arrow[r, "\alpha"] \arrow[d, "\beta"'] \arrow[dr, phantom, "\ulcorner" very near end] & {[n]} \arrow[d, "\gamma"] \\ {[n]} \arrow[r] & {[\ell]} \end{tikzcd}$$

the contravariant functor X maps this universal square to a *pullback* in `Set`:

$$\begin{tikzcd} X_k \arrow[dr, phantom, "\lrcorner" very near start] & X_n \arrow[l, "\alpha^*"'] \\ X_n \arrow[u, "\beta^*"] & X_\ell \arrow[l] \arrow[u, "\gamma^*"'] \end{tikzcd}$$

Then since γ is a split epimorphism, so is γ^*, which proves the existence of an ℓ-cell $z \in X_\ell$ such that $\gamma^* z = x$. However, since x is non-degenerate and n is minimal, this can only happen if $\ell = n$ so that $\alpha, \beta = \mathrm{id}_{[n]}$—implying on $(x,t) = (y,s)$.

Assume, for the hypothesis of induction, that equality of the points is stablished for some $N - 1 > 1$. We now work on the case $N > 1$: our goal will be to shorten the zig-zag sequence by one. Start by factoring the morphism β_1 as follows

$$\begin{tikzcd}
{[k_1]} \arrow[rr, "\beta_1"] \arrow[dr, two heads, "\varepsilon"'] & & {[m_1]} \\
& {[m'_1]} \arrow[ur, tail, "\delta"'] &
\end{tikzcd}$$

If we now consider the pushout of the pair of *surjective* maps (α_1, ε), the simplicial set takes this pushout to a *pullback* in `Set` and from an analogous argument as done above, we conclude that we must have $\varepsilon = \mathrm{id}_{[k_1]}$. Therefore the factorization of β_1 reduces solely to $\beta_1 = \delta$, which is a *monomorphism*. Since $\operatorname{im}\alpha_2 \cap \operatorname{im}\beta_1 \neq \varnothing$, then by Lemma 19.1.9 we know that the pullback of the pair (β_1, α_2) exists in $\mathbf{\Delta}$:

$$\begin{tikzcd}
{[k'_1]} \arrow[r, tail, "\theta"] \arrow[d, "\eta"'] \arrow[dr, phantom, "\lrcorner", very near start] & {[k_2]} \arrow[d, "\alpha_2"] \\
{[k_1]} \arrow[r, tail, "\beta_1"'] & {[m_1]}
\end{tikzcd} \tag{19.2}$$

Considering the simplex covariant functor $\Delta^{\bullet}_{\mathrm{top}}$, we have a corresponding square

$$\begin{tikzcd}
\Delta^{k'_1}_{\mathrm{top}} \arrow[r, tail, "\theta_*"] \arrow[d, "\eta_*"'] \arrow[dr, phantom, "\lrcorner", very near start] & \Delta^{k_2}_{\mathrm{top}} \arrow[d, "(\alpha_2)_*"] \\
\Delta^{k_1}_{\mathrm{top}} \arrow[r, tail, "(\beta_1)_*"'] & \Delta^{m_1}_{\mathrm{top}}
\end{tikzcd}$$

in `Top`, which is again a pullback.

$\boxed{\text{Prove that } \Delta^{\bullet}_{\mathrm{top}} \text{ does indeed preserve the pullback}}$

Therefore can find $c \in X_{k'_1}$ such that $c = \eta^* x_1 = \theta^* y_1$, and take the *unique* $w \in \Delta^{k'_1}_{\mathrm{top}}$ such that $\eta_* w = t_1$ and $\theta_* w = s_1$.

Notice that by the pullback Eq. (19.2) one has

$$\begin{tikzcd}
{[n]} & & {[m_1]} & & {[m_2]} \\
& {[k_1]} \arrow[ul, "\alpha_1"] \arrow[ur, tail, "\beta_1"] & & {[k_2]} \arrow[ul, "\alpha_2"'] \arrow[ur, "\beta_2"'] & \\
& & {[k'_1]} \arrow[ul, "\eta"] \arrow[ur, tail, "\theta"'] & &
\end{tikzcd}$$

so together with the pair $(c, w) \in X_{k'_1} \times \Delta^{k'_1}_{\mathrm{top}}$ and $[n] \xleftarrow{\alpha_1\eta} [k'_1] \xrightarrow{\beta_2\theta} [m_2]$ we can shorten the zig-zag length by one, yielding the case $N - 1$, which is true by hypothesis. Therefore $(x, t) = (y, s)$ and the lemma follows.

$\natural$

Theorem 19.2.25. Let X be a simplicial set. The geometric realization of X has a natural structure of a CW-complex with exactly one closed n-cell $\widehat{x}\colon \Delta^n_{\text{top}} \to |X|$ for each non-degenerate n-cell $x \in X_n$.

Proof. Let $(x,t),(y,s) \in \coprod_{x \in X^{\text{nd}}_n} \Delta^n_{\text{top}}$ be any two points. We analyse the case where $x \otimes t = y \otimes s$ are equivalent points of $|X|_n$ by means of Lemma 19.2.24:

- If both $t, s \in \operatorname{Int} \Delta^n_{\text{top}}$ then by the uniqueness of representatives we find that
$$(x,t) = (y,s).$$
- Now if for instance $t \in \partial \Delta^n_{\text{top}}$, then there exits a unique $(z,r) \in X_k \times \operatorname{Int} \Delta^k_{\text{top}}$ for some minimal $k < n$ such that $x \otimes t = z \otimes t$. Notice however that since k is minimal and unique with such property, then it must be the case that $s \in \partial \Delta^n_{\text{top}}$—therefore in this case one has both points
$$(x,t),(y,s) \in \coprod_{x \in X^{\text{nd}}_n} \partial \Delta^n_{\text{top}}.$$

From this we can conclude that the following square is a pushout in $\mathtt{Top}$:

$$\begin{tikzcd}
\coprod_{x \in X^{\text{nd}}_n} \partial \Delta^n_{\text{top}} \arrow[r] \arrow[d, hook] & {|X|_{n-1}} \arrow[d, hook] \\
\coprod_{x \in X^{\text{nd}}_n} \Delta^n_{\text{top}} \arrow[r, "(\widehat{x})_{x \in X^{\text{nd}}_n}"'] \arrow[ur, phantom, "\ulcorner" very near end] & {|X|_n}
\end{tikzcd}$$

which proves that $|X|$ is a CW-complex. $\natural$

Definition 19.2.26 (Skeletal filtrations in $\mathtt{sSet}$)**.** Let X be a simplicial set. We define the *skeletal filtration* of X to be the collection $(\operatorname{sk}_n X)_{n \in \mathbf{N}}$ of simplicial sets, where $\operatorname{sk}_n X$ is the *smallest subcomplex* of X containing every k-cell $x \in X_k$ for each $0 \leqslant k \leqslant n$. Explicitly, the collection of k-cells of $\operatorname{sk}_n X$ is
$$(\operatorname{sk}_n X)_k = \bigcup_{0 \leqslant j \leqslant k} \big\{ xf \in X_k : x \in X_j \text{ and } f \in \operatorname{Mor}_{\mathbf{\Delta}}([k],[j]) \big\}.$$

Therefore for each $n \in \mathbf{N}$ the simplicial set $\operatorname{sk}_n X$ is a *subcomplex* of $\operatorname{sk}_{n+1} X$.

From construction we see that
$$X = \operatorname{colim}_{n \in \mathbf{N}} \operatorname{sk}_n X,$$
which is, loosely, the *union* of all $\operatorname{sk}_n X$.

Remark 19.2.27 (On Definition 19.2.26)**.** Note by a subobject $S \in \mathtt{sSet}$ of the simplicial set X we actually mean an isomorphism class of *natural monomorphisms*—meaning, a class of natural transformations $\eta \colon S \Rightarrow X$ for which each associated morphism $\eta_{[n]} \colon S_n \rightarrowtail X_n$ is a monomorphism in $\mathtt{Set}$. We say that (S, η) is *equivalent* to another subobject $(B, \xi \colon B \Rightarrow X)$ of X if there exists a *natural isomorphism* $\sigma \colon S \overset{\cong}{\Rightarrow} B$ such that, for any $[n] \in \mathbf{\Delta}$, the diagram

$$\begin{tikzcd}
 & X_k & \\
S_k \arrow[ur, tail, "\eta_{[k]}"] \arrow[rr, "\simeq", "\sigma_{[k]}"'] & & B_k \arrow[ul, tail, "\xi_{[k]}"']
\end{tikzcd}$$

commutes in $\mathtt{Set}$.

Corollary 19.2.28 (Pushouts & pullbacks in $\mathtt{sSet}$)**.** Let X, Y and Z be simplicial sets. We define the following:

(a) Given morphisms of simplicial sets $f \colon X \to Z$ and $g \colon Y \to Z$, the pullback of the pair (f, g) to be the simplicial set P defined as follows: for each $n \in \mathbf{N}$, the set P_n is the pullback

$$\begin{tikzcd}
P_n \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & Y_n \arrow[d, "g_n"] \\
X_n \arrow[r, "f_n"'] & Z_n
\end{tikzcd}$$

in the category of sets.

(b) Given morphisms of simplicial sets $u \colon X \to Z$ and $v \colon Y \to Z$, the pushout of the pair (u, v) to be the simplicial set Q defined as follows: for each $n \in \mathbf{N}$, the set Q_n is the pushout

$$\begin{tikzcd}
Z_n \arrow[r, "v_n"] \arrow[d, "u_n"'] & Y_n \arrow[d] \\
X_n \arrow[r] & Q_n \arrow[ul, phantom, "\ulcorner", very near start]
\end{tikzcd}$$

in the category of sets.

Proof. These limits all exist in $\mathtt{Set}$, therefore each limit is well defined in $\mathtt{sSet}$. ♮

Proposition 19.2.29. Given a simplicial set X, the square

$$\begin{tikzcd}
\coprod_{x \in X_n^{\mathrm{nd}}} \partial \Delta^n \arrow[r] \arrow[d, tail] & \mathrm{sk}_{n-1} X \arrow[d, tail] \\
\coprod_{x \in X_n^{\mathrm{nd}}} \Delta^n \arrow[r] & \mathrm{sk}_n X \arrow[ul, phantom, "\ulcorner", very near start]
\end{tikzcd}$$

is a pushout in $\mathtt{sSet}$.

Proof. Let $P \in \mathtt{sSet}$ be the pushout of the given square, and define $p\colon P \to \mathrm{sk}_n X$ to be unique morphism of simplicial sets making the diagram

$$\begin{tikzcd}
\coprod_{x \in X_n^{\mathrm{nd}}} \partial \Delta^n \arrow[r] \arrow[d] & \mathrm{sk}_{n-1} X \arrow[d, tail] \arrow[ddr, tail, bend left] & \\
\coprod_{x \in X_n^{\mathrm{nd}}} \Delta^n \arrow[r] \arrow[drr, bend right] & P \arrow[ul, phantom, "\ulcorner" very near start] \arrow[dr, dashed, "p"] & \\
 & & \mathrm{sk}_n X
\end{tikzcd}$$

commute in $\mathtt{sSet}$. We'll show that p is an isomorphism of simplicial sets:

- (Epimorphism) Let $x \in (\mathrm{sk}\, X_n)_k$ be any k-cell, and let $(y, \alpha\colon [k] \twoheadrightarrow [m])$ be the unique representative pair of x with $y \in X_m$ being a non-degenerate m-cell such that $x = y\alpha$ and $m \leqslant k$. If it is the case that $m < n$, then x is a degenerate k-cell and also $y \in \mathrm{sk}_{n-1} X$—but then $x \in \mathrm{sk}_{n-1} X$ via α^*, showing that $x \in \operatorname{im} p_k$. For the case where $m = n$ we see that x is non-degenerate and therefore $(x, \Delta^n) \in \coprod_{x \in X_n^{\mathrm{nd}}} \Delta^n$, thus $x \in \operatorname{im} p_k$ by the commutativity of

$$\begin{tikzcd}
(x, \Delta^n) \arrow[r] \arrow[dr, "x"'] & P_k \arrow[d, "p_k"] \\
 & (\mathrm{sk}_n X)_k
\end{tikzcd}$$

  where $x\colon \Delta^n \to (\mathrm{sk}_n X)_k$ is the unique representative morphism of x.

- (Monomorphism) To prove that p is a monomorphism it suffices to show the following two properties:

  (i) Let $x \in X_n^{\mathrm{nd}}$ be any non-degenerate n-cell and consider the canonical morphisms $x\colon \Delta^n \to \mathrm{sk}_n X$ and $\mathrm{sk}_{n-1} X \rightarrowtail \mathrm{sk}_n X$—we'll show that the *pullback* of this pair of morphisms is $\partial \Delta^n$.

  Let $\alpha \in \Delta^n[k]$ be a k-cell such that the k-cell $x\alpha \in X_k$ is an element of $(\mathrm{sk}_{n-1} X)_k$. Suppose, for the sake of contradiction, that α is an epimorphism, so that $n \leqslant k$. Since $x\alpha$ a k-cell of the $(n-1)$-skeleton, the unique pair $(y, \beta\colon [k] \twoheadrightarrow [m])$—where $y \in X_m$ is non-degenerate with $m < n$ and $y\beta = x\alpha$—is such that $y \in (\mathrm{sk}_n X)_m$. Since α and β are epimorphisms in $\mathbf{\Delta}$, their pushout exists and is absolute:

$$\begin{tikzcd}
{[k]} \arrow[r, two heads, "\alpha"] \arrow[d, two heads, "\beta"'] & {[n]} \arrow[d, "\delta"] \\
{[m]} \arrow[r, "\gamma"'] & {[\ell]} \arrow[ul, phantom, "\ulcorner" very near start]
\end{tikzcd}$$

Since X is covariant, the corresponding square is a pullback in sSet:

$$\begin{tikzcd} X_k & X_n \arrow[l, two heads, "\alpha^*"'] \\ X_m \arrow[u, two heads, "\beta^*"] \arrow[ur, phantom, very near start, "\lrcorner"] & X_\ell \arrow[l, "\gamma^*"] \arrow[u, "\delta^*"'] \end{tikzcd}$$

Using the fact that $y\beta = x\alpha$, there exists $z \in X_\ell$ such that $z\delta = x$ and $z\gamma = y$. Since x is assumed to be non-degenerate, and δ is an epimorphism with $\ell \leqslant n$ and $z\delta = x$, then it must be the case that $\delta = \mathrm{id}_{[n]}$ and hence $\ell = n$. Note however that $m < n$ so that $\gamma \colon [m] \to [n]$ should not be able to be an epimorphism, which is a contradiction. From this it follows that α isn't an epimorphism and hence $\alpha \in \partial\Delta^n[k]$ is a k-cell of the boundary. We thus conclude that

$$\begin{tikzcd} \partial\Delta^n \arrow[r] \arrow[d, tail] \arrow[dr, phantom, very near start, "\lrcorner"] & \mathrm{sk}_{n-1} X \arrow[d, tail] \\ \Delta^n \arrow[r, "x"'] & \mathrm{sk}_n X \end{tikzcd}$$

(ii) Given any two *distinct* non-degenerate n-cells $x, y \in X_n^{\mathrm{nd}}$, let the square

$$\begin{tikzcd} Q \arrow[r, "v"] \arrow[d, "w"'] & \Delta^n \arrow[d, "x"] \\ \Delta^n \arrow[r, "y"'] & \mathrm{sk}_n X \end{tikzcd}$$

be a pullback in sSet. Our goal will be to show that v and w can be factored through the monomorphism $\partial\Delta^n \rightarrowtail \Delta^n$.

Consider a pair of k-cells $\alpha, \beta \in \Delta^n[k]$ such that $x\alpha = y\beta$. First, notice that if one of them was an identity map, then k would coincide with n and it would follow that both α and β are identities, yielding $x = y$, which contradicts the assumption of x and y being distinct cells. We can therefore assume that neither α nor β are identities. Consider the factorization

$$\begin{tikzcd} {[k]} \arrow[rr, "\alpha"] \arrow[dr, two heads, "\alpha_-"'] & & {[n]} \\ & {[m_1]} \arrow[ur, tail, "\alpha_+"'] & \end{tikzcd} \qquad \begin{tikzcd} {[k]} \arrow[rr, "\beta"] \arrow[dr, two heads, "\beta_-"'] & & {[n]} \\ & {[m_2]} \arrow[ur, tail, "\beta_+"'] & \end{tikzcd}$$

Since α_- and β_- are epimorphisms, their pushout exists and is absolute:

$$\begin{tikzcd} {[k]} \arrow[r, two heads, "\alpha_-"] \arrow[d, two heads, "\beta_-"'] & {[m_1]} \arrow[d, "\delta"] \\ {[m_2]} \arrow[r, "\gamma"'] & {[\ell]} \arrow[ul, phantom, very near start, "\ulcorner"] \end{tikzcd} \tag{19.3}$$

Since X is covariant, the corresponding square is a pullback in $\mathtt{sSet}$:

$$\begin{tikzcd}
X_k \arrow[dr, phantom, "\lrcorner", very near start] & X_{m_1} \arrow[l, two heads, "\alpha_-^*"'] \\
X_{m_2} \arrow[u, two heads, "\beta_-^*"] & X_\ell \arrow[l, "\gamma^*"] \arrow[u, "\delta^*"']
\end{tikzcd}$$

Therefore there exists $z \in X_\ell$ for which $z\alpha_-\gamma = x$ while $z\delta\beta_- = y$. Now, since both x and y are non-degenerate cells, it follows that all maps in Eq. (19.3) are identity morphisms. It then follows that $\alpha = \alpha_+$ and $\beta = \beta_+$ are monomorphisms. Since neither of them is an identity, it follows that $\alpha, \beta \in \partial\Delta^n$ are k-cells of the boundary. This shows that v and w can be both factorized through $\partial\Delta^n \rightarrowtail \Delta^n$.

$\natural$

Chapter 20

Wheeled Graphs

20.1 Colouring the Portrait

Definition 20.1.1 (Colours & Profiles)**.** Fix, for the remaining of the chapter, a *non-empty set* $\mathfrak{C}$—whose elements will be called *colours*. A $\mathfrak{C}$*-profile* is a *finite*[1] *sequence* of colours of $\mathfrak{C}$.

Notation 20.1.2. We denote a $\mathfrak{C}$-profile by $\underline{c} = (c_1, \dots, c_n)$ or also $c_{[1,n]} = \underline{c}$ whenthe indexing matters. Some of the operations on profiles are the following:

- The lenght of the profile is denoted by $|\underline{c}| = n$.
- Given $1 \leqslant j \leqslant n$, we define the notion of colour removal:

$$\underline{c} \smallsetminus c_j \coloneqq (c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_n).$$

- Given another $\mathfrak{C}$-profile $\underline{d}$, with $|\underline{d}| = m$, we define the *concatenation* of $\underline{d}$ and $\underline{c}$ to be the $\mathfrak{C}$-profile

$$(\underline{d}, \underline{c}) \coloneqq (d_1, \dots, d_m, c_1, \dots, c_n).$$

- Given a permutation $\sigma \in \mathrm{Sym}_n$, we define the action of σ on the profile $\underline{c}$ to be the $\mathfrak{C}$-profile $\sigma\underline{c}$ given by

$$\sigma\underline{c} \coloneqq (c_{\sigma(1)}, \dots, c_{\sigma(n)}).$$

Definition 20.1.3 ($\mathfrak{C}$-profile category)**.** We define a category $\mathtt{Prof}(\mathfrak{C})$ whose objects are $\mathfrak{C}$-profiles, and a morphism $\sigma \colon \underline{c} \to \underline{d}$ is a *permutation* such that $\sigma\underline{c} = \underline{d}$.

[1] Empty sequences are also admitted.

Part VI

Manifold Theory

Chapter 21

Differentiable Manifolds

21.1 Differentiable Manifolds

Charts & Atlases

Definition 21.1.1 (Atlas)**.** Let X be an n-dimensional topological manifold. An *atlas* of class C^p on X is a collection of charts $\{(U_j, \phi_j)\}_{j\in J}$ such that

(a) For all $j \in J$, $U_j \subseteq X$ is an open set, and the collection $\{U_j\}_{j\in J}$ is an *open cover* for X.

(b) For every $j \in J$, $\phi_j: U_j \to V_j$ is a *topological isomorphism* from the open set $U_j \subseteq X$ to an open set $V_j \subseteq R^n$.

(c) For each pair $i, j \in J$, the induced *change of coordinates* map

$$\phi_j\phi_i^{-1}: \phi_i(U_i \cap U_j) \xrightarrow{\simeq} \phi_j(U_i \cap U_j)$$

is of class C^p — that is, every chart of the atlas with intersecting domain is *compatible*.

Definition 21.1.2 (Chart & atlas compatibility)**.** Let X be a topological n-manifold. If $(U, \phi: U \xrightarrow{\simeq} V)$ and $(U', \psi: U' \xrightarrow{\simeq} V')$ are *charts* on X, we say that they are *C^p-compatible* if the two induced transition maps

$$\phi\psi^{-1}: \psi(U \cap U') \longrightarrow \phi(U \cap U') \quad \text{and} \quad \psi\phi^{-1}: \phi(U \cap U') \longrightarrow \psi(U \cap U')$$

are of class C^p.

From this definition, we say that a chart is said to be compatible with a given atlas if it is compatible with every chart of the atlas. Moreover, given two atlases, we say that they are compatible if every chart of one is compatible with the other atlas.

Lemma 21.1.3. Let $\mathcal{A} := \{(U_j, \phi_j)\}_{j\in J}$ be an atlas on a topological manifold X. If both (V, ψ) and (W, σ) are charts of X compatible with the atlas $\mathcal{A}$, then they are compatible with each other.

Proof. Let $p \in V \cap W$ be any point and let $j \in J$ be such that $p \in U_j$ — thus $p \in V \cap W \cap U_j$. Since $\phi_j\psi^{-1}$ and $\sigma\phi_j^{-1}$ are C^p maps, then $\sigma\psi = (\sigma\phi_j^{-1}) \circ (\phi_j\psi^{-1})$ is C^p when restricted to $\psi(V \cap W \cap U_j)$. Moreover, since $\psi(p) \in \psi(V \cap W \cap U_j)$, it follows that $\sigma\psi$ is C^p on

$\psi(p)$ — therefore $\sigma\psi$ is C^p on every point of its domain since p was chosen arbitrarily. The same analogous argument can be used to show that $\psi\sigma^{-1}$ is C^p. ♮

Proposition 21.1.4. The compatibility of atlases form an *equivalence* relation.

Proof. Clearly reflexivity and symmetry are satisfied. Let $\mathcal{U} := \{(U_j, \phi_j)\}_{j\in J}$ and $\mathcal{V} := \{(V_i, \psi_i)\}_{i\in I}$ be two compatible atlases for some topological manifold X. If $\mathcal{A} := \{(A_s, \mu_s)\}_{s\in S}$ is another atlas for X, which happens to be compatible to $\mathcal{U}$, then for every $s \in S$ the maps $\phi_j\mu_s^{-1}$ and $\mu_s\phi_j^{-1}$ are of class C^p for any $j \in J$. Since $\phi_j\psi_i^{-1}$ and $\psi_i\phi_j^{-1}$ are C^p for all $i \in I$, then in particular

$$\mu_s\psi_i^{-1} = (\mu_s\phi_j^{-1}) \circ (\phi_j\psi_i^{-1}) \quad \text{and} \quad \psi_i\mu_s^{-1} = (\psi_i\phi_j^{-1}) \circ (\phi_j\mu_s^{-1})$$

are both maps of class C^p. Therefore we conclude that $\mathcal{A}$ is compatible with $\mathcal{V}$. ♮

Corollary 21.1.5. Any atlas on a topological manifold is contained in a *unique maximal atlas* — an atlas is said to be maximal if it isn't contained in any atlas other than itself.

C^p-Manifolds

Classical Definition

In this chapter we shall mostly consider the case of C^∞-manifolds, also called *smooth* manifolds, but for generality we'll define differentiable manifolds for all $p \in \mathbf{N} \cup \{\infty\}$.

Definition 21.1.6 (C^p-manifold structure on X)**.** The equivalence classes of atlases of class C^p on a topological space X define what is called a *C^p-manifold structure on X*.

Alternative Definition

We now give another definition of a C^p-manifold structure on topological spaces, to do that, we first introduce the following concept.

Definition 21.1.7 (Functionally structured space)**.** Let X be a topological space. A *functional structure* on X is a map F_X on the collection of open sets of X such that, for any open set $U \subseteq X$, we have:

(a) $F_X(U)$ is a *subalgebra* of $C(U)$, the algebra of continuous real valued maps on U.

(b) $F_X(U)$ contains all constant maps.

(c) If $V \subseteq U$ is another open set of X, and $f \in F_X(U)$, then $f|_V \in F_X(V)$.

(d) If $U = \bigcup_{j\in J} U_j$, and $f\colon U \to \mathbf{R}$ is a continuous map such that $f|_{U_j} \in F_X(U_j)$ for each $j \in J$, then it follows that $f \in F_X(U)$.

The pair (X, F_X) is called a *functionally structured space*.

Let $U \subseteq X$ be open. For any open set $V \subseteq U$ we define

$$F_U(V) := F_X(V),$$

and hence (U, F_U) is a functionally structured space.

Definition 21.1.8 (Morphisms of functionally structured spaces)**.** A morphism

$$\phi\colon (X, F_X) \to (Y, F_Y)$$

between functionally structured spaces is a map $\phi\colon X \to Y$ such that, for any open set $V \subseteq Y$ and $f \in F_Y(V)$, we have $f\phi \in F_X(\phi^{-1}(V))$.

Definition 21.1.9 (Second definition of a C^p-manifold)**.** An n-dimensional differentiable manifold is second countable functionally structured Hausdorff space (M, F) which is locally isomorphic to $(\mathbf{R}^n, C^p)$.

The local isomorphism is equivalent to the requirement that, for each point $p \in M$, there exists a neighbourhood $U \subseteq M$ of p such that $(U, F_U) \simeq (V, C^p_V)$ as functionally structured spaces — for some open set $V \subseteq \mathbf{R}^n$.

Lemma 21.1.10. Let $U, V \subseteq R^n$ be open subspaces. An isomorphism between functionally structured spaces

$$\phi\colon (U, C^p_U) \xrightarrow{\simeq} (V, C^p_V)$$

is a map $\phi\colon U \to V$ of class C^p if and only if $f\phi \in C^p(U)$ for all $f \in C^p(V)$.

Proof. Clearly, if $\phi\colon U \to V$ is of class C^p then $f\phi$ is a composition of C^p maps, thus $f\phi \in C^p(U)$ for any $f \in C^p(V)$.

Conversely, if we have the hypothesis that $f\phi \in C^p(U)$ for all $f \in C^p(V)$, one may consider the canonical projections $\pi_j\colon \mathbf{R}^n \twoheadrightarrow \mathbf{R}$ and notice that $\pi_j\phi \in C^p(U)$. Therefore each component of ϕ is a $C^p(U)$ map, implying that ϕ itself is a $C^p(U, V)$ map. ♮

Lemma 21.1.11 (Equivalence of the definitions)**.** The constructions on Definition 21.1.6 and Definition 21.1.9 are equivalent.

Proof. Let (M, F) be a C^p-manifold in the sense of Definition 21.1.9. A chart on M will be interpreted as a pair $(U, \phi\colon U \xrightarrow{\simeq} V)$ such that ϕ is an isomorphism $(U, F) \simeq (V, C^p_V)$ of functionally structured spaces. Since every point of M has a neighbourhood from which one can define the above mentioned isomorphism, we see that these charts do cover the whole space M.

It remains to be proven that the transition maps are C^p. From our interpretation of chart, given any two charts $\phi\colon (U, F_U) \xrightarrow{\simeq} (V, C^p_V)$ and $\psi\colon (U', F_{U'}) \xrightarrow{\simeq} (V', C^p_{V'})$ in M, since $\phi\psi^{-1}$ and $\psi\phi^{-1}$ are isomorphisms of functionally structured spaces (V, C^p_V) and $(V', C^p_{V'})$ — by Lemma 21.1.10 they are C^p maps. Therefore, our collection of charts match the requirements of Definition 21.1.1.

For the converse, let M be a C^p-manifold with an atlas $\mathcal{A}$. For every chart $(U, \phi\colon U \xrightarrow{\simeq} V) \in \mathcal{A}$ (where ϕ is a topological morphism), define

$$F(U) := \{f\phi \in C^p(U) : f \in C^p(V)\}.$$

Let $x \in M$ be any point and consider a chart $(U, \psi\colon U \xrightarrow{\simeq} V) \in \mathcal{A}$, where $U \subseteq M$ is a neighbourhood of x. Notice that ϕ naturally induces a morphism of functionally structured spaces $\phi\colon (U, F_U) \to (V, C^p_V)$ — moreover, since ϕ is a topological isomorphism, then ϕ is an isomorphism $(U, F_U) \simeq (V, C^p_V)$. ♮

Definition 21.1.12 (Orientation)**.** A manifold M is said to be *oriented* if for any two charts ϕ and ψ on M, the *Jacobian* matrix of the *transition* map $\phi\psi^{-1}$ has a *positive determinant for all points of its domain*. This maximal atlas on M is said to define an *orientation* on M. A manifold endowed with such an atlas is said to be *orientable*.

Smooth Morphisms

Morphisms Between Manifolds

Remark 21.1.13. From now on, unless stated otherwise, all manifolds are assumed to be endowed with a smooth structure.

Remark 21.1.14 (Chart notation)**.** When convenient, a chart $(U, \phi: U \to V)$ shall simply be denoted by $\phi: U \to V$.

Definition 21.1.15 (Smooth map)**.** Let M and N be smooth manifolds of dimension m and n, respectively. A *continuous* map $f: M \to N$ is said to be *smooth at a point* $p \in M$ if there exists a chart (V, ψ) about $f(p)$ in N, and a chart (U, ϕ) about p in M such that the map

$$\psi f \phi^{-1}: \phi(f^{-1}(V) \cap U) \to \mathbf{R}^n,$$

where $\phi(f^{-1}(V) \cap U) \subseteq \mathbf{R}^m$, is smooth at $\phi(p)$. Naturally, f is said to be *smooth* when f is smooth at every point of M.

Remark 21.1.16. The requirement of continuity of $f: M \to N$ in Definition 21.1.15 is necessary to ensure that $f^{-1}(V) \subseteq M$ is open.

Lemma 21.1.17 (Smooth maps are choice-independent)**.** Let $f: M \to N$ be a smooth map at $p \in M$ between smooth manifolds. If (U, ϕ) is any chart about p in M and (V, ψ) is any chart about $f(p)$ in N, then the composition map $\psi f \phi^{-1}$ is C^∞ at $\phi(p)$. Since both M and N are endowed with C^∞ compatible charts, it follows that the transition maps $\phi_0\phi^{-1}: \phi(U_0 \cap U) \to \phi_0(U_0 \cap U)$ and $\psi\psi_0^{-1}: \psi_0(V_0 \cap V) \to \psi(V_0 \cap V)$ are both C^∞ maps. Therefore the composition

$$(\psi\psi_0^{-1}) \circ (\psi_0 f \phi_0^{-1}) \circ (\phi_0\phi^{-1}) = \psi f \phi^{-1}: \phi(U_0 \cap U) \longrightarrow \mathbf{R}^n,$$

where n is the dimension of N, is a C^∞ map. Since $p \in U_0 \cap U$, then $\psi f \phi^{-1}$ is C^∞ at $\phi(p)$.

Proof. From the definition of smoothness, let (U_0, ϕ_0) be a chart about p of M and (V_0, ψ_0) be a chart about $f(p)$ of N such that the map $\psi_0 f \phi_0^{-1}$ is C^∞ at $\phi_0(p)$. ♮

Proposition 21.1.18. The composition of smooth maps is smooth.

Proof. Let N, M and W be any three manifolds, and consider two smooth maps $f: N \to M$ and $g: M \to W$. Take any three charts (U, ϕ) of N, (V, ψ) of M, and (E, γ) of W. Notice that the composition

$$(\gamma g \psi^{-1}) \circ (\psi f \phi^{-1}) = \gamma(gf)\phi^{-1}: \phi(f^{-1}(V) \cap U) \longrightarrow \mathbf{R}^w,$$

where w is the dimension of W, is a C^∞ map since both $\gamma g \psi^{-1}$ and $\psi f \phi^{-1}$ are C^∞. Since the charts where chosen arbitrarily, it follows that gf is a smooth map. Notice that the last two conclusions came directly from Proposition 21.1.19. ♮

Equivalent Conditions for Smoothness

Proposition 21.1.19 (Equivalent conditions for smoothness)**.** Let $f\colon N \to M$ be a continuous map between manifolds N and M, with respective dimensions n and m. The following properties are equivalent:

(a) The continuous map $f\colon N \to M$ is a smooth map.

(b) There exists atlases $\mathcal{N}$ for N and $\mathcal{M}$ for M such that, for every chart $(U,\phi) \in \mathcal{N}$ and $(V,\psi) \in \mathcal{M}$, the composition

$$\psi f \phi^{-1}\colon \phi(f^{-1}(V) \cap U) \to \mathbf{R}^m$$

is a C^∞ map.

(c) For all pairs of charts (U,ϕ) on N and (V,ψ) on M, the composition

$$\psi f \phi^{-1}\colon (f^{-1}(V) \cap U) \to \mathbf{R}^m$$

is a C^∞ map.

Proof.
- (a) $\Rightarrow$ (c): Consider any pair of charts (U,ϕ) on N and (V,ψ) on M for which $f^{-1}(V) \cap U$ is non-empty. Take any $p \in f^{-1}(V) \cap U$. Since f is C^∞, it follows that the composition $\psi f \phi^{-1}$ is C^∞ at $\phi(p)$ — thus smooth for any point of its domain.
- (c) $\Rightarrow$ (b): It suffices to choose atlases with C^∞ compatible charts for N and M.
- (b) $\Rightarrow$ (a): If $p \in N$ is any point, choose a chart (U,ϕ) about p and (V,ψ) about $f(p)$. Property (b) ensures that the composition $\psi f \phi^{-1}$ is, in particular, continuous at $\phi(p)$, since $p \in f^{-1}(V) \cap U$.

♮

Corollary 21.1.20. Let M be an n-manifold and $f\colon M \to \mathbf{R}^d$ be a continuous map. The following properties are equivalent:

(a) The map $f\colon M \to \mathbf{R}^d$ is a C^∞-morphism.

(b) The manifold M is endowed with an atlas such that, for every chart $\phi\colon U \to V$ of M, the map $f\phi^{-1}\colon \phi(U) \to \mathbf{R}^d$ is a map of class C^∞, where $V \subseteq \mathbf{R}^n$.

(c) For all charts $\phi\colon U \to V$ of M, the map $f\phi^{-1}\colon \phi(U) \to \mathbf{R}^d$ is of class C^∞, where $V \subseteq \mathbf{R}^n$.

Proof. The whole proposition is simply a particular case of Lemma 21.1.26. ♮

Proposition 21.1.21 (Smoothness from projections)**.** Let M be an n-manifold. A map $f\colon M \to \mathbf{R}^d$ is a smooth map if and only if its projections $f_j\colon M \to \mathbf{R}$, for all $1 \leqslant j \leqslant d$, are smooth maps.

Proof. From definition, f is a smooth map if and only if, for each chart $\phi: U \to V$ of M, the map $f\phi^{-1}: V \to \mathbf{R}^d$ is of class C^∞. Moreover, from the definition of continuity on real spaces, we find that $f\phi^{-1}$ is C^∞ if and only if $f_j\phi^{-1}$ is C^∞ for all $1 \leqslant j \leqslant d$. ♮

Proposition 21.1.22. Let $f: N \to M$ be a continuous map between two manifolds of dimensions n and m, respectively. The following properties are equivalent:

(a) The map $f: N \to M$ is a C^∞-morphism.

(b) The manifold M is endowed with an atlas such that, for all charts (V, ψ) of M, the map $\psi f: f^{-1}(V) \to \mathbf{R}^m$ is a C^∞-morphism.

(c) For every chart (V, ψ) of M, the map $\psi f: f^{-1}(V) \to \mathbf{R}^m$ is a C^∞-morphism.

Proof. (b) $\Rightarrow$ (a): Together with the atlas of M, the continuous map f induces a structure on N as follows. For each chart (V, ψ) of M, construct a collection $\mathcal{F}_V$: for every chart (U, ϕ) in the atlas of N, define a chart $(U \cap f^{-1}(V), \phi|_{U \cap f^{-1}(V)})$ — define $\mathcal{F}_V$ to be the collection of all such chart. Then $\mathcal{F}_V$ is a smooth atlas for $f^{-1}(V) \subseteq N$. From (b) we know that $\psi f: f^{-1}(V) \to \mathbf{R}^m$ is of class C^∞, then by Corollary 21.1.20 we have that the map $\psi f \phi^{-1}: \phi(U \cap F^{-1}(V)) \to \mathbf{R}^m$ is of class C^∞. Now by Proposition 21.1.19 we obtain that f is a smooth morphism.

(a) $\Rightarrow$ (c): Coordinate charts are C^∞-morphisms and f is a C^∞-morphism by hypothesis. Therefore by Proposition 21.1.18 we find that ψf is a C^∞-morphism. The implication (c) $\Rightarrow$ (b) is immediate. ♮

Corollary 21.1.23. Let $f: N \to M$ be a continuous map between manifolds of dimensions n and m, respectively. The following properties are equivalent:

1. The map $f: N \to M$ is a C^∞-morphism.
2. The manifold M is endowed with an atlas such that, for all charts (V, ψ), the projections $\psi_j f: f^{-1}(V) \to \mathbf{R}$ of f relative to the chart, for $1 \leqslant j \leqslant m$, are all C^∞-morphisms.
3. For every chart (V, ψ) of M, the components of f with respect to the chart are smooth maps, that is, $\psi_j f: f^{-1}(V) \to \mathbf{R}$ are C^∞-morphisms.

Proof. The proposition is consequence of Proposition 21.1.21 together with Proposition 21.1.22. ♮

The Category of Smooth Manifolds

Definition 21.1.24 (Category of C^∞-manifolds)**.** We define `Man` to be the category of smooth manifolds and smooth morphisms between them, these will be interchangeably called C^∞-morphisms.

Corollary 21.1.25. An isomorphism in the category `Man` is a bijective C^∞-morphism of manifolds with a smooth inverse. Some call these isomorphisms by "diffeomorphisms", we shall call them C^∞-isomorphisms or smooth isomorphisms.

Lemma 21.1.26. Let (M, F_M) and (N, F_N) be smooth manifolds. A map $f: M \to N$ is smooth in the sense of Definition 21.1.15 if and only if f is smooth in the sense of Definition 21.1.8.

Proof. Assume that M and N are, respectively, m and n-dimensional spaces. First we consider f as a morphism of functionally structured spaces. Let $p \in M$ be any point and consider charts, in the sense of isomorphisms of functionally structured spaces:

- $\phi: U \xrightarrow{\simeq} V$ in M — where U is a neighbourhood of p and $V \subseteq R^m$
- $\psi: U' \xrightarrow{\simeq} V'$ in N — where U' is a neighbourhood of $f(p)$ and $V' \subseteq \mathbf{R}^n$.

Since ϕ is an isomorphism, one can consider its inverse and use the property that, for any open set $S \subseteq V$ and map $g \in F_M(S)$, then $g\phi^{-1} \in C^\infty(\phi(S))$. Since f is a morphism of functionally structured spaces, in particular $\pi_j \psi \in F_N(U')$ for all $1 \leqslant j \leqslant n$, then $\pi_j \psi f \in F_M(f^{-1}(U'))$. Therefore, since $p \in U$ and $f(p) \in U'$, the intersection $f^{-1}(U') \cap U \subseteq M$ is non-empty and we can conclude that

$$\pi_j \psi f \phi^{-1} \in C^\infty(\phi(f^{-1}(U') \cap U)).$$

Since this is the case for every $1 \leqslant j \leqslant n$, it follows that $\psi f \phi^{-1}$ is C^∞ — therefore f is smooth in the sense of Definition 21.1.15.

For the converse, suppose $f: M \to N$ is smooth as in Definition 21.1.15. From the last property of the functional structures on spaces, we can simply consider a chart $\psi: U' \to V'$ in N and define a functional structure on N as

$$F_N(U') := \{g\phi \in C^\infty(U') : g \in C^\infty(V')\}.$$

If $h \in F_N(U')$ is any map, then the analogous structure F_M on M has

$$F_M(f^{-1}(U')) = \{wf \in C^\infty(f^{-1}(U')) : w \in C^\infty(U')\},$$

therefore $hf \in F_M(f^{-1}(U'))$ as wanted. ♮

Definition 21.1.27 (Lie group)**.** A *Lie group* is a *smooth manifold* G together with a multiplicative structure $G \times G \to G$ and an inverse map $G \to G$, both of which are C^∞*-morphisms*, making G into a *group*.

C^∞-Isomorphisms

Proposition 21.1.28 (Coordinate maps are C^∞-isomorphisms)**.** Let $(U, \phi: U \to V)$ be a chart on an n-manifold M, then ϕ is a C^∞-isomorphism.

Proof. By the definition, the coordinate map ϕ is a topological isomorphism, we thus show that both ϕ and ϕ^{-1} are smooth maps. Lets endow U with the natural structure of a manifold via the smooth atlas $\{(U, \phi)\}$, now for V we canonically endow it with the smooth atlas $\{(V, \mathrm{id}_V)\}$. For the smoothness of ϕ, it suffices to observe that $\mathrm{id}_V \phi\phi^{-1} = \mathrm{id}_V$ is certainly C^∞, therefore by Proposition 21.1.19 we find that ϕ is smooth. Analogously, for the smoothness of $\phi^{-1}: V \to U$, we know that $\phi\phi^{-1}\,\mathrm{id}_V = \mathrm{id}_V$, thus ϕ^{-1} is smooth for the same reason. ♮

Proposition 21.1.29 (C^∞-isomorphisms are charts)**.** Let M be an n-manifold and $U \subseteq M$ be any subset. If $f: U \to V$, where $V \subseteq \mathbf{R}^n$, is a C^∞-isomorphism, the induced pair (U, f) is a *chart* in the smooth structure of M.

Proof. Given any chart (E, ϕ) on M with non-empty $E \cap U$, we have that $f\phi^{-1}$ and ϕf are compositions of smooth maps (by Proposition 21.1.28 and Proposition 21.1.18), therefore both are smooth maps. This shows that (U, f) is compatible with any intersecting chart of M, therefore by the maximality of the smooth atlas of M we find that (U, f) is a chart of M. ♮

Proposition 21.1.30. Let $U \subseteq M$ be an open set in the n-dimensional smooth manifold M. If $F: U \to B$ is a C^∞-isomorphism where $B \subseteq \mathbf{R}^n$ is an open set, then the pair (U, F) is a coordinate chart in the differentiable structure of M.

Proof. Given any chart (V, ϕ) of the manifold M where $U \cap V$ is non-empty, one has that the transition maps $F\phi^{-1}$ and ϕF^{-1} are both compostions of C^∞-isomorphisms—thus themselves C^∞-isomorphisms. Since the differentiable structure of M is defined to be maximal, it must be the case that (U, F) is a chart in the atlas of M. ♮

Local Coordinates

Let M be an smooth n-manifold and $f: M \to \mathbf{R}$ be any real valued map. If $x: U \to V$ is a chart of M, we consider the induced map $\overline{f} := f x^{-1}: V \to \mathbf{R}$, a multivariable real map from a subset of $\mathbf{R}^n$ to $\mathbf{R}$. For any $p \in U$, one has *"coordinates"* $x(p) = (x_1(p), \dots, x_n(p))$. These coordinates can be used to compute f via

$$f(p) = \overline{f}(x_1(p), \dots, x_n(p)),$$

working as some kind of local coordinates on the open set U. This an intuitive view of what the following definition states.

Definition 21.1.31 (Local coordinates)**.** If X is an $\mathbf{R}^n$ modelled manifold, and $\phi: U \to \mathbf{R}^n$ is a chart, where $U \subseteq X$ is open, then, we say that the collection $(\phi_j)_{j=1}^n$ are *local coordinates* for X on U.

Partial Derivatives

Definition 21.1.32 (Partial derivative)**.** Let M be an n-manifold, and (U, ϕ) be a chart of M. Given a C^∞-morphism $f: M \to \mathbf{R}$ we define, for all $p \in U$, the *partial derivative* of f with respect to ϕ_j at p to be

$$\left.\frac{\partial}{\partial \phi_j}\right|_p f := \frac{\partial f}{\partial \phi_j}(p) := \frac{\partial (f\phi^{-1})}{\partial \pi_j}(\phi(p)) := \left.\frac{\partial}{\partial \pi_j}\right|_{\phi(p)} (f\phi^{-1}).$$

where ϕ_j is the j-th projection of ϕ, and $\pi_j: \mathbf{R}^n \twoheadrightarrow \mathbf{R}$ is the j-th canonical projection, for $1 \leqslant j \leqslant n$. Since ϕ is a bijective map, we have that

$$\frac{\partial f}{\partial \phi_j} \circ \phi^{-1} = \frac{\partial (f\phi^{-1})}{\partial \pi_j}: \phi(U) \longrightarrow \mathbf{R}$$

is a map of class C^∞ on $\phi(U)$. Moreover, since its pullback is C^∞, then $\partial f/\partial\phi_j$ is a C^∞-morphism on U.

Proposition 21.1.33. Let M be an n-manifold and (U,ϕ) be a chart of M. Then we have that for all $1 \leqslant i, j \leqslant n$, the projections of ϕ satisfy

$$\frac{\partial\phi_i}{\partial\phi_j} = \delta_{ij}.$$

Proof. From the definition, one has

$$\frac{\partial\phi_i}{\partial\phi_j}(p) = \frac{\partial(\phi_i\phi^{-1})}{\partial\pi_j}(\phi(p)) = \frac{\partial((\pi_i\phi)\circ\phi^{-1})}{\partial\pi_j}(\phi(p)) = \frac{\partial\pi_i}{\partial\pi_j}(\phi(p)) = \delta_{ij}$$

for any $p \in U$. ♮

Definition 21.1.34 (Jacobian)**.** Let $f\colon N \to M$ be a C^∞-morphism — assume n and m are the respective dimensions of N and M — and consider a pair of charts (U,ϕ) of N and (V,ψ) of M, for which $f(U) \subseteq V$. We define the *i-th projection f in the chart* (V,ψ) to be the map

$$f_i \coloneqq \psi_i f\colon U \longrightarrow \mathbf{R},$$

for all $1 \leqslant i \leqslant m$. We also define the *Jacobian matrix of f in the chart* (V,ψ) to be the $m \times n$ matrix whose (i,j)-th component is $\partial f_i/\partial\phi_j$ — for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$.

If it is the case that N and M have the same dimension, we define the *Jacobian of f in the chart* (V,ψ) to be the determinant of its respective Jacobian matrix.

Inverse Map Theorem for Manifolds

Theorem 21.1.35 (Inverse map theorem)**.** Let $f\colon N \to M$ be a C^∞-morphism, $p \in N$ be any point, and both N and M be n-dimensional manifolds. Let (U,ϕ) be a chart about p in N, and (V,ψ) be a chart about $f(p)$ in M, such that $f(U) \subseteq V$. Then f is *locally invertible at p* if and only if the *determinant* $\det\left[\frac{\partial f_i}{\partial\phi_j}\right]_{ij}$ is *non-zero*—where $f_i \coloneqq \psi_i f$.

Proof. Let $\pi_j\colon \mathbf{R}^n \to \mathbf{R}$ denote the j-th canonical projection, so that $f_i = \psi_i f = \pi_i\psi f$. The local representation of f_{*p} relative to the charts (U,ϕ) and (V,ψ) is given by

$$\left[\frac{\partial f_i}{\partial\phi_j}\right]_{i,j=1}^n = \left[\frac{\partial(\pi_i\psi f)}{\partial\phi_j}\right]_{i,j=1}^n = \left[\frac{\partial(\pi_i\psi f\phi^{-1})}{\partial\pi_j}(\phi(p))\right]_{i,j=1}^n. \tag{21.1}$$

That is, the local representation of f_{*p} is the same as the Jacobian matrix of

$$\psi f\phi^{-1}\colon \phi(f^{-1}V \cap U) \to \mathbf{R}^m$$

at the point $\phi(p)$. By the inverse map theorem for Banach spaces (see Theorem A.5.8), we find that the determinant of the Jacobian Eq. (21.1) is non-zero if and only if $\psi f\phi^{-1}$ is a locally invertible map at $\phi(p)$. Notice however that both ϕ and ψ are C^∞-isomorphisms, thus $\psi f\phi^{-1}$ is locally invertible at the said point if and only if f is invertible at p. ♮

Corollary 21.1.36. Let M be an n-dimensional manifold, and (U, ϕ) be a chart of M about a given point $p \in M$. A C^∞-morphism $F\colon U \to \mathbf{R}^n$, in the coordinate chart (U, ϕ), forms a coordinate chart[1] about p if and only if its Jacobian determinant $\det \left[\frac{\partial F_i}{\partial \phi_j}\right]_{ij}$ is non-zero.

Proof. From the inverse map theorem, the determinant of the Jacobian of F at the point p is non-zero if and only if $F\colon U \to \mathbf{R}^n$ is locally invertible at p. Moreover, the condition for F to be locally invertible is equivalent to the existence of a neighbourhood $X \subseteq M$ of p such that the induced map $F\colon X \to F(X)$ is a C^∞-isomorphism, and thus (X, F) is a coordinate chart about p (see Proposition 21.1.30). ♮

21.2 Structures

Definition 21.2.1 (Induced structure)**.** Let (X, F_X) be a functionally structured space, and $\phi\colon X \to Y$ be a continuous map, where Y is a topological space. We define the *induced functional structure* on Y via F_X and ϕ to be given by

$$F_Y(U) := \{f \in C(U) : f\phi \in F_X(\phi^{-1}(U))\},$$

for any open set $U \subseteq Y$.

Definition 21.2.2 (Induced structure on subspace)**.** Let (X, F) be a functionally structured space, and $A \subseteq X$ be a subspace. We construct a functional structure F_A on A as follows: for all open sets $U \subseteq A$, a continuous map $f\colon U \to \mathbf{R}$ is contained in $F_A(U)$ if and only if for every $p \in U$ there exists a neighbourhood $W \subseteq X$ of p such that f is the restriction to $W \cap A$ of some map $g \in F(W)$.

Definition 21.2.3 (Binary product of manifolds)**.** Let M and N be manifolds of dimension m and n, respectively. If $\phi\colon U \to \mathbf{R}^m$ is a chart for M and $\psi\colon V \to \mathbf{R}^n$ is a chart for N, we take $\phi \times \psi\colon U \times V \to \mathbf{R}^{m+n}$ to be a chart for the product space $M \times N$. The maximal atlas consisting of such product charts make $M \times N$ into a product manifold of dimension $m + n$.

Definition 21.2.4 (Submanifold)**.** Let M be an n-manifold. A subset $N \subseteq M$ is said to be a *k-manifold* of M if for each $p \in N$ there exists a chart $\phi\colon U \to V$ of M about p for which

$$\phi(U \cap N) = V \cap (\mathbf{R}^k \times \{0\}).$$

Charts with this property are called *adapted* to N. The maximal atlas containing all adapted charts to N makes N into a manifold.

We define the *codimension* of N *in* M to be the difference $n - k$. The submanifold N is said to be *smooth* exactly when about each point there exists an adapted smooth chart to N — so that one can find a unique maximal smooth atlas for N.

[1] That is, there exists a neighbourhood $X \subseteq M$ of p such that the induced map $F\colon X \to F(X)$ is a C^∞-isomorphism—then (X, F) is a coordinate chart about p.

Definition 21.2.5 (Smooth embedding)**.** A C^∞-morphism $f: N \to M$ is said to be a *smooth embedding* if the image $f(N) \subseteq M$ is a *smooth submanifold* of M and the induced restriction $f: N \xrightarrow{\simeq} f(N)$ is a C^∞-isomorphism.

Example 21.2.6 (Sphere)**.** Consider the sphere $S^n \subseteq \mathbf{R}^{n+1}$, we'll define a smooth atlas on S^n using solely two charts. To that end, define two points $n := (0, \dots, 0, 1)$ and $s := (0, \dots, 0, -1)$ in $\mathbf{R}^{n+1}$ — these correspond to the north and south poles of S^n, respectively. Consider the *stereographic projection* $\phi_n: S^n \smallsetminus \{n\} \to \mathbf{R}^n$, from the sphere with its north pole cut out to the hyperplane $R^n \simeq \mathbf{R}^n \times \{0\}$ — the stereographic projection maps each point $x \in S^n$ to the intersection of the line passing through x and n, and the hyperplane above-mentioned. The inverse map of ϕ_n is $\pi_n: \mathbf{R}^n \to S^n \smallsetminus \{n\}$, mapping

$$x \longmapsto \frac{(2x, \|x\|^2 - 1)}{(1 + \|x\|)^2}.$$

We define the south stereographic projection $\phi_s: S^n \smallsetminus \{s\} \to \mathbf{R}^n$ equivalently. This construction has a smooth transition map

$$\phi_s \phi_n^{-1}(x) = \frac{y}{\|y\|^2}.$$

21.3 Partitions of Unity

Definition 21.3.1 (Adequate atlas)**.** Given an m-manifold M, we say that an atlas $(U_j, \phi_j)_{j \in J}$ is *adequate* if it is *locally finite*, and for all we have either $\phi_j(U_j) = \mathbf{R}^m$ or $\phi_j(U_j) = \mathbf{H}^m$, and

$$\bigcup_{j \in J} \phi_j^{-1}(\operatorname{Int} D^m) = M.$$

Since every topological manifold is paracompact, by Corollary 14.5.8 we obtain the following corollary.

Corollary 21.3.2. Let $\mathcal{U}$ be a covering of M. There exists an *adequate atlas* $(U_j, \phi_j)_{j \in J}$ such that $(U_j)_{j \in J}$ is a *refinement* of $\mathcal{U}$.

Definition 21.3.3 (Bump functions & partition of unity)**.** Let M be an m-manifold and $(U_j, \phi_j)_{j \in J}$ be an *adequate* atlas on M. Let $\lambda: \mathbf{R}^m \to \mathbf{R}$ be a *smooth non-negative* map such that $\lambda|_{D^m} := 1$ and $\lambda|_{\mathbf{R}^m \smallsetminus D^m(2)} := 0$.

- Define, for every $j \in J$, maps $\lambda_j: M \to \mathbf{R}$ where
  $$\lambda_j|_{U_j} := \lambda \phi_j \quad \text{and} \quad \lambda_j|_{M \smallsetminus U_j} := 0.$$
  We name the collection $(\lambda_j)_{j \in J}$ the *bump functions* associated to $(U_j)_{j \in J}$.
- Define, for each $j \in J$, maps $\mu_j: M \to \mathbf{R}$ given by
  $$\mu_j(p) := \frac{\lambda_j(p)}{\sum_{i \in J} \lambda_i(p)}$$
  for all $p \in M$. The collection $(\mu_j)_{j \in J}$ is said to be the *partition of unity* associated to the atlas $(U_j, \phi_j)_{j \in J}$.

21.4 Tangent Space

Algebra of Germs

Define an equivalence relation $\sim_p$ on the set of real-valued C^∞-morphisms in M, where $p \in M$ is any point, as follows: given neighbourhoods U and V of p, and C^∞-morphisms $f\colon U \to \mathbf{R}$ and $g\colon V \to \mathbf{R}$, we say that $f \sim_p g$ if and only if there exists a neighbourhood $Q \subseteq U \cap V$ of p such that $f|_Q = g|_Q$.

Definition 21.4.1 (Algebra of germs in a manifold)**.** Given a manifold M and a point $p \in M$, we define the *germs of real-valued C^∞-morphisms at p in M*, to be the natural $\mathbf{R}$-algebra on the following quotient:

$$C_p^\infty(M) := \{f \in \mathrm{Mor}_{\mathtt{Man}}(U, \mathbf{R}) : p \in U \subseteq M\} \Big/ {\sim_p}.$$

A *point-derivation* of $C_p^\infty(M)$ is an $\mathbf{R}$*-linear map* $X\colon C_p^\infty(M) \to \mathbf{R}$ such that

$$X(fg) = (Xf)g(p) + f(p)(Xg).$$

Definition 21.4.2 (Tangent vector & tangent space)**.** Let M be a manifold and $p \in M$ be any point. A *tangent vector* at p is a *point-derivation* on $C_p^\infty(M)$. The collection of tangent vectors at p, denoted by T_pM, together with a natural action $\mathbf{R} \times T_pM \to T_pM$, forms the $\mathbf{R}$-vector space called *tangent space of M at p*.

Differential of a Smooth Morphism

Definition 21.4.3 (Differential at a point)**.** Let $f\colon N \to M$ be a C^∞-morphism. At every point $p \in N$, the morphism f induces a pushforward

$$f_*\colon T_pN \longrightarrow T_{f(p)}M,$$

called *differential of f at p*, which is an $\mathbf{R}$-linear map between tangent spaces. For any tangent vector $X_p\colon C_p^\infty(N) \to \mathbf{R}$ of T_pN, we define the respective tangent vector $f_*(X_p)\colon C_{f(p)}^\infty(M) \to \mathbf{R}$ of $T_{f(p)}M$ as the $\mathbf{R}$-linear map given by

$$(f_*(X_p))(g) := X_p(gf) \in \mathbf{R},$$

for all $g \in C_{f(p)}^\infty(M)$.

Proposition 21.4.4. Let $f\colon N \to M$ and $g\colon M \to P$ be C^∞-morphisms of manifolds, and $p \in N$ be a given point. Then one has that the differential of the composition gf at p, the linear map $(gf)_{*p}\colon T_pN \to T_{f(p)}M$, satisfies

$$(gf)_{*p} = g_{*f(p)} f_{*p},$$

where $g_{*f(p)}$ denotes the differential of g at $f(p)$, while f_{*p} the differential of f at p.

Proof. Let $X_p \in T_p N$ be any tangent vector and $\ell \in C^\infty_{gf(p)}(P)$, then one has

$$((gf)_* X_p)(\ell) = X_p(\ell(gf)),$$
$$((g_* f_*) X_p)(\ell) = (g_*(f_* X_p))(\ell) = (f_* X_p)(\ell g) = X_p((\ell g) f) = X_p(\ell(gf)).$$

Where we used the fact that $f_* X_p \in T_{f(p)} M$. Therefore we conclude that $((gf)_* X_p)(\ell) = ((g_* f_*) X_p)(\ell)$. ♮

Corollary 21.4.5. Let $f: N \to M$ be a C^∞-isomorphism and $p \in N$ be any point, then the R-linear map $f_*: T_p N \to T_{f(p)} M$ is an isomorphism of R-vector spaces.

Proof. Since f is an isomorphism of manifolds, consider its inverse $f^{-1}: M \to N$. Using the chain rule, we find that

$$\mathrm{id}_{T_p N} = (\mathrm{id}_N)_{*p} = (f^{-1} f)_{*p} = f^{-1}_{*f(p)} f_{*p},$$
$$\mathrm{id}_{T_{f(p)} M} = (\mathrm{id}_M)_{*f(p)} = (f f^{-1})_{*f(p)} = f_{*f^{-1}f(p)} f^{-1}_{*f(p)} = f_{*p} f^{-1}_{*f(p)}.$$

This shows that f_{*p} and $f^{-1}_{*f(p)}$ are inverses of each other—thus isomorphisms. ♮

Corollary 21.4.6 (Dimension invariance)**.** If $U \subseteq \mathbf{R}^n$ is an open set C^∞-isomorphic to an open set $V \subseteq \mathbf{R}^m$, then $n = m$.

Proof. Let $f: U \to V$ be a C^∞-isomorphism. For any $p \in U$ one has that $f_{*p}: T_p U \to T_{f(p)} V$ is an isomorphism of vector spaces by Corollary 21.4.5. Since $\dim_{\mathbf{R}} T_p U = n$ and $\dim_{\mathbf{R}} T_{f(p)} V = m$, then necessarily $n = m$. ♮

Definition 21.4.7 (Immersions and submersions)**.** Let $f: N \to M$ be a C^∞-morphism, and $p \in N$ be any point—also assume that $\dim N = n$ and $\dim M = m$. We define the following two concepts concerning the differential $f_{*p}: T_p N \to T_{f(p)} M$:

(a) The map f is said to be an *immersion* at p if f_{*p} is *injective*—if that is the case, then $n \leqslant m$. If f is an immersion at every point, we simply classify f as an immersion.

(b) The map f is said to be an *submersion* at p if f_{*p} is *surjective*—if that is the case, then $m \leqslant n$. If f is a submersion at every given point, we call f a submersion.

Remark 21.4.8. Mind you: immersions *do not need to be injective*, neither do submersions need to be surjective.

We now restate Definition 21.2.4 and Definition 21.2.5 with this new terminology of immersions. Given a C^∞-morphism $f: M \to N$, if f is an immersion and is injective, then the pair (M, f) defines a *submanifold* of N. Consequently, $f: M \to fM$ is a C^∞-isomorphism, being an *embedding*. Later (see Definition 21.5.1) we shall define fM to be an *embedded submanifold* on N.

Tangent Space Basis at a Point

Lemma 21.4.9. Let M be an n-dimensional smooth manifold and (U,ϕ) be a chart about a point $p \in M$, then for every $1 \leqslant j \leqslant n$ we have

$$\phi_*\Big(\frac{\partial}{\partial \phi_j}\Big|_p\Big) = \frac{\partial}{\partial \pi_j}\Big|_{\phi(p)}$$

where $(\pi_j)_{j=1}^n$ are the local coordinates of $\mathbf{R}^n$—that is, the canonical projections.

Proof. Let $f \in C^\infty_{\phi(p)}(\mathbf{R}^n)$ be any smooth morphism, then we have

$$\Big(\phi_*\Big(\frac{\partial}{\partial \phi_j}\Big|_p\Big)\Big)(f) = \frac{\partial}{\partial \phi_j}\Big|_p (f\phi) = \frac{\partial}{\partial \pi_j}\Big|_{\phi(p)} (f\phi)\phi^{-1} = \frac{\partial}{\partial \pi_j}\Big|_{\phi(p)} f(\phi\phi^{-1}) = \frac{\partial}{\partial \pi_j}\Big|_{\phi(p)} f$$

for each $1 \leqslant j \leqslant n$. ♮

Proposition 21.4.10 (Tangent space basis at a point)**.** Let M be an n-dimensional smooth manifold, and (U,ϕ) be a chart about $p \in M$. Then the R-vector space T_pM has a basis

$$\Big(\frac{\partial}{\partial \phi_1}\Big|_p, \dots, \frac{\partial}{\partial \phi_n}\Big|_p\Big)$$

Proof. Via Lemma 21.4.9 we know that the map $\phi_*: T_pM \to T_{\phi(p)}\mathbf{R}^n$ is R-linear and sends each $\frac{\partial}{\partial \phi_j}\big|_p$ to the corresponding base element $\frac{\partial}{\partial \pi_j}\big|_{\phi(p)}$ of $T_{\phi(p)}\mathbf{R}^n$. Therefore ϕ_* is an $\mathbf{R}$-linear isomorphism and the collection of tangent vectors $\frac{\partial}{\partial \phi_j}\big|_p$ is a basis. ♮

Proposition 21.4.11 (Transition map matrix)**.** Let (U,ϕ) and (V,ψ) be coordinate charts for an n-dimensional manifold M. Then one has on the intersection $U \cap V$, for each $1 \leqslant j \leqslant n$,

$$\frac{\partial}{\partial \phi_j} = \sum_{i=1}^n \frac{\partial \psi_i}{\partial \phi_j}\frac{\partial}{\partial \psi_i}$$

Proof. Let $p \in U \cap V$ be any point and define $[a_{ij}]_{i,j=1}^n$ be a change of basis matrix from $\big(\frac{\partial}{\partial \psi_j}\big|_p\big)_{j=1}^n$ to $\big(\frac{\partial}{\partial \phi_j}\big|_p\big)_{j=1}^n$—that is, the coefficients of the matrix must satisfy

$$\frac{\partial}{\partial \phi_j} = \sum_{k=1}^n a_{kj}\frac{\partial}{\partial y_k}$$

Since $\frac{\partial \psi_i}{\partial \psi_j} = \delta_{ij}$ then by applying the component ψ_i to the above equation of tangent vectors, one has

$$\frac{\partial \psi_i}{\partial \phi_j} = \sum_{k=1}^n a_{kj}\frac{\partial \psi_i}{\partial \psi_k} = \sum_{k=1}^n a_{kj}\delta_{ik} = a_{ij}.$$

Which proves the proposition. ♮

Proposition 21.4.12 (Local expression for the differential of a map)**.** Let $f\colon N \to M$ be a C^∞-morphism, and consider a point $p \in N$—moreover, let $\dim N = n$ and $\dim M = m$. Given coordinate charts (U, ϕ) about p, and (V, ψ) about $f(p)$, the differential $f_{*p}\colon T_pN \to T_{f(p)}M$ can be locally represented—that is, with respect to the bases $\big(\frac{\partial}{\partial\phi_j}\big|_p\big)_{j=1}^n$ and $\big(\frac{\partial}{\partial\psi_i}\big|_{f(p)}\big)_{i=1}^m$—by the matrix

$$\Big[\frac{\partial f_i}{\partial \phi_j}(p)\Big]_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}}$$

where $f_i \coloneqq \psi_i f$ is the i-th component of f with respect to ψ, for each $1 \leqslant i \leqslant m$.

Proof. Given the said coordinate charts, f_{*p} will be completely determined by its image on the local basis of T_pN given by (U, ϕ). That is, f_{*p} is given by the matrix $[a_{ij}]_{i,j=1}^n$ whose coefficients satisfy

$$f_*\Big(\frac{\partial}{\partial\phi_j}\Big|_p\Big) = \sum_{k=1}^m a_{kj}\frac{\partial}{\partial\psi_k}\Big|_{f(p)}.$$

Moreover, if we apply ψ_i to the above equation concerning tangent vectors and obtain

$$a_{ij} = \Big(\sum_{k=1}^m a_{kj}\frac{\partial}{\partial\psi_k}\Big|_{f(p)}\Big)(\psi_i) = f_*\Big(\frac{\partial}{\partial\phi_j}\Big|_p\Big)(\psi_i) = \frac{\partial}{\partial\phi_j}\Big|_p(\psi_i f) = \frac{\partial f_i}{\partial\phi_j}(p).$$

♮

Rank, & Critical & Regular points

Definition 21.4.13 (Rank at a point)**.** Let $f\colon N \to M$ be a C^∞-morphism, and $p \in N$ be any point. We define the *rank of f at the point p* to be

$$\operatorname{rank}_p f \coloneqq \operatorname{rank} f_{*p}.$$

Locally, by the independence of the choice of coordinate charts—given a chart (U, ϕ) about p, and a chart (V, ψ) about $f(p)$—one has

$$\operatorname{rank}_p f = \operatorname{rank}\Big[\frac{\partial f_i}{\partial\phi_j}(p)\Big].$$

Definition 21.4.14 (Critical & Regular)**.** Let $f\colon N \to M$ be a C^∞-morphism, and consider any two points $p \in N$ and $q \in M$. We define the following concerning the classification of the point p:

(a) We say that p is a *critical point* of f if the differential f_{*p} *isn't surjective*. The point q is said to be a *critival value* if it is the *image* of a critical point.

(b) We say that p is a *regular point* of f if the differential f_{*p} *is surjective*—that is, f is a *submersion at p*.

Finally, if q is not a critical value, we say that it is a *regular value*—we do *not* impose q to be an element of the image of f.

Proposition 21.4.15. Let $f: M \to \mathbf{R}$ be a C^∞-morphism, and $p \in M$ be any point. Then p is a *critical point* of f if and only if there exists a chart (U, ϕ) about p such that

$$\frac{\partial f}{\partial \phi_j}(p) = 0$$

for all $1 \leqslant j \leqslant n$.

Proof. Let (V, ψ) be *any* chart about p. Since $T_{f(p)}\mathbf{R} \simeq \mathbf{R}$, one has that the differential $f_{*p}: T_p M \to \mathbf{R}$ has rank $f_{*p} = 1$—which means that f_{*p} is surjective—or rank $f_{*p} = 0$, that is, $f_{*p} = 0$. Therefore f_{*p} isn't surjective (i.e. f *isn't* a submersion at p) if and only if f_{*p} is zero—which is equivalent to the vanishing of all partial derivatives. ♮

Tangent Bundle

Given an n-manifold M, define

$$TM := \bigcup_{p \in M} T_p M$$

which will be associated with a natural projection $\pi: TM \twoheadrightarrow M$ mapping each pair $(p, \xi) \in TM$, where $\xi \in T_p M$, to the point $p \in M$.

We wish to define a smooth structure on TM. To that end, let $p \in M$ be any point and $\phi: U \to U' \subseteq \mathbf{R}^n$ be a chart of M about p, forming a local basis $\big(\frac{\partial}{\partial \phi_j}\big)_{j=1}^n$ for the spaces $T_x M$ for any $x \in U$. From this we conclude that there exists a linear isomorphism $\pi^{-1}U \simeq U \times \mathbf{R}^n \simeq U' \times \mathbf{R}^n$ associating each pair $\big(p, \xi = \sum_{j=1}^n a_j \frac{\partial}{\partial \phi_j}\big) \in TM$ to $(p, \phi_* \xi) \in U \times \mathbf{R}^n$ or $(\phi p, \phi_* \xi) \in U' \times \mathbf{R}^n$. We shall take the latter, which has the form

$$(\phi\pi) \times \phi_*: \pi^{-1}U \longrightarrow U' \times \mathbf{R}^n,$$

as a chart on TM. To see that this gives a smoothly compatible structure to TM, let $\psi: V \to \mathbf{R}^n$ be another chart on M and consider the transition map $\theta := \psi\phi^{-1}: \phi(U \cap V) \to \psi(U \cap V)$. The corresponding transition map between the charts $(\phi\pi) \times \phi_*$ and $(\psi\pi) \times \psi_*$ is then $\theta \times \theta_*$, which is smooth. This proves that the constructed atlas gives to TM a smooth structure of a $2n$-dimensional manifold. We shall refer to such manifold as the *tangent bundle* of M.

Definition 21.4.16 (Vector field)**.** Let M be a manifold. A *vector field* on M is a *section* of the natural projection $\pi: TM \twoheadrightarrow M$—that is, a C^∞-morphism $\xi: M \to TM$ for which $\pi\xi = \mathrm{id}_M$. More explicitly, given a point $p \in M$ and a chart ϕ of M about p, we have

$$\xi p = \sum_{j=1}^n \alpha_j(p) \frac{\partial}{\partial \phi_j},$$

where $\alpha_j: M \to \mathbf{R}$ are C^∞-morphisms. The collection of all vector fields on M is denoted by $\mathfrak{X} M$.

Definition 21.4.17 (Flow)**.** Let M be an m-manifold. A *smooth flow* on M is a *smooth action* of the abelian topological group $(\mathbf{R}, +)$ on M, that is, a C^∞-morphism $\theta\colon \mathbf{R} \times M \longrightarrow M$ satisfying:

- For all $x \in M$, we have $\theta(0, x) = x$.
- For every pair $s, t \in \mathbf{R}$, and point $x \in M$, one has $\theta(s + t, x) = \theta(s, \theta(t, x))$.

Definition 21.4.18 (Tangent field to a flow)**.** Let θ be a flow on a manifold M. We define a *tangent field* of θ to be the vector field $\xi\colon M \longrightarrow TM$ given by

$$\xi p := \theta_* \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{(0,p)}$$

that is, the vector $\xi p \in TM$ is tangent to the curve $\gamma(t) := \theta(t, p)$ at the point $t = 0$.

Example 21.4.19. Given a sphere $S^{2n-1} \hookrightarrow \mathbf{C}^n$, consider the flow $\theta\colon \mathbf{R} \times S^{2n-1} \longrightarrow S^{2n-1}$ given by

$$\theta(t, z) := e^{it} z.$$

The tangent field $\xi\colon S^{2n-1} \longrightarrow TS^{2n-1}$ to θ is then

$$\xi(z) = \theta_* \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} z = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} e^{it} z = iz.$$

Therefore ξ is a unitary vector field on the $(2n-1)$-sphere.

Definition 21.4.20 (Parallelizable manifold)**.** An m-manifold M is said to be *parallelizable* if there exists a C^∞*-isomorphism* $TM \xrightarrow{\simeq} M \times \mathbf{R}^n$ for which the restriction $T_pM \longrightarrow p \times \mathbf{R}^n$ is an $\mathbf{R}$*-linear isomorphism* for each $p \in M$.

Example 21.4.21. The circle is parallelizable. Indeed, we can use $\xi\colon M \longrightarrow TM$ as the tangent vector field to the flow $\mathbf{R} \times S^1 \longrightarrow S^1$ mapping $(t, z) \mapsto e^{it} z$.

Proposition 21.4.22. The n-torus is parallelizable.

Proposition 21.4.23. If n is odd, then $S^n \times S^k$ is parallelizable for all $k \geqslant 1$.

To prove: n-torus and product of spheres are parallelizable

21.5 Submanifolds

Embedded Submanifolds

We shall give two descriptions of what is called an embedded submanifold:

Definition 21.5.1 (Embedded submanifold)**.** A subspace $S \subseteq N$ is said to be an *embedded submanifold* on a manifold N if the canonical inclusion $S \hookrightarrow N$ is a *smooth embedding*.

Definition 21.5.2 (Embedded submanifold, again[2])**.** Let $S \subseteq N$ be a subspace of the n-dimensional manifold N. We say that S is a *k-dimensional embedded submanifold* of N if for every $p \in S$ there exists a coordinate chart (U, ϕ) about p such that $n - k$ coordinates of ϕ vanish at the intersection $U \cap S$. By the possible rearrangement of the indices, we may assume that

$$\phi|_{U \cap S} = (\phi_1, \dots, \phi_k, 0, \dots, 0).$$

The chart (U, ϕ) is then refered to as an *adapted chart relative to S*. We shall denote by $\phi_S : U \cap S \to \mathbf{R}^k$ the induced C^∞-isomorphism given by

$$\phi_S := (\phi_1, \dots, \phi_k),$$

where $(U \cap S, \phi_S)$ is a chart for S in the subspace topology.

The proof of the equivalence between Definition 21.5.1 and Definition 21.5.2 shall be posponed to Theorem 21.6.6.

Definition 21.5.3 (Codimension)**.** Let $S \subseteq N$ be a k-dimensional submanifold of the n-dimensional manifold N. We define the *codimension* of S as a submanifold of N to be

$$\operatorname{codim}_N S := n - k.$$

Proposition 21.5.4 (Embedded submanifolds are manifolds)**.** Let S be a embedded submanifold of N and let $(U_\gamma, \phi_\gamma)_{\gamma \in \Gamma}$ be a collection of adapted charts of N relative to S, such that the family of charts *cover* S. Then the induced collection $(U_\gamma \cap S, \phi_{\gamma\, S})_{\gamma \in \Gamma}$ is an *atlas* for S, and therefore S is itself a manifold. If $\dim N = n$ and S is locally defined by the vanishing of $n - k$ coordinates, then $\dim S = k$.

Proof. Consider two intersecting adapted charts (U, ϕ) and (V, ψ) of N contained in the above-mentioned collection. Let $p \in (U \cap V) \cap S$ be any point, then if $\phi(p) = (x_1, \dots, x_k, 0, \dots, 0)$ and $\psi(p) = (y_1, \dots, y_k, 0, \dots, 0)$, we obtain

$$\psi_S \phi_S^{-1}(x_1, \dots, x_k) = \psi_S(p) = (y_1, \dots, y_k).$$

Since the natural projections $\pi_j(\psi_S \phi_S^{-1}) : (U \cap V) \cap S \to \mathbf{R}$ are C^∞-morphisms, then one concludes that the transition map $\psi_S \phi_S$ is itself a C^∞-morphism, showing that the charts are compatible. Therefore the induced family of charts $(U_\gamma \cap S, \phi_{\gamma\, S})_{\gamma \in \Gamma}$ is a smooth atlas for S. It is immediate that S is a k-dimensional smooth manifold. ♮

Fibre of C^∞-Morphisms

Definition 21.5.5 (Fibres & levels)**.** Given a C^∞-morphism $f : N \to M$, we say that a point $q \in M$ is the *level* of the fibre $f^{-1}q$—in particular, if q is a regular value, then we say that $f^{-1}q$ is a *regular fibre*.

[2] This is the common description when authors call embedded submanifolds by "regular submanifolds"

Lemma 21.5.6. Let $g\colon N \to \mathbf{R}$ be a C^∞-morphism, and $q \in \mathbf{R}$ be any regular value. If $f\colon N \to \mathbf{R}$ is defined by $f := g - q$, then 0 is a regular value of f and $g^{-1}q = f^{-1}0$. That is, one can construct a smooth map such that its *regular zero fibre* equals to a given regular fibre of g.

Proof. From construction, it is obvious that $g^{-1}q = f^{-1}0$. To show that 0 is a regular value of f, notice that for all $p \in N$ one has $f_{*p} = g_{*p}$—therefore f and g have the exact same critical points, and thus the same regular values. It follows that $f^{-1}0$ is indeed a regular fibre of f. ♮

Lemma 21.5.7. Let $g\colon N \to \mathbf{R}$ be a C^∞-morphism, and let $q \in \operatorname{im} g$ be a regular value. Then the non-empty regular fibre $g^{-1}q$ is a *regular submanifold of* N with codimension 1.

Proof. Assume that $\dim N = n$. As in Lemma 21.5.6, define $f\colon N \to \mathbf{R}$ to be the C^∞-morphism given by $f := g - q$, so that $g^{-1}(q) = f^{-1}(0)$ and 0 is a regular value of f. Given any $p \in f^{-1}(0)$, by definition we have that p is a regular point of f and by Proposition 21.4.15 we know that any chart (U, ϕ) about p has a *non-vanishing* partial derivative $\frac{\partial f}{\partial \phi_j}(p) \neq 0$ for some index j. By permutation of the indices, one may assume without much disturbance that $\frac{\partial f}{\partial \phi_1}(p) \neq 0$ is a non-vanishing partial derivative. Consider the C^∞-morphism $\Phi\colon U \to \mathbf{R}^n$ defined by $\Phi := (f, \phi_2, \dots, \phi_n)$, whose jacobian matrix equals

$$\operatorname{Jac}\Phi = \begin{bmatrix} \frac{\partial f}{\partial \phi_1} & \frac{\partial f}{\partial \phi_2} & \cdots & \frac{\partial f}{\partial \phi_{n-1}} & \frac{\partial f}{\partial \phi_n} \\ \frac{\partial \phi_2}{\partial \phi_1} & \frac{\partial \phi_2}{\partial \phi_2} & \cdots & \frac{\partial \phi_2}{\partial \phi_{n-1}} & \frac{\partial \phi_2}{\partial \phi_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial \phi_n}{\partial \phi_1} & \frac{\partial \phi_n}{\partial \phi_2} & \cdots & \frac{\partial \phi_n}{\partial \phi_{n-1}} & \frac{\partial \phi_n}{\partial \phi_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial \phi_1} & \frac{\partial f}{\partial \phi_2} & \cdots & \frac{\partial f}{\partial \phi_{n-1}} & \frac{\partial f}{\partial \phi_n} \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n \end{bmatrix}$$

Which has determinant $\det(\operatorname{Jac}\Phi(p)) = \frac{\partial f}{\partial \phi_1}(p) \neq 0$ at the point p. From the inverse map theorem (see Corollary 21.1.36) there must exist a neighbourhood $X \subseteq M$ of p such that $(X, \Phi\colon X \to \Phi(X))$ is a chart about p for M. Notice that by construction one has $\Phi|_{X \cap f^{-1}(0)} = (0, \Phi_2, \dots, \Phi_n)$, which shows that (X, Φ) is an adapted chart of M relative to $f^{-1}(0) = g^{-1}(q)$. This shows that every point of $g^{-1}(q)$ has a corresponding adapted chart in M, ensuring that $g^{-1}(q)$ is a regular submanifold with $\operatorname{codim} g^{-1}(q) = 1$ ♮

Theorem 21.5.8 (Regular fibres are embedded submanifolds)**.** Let $f\colon N \to M$ be a C^∞-morphism, and let $\dim N = n$ while $\dim M = m$. Given a regular value $q \in \operatorname{im} f$, the non-empty regular fibre $f^{-1}q$ is a embedded submanifold of N with $\dim f^{-1}q = n - m$.

Proof. We could simply use Lemma 21.5.6, but we shall construct the proof from the ground level. Let (V, ψ) be a chart of M about q satisfying $\psi q = 0 \in \mathbf{R}^m$ (that is, (V, ψ) is centred at q)—so that $f^{-1}V$ is a neighbourhood of $f^{-1}q$ in N, and one has

$$f^{-1}V \cap f^{-1}q = f^{-1}q = (\psi f|_{f^{-1}V})^{-1}0.$$

Therefore $f^{-1}q$ is the zero fibre of the map ψf, which we can analyse locally as the zero fibres of the collection of local projections $(f_j = \pi_j(\psi f|_{f^{-1}V}))_{j=1}^m$. Since $q \in \operatorname{im} f$ is a regular value, every point $p \in f^{-1}q$ is regular and thus has a surjective f_{*p}—that is, $n \geqslant m$ and $\operatorname{rank} f_{*p} = m$. Let (U, ϕ) be a chart about p in N where $U \subseteq f^{-1}V$. Consider the local Jacobian representation of f_{*p} at (U, ϕ), and assume that the first $m \times m$ block of $\left[\frac{\partial f_i}{\partial \phi_j} p\right]$ is invertible—if that is not the case, simply rearrange the indices so that this is true.

Define a C^∞-morphism $\phi' \colon U \to \mathbf{R}^n$ by making $\phi'_j := f_j = \pi_j(\psi f|_{f^{-1}V})$ for $1 \leqslant j \leqslant m$, while $\phi'_j := \phi_j$ for the remaining $m < j \leqslant n$. To show that ϕ' is locally invertible at p, notice that the local Jacobian matrix of f about ϕ' at p is

$$\operatorname{Jac}_{\phi'} f|_p = \begin{bmatrix} \frac{\partial f_i}{\partial \phi_j} p & \frac{\partial f_i}{\partial \phi_\beta} p \\ \frac{\partial \phi_\alpha}{\partial \phi_j} p & \frac{\partial \phi_\alpha}{\partial \phi_\beta} p \end{bmatrix} = \begin{bmatrix} \frac{\partial f_i}{\partial \phi_j} p & \frac{\partial f_i}{\partial \phi_\beta} p \\ 0 & \mathrm{id} \end{bmatrix}$$

where $1 \leqslant i, j \leqslant m$ and $m < \alpha, \beta \leqslant n$. This shows that

$$\det(\operatorname{Jac}_{\phi'} f|_p) = \det \left[\frac{\partial f_i}{\partial \phi_j} p\right]_{1 \leqslant i,j \leqslant m} \neq 0,$$

that is, ϕ' is locally invertible at p and by Corollary 21.1.36 there exists $U_p \subseteq U$ such that $\phi' \colon U_p \to \phi'(U_p)$ is a C^∞-isomorphism, and thus (U, ϕ') is a chart about p in N. Notice that, by construction, $f^{-1}q$ is the set obtained by imposing $\phi'_j = f_j = 0$ for each $1 \leqslant j \leqslant m$—thus (U_p, ϕ') is an *adapted chart* for N relative to $f^{-1}q$. Therefore $f^{-1}q$ admits an adapted chart for each of its points, showing that it is a embedded submanifold of dimension $n - m$. ♮

Corollary 21.5.9 (Producing adapted charts)**.** Let $f \colon N \to \mathbf{R}^m$ be a C^∞-morphism and $\dim N = n$, consider the zero fibre $f^{-1}0$. If there exists a chart (U, ϕ) about $p \in f^{-1}0$ such that

$$\det \left[\frac{\partial f_i}{\partial \phi_j} p\right]_{1 \leqslant i,j \leqslant m} \neq 0,$$

then there exists a neighbourhood $U_p \subseteq N$ of p such that the pair (U_p, ϕ'), where

$$\phi' := (f_1, \ldots, f_m, \phi_{m+1}, \ldots, \phi_n) \colon U_p \longrightarrow \phi' U_p$$

is a C^∞-isomorphism, forms an *adapted chart* for N relative to $f^{-1}0$.

Lemma 21.5.10. Let M be a manifold and $S \subseteq M$ be a subset. Then S is a embedded submanifold of M with codimension k if and only if every point of S has a neighbourhood $U \subseteq M$ such that $U \cap S$ is the regular zero fibre of a C^∞-isomorphism $\phi \colon U \to \mathbf{R}^k$.

Proof. Let $\dim M := m$. ($\Rightarrow$) Assume S is a regular submanifold. For any $p \in S$, take (U, ϕ) to be an adapted chart of M with respect to S, so that

$$\phi|_S = (\phi_1, \ldots, \phi_{m-k}, 0, \ldots, 0).$$

Define $\widehat{\phi}\colon U \to \mathbf{R}^k$ to be the C^∞-isomorphism given by $\widehat{\phi} := (\phi_{m-k+1}, \ldots, \phi_m)$, so that one has

$$S \cap U = \widehat{\phi}^{-1}0.$$

($\Leftarrow$) For the converse, suppose that S admits a pair $(U, \widehat{\phi}\colon U \to \mathbf{R}^k)$ about each point $p \in S$ such that $U \cap S$ is the regular zero fibre of the C^∞-isomorphism $\widehat{\phi}$. Therefore, any C^∞-isomorphism $\phi\colon U \to \mathbf{R}^m$ such that

$$\begin{tikzcd}
U \arrow[rr, "\widehat{\phi}"] \arrow[dr, "\phi"'] & & \mathbf{R}^k \\
& \mathbf{R}^m \arrow[ur, two heads] &
\end{tikzcd}$$

commutes, forms an adapted chart (U, ϕ) relative to S about p. This shows that S is a relative submanifold of M. ♮

21.6 Rank of C^∞-Morphisms

Theorem 21.6.1 (Constant rank, then the fibre is a embedded submanifold)**.** Let $f\colon N \to M$ be a C^∞-morphism and $q \in M$ be any point. If f has has a locally constant rank k at $f^{-1}q$, then the fibre $f^{-1}q$ is a embedded submanifold of N of codimension k.

Proof. Take any $p \in f^{-1}q$ and, by the constant rank theorem, let (U, ϕ) be a chart centred at p and (V, ψ) be a chart centred at $fp = q$ such that

$$\psi f \phi^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_k, 0, \ldots, 0) \in \mathbf{R}^m$$

The zero fibre $(\psi f \phi^{-1})^{-1}0$ is therefore the common zero fibre of the coordinates $\bigcap_{j=1}^k (\psi f \phi^{-1})_j^{-1}0$. Notice that

$$\phi(f^{-1}q) = \phi(f^{-1}(\psi^{-1}0)) = (\psi f \phi^{-1})^{-1}0,$$

therefore $U \cap f^{-1}q = \bigcap_{j=1}^k \phi_j^{-1}0$, proving that (U, ϕ) is an adapted chart of N about p with respect to $f^{-1}q$. ♮

Theorem 21.6.2 (Constant rank)**.** Let N and M be manifolds of dimensions n and m, respectively. Given a C^∞-morphism $f\colon N \to M$ with local constant rank k in a neighbourhood of a $p \in N$, there are charts (U, ϕ) centred at p and (V, ϕ) centred at fp such that

$$\pi_j(\psi f \phi^{-1})|_{\phi U} = 0$$

for all $k < j \leqslant n$.

Proof. Let (U', ϕ') be a chart about p, and (V', ψ') be a chart about fp, and consider the C^∞-morphism

$$\psi' f \phi'^{-1}\colon \phi'(f^{-1}V \cap U) \longrightarrow \mathbf{R}^m.$$

Since both ϕ' and ψ' are C^∞-isomorphism, it follows that $\psi' f \phi'^{-1}$ has a *local* constant rank k at $\phi' p$. By Theorem A.5.11 we conclude that there exists a C^∞-isomorphism ϕ'' of a neighbourhood of $\phi' p \in \mathbf{R}^n$ and a C^∞-isomorphism ψ'' of a neighbourhood of $\psi f p \in \mathbf{R}^m$ for which

$$\psi''(\psi' f \phi'^{-1})\phi''^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_k, 0, \ldots, 0).$$

Therefore by letting $\phi := \phi''\phi'$ and $\psi := \psi''\psi'$ we obtain the desired result. ♮

The following theorems are mere corollaries of the constant rank theorem: given a smooth map $f\colon M \to N$, if f is an immersion at a point p, then it has a locally constant rank m near p, on the other hand, if f is a submersion at a point q, it has a locally constant rank n near q.

Theorem 21.6.3 (Immersion)**.** Let $\theta\colon M \to N$ be a C^∞-morphism, and $\dim M = m$ while $\dim N = n$. If θ is an immersion at $p \in M$, then there are charts ϕ about p and ψ about θp such that the diagram

$$\begin{tikzcd} M \arrow[r, "\theta"] & N \arrow[d, "\psi"] \\ \mathbf{R}^m \arrow[u, "\phi^{-1}"] \arrow[r, hook] & \mathbf{R}^n \end{tikzcd}$$

commutes in `Man`. That is, θ locally acts as an inclusion about p.

Theorem 21.6.4. Let $\theta\colon M \to N$ be a C^∞-morphism. If θ is a submersion at some point $p \in M$, there are charts ϕ about p and ψ about θp for which

$$\begin{tikzcd} M \arrow[r, "\theta"] & N \arrow[d, "\psi"] \\ \mathbf{R}^m \arrow[u, "\phi^{-1}"] \arrow[r, two heads] & \mathbf{R}^n \end{tikzcd}$$

commutes in `Man`. That is, θ locally acts as a projection about p.

Corollary 21.6.5. A smooth submersion is an *open map*.

Proof. Let $f\colon M \to N$ be a smooth submersion and $U \subseteq M$ be an open subset. Let $fp \in fU$ be any point for which $p \in U$. Applying the submersion theorem to f, we can find a neighbourhood $V \subseteq U$ of p for which $f|_V$ acts as a projection, which implies that $fV \subseteq fU$ is a neighbourhood for p—since projections are open morphisms. This proves that fU is an open set. ♮

Theorem 21.6.6. Let N be an n-manifold, and M be an m-manifold. Then:

(a) If $f\colon N \to M$ is a smooth embedding, then the image fN is an embedded submanifold of M in the sense of Definition 21.5.2.

(b) If N is an embedded submanifold of M in the sense of Definition 21.5.2, then the canonical inclusion $\iota\colon N \hookrightarrow M$ is a smooth embedding—thus N is an embedded submanifold in the sense of Definition 21.5.1.

This shows that the two definitions for embedded submanifolds are equivalent.

Proof. (a) By the immersion theorem, let (U, ϕ) be a chart centred at p and (V, ψ) be a chart centred at fp such that

$$\begin{tikzcd} U \arrow[r, "f"] & V \\ \phi U \arrow[u, "\phi^{-1}"] \arrow[r, hook] & \psi V \arrow[u, "\psi^{-1}"'] \end{tikzcd}$$

commutes—where $\phi U \subseteq \mathbf{R}^n$ and $\phi V \subseteq \mathbf{R}^m$. Then fU is simply the common zero fibres of the last $m-n$ coordinates of ψ, that is, $fU = \bigcap_{j=n+1}^{m} \psi_j^{-1} 0$. Since f is injective, it follows that fN—endowed with the subspace topology—is isomorphic to N. Therefore fU is open in fN, thus there must exist an open set $V' \subseteq M$ such that $V' \cap fN = fU$ and hence V' is a neighbourhood of fp. Since

$$(V \cap V') \cap fN = V \cap fU = fU,$$

it follows that $(V \cap V', \psi)$ is an adapted chart of M about fp with respect to fN. Therefore fN is an embedded submanifold in the sense of Definition 21.5.2.

(b) Recall that both ιN and N are endowed with the subspace topology. Since ι is injective, the induced map $\iota\colon N \to \iota N$ is a topological isomorphism—therefore ι is a topological embedding. In order to show that ι is an immersion, let $p \in N$ be any point, and take an adapted chart (V, ψ) of M about p with respect to N—that is, $V \cap N = \bigcap_{j=n+1}^{m} \psi_j^{-1} 0$. Define $\psi_N := (\psi_1, \dots, \psi_n)\colon V \cap N \to \psi_N(V \cap N) \subseteq \mathbf{R}^n$, then

$$\begin{tikzcd} V \cap N \arrow[r, hook, "\iota"] & V \\ \psi_N(V \cap N) \arrow[u, "\psi_N^{-1}"] \arrow[r, hook] & \psi V \arrow[u, "\psi^{-1}"'] \end{tikzcd}$$

commutes, showing that ι is an immersion.

♮

21.7 Transversality

Definition 21.7.1 (Transversal C^∞-morphism)**.** Let M be a manifold and $S \subseteq M$ be an embedded submanifold. A C^∞-morphism $f\colon N \to M$ is said to be *transversal* to S, denoted $f \pitchfork S$, if for every $p \in f^{-1}S$ one has

$$f_*(T_p N) + T_{fp} S = T_{fp} M.$$

Moreover, we can equally define this notion for embedded submanifolds: let N and L be embedded submanifolds of M. We say that N *intersects* L *transversely*, denoted $N \pitchfork L$, if for any $p \in N \cap L$ we have

$$T_pN + T_pL = T_pM.$$

Theorem 21.7.2 (Transversality)**.** Let M be a manifold, and $S \subseteq M$ be a embedded submanifold with $\operatorname{codim}_M S = k$. Then every C^∞-morphism $f\colon N \to M$ that is *transversal* to S is such that $f^{-1}S$ is a embedded submanifold of codimension k in N.

Proof. Let $\dim M := m$ and $\dim N := n$. Since S is a regular submanifold of M, given any point $p \in f^{-1}S$, we can find an adapted chart (U, ϕ) of M with respect to S that is centred at $fp \in S$ and such that

$$U \cap S = \{q \in M : \phi_j q = 0 \text{ for each } m - k < j \leqslant m\},$$

that is, $U \cap S$ is the zeros of the last $m - k$ coordinates of ϕ.

Define $g\colon U \to \mathbf{R}^k$ as the C^∞-isomorphism given by those last $m - k$ coordinates:

$$g := (\phi_{m-k+1}, \dots, \phi_m),$$

so that $g|_S = 0$. Since $fp \in U$ then $f^{-1}U \cap f^{-1}S$ is a neighbourhood for p in N, and $p \in (gf|_{f^{-1}U})^{-1}0$ because $fp \in S$, therefore

$$f^{-1}U \cap f^{-1}S = (gf|_{f^{-1}U})^{-1}0.$$

We shall prove that 0 is a regular value of $gf|_{f^{-1}U}$ so that $f^{-1}U \cap f^{-1}S$ is a regular zero fibre of the map. To that end, let $z \in T_0\mathbf{R}^k$ and $x \in (gf|_{f^{-1}U})^{-1}0$ be any two elements. Since g is an isomorphism, 0 is a regular value of g, therefore there exists $y \in T_{fx}M$ for which $g_{*fx}(y) = z$. Using the fact that f is transversal to S, there exists $y_0 \in T_{fx}S$ and $v \in T_xN$ for which $y = y_0 + f_{*x}v$. Recall that $g|_S = 0$ from construction, therefore $g_{*fx}(y_0) = 0$. Hence one has

$$z = g_{*fx}y = g_{*fx}(y_0 + f_{*x}v) = g_{*fx}f_{*x}v = (gf)_{*x}v,$$

showing that $(gf)_{*x}$ is surjective—hence x is a regular point of $gf|_{f^{-1}U}$. This shows that 0 is a regular value of $gf|_{f^{-1}U}$, therefore the subset $f^{-1}U$ is a neihbourhood of p satisfying the desired condition specified in Lemma 21.5.10—moreover, since p was any point of $f^{-1}S$, the theorem follows by just mentioned lemma. ♮

Theorem 21.7.3. Let M be an m-manifold, and both N and L be submanifolds—whose dimensions are n and ℓ, respectively. If $N \pitchfork L$ in M, then $N \cap L$ is a submanifold of M with dimension $\dim N \cap L = n + \ell - m$.

Proof. Let $p \in N \cap L$ be any point. Let (U, ϕ) and (V, ψ) be adapted charts of M about p with respect to N and L. Let $W := U \cap V$ and define the induced maps

$$\eta := (\phi_{n+1}, \dots, \phi_m)\colon W \longrightarrow \mathbf{R}^{m-n} \quad \text{and} \quad \lambda := (\psi_{\ell+1}, \dots, \psi_m)\colon W \longrightarrow \mathbf{R}^{m-\ell}$$

so that $W \cap N = \eta^{-1}0$ and $W \cap L = \lambda^{-1}0$ are regular fibres of η and λ, respectively.

Consider the product map $\eta \times \lambda \colon W \to \mathbf{R}^{m-n} \times \mathbf{R}^{m-\ell}$—we wish to show that $(\eta \times \lambda)^{-1}0 = W \cap (N \cap L)$ is a regular fibre of $\eta \times \lambda$, so that $(W, \phi \times \psi)$ is an adapted chart of M about p with respect to $N \cap L$. To that end, we shall consider the differential $(\eta \times \lambda)_* \colon T_p M \to \mathbf{R}^{m-n} \times \mathbf{R}^{m-\ell}$, whose kernel is

$$\ker(\eta \times \lambda)_* = \ker \eta_* \cap \ker \lambda_* = T_p N \cap T_p L.$$

On the other hand, we know that

$$\dim \ker(\eta \times \lambda)_* = \dim T_p N + \dim T_p L - \dim T_p M = n + \ell - m,$$

therefore by the rank plus nullity theorem one has

$$\operatorname{rank}(\eta \times \lambda)_* = m - (n + \ell - m) = 2m - n - \ell = \dim(\mathbf{R}^{m-n} \times \mathbf{R}^{m-\ell})$$

which shows that $(\eta \times \lambda)_*$ is an epimorphism at p—thus 0 is a regular value of $\eta \times \lambda$. This concludes the proof that $N \cap L$ is an embedded submanifold of M with dimension $n + \ell - m$.

$\natural$

21.8 Classification of 1-Manifolds

Definition 21.8.1 (Parametrization by arc-length)**.** Let $I \subseteq \mathbf{R}$ be an interval, and M be a smooth manifold. We say that a C^∞-morphism $f \colon I \to M$ is a *parametrization by arc-length* if the restriction $f \colon I \to fI$ is a C^∞-isomorphism, and if the velocity vector $f_{*s}1 \in T_{fs}M$ is *unitary* for each $s \in I$.

Lemma 21.8.2. Consider a pair of arc-length parametrizations $I \xrightarrow{f} M \xleftarrow{g} J$. Then $fI \cap gJ$ has at most two connected components. One has the following properties concerning the number of connected components:

(a) If $fI \cap gJ$ has only *one* connected component, then one can *extend* f to an arc-length parametrization of $fI \cup gJ$.

(b) If $fI \cap gJ$ has *two* components, then $M \simeq S^1$.

Proof. Consider the C^∞-isomorphism

$$g^{-1}f \colon f^{-1}(fI \cap gJ) \xrightarrow{\simeq} g^{-1}(fI \cap gJ),$$

sending open sets of I to open sets of J, and with derivative ± 1 everywhere—from the definition of the parametrization. Let Γ be the pullback of the pair (f, g)—that is, composed of pairs $(s, t) \in I \times J$ such that $fs = gt$. Therefore Γ is a closed subset of $I \times J$, with the product topology, and consists of line segments with slope ± 1 by the behaviour of $g^{-1}f$. Since $g^{-1}f$ is an isomorphism, it must be the case that the line segments don't end abruptly, but extend from edge to edge of $I \times J$. From the fact that

$g^{-1}f$ is injective and constant derivative, it must be the case that Γ is composed of at most two line segments—furthermore, if Γ has two components, then they must have the same slope and the start and end edges has to be distinct. We analyse the number of components of Γ:

- If Γ has a single component, then one can extend $g^{-1}f$ to a line $\ell\colon \mathbf{R} \to \mathbf{R}$. Therefore the map $g\ell\colon I \cup \ell^{-1}J \to fI \cup gJ$ forms an extension of f.
- If Γ has two components, let those be the line segments connecting $(a,\alpha) \to (b,\beta)$ and $(c,\gamma) \to (d,\delta)$—whose points lie in $I \times J$, with ends in the boundary of the square, that is, $I = (a,d)$ and $J = (\gamma,\delta)$ for instance. By merely a translation of J, we may assume that $\gamma = c$ and $\delta = d$, for which we obtain the relations

$$a < b \leqslant c < d \leqslant \alpha < \beta.$$

We may define a continuous maps $\theta\colon [a,\beta] \to \mathbf{R}$ given by $t \mapsto \frac{2\pi t}{\alpha - a}$, and $h\colon S^1 \to M$ mapping

$$h(\cos(\theta t), \sin(\theta t)) := \begin{cases} ft, & \text{if } a < t < d \\ gt, & \text{if } c < t < \beta \end{cases}$$

which is well defined since f and g agree on $[c,d]$ due to the translation of J. Since hS^1 is a compact open set of M, it follows that it must be the case that $hS^1 = M$. Therefore, since the restrictions of f and g are C^∞-isomorphisms, it follows that h is a C^∞-isomorphism.

♮

Theorem 21.8.3 (Classification of 1-manifolds)**.** Any smooth connected 1-manifold M is C^∞-isomorphic to either the *circle* S^1 or to some *interval* of real numbers: $[0,1]$, $[0,1)$, $(0,1]$ or $(0,1)$. In fact, the following is a complete classification list:

(1) If M is compact without boundary, then M is C^∞-isomorphic to a circle.

(2) If M is compact with boundary, then M is C^∞-isomorphic to a closed interval.

(3) If M is non-compact without boundary, then M is C^∞-isomorphic to an open interval.

(4) If M is non-compact with boundary, then M is C^∞-isomorphic to a half-open interval.

Proof. Given a parametrization by arc-length $f'\colon J \to M$, via Lemma 21.8.2 one can extend f' to a *maximal* arc-length parametrization $f\colon I \to M$—so that I is the maximal interval f can be extended to.

Assuming $M \not\cong S^1$, suppose that there exists a limit point x of fI with $x \in M \setminus fI$, so that f is not surjective. Let U be a neigbourhood of x and $g\colon I' \to U$ be an arc-length parametrization of U. From Lemma 21.8.2 we can use g to extend f to a parametrization $fI \cup gJ$, contradicting the hypothesis of maximality of f. ♮

Corollary 21.8.4. The boundary of a compact 1-manifold has an *even* number of points.

Proof. Indeed, if M is a compact 1-manifold with boundary, then it's isomorphic to the disjoint union of a collection of closed intervals, each of which has two boundary points, thus M has an even number of boundary points. ♮

Chapter 22

Cobordism

22.1 Cobordisms

Unoriented Cobordisms

Definition 22.1.1 (Unoriented cobordism)**.** Given a pair Σ_0 and Σ_1 of smooth compact $(n-1)$-manifolds without boundary, we define a *cobordism between* Σ_0 *and* Σ_1 to be a smooth compact n-manifold M whose boundary is $\partial M = \Sigma_0 \amalg \Sigma_1$. We thus call the manifolds Σ_0 and Σ_1 *cobordant*.

Example 22.1.2. Two interesting cobordisms are formed from the empty manifold to the circle, which shall be called *birth-of-a-circle*, and from the circle to the empty manifold, so called *death-of-a-circle*.

Lemma 22.1.3 (Cobordant zero and one dimensional manifolds)**.** Two given compact 0-manifolds without boundary are cobordant if and only if they have the same number of points modulo 2. Moreover, any two compact 1-manifolds without boundary are cobordant.

Proof. Let's consider the case of a pair of 0-manifolds Σ_0 and Σ_1. Notice that since every pair of points can be connected by a smooth curve, and every 1-manifold with boundary has an even number of boundary points[1], it follows that Σ_0 and Σ_1 are cobordant if and only if the disjoint union $\Sigma_0 \amalg \Sigma_1$ has an even number of points.

For the second statement, one should recall that a compact 1-manifold is the disjoint union of circles. Then we can choose one of the manifolds to attach copies of the death-of-a-circle cobordism for each of its circles, and attach birth-of-a-circle cobordisms for each of its respective circles of the other 1-manifold. This construction yields a cobordism between them. ♮

[1] This is due to the fact that 1-manifolds are C^∞-isomorphic to a finite disjoint union of circles or intervals (see Corollary 21.8.4).

Oriented Cobordisms

Consider the following setup: let Σ be a closed submanifold of M with codimension 1, where $\dim M = n$. Assume both manifolds to be oriented.

Definition 22.1.4 (Positive normal)**.** Let $[v_1, \ldots, v_{n-1}]$ be a positive basis for $T_x\Sigma$ for any given point $x \in \Sigma$. We say that a tangent vector $v \in T_xM$ is a *positive normal* if the induced basis $[v_1, \ldots, v_{n-1}, v]$ for T_xM is positive.

Definition 22.1.5 (In and out boundaries)**.** If Σ is a connected component of ∂M, we call Σ an *in-boundary* if a positive normal points inwards relative to M, and otherwise an *out-boundary*—when a positive normal points outward relative to M.

The notion of an in and out boundary allows us to define the notion of an oriented cobordism. From now on, a cobordism will always mean an *oriented* one, unless stated otherwise.

Definition 22.1.6 (Oriented cobordism)**.** Let Σ_{in} and Σ_{out} be compact $(n-1)$-manifolds without boundary. We define an *oriented cobordism* between them to be a triple $(M, \iota_{\text{in}}, \iota_{\text{out}})$, where M is a smooth compact oriented n-manifold, and arrows

$$\Sigma_{\text{in}} \xrightarrow{\iota_{\text{in}}} M \xleftarrow{\iota_{\text{out}}} \Sigma_{\text{out}}$$

which are C^∞-isomorphisms when restricted to the in and out boundary of M, respectively. We shall denote the oriented cobordism M as an arrow $M\colon \Sigma_{\text{in}} \Rightarrow \Sigma_{\text{out}}$.

Definition 22.1.7 (Equivalence of cobordisms)**.** Given two cobordisms

$$\begin{tikzcd} \Sigma_{\text{in}} \arrow[r, Rightarrow, shift left, "M"] \arrow[r, Rightarrow, shift right, "N"'] & \Sigma_{\text{out}} \end{tikzcd}$$

we say that M is *equivalent* to the cobordism N if there exists an orientation-preserving C^∞-isomorphism $\phi\colon M \xrightarrow{\simeq} N$ such that the following diagram commutes in `Man`:

$$\begin{tikzcd} & N & \\ \Sigma_{\text{in}} \arrow[ur] \arrow[dr] & & \Sigma_{\text{out}} \arrow[ul] \arrow[dl] \\ & M \arrow[uu, "\phi", "\simeq"'] & \end{tikzcd}$$

22.2 Elements of Morse Theory

Definition 22.2.1. Let $f\colon M \to I$ be a C^∞-morphism, and $p \in M$ be a critical point of f. We call p a *non-degenerate* point if there exists a chart about p for which the local Hessian of f is invertible. Furthermore, define the *index of f at p* to be the number of *negative eigenvalues* of the local Hessian.

Definition 22.2.2 (Morse maps)**.** Given a smooth manifold M, we say that a C^∞-morphism $f\colon M \to I$ is a *Morse map* if every critical point of f is non-degenerate. If it happens to be the case that M is a manifold with boundary, we shall require that $f^{-1}\partial I = \partial M$ and that the boundary points $\partial I = \{0, 1\}$ are regular values of f—this ensures that ∂M contains no critical points.

The existence of Morse maps is ensured by the following theorem:

Theorem 22.2.3. For any manifold M and integer $2 \leqslant r \leqslant \infty$, the collection of Morse maps $M \to I$ is dense in $C^r(M, I)$.

The following is a generalization of the construction of attaching spaces:

Definition 22.2.4 (Gluing)**.** Let $f\colon X \to Y$ and $g\colon X \to Z$ be topological morphisms. We define the *gluing of Y and Z along X* to be the pushout

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "g"'] & Y \arrow[d] \\ Z \arrow[r] & Y \amalg_X Z \arrow[ul, phantom, "\ulcorner", very near start] \end{tikzcd}$$

Explicitly, $Y \amalg_X Z$ is the quotient space of $Y \amalg Z$ where $y \sim z$ if and only if there exists a common $x \in X$ such that $fx = y$ and $gx = z$.

Part VII

Analysis

Chapter 23

Normed Vector Spaces

Remark 23.0.1. Throughout this whole chapter we shall denote by k a field that is either $\mathbf{C}$ or $\mathbf{R}$ — we'll adopt the use of $|-|\colon k \to \mathbf{R}_{\geqslant 0}$ for the standard norm of the underlying field in order to distinguish that from the norm of the vector spaces.

23.1 Norms on Spaces

Definition 23.1.1 ((Pre)Norm and (pre)normed vector spaces)**.** Let E be a k-vector space. We say that a map $\|-\|\colon E \to \mathbf{R}_{\geqslant 0}$ is a *pre-norm* in E if for all $x, y \in E$ and $\lambda \in k$ the map satisfies the following properties

(a) Product by scalar: $\|\lambda x\| = |\lambda|\,\|x\|$.

(b) Triangle inequality: $\|x + y\| \leqslant \|x\| + \|y\|$.

Moreover, if $\|-\|$ satisfies the following additional condition, it is called a *norm*.

(c) If $\|x\| = 0$ then $x = 0$.

The vector space E endowed with the (pre)norm $\|-\|$ is called a *(pre)normed vector space*.

It should be noted immediately that the first condition for a pre-norm implies in $\|0\| = 0$, hence the last condition for $\|-\|$ to be a norm can be substituted equivalently by "$\|x\| = 0$ if and only if $x = 0$". Moreover, in a prenormed space, it's not possible to assert that a sequence has a unique limit, property which is only ensured by the last condition.

A simple property that can be extracted from Definition 23.1.1, is that any subspace of a (pre)normed k-vector space is itself (pre)normed with the naturally inherited norm.

Example 23.1.2 (Norms from maps)**.** Let E be a k-vector space and let's consider any functional $f \in E^*$. We can build a *pre-norm* from the linear map f by defining a map $\|-\|_f\colon E \to \mathbf{R}_{\geqslant 0}$ given by $\|x\|_f := |f(x)|$ — which clearly satisfies all of the required conditions for a norm. Note however that for $\|-\|_f$ to satisfy the last condition, it must be true that $\ker f = 0$, that is, $\dim_k E = 1$ necessarily for $\|-\|_f$ to be a *norm*.

Example 23.1.3 (Metrics from norms)**.** Given a (pre)normed k-vector space $(E, \|-\|)$, we can naturally define a metric in E to be a map $d\colon E\times E \to \mathbf{R}_{\geqslant 0}$ given by $d(x, y) := \|x-y\|$ — which makes E into a (pre)metric space (refer to Definition 12.3.1). We now verify each of the conditions for a pre-metric:

(a) If $x, y \in E$ are any elements, then

$$d(x, y) = \|x - y\| = |-1|\,\|x - y\| = \|y - x\| = d(y, x).$$

(b) If $z \in E$ is another element, then

$$d(x, z) = \|x - z\| \leqslant \|x - y\| + \|y - z\| = d(x, y) + d(y, z).$$

Moreover, if $\|-\|$ is a norm, then we are also able to satisfy:

(c) Given $x, y \in E$ such that $d(x, y) = 0$, then $\|x - y\| = 0$ and since $\|-\|$ is a norm, we obtain $x - y = 0$, thus $x = y$ as wanted.

Proposition 23.1.4. Let $(E, \|-\|)$ be a (pre)normed k-vector space, and $d\colon E \times E \to \mathbf{R}_{\geqslant 0}$ be the (pre)metric induced by $\|-\|$. Then, E is a *Hausdorff* space if and only if $\|-\|$ is a *norm*.

Proof. Suppose E is a Hausdorff space and let $x, y \in E$ be such that $\|x - y\| = 0$, then in particular $d(x, y) = 0$ and therefore any open ball $B_x(r)$ centred at x, for $r > 0$, does also contain y — but since E is Hausdorff, this can only be true if $x = y$, thus $x - y = 0$ and we obtain that $\|-\|$ is a norm, and d is a metric.

Now, assume that $\|-\|$ is a norm, thus d is a metric and since any metric space is Hausdorff (see Proposition 12.4.9), in particular, E is Hausdorff. ♮

Lemma 23.1.5. Let $(E, \|-\|)$ be a normed k-vector space, then the following is true for any given $x, y, z \in E$:

(a) If $d\colon E \times E \to \mathbf{R}_{\geqslant 0}$ is the metric induced by $\|-\|$, then $d(x + y, y + z) = d(x, z)$.

(b) The inequality $\big|\|x\| - \|y\|\big| \leqslant \|x - y\|$ holds.

Proof. (a) $d(x + y, y + z) = \|(x + y) - (y + z)\| = \|x - z\| = d(x, z)$.

(b) Notice that $\|x\| = \|(x - y) + y\| \leqslant \|x - y\| + \|y\|$, therefore $\|x\| - \|y\| \leqslant \|x - y\|$. Moreover, symmetrically we have that $\|y\| - \|x\| \leqslant \|y - x\| = \|x - y\|$, therefore,

$$-\|x - y\| \leqslant \|x\| - \|y\| \leqslant \|x - y\|$$

as wanted.

♮

Proposition 23.1.6. Any norm in an $\mathbf{R}$-vector space is *continuous*.

Proof. This follows directly from the inequality obtained in item (b) of Lemma 23.1.5. ♮

Proposition 23.1.7. Given a (pre)normed k-vector space E, if we endow the products $E \times E$ and $k \times E$ with the product topology, *addition* of vectors and *multiplication* of vectors by scalars are both *continuous* maps.

Proof. Given sequences $x_j \to x$ and $y_j \to y$ for any $x, y \in E$, then $x_j + y_j \to x + y$, let $\varepsilon > 0$ be any bound, then choose a common $n \in \mathbf{N}$ for which $\|x - x_j\| < \frac{\varepsilon}{2}$ and $\|y - y_j\| < \frac{\varepsilon}{2}$ for all $j \geqslant n$. Therefore, by the triangle inequality

$$\|(x + y) - (x_j - y_j)\| \leqslant \|x - x_j\| - \|y - y_j\| < \varepsilon \text{ for all } j > n,$$

where we thus conclude that $x_j + y_j \to x + y$ — all limits are well defined since E is Hausdorff.

Let $\lambda_j \to \lambda$ be a convergent sequence in k for any given $\lambda \in k$. Since convergent sequences are bounded, let $M \in k$ be a bound for $(x_j)_{j\in\mathbf{N}}$, that is, $|x_j| < M$ — also, define $M' := \max(M, |\lambda|)$. Given any $\varepsilon > 0$ we choose a common $n \in \mathbf{N}$ for which $|\lambda - \lambda_j| < \frac{\varepsilon}{2M'}$ and $\|x - x_j\| < \frac{\varepsilon}{2M'}$ for all $j \geqslant n$. With this we can use the triangle inequality to find

$$\begin{aligned}
\|\lambda x - \lambda_j x_j\| &= \|(\lambda - \lambda_j)x_j + \lambda(x - x_j)\| \\
&\leqslant \|(\lambda - \lambda_j)x_j\| + \|\lambda(x - x_j)\| \\
&= |\lambda - \lambda_j|\,\|x_j\| + |\lambda|\,\|x - x_j\| \\
&= \frac{\varepsilon}{2M'}\|x_j\| + |\lambda|\frac{\varepsilon}{2M'} \\
&< \frac{\varepsilon}{2M'}M' + M'\frac{\varepsilon}{2M'} \\
&= \varepsilon,
\end{aligned}$$

that is, $\lambda_j x_j \to \lambda x$ as wanted. ♮

Definition 23.1.8 (Equivalence of norms)**.** Given vector space E and two norms, $\| - \|_1$ and $\| - \|_2$, on E, we say that such norms are equivalent if there exists $a, b > 0$ such that for all $v \in E$ we have the inequalities

$$a\|v\|_1 \leqslant \|v\|_2 \leqslant b\|v\|_1.$$

Proposition 23.1.9. Two norms are equivalent if and only if they induce the same topology.

Proof. Let E be a k-vector space and both $\| - \|_1$ and $\| - \|_2$ be norms on E. Let $U \subseteq (E, \| - \|_1)$ be open and, for any given $x \in U$, let $B_x(r) \subseteq (E, \| - \|_1)$ be an open ball centred at x. Notice that if $b > 0$ is the right constant in Definition 23.1.8, then by choosing the open ball $B_x(r/b) \subseteq (E, \|-\|_2)$, we are able to obtain that for all $y \in B_x(r/b)$ we have

$$\|x - y\|_1 \leqslant b\|x - y\|_2 < r$$

thus $y \in B_x(r)$ in $\| - \|_1$ — this implies in $B_x(cr) \subseteq B_x(r)$ in $\| - \|_2$ and therefore we conclude that $B_x(r)$ is open in $(E, \| - \|_2)$. Notice that the proof can be mirrored for the other case, thus we are done. ♮

Quotient Space

Example 23.1.10 (Pre-norm on quotients of normed spaces)**.** Let $(E, \|-\|)$ be a *pre-normed* k-vector space and $F \subseteq E$ a subspace. We can define a *pre-norm* $\| - \|_{\sim} \colon E/F \to \mathbf{R}_{\geqslant 0}$ on the quotient space E/F as

$$\|[v]\|_{\sim} := \inf_{u \in [v]} \|u\|,$$

that is, the infimum of the norm of the representatives of the class. This indeed defines a pre-norm since, if $[v], [w] \in E/F$, then

$$\begin{aligned} \|[v] + [w]\|_{\sim} &= \inf\{\|x + y\| : x + y \in [v] + [w]\} \\ &\leqslant \inf\{\|x\| + \|y\| : x \in [v] \text{ and } y \in [w]\} \end{aligned} \tag{23.1}$$

We claim that $\inf\{\|x\| + \|y\|\} \leqslant \inf \|x\| + \inf \|y\|$. To prove that, assume that both $\inf \|x\|$ and $\inf \|y\|$ are finite — otherwise, if one of them is infinite, the inequality is trivially true. For any $\varepsilon > 0$ we can find $x' \in [v]$ and $y' \in [w]$ for which

$$\begin{aligned} \inf_{x \in [v]} \|x\| \leqslant \|x'\| \leqslant \inf_{x \in [v]} \|x\| + \varepsilon, \\ \inf_{y \in [v]} \|y\| \leqslant \|y'\| \leqslant \inf_{y \in [v]} \|y\| + \varepsilon. \end{aligned}$$

Therefore we find that $\inf\{\|x\|+\|y\|\} \leqslant \|x'\|+\|y'\|$ therefore $\inf\{\|x\|+\|y\|\} \leqslant \inf \|x\| + \inf \|y\| + 2\varepsilon$, but since ε may be indefinitely little, we find that indeed $\inf\{\|x\| + \|y\|\} \leqslant \inf \|x\| + \inf \|y\|$. Hence, we can turn to Eq. (23.1) and conclude that

$$\begin{aligned} \|[v] + [w]\|_{\sim} &\leqslant \inf\{\|x\| + \|y\| : x \in [v] \text{ and } y \in [w]\} \\ &\leqslant \inf_{x \in [v]} \|x\| + \inf_{y \in [w]} \|y\| \\ &= \|[v]\|_{\sim} + \|[w]\|_{\sim}, \end{aligned}$$

thus satisfying the triangle inequality. The condition that $\|\lambda v\|_{\sim} = |\lambda| \, \|v\|_{\sim}$ is trivially obtained — hence $\| - \|_{\sim}$ is a pre-norm in E/F.

Proposition 23.1.11 (Quotient norm)**.** The pre-norm $\|-\|_{\sim}$ described in Example 23.1.10 is a *norm* if and only if the subspace F is closed in E.

Proof. Let $[v] \in E/F$ be any class such that $\|[v]\|_{\sim} = 0$, therefore, we must be able to find a sequence $(v_j)_{j \in \mathbf{N}}$ of elements $v_j \in [v]$ such that $\|v_j\| \to 0$ as $j \to \infty$. Moreover, choosing $v \in [v]$ to be any representative, since $v_j - v \in F$ and F is closed normed space, the convergence of the norm to zero implies that $v_j \to 0$ for $j \to \infty$ in F — therefore, we conclude that $F \ni v_j - v \to -v$ and thus $-v \in F$ from the closeness property, in particular, we find that $v \in F$.

We now claim that since $v \in F$, then $[v] = [0]$. Let $u \in [v]$ be any representative, then $u - v \in F$ and since $v \in F$ by assumption, then in particular $u \in F$, hence $u \in [0]$ — that is, $[v] \subseteq [0]$. Now, if $w \in [0]$, from definition we obtain $w \in F$, but since $v \in F$ then in particular $w - v \in F$ hence $w \in [v]$ — which implies that $[0] \subseteq [v]$ and thus $[v] = [0]$ for all $v \in F$. Therefore, from this claim we obtain that $\|[v]\|_{\sim} = 0$ implies $[v] = [0] = F$, the zero element of E/F. ♮

Proposition 23.1.12 (Norm out of pre-norm)**.** Let $(E, \|-\|)$ be a *pre-normed* k-vector space and $E_0 := \{x \in E : \|x\| = 0\}$. Then, the map $\|-\|_\sim \colon E/E_0 \to \mathbf{R}_{\geqslant 0}$ defined by $\|[x]\|_\sim := \|x\|$ is well defined and is a *norm* for the space E/E_0.

Proof. Let $[x] \in E/E_0$ be a class such that $\|[x]\|_\sim = 0$. Choose any representative $x \in [x]$ and notice that since $\|x\| = 0$ then $x \in E_0$, that is, $[x] \subseteq E_0$ — moreover, if $y \in E_0$, then surely $y \in [x]$, that is, $E_0 \subseteq [x]$. This shows that $[x] = E_0$, where $E_0 = [0] \in E/E_0$. ♮

Examples of Normed Spaces

The following is an immediate proposition, so we won't bother to write down the proof.

Lemma 23.1.13 (Complex conjugate space norm)**.** Let $(E, \|-\|)$ be a pre-normed $\mathbf{C}$-vector space. Then the naturally induced map $\|-\| \colon \overline{E} \to \mathbf{R}_{\geqslant 0}$ is a pre-norm for the complex conjugate space $\overline{E}$.

Proposition 23.1.14. Let E be a pre-normed k-vector space, $F \subseteq E$ be a closed subspace, and $x \in E \setminus F$. Then there exists a scalar $C > 0$ for which, for every given scalar $\lambda \in k$ and vector $y \in F$, we have
$$|\lambda| \leqslant C\|\lambda x + y\|.$$

Proof. If $\lambda = 0$ then the proposition follows trivially. Otherwise, let $\lambda \neq 0$ and notice that since F is closed, there must exist $\theta > 0$ for which $\|x - y\| \geqslant \theta$ for every given $y \in F$ — since x lies outside of F. In particular, since $-\frac{1}{\lambda} y \in F$, then $\|x - (-\frac{1}{\lambda} y)\| \geqslant \theta$ — notice that such choice of vector was not made arbitrarily since
$$\|\lambda x + y\| = |\lambda| \, \|x - (-y/\lambda)\| \geqslant |\lambda|\theta,$$
therefore, if we choose $C := \frac{1}{\theta}$, we obtain the desired inequality. ♮

Example 23.1.15 (p-norms)**.** The following are recurrent norms on two of the most relevant spaces to our analytical study of normed k-vector spaces:

(a) For every integer $1 \leqslant p < \infty$ we define a norm $\|-\|_p \colon k^n \to \mathbf{R}_{\geqslant 0}$ defined by, for all $x \in k^n$,
$$\|x\|_p := \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p}.$$

(b) The infinite case for the space k^n is defined by a map $\|-\|_\infty \colon k^n \to \mathbf{R}_{\geqslant 0}$ given by
$$\|x\|_\infty := \max_{1 \leqslant j \leqslant n} |x_j|.$$

(c) Let $\ell^p(J) \subseteq k^J$ be the k-vector subspace consisting of maps $f \colon J \to k$ with *countable support* — that is, $\sum_{t \in J} |f(t)|^p < \infty$. For each integer $1 \leqslant p < \infty$, we define a norm $\|-\|_p \colon \ell^p(J) \to \mathbf{R}_{\geqslant 0}$ given by
$$\|f\|_p := \left(\sum_{t \in J} |f(t)|^p \right)^{1/p}.$$

(d) For the infinite case, we define $\ell^\infty(J) \subseteq k^J$ to be the k-vector subspace consisting of maps $f\colon J \to k$ such that $\sup_{t\in J} |f(t)| < \infty$. We define the norm $\|-\|_\infty\colon \ell^\infty(J) \to \mathbf{R}_{\geqslant 0}$ by

$$\|f\|_\infty := \sup_{t\in J} |f(t)|.$$

In fact, each one of the above spaces is Banach, but we'll prove this later. We need now to prove that these are indeed normed vector spaces.

Proof. (a) Let $x \in k^n$ be any vector and $\lambda \in k$ any scalar, then

$$\|\lambda x\|_p = \left(\sum_j |\lambda x_j|^p\right)^{1/p} = \left(|\lambda|^p \sum_j |x_j|^p\right)^{1/p} = |\lambda|\, \|x\|_p.$$

Moreover, if $y \in k^n$ is another vector, then since $p \geqslant 1$ we may just use Minkowski's inequalities (see Proposition A.1.9) in order to obtain the triangle inequality — it should be noted that $p < 1$ does not yield a valid triangle inequality, hence justifying the restriction $p \geqslant 1$. Also, if $\|z\| = 0$ for some $z \in k^n$, then $z_j = 0$ for all $1 \leqslant j \leqslant n$ — since $|z_j| = 0$ — and thus $z = 0$.

(b) The map $\| - \|_\infty\colon k^n \to \mathbf{R}_{\geqslant 0}$ is clearly a norm.

(c)

(d) Let $\lambda \in k$ and $f, g \in \ell^\infty(J)$ be any elements. Since $\sup_t |f(t)| < \infty$, we conclude that $\|\lambda f\| = \sup_t |\lambda f(t)| = \sup_t |\lambda|\,|f(t)| = |\lambda| \sup_t |f(t)|$, thus $\|\lambda f\| = |\lambda|\, \|f\|$. Moreover, we have $\sup_t |f(t)+g(t)| \leqslant \sup_t(|f(t)|+|g(t)|) \leqslant \sup_t |f(t)|+\sup_t |g(t)| < \infty$, therefore $\|f + g\| \leqslant \|f\| + \|g\|$. Also, if $h \in \ell^\infty(J)$ is a map such that $\|h\| = 0$, then $\sup_t |h(t)| = 0$, which by the definition of the supremum of a set and since $|h(t)| \geqslant 0$, we conclude that $h(t) = 0$ for every $t \in J$ — that is, $h = 0$.

Prove that the maps define norms for ℓ^p and ℓ^∞, moreover, prove that each of the mentioned spaces is Banach.

♮

Lemma 23.1.16. Let $\| - \|_p, \| - \|_\infty\colon k^n \rightrightarrows \mathbf{R}_{\geqslant 0}$ be the norms defined in Example 23.1.15, then the following inequality holds for all $1 \leqslant p < \infty$ and all $x \in k^n$:

$$\|x\|_\infty \leqslant \|x\|_p \leqslant n^{1/p}\|x\|_\infty.$$

Therefore $\| - \|_p$ and $\| - \|_\infty$ are equivalent norms.

Proof. Since $(\sum_j |x_j|^p)^{1/p} \leqslant (\sum_j (\max_j |x_j|)^p)^{1/p} = n^{1/p} \max_j |x_j|$, we obtain $\|x\|_p \leqslant n^{1/p}\|x\|_\infty$. Moreover, clearly $\|x\|_\infty \leqslant \|x\|_p$, thus the proposition follows. ♮

Example 23.1.17 (ℓ^∞ space of maps)**.** Let X be any set and define $\ell^\infty(X)$ to be a k-vector subspace of k^X, consisting of all bounded maps. We define a norm $\|-\|_\infty\colon \ell^\infty(X) \to \mathbf{R}_{\geqslant 0}$ to be the map given by

$$\|f\|_\infty := \sup_{t\in X} |f(t)|.$$

Such norm is called the *uniform norm* on the function space.

Example 23.1.18 (Uniform convergence norm)**.** Let X be a *Hausdorff* space and $\Omega \subseteq X$ be a *compact* set. If $C(\Omega, k)$ denotes the k-vector space of continuous functionals $\Omega \to k$, the map $\| - \|_\infty \colon C(\Omega, k) \to \mathbf{R}_{\geqslant 0}$ given by

$$\|f\|_\infty := \sup_{t \in \Omega} |f(t)| = \max_{t \in \Omega} |f(t)|$$

defines a *norm* in $C(\Omega, k)$. Moreover, the normed k-vector space $(C(\Omega, k), \| - \|_\infty)$ is *Banach*.

Proof.

Prove

$\natural$

23.2 Properties of Normed Vector Spaces

Finite Dimensional

Lemma 23.2.1. Let E be an n-dimensional normed k-vector space and $\{x_1, \dots, x_n\}$ be a basis of E. There exists a scalar $C > 0$ such that, for every choice of scalars $\{\lambda_1, \dots, \lambda_n\} \subseteq k$, we have

$$\left\| \sum_{j=1}^{n} \lambda_j x_j \right\| \geqslant C \sum_{j=1}^{n} |\lambda_j|.$$

Proof. For the sake of brevity, denote $S := \sum_{j=1}^{n} |\lambda_j|$. If $S = 0$ then the lemma follows trivially. Suppose that $S > 0$ and, for the sake of contradiction, that there exists no scalar $C > 0$ such that $\| \sum_j |\lambda_j x_j| \| \geqslant C \sum_j |\lambda_j|$ — this implies that for all integer $p \geqslant 1$ there exists a point $y_p := \sum_{j=1}^{n} \lambda_j(p) x_j$ such that

$$\sum_{j=1}^{n} |\lambda_j(p)| \geqslant p \|y_p\|. \tag{23.2}$$

Moreover, for each $p \geqslant 1$, we may as well construct $z_p := \sum_{j=1}^{n} \alpha_j(p) x_j$ — where $\alpha_j \colon \mathbf{N} \to k$ for all $1 \leqslant j \leqslant n$ — for which

$$\alpha_j(p) := \frac{\lambda_j(p)}{\sum_{j=1}^{n} |\lambda_j(p)|}.$$

This ensures us that $\sum_j \alpha_j(p) = 1$ — hence $|\alpha_j(p)| \leqslant 1$. Dividing Eq. (23.2) by $\sum_j |\lambda_j(p)|$ we find

$$\frac{1}{p} \geqslant \|y_p\|. \tag{23.3}$$

For every fixed $1 \leqslant j \leqslant n$, the sequence $(\alpha_j(p))_p$ is bounded and since we are working either with the complex or real numbers as the underlying field, we can conclude from Bolzano-Weirstraß theorem (see Theorem 14.3.2) that from the sequence of bounded scalars one can extract a convergent subsequence $(\alpha_j(p'))_{p'}$ — assume that $\alpha_j(p') \to \beta_j$ for some $\beta_j \in k$. This induces a subsequence of $(y_p)_p$, given by $(y_{p'})_{p'}$, so that $y_{p'} \to \sum_j \beta_j x_j$ — moreover, $\sum_j |\beta_j| = 1$. However, from Eq. (23.3) we find that $y_p \to 0$, thus also $y_{p'} \to 0$. Since E is Hausdorff, the limit of the sequence $(y_{p'})_{p'}$ must be unique, hence $\sum_j \beta_j x_j = 0$ — which contradicts the initial hypothesis that the set $\{x_1, \dots, x_n\}$ was linearly independent in E.

This shows one cannot build a sequence $(y_p)_p$ for which Eq. (23.2) is satisfied, hence the proposition follows. ♮

Corollary 23.2.2. The unit ball is compact in a finite dimensional normed vector space.

Proof. Let E be an n-dimensional k-vector space and $\{e_1, \dots, e_n\}$ be a basis for E. Let $(x_p)_{p\in\mathbf{N}}$ be a sequence of points in the unit ball $B_0(1)$ (that is, $\|x_p\| \leqslant 1$) — we'll define for each $p \in \mathbf{N}$ that $x_p := \sum_{j=1}^n \alpha_j(p)e_j$, where $\alpha_j : \mathbf{N} \to k$ for all $1 \leqslant j \leqslant n$. One concludes that, since $\|\alpha_j(p)\| \leqslant 1$ for each $1 \leqslant j \leqslant n$, we can use the Bolzano-Weierstraß theorem (see Theorem 14.3.2) to conclude that we can extract a convergent subsequence $(\alpha_j(p'))_{p'}$ from $(\alpha_j(p))_{p\in\mathbf{N}}$ — for instance, assume that $\alpha_j(p) \to \beta_j$ for some $\beta_j \in k$. Then the induced subsequence $(x_{p'})_{p'}$ is such that $x_{p'} \to \sum_j \beta_j e_j := x$. Moreover, for each p' we have the inequality $\|x\| \leqslant \|x - x_{p'}\| + \|x_p\| \leqslant \|x - x_{p'}\| + 1$, thus as $p' \to \infty$ we have $\|x - x_{p'}\| \to 0$ and therefore $\|x\| \leqslant 1$ — that is, every sequence in $B_0(1)$ has a convergent subsequence in $B_0(1)$, which by Theorem 14.3.8 implies that $B_0(1)$ is compact in E. ♮

Lemma 23.2.3. If E is a *finite* dimensional k-vector space, then *all* norms in E are *equivalent*.

Proof. Assume that $\{e_1, \dots, e_n\}$ is a basis for E. We'll show that every norm is equivalent to $\| - \| : E \to \mathbf{R}_{\geqslant 0}$ given by $\|\sum_{j=1}^n \lambda_j e_j\| := \sum_{j=1}^n |\lambda_j|$. Let $\| - \|' : E \to \mathbf{R}_{\geqslant 0}$ be any other norm — then from the triangle inequality we have, for any point $x := \sum_j \lambda_j e_j$ in E:

$$\|x\|' \leqslant \sum_{j=1}^n \|\lambda_j e_j\| \leqslant \max_{1\leqslant j\leqslant n} \|e_j\| \sum_{j=1}^n |\lambda_j| = \max_{1\leqslant j\leqslant n} \|e_j\| \, \|x\|.$$

Defining $C := \max_j \|e_j\|$, we have shown that $\|x\|' \leqslant C\|x\|$.

For the last part, we must prove the existence of a scalar $B > 0$ such that $\|x\| \leqslant B\|x\|'$. We proceed by contradiction, that is, assuming we can choose a sequence of points $(x_p)_{p\in\mathbf{N}}$ such that

$$\|x_p\| \geqslant p\|x_p\|' \tag{23.4}$$

for all $p \in \mathbf{N}$. Let $(y_p)_{p\in\mathbf{N}}$ be the sequence defined by $y_p := x_p/\|x_p\|$, so that $\|y_p\| = 1$ — thus, by Eq. (23.4) we obtain the bound $\|y_p\|' \leqslant 1/p$, thus $y_p \to 0$. If we let $y_p := \sum_j \alpha_j(p)e_j$, we obtain that $\sum_j |\alpha_j(p)| = 1$ implies $|\alpha_j(p)| \leqslant 1$ — hence, for every fixed $1 \leqslant j \leqslant n$, the sequence $(\alpha_j(p))_{p\in\mathbf{N}}$ contains a convergent subsequence $(\alpha_j(p'))_{p'}$ such that $\alpha_j(p') \to \beta_j$ for some $\beta_j \in k$. Moreover, the limits are such that $\sum_j |\beta_j| = 1$. Therefore, the induced subsequence $(y_{p'})_{p'}$ is such that $y_{p'} \to \sum_j \beta_j e_j$ — however, since E is Hausdorff, the sequence must have a unique limit, hence $\sum_j \beta_j e_j = 0$, which is only possible if $\beta_j = 0$ for all $1 \leqslant j \leqslant n$ — this contradicts the fact that $\sum_j |\beta_j| = 1$. We conclude that a sequence $(x_p)_{p\in\mathbf{N}}$ satisfying Eq. (23.4) must not exist, thus implying that there exists $B > 0$ such that $\|x\| \leqslant B\|x\|'$ for all $x \in E$. ♮

Corollary 23.2.4. Every *finite* dimensional normed k-vector space is Banach.

Proof. We'll work on a n-dimensional k-vector space $(E, \|-\|)$ — where we let $\{e_1, \dots, e_n\}$ be any basis and $\|-\|'$ be the norm $\|\sum_j \lambda_j e_j\|' := \sum_j |\lambda_j|$. Let $(x_p)_{p\in\mathbf{N}}$ be any Cauchy sequence with respect to $\|-\|$, and define $x_p := \sum_j \lambda_j(p) e_j$ for some map $\lambda_j\colon \mathbf{N} \to k$. From Lemma 23.2.3 we find that there exists $C > 0$ for which $\|x\|' \leqslant C\|x\|$ for every $x \in E$ — therefore $(x_p)_{p\in\mathbf{N}}$ is also Cauchy with respect to $\|-\|'$. Hence, for all $\varepsilon > 0$ there exists $N \in \mathbf{N}$ such that for all $p, q \geqslant N$ we have $\|x_p - x_q\|' = \sum_j |\lambda_j(p) - \lambda_j(q)| < C\varepsilon$ — therefore for every $1 \leqslant j \leqslant n$ we have $|\lambda_j(p) - \lambda_j(q)| < C\varepsilon$, which implies that $(\lambda_j(p))_{p\in\mathbf{N}}$ is Cauchy in k. Let for instance $\lambda_j(p) \to \alpha_j$ for some $\alpha_j \in k$. Moreover, notice that $\|x_p - \sum_j \alpha_j e_j\|' = \sum_j |\lambda_j(p) - \alpha_j|$, which converges to zero as $p \to \infty$ — thus $x_p \to \sum_j \alpha_j e_j$, which proves that $(x_p)_{p\in\mathbf{N}}$ converges in E. ♮

Proposition 23.2.5 (Compact sets)**.** If E is a *finite* dimensional normed k-vector space, a subset $\Omega \subseteq E$ is *compact* if and only if Ω is *bounded and closed*.

Proof. The first part comes from Proposition 14.2.20. For the second, let Ω be bounded and closed. If $\dim E = n$, we let $\{e_1, \dots, e_n\}$ be a basis for E. Let $(x_p)_{p\in\mathbf{N}}$ be any sequence of points in Ω and assume that each x_p has the form $x_j := \sum_{i=1}^n \lambda_i(p) e_i$ for scalars $\lambda_i(p) \in k$. From the boundness of Ω one can find $M > 0$ such that $\|x_p\| \leqslant C$ for every index $p \in \mathbf{N}$. Evoking Lemma 23.2.1 for each index $p \in \mathbf{N}$, we find that

$$M \geqslant \|x_p\| \geqslant C \sum_{i=1}^n |\lambda_i(p)|,$$

for some $C > 0$. Fixing any $1 \leqslant i \leqslant n$, one concludes that $|\lambda_i(p)| \leqslant M$ and therefore the sequence $(\lambda_i(p))_{p\in\mathbf{N}}$ is bounded. Using Bolzano-Weierstraß theorem (see Theorem 14.3.2) there exists a convergent subsequence $(\lambda_i(p'))_{p'}$ — assume for instance that $\lambda_i(p') \to \lambda_i$ for some $\lambda_i \in k$. Such subsequence induces another subsequence $(x_{p'})_{p'}$, for which $x_{p'} \to \sum_i \lambda_i e_i := x$ — but since Ω is closed, we find $x \in \Omega$. Thus any sequence in Ω has a convergent subsequence in Ω — this shows that Ω is compact (see Theorem 14.3.8). ♮

$\ell^\infty(\mathbf{N})$ and $\ell^p(\mathbf{N})$ are Banach Spaces

Lemma 23.2.6. Every Cauchy sequence in a normed k-vector space is bounded.

Proof. Let $(x_j)_{j\in\mathbf{N}}$ be a Cauchy sequence in a normed k-vector space E. If $\varepsilon = 1$, then there exists $N \in \mathbf{N}$ for which $i, j \geqslant N$ implies $\|x_j - x_i\| < 1$ and therefore $\|x_j\| \leqslant \|x_j - x_N\| + \|x_N\| < 1 + \|x_N\|$. We conclude that the sequence is bounded:

$$\|x_j\| \leqslant 1 + \max_{0\leqslant i\leqslant N} \|x_i\|.$$

♮

Proposition 23.2.7. The space $\ell^\infty(\mathbf{N})$ is Banach.

Proof. Let $x := (x^p)_{p\in\mathbf{N}}$ denote any Cauchy sequence of points $x^p \in \ell^\infty(\mathbf{N})$ — that is, $x^p := (x_j(p))_{j\in\mathbf{N}}$ is itself a sequence — therefore for all $\varepsilon > 0$ there exists $N \in \mathbf{N}$ such that for all $p, q \geqslant N$ we have

$$\|x^p - x^q\| = \sup_{j\in\mathbf{N}} |x_j(p) - x_j(q)| < \varepsilon \tag{23.5}$$

We first construct a candidate for the limit of x. Let $j_0 \in \mathbf{N}$ be any fixed index — we'll show that the sequence $(x_{j_0}(p))_{p\in\mathbf{N}}$ forms a Cauchy sequence in k. Notice that from Eq. (23.5) it is clear that $\sup_{j\in\mathbf{N}} |x_j(p) - x_j(q)| < \varepsilon$ implies in $|x_{j_0}(p) - x_{j_0}(q)| < \varepsilon$ for all $p, q > N$ — therefore $(x_{j_0}(p))_{p\in\mathbf{N}}$ is indeed Cauchy, and since k (either $\mathbf{R}$ or $\mathbf{C}$) is complete, there exists $y_{j_0} \in k$ for which $x_{j_0}(p) \to y_{j_0}$. Our candidate sequence will thus be formed by $y := (y_j)_{j\in\mathbf{N}}$ — where each y_j is constructed just as above. Indeed, $y \in \ell^\infty(\mathbf{N})$, since for every $j \in \mathbf{N}$ we have $|y_j| = \lim_{p\to\infty} |x_j(p)| \leqslant \lim_{p\to\infty} \|x^p\|_\infty$ and since x is Cauchy, from Lemma 23.2.6, we find that there exists $C > 0$ such that $|y_j| \leqslant \|x^p\|_\infty < C$ — hence $\|y\|_\infty = \sup_{j\in\mathbf{N}} |y_j| < \infty$.

We now show that $(y_j)_{j\in\mathbf{N}}$ is the limit of $(x^p)_{p\in\mathbf{N}}$. Let $\varepsilon > 0$ be any bound and $N \in \mathbf{N}$ be such that $p, q > N$ implies $\|x^p - x^q\|_\infty < \varepsilon$ — that is, for any fixed $j_0 \in \mathbf{N}$ we have $|x^p_{j_0} - x^q_{j_0}| < \varepsilon$. Moreover, if we let $q \to \infty$ we'll find that $|x^p_{j_0} - x^q_{j_0}| \to |x^p_{j_0} - y_{j_0}| < \varepsilon$ — hence, since this must be true for any $j_0 \in \mathbf{N}$, we obtain that $\|x\|_\infty = \sup_{j\in\mathbf{N}} |x^q_j - y_j| < \varepsilon$. This proves that $x_j(p) \to y_j$ as $p \to \infty$ and therefore $x^p \to y$ as wanted. Thus any Cauchy sequence converges and the limit is given by the sequence of the limit of the components. ♮

Proposition 23.2.8. The space $\ell^p(\mathbf{N})$ is Banach for all $1 \leqslant p < \infty$.

prove

Common Properties Disregarding the Dimension of the Space

Proposition 23.2.9. Let $f\colon X \to Y$ be an $\mathbf{R}$-linear map between normed $\mathbf{R}$-vector spaces X and Y. Then, f is contiguous if and only if there exists a scalar $C > 0$, called *bound*, for which $\|f(x)\|_Y \leqslant C\|x\|_X$ for all $x \in X$.

Proof. If f is bounded by C, let B be a basis for X and consider any element $x := \sum_{v\in B} a_v v$. From linearity we have $f(x) = \sum_{v\in B} a_v f(v)$, thus

$$\|f(x)\|_Y = \left\| \sum_{v\in B} a_v f(v) \right\|_Y \leqslant \sum_{v\in B} \|a_v\|_{\mathbf{R}} \|f(v)\|_Y \leqslant C\|x\|_X \sum_{v\in B} \|a_v\|_{\mathbf{R}}.$$

This boils down to $f = O(\mathrm{id}_X)$ — which implies in $f(x - x_0) = f(x) - f(x_0) \to 0$ as $x \to x_0$, where $x_0 \in X$ is any point, that is, f is continuous at any point of X. Even better than that, we can show that f is uniformly continuous (I won't carry it out since it's equivalent to what we wrote in Proposition A.2.37).

For the opposite, suppose f is continuous at 0, then there will surely exist $\delta > 0$ for which $\|x\|_X < \delta$ implies in $\|f(x)\|_Y < 1$. Therefore, for any choice of non-zero $x \in X$, we find

$$\left\| f\left(\frac{\delta}{\|x\|_X} x\right) \right\|_Y = \frac{\delta}{\|x\|_X} \|f(x)\|_Y < 1,$$

therefore $\|f(x)\| < \frac{\|x\|_X}{\delta}$, thus f is indeed bounded. ♮

Proposition 23.2.10. Let E, F and G be normed vector spaces, and let $u \colon E \to F$ and $v \colon F \to G$ be continuous linear maps. Then, $vu \colon E \to G$ is also a continuous linear map, and

$$\|vu\| \leqslant \|v\| \, \|u\|,$$

where $\|f\| \coloneqq \sup_{x \in X} \|f(x)\|$ for a continuous linear map $f \colon X \to Y$ between normed vector spaces[1].

Proof. The first assertion is trivial, since the composition of continuous maps is continuous, and the same is true for linear maps. Let $x \in E$ be any element, notice that

$$\|vu(x)\|_G \leqslant \|v\| \, \|u(x)\|_F \leqslant \|v\| \, \|u\| \, \|x\|_E,$$

thus the inequality holds. ♮

Proposition 23.2.11. A multilinear map $\phi \colon \prod_{j=1}^n E_j \to F$ between normed vector spaces $E_1, \dots, E_n$, and F, is continuous if and only if there exists a bound $C > 0$ such that, for every $x \in \prod_{j=1}^n E_j$,

$$\|\phi(x)\|_F \leqslant C \prod_{j=1}^n \|x_j\|_{E_j}.$$

Proposition 23.2.12. Let $E_1, \dots, E_r$, and F be normed vector spaces. There exists a canonical map from repeated continuous linear maps to the continuous multilinear maps, which is a continuous linear isomorphism, and is norm-preserving — that is, the canonical map

$$\Phi \colon L(E_1, L(E_2, \dots, L(E_r, F), \dots)) \xrightarrow{\simeq} L^n(E_1, \dots, E_n; F)$$

is a *Banach isomorphism*.

Proof. We define Φ by the following: if $\lambda \in L(E_1, L(E_2, \dots, L(E_n, F) \dots))$ is given by

$$\lambda(x_1) = \lambda_2, \ \text{ where } \ \lambda_2(x_2) = \lambda_3, \ \dots, \ \lambda_n(x_n) = y \in F,$$

we define $\Phi(\lambda) \coloneqq \overline{\lambda} \in L(E_1, \dots, E_n; F)$ by the mapping

$$\overline{\lambda}(x_1, \dots, x_n) \coloneqq \lambda(x_1)(x_2) \dots (x_n),$$

[1] Beware! This is not the norm we shall adopt for our studies on banachable topological vector spaces. For the latter, see Definition 23.3.4

where $\lambda_j(x_j)(x_{j+1})\dots(x_n) := \lambda_{j-1}(x_{j-1})\dots(x_n)$ for every $1 \leqslant j \leqslant n$ — where $\lambda_1 := \lambda$.

Given $\lambda \in L(E_1, L(E_2, \dots, L(E_n, F), \dots))$, the map $\overline{\lambda}$ is surely multilinear since each of the recursive arguments are linear. Moreover, notice that, for any $x \in \prod_{j=1}^n E_j$, we have

$$\|\overline{\lambda}(x)\|_F \leqslant \|\lambda(x_1)(x_2)\dots(x_n)\|_F \leqslant \|\lambda\| \prod_{j=1}^n \|x_j\|_{E_j},$$

thus $\|\overline{\lambda}\| \leqslant \|\lambda\|$.

On the other hand, given $\overline{\phi} \in L(E_1, \dots, E_n; F)$, define the map $\phi = \Phi^{-1}(\overline{\phi})$ by

$$\phi(x_1)(x_2)\dots(x_n) := \overline{\phi}(x_1, \dots, x_n).$$

Therefore

$$\|\phi(x_1)(x_2)\dots(x_n)\|_F \leqslant \|\overline{\phi}\| \prod_{j=1}^n \|x_j\|_{E_j},$$

which shows that $\|\phi\| \leqslant \|\overline{\phi}\|$. We conclude that $\Phi(\lambda) = \overline{\lambda}$ for all repeating map λ. ♮

Theorem 23.2.13 (Hahn-Banach)**.** Let E be a normed $\mathbf{R}$-vector space, and $F \subseteq E$ be a subspace. Let $\lambda \in F^*$ be a functional with bound $C > 0$. Then there exist an extension of λ to a functional on E with the same bound C — that is, a map $\overline{\lambda}\colon E \to \mathbf{R}$ such that $\overline{\lambda}|_F = \lambda$ and $\|\overline{\lambda}(x)\|_{\mathbf{R}} \leqslant C\|x\|_E$ for all $x \in E$.

Corollary 23.2.14 (Hahn-Banach)**.** Let E be a Banach space and $x \in E$ be a non-zero element. There exists a continuous linear map $\phi \in E^*$ such that $\phi(x) \neq 0$.

Properties of Banach Spaces

Definition 23.2.15. We define a *Banach isomorphism* to be a continuous linear map $u\colon E \to F$, between Banach spaces E and F, that is both invertible (there exists a continuous linear map $u^{-2}\colon F \to E$ that is the two-sided inverse of u), and norm preserving — that is, given any $x \in E$, we have $\|u(x)\|_F = \|x\|_E$. Banach isomorphisms may also be referenced to isometries in the literature.

Proposition 23.2.16 (Bijections are isomorphisms)**.** Every continuous bijective $\mathbf{R}$-linear map between topological vector spaces is an isomorphism.

Proposition 23.2.17 (Splitting)**.** Let E be a Banach space, and F and G be complementary closed subspaces of E — that is, $E = F + G$ and $F \cap G = 0$. Then the morphism $F \times G \to E$ given by $(f, g) \mapsto f + g$ is a continuous linear isomorphism.

23.3 Topological Vector Spaces

Definition 23.3.1. A topological vector space is a k-vector space together with a topology such that addition of vectors and the product by scalars are both continuous k-linear maps.

We denote by $\mathtt{TVect}_{\mathbf{R}}$ the category consisting of topological $\mathbf{R}$-vector spaces together with morphisms, which are continuous $\mathbf{R}$-linear maps (which may also be referenced to by the term "top-linear").

Let E be a topological vector space. The continuous $\mathbf{R}$-linear maps corresponding to the dual space $E^* = \mathrm{Mor}_{\mathtt{TVect}_{\mathbf{R}}}(E, \mathbf{R})$, of a topological $\mathbf{R}$-vector space E, are called $\mathbf{R}$ *forms*. The collection of forms of the form $E \to \mathbf{R}$ will be conveniently separated in classes and denoted:

- $L(E)$: the collection of continuous linear maps $E \to \mathbf{R}$.
- $L^r(E)$: the collection of continuous r-multilinear maps $E^r \to \mathbf{R}$.
- $L^r_{\mathrm{Sym}}(E)$: the collection of continuous r-multilinear symmetric maps $E^r \to \mathbf{R}$.
- $L^r_{\mathrm{Alt}}(E)$: the collection of continuous r-multilinear alternating maps $E^r \to \mathbf{R}$.

Definition 23.3.2 (Locally convex)**.** A topological vector space E is said to be locally convex if, for every open set $U \subseteq E$, any pair of points $x, y \in U$ are such that $tx+(1-t)y \in U$ for all $t \in [0,1]$.

Definition 23.3.3 (Banachable)**.** A topological $\mathbf{R}$-vector space E is said to be banachable if E is complete and its topology can be defined by a norm.

As a point of order, *every time* we mention a topological $\mathbf{R}$-vector space in the course of this chapter, we shall mean a *banachable space*.

Definition 23.3.4 (Norm of a morphism)**.** Let E and F be topological $\mathbf{R}$-vector spaces. In order to make $\mathrm{Mor}_{\mathtt{TVect}_{\mathbf{R}}}(E, F)$ into a topological $\mathbf{R}$-vector space, we can construct a norm for which, given a morphism $A \colon E \to F$, define $K \coloneqq \{k \in \mathbf{R} : \|Ax\|_F \leqslant k\|x\|_E, \text{ for all } x \in E\}$, the norm of A is

$$\|A\| \coloneqq \sup_{k \in K} k.$$

If $\mathrm{Mor}_{\mathtt{TVect}_{\mathbf{R}}}(E_1, \ldots, E_n; F)$ is the collection of continuous $\mathbf{R}$-multilinear maps, then we define similarly the norm of a continuous multilinear map $B\colon \prod_{j=1}^n E_j \to F$ as

$$\|B\| \coloneqq \sup_{m \in M} m,$$

where $M \coloneqq \{m \in \mathbf{R} : \|Bx\|_F \leqslant m \prod_{j=1}^n \|x_j\|_{E_j}, \text{ for all } x \in E\}$.

Remark 23.3.5. From now on, C^p*-morphism* will refer to a map $f\colon U \to V$ between open subsets of Banach spaces such that f is a continuous map of class C^p, where $p \leqslant \infty$.

Part VIII

Statistics & Probability

Chapter 24

Probability Theory

24.1 Probability

Definition 24.1.1 (Sample space & events)**.** We define the *sample space* of an experiment to be the set Ω composed of all possible outcomes.

An *event* is defined to be any subset $A \subseteq \Omega$ of the sample space. The event A is said to have occurred whenever the outcome inhabits A. Two events $A, B \subseteq \Omega$ are said to be *mutually exclusive* whenever A and B are disjoint.

Definition 24.1.2 (σ-algebra)**.** Let Ω be a set. A family of subsets $\Sigma \subseteq 2^{\Omega}$ is called a *σ-algebra* if it satisfies the following conditions:

(a) The empty set is an element of Σ.

(b) If $A \in \Sigma$, then the complement A^{c} is an element of Σ.

(c) If $\{A_j\}_{j\in J}$ is a collection of elements of Σ indexed by a *countable set* J, then the union $\bigcup_{j\in J} A_j$ is also contained in Σ.

From condition (b) it follows immediately that $\Omega \in \Sigma$.

Definition 24.1.3 (Probability function)**.** Given a sample space Ω and an associated σ-algebra Σ, we define a *probability function* on Σ to be a map $\mathbb{P}\colon \Sigma \to \mathbf{R}$ such that

(a) The map $\mathbb{P}$ is non-negative.

(b) The probability of the whole sample space is 1, that is, $\mathbb{P}(\Sigma) = 1$.

(c) Given a *countable set* of pairwise *disjoint* events $\{A_j\}_j \subseteq \Sigma$, then

$$\mathbb{P}\Big(\bigcup_{j\in J} A_j\Big) = \sum_{j\in J} \mathbb{P}(A_j).$$

Lemma 24.1.4. Let $\Omega = \{s_j\}_{j\in J}$ be a countable sample space and Σ be an associated σ-algebra. If $\mathbb{P}\colon \Sigma \to \mathbf{R}$ is a mapping associated with non-negative real numbers $\{p_j\}_{j\in J}$ with $\sum_{j\in J} p_j = 1$, for which

$$\mathbb{P}(A) := \sum_{j: s_j \in A} p_j$$

for each $A \in \Sigma$, then $\mathbb{P}$ is a probability function on Σ.

Theorem 24.1.5. Let Σ be a σ-algebra associated to a sample space Ω, and $\mathbb{P}\colon \Sigma \to \mathbf{R}$ be a probability function. If $A, B \in \Sigma$ are any sets then the following is holds:

(a) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

(b) $\mathbb{P}(\varnothing) = 0$.

(c) $\mathbb{P}(A) \leqslant 1$, therefore $\mathbb{P}(\Sigma) \subseteq [0, 1]$.

(d) $\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

(e) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$, hence $\mathbb{P}(A \cup B) \geqslant \mathbb{P}(A) + \mathbb{P}(B) - 1$, which is known as the *Bonferroni's inequality*.

(f) If $A \subseteq B$ then $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.

Moreover, if $\{C_j\}_{j\in J}$ is a countable partition of Σ, and $\{A_i\}_{i\in\mathbf{N}}$ is any family of elements of Σ, we also have

(g) $\mathbb{P}(A) = \sum_{j\in J} \mathbb{P}(A \cap C_j)$.

(h) $\mathbb{P}(\bigcup_{i\in\mathbf{N}} A_i) \leqslant \sum_{i\in\mathbf{N}} \mathbb{P}(A_i)$.

Proof. (a) Notice that $A \cup A^c = \Sigma$, therefore $\mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) = 1$, which proves the statement.

(b) Since $\Sigma^c = \varnothing$, then $\mathbb{P}(\varnothing) = 1 - \mathbb{P}(\Sigma) = 0$.

(c) Since $\mathbb{P}$ is a non-negative map, then $\mathbb{P}(A^c) \geqslant 0$, hence $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$ implies in $\mathbb{P}(A) \leqslant 1$.

(d) Notice that in general $B = (B \cap A) \cup (B \cap A^c)$, therefore $\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c)$, from which the formula follows.

(e) Since $A \cup B = A \cup (B \cap A^c)$, and from the fact that the sets A and $B \cap A^c$ are disjoint, then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c)$. Using the result from the last item for $\mathbb{P}(B \cap A^c)$ we obtain the required formula. For the inequality, it suffices to see that $\mathbb{P}(A \cap B) \leqslant 1$.

(f) Since $A \cap B = A$ then by the result of item (d) we obtain

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(B \cap A^c)$$

and since $\mathbb{P}(B \cap A^c) \in [0, 1]$, then $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.

(g) Since $\bigcup_{j\in J} C_j = \Sigma$ we have

$$A = A \cap \Sigma = A \cap \Big(\bigcup_{j\in J} C_j\Big) = \bigcup_{j\in J} A \cap C_j,$$

therefore $\mathbb{P}(A) = \mathbb{P}(\bigcup_{j\in J} A \cap C_j) = \sum_{j\in J} \mathbb{P}(A \cap C_j)$.

(h) We shall construct a collection $\{A'_i\}_{i \in \mathbf{N}}$ of disjoint sets partitioning $\bigcup_{i \in \mathbf{N}} A_i$. To do so, define $A'_0 := A_0$ and for any other $i \in \mathbf{N}_{>0}$ we take $A'_i := A_i \setminus \bigcup_{j=0}^{i-1} A_j$. To see that such collection is indeed a partition, let $i, j \in \mathbf{N}$ be any two distinct indices and notice that

$$\begin{aligned} A'_i \cap A'_j &= \Big(A_i \setminus \bigcup_{k=0}^{i-1} A_k\Big) \cap \Big(A_j \setminus \bigcup_{k=0}^{j-1} A_k\Big) \\ &= \Big(A_i \cap \Big(\bigcup_{k=0}^{i-1} A_k\Big)^{\mathrm{c}}\Big) \cap \Big(A_j \cap \Big(\bigcup_{k=0}^{j-1} A_k\Big)^{\mathrm{c}}\Big) \\ &= \Big(A_i \cap \Big(\bigcap_{k=0}^{i-1} A_k^{\mathrm{c}}\Big)\Big) \cap \Big(A_j \cap \Big(\bigcap_{k=0}^{j-1} A_k^{\mathrm{c}}\Big)\Big), \end{aligned}$$

from which, if $i > j$ then $A_j^{\mathrm{c}} \subseteq A_i \cap \bigcap_{k=0}^{i-1} A_k$, thus the first term is disjoint from the second—the case for $i < j$ is symmetric. Since our new collection satisfies the pairwise disjoint condition, we find $\mathbb{P}(\bigcup_{i \in \mathbf{N}} A_i) = \sum_{i \in \mathbf{N}} \mathbb{P}(A'_i)$. From our construction we know that $A'_i \subseteq A_i$ thus $\mathbb{P}(A'_i) \leqslant \mathbb{P}(A_i)$ and hence $\sum_{i \in \mathbf{N}} \mathbb{P}(A'_i) \leqslant \sum_{i \in \mathbf{N}} \mathbb{P}(A_i)$, which proves the statement.

$\natural$

Definition 24.1.6 (Conditional probability)**.** Let A and B be events in a sample space Ω, with $\mathbb{P}(B) > 0$ —where $\mathbb{P}$ is a probability function. We define the *conditional probability of A given B* to be

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

In this case, the map $\mathbb{P}(- \mid B) \colon \Sigma \to \mathbf{R}$ is a probability function associated to $\mathbb{P}$.

Moreover, by symmetry we have $\mathbb{P}(B \mid A) = \mathbb{P}(A \cap B)/\mathbb{P}(A)$, therefore one obtains

$$\mathbb{P}(A \mid B) = \mathbb{P}(B \mid A)\frac{\mathbb{P}(A)}{\mathbb{P}(B)}.$$

When A and B are unrelated events—that is, $A \cap B = \varnothing$ —one has both $\mathbb{P}(A \mid B) = \mathbb{P}(B \mid A) = 0$ since $\mathbb{P}(A \cap B) = 0$.

Theorem 24.1.7 (Bayes' rule)**.** Let $\{A_j\}_{j \in J}$ be a countable partition of the sample space, and let B be any event. Then, for each $j \in J$ on has

$$\mathbb{P}(A_j \mid B) = \frac{\mathbb{P}(B \mid A_j)\,\mathbb{P}(A_j)}{\sum_{j \in J} \mathbb{P}(B \mid A_j)\,\mathbb{P}(A_j)}.$$

Definition 24.1.8 (Statistically independent events)**.** A pair of events A and B is said to be statistically independent from each other if it is the case that $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

Theorem 24.1.9. If A and B are statistically independent events, then the following pairs of events are also independent:

(a) A and B^c.

(b) A^c and B.

(c) A^c and B^c.

Proof. (a) One has

$$\begin{aligned}\mathbb{P}(A \cap B^c) &= \mathbb{P}(A) - \mathbb{P}(A \cap B)\\ &= \mathbb{P}(A) - \mathbb{P}(A)\,\mathbb{P}(B)\\ &= \mathbb{P}(A)(1 - \mathbb{P}(B))\\ &= \mathbb{P}(A)\,\mathbb{P}(B^c).\end{aligned}$$

(b) The symmetric argument can be applied to the case of A^c and B.

(c) For both complements, we have

$$\begin{aligned}\mathbb{P}(A^c \cap B^c) &= \mathbb{P}((A \cup B)^c)\\ &= 1 - \mathbb{P}(A \cup B)\\ &= 1 - (\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B))\\ &= 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A)\,\mathbb{P}(B)\\ &= (1 - \mathbb{P}(A))(1 - \mathbb{P}(B))\\ &= \mathbb{P}(A^c)\,\mathbb{P}(B^c).\end{aligned}$$

$\natural$

Remark 24.1.10. Let A, B and C be three events in our sample space such that

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C).$$

It is not necessarily true that A, B and C are pairwise disjoint—that is, although the probability of them occurring simultaneously splits, it does not mean that the events are pairwise independent! Moreover, one *cannot* define the independence of the events A, B and C by requiring pairwise independence without running into problems—unfortunately the generalisation of Definition 24.1.8 is not that simple, but here we come to the rescue.

Definition 24.1.11 (Mutually independent events)**.** Let $(A_1, \ldots, A_n)$ be a collection of events. We say that they are *mutually independent* if for any subcollection $(A_{j_1}, \ldots, A_{j_k})$ one has

$$\mathbb{P}\Big(\bigcap_{i=1}^{k} A_{j_i}\Big) = \prod_{i=1}^{k} \mathbb{P}(A_{j_i}).$$

24.2 Random Variables

Definition 24.2.1 (Random variable)**.** A random variable is a map $X\colon \Omega \to \mathbf{R}$, where Ω is our ambient sample space. This random variable induces a new sample space $\mathcal{X} = X(\Omega)$, on which we can define a new probability function.

Suppose that Ω is countable. If Σ and Σ_X are σ-algebras associated to Ω and $\mathcal{X}$ respectively, then we define $\mathbb{P}_X\colon \Sigma_X \to \mathbf{R}$ given by

$$\mathbb{P}_X(X = x) := \mathbb{P}(\{y \in \Omega : X(y) = x\}),$$

where $\mathbb{P}\colon \Sigma \to \mathbf{R}$ is a probability function. To ease the notation, we shall merely write $\mathbb{P}(X = x)$ rather than $\mathbb{P}_X(X = x)$.

Now, if Ω is an uncountable sample space, we define $\mathbb{P}_X$ as follows: for any $A \subseteq \mathcal{X}$ let

$$\mathbb{P}_X(X \in A) := \mathbb{P}(\{s \in \Omega : X(s) \in A\}).$$

24.3 Distribution Functions

Definition 24.3.1 (Cumulative distribution function)**.** Let $X\colon \Omega \to \mathbf{R}$ be a random variable. We define a *cumulative distribution function* (cdf) to be map $F_X\colon \Omega \to \mathbf{R}$ given by

$$F_X(x) := \mathbb{P}_X(X \leqslant x)$$

for each $x \in \Omega$.

Part IX

Appendix

Appendix A

Calculus on Several Variables

A.1 Prelude

Sequences

Add important facts about sequences as they come up, just as a way to collect important results

Behaviour of Maps

Before we start our journey through differential calculus on several variables, I would like to point out some really important definitions for classifying the behaviour of maps — of which we'll use extensively.

Definition A.1.1 (Ultimately)**.** We say that a property P is ultimately satisfied by a function f over a filter base $\mathcal{B}$ if there exists a $B \in \mathcal{B}$ such that $P(f|_B)$.

Definition A.1.2 (Little-oh)**.** A function f is said to be little-oh (or infinitesimal) of another function g, which we write as $f =_{\mathcal{B}} o(g)$, if there exists a function α such that $f(x) = \alpha(x)g(x)$ holds ultimately over $\mathcal{B}$, and α is infinitesimal over $\mathcal{B}$.

Definition A.1.3 (Big-oh)**.** Given functions f and g, we say that f is big-oh of g, and write $f =_{\mathcal{B}} O(g)$, if there exists a function β such that ultimately over $\mathcal{B}$ we have $f(x) = \beta(x)g(x)$, and β is ultimately bounded over $\mathcal{B}$.

Definition A.1.4 (Order over base)**.** We say that functions f and g have the same order over $\mathcal{B}$, and write $f \asymp_{\mathcal{B}} g$, if $f =_{\mathcal{B}} O(g)$ and $g =_{\mathcal{B}} O(f)$, or equivalently, if exists $a, b > 0$ such that for some $B \in \mathcal{B}$ we have $a|g(x)| \leqslant |f(x)| \leqslant b|g(x)|$.

Definition A.1.5. Given functions f, g, we say that f behaves asymptotically like g over $\mathcal{B}$, and write $f \sim_{\mathcal{B}} g$, if there exists a function γ such that $\lim_{\mathcal{B}} \gamma(x) = 1$ and $f(x) = \gamma(x)g(x)$ ultimately over $\mathcal{B}$.

Fundamental Inequalities

Now we take a look at some fundamental inequalities that are used in some of the proofs of Appendix A.3.

Lemma A.1.6. For $x > 0$ we have

$$x^{\alpha} - \alpha x + \alpha - 1 \leqslant 0, \quad \text{for } 0 < \alpha < 1, \tag{A.1}$$

$$x^{\alpha} - \alpha x + \alpha - 1 \geqslant 0, \quad \text{for } \alpha < 0 \text{ or } 1 < \alpha. \tag{A.2}$$

Proof. Let $f(x) = x^{\alpha} - \alpha x + \alpha - 1$, then $f'(x) = \alpha(x^{\alpha-1} - 1)$. Notice that $f'(1) = 0$ and that for $\alpha \in (0,1)$ we have that for some $\delta > 0$, $f'(1-\delta) > 0$ and $f'(1+\delta) < 0$, which shows that $x = 1$ is a strict maximum. In the case where $\alpha < 0$ or $\alpha > 1$, $f'(1-\delta) < 0$ and $f'(1+\delta) > 0$, showing that $x = 1$ is a strict minimum. The strictness comes from the fact that f is monotone in the intervals $x \in (0,1)$ and $x > 1$. Since $f(1) = 0$, then for $\alpha \in (0,1)$ the function is non-positive, and for $\alpha < 0$ or $\alpha > 1$ the function is non-negative. ♮

Proposition A.1.7 (Young's inequalities)**.** Let $a, b > 0$ and $p, q \notin \{0,1\}$ such that $p^{-1} + q^{-1} = 1$. Then

$$a^{p^{-1}} b^{q^{-1}} \leqslant \frac{a}{p} + \frac{b}{q}, \quad \text{if } p > 1, \tag{A.3}$$

$$a^{p^{-1}} b^{q^{-1}} \geqslant \frac{a}{p} + \frac{b}{q}, \quad \text{if } p < 1. \tag{A.4}$$

The equality of such relations hold only when $a = b$.

Proof. Let $\alpha = p^{-1}$ and set $x = \frac{a}{b}$. From Eq. (A.1) we have

$$\begin{aligned} 0 \geqslant \left(\frac{a}{b}\right)^{\frac{1}{p}} - \frac{1}{p}\frac{a}{b} + \frac{1}{p} - 1 &= \left(\frac{a}{b}\right)^{\frac{1}{p}} - \frac{1}{p}\frac{a}{b} - \frac{1}{q} \\ \frac{1}{p}\frac{a}{b} + \frac{1}{q} &\geqslant \left(\frac{a}{b}\right)^{\frac{1}{p}} \\ \frac{a}{p} + \frac{b}{q} &\geqslant a^{\frac{1}{p}} b^{1-\frac{1}{p}} = a^{p^{-1}} b^{q^{-1}}. \end{aligned}$$

Now, from Eq. (A.1) we have equivalently that

$$\frac{a}{p} + \frac{b}{q} \geqslant a^{p^{-1}} b^{q^{-1}}.$$

♮

Proposition A.1.8 (Hőlder's inequalities)**.** Let $x_j, y_j \geqslant 0$ for $1 \leqslant j \leqslant n$ and $p^{-1} + q^{-1} = 1$. Then

$$\sum_{1\leqslant j\leqslant n} x_j y_j \leqslant \Big(\sum_{1\leqslant j\leqslant n} x_j^p\Big)^{\frac{1}{p}} \Big(\sum_{1\leqslant j\leqslant n} y_j^q\Big)^{\frac{1}{q}} \quad \text{for } p > 1, \tag{A.5}$$

$$\sum_{1\leqslant j\leqslant n} x_j y_j \geqslant \Big(\sum_{1\leqslant j\leqslant n} x_j^p\Big)^{\frac{1}{p}} \Big(\sum_{1\leqslant j\leqslant n} y_j^q\Big)^{\frac{1}{q}} \quad \text{for } p < 1 \text{ and } p \neq 0. \tag{A.6}$$

If $p < 0$, then we need the strictness $x_j > 0$ for all $1 \leqslant j \leqslant n$. Equality is obtained for the case where $(x_j^p)_{j=1}^n$ and $(y_j^q)_{j=1}^n$ are linearly dependent.

Proof. Define $x = \sum_{j=1}^n x_j > 0$ and $y = \sum_{j=1}^n y_j > 0$. We can use Eq. (A.3) with $a = \frac{x_j^p}{x}$ and $b = \frac{y_j^q}{y}$, for which we find that

$$\frac{x_j}{x^{p^{-1}}}\frac{y_j}{y^{q^{-1}}} \leqslant \frac{1}{p}\frac{x_j^p}{x} + \frac{1}{q}\frac{y_j^q}{y}$$

hence, summing such inequality over $1 \leqslant j \leqslant n$ we find

$$\frac{\sum_{j=1}^n x_j y_j}{x^{p^{-1}}y^{q^{-1}}} \leqslant \frac{1}{p}\frac{\sum_{j=1}^n x_j^p}{x} + \frac{1}{q}\frac{\sum_{j=1}^n y_j^p}{y} = \frac{1}{p} + \frac{1}{q} = 1$$

and finally Eq. (A.5) is shown

$$\sum_{j=1}^n x_j y_j \leqslant x^{p^{-1}} y^{q^{-1}} = \Big(\sum_{j=1}^n x_j^p\Big)^{\frac{1}{p}} \Big(\sum_{j=1}^n y_j^q\Big)^{\frac{1}{q}}.$$

The same equivalent proof can be made with Eq. (A.4) for Eq. (A.6). Since the equality of the Young's inequalities occurs only for $a = b$, we find that the linear dependence $x_j^p = \frac{x}{y} y_j^q$ implies in the equality of Hőlder's inequalities. ♮

Proposition A.1.9 (Minkowski's inequalities)**.** Let $x_j, y_j \geqslant 0$ for $1 \leqslant j \leqslant n$. Then

$$\Big(\sum_{1\leqslant j\leqslant n} (x_j + y_j)^p\Big)^{\frac{1}{p}} \leqslant \Big(\sum_{1\leqslant j\leqslant n} x_j^p\Big)^{\frac{1}{p}} + \Big(\sum_{1\leqslant j\leqslant n} y_j^p\Big)^{\frac{1}{p}} \text{ for } p > 1, \tag{A.7}$$

$$\Big(\sum_{1\leqslant j\leqslant n} (x_j + y_j)^p\Big)^{\frac{1}{p}} \geqslant \Big(\sum_{1\leqslant j\leqslant n} x_j^p\Big)^{\frac{1}{p}} + \Big(\sum_{1\leqslant j\leqslant n} y_j^p\Big)^{\frac{1}{p}} \text{ for } p < 1, \text{ and } p \neq 0 \tag{A.8}$$

The equality occurs when $(x_j)_{1\leqslant j\leqslant n}$ and $(y_j)_{1\leqslant j\leqslant n}$ are linearly dependent.

Proof. Notice that

$$\sum_{j=1}^n (x_j + y_j)^p = \sum_{j=1}^n (x_j + y_j)(x_j + y_j)^{p-1} = \sum_{j=1}^n x_j (x_j + y_j)^{p-1} \sum_{j=1}^n y_j (x_j + y_j)^{p-1}$$

If $p > 0$, then applying Eq. (A.5) we find (noting that $q = \frac{p}{p-1}$)

$$\sum_{j=1}^n (x_j + y_j)^p \leqslant \Big[\Big(\sum_{j=1}^n x_j^p\Big)^{\frac{1}{p}} + \Big(\sum_{j=1}^n y_j^p\Big)^{\frac{1}{p}}\Big]\Big(\sum_{j=1}^n (x_j + y_j)^p\Big)^{\frac{1}{q}}.$$

For $p < 1$ with $p \neq 0$, from Eq. (A.6) we get

$$\sum_{j=1}^{n}(x_j + y_j)^p \geqslant \left[\left(\sum_{j=1}^{n} x_j^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} y_j^p\right)^{\frac{1}{p}}\right]\left(\sum_{j=1}^{n}(x_j + y_j)^p\right)^{\frac{1}{q}}.$$

Now, dividing both inequalities by the term $\left(\sum_{j=1}(x_j + y_j)^p\right)^{\frac{1}{q}}$ we find respectively Eq. (A.7) and Eq. (A.8). The equality occurs the same as with Hőlder inequalities. ♮

Fixed Points and Banach Spaces

Definition A.1.10 (Fixed point)**.** Let $f\colon X \to X$ be any map. A fixed point of f is an element $x \in X$ such that $f(x) = x$.

Theorem A.1.11 (Fixed point theorem)**.** Let $I \subseteq \mathbf{R}$ be a closed set and $f\colon I \to \mathbf{R}$ a function such that $f(I) \subseteq \mathbf{R}$ and for some fixed $\theta \in [0, 1)$ we have, for all $x, y \in \mathbf{R}$:

$$|f(x) - f(y)| \leqslant \theta|x - y|.$$

Then there exists a unique fixed point $c \in I$, that is $f(c) = c$.

Proof. Let $x_0 \in I$ and define the sequence $x_n := f(x_{n-1})$ for all $n \geqslant 1$. We first show that $(x_n)_{n\in\mathbf{N}}$ is Cauchy. Let $n > m \geqslant 1$, then

$$|x_n - x_m| = \left|\sum_{k=m}^{n} x_{k+1} - x_k\right| \leqslant \sum_{k=m}^{n} |x_{k+1} - x_k| = \sum_{k=m}^{n} |f(x_k) - f(x_{k-1})| \tag{A.9}$$

Since $|f(x_k) - f(x_{k-1})| \leqslant \theta|x_k - x_{k-1}|$ and $f(x_k) = x_{k+1}$, we can make $\prod_{i=1}^{k} |f(x_i) - f(x_{i-1})| \leqslant \theta^k \prod_{i=1}^{n} |x_i - x_{i-1}|$ and divide both the inequality by $\prod_{i=1}^{k} |f(x_i) - f(x_{i-1})| = \prod_{i=1}^{k} |x_i - x_{i-1}|$ in order to obtain

$$|f(x_k) - f(x_{k-1})| \leqslant \theta^k |x_1 - x_0|. \tag{A.10}$$

Now we can substitute Eq. (A.10) in Eq. (A.9), then

$$|x_n - x_m| \leqslant \sum_{k=m}^{n-1} \theta^k |x_1 - x_0|$$

Moreover, since $\theta \in [0, 1)$ we have from the geometric series that $\sum_{k=0}^{\infty} \theta^k = \frac{1}{1-\theta}$, so we can conclude that

$$|x_n - x_m| \leqslant \frac{\theta^m}{1-\theta}|x_1 - x_0|.$$

and thus $(x_n)_{n\in\mathbf{N}}$ is indeed a Cauchy sequence.

Let $x_n \to c \in I$, since I is closed and thus $\operatorname{Cl} I = I$. Since $|f(x) - f(y)| \leqslant \theta|x - y|$ the function is Lipschitz continuous, hence

$$\lim_{n\to\infty} f(x_n) = f(c) = \lim_{n\to\infty} x_{n+1} = c$$

and therefore c is a fixed point of f.

For the uniqueness of the fixed point, let c_1, c_2 be fixed points of f, then $|f(c_1) - f(c_2)| = |c_1 - c_2| \leqslant \theta |c_1 - c_2|$ and thus $(1-\theta)|c_1 - c_2| \leqslant 0$, but $\theta \in [0, 1)$, hence $c_1 = c_2$. ♮

Corollary A.1.12. Let $I \subseteq \mathbf{R}$ be closed and $f\colon I \to I$ be a differentiable function such that exists $\theta \in [0,1)$ for which $|f'(x)| \leqslant \theta$, for all $x \in I$. Then there exists a unique fixed point of f.

Proof. Choose any distinct points $x, y \in I$, from the mean value theorem, there exists $x_0 \in (x, y)$ such that $f(x) - f(y) = f'(x_0)(x - y)$, then $|f(x) - f(y)| \leqslant \theta |x - y|$, which satisfies the condition of Theorem A.1.11, hence the proposition holds. ♮

Definition A.1.13. Let $(V, \|\cdot\|)$ be a normed vector space. We say that a sequence $(x_n)_{n \in \mathbf{N}} \subseteq V$ is Cauchy with respect to the norm $\|\cdot\|$ if for all $\varepsilon > 0$ there exists an index $N \in \mathbf{N}$ such that, for all $n, m \geqslant N$, we have $\|x_n - x_m\| < \varepsilon$.

Definition A.1.14 (Banach space)**.** A normed vector space $(V, \|\cdot\|)$ is a Banach space if every Cauchy sequence converges with respect to $\|\cdot\|$.

Definition A.1.15. Let B be a Banach space. A subset $A \subseteq B$ is said to be closed if the limit of every convergent sequence in A belongs to A.

Definition A.1.16 (Contraction)**.** Let B be a Banach space and $0 < \theta < 1$, then a map $f\colon B \to B$ is said to be a θ-contraction if for all $v, w \in B$ we have

$$\|f(v) - f(w)\| \leqslant \theta \|v - w\|.$$

Theorem A.1.17 (Banach fixed point)**.** Let B be a Banach space and $A \subseteq B$ be a closed subset. Let $f\colon A \to A$ be a θ-contraction. Then f has a unique fixed point.

Proof. The proof of the Banach fixed point is merely the same analogous proof as the one developed in Theorem A.1.11. ♮

Proposition A.1.18 (Fixed point stability)**.** Let $A \subseteq B$ be a closed subspace of the Banach space B. Let $\Omega \subseteq B$ e an open subspace of B. Consider the collection $\{f_x \in B(A, A) : x \in \Omega\}$ of θ-contractions such that the map $x \mapsto f_x(y)$ is continuous — that is, $\lim_{x \to x_0} f_x(y) = f_{x_0}(y)$. Then the solution map $s\colon \Omega \to A$ defined as

$$s(x) = y \text{ if and only if } f_x(y) = y$$

is continuous at x_0 — that is, $\lim_{x \to x_0} s(x) = s(x_0)$.

Proof. We know from Theorem A.1.17 that — given any $x \in \Omega$ — the fixed point (unique) solution can be obtained as the limit of a sequence recursively defined as $y_j = f_x(y_{j-1})$ and $y_0 \in A$ being any element. This way, consider such sequence $(y_j)_{j=1}^{\infty}$ but define $y_0 = s(x_0)$. Notice that since $\sum_{j=1}^{n} y_j - y_{j-1} = y_n - y_0$, then we can write y_n in the following form

$$y_n = \sum_{j=1}^{n} (y_j - y_{j-1}) + y_0 = \sum_{j=2}^{n} (f_x(y_{j-1}) - f_x(y_{j-2})) + y_0 = \sum_{j=1}^{n} \left(f_x^{j-1}(y_1) - f_x^{j-1}(y_0) \right) + y_0.$$

Now observe that

$$\sum_{j=1}^{n} \|f_x^{j-1}(y_1) - f_x^{j-1}(y_0)\| \leqslant \sum_{j=1}^{n} \theta^{j-1}\|y_1 - y_0\| = \frac{\|y_1 - y_0\|}{1-q}.$$

That is, $\|y_n - y_0\| \leqslant \frac{\|y_1 - y_0\|}{1-q}$, hence — since $(y_j)_{j=1}^{\infty}$ converges to the fixed point of f_x, we have

$$\|s(x) - s(x_0)\| = \|f(s(x)) - f(s(x_0))\| \leqslant \frac{\|y_1 - y_0\|}{1-q} = \frac{\|f_x(s(x_0)) - f_{x_0}(s(x_0))\|}{1-q}.$$

On the other hand we also know that $\lim_{x\to x_0} f_x(y) = f_{x_0}(y)$ thus

$$\lim_{x\to x_0} \|f_x(s(x_0)) - f_{x_0}(s(x_0))\| = 0.$$

This shows that $\lim_{x\to x_0} \|s(x) - s(x_0)\| = 0$ and therefore

$$\lim_{x\to x_0} s(x) = s(x_0).$$

$\natural$

Definition A.1.19. Let V and W be Banach spaces. We define the set $B(V, W)$ as the collection of all linear maps $f\colon V \to W$.

A.2 Continuity

Remark A.2.1. This part will be mainly concerned with the euclidean metric space given by $\mathbf{R}^n$ and the metric

$$d(x, y) = \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2}$$

where $x = (x_j)_{j=1}^n, y = (y_j)_{j=1}^n \in \mathbf{R}$.

Compact sets in $\mathbf{R}^n$

Definition A.2.2. A set $K \subseteq \mathbf{R}^n$ is compact if for every open cover C of K there exists a finite subcover $\mathcal{U} \subseteq C$.

Definition A.2.3 (General closed interval)**.** We define a closed interval I in $\mathbf{R}^n$ to be the set

$$I = \{x \in R^n : x_j \in [a_j, b_j], 1 \leqslant j \leqslant n\}.$$

where $a, b \in \mathbf{R}^n$ are the boundaries of the interval I.

Proposition A.2.4. A closed interval in $\mathbf{R}^n$ is compact.

Proof. Suppose for the sake of contradiction that $\mathcal{U}$ is a cover of I, closed interval in $\mathbf{R}^n$, such that $\mathcal{U}$ doesn't admit a finite subcover. Consider the set of bisections of I in which for each component of the vectors $x \in I$, that is, we create the sets $I_j^1 = \{x \in I : x_j \in [a_j, (a_j + b_j)/2]\}$ and $I_j^2 = \{x \in I : x_j \in [(a_j + b_j)/2, b_j]\}$ for each $1 \leqslant j \leqslant n$, generating 2^n subsets of I. Notice that since these sets are contained in I, at least one of those should not admit a finite subcover from $\mathcal{U}$, otherwise I would be compact. Hence define such set as I_1. Now recursively do the same bisection process for I_1. We end up with a chain of nested intervals

$$I \supset I_1 \supset I_2 \supset \dots$$

each of which does not admit a finite subcover from $\mathcal{U}$. Consider the interval $I_m = \{x \in \mathbf{R}^n : x_j \in [a_j^m, b_j^m], 1 \leqslant j \leqslant n\}$ from the nested chain. Notice that for each $1 \leqslant j \leqslant n$ we have that the coordinate closed intervals form again a chain of nested intervals

$$[a_j, b_j] \supseteq [a_j^1, b_j^1] \supseteq [a_j^2, b_j^2] \supseteq \dots$$

hence $\lim_{t\to\infty} d(a_j^t, b_j^t) = 0$. Since they form a nested chain, their intersection is non-empty and therefore there exists a point $\zeta_j \in [a_j^m, b_j^m]$ common to all such intervals. In doing so for $1 \leqslant j \leqslant n$ we find a point $\zeta = (\zeta_j)_{j=1}^n \in \mathbf{R}^n$ such that $\zeta \in I_i$ for all $i \geqslant 1$ and $\zeta \in I$. From the last assertion one sees that there must exist $U \in \mathcal{U}$ with $\zeta \in U$, hence exists $\varepsilon > 0$ such that $B_\zeta(\varepsilon) \subseteq U$. Now, since $\lim_{t\to\infty} d(a_j^t, b_j^t) = 0$, it must be true that there exists $M > 0$ such that for all $m > M$ we have $I_m \subseteq B_\zeta(\varepsilon) \subseteq U$, which clearly covers finitely I_m. This contradicts the assumption that all I_m couldn't be finitely covered by a subcover of $\mathcal{U}$. This shows that we cannot pick a subset of I with such property, implying that I itself should be compact. ♮

Proposition A.2.5. Let K be a compact set of $\mathbf{R}^n$, then

(a) K is closed in $\mathbf{R}^n$.

(b) Any closed subset of K is compact.

Proof. (a) Let $y \in \mathbf{R}^n$ be any limit point of K. Suppose that $y \notin K$. For each point in K, say x, denote U_x a neighbourhood. Consider the collection $\mathcal{U} = \{U_x : x \in K\}$, which covers the set K. Since K is said to be compact, there exists a finite subcover $\mathcal{U}' = \{U_{x_1}, \dots, U_{x_m}\} \subseteq \mathcal{U}$. From the hypothesis y does not belong to K, we can find a neighbourhood V_j of y for $1 \leqslant j \leqslant m$ such that $U_{x_j} \cap V_j = \varnothing$. Consider now the neighbourhood $V = \bigcap_{1\leqslant j\leqslant m} V_j$ of y. Since $K = \bigcup \mathcal{U}'$, we have $K \cap V = \varnothing$, therefore y cannot be a limit point of K, which is a contradiction. This implies that $y \in K$, if not, problematic neighbourhoods V_j can be chosen.

(b) Let $C \subseteq \mathbf{R}^n$ be a closed set and $C \subseteq K$. Let $\mathcal{G}$ be an open cover of C in $\mathbf{R}^n$. Notice that $\mathcal{U} = \mathcal{G} \cup (\mathbf{R}^n \setminus C)$ is an open cover of $\mathbf{R}^n$, which clearly covers K. Therefore there exists a finite subcover of K, $\mathcal{U}' \subseteq \mathcal{U}$, but since $C \subseteq K$, then $\mathcal{U}'$ also covers C. Since $(\mathbf{R}^n \setminus C) \cap C = \varnothing$, then $\mathcal{U}' \setminus \{\mathbf{R}^n \setminus C\}$ is a finite subcover of C from $\mathcal{G}$, therefore C is a compact set.

♮

Definition A.2.6. The diameter of a set $A \subseteq \mathbf{R}^n$ is defined to be

$$d(A) = \sup_{x,y \in A} d(x,y).$$

Definition A.2.7. A set $A \subseteq \mathbf{R}^n$ is said to be bounded if $d(A)$ is finite.

Proposition A.2.8. If $K \subseteq \mathbf{R}^n$ is compact, then K is also bounded in $\mathbf{R}^n$.

Proof. Let $\mathcal{B}$ be the collection of all open balls around a given point $x \in \mathbf{R}^n$, the set $\mathcal{B}$ covers $\mathbf{R}^n$ and therefore also covers K. Notice that since K is compact, there exists a finite number of open balls $\mathcal{B}' \subseteq \mathcal{B}$ that cover K, hence the distance between any elements of K must be finite. ♮

Theorem A.2.9 (Heine-Borel)**.** Let $K \subseteq \mathbf{R}^n$ be any set. The following statements are equivalent:

(a) K is closed and bounded.

(b) K is compact.

Proof. Notice that the implication (b) $\Rightarrow$ (a) is already proven by the last two propositions (Proposition A.2.5 and Proposition A.2.8). Suppose now that K is a closed and bounded set. Since K is bounded, there exists a closed interval $I \supset K$. Since I is compact (Proposition A.2.4) and K is closed we find that K itself is compact (Proposition A.2.5). ♮

Limits

Remark A.2.10. In this subsection we shall denote a general set as X.

Definition A.2.11 (Limit)**.** Let $f\colon X \to \mathbf{R}^n$ be a map. We say that $x \in \mathbf{R}^n$ is the limit of f over a filter base $\mathcal{B} \subseteq 2^X$ if for every neighbourhood V of x there exists $B \in \mathcal{B}$ for which $f(B) \subseteq V$.

Definition A.2.12 (Bounded)**.** A map $f\colon X \to \mathbf{R}^n$ is said to be bounded if $f(X) \subseteq \mathbf{R}^n$ is bounded.

Definition A.2.13 (Ultimately bounded)**.** Given a filter base $\mathcal{B} \subseteq 2^X$, a map $f\colon X \to \mathbf{R}^n$ is ultimately bounded over the base $\mathcal{B}$ if there exists $B \in \mathcal{B}$ for which f is bounded.

Proposition A.2.14 (Unique limit)**.** A map can have at most one limit over a filter base.

Proof. Let $f\colon X \to \mathbf{R}^n$ be a map and suppose that $\lim_{\mathcal{B}} f(x) = a$ and $\lim_{\mathcal{B}} f(x) = b$, where $a \neq b$. Since they are distinct points, there must exist neighbourhoods V_a and V_b for which $V_a \cap V_b = \varnothing$. Now, remember from the definition that there must exist $B_a, B_b \in \mathcal{B}$ such that $f(B_a) \subseteq V_a$ and $f(B_b) \subseteq V_b$. From the downward direction property of filter bases, there exists $B \subseteq B_a \cap B_b$ in $\mathcal{B}$. Since $\varnothing \notin \mathcal{B}$, then $B \neq \varnothing$, hence $f(B) \subseteq V_a \cap V_b$ is non-empty, contradicting the assumption that there could be chosen non-intersecting neighbourhoods of a and b, which implies that $a = b$ in $\mathbf{R}^n$. ♮

Proposition A.2.15. If a map has a limit over a given filter base, then the map is ultimately bounded over that filter base.

Proof. Let $f\colon X \to \mathbf{R}^n$ be a map and $\mathcal{B} \subseteq 2^X$ be a filter base. Assume that $\lim_{\mathcal{B}} f(x) = \ell \in \mathbf{R}^n$. For the sake of contradiction, suppose that f is not ultimately bounded over $\mathcal{B}$. Let $B_\ell(\varepsilon)$ be an open ball centred at ℓ with radius $\varepsilon > 0$. From the definition of a limit over a filter, there exists $B \in \mathcal{B}$ for which $f(B) \subseteq B_\ell(\varepsilon)$, but $\sup_{x,y\in B_\ell(\varepsilon)} d(x,y) = 2\varepsilon$, which contradicts the fact that $d(f(B))$ is not bounded in $\mathbf{R}$. Therefore, f needs to be ultimately bounded over $\mathcal{B}$. ♮

Corollary A.2.16. Let $f\colon X \to \mathbf{R}^n$ and $\mathcal{B}$ be a filter base over X. The map has a limit y over $\mathcal{B}$ if and only if each of the projection functions $\pi_j f$ have limit y_j. That is,

$$\lim_{\mathcal{B}} f(x) = y \Leftrightarrow \lim_{\mathcal{B}} \pi_j f(x) = y_j, \quad \text{for } 1 \leqslant j \leqslant n$$

Definition A.2.17 (Cauchy sequence)**.** A sequence $(x_j)_{j\in\mathbf{N}}$ of points in $\mathbf{R}^n$ is a Cauchy sequence if for every $\varepsilon > 0$ there exists an index $N \in \mathbf{N}$ for which $d(x_i, x_j) < \varepsilon$ for all $i, j > N$.

Proposition A.2.18. A sequence $(x_j)_{j\in\mathbf{N}}$ of points in $\mathbf{R}^n$ is Cauchy if and only if $(x_j^i)_{j\in\mathbf{N}}$ is Cauchy for all $1 \leqslant i \leqslant n$, where $x_j = (x_j^i)_{i=1}^n$.

Proof. Notice that since

$$d(x_j^i, x_k^i) \leqslant d(x_j, x_k) \leqslant \sqrt{n} \max_{1\leqslant i\leqslant n} d(x_j^i, x_k^i)$$

then, if $(x_j)_{j\in\mathbf{N}}$ is Cauchy, we have that the inequality $d(x_j^i, x_k^k) \leqslant d(x_j, x_k)$ implies that $(x_j^i)_{j\in\mathbf{N}}$ is Cauchy for each $1 \leqslant i \leqslant n$. Now, if in turn we have that for each component the sequence $(x_j^i)_{j\in\mathbf{N}}$ is Cauchy, then since $d(x_j, x_k) \leqslant \sqrt{n} \max_{1\leqslant i\leqslant n} d(x_j^i, x_k^i)$ we find that $(x_j)_{j\in\mathbf{N}}$ is Cauchy. ♮

Proposition A.2.19. A sequence in $\mathbf{R}^n$ is convergent if and only if the sequence is Cauchy.

Proof. Suppose that $(x_j)_{j\in\mathbf{N}}$ is a convergent sequence in $\mathbf{R}^n$ with $x_j \to x$. Let $\varepsilon > 0$, choose any neighbourhood U of x such that $d(U) = \varepsilon$. Since the sequence converges, there exists $N \in \mathbf{N}$ for which $x_j \in U$ for all $j \geqslant N$, that is, for all $j, k \geqslant N$ we have $d(x_j, x_k) < \varepsilon$. Hence we conclude that $(x_j)_{j\in\mathbf{N}}$ is Cauchy. For the opposite case, let $(x_j)_{j\in\mathbf{N}}$ be a Cauchy sequence. Then clearly there exists an element x for which every neighbourhood contains infinitely many points of $(x_j)_{j\in\mathbf{N}}$. ♮

Definition A.2.20 (Oscillation)**.** The oscillation of $f\colon X \to \mathbf{R}^n$ on $E \subseteq X$ is given by

$$\omega(f, E) = d(f(E)).$$

Theorem A.2.21 (Cauchy criterion for several variables)**.** Let $f\colon X \to \mathbf{R}^n$ be a map and $\mathcal{B}$ be a filter base over X. The map f has a limit over $\mathcal{B}$ if and only if for all $\varepsilon > 0$ there exist $B \in \mathcal{B}$ such that $\omega(f, B) < \varepsilon$.

Proof. Apply the Cauchy criterion for single variable maps on each of $\pi_j f$ for $1 \leqslant j \leqslant n$, now, using Corollary A.2.16, we see that theorem is true for f. ♮

Theorem A.2.22. Let $g: Y \to R^n$ and $f: X \to Y$ be mappings. Let filter basis $\mathcal{B}_Y$ on Y — such that g has a limit over $\mathcal{B}_Y$ — and $\mathcal{B}_X$ on X such that for all $B_Y \in \mathcal{B}_Y$ there exists $B_X \in \mathcal{B}_X$ for which $f(B_X) \subseteq B_Y$. Then the composition $gf: X \to \mathbf{R}^n$ has a limit over $\mathcal{B}_X$ and we have the relation

$$\lim_{\mathcal{B}_X} gf(x) = \lim_{\mathcal{B}_Y} g(y).$$

Proof. Apply the property of the limit of the composition of single variable maps to each of the $\pi_j g$ and $\pi_j f$. From Corollary A.2.16 we see that the theorem is true for g and f. ♮

Definition A.2.23 (Limit at infinity)**.** Let $f: E \to \mathbf{R}^n$, where $E \subseteq \mathbf{R}^m$. The filter base that yields the limit $x \to \infty$ is given by $\mathcal{B}_\infty = \{\mathbf{R}^m \setminus B_a(r) : r \in \mathbf{R}\}$ for any fixed point $a \in \mathbf{R}^m$.

Definition A.2.24 (Limit to infinity)**.** Let $f: E \to \mathbf{R}^n$, where $E \subseteq \mathbf{R}^m$, and a filter base $\mathcal{B}$ on E. We say that $f(x) \to_{\mathcal{B}} \infty$ if — given any fixed point $y \in \mathbf{R}^n$ — any open ball $B_y(r) \subseteq \mathbf{R}^n$ is such that there exists $B \in \mathcal{B}$ for which $f(B) \subseteq \mathbf{R}^n \setminus B_y(r)$.

Continuity

Remark A.2.25. Throughout this subsection we shall assume that E is a subset of $\mathbf{R}^m$.

Definition A.2.26 (Continuous)**.** A map $f: E \to \mathbf{R}^n$ is said to be continuous at a point $x \in E$ if for every neighbourhood V of $f(x)$ there exists a neighbourhood $U \subseteq E$ of x such that $f(U) \subseteq V$.

Corollary A.2.27. A map $f: E \to \mathbf{R}^n$ is continuous at a point x if and only if $\pi_j f$ is continuous at x for each $1 \leqslant j \leqslant n$.

Definition A.2.28 (Path)**.** We define a path in $\mathbf{R}^n$ to be a continuous map between an interval $I \subseteq \mathbf{R}$ and $\mathbf{R}^n$.

Definition A.2.29 (Support)**.** We define the support of a path $\gamma: I \to \mathbf{R}^n$ to be the image $\gamma(I)$.

Definition A.2.30 (Oscillation at a point)**.** Let $f: E \to \mathbf{R}^n$ be a map and $x \in E$. We define the oscillation of f at the point x as the limit

$$\omega(f, x) = \lim_{r \to +0} \omega(f, B_a(r) \cap E).$$

Proposition A.2.31 (Local properties)**.** Let $f: E \to \mathbf{R}^n$ be a map.

(a) f is continuous at $x \in E$ if and only if $\omega(f, x) = 0$.

(b) If f is continuous at a point $x \in E$, then f is bounded in some neighbourhood $U_x \cap E$ of x.

(c) Let set $X \subseteq \mathbf{R}^m$ and $Y \in \mathbf{R}^n$. Let $g: Y \to \mathbf{R}^k$ be a continuous map at $y \in Y$. Let $f: X \to Y$ be continuous at $x \in X$ and $f(x) = y$. Then the map $gf: X \to R^k$ is continuous at x.

If the map is real valued, we also have more properties. Let $f, g: E \to \mathbf{R}$, then

(a) If f is continuous at a point $\bar{x} \in E$, there exists a neighbourhood $U \cap E$ of $\bar{x}$ such that $f(x)f(\bar{x}) > 0$ for all $x \in U \cap E$.

(b) If f and g are continuous at a point $x \in E$, then for any $\alpha, \beta \in \mathbf{R}$ we have that the linear combination $\alpha f + \beta g: E \to \mathbf{R}$, their product $f \cdot g: E \to \mathbf{R}$ and — if $g(x) \neq 0$ — the quotient $\frac{f}{g}: E \to \mathbf{R}$ are all continuous at the point x.

Proof.

Write proofs: local properties

♮

Definition A.2.32 (Uniformly continuous)**.** Let $f: X \to Y$ be a map between metric spaces. We say that f is uniformly continuous on X if for every $\varepsilon > 0$ there exists $\delta > 0$ such that, for every subset $E \subseteq X$ of diameter $d(E) < \delta$, we have an oscillation $\omega(f, E) < \varepsilon$.

Theorem A.2.33 (Heine-Cantor theorem)**.** Let $f: X \to Y$ be a continuous map between metric spaces X and Y. If X is compact, then f is uniformly continuous.

Proof. Let any $\varepsilon > 0$. Since f is continuous, there exists, for any $x \in X$ a $\delta_x > 0$ for which, if $d_X(x, y) < \delta_x$, then $d_Y(f(x), f(y)) < \varepsilon/2$. Lets consider the open cover $\mathcal{U}$ of X consisting of the neighbourhoods $U_x := \{y \in X : d_X(x, y) < \frac{\delta_x}{2}\}$ for each $x \in X$. From definition, if X is compact, then there exists a finite subcover $\{U_{x_j}\}_{j=1}^n \subseteq \mathcal{U}$ of X. Define the minimum radius of the given neighbourhoods as $\delta := \min_{1 \leqslant j \leqslant n} \delta_{x_j}/2$.

Let $x, y \in X$ be any points such that $d_X(x, y) < \delta$. From the finite subcover, we have that there exists $1 \leqslant j_0 \leqslant n$ such that $x \in U_{x_{j_0}}$, which implies that $d_X(x, x_{j_0}) < \delta_{x_{j_0}}/2$, thus $d_Y(f(x), f(x_{j_0})) < \varepsilon/2$, from construction. Notice however that

$$d_X(y, x_{j_0}) \leqslant d_X(y, x) + d_X(x, x_{j_0}) < \delta + \delta_{x_{j_0}} \leqslant \delta_{x_{j_0}},$$

therefore it follows that $d_Y(f(x_{j_0}), f(y)) < \varepsilon/2$. Using again the triangle inequality we observe that

$$d_Y(f(x), f(y)) \leqslant d_Y(f(x), f(x_{j_0})) + d_Y(f(x_{j_0}), f(y)) < \varepsilon.$$

Therefore, f is indeed uniformly continuous with constant δ. ♮

Definition A.2.34 (Pathwise connected)**.** A set $X \subseteq \mathbf{R}^n$ is pathwise connected if for all $x, y \in E$ there exists a path $\gamma: I \to E$ with endpoints at x and y and support in E.

Definition A.2.35 (Domain)**.** A domain in $\mathbf{R}^n$ is an open pathwise connected subset of $\mathbf{R}^n$.

Proposition A.2.36 (Global properties)**.** The following are global properties on continuous maps of several variables. Let $K \subseteq \mathbf{R}^m$ be a compact set and $E \subseteq \mathbf{R}^m$ be pathwise connected.

(a) A continuous map $f\colon K \to \mathbf{R}^n$ is uniformly continuous.

(b) A continuous map $f\colon K \to \mathbf{R}^n$ is bounded on K.

(c) A continuous map $f\colon K \to \mathbf{R}$ assumes its maximal and minimal values at least once in K.

(d) Let $f\colon E \to \mathbf{R}$ be a continuous map and assume $f(a) = A$ and $f(b) = B$ at $a, b \in E$. For any $C \in [A, B] \subseteq \mathbf{R}$ there exists $c \in E$ such that $f(c) = C$.

Proof.

Write proofs: global properties

(d) From the connectedness of E, let $\gamma\colon [x, y] \to E$ be a continuous path such that $\gamma(x) = a$ and $\gamma(y) = b$. Consider the composition of continuous maps $f\gamma\colon I \to \mathbf{R}$. Since $f\gamma(a) = A$ and $f\gamma(b) = B$, for any given $C \in [A, B]$, there exists $z \in [x, y]$ such that $f\gamma(z) = C$, hence there exists $c = \gamma(z) \in E$ for which $f(c) = C$. ♮

Proposition A.2.37. Every linear map of the form $L\colon \mathbf{R}^m \to \mathbf{R}^n$ is continuous. Moreover, it is uniformly continuous.

Proof. Let $f\colon E \to \mathbf{R}^n$ be any map, then for all $1 \leqslant j \leqslant n$ we have

$$\|\pi_j f(x)\|_{\mathbf{R}} \leqslant \|f(x)\|_{\mathbf{R}^n} \leqslant \sum_{j=1}^{n} \|\pi_j f(x)\|_{\mathbf{R}}$$

In particular, for the linear map L we have that — given any $x \in \mathbf{R}^m$

$$\|L(x)\|_{\mathbf{R}^n} = \|\sum_{j=1}^{m} x_j L(e_j)\|_{\mathbf{R}^n} \leqslant \sum_{j=1}^{m} \|x_j\|_{\mathbf{R}} \|L(e_j)\|_{\mathbf{R}^n} \leqslant \|x\|_{\mathbf{R}^m} \left(\sum_{j=1}^{m} \|L(e_j)\|_{\mathbf{R}^n} \right).$$

Hence $L = O(\mathrm{id})$ — where $\mathrm{id}\colon \mathbf{R}^m \to \mathbf{R}^m$ mapping $x \mapsto x$. It follows from this that as $x \to x_0$, we have $L(x - x_0) = L(x) - L(x_0) \to 0$. This shows that L is continuous at any point of $\mathbf{R}^m$. Notice that given any $\varepsilon > 0$ if $\|x - y\|_{\mathbf{R}^m} < \frac{\varepsilon}{\ell}$ — where $\ell := \sum_{j=1}^{m} L(e_j)$ — then

$$\|L(x) - L(y)\|_{\mathbf{R}^n} = \|\sum_{j=1}^{m} (x_j - y_j) L(e_j)\|_{\mathbf{R}^n} \leqslant \sum_{j=1}^{m} \|L(e_j)\|_{\mathbf{R}^n} \|x_j - y_j\|_{\mathbf{R}} \leqslant \ell \|x - y\|_{\mathbf{R}^m} < \ell \frac{\varepsilon}{\ell} = \varepsilon$$

where we conclude that L is uniformly continuous. ♮

A.3 Differentiable Maps

Remark A.3.1. Throughout this chapter, we'll denote by $\|\|\colon \mathbf{R}^n \to \mathbf{R}$ the standard norm in $\mathbf{R}^n$, given by

$$\|x\| = \sqrt{\sum_{j=1}^{n} x_j^2}.$$

Moreover, the set E is a any subset of $\mathbf{R}^m$.

Definition A.3.2 (Little-Oh)**.** Given maps $f\colon X \to \mathbf{R}^n$ and $g\colon X \to \mathbf{R}^m$, we say that f is little-Oh of g over a filter base $\mathcal{B} \subseteq 2^X$ — and write $f = o(g)$ — if $\|f(x)\|_{\mathbf{R}^n} = o(\|g(x)\|_{\mathbf{R}^m})$ over $\mathcal{B}$.

Proposition A.3.3. Let $L\colon \mathbf{R}^m \to \mathbf{R}^n$ be a linear map, then for $h \in \mathbf{R}^m$, as $h \to 0$ we have

$$L(o(h)) = o(h).$$

Proof. Define $f(h) = \alpha(h)h = o(h)$ — that is $\alpha\colon \mathbf{R}^n \to \mathbf{R}$, where $\alpha(h) \to 0$ as $h \to 0$. This yields $L(f(h)) = L(\alpha(h)h) = \alpha(h)\sum_{j=1}^{n} h_j L(e_j)$, therefore

$$\begin{aligned}
\|L(f(h))\|_{\mathbf{R}^n} &= \|\alpha(h)\|_{\mathbf{R}}\|\sum_{j=1}^{n} h_j L(e_j)\|_{\mathbf{R}^n} \\
&\leqslant \|\alpha(h)\|_{\mathbf{R}} \sum_{j=1}^{n} \|L(e_j)\|_{\mathbf{R}^n}\|h_j\|_{\mathbf{R}} \\
&\leqslant \left(\|\alpha(h)\|_{\mathbf{R}} \sum_{j=1}^{n} \|L(e_j)\|_{\mathbf{R}^n}\right)\|h\|_{\mathbf{R}^m}
\end{aligned}$$

and since $\|\alpha(h)\|_{\mathbf{R}} \sum_{j=1}^{n}\|L(e_j)\|_{\mathbf{R}^n} \to 0$ as $h \to 0$, we find that $L(o(h)) = o(h)$. ♮

Definition A.3.4 (Big-Oh)**.** Let maps $f\colon X \to \mathbf{R}^n$ and $g\colon X \to \mathbf{R}^m$. f is said to be big-Oh of g over a filter base $\mathcal{B} \subseteq 2^X$ — and denote by $f = O(g)$ — if $\|f(x)\|_{\mathbf{R}^n} = O(\|g(x)\|_{\mathbf{R}^m})$ over $\mathcal{B}$.

Proposition A.3.5. Let $L\colon \mathbf{R}^m \to \mathbf{R}^n$ be a linear map. For any $h \in \mathbf{R}^m$ we have that — as $h \to 0$

$$L(h) = O(h).$$

Proof. Consider $h = \sum_{j=1}^{m} h_j e_j$, then $L(h) = \sum_{j=1}^{m} h_j L(e_j)$. From Minkowski's inequalities (see Proposition A.1.9) we have

$$\|L(h)\|_{\mathbf{R}^n} = \|\sum_{j=1}^{m} h_j L(e_j)\|_{\mathbf{R}^n} \leqslant \sum_{j=1}^{m}\|L(e_j)\|_{\mathbf{R}^n}\|h_j\|_{\mathbf{R}} \leqslant \left(\sum_{j=1}^{m}\|L(e_j)\|_{\mathbf{R}^n}\right)\|h\|_{\mathbf{R}^m} = M\|h\|_{\mathbf{R}^m}$$

Thus, as $h \to 0$ that $L(h) = O(h)$. ♮

Definition A.3.6 (Tangent space)**.** Let $x \in \mathbf{R}^n$. We define the tangent space of x — denoted by $T_x\mathbf{R}^n$ — to be the R-vector space spanned by x.

Definition A.3.7 (Differentiable)**.** A map $f\colon E \to \mathbf{R}^n$ is differentiable at a point $x \in E$ — where x is a limit point of E — if

$$f(x+t) - f(x) = L(x)(t) + \alpha(x,t) \tag{A.11}$$

where $L(x)\colon \mathbf{R}^m \to \mathbf{R}^n$ is a linear map — on the variable t — and $\alpha(x,t) = o(t)$ as $t \to 0$. We call the linear map $L(x)$ the *differential* (or tangent map) of f at the point $x \in E$ — we shall normally denote $L(x)$ by $\mathrm{d}f(x)\colon T_x\mathbf{R}^m \to T_{f(x)}\mathbf{R}^n$. We can clarify the last equation by writing it equivalently as

$$f(x+t) - f(x) = \mathrm{d}f(x)(t) + o(t). \tag{A.12}$$

Proposition A.3.8. A map $f\colon E \to \mathbf{R}^n$ is differentiable at a limit point $x \in E$ if and only if, for all $1 \leqslant j \leqslant n$, the maps $\pi_j \circ f\colon E \to \mathbf{R}$ are differentiable at x.

Proof. Notice that Eq. (A.12) can be rewritten coordinate-wise — as $t \to 0$

$$\pi_j \circ f(x+t) - \pi_j \circ f(x) = (\pi_j \circ \mathrm{d}f(x))(t) + o(t). \tag{A.13}$$

Assume that $\pi_j \circ \mathrm{d}f(x)$ has the general form of $(\pi_j \circ \mathrm{d}f(x))(t) \coloneqq \sum_{j=1}^m a_x^j t_j$, where $a_x^j \in \mathbf{R}$ are scalars dependent on $x \in E$ for all $1 \leqslant j \leqslant m$. Consider, for each $1 \leqslant j \leqslant m$, the displacements $t = t_j e_j$ so that $\|t\|_{\mathbf{R}^m} = \|t_j\|_{\mathbf{R}}$. For such displacements, we find — as $t_j \to 0$

$$\pi_j \circ f(x + t_j e_j) - \pi_j \circ f(x) = a_x^j t_j + o(t_j)$$

therefore we find that each of the scalar corresponding to displacements on the jth coordinate — of the jth differential $\pi_j \circ f$ of f — can be written as

$$a_x^j = \lim_{t_j \to 0} \frac{\pi_j \circ f(x + t_j e_j) - \pi_j \circ f(x)}{t_j}. \tag{A.14}$$

The equivalence between Eq. (A.12) and Eq. (A.13) shows that the map f is differentiable at x if and only if each of its coordinate decompositions are differentiable at x. ♮

Partial Derivative and Differential of Real Valued Maps

Definition A.3.9 (Partial derivative)**.** Let $f\colon E \to \mathbf{R}$, where $E \subseteq \mathbf{R}^m$. We define the partial derivative of f at the point $x \in \operatorname{Int} E$, with respect to its jth coordinate, to be the real value (if existent)

$$\partial_j f(x) = \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t}$$

Notice that, given a map $f\colon E \to \mathbf{R}^n$, the partial derivative of each of its coordinate-wise maps $\pi_j \circ f$ is the real value

$$\partial_j(\pi_j \circ f)(x) = \lim_{\mathbf{R}\ni t\to 0} \frac{\pi_j \circ f(x+te_j) - \pi_j \circ f(x)}{t}.$$

Consider the jth projection map $\pi_j\colon \mathbf{R}^m \to \mathbf{R}$. From the definition Definition A.3.7, the differential of π_j at a point $x \in \mathbf{R}^m$ is given by $\mathrm{d}\pi_j(x)(t) = \pi_j(x+t) - \pi_j(x) = t_j$, which is independent of x. We can define now the following operator, that will stand as a clever notation for the differential of the projection maps.

Notation A.3.10. Given a point $x \in \mathbf{R}^m$, we denote the linear map $\mathrm{d}x_j\colon T_x\mathbf{R}^m \to T_{\pi_j(x)}\mathbf{R}$ as

$$\mathrm{d}x_j(t) = t_j.$$

Proposition A.3.11 (Real valued differential)**.** Let $f\colon E \to \mathbf{R}$ be a differentiable map at an interior point $x \in \operatorname{Int} E$. Then f has a partial derivative at x for each of its variables. The differential of f at the point x — the map $\mathrm{d}f(x)\colon T_x\mathbf{R}^m \to T_{f(x)}\mathbf{R}^n$ — is uniquely defined as the sum of the partial derivatives of f at x — that is, for any $t \in T_x\mathbf{R}^m$, we have

$$\mathrm{d}f(x)(t) = \sum_{1\leqslant j\leqslant m} \partial_j f(x) t_j.$$

In view of the introduced Notation A.3.10, we can rewrite the differential as the map

$$\mathrm{d}f(x) = \sum_{1\leqslant j\leqslant m} \partial_j f(x)\mathrm{d}x_j.$$

Notation A.3.12. For the sake of brevity — given a function $f\colon E \to R^n$ — we define $f_j \coloneqq \pi_j \circ f$ for each $1 \leqslant j \leqslant n$.

Differential of a Map $\mathbf{R}^m \to \mathbf{R}^n$

Now we can generalize the results for real valued functions $E \to \mathbf{R}$ to functions of the type $E \to \mathbf{R}^n$.

Corollary A.3.13 (Several variables differential)**.** Let $f\colon E \to \mathbf{R}^n$ be differentiable at the interior point $x \in \operatorname{Int} E$. The differential of f at x, $\mathrm{d}f(x)\colon T_x\mathbf{R}^m \to T_{f(x)}\mathbf{R}^m$, exists and is uniquely given by — for any given $t \in T_x\mathbf{R}^m$

$$\mathrm{d}f(x)(t) = \begin{bmatrix} \mathrm{d}f_1(x)(t) \\ \vdots \\ \mathrm{d}f_n(x) \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^m \partial_j f_1(x) t_j \\ \vdots \\ \sum_{j=1}^m \partial_j f_n(x) t_j \end{bmatrix} = \begin{bmatrix} \partial_1 f_1(x) & \dots & \partial_m f_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(x) & \dots & \partial_m f_n(x) \end{bmatrix} \begin{bmatrix} t_1 \\ \vdots \\ t_m \end{bmatrix} \tag{A.15}$$

Definition A.3.14 (Jacobi matrix)**.** The matrix given by the partial derivatives of the projections of $f\colon E \to \mathbf{R}^n$ — $[\partial_i f_j(x)]_{i,j}$, with $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$ — is called the Jacobi matrix f at x.

Notation A.3.15 (Jacobi matrix)**.** We'll denote the Jacobi matrix of a map $f\colon E \to \mathbf{R}^n$ at an interior point x — where f is differentiable — as $f'(x)$.

Definition A.3.16 (Jacobian)**.** Let $E \subseteq \mathbf{R}^n$ and a map $f\colon E \to \mathbf{R}^n$ — where f is differentiable at an interior point $x \in \operatorname{Int} E$. The Jacobi matrix of f at the point x is a $n \times n$ square matrix and its determinant $\det[\partial_i f_j(x)]_{i,j}$ is called the Jacobian of f at x.

Connections Between Differentiability and Continuity

Corollary A.3.17 (Differentiable implies continuous)**.** Let $f\colon E \to \mathbf{R}^n$ be a differentiable map at a point $x \in E$. Then f is continuous at x.

Proof. By definition, the differential of f at x exists and is a linear map. By Proposition A.2.37 we find that $\mathrm{d}f(x)$ is continuos and hence $\mathrm{d}f(x)(t) \to 0$ as $t \to 0$. Then as $t \to 0$ we have $f(x+t) - f(x) = \mathrm{d}f(x)(t) + o(t) \to 0$ so $\lim_{t\to 0} f(x+t) = f(x)$ — the map is continuous at x. ♮

Remark A.3.18 (Differential and partial derivatives)**.** The Proposition A.3.11 shows that if a map is differentiable at an interior point of its domain, then the partial derivatives exist at the given point. However, the converse does not hold. For instance, consider the map $g\colon \mathbf{R}^2 \to \mathbf{R}$ given by

$$g(x,y) := \begin{cases} \frac{xy}{x^2+y^2}, & x^2+y^2 \neq 0 \\ 0, & x^2+y^2 = 0 \end{cases}$$

Notice that $g(0,y) = g(x,0) = 0$ but $g(x,x) = \frac{1}{2}$ when $x \neq 0$, therefore g has no limit as $(x,y) \to 0$. On the other hand, g has partial derivatives for all points over the plane

$$\partial_1 g(x,y) = \frac{y(y^2-x^2)}{(x^2+y^2)^2} \quad \text{and} \quad \partial_2 g(x,y) = \frac{x(x^2-y^2)}{(x^2+y^2)^2}.$$

Laws of Differentiability

Theorem A.3.19. Let $f, g\colon E \to \mathbf{R}^n$ be differentiable maps at the point $x \in E$, then the linear combination $\lambda f + \gamma g\colon E \to \mathbf{R}^n$, where $\lambda, \gamma \in \mathbf{R}$ are scalars, is differentiable at x and the differential $\mathrm{d}(\lambda f + \gamma g)\colon T_x\mathbf{R}^m \to T_{\lambda f(x)+\gamma g(x)}\mathbf{R}^n$ is given by

$$\mathrm{d}(\lambda f + \gamma g)(x) = (\lambda \mathrm{d}f + \gamma \mathrm{d}g)(x).$$

Proof. Note that

$$\begin{aligned}(\lambda f + \gamma g)(x+t) - (\lambda f + \gamma g)(x) &= \lambda(f(x+t) - f(x)) + \gamma(g(x+t) + g(x)) \\ &= \lambda(\mathrm{d}f(x)(t) + o(t)) + \gamma(\mathrm{d}g(x)(t) + o(t)) \\ &= (\lambda \mathrm{d}f(x) + \gamma \mathrm{d}g(x))(t) + o(t)\end{aligned}$$

thus $\lambda f + \gamma g$ is differentiable at x. Taking the limit as $t \to 0$, we find

$$\mathrm{d}(\lambda f + \gamma g)(x)(t) = \lambda \mathrm{d}f(x)(t) + \gamma \mathrm{d}g(x)(t).$$

If we also assume that $x \in \operatorname{Int} E$, then we can rewrite the last relation as

$$\begin{bmatrix} \partial_1(\lambda f + \gamma g)_1(x) & \dots & \partial_m(\lambda f + \gamma g)_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1(\lambda f + \gamma g)_n(x) & \dots & \partial_m(\lambda f + \gamma g)_n(x) \end{bmatrix} = \begin{bmatrix} \lambda\partial_1 f_1(x) + \gamma\partial_1 g_1(x) & \dots & \lambda\partial_m f_1(x) + \gamma\partial_m g_1(x) \\ \vdots & \ddots & \vdots \\ \lambda\partial_1 f_n(x) + \gamma\partial_1 g_n(x) & \dots & \lambda\partial_m f_n(x) + \gamma\partial_m g_n(x) \end{bmatrix}$$

thus the following equality is obtained

$$\partial_i(\lambda f + \gamma g)_j(x) = \lambda\partial_i f_j(x) + \gamma\partial_i g_j(x)$$

$\natural$

Theorem A.3.20. Let $h, \ell\colon E \to \mathbf{R}$ be differentiable maps at $x \in E$. Then

(a) The product $h \cdot \ell\colon E \to \mathbf{R}$ is differentiable at x and the differential $\mathrm{d}(h \cdot \ell)\colon T_x\mathbf{R}^m \to T_{(h\cdot\ell)(x)}\mathbf{R}$ is given by

$$\mathrm{d}(h \cdot \ell)(x) = h(x)\mathrm{d}\ell(x) + \ell(x)\mathrm{d}h(x). \tag{A.16}$$

If we also assume $x \in \operatorname{Int} E$, then the relation can be rewritten matricially as

$$\begin{bmatrix} \partial_1(h \cdot \ell)(x) & \cdots & \partial_m(h \cdot \ell)(x) \end{bmatrix} = \begin{bmatrix} h(x)\partial_1\ell(x) + \ell\partial_1 h(x) & \cdots & h(x)\partial_m\ell(x) \end{bmatrix}$$

thus the following relation is obtained

$$\partial_i(h \cdot \ell)(x) = h(x)\partial_i\ell(x) + \ell(x)\partial_i h(x) \quad \text{for } 1 \leqslant i \leqslant m \tag{A.17}$$

(b) If $\ell(x) \neq 0$, the quotient $\frac{h}{\ell}\colon E \to \mathbf{R}$ is differentiable at x and the differential of such quotient at x, $\mathrm{d}\frac{h}{\ell}\colon T_x\mathbf{R}^m \to T_{\frac{h}{\ell}(x)}\mathbf{R}$, is given by

$$\mathrm{d}\frac{h}{\ell}(x) = \frac{\ell(x)\mathrm{d}h(x) - h(x)\mathrm{d}\ell(x)}{\ell^2(x)}. \tag{A.18}$$

If we also assume $x \in \operatorname{Int} E$, then the relation can be rewritten matricially as

$$\begin{bmatrix} \partial_1\frac{h}{\ell}(x) & \cdots & \partial_m\frac{h}{\ell}(x) \end{bmatrix} = \frac{1}{\ell^2(x)} \begin{bmatrix} \ell(x)\partial_1 h(x) - h(x)\partial_1\ell(x) & \cdots & \ell(x)\partial_m h(x) - h(x)\partial_m\ell(x) \end{bmatrix}$$

thus we obtain

$$\partial_i\frac{h}{\ell}(x) = \frac{\ell(x)\partial_i h(x) - h(x)\partial_i\ell(x)}{\ell^2(x)} \tag{A.19}$$

Theorem A.3.21 (Composition)**.** Let $A \subseteq \mathbf{R}^m$ and $B \subseteq \mathbf{R}^n$ be any sets and define maps $f\colon A \to B$ and $g\colon B \to R^k$. Let f and g be such that f is differentiable at $x \in A$ and g is differentiable at $f(x) \in B$. Then the map $g \circ f\colon A \to \mathbf{R}^k$ is differentiable at $x \in A$. Moreover, the differential $\mathrm{d}(g \circ f)(x)\colon T\mathbf{R}^m_x \to T\mathbf{R}^k g(f(x))$ is

$$\mathrm{d}(g \circ f)(x) = \mathrm{d}g(f(x)) \circ \mathrm{d}f(x). \tag{A.20}$$

Also, if $x \in \operatorname{Int} A$ and $f(x) \in \operatorname{Int} B$, then the theorem can be rewritten in terms of the Jacobian matrix

$$\begin{bmatrix} \partial_1(g_1 \circ f)(x) & \dots & \partial_m(g_1 \circ f)(x) \\ \vdots & \ddots & \vdots \\ \partial_1(g_k \circ f)(x) & \dots & \partial_m(g_k \circ f)(x) \end{bmatrix} = \begin{bmatrix} \partial_1 g_1(f(x)) & \dots & \partial_n g_1(f(x)) \\ \vdots & \ddots & \vdots \\ \partial_1 g_k(f(x)) & \dots & \partial_n g_k(f(x)) \end{bmatrix} \begin{bmatrix} \partial_1 f_1(x) & \dots & \partial_m f_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(x) & \dots & \partial_m f_n(x) \end{bmatrix}$$

hence we find the relation — for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant k$

$$\partial_i(g_j \circ f)(x) = \sum_{1 \leqslant r \leqslant n} (\partial_r g_j(f(x))) \cdot (\partial_i f_r(x)) \tag{A.21}$$

Proof. Consider the filter base given by $h \to 0$, where $f(x+h) - f(x) = \mathrm{d}f(x)(h) + o(h)$, and as $t \to 0$ we find $g(f(x)+t) - g(f(x)) = \mathrm{d}g(f(x)) + o(t)$. Therefore by defining $t := f(x+h) - f(x)$ we can write — as $h, t \to 0$

$$\begin{aligned} g(f(x+h)) - g(f(x)) &= g(f(x)+t) - g(f(x)) \\ &= \mathrm{d}g(f(x))(t) + o(t) \\ &= \mathrm{d}g(f(x))(f(x+h) - f(x)) + o(f(x+h) - f(x)) \\ &= \mathrm{d}g(f(x))(\mathrm{d}f(x)(h) + o(h)) + o(f(x+h) - f(x)) \\ &= \mathrm{d}g(f(x))(\mathrm{d}f(x)(h)) + \mathrm{d}g(f(x))(o(h)) + o(f(x+h) - f(x)) \end{aligned}$$

From definition we have that — as $h \to 0$

$$\mathrm{d}g(f(x))(o(h)) = g(f(x) + o(h)) - g(f(x)) - o(h),$$

thus $\mathrm{d}g(f(x))(o(h)) \to 0$ as $h \to 0$, that is, $\mathrm{d}g(f(x))(o(h)) = o(h)$. Also, as $h \to 0$, we have — recalling Proposition A.3.5

$$t = f(x+h) - f(x) = \mathrm{d}f(x)(h) + o(h) = O(h) + o(h) = O(h).$$

Notice that $o(O(h)) = o(h)$, thus $o(f(x+h)-f(x)) = o(h)$, and hence we finally conclude that

$$\begin{aligned} g(f(x+h)) - g(f(x)) &= \mathrm{d}g(f(x))(\mathrm{d}f(x)(h)) + \mathrm{d}g(f(x))(o(h)) + o(f(x+h) - f(x)) \\ &= \mathrm{d}g(f(x))(\mathrm{d}f(x)(h)) + o(h) + o(h) \\ &= \mathrm{d}g(f(x))(\mathrm{d}f(x)(h)) + o(h) \end{aligned}$$

which from definition implies that $g \circ f$ is differentiable at x and $\mathrm{d}(g \circ f)(x) = \mathrm{d}g(f(x))(\mathrm{d}f(x))$. ♮

Definition A.3.22 (Directional derivative)**.** Let $f : E \to \mathbf{R}$ and $x \in \operatorname{Int} E$, and $v \in T\mathbf{R}^m_x$. If the following limit exists

$$\mathrm{D}_v f(x) := \lim_{t \to 0} \frac{f(x+vt) - f(x)}{t} \tag{A.22}$$

then the real quantity $\mathrm{D}_v f(x) \in T\mathbf{R}_{f(x)}$ is called the directional derivative of f at the point x evaluated at the vector v.

Definition A.3.23 (Gradient)**.** Let $E \subseteq \mathbf{R}^m$ and $f\colon E \to \mathbf{R}$ be a real map. We define the gradient of f as the vector field $\operatorname{grad} f\colon \mathbf{R}^m \to \mathbf{R}^m$ written as

$$\operatorname{grad} f = (\partial_1 f, \ldots, \partial_m f)\,.$$

Corollary A.3.24. Let $E \subseteq \mathbf{R}^m$. If $f\colon E \to \mathbf{R}$ is differentiable at $x \in \operatorname{Int} E$, then

$$\mathrm{D}_v f(x) = \mathrm{d}f(x)(v) = \sum_{1 \leqslant j \leqslant m} \partial_j f(x) v_j.$$

Written in terms of the euclidean inner product in $\mathbf{R}^m$:

$$\mathrm{D}_v f(x) = \langle \operatorname{grad} f, v \rangle.$$

Proof. Let $\ell\colon \mathbf{R} \to E$ be a map $\ell(t) = x + vt$. Then the composition $f \circ \ell$ is such that the differential at the point $t = 0$ — that is $\mathrm{d}(f \circ \ell)(0)\colon T\mathbf{R}_0 \to T\mathbf{R}_x$ — is given by $\mathrm{d}(f \circ \ell)(0) = \mathrm{d}f(x) \circ \mathrm{d}\ell(0)$, from the composition theorem. Thus, if $x \in \operatorname{Int} E$, then

$$\partial(f \circ \ell)(0) = \sum_{1 \leqslant r \leqslant m} (\partial_r f(x)) \cdot (\partial \ell_r(0)).$$

Moreover, $\partial \ell_r(0) = \lim_{\mathbf{R} \ni h \to 0} \frac{\ell_r(0+h) - \ell_r(0)}{h} = \lim_{h \to 0} \frac{(x_r + v_r h) - x_r}{h} = v_r$, which implies that

$$\partial(f \circ \ell)(0) = \sum_{1 \leqslant r \leqslant m} \partial_r f(x) v_r = \mathrm{d}f(x)(v).$$

In order to make the connection to the directional derivative, it suffices to observe that since $f \circ \ell\colon \mathbf{R} \to \mathbf{R}$ then $\partial(f \circ \ell)$ is just the single variable derivative of the composite map — that is

$$\partial(f \circ \ell)(0) = \lim_{\mathbf{R} \ni t \to 0} \frac{(f \circ \ell)(0 + t) - f(\ell(0))}{t} = \lim_{\mathbf{R} \ni t \to 0} \frac{f(x + vt) - f(x)}{t} = \mathrm{D}_v f(x).$$

♮

Let $f\colon E \to \mathbf{R}$ be a differentiable map at $x \in E$. Define ϕ to be the angle between $\operatorname{grad} f(x)$ and a given unit vector $u \in T\mathbf{R}^m_x$. Lets analyse the implication of certain choices of u:

- If $u = \frac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|}$ then $\phi = 0$ and hence

  $$\mathrm{D}_u f(x) = \langle \operatorname{grad} f(x), u \rangle = \|\operatorname{grad} f(x)\| \|u\| \cos(\phi) = \|\operatorname{grad} f(x)\|,$$

  which indicates that for this choice of u the value $\mathrm{D}f(x) \in \mathbf{R}$ is maximized.

- If u is any vector perpendicular to $\operatorname{grad} f(x)$ then $\mathrm{D}_u f(x) = 0$.

- If $u = -\frac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|}$, then $\phi = \frac{\pi}{2}$ and thus

  $$\mathrm{D}_u f(x) = -\|\operatorname{grad} f(x)\|,$$

  which is the minimum value for the directional derivative over a unit vector.

Theorem A.3.25 (Inverse mapping differential)**.** Let $x, y \in \mathbf{R}^n$ be any points and consider the map $f: U_x \to V_y$, where $U_x \subseteq \mathbf{R}^n$ is a neighbourhood of x and $V_y \subseteq \mathbf{R}^n$ is a neighbourhood of y — and define $f(x) = y$. Let f be continuous at x and have a continuous inverse mapping $f^{-1}: V_y \to U_x$ at the point y. If the map f is differentiable at x and $\mathrm{d}f(x): T\mathbf{R}^n_x \to T\mathbf{R}^n_y$ has an inverse $\mathrm{d}f(x)^{-1}: T\mathbf{R}^n_y \to T\mathbf{R}^n_x$, then the map f^{-1} is differentiable at y and the differential $\mathrm{d}f^{-1}(y): T\mathbf{R}^n_y \to T\mathbf{R}^n_x$ is such that

$$\mathrm{d}f^{-1}(y) = (\mathrm{d}f(x))^{-1}.$$

Proof. From the continuity of f, for any neighbourhood $V \subseteq V_y$ of y, the preimage $f^{-1}(V) \subseteq U_x$ is open. Let $y + t \in V$ be some element, then there exists $x + h \in U_x$ for which $f(x + h) = y + t$. Moreover, since f is continuous at x, — as $h \to 0$ — we have $t = f(x + h) - y \to 0$ and since f^{-1} is continuous at y, — as $t \to 0$ — we have $h = f^{-1}(y + t) - x$.

From the differentiability of f at x we have that — as $h \to 0$, from Proposition A.3.5

$$t = \mathrm{d}f(x)(h) + o(h) = O(h) + o(h) = O(h). \tag{A.23}$$

Since $\mathrm{d}f(x)$ is invertible, from Eq. (A.23) we find — as $h \to 0$

$$(\mathrm{d}f(x))^{-1}(t) = \left(\mathrm{d}f(x))^{-1} \circ \mathrm{d}f(x)\right)(h) + (\mathrm{d}f(x))^{-1}(o(h)) = h + o(h) \tag{A.24}$$

Where $(\mathrm{d}f(x))^{-1}(o(h)) = o(h)$ comes from Proposition A.3.3. Now let $\delta > 0$ be such that, if $\|h\| < \delta$, then $\|o(h)\| \leqslant \frac{1}{2}\|h\|$. For such δ we find that — since $(\mathrm{d}f(x))^{-1}(t) \geqslant h - o(h)$

$$\|(\mathrm{d}f(x))^{-1}(t)\| = \|h - o(h)\| \geqslant \|h\| - \|o(h)\| \geqslant \frac{1}{2}\|h\|$$

Hence, using the fact that $h = f^{-1}(y+t)+x \to 0$ as $t \to 0$, then the inequality obtained above — that is

$$\|h\| \leqslant 2\|(\mathrm{d}f(x))^{-1}(t)\|$$

together with the fact that $(\mathrm{d}f(x))^{-1}(t) \to 0$ as $t \to 0$ — since linear maps are continuous (see Proposition A.2.37) and vanish at 0 — we conclude that $h = O(t)$. Therefore h and t have the same order over $h, t \to 0$, that is $h \asymp t$ and, equivalently, there exists $a, b > 0$ constants for which exists a neighbourhood $U \subseteq \mathbf{R}^n$ of 0 such that $a\|t\| \leqslant \|h\| \leqslant b\|t\|$ — see Definition A.1.4. This shows that $o(h) = o(t)$. Henceforth, by the use of Eq. (A.24) we find

$$f^{-1}(y + t) + f^{-1}(y) = (\mathrm{d}f(x))^{-1}(t) + o(t).$$

Thus indeed f^{-1} is differentiable at y and $\mathrm{d}f^{-1}(x) = (\mathrm{d}f(x))^{-1}$. ♮

Mean Value Theorem

Remark A.3.26. In what remains of this chapter we'll denote a general domain by $G \subseteq \mathbf{R}^m$. Moreover, if $x, y \in \mathbf{R}^n$, we'll denote by $[x, y]$ the line segment $\gamma([0,1])$ — where $\gamma: [0,1] \to \mathbf{R}^n$ maps $t \overset{\gamma}{\mapsto} (t-1)x + ty$. On the other hand, (x, y) denotes $\gamma((0,1))$.

Theorem A.3.27 (Mean Value Theorem for real valued maps)**.** Let $f\colon E \to \mathbf{R}$ be a map. If f is continuous on the line segment $[x,y]$ and differentiable on (x,y) — where $x, y \in E$ — then there exists $z \in (x,y)$ such that

$$f(y) - f(x) = \mathrm{d}f(z)(y-x).$$

Proof. Define $h(t) = (1-t)x + ty$. Let $g\colon [0,1] \to \mathbf{R}$ defined by $g = f \circ h$. Since h and f are both continuous maps on, respectively, $[0,1]$ and $h([0,1]) = [x,y]$ — together with the fact that the composition of continuous maps yield a continuous map — we find that g is continuous on $[0,1]$. Since h and f are differentiable at, respectively $(0,1)$ and $h((0,1)) = (x,y)$ — by the composition of differentiable maps — we get that g is differentiable on $(0,1)$.

We can now apply the Mean Value Theorem on g to ensure that there exists $\bar{t} \in [0,1]$ such that

$$g(1) - g(0) = g'(\bar{t}).$$

Since $g'(\bar{t}) = [\mathrm{d}f(h(\bar{t}))](h'(\bar{t})) = \mathrm{d}f((1-\bar{t})x + \bar{t}y)(y-x)$, making $z := (1-\bar{t})x + \bar{t}y \in (x,y)$ and substituting the values for g in terms of f we prove the theorem. ♮

Corollary A.3.28 (Constant real valued map)**.** Let $f\colon G \to \mathbf{R}$ be a map — where $G \subseteq \mathbf{R}^m$ is a domain. If f is differentiable at every point of G and its differential vanishes at all points of G, then f is constant on the domain G.

Proof. Since for all $x \in G$ the differential $\mathrm{d}f(x)$ vanishes for all $v \in T\mathbf{R}^m_x$, then $\dim\ker(\mathrm{d}f(x)) = m$ and $\mathrm{d}f(x)$ is the null mapping. Since G is open, the partial derivatives $\partial_j f(x)$ exist — where $1 \leqslant j \leqslant m$ — and $\partial_j f(x) = 0$.

Let $x \in G$ be any point and consider $B_x(r) \subseteq G$ an open ball centred in x. Let $y \in B_x(r)$ be any point, then from Theorem A.3.27 — since $[x,y] \subseteq B_x(r)$ — we have the existence of $z \in [x,y]$ such that

$$f(y) - f(x) = \mathrm{d}f(z)(y-x) = 0, \tag{A.25}$$

that is, $f(y) = f(x)$ for all points of the open ball — the map is constant at any open ball contained in G of every point of G.

Let $x, y \in G$ be any points. Since G is path-connected, there exists a continuous map $\gamma\colon [0,1] \to G$ such that $\gamma(0) = x$ and $\gamma(1) = y$. Consider any open ball $B_x(r) \subseteq G$ centred in x. Let $\Delta = \{\delta \in [0,1] : \gamma([0,\delta]) \in B_x(r)\}$. From the continuity of γ and the fact that $\gamma(0) = x \in B_x(r)$, the collection Δ is non-empty. If $\delta \in \Delta$ — from Eq. (A.25) — we find that $(f \circ \gamma)([0,\delta] = f(x)$. Since Δ is limited, we can define $\varepsilon = \sup\Delta$. From the continuity of γ we have that $(f \circ \gamma)(\varepsilon) = f(x)$. We now show that, in fact, $\varepsilon = 1$. Suppose — for the sake of contradiction — that $\varepsilon < 1$, then there would exist some open ball $B_{\gamma(\varepsilon)}(d)$ such that $f(B_{\gamma(\varepsilon)}(d)) = f(\gamma(\varepsilon)) = f(x)$ and, for instance, $\varepsilon + d \in \Delta$ and $\varepsilon < \varepsilon + d$, which contradicts the hypothesis that $\varepsilon = \sup\Delta$. From this, we conclude that $\varepsilon = 1$ and hence $f(\gamma([0,1])) = f(x) = f(y)$ which closes the proof. ♮

Lemma A.3.29. Let $[a,b] \subseteq \mathbf{R}$ be a closed interval and $f\colon [a,b] \to \mathbf{R}^n$ be a map. If f is continuous on $[a,b]$ and differentiable on (a,b), then there exists $x \in [a,b]$ such that

$$\|f(b) - f(a)\|_{\mathbf{R}^n} \leqslant (b-a)\|\mathrm{d}f(x)\|_{\mathbf{R}^n}.$$

Proof. Let $\phi\colon [a,b] \to \mathbf{R}$ be defined as $\phi(t) = \langle f(b)-f(a), f(t)\rangle_{\mathbf{R}^n}$ — thus ϕ is continuous on $[a,b]$ and differentiable on (a,b). Applying the Mean Value Theorem we ensure the existence of some $x \in (a,b)$ such that

$$\phi(b) - \phi(a) = (b-a)\phi'(x) = (b-a)\langle f(b) - f(a), \mathrm{d}f(x)\rangle_{\mathbf{R}^n}. \tag{A.26}$$

Since

$$\begin{aligned}\phi(b) - \phi(a) &= \langle f(b) - f(a), f(b)\rangle_{\mathbf{R}^n} - \langle f(b) - f(a), f(a)\rangle_{\mathbf{R}^n} \\ &= \langle f(b) - f(a), f(b) - f(a)\rangle_{\mathbf{R}^n} \\ &= \|f(b) - f(a)\|^2_{\mathbf{R}^n} \end{aligned} \tag{A.27}$$

We can substitute Eq. (A.27) into Eq. (A.26) to find

$$\|f(b) - f(a)\|^2_{\mathbf{R}^n} = (b-a)\langle f(b) - f(a), \mathrm{d}f(x)\rangle_{\mathbf{R}^n}.$$

From Hőlder's inequalities (see Proposition A.1.8) we know that $\langle v, u\rangle_{\mathbf{R}^n} \leqslant \|v\|_{\mathbf{R}^n}\|u\|_{\mathbf{R}^n}$. Using such fact and assuming $f(a) \neq f(b)$, we find that

$$\|f(b) - f(a)\|^2_{\mathbf{R}^n} = (b-a)\langle f(b) - f(a), \mathrm{d}f(x)\rangle_{\mathbf{R}^n} \leqslant (b-a)\|f(b) - f(a)\|_{\mathbf{R}^n}\|\mathrm{d}f(x)\|_{\mathbf{R}^n}$$

implies $\|f(b) - f(a)\|_{\mathbf{R}^n} \leqslant (b-a)\|\mathrm{d}f(x)\|_{\mathbf{R}^n}$. ♮

Theorem A.3.30 (Mean Value Theorem for $\mathbf{R}^n$ valued maps)**.** Let $f\colon E \to \mathbf{R}^n$ be a continuous map on the line segment $[x,y]$ and differentiable on (x,y) — with $x, y \in E$. If for all $z \in (x,y)$ and all $v \in T\mathbf{R}^n_z$ we have $\|\mathrm{d}f(z)(v)\|_{\mathbf{R}^n} \leqslant M\|v\|_{\mathbf{R}^n}$ — for some fixed $M \in \mathbf{R}$ — then

$$\|f(y) - f(x)\|_{\mathbf{R}^n} \leqslant M\|y - x\|_{\mathbf{R}^n}.$$

Proof. Let $h\colon [0,1] \to E$ be defined as $h(t) = (1-t)x + ty$, and $g\colon [0,1] \to \mathbf{R}^n$ defined by $g = f \circ h$. Notice that g is the composition of continuous maps on $[0,1]$ and differentiable maps on $(0,1)$, so g inherits both properties. By applying Lemma A.3.29 we ensure the existence of some $\bar{t} \in [0,1]$ such that

$$\|g(1) - g(0)\| \leqslant \|\mathrm{d}g(\bar{t})\| \tag{A.28}$$

From the composition theorem we have that — for any $t \in (0,1)$

$$\mathrm{d}g(t) = [\mathrm{d}f(h(t))](\mathrm{d}h(t)) = [\mathrm{d}f(h(t))](y - x).$$

Since $\|\mathrm{d}f(z)(v)\| \leqslant M\|v\|$, we get

$$\|\mathrm{d}g(t)\| = \|\mathrm{d}f(h(t))(y-x)\| \leqslant M\|y - x\|. \tag{A.29}$$

By substituting Eq. (A.29) into Eq. (A.28) and noting that $g(1) = f(y)$ and $g(0) = f(x)$ we can finally conclude that

$$\|f(y) - f(x)\| \leqslant M\|y - x\|.$$

♮

Corollary A.3.31. Let $f\colon G \to \mathbf{R}^n$ be a differentiable map in G. If $\mathrm{d}f(x)$ is the zero-map — that is, $\mathrm{d}f(x)(y) = 0$ for all $y \in T\mathbf{R}^m_x$ — for all $x \in G$, then f is constant.

Proof. Let $M = 0$ and use Theorem A.3.30. ♮

Sufficient Condition for Differentiability

Theorem A.3.32. Let $x \in \mathbf{R}^m$ be any point and let $U \subseteq \mathbf{R}^m$ be a neighbourhood of x. Let $f\colon U \to \mathbf{R}$ be a map. If for all $1 \leqslant j \leqslant m$ the partial derivatives $\partial_j f$ exist for each point of U, and $\partial_j f$ is continuous at x, then f is differentiable at x.

Proof. Since the collection of open balls in $\mathbf{R}^m$ forms a base for the standard real topology, we may assume that $U = B_x(r)$ for some $r > 0$. Consider a point $x + h \in B_x(r)$, then for all $0 \leqslant k \leqslant m$ we have that $y_k = x + \sum_{j=1}^{k} h_j e_j \in B_x(r)$. Moreover, if $0 \leqslant \ell, k \leqslant m$ then the line segment $[y_\ell, y_k] \subseteq U$. Notice that

$$\begin{aligned} f(x+h) - f(x) &= f(x+h) + \left[\sum_{k=1}^{m-1} f\left(x + \sum_{j=1}^{k} h_i e_j\right) - f\left(x + \sum_{j=1}^{k} h_i e_j\right)\right] - f(x) \\ &= \sum_{k=1}^{m} \left[f\left(x + \sum_{j=1}^{k} h_j e_j\right) - f\left(x + \sum_{j=1}^{k-1} h_j e_j\right)\right] \\ &= \sum_{k=1}^{m} f(y_k) - f(y_{k-1}) \end{aligned} \tag{A.30}$$

Since f has partial derivatives for every point of U, for each $1 \leqslant k \leqslant m$, we can apply the Mean Value Theorem A.3.27 for the pair of points points y_k and y_{k-1}, to find a point $z_k = x + \theta_k h_k e_k$ belonging to the segment line $[y_{k-1}, y_k] \subseteq U$ — with $\theta_j \in [0, 1]$ — such that

$$f(y_k) - f(y_{k-1}) = \partial_k f(z_k) h_k \tag{A.31}$$

Substituting Eq. (A.31) into Eq. (A.30) we get

$$f(x+h) - f(x) = \sum_{k=1}^{m} \partial_k f(z_k) h_k \tag{A.32}$$

Since all partial derivatives are continuous at x, for all $1 \leqslant k \leqslant m$ there exists a map $\alpha_k\colon U \to \mathbf{R}$ such that $\alpha_k(h_k) \to 0$ as $h_k \to 0$ and we can write

$$\partial_k f(z_k) = \partial_k f(x + \theta_k h_k e_k) = \partial_k f(x) + \alpha_k(h_k) \tag{A.33}$$

Substituting Eq. (A.33) into Eq. (A.32) we get the following relation

$$f(x+h) - f(x) = \sum_{k=1}^{m} \partial_k f(x) h_k + \alpha_k(h_k) h_k = \sum_{k=1}^{m} \partial_k f(x) h_k + \sum_{k=1}^{m} \alpha_k(h_k) h_k$$

Notice that as $h \to 0$ we have that $\alpha_k(h_k) h_k \to 0$ for all $1 \leqslant k \leqslant m$, hence $\alpha_k(h_k) h_k = o(h)$ and thus

$$f(x+h) - f(x) = \sum_{k=1}^{m} \partial_k f(x) h_k + o(h)$$

From Definition A.3.7 the map f is differentiable at x — and $\mathrm{d}f(x)(h) = \sum_{k=1}^{m} \partial_k f(x) h_k$. ♮

Definition A.3.33 (Continuously differentiable map)**.** We denote by $C^1(G, \mathbf{R})$ the vector space of real valued continuously differentiable maps $\mathbf{R}^m \supseteq G \to \mathbf{R}$ — that is, with continuous partial derivatives in G.

Higher Order Partial Derivatives

Definition A.3.34 (k-th partial derivative)**.** Let $f\colon G \to \mathbf{R}$ be a map — over a domain $G \subseteq \mathbf{R}^m$ — partially differentiable over its j_1-th variable. If in turn $\partial_{j_1} f\colon G \to \mathbf{R}$ is partially differentiable over its j_2-th variable, we say that $\partial_{j_2}(\partial_{j_1} f) = \partial_{j_2 j_1} f\colon G \to \mathbf{R}$ is said to be the second partial derivative of f with respect to (j_1, j_2) variables. If such property holds k times over variables the $(j_1, j_2, \dots, j_k)$, we say that

$$\partial_{j_k j_{k-1} \dots j_2 j_1} f\colon G \to \mathbf{R}$$

is the k-th partial derivative of f with respect to variables $(j_i)_{1 \leqslant i \leqslant k}$.

Definition A.3.35 (k-continuously differentiable)**.** A map $f\colon G \to \mathbf{R}$ is said to be k-continuously differentiable if for all tuples of indices $I_i = (j_1, \dots, j_i)$ and for all $1 \leqslant i \leqslant k$ the map $\partial_{I_i} f$ exists and is continuous at G.

Lemma A.3.36. Let $E \subseteq \mathbf{R}^m$ be an open set and $f\colon E \to \mathbf{R}$ be a map with existing partial derivatives $\partial_i f$ and $\partial_{ji} f$ at every point of E. Let $y \in E$ and a map $\phi_{ji}\colon V \to \mathbf{R}$ — where $V = \{(t, h) \in \mathbf{R}^2 : y + te_i + he_j \in E\}$ — being defined by

$$\phi_{ji}(t, h) = f(y + te_i + he_j) - f(y + te_i) + f(y + he_j) - f(y).$$

Then for all pairs $(t, h) \in V$ — $t, h \neq 0$ — there exists a point $x = y + \theta_i te_i + \theta_j he_j \in E$, where $\theta_i, \theta_j \in (0, 1)$, such that

$$\phi(t, h) = th[\partial_{21} f(x)].$$

Proof. Since E is open, V is non-empty. Choose any non-zero tuple $(t, h) \in V$. Let the map $\omega\colon [0, 1] \to \mathbf{R}$ be defined by

$$\omega(\theta) = f(y + \theta te_i + he_j) - f(y + \theta te_i).$$

Notice that ω is continuous on $[0, 1]$ and differentiable on $(0, 1)$ — since f has partial derivative $\partial_i f$. We can use the Mean Value Theorem to find $\theta_i \in (0, 1)$ such that

$$\begin{aligned}
\omega(1) - \omega(0) &= \omega'(\theta_i) \\
&= \left[\frac{\partial}{\partial \theta}\left(f(y + \theta te_i + he_j) - f(y + \theta te_i)\right)\right]_{\theta = x_1} \\
&= \left[\theta \partial_i f(y + \theta te_i + he_j) - \theta \partial_i f(y + \theta te_i)\right]_{\theta = \theta_1} && \text{(A.34)} \\
&= t(\partial_1 f(y + \theta_i te_i + he_j) - \partial_1 f(y + \theta_i te_i)) && \text{(A.35)}
\end{aligned}$$

Since $\partial_i f$ is partially differentiable — with respect to its j-th variable — on every point of E, we conclude that $\partial_1 f$ continuous on the line segment $[y, y + te_i + he_j] \subseteq E$ and is differentiable on the line segment $(y, y + te_i + he_j) \subseteq E$. Let a map $\gamma\colon [0,1] \to \mathbf{R}$ defined by

$$\gamma(\theta) = \partial_1 f(y + \theta_i te_i + \theta e_j)$$

which, from $\partial_i f$ properties, is continuous on $[0,1]$ and differentiable on $(0,1)$. We can apply the Mean Value Theorem on γ to find an element $\theta_j \in (0,1)$ for which

$$\begin{aligned} \gamma(1) - \gamma(0) &= \gamma'(\theta_j) \\ &= \left[\frac{\partial}{\partial\theta}\partial_1 f(y + \theta_i te_i + \theta e_j)\right]_{\theta=\theta_j} \\ &= \left[h\partial_{21} f(y + \theta_i te_i + \theta he_j)\right]_{\theta=\theta_j} \\ &= h\partial_{21} f(y + \theta_i te_i + \theta_j he_j) \end{aligned} \tag{A.36}$$

Therefore, substituting Eq. (A.36) into Eq. (A.34) we get

$$\phi(t,h) = \omega(1) - \omega(0) = th[\partial_{21} f(y + \theta_i te_i + \theta_j he_j)].$$

$\natural$

Theorem A.3.37 (Order of the partial derivative variables)**.** Let $f\colon G \to \mathbf{R}$ be a map with partial derivatives $\partial_{ij} f\colon G \to \mathbf{R}$ and $\partial_{ji} f\colon G \to \mathbf{R}$. If $\partial_{ij} f$ and $\partial_{ji} f$ are both continuous over the point $x \in G$, then

$$\partial_{ij} f(x) = \partial_{ji} f(x).$$

Proof. Since G is open, there exists $h \in \mathbf{R}^m$ such that $x + h \in G$ — where we impose that $h_i, h_j \neq 0$. Define the map

$$\Delta(f,h) = [f(x + h_i e_i + h_j e_j) - f(x + h_i e_i)] - [f(x + h_j e_j) - f(x)] \tag{A.37}$$

Define maps $\phi_{ji}, \phi_{ij}\colon [0,1] \to \mathbf{R}$ by

$$\phi_{ji}(\theta) = f(x + \theta h_i e_i + h_j e_j) - f(x + \theta h_i e_i) \tag{A.38}$$

$$\phi_{ij}(\theta) = f(x + h_i e_i + \theta h_j e_j) - f(x + \theta h_j e_j) \tag{A.39}$$

Applying Lemma A.3.36 on Eq. (A.38) we find $\theta_i, \theta_j \in (0,1)$ such that

$$\Delta(f,h) = \phi_{ji}(1) - \phi_{ji}(0) = h_j h_i[\partial_{ji} f(x + \theta_i h_i e_i + \theta_j h_j e_j)] \tag{A.40}$$

Analogously, we can apply Lemma A.3.36 on Eq. (A.39) to find $\overline{\theta}_i, \overline{\theta}_j \in (0,1)$ such that

$$\Delta(f,h) = \phi_{ij}(1) - \phi_{ij}(0) = h_i h_j[\partial_{ij} f(x + \overline{\theta}_i h_i e_i + \overline{\theta}_j h_j e_j)] \tag{A.41}$$

Equating both Eq. (A.40) and Eq. (A.41) we get

$$h_j h_i[\partial_{ji} f(x + \theta_i h_i e_i + \theta_j h_j e_j)] = h_i h_j[\partial_{ij} f(x + \overline{\theta}_i h_i e_i + \overline{\theta}_j h_j e_j)]$$

Since $h_i, h_j \neq 0$, we have

$$\partial_{ji} f(x + \theta_i h_i e_i + \theta_j h_j e_j) = \partial_{ij} f(x + \overline{\theta}_i h_i e_i + \overline{\theta}_j h_j e_j).$$

Since $\partial_{ji}f$ and $\partial_{ij}f$ are continuous at x, as $h \to 0$ we get

$$\partial_{ji}f(x) = \partial_{ij}f(x).$$

♮

Corollary A.3.38. If $f \in C^k(G, \mathbf{R})$, then the maps $\partial_{i_k \dots i_1}f$ for any k-tuple $I_k = (i_1, \dots, i_k)$ — where $1 \leqslant i_j \leqslant m$ for all $1 \leqslant j \leqslant k$ — are the same for any permutation of I_k.

Proof. We proceed by induction on k. If $k = 2$ then the proposition is true, from Theorem A.3.37. Assume the proposition is true for some $k > 2$. Let σ be any permutation on I_k, since any permutation can be written as the composition of finitely many elementary transpositions, it suffices to observe that — from the inductive hypothesis

$$\partial_{I_{k+1}}f = \partial_{i_{k+1}}\left(\partial_{I_k}f\right) = \partial_{i_{k+1}}\left(\partial_{\sigma(I_k)}f\right). \tag{A.42}$$

From Theorem A.3.37 following relation is verified

$$\partial_{I_{k+1}}f = \partial_{i_{k+1}i_k}\left(\partial_{I_{k-1}}f\right) = \partial_{i_k i_{k+1}}\left(\partial_{I_{k-1}}f\right) = \partial_{i_k i_{k+1} i_{k-1} \dots i_1}f. \tag{A.43}$$

Thus, with Eq. (A.42) and Eq. (A.43) we find that the the proposition is true for $k+1$. This finishes the induction proof. ♮

Example A.3.39. Let $f \in C^k(G, \mathbf{R})$ and, for some given $x \in G$, let $h \in \mathbf{R}^m$ be such that $x + h \in G$. Consider the map $\phi\colon [0,1] \to \mathbf{R}$, defined as

$$\phi(t) = f(x + th).$$

Then $\phi \in C^k([0,1], \mathbf{R})$ and its k-th derivative is given by

$$\phi^{(k)}(t) = (h_1\partial_1 + \dots + h_m\partial_m)^k f(x + th).$$

More generally, for all $1 \leqslant j \leqslant k$ we have

$$\phi^{(j)}(t) = (h_1\partial_1 + \dots + h_m\partial_m)^j f(x + th).$$

Definition A.3.40 (Hessian matrix)**.** Let $f\colon E \to \mathbf{R}$ be twice partially differentiable with respect to all variables at the point $x \in \operatorname{Int} E$. The Hessian of f is defined to be the matrix

$$\operatorname{Hess} f(x) = \begin{bmatrix} \partial_{11}f(x) & \dots & \partial_{1m}f(x) \\ \vdots & \ddots & \vdots \\ \partial_{m1}f(x) & \dots & \partial_{mm}f(x) \end{bmatrix}$$

Definition A.3.41. The Hessian of f is defined to be the second order differential of f at the interior point x that is, the multilinear map

$$\mathrm{d}^2 f(x)\colon T\mathbf{R}^m_x \times T\mathbf{R}^m_x \to \mathbf{R}, \text{ mapping } (y, z) \xmapsto{f} \sum_{i,j=1}^{m} \partial_{ij}f(x) y_i z_i$$

Taylor's Formula

Theorem A.3.42 (Taylor's formula)**.** Let $x \in \mathbf{R}^m$ be a point and let $U \subseteq \mathbf{R}^m$ be a neighbourhood of x. Consider a point $h \in \mathbf{R}^m$ such that the line segment $[x, x+h]$ is contained in U. Let $f\colon U \to \mathbf{R}$ be a map $f \in C^{k+1}(U, \mathbf{R})$. Then the following equality holds

$$f(x+h) - f(x) = \sum_{\ell=1}^{k} \frac{1}{\ell!}(h_1\partial_1 + \cdots + h_m\partial_m)^\ell f(x) + r_k(x,h)$$

Where the polynomial term is called Taylor polynomial of order k of f on x, and r_k is the k-th order remainder of f on x — which can be written in the following forms:

(i) Integral form

$$r_k(x,h) = \int_0^1 \frac{(1-t)^k}{k!}(h_1\partial_1 + \cdots + h_m\partial_m)^{k+1} f(x+th)\,\mathrm{d}t.$$

(ii) Lagrange form — for some $\theta \in (0,1)$

$$r_k(x,h) = \frac{1}{(k+1)!}(h_1\partial_1 + \cdots + h_m\partial_m)^{k+1} f(x+\theta h).$$

(iii) Peano form — as $h \to 0$

$$r_k(x,h) = \frac{1}{(k+1)!}(h_1\partial_1 + \cdots + h_m\partial_m)^{k+1} f(x) + o(\|h\|_{\mathbf{R}^m}^{k+1}).$$

and we rewrite the Taylor formula as

$$f(x+h) - f(x) = \sum_{\ell=1}^{k+1} \frac{1}{\ell!}(h_1\partial_1 + \cdots + h_m\partial_m)^\ell f(x) + o(\|h\|_{\mathbf{R}^m}^{k+1}).$$

Proof. Let $\phi\colon [0,1] \to \mathbf{R}$ be defined by $\phi(t) = f(x+th)$. Notice that this implies in $\phi \in C^{k+1}([0,1], \mathbf{R})$ and — using the Taylor's formula for one variable and Example A.3.39 — we find, for $\tau \in [0,1]$

$$\phi(\tau) = f(x+\tau h) = \sum_{\ell=1}^{k} \frac{\phi^{(\ell)}(0)}{\ell!}\tau^\ell + R_k(\tau)$$

Notice that $\phi(1) - \phi(0) = f(x+h) - f(x)$, hence

$$f(x+h) - f(x) = \sum_{\ell=1}^{k} \frac{1}{\ell!}(h_1\partial_1 + \cdots + h_m\partial_m)^\ell f(x) + r_k(x,h).$$

We can now analyse the possible forms for the remainders.

(i) (Integral form) We have R_k given by

$$R_k(\tau) = \int_0^1 \frac{(1-t)^k}{k!}\phi^{(k+1)}(t\tau)t^{k+1}\,\mathrm{d}t.$$

Hence

$$r_k(x,h) = \int_0^1 \frac{(1-t)^k}{k!}(h_1\partial_1 + \cdots + h_m\partial_m)^{k+1} f(x + t\tau h)\tau^{k+1}\,\mathrm{d}t.$$

(ii) (Lagrange form) For some $\theta \in (0,1)$, R_k is given by

$$R_k(\tau) = \frac{1}{(k+1)!}\phi^{(k+1)}(\theta).$$

Hence

$$r_k(x,h) = \frac{1}{(k+1)!}(h_1\partial_1 + \cdots + h_m\partial_m)^{k+1} f(x + \theta h).$$

(iii) (Peano form) From the Lagrange form and as $h \to 0$

$$\begin{aligned} r_k(x,h) &= \frac{1}{(k+1)!}(h_1\partial_1 + \cdots + h_m\partial_m)^{k+1} f(x + \theta h) \\ &= \frac{1}{(k+1)!}(h_1\partial_1 + \cdots + h_m\partial_m)^{k+1} f(x) + o(\|h\|_{\mathbf{R}^m}^{k+1}). \end{aligned}$$

This finishes the proof. ♮

A.4 Extrema on Several Variables

Definition A.4.1 (Extrema points)**.** Let $f\colon E \to \mathbf{R}$ be a map. The map f is said to have a local maximum (or local minimum) at$x_0 \in \operatorname{Int} E$ if there exists a neighbourhood $U \subseteq E$ of x_0 for which $f(x) \leqslant f(x_0)$ (or $f(x) \geqslant f(x_0)$) for all $x \in U$. The local maximum (or minimum) is said to be strict if the strict inequality holds for all $x \in U \setminus x_0$. A local maximum (or minimum) is said to be a local extrema of f.

Theorem A.4.2 (Necessary condition for local extrema)**.** Let $f\colon U \to \mathbf{R}$ be a map defined on the neighbourhood $U \subseteq \mathbf{R}^m$ of a point x_0. Assume that f is partially differentiable at x_0. For x_0 to be a local extrema of f it is **necessary** that for all $1 \leqslant j \leqslant m$ we have $\partial_j f(x_0) = 0$, that is

$$\operatorname{grad} f(x_0) = 0.$$

Proof. Let $1 \leqslant j \leqslant m$ be any index and consider the map $g_j\colon \mathbf{R} \to \mathbf{R}$ given by

$$x \overset{g_j}{\longmapsto} f(x_0^1, \ldots, x_0^{j-1}, x, x_0^j, \ldots, x_0^m).$$

A necessary condition for x_0^j to be an extrema of g_j is that its derivative at x_0^j needs to be zero. Moreover $g_j' = \partial_j f$, thus the necessary condition for g_j implies $\partial_j f(x_0) = 0$. ♮

Definition A.4.3. Let A be a $(n \times n)$-symmetric matrix. We define the following:

- A is said to be positive definite if for all $v \in \mathbf{R}^n \setminus 0$

$$\langle v, Av \rangle > 0.$$

  On the other hand, A is negative definite if for all $v \in \mathbf{R}^n \setminus 0$

$$\langle v, Av \rangle < 0.$$

- A is positive semidefinite if for all $v \in \mathbf{R}^n$

$$\langle v, Av \rangle \geqslant 0.$$

  On the other hand, A is negative semidefinite if for all $v \in \mathbf{R}^n$ we have

$$\langle v, Av \rangle \leqslant 0.$$

Lemma A.4.4. Let A be a symmetric $n \times n$ matrix. A is positive definite if and only if there exists $\lambda > 0$ such that for all $v \in \mathbf{R}^n \setminus 0$

$$\langle v, Av \rangle \geqslant \lambda \|v\|^2.$$

On the other hand, A is negative definite if and only if there exists $\lambda > 0$ such that for all $v \in \mathbf{R}^n \setminus 0$

$$\langle v, Av \rangle \leqslant -\lambda \|v\|^2.$$

Proof. Lets treat only the case for positive definite, the other is analogous and would be boring to repeat myself. Suppose there exists such $\lambda > 0$ for which all $v \in \mathbf{R}^n$ satisfy $\langle v, Av \rangle \geqslant \lambda \|v\|^2$. Then since $v \neq 0$ and hence $\|v\| > 0$, it follows immediately that A is positive definite.

Suppose A is positive definite. Consider the unitary $n-1$-sphere $S^{n-1} = \{v \in \mathbf{R}^n : \|v\| = 1\}$. Since S^{n-1} is closed and bounded, we conclude that S^{n-1} is compact by Theorem A.2.9. Consider the map $f \colon S^{n-1} \to \mathbf{R}$ given by the mapping $v \mapsto \langle v, Av \rangle$ which is continuous. Therefore, from the global properties enunciated at Proposition A.2.36 we find that there exists $v_0 \in S^{n-1}$ where f assumes a minimum value. Let $f(v_0) = \lambda$ and since A is positive definite, we have that λ is necessarily positive. Let $v \in \mathbf{R}^n \setminus 0$ be any element, then from the definition of a minimum, we find $f(\frac{v}{\|v\|}) \geqslant \lambda$. Therefore we conclude that $\langle v, Av \rangle \geqslant \lambda \|v\|^2$, which finishes the proof. ♮

Definition A.4.5 (Critical point)**.** Let $f \colon U \to \mathbf{R}^n$ be a map defined on a neighbourhood $U \subseteq \mathbf{R}^m$ of a point x_0. Assume that f is differentiable at x_0. We say that x_0 is a critical point of f if the rank of the Jacobi matrix $f'(x_0)$ has a rank less than $\min(m, n)$ — where $\min(m, n)$ is the maximum possible value of the rank.

Theorem A.4.6 (Classification of critical points of real valued maps)**.** Let $f \colon U \to \mathbf{R}$ be twice continuously differentiable and let $x_0 \in U \subseteq \mathbf{R}^m$ be an internal point such that $\operatorname{grad} f(x_0) = 0$. If the matrix $\operatorname{Hess} f(x_0)$:

1. is positive definite, then x_0 is a local minimum of f.
2. is negative definite, then x_0 is a local maximum of f.
3. is indefinite, then x_0 is not an extremum point of f.

Proof. Let $h \in T_{x_0}\mathbf{R}^m$ and consider the Taylor polynomial of f of order 2 at the point x_0 in the Peano form:

$$f(x_0 + h) - f(x_0) = \sum_{\ell=1}^{2} \frac{1}{\ell!}(h_1\partial_1 + \cdots + h_m\partial_m)^\ell f(x_0) + o(\|h\|^2).$$

Notice however that for x_0 to be an extremum point candidate of f we have from Theorem A.4.2 that $\operatorname{grad} f(x_0) = 0$ thus the Taylor polynomial only has its second order factor. Moreover, we can rewrite the second order term as an inner product of $\operatorname{Hess} f(x_0)$ and h in the following manner

$$\sum_{i,j=1}^{m} \partial_{ij} f(x_0) h_i h_j = \langle h, \operatorname{Hess}(f(x_0))h \rangle .$$

Thus we can now rewrite the Taylor polynomial in a way that lets us analyse the hessian of the map

$$f(x_0 + h) - f(x_0) = \langle h, \operatorname{Hess}(f(x_0))h \rangle + o(\|h\|^2).$$

Suppose $\operatorname{Hess} f(x_0)$ is positive definite, then using Lemma A.4.4 we find some $\lambda > 0$ such that for all $v \in \mathbf{R}^n$ we have $\langle v, \operatorname{Hess}(f(x_0))v \rangle \geqslant \lambda \|v\|^2$. In particular, we can choose $\delta > 0$ for which $|o(\|h\|^2)| \leqslant \frac{\lambda}{4}\|h\|^2$ for all $h \in T_{x_0}\mathbf{R}^n$ such that $\|h\| \leqslant \delta$. Then for all $\|h\| \leqslant \delta$ we have

$$f(x_0 + h) - f(x_0) = \frac{1}{2} \langle h, \operatorname{Hess}(f(x_0))h \rangle + o(\|h\|^2) \geqslant \frac{1}{2}\lambda \|h\|^2 - \frac{\lambda}{4}\|h\|^2 \geqslant \frac{\lambda}{4}\|h\|^2$$

And since $\frac{\lambda}{4}\|h\|^2 > 0$ then $f(x_0 + h) - f(x_0) > 0$ and therefore — at least in the neighbourhood $B_{x_0}(\delta)$ — we are ensured that x_0 is a local minimum of f.

If on the other hand we have $\operatorname{Hess} f(x_0)$ negative definite, then $\operatorname{Hess}(-f(x_0))$ is positive definite and the proposition follows.

Suppose $\operatorname{Hess} f(x_0)$ is indefinite and let $u \in S^{n-1}$ be the point where the mapping given by $\ell \mapsto \langle \ell, \operatorname{Hess}(f(x_0))\ell \rangle$ assumes its minimum $m < 0$ and $v \in S^{n-1}$ be the point where the map assumes it maximum $M > 0$. Define $t > 0$ so that $x_0 + ut \in U$, then

$$f(x_0 + ut) - f(x_0) = \langle ut, \operatorname{Hess}(f(x_0))ut \rangle + o(\|ut\|^2) = \frac{1}{2!} m t^2 + o(t^2)$$

Then, for $t \to 0$ we'll have $f(x_0 + tu) - f(x_0) < 0$ in some neighbourhood of x_0, since $m < 0$ — this implies that x_0 is a local minimum of such neighbourhood. Moreover, if we now set $t > 0$ so that $x_0 + vt \in U$, it follows that

$$f(x_0 + vt) - f(x_0) = \frac{1}{2!} \langle vt, \operatorname{Hess}(f(x_0))vt \rangle + o(\|vt\|^2) = \frac{1}{2!} M t^2 + o(t^2)$$

so that, as $t \to 0$ we have $f(x_0 + tv) - f(x_0) > 0$ for some neighbourhood of x_0, since $M > 0$ — which now implies that x_0 is in fact a local maximum in some neighbourhood, which is a direct contradiction to the assertion that x_0 was a local minimum. This shows us that for $\operatorname{Hess} f(x_0)$ indefinite, x_0 is not an extremum of f. ♮

A.5 Implicit Map Theorem

Definition A.5.1 (Level curve)**.** Let $f\colon \Omega \to \mathbf{R}$, where $\Omega \subseteq \mathbf{R}^n$, be any map. For every $c \in \mathbf{R}$ we define the set $N_f(c) = \{x \in \Omega : f(x) = c\}$ to be the c-level curve of f. Moreover, if $n = 2$, we can call $N_f(c)$ the c-contour line of f with value c.

Theorem A.5.2 (Implicit Theorem)**.** Let V and L be normed vector spaces and W be a Banach space. Define $\Omega \subseteq V \times W$ to be an open set and $(x_0, y_0) \in \Omega$. Consider $F\colon \Omega \to L$ to be a mapping such that

- $F(x_0, y_0) = c$ for some $c \in L$.
- F is continuous at (x_0, y_0).
- F is differentiable and its differential $\mathrm{d}F\colon \Omega \to L$ is continuous at (x_0, y_0).
- $\partial_2 F(x_0, y_0)\colon W \to L$ is an isomorphism, that is, it is invertible.

Then there exists neighbourhoods $U_{x_0} \subseteq V$ and $U_{y_0} \subseteq W$ such that $U_{x_0} \times U_{y_0} \subseteq \Omega$, and a map $f\colon U_{x_0} \to U_{y_0}$ such that

- $f(x_0) = y_0$.
- f is continuous at x_0.
- $F(x, y) = 0$ if and only if $f(x) = y$, for $x, y \in U_{x_0} \times U_{y_0}$.

Proof. To ease our lives, lets assume that Ω has the following form

$$\Omega = \{(x, y) \in V \times W : \|x - x_0\|_V < \alpha \text{ and } \|y - y_0\|_W < \beta\}.$$

If that is not the case, since Ω is open — and hence (x_0, y_0) is an internal point — we can merely choose an open set contained in Ω that satisfies the above form.

Define a collection $\{g_x : x \in B_{x_0}(\alpha)\}$ of maps

$$g_x(y) = y - [\partial_2 F(x_0, y_0)]^{-1}(F(x, y))$$

The domain of each g_x is defined to be the collection $B_{y_0}(\beta) = \{y \in W : \|y - y_0\|_W < \beta\}$. The maps g_x are well defined since $[\partial_2 F(x_0, y_0)]^{-1}$ exists and is a continuous linear map — moreover, the domain of g_x is the normed vector space W — that is $g_x\colon B_{y_0}(\beta) \to W$.

Suppose y_x is a fixed point of g_x, then $g_x(y_x) = y_x - [\partial_2 F(x_0, y_0)]^{-1}(F(x, y_x)) = y_x$ hence clearly y_x is indeed a fixed point of g_x if and only if $[\partial_2 F(x_0, y_0)]^{-1}(F(x, y_x)) = 0$, that is, $F(x, y_x) \in \ker[\partial_2 F(x_0, y_0)]^{-1}$ — but $\partial_2 F(x_0, y_0)$ is an isomorphism, so clearly $F(x, y_x) = 0$.

Let $x \in B_{x_0}(\alpha)$ be any element and consider the map g_x. Notice that since F is differentiable and g_x is therefore a composition of differentiable maps, it follows that

g_x is differentiable and since $\partial_2 F(x_0, y_0)$ is continuous and linear we get

$$\begin{aligned}
\partial_2 \left([\partial_2 F(x_0, y_0)]^{-1}(F(x,y)) \right) &= \lim_{t \to 0} \frac{[\partial_2 F(x_0, y_0)]^{-1}(F(x, y+t)) - [\partial_2 F(x_0, y_0)]^{-1}(F(x,y))}{t} \\
&= \lim_{t \to 0} \frac{[\partial_2 F(x_0, y_0)]^{-1}(F(x, y+t) - F(x,y))}{t} \\
&= \lim_{t \to 0} [\partial_2 F(x_0, y_0)]^{-1} \left(\frac{F(x, y+t) - F(x,y)}{t} \right) \\
&= [\partial_2 F(x_0, y_0)]^{-1} \left(\lim_{t \to 0} \frac{F(x, y+t) - F(x,y)}{t} \right) \\
&= [\partial_2 F(x_0, y_0)]^{-1} (\partial_2 F(x,y))
\end{aligned}$$

Therefore we can write the differential of g_x as

$$\mathrm{d}g_x(y) = 1_W - \partial_2 \left([\partial_2 F(x_0, y_0)]^{-1}(F(x,y)) \right) = [\partial_2 F(x_0, y_0)]^{-1}(\partial_2 F(x_0, y_0) - \partial_2 F(x,y)).$$

By the continuity of the map $\partial_2 F(x_0, y_0)$ at the point (x_0, y_0), we find that there exists $0 < \gamma < \min(\alpha, \beta)$ such that in the neighbourhood $B_{x_0}(\gamma) \times B_{y_0}(\gamma) \subseteq \Omega$ we have

$$\|\mathrm{d}g_x(y)\| \leqslant \|[\partial_2 F(x_0, y_0)]^{-1}\| \cdot \|\partial_2 F(x_0, y_0) - \partial_2 F(x,y)\| < \frac{1}{2}.$$

Let $x \in B_{x_0}(\gamma)$ be any element and take any two $y_1, y_2 \in B_{y_0}(\gamma)$. Then we have by the generalization of the mean value theorem that

$$\|g_x(y_1) - g_x(y_2)\|_W \leqslant \sup_{t \in (y_1, y_2)} \|\mathrm{d}g(t)\| \|y_1 - y_2\|_W < \frac{1}{2}\|y_1 - y_2\|.$$

That is, g_x is Lipschitz continuous.

From the definition of g_x and the fact that $g_{x_0}(y_0) = y_0$ — since $F(x_0, y_0) = 0$ — we have

$$\begin{aligned}
\|g_x(y) - y_0\|_W &= \|g_x(y) - g_{x_0}(y_0)\|_W \\
&\leqslant \|g_x(y) - g_x(y_0)\|_W + \|g_x(y_0) - g_{x_0}(y_0)\|_W \\
&\leqslant \frac{1}{2}\|y - y_0\|_W + \|[\partial_2 F(x_0, y_0)]^{-1}(F(x, y_0) - F(x_0, y_0))\|_W \\
&= \frac{1}{2}\|y - y_0\|_W + \|[\partial_2 F(x_0, y_0)]^{-1}(F(x, y_0))\|_W
\end{aligned}$$

Since F is continuous at (x_0, y_0) — so is the projection map $x \mapsto F(x, y_0)$ — then for all $\varepsilon' \in (0, \gamma)$ there exists $\delta \in (0, \gamma)$ such that $\|x - x_0\|_V < \delta$ implies $\|F(x, y_0) - F(x_0, y_0)\|_L = \|F(x, y_0)\|_L < \varepsilon'$. In particular, choose $\varepsilon' = \frac{\varepsilon}{2\|[\partial_2 F(x_0, y_0)]^{-1}\|}$ for any given $\varepsilon > 0$, then for all $\|x - x_0\|_V < \delta$ and for all $\|y - y_0\|_W \leqslant \varepsilon$ we have — using the property that the norm of linear maps is sub-multiplicative

$$
\begin{aligned}
\|g_x(y) - y_0\|_W &\leqslant \frac{1}{2}\|y - y_0\|_W + \|[\partial_2 F(x_0, y_0)]^{-1}(F(x, y_0))\|_W \\
&\leqslant \frac{1}{2}\varepsilon + \|[\partial_2 F(x_0, y_0)]^{-1}\| \cdot \|F(x, y_0)\|_W \\
&\leqslant \frac{1}{2}\varepsilon + \|[\partial_2 F(x_0, y_0)]^{-1}\| \frac{\varepsilon}{2\|[\partial_2 F(x_0, y_0)]^{-1}\|} \\
&= \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon
\end{aligned}
$$

The relation above is equivalent to: for all $\|x - x_0\|_V < \delta$ we have

$$g_x(\mathrm{Cl}(B_{y_0}(\varepsilon))) \subseteq B_{y_0}(\varepsilon)$$

Since $B_{y_0}(\varepsilon) \subseteq W$ is a closed set, it follows from Theorem A.1.17 that there exists a unique fixed point $y_x \in \mathrm{Cl}(B_{y_0}(\varepsilon))$ of g_x. Define now the map $f : B_{x_0}(\delta) \to B_{y_0}(\varepsilon)$ where $f(x) = y_x$ and y_x is the corresponding fixed point of each g_x.

From construction $B_{x_0}(\delta) \times B_{y_0}(\varepsilon) \subseteq \Omega$. Clearly, for all $x \in B_{x_0}(\delta)$ and $y \in B_{y_0}(\varepsilon)$, $F(x, y) = 0$ if and only if $f(x) = y$. Moreover, we have immediately that $f(x_0) = y_0$. For the continuity of f at the point x_0, we can observe that for all $\varepsilon \in (0, \gamma)$ there exists a $\delta \in (0, \gamma)$ such that $\|g_x(y_x) - y_0\| = \|f(x) - y_0\| < \varepsilon$, thus we are finally done. ♮

We'll now extend the Implicit Map Theorem for cases 3 where we have more special conditions, allowing for additional properties for the implicit map f.

Lemma A.5.3 (Continuity of the implicit map)**.** Let F satisfy the properties described in Theorem A.5.2 and additionally suppose that there exists a neighbourhood of (x_0, y_0) where the map $\partial_2 F(x_0, y_0) : W \to L$ is continuous. Then the implicit map $f : U \to V$ is such that there exists a neighbourhood of x_0 such that f is continuous.

Proof. ♮

Lemma A.5.4 (Differentiability of the implicit map)**.** Let F satisfy the properties described in Theorem A.5.2 and additionally suppose that the partial derivative $\partial_1 F(x, y) : V \to L$ exists in some neighbourhood of the point (x_0, y_0) and is continuous at (x_0, y_0). Then the implicit map f is differentiable at x_0 and

$$\mathrm{d}f(x_0) = -[\partial_2 F(x_0, y_0)]^{-1} \partial_1 F(x_0, y_0).$$

Proof. ♮

Lemma A.5.5 (Continuous differentiability of the implicit map)**.** Let F satisfy the properties described in Theorem A.5.2 and additionally suppose that the partial derivatives of F are continuous in some neighbourhood of (x_0, y_0). Then the map f is continuously differentiable in some neighbourhood of x_0 and its differential in this neighbourhood is given by

$$\mathrm{d}f(x) = -[\partial_1 F(x, f(x))]^{-1} \partial_1 F(x, f(x)).$$

Proof. ♮

Lemma A.5.6 (C^k implicit map)**.** Let F satisfy the properties described in Theorem A.5.2 and additionally suppose that $F \in C^k(\Omega, L)$. Then the implicit map f is a member of the class $C^k(U, W)$ in some neighbourhood $U \subseteq V$ of x_0.

Proof. ♮

Write down the proof of the extensions of the basic Implicit Map Theorem

Corollaries of the Implicit Map Theorem

Inverse Map Theorem

Definition A.5.7 (Diffeomorphisms)**.** Let U and V be open subsets of $\mathbf{R}^m$. A map $f: U \to V$ is said to be a isomorphism of manifolds of class C^p (or diffeomorphism) — for $p \in \mathbf{N} \cup \{\infty\}$ — if the following conditions are satisfied

- $f \in C^p(U, V)$.
- f is bijective and $f^{-1} \in C^p(V, U)$.

Theorem A.5.8 (Inverse Map Theorem)**.** Let V and W be Banach spaces and $\Omega \subseteq V$ be an open set. Consider a point $x_0 \in \Omega$ and a map $f: \Omega \to W$ such that the following conditions are satisfied

- $f \in C^1(\Omega, W)$.
- $\mathrm{d}f(x_0)$ is invertible and $[\mathrm{d}f(x_0)]^{-1}$ is a continuous map.

Then there exists an open neighbourhood $X \subseteq \Omega$ of x_0 and an open neighbourhood $Y \subseteq W$ of the point $y_0 = f(x_0)$, for which the restriction $f: X \to Y$ is bijective. The inverse map $f^{-1}: Y \to X$ is differentiable and its differential is given by

$$\mathrm{d}f^{-1}(y_0) = [\mathrm{d}f(x_0)]^{-1}.$$

Proof. Let $F: N \to W$ be a map defined on $N \subseteq V \times W$ where N is a neighbourhood of (x_0, y_0) and let $F(x, y) = f(x) - y$. Since F is the composition of the restriction of maps that are continuously differentiable, it follows that F is continuously differentiable, moreover, $\partial_1 F(x_0, y_0) = \mathrm{d}f(x_0)$. Moreover, since $\mathrm{d}f(x_0)$ is invertible then $\partial_1 F(x_0, y_0)$ is also invertible. We have $F(x_0, y_0) = 0$ by construction, since $f(x_0) = y_0$.

We can now see that F satisfies the requirements for the Implicit Map Theorem, thus there exists a neighbourhood Y of y_0 and a continuously differentiable map $g: Y \to V$ (where we use extension Lemma A.5.6) for which $g(Y)$ is contained in a neighbourhood $X' \subseteq V$ of x_0. Moreover, $F(x, y) = 0$ if and only if $g(y) = x$, that is, $F(g(y), y) = 0$ and therefore $fg(y) = y$ for any $y \in Y$, that is, g is injective on Y — also $g(y_0) = x_0$. The map g has a differential given by (using the extension Lemma A.5.4)

$$\mathrm{d}g(y) = [\partial_1 F(x, y)]^{-1}[\partial_2 F(x, y)], \text{ for all } (x, y) \in X' \times Y.$$

From the definition of F we find that

$$\mathrm{d}g(y) = [\mathrm{d}f(x)]^{-1}, \text{ for all } (x, y) \in X' \times Y.$$

Lets consider the restriction $f\colon g(Y) \to W$. Since g is injective, the restriction $g\colon Y \to g(Y)$ is a bijection. Since f is continuous and Y is open, then $f^{-1}(Y) = g(Y)$ is open. Define $X = g(Y)$, so that $f\colon X \to Y$ is a bijection and clearly $g = f^{-1}$ for such restriction. Hence we conclude that

$$\mathrm{d}f^{-1}(y) = [\mathrm{d}f(x)]^{-1}, \text{ for all } (x, y) \in X \times Y.$$

$\natural$

Theorem A.5.9 (Open Map Theorem)**.** Let $\Omega \subseteq \mathbf{R}^n$ be an open set and $f\colon \Omega \to \mathbf{R}^n$ be a continuously differentiable map. If $\mathrm{d}f(x)$ is invertible for all $x \in \Omega$, then the map f is an open mapping — that is, maps open subsets of Ω to open subsets of $\mathbf{R}^n$.

Proof. Let $x \in \Omega$ be any point. From hypothesis, $\mathrm{d}f(x)$ is invertible, hence we can apply Theorem A.5.8 in order to obtain an open neighbourhood $V_x \subseteq \Omega$ of x and $V_{f(x)} \subseteq \mathbf{R}^n$ such that the map $f\colon V_x \to V_{f(x)}$ is a local bijection and hence $f(V_x)$ is open. With this in our hands, we can create an open cover $\mathcal{U} = \{V_x \subseteq \Omega : x \in \Omega\}$ of such neighbourhoods — that is, given any open set $U \subseteq \Omega$, there exists a collection of open sets $\mathcal{V} \subseteq \mathcal{U}$ such that $U = \bigcup_{V \in \mathcal{V}} V$ and since $f(U) = \bigcup_{V \in \mathcal{V}} f(V)$ is the union of open sets, then $f(U)$ is open. $\natural$

Theorem A.5.10 (Maximal Rank Theorem)**.** Let $\Omega \subseteq \mathbf{R}^n$ be an open set and $x_0 \in \Omega$. Let $f\colon \Omega \to \mathbf{R}^m$ be a continuously differentiable map. Define $y_0 \in \mathbf{R}^m$ so that $f(x_0) = y_0$. The following holds

(a). Suppose $n \leqslant m$ and that $\mathrm{d}f(x_0)$ has maximal $\operatorname{rank}(\mathrm{d}f(x_0)) = n$. Then there exists open sets $\Omega_{y_0} \subseteq \mathbf{R}^m$ and $\Omega_{x_0} \subseteq \Omega \subseteq \mathbf{R}^n$, respectively neighbourhoods of the points y_0 and x_0 with $f(\Omega_{x_0}) \subseteq \Omega_{y_0}$, and a differentiable map $g\colon \Omega_{y_0} \to \mathbf{R}^m$ such that the following diagram commutes

$$\begin{tikzcd}
\Omega_{x_0} \arrow[rr, "f"] \arrow[dr, "\iota"'] & & \Omega_{y_0} \arrow[dl, "g"] \\
& \mathbf{R}^m &
\end{tikzcd}$$

Where $\iota\colon \mathbf{R}^n \hookrightarrow \mathbf{R}^m$ is the canonical inclusion map.

(b). Suppose $n \geqslant m$ and that $\mathrm{d}f(x_0)$ has maximal $\operatorname{rank}(\mathrm{d}f(x_0)) = m$. Then there exists $\Omega_{x_0} \subseteq \Omega \subseteq \mathbf{R}^n$, neighbourhood of x_0, and a differentiable map $g\colon \Omega_{x_0} \to \Omega$ such that the following diagram commutes

$$\begin{tikzcd}
& \Omega_{x_0} \arrow[dl, "g"'] \arrow[dr, two heads, "\pi"] & \\
\Omega \arrow[rr, "f"] & & \mathbf{R}^m
\end{tikzcd}$$

Where $\pi\colon \mathbf{R}^n \twoheadrightarrow \mathbf{R}^m$ is the canonical projection map.

Proof. 1. Since $\operatorname{rank}(\mathrm{d}f(x_0)) = n$, then, from the rank plus nullity theorem we find that $\ker(\mathrm{d}f(x_0)) = 0$ and therefore $\mathrm{d}f(x_0)$ is injective. Consider the matrix representation $f'(x_0)$ of the differential $\mathrm{d}f(x_0)$. From the injective property of f, there must exist n linearly independent rows in $f'(x_0)$. Let A be the $(n \times n)$-matrix containing these linearly independent rows and C the $((m-n)\times n)$-matrix containing the remaining rows of $f'(x_0)$. Do row operations on $f'(x_0)$ so that we see it as equivalent to the matrix

$$\begin{bmatrix} A \\ C \end{bmatrix}$$

Notice that the collection of rows of A form a basis for $\mathbf{R}^n$, thus A is invertible and hence $\det A \neq 0$. Define a map $F\colon \Omega \times \mathbf{R}^{m-n} \to \mathbf{R}^m$ given by the mapping $(x, y) \mapsto f(x) + (0, y)$, then we obtain

$$F'(x_0, 0) = \begin{bmatrix} A & 0 \\ C & I_{m-n} \end{bmatrix}$$

which in particular makes the, otherwise dependent, rows of C into a collection of linearly independent vectors, by attaching the canonical base of the space $\mathbf{R}^{m-n}$ into each of them. This makes $F'(x_0, 0)$ an invertible matrix. By applying Theorem A.5.8 we obtain a neighbourhood $U \subseteq \Omega \times \mathbf{R}^{m-n}$ of $(x_0, 0)$ and a neighbourhood $\Omega_{y_0} \subseteq \mathbf{R}^m$ of $F(x_0, 0) = f(x_0) = y_0$ for which the restriction map $F\colon U \to \Omega_{y_0}$ is an isomorphism of manifolds. Let $g\colon \Omega_{y_0} \to U$ be the continuously differentiable inverse of F, and define $\Omega_{x_0} = f^{-1}(\Omega_{y_0}) \cap \Omega$, which is clearly a neighbourhood of x_0. Notice that the composition $gf(x) = gF(x, 0) = (x, 0) = \iota(x)$, thus we are done.

2. Since $\operatorname{rank}(\mathrm{d}f(x_0))$ equals the dimension of its codomain, it follows that $\mathrm{d}f(x_0)$ is a surjective linear map. Since $\mathrm{d}f(x_0)$ has rank m, then its matrix representation $f'(x_0)$ has m linearly independent columns. Let D be the $(m \times m)$-matrix whose columns are those of $f'(x_0)$ that are linearly independent and C be the $(m \times (n - m))$-matrix composed of the remaining columns of $f'(x_0)$. Operate on the matrix $f'(x_0)$ via column operations so that its final equivalent matrix is

$$\begin{bmatrix} D & C \end{bmatrix}$$

Since D is composed of linearly independent vectors, D is invertible. Define the map $F\colon \Omega \to \mathbf{R}^m \times \mathbf{R}^{n-m}$ by $(x, y) \mapsto (f(x), y)$. Clearly, F is differentiable at x_0 and its matrix representation is

$$F'(x_0) = \begin{bmatrix} 0 & I_{n-m} \\ D & C \end{bmatrix}$$

which is invertible since the attachment of the canonical basis of $\mathbf{R}^{n-m}$ into the column vectors of C transforms the collection of the last $n - m$ column vectors of $F'(x_0)$ into a linearly independent set. Applying Theorem A.5.8 we are able to obtain a neighbourhood $\Omega_{x_0} \subseteq \Omega$ such that the restriction map $F\colon \Omega_{x_0} \to \mathbf{R}^m \times$

$\mathbf{R}^{n-m}$ is an isomorphism of manifolds. Let $g\colon \mathbf{R}^m \times \mathbf{R}^{n-m} \to \Omega_{x_0}$ be its continuously differentiable inverse map, then the composition $fg(x,y) = x = \pi(x,y)$ is merely the canonical projection map, as we expected.

♮

Theorem A.5.11 (Constant rank)**.** Let $U \subseteq \mathbf{R}^n$ be an open set, and $f\colon U \to \mathbf{R}^m$ be a C^∞-map with a locally constant rank k at a neighbourhood $V \subseteq U$ of a point $p \in U$. Then after a possible permutation of coordinates near p and fp, the map f assumes the form

$$f|_U = (f_1, \dots, f_k, 0, \dots, 0).$$

In other words, there exists a C^∞-isomorphism $\phi\colon Q \to \mathbf{R}^n$ centred at p (that is, with $\phi p = 0$), and a C^∞-isomorphism $\psi\colon W \to \mathbf{R}^m$—for some neighbourhood $W \subseteq \mathbf{R}^m$ of fp—centred at fp such that

$$\psi f \phi^{-1}(x_1, \dots, x_n) = (x_1, \dots, x_k, 0, \dots, 0) \in \mathbf{R}^m. \tag{A.44}$$

Proof. Since f has a locally constant rank k at p, the local representation of the Jacobian matrix of f at V can be rearranged so that

$$\det\left[\partial_j f_i x\right]_{1 \leqslant i,j \leqslant k} \neq 0$$

for any $x \in V$. Define $\phi\colon U \to \mathbf{R}^n$ to be the map $\phi \coloneqq (f_1, \dots, f_k, \pi_{k+1}, \dots, \pi_n)$, so that one has a local Jacobian matrix at V given by

$$\operatorname{Jac}\phi = \begin{bmatrix} \partial_1 f_1 & \dots & \partial_k f_1 & \partial_{k+1} f_1 & \dots & \partial_n f_1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_k & \dots & \partial_k f_k & \partial_{k+1} f_k & \dots & \partial_n f_k \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{bmatrix}$$

Then $\det(\operatorname{Jac}\phi x) = \det\left[\partial_j f_i x\right]_{1 \leqslant i,j \leqslant k}$ is non-zero for every $x \in V$. Using the inverse map theorem, ϕ is a C^∞-isomorphism $\phi\colon U_p \to U_{\phi p}$ for some neighbourhoods $U_p \subseteq V$ of p and $U_{\phi p} \subseteq \mathbf{R}^n$ of ϕp, respectively.

Consider the map $g \coloneqq f\phi^{-1}\colon U_{\phi p} \to \mathbf{R}^m$. Since ϕ^{-1} has a constant maximal rank n at $U_{\phi p}$, and $f\colon U_p \to \mathbf{R}^m$ has constant rank k at U_p, it follows that g has constant rank k at $U_{\phi p}$. By construction, the Jacobian matrix of g is

$$\operatorname{Jac} g = \begin{bmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ \partial_1 g_{k+1} & \dots & \partial_k g_{k+1} & \partial_{k+1} g_{k+1} & \dots & \partial_n g_{k+1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial_1 g_m & \dots & \partial_k g_m & \partial_{k+1} g_m & \dots & \partial_n g_m \end{bmatrix}$$

Therefore in the open set $U_{\phi p}$ one has $\partial_j g_i = 0$ for every $k+1 \leqslant i \leqslant m$ and $k+1 \leqslant j \leqslant n$. Restricting the neighbourhood $U_{\phi p}$ to some inner open ball about ϕp—so that the neighbourhood is a convex set—we see that g_i is independent of the last $m-k$ variables, for each $k+1 \leqslant i \leqslant m$.

Define a C^∞-morphism $\psi\colon U_{\phi p} \to \mathbf{R}^m$ by making $\psi_i := \pi_i$ for each $1 \leqslant i \leqslant k$, and $\psi_i := \pi_i - g_i$ for each $k+1 \leqslant i \leqslant m$. The Jacobi matrix of ψ has the form

$$\operatorname{Jac}\psi = \begin{bmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ -\partial_1 g_{k+1} & \dots & -\partial_k g_{k+1} & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\partial_1 g_m & \dots & -\partial_k g_m & 0 & \dots & 1 \end{bmatrix}$$

Which has $\det(\operatorname{Jac}\psi) = 1$—hence ψ is a C^∞-isomorphism $\psi\colon U_{fp} \to U_{\psi fp}$ for some neighbourhoods $U_{fp}, U_{\psi fp} \subseteq \mathbf{R}^m$ of fp and ψfp, respectively. Considering the neighbourhood $g^{-1}U_{fp}$ we find that the map

$$\psi f \phi^{-1}\colon g^{-1}U_{fp} \longrightarrow \mathbf{R}^m$$

is a C^∞-morphism. Defining $Q := \phi^{-1}U_{\phi p}$ and $W := \psi^{-1}U_{\psi fp}$ we obtain induced C^∞-isomorphisms $\phi\colon Q \to \phi Q$ and $\psi\colon W \to \psi W$ such that $\psi f \phi^{-1}$ satisfies the required Eq. (A.44). ♮

A.6 Extrema With Constraints

Theorem A.6.1 (Existence of Lagrange multipliers)**.** Let $f\colon \Omega \to \mathbf{R}$ be a map, where $\Omega \subseteq \mathbf{R}^d$, and a map $F\colon \Omega \to \mathbf{R}^m$, where $m < d$ — which will be called constraint map. Suppose x_0 is a local extremum of f in the surface $S \subseteq \Omega$ defined by the constraint $S = \{x \in \Omega : F(x) = 0\}$. Suppose additionally that $\operatorname{rank} F'(x) = m$ for every $x \in \Omega$. Then, there exists a vector $\lambda \in \mathbf{R}^m$ — whose components are called Lagrange multipliers — such that

$$f'(x_0) = F'(x_0)\lambda.$$

In the particular case where $m = 1$, the constraint map generates a 1 dimensional surface and therefore $\operatorname{grad} f(x_0) = \lambda \operatorname{grad} F(x_0)$, where $\lambda \in \mathbf{R}$.

Proof. Since $F'(x_0)$ has rank m, let D be the $(m \times m)$-matrix whose columns are the m linearly independent columns of $F'(x_0)$. Define C to be the $(m \times d - m)$-matrix whose columns are the remaining columns of $F'(x_0)$. By means of column operations, arrange $F'(x_0)$ into an equivalent matrix of the form

$$\begin{bmatrix} C & D \end{bmatrix}$$

That is, the equivalent matrix has an invertible principal minor D defined on its last m columns. If we now identify points $x \in \mathbf{R}^d$ with points $(x_1, x_2) = x \in \mathbf{R}^{d-m} \times \mathbf{R}^m$, we find that $\partial_2 F(x_0^1, x_0^2)$ is an invertible map (it corresponds to the matrix D).

We can now apply the Implicit Map Theorem to find neighbourhoods $\Omega_{x_0^1} \subseteq \mathbf{R}^{d-m}$ and $\Omega_{x_0^2} \subseteq \mathbf{R}^m$ of x_0^1 and x_0^2, respectively, and a map $g \colon \Omega_{x_0^1} \to \Omega_{x_0^2}$ for which $F(x_1, x_2) = 0$ if and only if $g(x_1) = x_2$. In particular, the mapping $x_1 \mapsto F(x_1, g(x_1))$ is identically zero and therefore $\Omega_{x_0^1} \times g(\Omega_{x_0^1}) \subseteq S$. From the fact that $x_0 = (x_0^1, x_0^2) = (x_0^1, g(x_0^1))$ is a local extremum of f, defining $\phi \colon \Omega_{x_0^1} \to \mathbf{R}$ by $\phi(x_1) = f(x_1, g(x_1))$, we conclude that x_0^1 is a local maximum of ϕ. In particular, it is necessary that $\mathrm{d}\phi(x_0^1) = 0$ and therefore

$$\mathrm{d}\phi(x_0^1) = \sum_{j=1}^{d-m} \partial_j f(x_0^1, g(x_0^1)) + \sum_{i=d-m+1}^{m} \partial_i f(x_0^1, g(x_0^1)) \partial_i g(x_0^1) = 0.$$

If we now regard f as map of the form $\mathbf{R}^{d-m} \times \mathbf{R}^m \to \mathbf{R}^m$, then we see that

$$\mathrm{d}\phi(x_0^1) = \partial_1 f(x_0^1, g(x_0^1)) + \partial_2 f(x_0^1, g(x_0^1)) \mathrm{d}g(x_0^1) = 0.$$

From the Implicit Map Theorem, g was constructed so that

$$\mathrm{d}g(x_0^1) = -[\partial_2 F(x_0^1, g(x_0^1))]^{-1} \partial_1 F(x_0^1, g(x_0^1)).$$

Hence, substituting into the above equation we find

$$\partial_1 f(x_0^1, g(x_0^1)) - \partial_2 f(x_0^1, g(x_0^1)) [\partial_2 F(x_0^1, g(x_0^1))]^{-1} \partial_1 F(x_0^1, g(x_0^1)) = 0$$

Now, if we define $\lambda \in \mathbf{R}^m$ by $\lambda = \partial_2 f(x_0^1, g(x_0^1)) [\partial_2 F(x_0^1, g(x_0^1))]^{-1}$, we find that

$$\begin{aligned} \partial_1 f(x_0^1, g(x_0^1)) &= \partial_1 F(x_0^1, g(x_0^1)) \lambda \\ \partial_2 f(x_0^1, g(x_0^1)) &= \partial_2 F(x_0^1, g(x_0^1)) \lambda \end{aligned}$$

So that we can conclude

$$\mathrm{d}f(x_0) = \mathrm{d}F(x_0) \lambda.$$

♮

Theorem A.6.2 (Sufficient condition for a constraint extremum)**.** Let $\Omega \subseteq \mathbf{R}^d$ and maps $f \colon \Omega \to \mathbf{R}$, and $F \colon \Omega \to \mathbf{R}^m$, both of which are $C^2(\Omega, \mathbf{R})$. Let $S \subseteq \Omega$ be the surface defined by $S = \{x \in \Omega : F(x) = 0\}$. Consider $x_0 \in S$ as a possible candidate of local extremum of f in the surface S. Suppose additionally that $\operatorname{rank} F'(x) = m$ for every $x \in \Omega$. Define the Lagrange multipliers $\lambda \in \mathbf{R}^m$ so that the map $L \colon S \to \mathbf{R}$ given by

$$L(x) = f(x) - \langle \lambda, F(x) \rangle,$$

satisfy $\operatorname{grad} L(x) = 0$ for all $x \in S$ and $\partial_\lambda L(x) = 0$.

It's sufficient for $x_0 \in S$ to be a local extremum in S if $\operatorname{Hess} L(x_0)$ is either positive definite or negative definite. If $\operatorname{Hess} L(x_0)$ is not definite, then x_0 cannot be an extremum point. Moreover, if $\operatorname{Hess} L(x_0)$ is positive definite, then x_0 is a local minimum on S, on the other hand, if $\operatorname{Hess} L(x_0)$ is negative definite, then x_0 is a local maximum on S.

Proof. Since $x_0 \in S$, then in particular $\operatorname{grad} L(x_0) = 0$, thus, as $S \ni x \to x_0$ we have the polynomial approximation

$$\begin{aligned} L(x) - L(x_0) &= \frac{1}{2!} \sum_{i,j=1}^{d} \partial_{ij} L(x_0)(x_i - x_0^i)(x_j - x_0^j) + o(\|x - x_0\|^2) \\ &= \frac{1}{2!} \langle x - x_0, \operatorname{Hess}(L(x_0))(x - x_0) \rangle + o(\|x - x_0\|_{\mathbf{R}^d}^2). \end{aligned}$$

We'll assume that S — which is a $(d-m)$-dimensional surface, since $\operatorname{rank} F(x) = m$ — can be parametrically defined in some neighbourhood of $x_0 \in S$ by a smooth map $\mathbf{R}^{d-m} \ni t \mapsto x(t) \in \mathbf{R}^d$ such that $x(0) = x_0$ and that exists a neighbourhood of $0 \in \mathbf{R}^{d-m}$ for which the parametrization is bijective. Since the mapping is smooth, as $t \to 0$ we have

$$x(t) - x(0) = \mathrm{d}x(0)(t) + o(\|t\|_{\mathbf{R}^{d-m}}).$$

Which implies that as $t \to 0$ we have $\|x(t) - x(0)\|_{\mathbf{R}^d} = O(\|t\|_{\mathbf{R}^d})$.

We can now exploit the parametrization of the surface S so that, as $t \to 0$ we have

$$\begin{aligned} L(x(t)) - L(x_0) &= \frac{1}{2!} \langle t, \operatorname{Hess}(L(x(0)))t \rangle + o(O(\|t\|_{\mathbf{R}^d})) \\ &= \frac{1}{2!} \langle t, \operatorname{Hess}(L(x(0)))t \rangle + o(\|t\|_{\mathbf{R}^d}). \end{aligned}$$

Where the entries of the Hessian are of the form $\partial_{ij} L(x(0)) = \partial_{ij} L(x(0)) \partial_i x(0) \partial_j x(0)$. Then, if $\operatorname{Hess} L(x(0))$ is positive or negative definite, by Theorem A.4.6, we obtain that $t = 0$ is an extremum of $L(x(t))$. On the other hand, since there exists a neighbourhood of $t = 0$ for which $x(t)$ is a bijective parametrization, it follows that L has an extremum at x_0 — and hence x_0 is an extremum of f in S and the classifications of maximum or minimum come again from the same theorem. If $\operatorname{Hess} L(x(0))$ is indefinite, by Theorem A.4.6 we conclude that $L(x(t))$ has no extremum at $t = 0$ — and with the same analogous arguments as before, we argue that L has no extrema at x_0 and neither does f. ♮

A.7 Riemann Integration of Real Valued Maps

Primary Definitions

The main setting we are going to be working in this section, which will encompass the study of multiple Riemann integrals, is the standard euclidean space $\mathbf{R}^n$ and the n-dimensional closed intervals $I = [a, b] = \{x \in \mathbf{R}^n : a_j \leqslant x_j \leqslant b_j \text{ for } 1 \leqslant j \leqslant n\}$. If it seems fit, we can denote that a point $x \in \mathbf{R}^n$ lies in the interval generated by given points $a, b \in \mathbf{R}^n$ by simply saying that $a \leqslant x \leqslant b$. Another terminology we are going to adopt is that, the interval I is non-degenerate if $a_j < b_j$ for all $1 \leqslant j \leqslant n$.

Notation A.7.1. In *this* section we denote by $\mathcal{I}^n$ the collection of all *closed* intervals of $\mathbf{R}^n$.

Definition A.7.2 (Interval measure)**.** We define the map vol: $\mathcal{I}^n \to \mathbf{R}$ as the *measure* (or volume) of the n-dimensional closed intervals of $\mathbf{R}^n$, it's defined as the product of the interval sides, that is

$$\operatorname{vol} I := \prod_{j=1}^{n} b_j - a_j, \quad \text{for } \ I = [a, b] \in \mathcal{I}^n.$$

Corollary A.7.3 (Measure of intervals in $\mathbf{R}^n$)**.** Let $I := [a, b] \subseteq \mathbf{R}^n$ be an n-dimensional closed interval, then the following properties are satisfied concerning the measure vol:

(a) (Homogeneity) Let $\gamma \geqslant 0$ be a scalar and define the multiplication of the interval by γ as $\gamma I := [\gamma a, \gamma b]$. Then we have that

$$\operatorname{vol}(\lambda I) = \lambda^n \operatorname{vol} I.$$

(b) (Additivity) Given a finite collection of closed intervals $\{I_j \subseteq \mathbf{R}^n\}_{j=1}^p$, we have that

$$\operatorname{vol} \bigcup_{j=1}^{p} I_j = \sum_{j=1}^{p} \operatorname{vol} I_j.$$

(c) (Cover inequality) Given a finite closed cover $\{I_j\}^p$, by n-dimensional closed intervals, of I — that is $I \subseteq \bigcup_{j=1}^p I_j$ — then

$$\operatorname{vol} I \leqslant \sum_{j=1}^{p} \operatorname{vol} I_j.$$

Definition A.7.4 (Partition)**.** Let $I \subseteq \mathbf{R}^n$ be a closed interval. A *partition* on I is a *finite* collection of closed intervals $\{I_j\}_{j=1}^p$ such that $I = \bigcup_{j=1}^p I_j$. The intervals pertaining to the partition are said to be *finer* than I.

Definition A.7.5 (Partition mesh)**.** Given a partition $P \in 2^{\mathcal{I}^n}$, we define the *mesh* of P as the maximum diameter (recall Definition A.2.6) of the elements of the partition. That is, mesh: $2^{\mathcal{I}^n} \to \mathbf{R}$ is a map defined by

$$\operatorname{mesh}(P) := \max_{I \in P} d(I).$$

Definition A.7.6 (Distinguished points)**.** Given a partition $P = \{I_j\}_{j=1}^p \in 2^{\mathcal{I}^n}$, we define a collection of *distinguished points* of the partition as a collection of points $\xi := \{\xi_j \in I_j\}_{j=1}^p$. The partition P together with the distinguished points ξ will be denoted as the pair (P, ξ) — the collection of pairs (P, ξ) will be denoted by $\mathcal{P}$.

An important filter base $\mathcal{B} \subseteq 2^{\mathcal{P}}$ is defined as the collection of sets B_d, where $d > 0$ is a scalar, such that $B_d := \{(P, \xi) \in \mathcal{P} : \operatorname{mesh}(P) < d\}$. We'll commonly denote $\mathcal{B}$ by $\operatorname{mesh}(P) \to 0$.

Riemann Sums and Integrals

Definition A.7.7 (Riemann sum)**.** Let $f\colon I \to \mathbf{R}$ be a map where $I \in \mathcal{I}^n$. Consider the partition together with distinguished points $(P, \xi) \in \mathcal{P}$, then, we define the *Riemann sum* $\sigma\colon \mathbf{R}^I \times \mathcal{P} \to \mathbf{R}$ by

$$\sigma(f, P, \xi) := \sum_{j=1}^{p} f(\xi_j) \operatorname{vol}(I_j),$$

where $P := \{I_j\}_{j=1}^p$. We say that $\sigma(f, P, \xi)$ is the Riemann sum of the map f with respect to the partition P and distinguished points ξ.

Definition A.7.8 (Riemann integrable maps)**.** A map $f\colon I \to \mathbf{R}$ is said to be *Riemann integrable* if the limit

$$\lim_{\operatorname{mesh}(P)\to 0} \sigma(f, P, \xi)$$

exists in $\mathbf{R}$. We'll denote the $\mathbf{R}$-vector space of Riemann integrable maps with a given domain $E \subseteq \mathbf{R}^n$ by $\mathcal{R}(E)$[1].

Proposition A.7.9 (Boundness of Riemann integrable maps)**.** Let $f\colon I \to \mathbf{R}$ be a Riemann integrable map. Then, f is bounded on I.

Proof. We prove the contrapositive proposition. Suppose that f is unbounded on I and let P be any partition of the interval I. In particular, since P covers I, then there exists an interval $I_k \in P$ for which f is unbounded. Let ξ be any collection of distinguished points of P and define ξ' as the collection of distinguished points $\xi'_j := \xi_j$ for $j \neq m$, and $\xi'_k \in I_k$ to be such that $\xi'_k \neq \xi_k$. Then, from construction, it follows that $\sigma(f, P, \xi) - \sigma(f, P, \xi') = (f(\xi_k) - f(\xi'_k)) \operatorname{vol}(I_k)$. Since from hypothesis f is unbounded in ξ_k, for every $M > 0$, there exists $\xi'_k \in I_k$ such that $\|f(\xi_k) - f(\xi'_k)\| > M$ — that is, $\|f(\xi_k) - f(\xi'_k)\|$ is obviously unbounded, which implies in the divergence of the Riemann sums, hence f is non-Riemann integrable. ♮

Definition A.7.10 (Riemann integral)**.** The Riemann integral of real valued maps is an $\mathbf{R}$-linear map $\int\colon \mathcal{R} \to \mathbf{R}$ defined by mapping any $f \in \mathcal{R}(I)$ to

$$\int_I f(x)\,\mathrm{d}x := \lim_{\operatorname{mesh}(P)\to 0} \sigma(f, P, \xi).$$

Sets of Lebesgue Measure Zero

Definition A.7.11 (Set of Lebesgue measure zero)**.** A set $E \subseteq \mathbf{R}^n$ is said be of Lebesgue *measure zero* if for every $\varepsilon > 0$ there exists a *countable open cover* $\mathcal{U}$ of E by n-dimensional *open* intervals whose total volume $\sum_{I\in\mathcal{U}} \operatorname{vol} \operatorname{Cl} I$ does *not* exceed ε.

Corollary A.7.12. A compact subset $E \subseteq \mathbf{R}^n$ is of measure zero if and only if, for all ε, there exists a *finite* open cover $\mathcal{U}$ of E by open n-dimensional intervals such that $\sum_{I\in\mathcal{U}} \operatorname{vol} \operatorname{Cl} I \leqslant \varepsilon$.

[1] For the time being, we have only defined the case where E is an interval, but I'm already generalizing the notation for its uses in the following subsections.

Proof. If E satisfies the last property, then clearly E is a set of measure zero. On the other hand, if we assume that E is of measure zero, given any $\varepsilon > 0$, let C be a *countable* cover of E for which $\sum_{I \in C} \operatorname{vol} \operatorname{Cl} I \leqslant \varepsilon$. Since E is compact, there exists a finite subcover $\mathcal{U} \subseteq C$ of E, and since $\sum_{I \in \mathcal{U}} \operatorname{vol} \operatorname{Cl} I \leqslant \sum_{I \in C} \operatorname{vol} \operatorname{Cl} I \leqslant \varepsilon$, thus $\mathcal{U}$ is the wanted finite cover. ♮

Lemma A.7.13. The following are properties of sets of measure zero:

(a) A subset of a set of measure zero is of measure zero.

(b) The countable union of sets of measure zero is of measure zero.

(c) A countable set is of measure zero.

(d) A non-degenerate interval is *not* a set of measure zero.

Proof. (a) Let $E \subseteq \mathbf{R}^n$ be a set of measure zero and $A \subseteq E$ be a subset. If $\mathcal{U}$ is a closed cover by intervals satisfying the measure zero condition, then in particular $\mathcal{U}$ covers A therefore A is of zero measure.

(b) Let $\{E_j \subseteq \mathbf{R}^n\}_{j \in J}$ be a countable collection of sets of measure zero, and let $\{\mathcal{U}_j\}_{j \in J}$ be a collection where $\mathcal{U}_j$ is the corresponding closed cover by intervals for E_j. Notice that the countable union $E := \bigcup_{j \in J} E_j$ can be covered by $\mathcal{U} := \bigcup_{j \in J} \mathcal{U}_j$, therefore, since the union of countable collections is countable, it follows that $\mathcal{U}$ is a countable cover for E which satisfies the wanted property.

(c) We initially consider a single point in space. Notice that, for any given ε, there exists a closed interval (for instance, one could choose an interval of equal sides containing the point, whose sides have length less than $\varepsilon^{1/n}$), whose volume is less than ε, contaning the given point — that is, this one interval is sufficient to cover the point. We conclude that a singleton is of measure zero. Using the last item, we find that a countable set is of measure zero.

(d) Let $I = [a, b] \subseteq \mathbf{R}^n$ be a non-degenerate interval. Since $\mathbf{R}^n$ is Lindelöf, every cover of I has a finite subcover so, we can proceed by induction on the cardinality $m \in \mathbf{N}$ of the open cover. For $m = 1$, let $(\alpha, \beta) \subseteq \mathbf{R}^n$ be an open interval covering I. Notice that every $x \in I$ is such that $a_j \leqslant x_j \leqslant b_j$, for all $1 \leqslant j \leqslant n$, then, since x must lie at (α, β), we necessarily have $\alpha_j < a \leqslant x_j \leqslant b < \beta_j$ so that $b_j - a_j < \beta_j - \alpha_j$ and hence $\operatorname{vol} I < \operatorname{vol}[\alpha, \beta]$. For the hypothesis of induction, suppose the proposition holds for a cover of cardinality $n - 1 \in \mathbf{N}_{>1}$. Let $\{(\alpha^i, \beta^i) \subseteq \mathbf{R}^n\}_{i=1}^m$ be a cover of I by open intervals. Let $1 \leqslant k \leqslant n$ be such that $a \in (\alpha^k, \beta^k)$, that is, $\alpha_k < a_j < \beta_j^k$. If for some index $1 \leqslant j_0 \leqslant n$ we have $b_{j_0} > \beta_{j_0}^k$, we define the point β' to be such that, if $\beta_j^k < b_j$ then $\beta'_j = \beta_j^k$, otherwise, if $\beta_j^k \geqslant b_j$, we let $a_j < \beta'_j < b_j$ — that is, we constructed a point so that the closed interval $[\beta', b]$ is non-degenerate. From the hypothesis of induction, every cover of $[\beta', b]$ with cardinality $m - 1$ has total volume strictly greater than $\operatorname{vol}[\beta', b]$. In particular, since $[\beta', b] \subset [a, b]$ then $\{(\alpha^i, \beta^i)\}_{i=1}^m$ is a cover of $[\beta', b]$, notice that $[\beta', b] \cap (\alpha^k, \beta^k) = \varnothing$, thus the cover $\{(\alpha^i, \beta^i)\}_{i=1, i \neq k}^m$ is a cover of $[\beta', b]$ with cardinality $n - 1$, hence $\operatorname{vol}[\beta', b] < \sum_{i=1, i \neq k}^m \operatorname{vol}[\alpha^i, \beta^i]$. Then we find that

$$\begin{aligned} b_j - a_j &< (b_j - \beta'_j) + (\beta'_j - a_j) \\ &\leqslant (b_j - \beta'_j) + (\beta^k_j - a_j) \\ &< (b_j - \beta'_j) + (\beta^k_j - \alpha^k_j). \end{aligned}$$

Therefore we conclude that

$$\begin{aligned} \operatorname{vol}[a, b] &< \operatorname{vol}[\beta', b] + \operatorname{vol}[\alpha^k, \beta^k] \\ &< \sum_{i=1, i\neq k}^{m} \operatorname{vol}[\alpha^i, \beta^i] + \operatorname{vol}[\alpha^k, \beta^k] \\ &= \sum_{i=1}^{m} \operatorname{vol}[\alpha^i, \beta^i], \end{aligned}$$

which proves that the proposition is true for all $m \in \mathbf{N}$.

$\natural$

Example A.7.14. Let $f\colon I \to \mathbf{R}$ be a continuous map on the interval $I \subseteq \mathbf{R}^{n-1}$. The graph of f is a n-dimensional set of Lebesgue measure zero.

Proof. Let Γ denote the graph of f over I. To see this, notice that since I is closed, then f is uniformly continuous on I. For any $\varepsilon > 0$, let $\delta > 0$ be such that $\|f(x) - f(y)\|_{\mathbf{R}} < \varepsilon$ for every $x, y \in I$ such that $\|x - y\|_{\mathbf{R}^{n-1}} < \delta$. Let P be a partition of I with mesh $P < \delta$. For any point $x_0 \in R$ of each $R \in P$, we have an induced interval $R' := R \times [f(x_0) - \varepsilon, f(x_0) + \varepsilon]$ which is such that $\Gamma \subseteq R'$, since we have an oscillation $\omega(f, R) < \varepsilon$ from the construction of P. Let P' be the collection of induced intervals R' of the partition P. From the last observation we have that P' is a closed cover of Γ by closed n-dimensional intervals. Moreover, $\sum_{R' \in P'} \operatorname{vol} R' = \sum_{R \in P} 2\varepsilon \operatorname{vol} R = 2\varepsilon \operatorname{vol} I$ — therefore Γ indeed is of Lebesgue measure zero $\natural$

Notation A.7.15. Given a set X and a property P, we say that P holds *almost everywhere* on X if the subset $A \subseteq X$, such that P is not true, is a set of measure zero.

Theorem A.7.16 (Lebesgue's criterion)**.** A map $f\colon I \to \mathbf{R}$ is Riemann integrable if and only if f is *bounded* on I and f is *continuous almost everywhere* on I.

Proof. (*Necessity*) Let f be Riemann integrable, then, from Proposition A.7.9, f is bounded on I. For the sake of contradiction, let $E \subseteq I$ be the set composed of the points of discontinuity of f, we'll suppose that E doesn't have measure zero. Notice that if $x \in E$, then there exists $n \in \mathbf{N}$ for which $\omega(f, x) \geqslant 1/n$ — that is, f does not converge to a value in x. We can then define $E_n := \{x \in I : \omega(f, x) \geqslant 1/n\}$ for every $n \in \mathbf{N}$ so that $E_n \subseteq E$ and thus $E = \bigcup_{n \in \mathbf{N}} E_n$. Since E isn't of measure zero from assumption, it follows that there necessarily exists at least one $n_0 \in \mathbf{N}$ such that E_{n_0} isn't of measure zero.

Let P be a partition of I, we'll consider two subsets of this partition:

$$A := \{R \in P : R \cap E_{n_0} \neq \varnothing \text{ and } \omega(f, R) \geqslant 1/(2n_0)\},$$

and $B := P \setminus A$. Since P partitions the interval, for any $x \in E_{n_0}$, there exists $R \in P$ such that $x \in \operatorname{Int} R$ or $x \in \partial R$ — we now analyse both cases:

- In the case where x is an interior point, since $\omega(f, x) \geqslant 1/n_0$, it follows that $R \in A$.
- Otherwise, if x is a boundary point, then there actually exists at least another $R' \in P$ such that $x \in \partial R \cap \partial R'$. Suppose, for the sake of contradiction, that every interval $R' \in P$ containing x as a boundary point is such that f has an oscillation $\omega(f, R') < 1/(2n_0)$, then, if we take any ball $B_x(r) \cap I$, neighbourhood of x, we find that, $\omega(f, B_x(r)) < 1/(2n_0) + 1/(2n_0) = 1/n_0$, that is, the limit $\omega(f, x) < 1/n_0$, which is a contradiction to the assumption that $x \in E_n$ — thus there must exist $R' \in P$ with $x \in \partial R'$ such that $\omega(f, R') \geqslant 1/(2n_0)$ so that $R' \in A$.

This implies that A covers the interval E_{n_0} by closed intervals and, by assumption, $\sum_{R \in A} \operatorname{vol} R > \operatorname{vol} E_{n_0}$.

We are now going to consider any two distinct choices of distinguished points ξ and ξ' of P such that, if ξ_j andξ'_j are elements of a common interval of B then $\xi_j = \xi'_j$, and if ξ_j and ξ'_j are elements common to an interval of A, we choose ξ_j and ξ_j to be any points such that $f(\xi_j) - f(\xi'_j) > 1/(3n_0)$ — which is always possible from the construction of A. Notice that we have

$$\begin{aligned}
\|\sigma(f, P, \xi) - \sigma(f, P, \xi')\| &= \|\sum_{R_j \in A} (f(\xi_j) - f(\xi'_j)) \operatorname{vol} R_j\| \\
&> \frac{1}{3n_0} \sum_{R_j \in A} \operatorname{vol} R_j \\
&> \frac{1}{3n_0} \operatorname{vol} E_{n_0} > 0,
\end{aligned}$$

and therefore the limit $\lim_{\operatorname{mesh}(P) \to 0} \sigma(f, P, \xi)$ does not converge whenever E_{n_0} is of measure zero — that is, if the set of discontinuities of f isn't of measure zero, then f is non-Riemann integrable.

(*Sufficiency*) Suppose that f is bounded and continuous almost everywhere on I. Let $\varepsilon > 0$ be any constant and consider the subset $E_\varepsilon := \{x \in I : \omega(f, x) \geqslant \varepsilon\}$ of discontinuous points of f. If E is the set of points of discontinuity of f, then by hypothesis E is of measure zero — since $E_\varepsilon \subseteq E$, then E_ε is also of measure zero. Moreover, E_ε is necessarily closed, which implies in E_ε compact. Let $\{I_j\}_{j=1}^k$ be a *finite* open cover of E_ε by open intervals $I_j \subseteq \mathbf{R}^n$ such that $\sum_{i=1}^k \operatorname{vol} \operatorname{Cl} I < \varepsilon$ (recall that this is possible because of Corollary A.7.12).

We now define the λ-dilations of intervals of the cover $\{I_j\}_{j=1}^k$ but with centre unchanged — that is, if $I_j = (x, y)$, then the dilation $I_j^\lambda := (z, w)$ must be such that $w_i - z_i = \lambda^{1/n}(y_i - x_i)$ and $w_i + z_i = y_i + x_i$, for all $1 \leqslant i \leqslant n$. Solving such system we obtain $z_i = \frac{1}{2}[x_i(1+\lambda^{1/n}) + y_i(1-\lambda^{1/n})]$ and $w_i = \frac{1}{2}[x_i(1-\lambda^{1/n}) + y_i(1+\lambda^{1/n})]$. This way we have

$$\sum_{j=1}^{k} \operatorname{vol} \operatorname{Cl} I_j^\lambda = \lambda^n \sum_{j=1}^{k} \operatorname{vol} \operatorname{Cl} I_j < \lambda^n \varepsilon.$$

We'll define the sets $C_\lambda := (\bigcup_{j=1}^{k} I_j^\lambda) \cap I$ — moreover, d will be the minimum distance between the boundaries of C_2 and C_3.

From construction, we have $E_\varepsilon \subseteq \operatorname{Int} C_2$, now if we consider the compact set $K := I \setminus (\operatorname{Int} C_2)$, then, for every $x \in K$, we have $\omega(f, x) < \varepsilon$. Using Theorem A.2.33, we find that there exists a $\delta > 0$ such that $\|f(x) - f(y)\| < 2\varepsilon$, for all $x, y \in K$ with $\|x - y\| < \delta$. We'll now finally show the convergence of the Riemann sums. Let $P' := \{R'_\alpha\}_{\alpha \in A}$ and $P'' := \{R''_\beta\}_{\beta \in B}$ be any two partitions of I with $\operatorname{mesh}(P'), \operatorname{mesh}(P'') < \min(d, \delta)$. Define $P := \{R_{(\alpha,\beta)} := R'_\alpha \cap R''_\beta\}_{(\alpha,\beta) \in A \times B}$, which clearly is a partition of I. Since we are dealing with partitions, for every $R'_\alpha \in P'$ we have $R'_\alpha = \bigcup_{\beta \in B} R_{(\alpha,\beta)}$ — thus $\operatorname{vol} R'_\alpha = \sum_{\beta \in B} \operatorname{vol} R_{(\alpha,\beta)}$. Let ξ' and ξ be any two distinguished points of the partitions P' and P, respectively, then

$$\begin{aligned}
\|\sigma(f, P', \xi') - \sigma(f, P, \xi)\| &= \Big\| \sum_{(\alpha,\beta) \in A \times B} (f(\xi'_\alpha) - f(\xi_{(\alpha,\beta)})) \operatorname{vol} R_{(\alpha,\beta)} \Big\| \\
&\leqslant \sum_{(\alpha,\beta) \in A' \times B'} \|f(\xi'_\alpha) - f(\xi_{(\alpha,\beta)})\| \operatorname{vol} R_{(\alpha,\beta)} \\
&\quad + \sum_{(\alpha,\beta)(A \times B) \setminus (A' \times B')} \|f(\xi'_\alpha) - f(\xi_{(\alpha,\beta)})\| \operatorname{vol} R_{(\alpha,\beta)},
\end{aligned}$$

where we define

$$A' \times B' := \{(\alpha, \beta) \in A \times B \colon R'_\alpha \subseteq C_3\}.$$

Moreover, since the diameter $d(R'_\alpha) < d$, then for every R'_α not entirely contained in C_3, the intersection with any R''_β, namely $R_{(\alpha,\beta)}$, cannot lie in the interior of C_2 — that is, for all $(\alpha, \beta) \in (A \times B) \setminus (A' \times B')$ we'll have $R_{(\alpha,\beta)} \subseteq K$ and surely $\xi'_\alpha, \xi_{(\alpha,\beta)} \in K$, which satisfy $\|f(\xi'_\alpha) - f(\xi_{(\alpha,\beta)})\| < 2\varepsilon$ since $\operatorname{mesh} P' < \delta$. Assuming f is bounded by $M > 0$ in I, that is $\|f\| \leqslant M$, then $\|f(\xi'_\alpha) - f(\xi_{(\alpha,\beta)})\| < 2M$ and thus

$$\begin{aligned}
\|\sigma(f, P', \xi') - \sigma(f, P, \xi)\| &\leqslant \sum_{(\alpha,\beta) \in A' \times B'} 2M \operatorname{vol} R_{(\alpha,\beta)} + \sum_{(\alpha,\beta) \in (A \times B) \setminus (A' \times B')} 2\varepsilon \operatorname{vol} R_{(\alpha,\beta)} \\
&\leqslant 2M(3^n \varepsilon) + 2\varepsilon \operatorname{vol} I \\
&= 2\varepsilon(3^n M + \operatorname{vol} I).
\end{aligned}$$

Therefore, since the same construction is applicable for P'', that is, $\|\sigma(f, P'', \xi'') - \sigma(f, P, \xi)\| \leqslant 2\varepsilon(3^n M + \operatorname{vol} I)$, thus

$$\|\sigma(f, P', \xi') - \sigma(f, P'', \xi'')\| \leqslant 4\varepsilon(3^n M + \operatorname{vol} I).$$

Since P' and P'' where chosen arbitrarily, it follows that the sequence of Riemann sums converge by the Cauchy criterion and thus f is Riemann integrable from definition. ♮

Integrating Over Sets

Definition A.7.17. A set $E \subseteq \mathbf{R}^n$ is said to be *admissible* if it is bounded and has measure zero boundary.

Corollary A.7.18 (Operations on admissible sets)**.** The finite union or intersection of admissible sets is admissible, and the difference of admissible sets is admissible.

Notation A.7.19. We denote the characteristic map of a given set E to be $\chi_E : E \to \{0, 1\}$, where $\chi_E(x) := 1$ if $x \in E$, and $\chi_E(x) := 0$ if $x \notin E$. Moreover, given any map $f : E \to \mathbf{R}$, we define the map $f_{\chi_E} : \mathbf{R}^n \to \mathbf{R}$ as $f_{\chi_E}|_E := f$ and f_{χ_E} is zero everywhere else.

Definition A.7.20 (Riemann integral over a set)**.** Let $E \subseteq \mathbf{R}^n$ be an admissible set and $f : I \to \mathbf{R}$ for some $I \supseteq E$, then we define the integral of f over the set E as

$$\int_E f(x)\,\mathrm{d}x := \int_I f_{\chi_E}(x)\,\mathrm{d}x.$$

Lemma A.7.21 (The integral is well-defined)**.** Let $f : E \to \mathbf{R}$ be a map. The integral of f over E is independent of the choice of the interval containing E.

Proof. Let $I, I' \subseteq \mathbf{R}^n$ be any two intervals containing E. Define an auxiliary interval $I_0 := I \cap I'$. It follows from construction that every point of discontinuity of f_{χ_E} is contained in I_0. From Theorem A.7.16 we see that, if the collection of points of discontinuity of f_{χ_E} is not of measure zero, then both integrals of f_{χ_E} over I and I' fail to exist — while, if the collection is of measure zero, both integrals exist simultaneously.

Suppose that the points of discontinuity form a set of measure zero. Given any $\varepsilon > 0$, we consider a partition P_0 of I with mesh $P_0 < \varepsilon$. Define partitions P and P' of I and I', respectively, such that $P_0 \subseteq P \cap P'$ — that is, inside I_0, they share the exact same collection of intervals as P_0. Given any distinguished points ξ and ξ' of P and P', respectively, we have that $f_{\chi_E}(x) = 0$ for every $x \in \xi \cup \xi'$ such that $x \notin I_0$, therefore the Riemann sums of f_{χ_E} under the partitions P and P' are always reduced to Riemann sums of f_{χ_E} under the partition P_0 and the corresponding distinguished points. This implies that the limit of the Riemann sums are equal for both partitions and therefore we have

$$\int_I f_{\chi_E}(x)\,\mathrm{d}x = \int_{I'} f_{\chi_E}(x)\,\mathrm{d}x.$$

♮

Corollary A.7.22. Let $E \subseteq \mathbf{R}^n$ be an admissible set and $f : E \to \mathbf{R}$. Then, f is Riemann integrable if and only if is bounded and continuous at almost all points of E.

Proof. Notice that the discontinuities of the corresponding map f_{χ_E} are those of f and perhaps a collection of points of ∂E, where χ_E changes its value. Since ∂E is of measure zero, it doesn't interfere in the use of Lebesgue's criterion and thus the proposition follows. ♮

Jordan Measure

Definition A.7.23 (Jordan measure)**.** Given a bounded set $E \subseteq \mathbf{R}$, we define the Jordan measure of E as the map

$$\mu(E) := \int_E \mathrm{d}x.$$

Notice that the integral over E only exists for admissible sets, thus the Jordan measure μ is only defined for admissible sets.

Definition A.7.24 (Jordan's measure zero sets)**.** A set $E \subseteq \mathbf{R}^n$ is said to be of Jordan measure zero (or of content zero) if for every $\varepsilon > 0$ there exists a *finite* open cover by intervals whose total volume is less than ε.

Definition A.7.25 (Jordan-measurable sets)**.** A set $E \subseteq \mathbf{R}^n$ is said to be Jordan-measurable if it is bounded and ∂E is a set of Jordan measure zero.

Riemann Integral Properties

Corollary A.7.26. The collection of Riemann integrable maps $E \to \mathbf{R}$ for an $\mathbf{R}$-vector space. Moreover, the Riemann integral is an $\mathbf{R}$-linear functional $\int_E : \mathcal{R}(E) \to \mathbf{R}$ of the dual space $\mathcal{R}(E)^*$.

Proposition A.7.27. Let $E, S \subseteq \mathbf{R}^n$ be admissible sets and $f : E \cup S \to \mathbf{R}$. Then

(a) f is Riemann integrable over $E \cup S$ if and only if it's integrable over both E and S simultaneously. Moreover, if such condition is met, the f is also Riemann integrable over $E \cap S$.

(b) If f is Riemann integrable over $E \cup S$ *and* $\mu(E \cap S) = 0$, then

$$\int_{E\cup S} f(x)\,\mathrm{d}x = \int_E f(x)\,\mathrm{d}x + \int_S f(x)\,\mathrm{d}x.$$

Proof. (a) If f is Riemann integrable over $E \cup S$, then f is bounded and continuous at almost all points of $E \cup S$, which implies that f is Riemann integrable over both E and S. Since $E \cap S \subseteq E \cup S$ then f is also Riemann integrable over the intersection.

(b) Notice that $\chi_{E\cup S}(x) = \chi_E(x) + \chi_S(x) - \chi_{E\cap S}(x)$, therefore by definition — if $I \supseteq E \cup S$ is an interval, then

$$\begin{aligned}
\int_{E\cup S} f(x)\,\mathrm{d}x &= \int_I f_{\chi_{E\cup S}}(x)\,\mathrm{d}x \\
&= \int_I f_{\chi_E}(x) + f_{\chi_S}(x) - f_{\chi_{E\cap S}}(x)\,\mathrm{d}x \\
&= \int_I f_{\chi_E}(x)\,\mathrm{d}x + \int_I f_{\chi_S}(x)\,\mathrm{d}x - \int_I f_{\chi_{E\cap S}}(x)\,\mathrm{d}x \\
&= \int_I f_{\chi_E}(x)\,\mathrm{d}x + \int_I f_{\chi_S}(x)\,\mathrm{d}x
\end{aligned}$$

where we used the hypothesis that $\mu(E \cap S) = 0$ to conclude that

$$\int_I f_{\chi_{E\cap S}}(x)\,\mathrm{d}x = \int_{E\cap S} f(x)\,\mathrm{d}x = 0.$$

$\natural$

Proposition A.7.28 (Estimate)**.** Let $f \in \mathcal{R}(E)$, where $E \subseteq \mathbf{R}^n$ is an admissible set, then $\|f\| \in \mathcal{R}(E)$ and also

$$\|\int_E f(x)\,\mathrm{d}x\| \leqslant \int_E \|f(x)\|\,\mathrm{d}x.$$

Proof. Notice that if f is Riemann integrable, then f is bounded and and continuous at almost all points of E, thus surely $\|f\|$ is bounded and continuous — hence Riemann integrable. Moreover, since $f(x) \leqslant \|f(x)\|$ for all $x \in E$, it follows that, for any given partition P and distinguished points χ, we have the inequality $\sigma(f, P, \chi) \leqslant \sigma(\|f\|, P, \chi)$, thus the integral inequality holds by taking the limit mesh $P \to 0$. $\natural$

Proposition A.7.29 (Non-negative real valued maps)**.** Let $f\colon E \to \mathbf{R}$ be a Riemann integrable map over the admissible set $E \subseteq \mathbf{R}^n$. If f is non-negative over E, that is, $f(x) \geqslant 0$ for all $x \in E$, then

$$\int_E f(x)\,\mathrm{d}x \geqslant 0.$$

Proof. Just observe that for any partition P and distinguished points χ, we have $\sigma(f, P, \chi) \geqslant 0$, thus $\lim_{\mathrm{mesh}\,P\to 0} \sigma(f, P, \chi) \geqslant 0$. $\natural$

Corollary A.7.30. Let $E \subseteq R^n$ be an admissible set and $f, g\colon E \to \mathbf{R}$ be two given maps. The following are miscellaneous immediate implications of Corollary A.7.30:

(a) If both f and g are Riemann integrable, and $f(x) \leqslant g(x)$ for all $x \in E$, then

$$\int_E f(x)\,\mathrm{d}x \leqslant \int_E g(x)\,\mathrm{d}x.$$

(b) If f is Riemann integrable over E and, for some constants $m, M \in \mathbf{R}$ we have $m \leqslant f(x) \leqslant M$ for all $x \in E$, then

$$m\mu(E) \leqslant \int_E f(x)\,\mathrm{d}x \leqslant M\mu(E).$$

(c) If f is Riemann integrable over E, define constants $m := \inf_{x\in E} f(x)$ and $M := \sup_{x\in E} f(x)$. Then, there exist a constant $m \leqslant \omega \leqslant M$ for which

$$\int_E f(x)\,\mathrm{d}x = \theta\mu(E).$$

(d) If additionally E is connected and f is continuous, then f is Riemann integrable over E and there exists $y \in E$ for which

$$\int_E f(x)\,\mathrm{d}x = f(y)\mu(E).$$

(e) If f and g are Riemann integrable over E and, for some constants $m, M \in \mathbf{R}$ we have $m \leqslant f(x) \leqslant M$ for all $x \in E$, and $g(x) \geqslant 0$ for all $x \in E$, then

$$m \int_E g(x)\,\mathrm{d}x \leqslant \int_E f(x)g(x)\,\mathrm{d}x \leqslant M \int_E g(x)\,\mathrm{d}x.$$

Lemma A.7.31. Let $f : E \to \mathbf{R}$ be a non-negative map that is Riemann integrable over E, where E is an admissible set. If $\int_E f(x)\,\mathrm{d}x = 0$, then $f(x) = 0$ at almost all points of E.

Proof. Let $x_0 \in E$ be any point of continuity of f and, for the sake of contradiction, assume that $f(x_0) > 0$ and consider a constant $c > 0$ with $f(x_0) \geqslant c > 0$. Let $U \subseteq E$ be any neighbourhood of x_0 for which $f(x) \geqslant c > 0$ for every $x \in U$. Then we obtain the following contradiction — where $I \supseteq E$ is an interval:

$$\int_E f(x)\,\mathrm{d}x = \int_I f\chi_E(x)\,\mathrm{d}x = \int_U f\chi_E(x)\,\mathrm{d}x + \int_{I \setminus U} f\chi_E(x)\,\mathrm{d}x \geqslant \int_U f\chi_E(x)\,\mathrm{d}x \geqslant c\mu(U) > 0.$$

This cannot be true, thus $f(x_0) = 0$ for every point $x_0 \in E$ where f is continuous. Since f is continuous almost everywhere, it follows that f is zero almost everywhere. $\natural$

Corollary A.7.32. Let $\sim$ be the equivalence relation on $\mathcal{R}(E)$ as follows: $f \sim g$ if and only if the collection of points where the map $f - g$ is non-zero forms a set of Lebesgue measure zero. Then the map $\|-\| : \mathcal{R}(E)/\!\sim\; \to \mathbf{R}$ defined by

$$\|f\| := \int_E \|f(x)\|_{\mathbf{R}^n}\,\mathrm{d}x$$

is a norm on the vector space $\mathcal{R}(E)/\!\sim$.

A.8 Fubini's Theorem

Upper and Lower Darboux Integrals

In order to proceed to the theorem of Fubini type, we first need to define the concept of Darboux lower and upper integrals.

Definition A.8.1 (Darboux sums)**.** Let $f : I \to \mathbf{R}$ be a map defined on the interval $I \subseteq \mathbf{R}^n$. If $P := \{I_j\}_{j \in J}$ is a partition of I by intervals, then we define the following

- The lower Darboux sum of f over the interval I corresponding to the partition P is
$$s(f, P) := \sum_{j \in J} \mathrm{vol}(I_j) \inf_{x \in I_j} f(x).$$
- The upper Darboux sum of f over the interval I corresponding to the partition P is
$$S(f, P) := \sum_{j \in J} \mathrm{vol}(I_j) \sup_{x \in I_j} f(x).$$

Definition A.8.2 (Upper and lower Darboux integrals)**.** Let $f\colon I \to \mathbf{R}$ be a map over the interval $I \subseteq \mathbf{R}^n$. We define:

- The lower Darboux integral of f is $\underline{\mathcal{J}} := \sup_P s(f, P)$, where the P goes through all partitions of I.
- The upper Darboux integral of f is $\underline{\mathcal{J}} := \inf_P S(f, P)$, where the P goes through all partitions of I.

Theorem A.8.3 (Darboux criterion)**.** A map $f\colon I \to \mathbf{R}$, where $I \subseteq \mathbf{R}^n$ is an interval, is integrable over I if and only if f is bounded on I and the upper and lower Darboux integrals agree over I.

Add proof of Darboux criterion later

Fubini

Theorem A.8.4. Let $X \times Y \subseteq \mathbf{R}^{m+n}$ be an interval, and $f\colon X \times Y \to \mathbf{R}$ be a Riemann integrable map over $X \times Y$. Then we have the following equality:

$$\int_{X\times Y} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_X \left(\int_Y f(x, y)\,\mathrm{d}y \right) \mathrm{d}x = \int_Y \left(\int_X f(x, y)\,\mathrm{d}x \right) \mathrm{d}y. \tag{A.45}$$

Proof. In order to prove the theorem, we first construct the following maps:

- $F_X\colon X \to \mathbf{R}$ is defined as $F_X(x_0) := \int_Y f(x_0, y)\,\mathrm{d}x$ whenever such integral exists, in case not, we let $\underline{\mathcal{J}_{x_0}} \leqslant F_X(x_0) \leqslant \overline{\mathcal{J}_{x_0}}$ be any element between the Darboux integrals of the map $f(x_0, -)\colon Y \to \mathbf{R}$ given by $y \mapsto f(x_0, y)$.
- $F_Y\colon Y \to \mathbf{R}$ is defined as $F_Y(y_0) := \int_x f(x, y_0)\,\mathrm{d}x$ whenever such integral exists, in case not, we let $\underline{\mathcal{J}_{y_0}} \leqslant F_Y(y_0) \leqslant \overline{\mathcal{J}_{y_0}}$ be any element between the Darboux integrals of the map $f(-, y_0)\colon X \to \mathbf{R}$ given by $x \mapsto f(x, y_0)$.

We'll carry the proof for the first equality, but, knowing the construction of F_Y, the idea for the second equality is the exact analogue.

Let $P := \{X_i\}_{i\in I} \times \{Y_j\}_{j\in J}$ be any partition of the interval $X \times Y$. Notice that

$$\begin{aligned}
s(f,P) &= \sum_{(i,j)\in I\times J} \inf_{(x,y)\in X_i\times y_j} f(x,y)\,\mathrm{vol}(X_i\times Y_j)\\
&= \sum_{(i,j)\in I\times J} \inf_{(x,y)\in X_i\times y_j} f(x,y)\,\mathrm{vol}(X_i)\,\mathrm{vol}(Y_j)\\
&\leqslant \sum_{i\in I} \inf_{x\in X_i}\left[\sum_{j\in J} f(x,y)\,\mathrm{vol}(Y_j)\right]\mathrm{vol}\,X_i\\
&\leqslant \sum_{i\in I} \inf_{x\in X_i}\left[\underline{\int_Y} f(x,y)\,\mathrm{d}y\right]\mathrm{vol}\,X_i\\
&\leqslant \sum_{i\in I} \inf_{x\in X_i} F_X(x)\,\mathrm{vol}\,X_i\\
&\leqslant \sum_{i\in I} \sup_{x\in X_i} F_X(x)\,\mathrm{vol}\,X_i\\
&\leqslant \sum_{i\in I} \sup_{x\in X_i}\left[\overline{\int_X} f(x,y)\,\mathrm{d}x\right]\mathrm{vol}\,X_i\\
&\leqslant \sum_{i\in I} \sup_{x\in X_i}\left[\sum_{j\in J} f(x,y)\,\mathrm{vol}\,Y_j\right]\mathrm{vol}\,X_i\\
&\leqslant \sum_{(i,j)\in I\times J} f(x,y)\,\mathrm{vol}(X_i\times Y_j)\\
&= S(f,P).
\end{aligned}$$

By hypothesis, f is Riemann integrable over $X \times Y$, thus

$$\lim_{\mathrm{mesh}\,P\to 0} s(f,P) = \lim_{\mathrm{mesh}\,P\to 0} S(f,P) = \int_{X\times Y} f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

therefore F_X is Riemann integrable over X and

$$\int_{X\times Y} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_X F_X(x)\,\mathrm{d}x.$$

♮

Corollary A.8.5. Let $f\colon X \times Y \to \mathbf{R}$ be Riemann integrable. Then for almost all points $x_0 \in X$ and almost all $y_0 \in Y$ the integrals $\int_Y f(x_0,y)\,\mathrm{d}y$ and $\int_X f(x,y_0)\,\mathrm{d}x$ exist.

Proof. Since $s(f,P) \leqslant S(f,P)$ for all partitions P of $X \times Y$, then $\underline{\int_Y} f(x_0,y)\,\mathrm{d}y \leqslant \overline{\int_X} f(x_0,y)\,\mathrm{d}y$ for every $x_0 \in X$. From Theorem A.8.4 we see that

$$\int_X \left(\overline{\int_Y} f(x,y)\,\mathrm{d}y - \underline{\int_Y} f(x,y)\,\mathrm{d}y\right)\mathrm{d}x = 0,$$

therefore by Lemma A.7.31 we conclude that the map $x_0 \mapsto \overline{\int_Y} f(x_0, y)\,\mathrm{d}y - \underline{\int_Y} f(x_0, y)\,\mathrm{d}y$ equals zero for almost all $x_0 \in X$ — that is, $\int_Y f(x_0, y)\,\mathrm{d}y$ exists for almost all $x_0 \in X$. The exact same argument can be used to show this for $\int_X f(x, y_0)\,\mathrm{d}x$. ♮

Corollary A.8.6 (Iterated Fubini's theorem)**.** Let $I := \prod_{j=1}^{n} [a_j, b_j] \subseteq \mathbf{R}^n$ be an interval and $f\colon I \to \mathbf{R}$ be a Riemann integrable map over I. Then

$$\int_I f(x)\,\mathrm{d}x = \int_{a_n}^{b_n} \left(\int_{a_{n-1}}^{b_{n-1}} \left(\dots \left(\int_{a_1}^{b_1} f(x, y)\,\mathrm{d}x_1 \right) \dots \right) \mathrm{d}x_{n-1} \right) \mathrm{d}x_n.$$

Moreover, the result of the integral is invariant under any permutation of the order of the sub-intervals $[a_j, b_j]$ of integration.

Proof. This is just a direct application of Theorem A.8.4, moreover, the second part of the proposition can be obtained by seeing that the second equality of Eq. (A.45) is nothing but a transposition on the set $\{X, Y\}$ — generalizing this for transpositions on $\{[a_j, b_j]\}_{j=1}^n$ imply that any permutation of the collection satisfies the integral equality. ♮

Measuring Volumes Between Graphs of Maps

Corollary A.8.7. Define $D \subseteq \mathbf{R}^{n-1}$ to be a bounded set and let maps $g_1, g_2\colon D \to \mathbf{R}$ such that $g_1 \leqslant g_2$. Let $f\colon E \to \mathbf{R}$ be a Riemann integrable map over $E := \{(x, y) \in \mathbf{R}^{n-1} \times \mathbf{R} : x \in D \text{ and } g_1(x) \leqslant y \leqslant g_2(x)\}$. Then

$$\int_E f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_D \left(\int_{g_1(x)}^{g_2(x)} f(x, y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Proof. For each $x \in D$ define the set $E_x := \{y \in \mathbf{R} : g_1(x) \leqslant y \leqslant g_2(x)\}$, and observe that, from construction, $\chi_E(x, y) = \chi_D(x)\chi_{E_x}(y)$. Given any interval $I := I_x \times I_y \supseteq E$, we have

$$\int_E f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_I f_{\chi_E}(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_{I_x} \left(\int_{I_y} f_{\chi_E}(x, y)\,\mathrm{d}y \right) \mathrm{d}x$$

♮

Corollary A.8.8 (Measuring the set between the graph of continuous maps)**.** Let $D \subseteq \mathbf{R}^{n-1}$ be a Jordan-measurable set and $g_1, g_2\colon D \to \mathbf{R}$ be continuous maps such that $g_1 \leqslant g_2$. Then the set $E := \{(x, y) \in \mathbf{R}^{n-1} \times \mathbf{R} : x \in D \text{ and } g_1(x) \leqslant y \leqslant g_2(x)\} \subseteq \mathbf{R}^n$ is Jordan-measurable and its Jordan measure is

$$\mu(E) = \int_D g_2(x) - g_1(x)\,\mathrm{d}x.$$

Proof. Denote by Γ_{g_1} and Γ_{g_2} the graphs of g_1 and g_2, respectively. By Example A.7.14, Γ_{g_1} and Γ_{g_2} are sets of Lebesgue measure zero. Define a set

$$Z := \partial D \times [\inf_{x \in D} g_1(x), \sup_{x \in D} g_2(x)] \subseteq \mathbf{R}^n$$

and the length $\ell := \sup_{x \in d} g_2(x) - \inf_{x \in D} g_1(x)$. Since D is a Jordan-measurable set, it follows that ∂D can be covered by a finite collection of open sets with total volume arbitrarily small — in particular, smaller than ε/ℓ. In particular, it follows that Z is a set of Jordan measure zero. Notice that $\partial E \subseteq \Gamma_{g_1} \cup \Gamma_{g_2} \cup Z$ and thus ∂E is a set of Jordan measure zero — thus E is Jordan-measurable. The measure formula for E follows immediately by Corollary A.8.7 $\natural$

Corollary A.8.9. Let $E \subseteq \mathbf{R}^n$ be a Jordan-measurable set and $I := I_x \times I_y \subseteq \mathbf{R}^{n-1} \times \mathbf{R}$ be an interval containing E. For each $y_0 \in I_y$, define the section $E_{y_0} := \{(x, y_0)\} \subseteq E$. Then, for almost all $y_0 \in I_y$ the set E_{y_0} is Jordan-measurable, and

$$\mu(E) = \int_{I_y} F_\mu(y)\, \mathrm{d}y,$$

where the map $F_\mu \colon I_y \to \mathbf{R}$ is defined as $F_\mu(y) := \mu(E_y)$ if E_y happens to be Jordan-measurable, otherwise, we let $\underline{\int_{E_y}} \mathrm{d}y \leqslant F_\mu(y) \leqslant \overline{\int_{E_y}} \mathrm{d}y$ be any value between the lower and upper Darboux integrals of the measure of E_y.

Corollary A.8.10 (Cavalieri's principle)**.** Let A and B be two Jordan-measurable solids in $\mathbf{R}^3$. For every $z_0 \in \mathbf{R}$, define the sets $A_{z_0} := \{(x, y, z_0) \in A\}$ and $B_{z_0} := \{(x, y, z_0) \in B\}$. If for every $z_0 \in \mathbf{R}$ the sets A_{z_0} and B_{z_0} have the same area, that is, $\mu(A_{z_0}) = \mu(B_{z_0})$, then the solids have the same volumes — $\mu(A) = \mu(B)$.

A.9 Change of Variables

Remark A.9.1. Throughout this section we'll use always the variable names D_t and D_x for *domains* — recall Definition A.2.35.

Definition A.9.2 (Support of a real valued map)**.** Given a map $f \colon D \to \mathbf{R}$, where $D \subseteq \mathbf{R}^n$ is a domain, the support of f, denoted by $\operatorname{supp} f$, is the closure of the collection of points of D with non-zero image under f, that is

$$\operatorname{supp} f := \mathrm{Cl}(\{x \in D : f(x) \neq 0\}).$$

The goal of this section will be to prove the theorem on the change of variables in the context of multiple integrals. Our approach will be to prove that such theorem holds for elementary C^1-isomorphisms and also for the compositions of those. We'll decompose the general Riemann integrable map $f \colon D_x \to \mathbf{R}$ into the composition of such elementary C^1-isomorphisms and hence prove the theorem. For the proof of the theorem, head to Appendix A.9.

Theorem A.9.3 (Change of variables)**.** Let $D_x, D_t \subseteq \mathbf{R}^n$ be bounded open domains, and $\phi\colon D_t \xrightarrow{\simeq} D_x$ a C^1-isomorphism. Let $f\colon D_x \to \mathbf{R}$ be a Riemann integrable map over D_x for which $\operatorname{supp} f$ is compact on D_x. Then the map given by $t \mapsto f\phi(t)\|\det(\operatorname{Jac}\phi(t))\|$ is Riemann integrable on D_t and

$$\int_{D_x} f(x)\,\mathrm{d}x = \int_{D_t} f\phi(t)\|\det(\operatorname{Jac}\phi(t))\|\,\mathrm{d}t.$$

Measurable Sets and Smooth Mappings

Lemma A.9.4. Let $\phi\colon D_t \xrightarrow{\simeq} D_x$ be a C^1-isomorphism. Then the following is true:

(a) Given any subset $E_t \subseteq D_t$ of Lebesgue measure zero, the image $\phi(E_t) \subseteq D_x$ is also a set of Lebesgue measure zero.

(b) Let $E_t \subseteq D_t$ be any open subset with closure $\operatorname{Cl} E_t \subseteq D_t$ of Jordan measure zero. Then its image $\phi(E_t)$ and the closure $\operatorname{Cl}(\phi(E_t))$ are both contained in D_x and the closure is a set of Jordan measure zero.

(c) Let $E_t \subseteq D_t$ be a Jordan-measurable set, and $\operatorname{Cl} E_t \subseteq D_t$. The image $\phi(E_t)$ is Jordan measurable and $\operatorname{Cl}(\phi(E_t)) \subseteq D_x$.

Proof. (a) Let $I \subseteq D_t$ be a closed interval with $E_t \subseteq I$. Since ϕ is continuously differentiable, there exists $M > 0$ such that $\|\mathrm{d}\phi(t)(x)\| \leqslant M$ for all $x \in I$. Using Theorem A.3.30, we obtain that for all pairs $x, y \in I$, we have the relation $\|f(x) - f(y)\|_{\mathbf{R}^n} \leqslant M\|x - y\|_{\mathbf{R}^n}$. Given any $\varepsilon > 0$, since E_t is of Lebesgue measure zero, let $\{I_j\}_{j\in J}$ be a countable collection of open intervals covering E_t with $\sum_{j\in J} \operatorname{vol} I_j < \varepsilon$ and we may assume that $I_j \subseteq I$ for all $j \in J$. Consider the countable collection $\{\phi(I_j)\}_{j\in J}$, which forms a cover for $\phi(E_t)$. Let $t_j \in I_j$ be the centre point of the interval corresponding to the index $j \in J$ — notice that, since the image of pairs of points of D_t under ϕ is bounded by the product of their distance and M — defining $x_j := \phi(t_j)$, we can consider the open interval $I'_j \subseteq E_t$ with x_j being its centre point, with dimension M times of I_j, then clearly I'_j is able to cover the whole set $\phi(I_j)$. Let $\{I'_j\}_j$ be the collection of those intervals we constructed and notice that the image $\phi(E_t)$ is completely encompassed by the union $\bigcup_{j\in J} I'_j$, that is, $\{I'_j\}_{j\in J}$ covers $\phi(E_t)$. Moreover, we the volume of such collection was constructed so that $\sum_{j\in J} \operatorname{vol} I'_j = \sum_{j\in J} M^n \operatorname{vol} I_j < M^n\varepsilon$, that is, $\phi(E_t)$ can be covered by a collection of open intervals whose total volume can be chosen to be arbitrarily small volume — thus $\phi(E_t)$ is of Lebesgue measure zero.

(b) Since $\operatorname{Cl} E_t$ is a set of Jordan measure zero, it follows it follows that $\operatorname{Cl} E_t$ is also of Lebesgue measure zero. Notice that, by the above item, the image $\phi(\operatorname{Cl} E_t)$ will also be of Lebesgue measure zero. Notice that $\operatorname{Cl}(\phi(\operatorname{Cl} E_t))$ is a compact set of $\mathbf{R}^n$ and therefore we can apply Corollary A.7.12 to conclude that $\operatorname{Cl}(\phi(\operatorname{Cl} E_t))$ is a set of Jordan measure zero. Since any subset of a set of Jordan measure zero also has measure zero, it follows that $\phi(\operatorname{Cl} E_t)$ is of measure zero.

(c) Since E_t is Jordan-measurable, the boundary ∂E_t is of Jordan measure zero, therefore so is its image $\phi(\partial E_t)$ by the use of the last item. Since ϕ is a C^1-isomorphism, given any $p \in \operatorname{Int} E_t$, we have $\phi(p) \in \operatorname{Int} \phi(E_t)$, therefore $\partial \phi(E_t) = \phi(\partial E_t)$ — thus $\phi(E_t)$ has boundary of Jordan measure zero and therefore is Jordan-measurable.

$\natural$

Corollary A.9.5. In the context of Theorem A.9.3, the integral

$$\int_{D_t} f\phi(t) \|\det(\operatorname{Jac}\phi(t))\| \,\mathrm{d}t$$

exists.

Proof. Since ϕ is a C^1-isomorphism, its Jacobian $\operatorname{Jac}\phi(t)$ is invertible for every $t \in D_t$, thus $\det(\operatorname{Jac}\phi(t)) \neq 0$. If $g\colon D_t \to \mathbf{R}$ is defined as $g(t) := f\phi(t)\|\det(\operatorname{Jac}\phi(t))\|$, then for every $t \in D_t$ such that $\phi(t) \notin \operatorname{supp} f$ we have $t \notin \operatorname{supp} g$ — thus $\operatorname{supp} g = \operatorname{supp}(f\phi\phi^{-1}(\operatorname{supp} f))$, and $\operatorname{supp} g$ is a compact subset of D_t since it is closed. From this we can conclude that the map $g_{\chi_{D_t}}$ has points of discontinuity in the fibers $\phi^{-1}(x)$, where $x \in D_x$ is a point of discontinuity of f. Since f is Riemann integrable in D_x, such points form a Lebesgue measure zero set — thus the points of discontinuity of $g_{\chi_{D_t}}$ are also of Lebesgue measure zero since ϕ preserves Lebesgue measure zero sets by Lemma A.9.4. Using the Lebesgue's criterion, we conclude that $g_{\chi_{D_t}}$ is Riemann integrable over any closed interval I_t such that $D_t \subseteq I_t$ and thus the said integral does exist. $\natural$

One-Dimensional Change of Variables

Lemma A.9.6 (Changing variables in one dimension)**.** Let $\phi\colon I_t \xrightarrow{\simeq} I_x$ be a C^1-isomorphism over the 1-dimensional intervals $I_t, I_x \subseteq \mathbf{R}$ and let $f\colon I_x \to \mathbf{R}$ be a Riemann integrable map over I_x. Then the map $I_t \ni t \mapsto f\phi(t)\|\phi'(t)\|$ is a Riemann integrable map over I_t and the following relation holds

$$\int_{I_x} f(x)\,\mathrm{d}x = \int_{I_t} f\phi(t)\|\phi'(t)\|\,\mathrm{d}t.$$

Proof. All maps f, ϕ and ϕ' are bounded, thus the map $g\colon I_t \to \mathbf{R}$ given by $g(t) := f\phi(t)\|\phi'(t)\|$ is bounded. As pointed out in the proof of Corollary A.9.5, the points of discontinuity of g are only those lying in the fibers $\phi^{-1}(x)$, where $x \in I_x$ is a point of discontinuity of f — since f is Riemann integrable and ϕ^{-1} is a C^1-isomorphism, the points of discontinuity of g form a set of Lebesgue measure zero, thus g is Riemann integrable over I_t. Let $P_x := \{[x_{j-1}, x_j]\}_{j\in J}$ be a partition of I_x and define $P_t := \{\phi^{-1}([x_{j-1}, x_j])\}_{j\in J}$ as the partition of I_t. Since ϕ and ϕ^{-1} are both continuous functions over compact sets, it follows that they are both uniformly continuous — thus, taking a sequence of partitions of I_x and the induced partitions of I_t, we have $\operatorname{mesh} P_x \to 0$ if and only if $\operatorname{mesh} P_t \to 0$. Let ξ be a collection of distinguished points of P_x and $\tau := \{\phi^{-1}(\xi_j)\}_{j\in J}$ be the corresponding collection of distinguishes points of P_t. Notice that, from construction, we have the following

$$\sum_{j\in J} f(\xi_j)\|x_j - x_{j-1}\| = \sum_{j\in J} f\phi(\tau_j)\|\phi(t_j) - \phi(t_{j-1})\|.$$

Since both f and g are Riemann integrable (over their respective domains), we find that the choice of distinguished points is arbitrary, thus we may choose ξ so that, when applying the mean value theorem to ϕ, we have $\phi'(\tau_j) = \frac{\phi(t_j)-\phi(t_{j-1})}{t_j - t_{j-1}}$ — where, as before, $\tau_j := \phi^{-1}(\xi_j)$. This way we may rewrite the equation as

$$\sum_{j\in J} f(\xi_j)\|x_j - x_{j-1}\| = \sum_{j\in J} f\phi(\tau_j)\|\phi'(\tau_j)\|\,\|t_j - t_{j-1}\|.$$

Now, taking the limit $\operatorname{mesh} P_x \to 0$, which implies $\operatorname{mesh} P_t \to 0$, we find the integral equality just wanted. ♮

Corollary A.9.7. Let $\phi\colon I_t \xrightarrow{\simeq} I_x$ be a C^1-isomorphism between closed 1-dimensional intervals $I_t, I_x \subseteq \mathbf{R}$, and $f\colon I_x \to \mathbf{R}$ be a Riemann integrable map over I_x. Then we have the following relations

$$\overline{\int_{I_x}} f(x)\,\mathrm{d}x = \overline{\int_{I_t}} f\phi(t)\|\phi'(t)\|\,\mathrm{d}t \quad\text{and}\quad \underline{\int_{I_x}} f(x)\,\mathrm{d}x = \underline{\int_{I_t}} f\phi(t)\|\phi'(t)\|\,\mathrm{d}t.$$

Proof. We'll verify the first relation, the second may be obtained with an analogous proof. Construct partitions P_x and P_t for I_x and I_t, respectively, just as we've done for the proof of Lemma A.9.6. Assume for the time being that f is a non-negative map (we'll take hold of the general case just in a bit) and let $M > 0$ be a constant bounding f. Define $\varepsilon := \sup_{j\in J}\omega(\phi, [t_{j-1}, t_j])$, that is, the supremum oscillation of ϕ over each of the intervals of the partition P_t. Notice that we have

$$\begin{aligned}
\sum_{j\in J} \sup_{x\in[x_{j-1},x_j]} f(x)\|x_j - x_{j-1}\| &\leqslant \sum_{j\in J} \sup_{t\in[t_{j-1},t_j]} f\phi(t) \sup_{t\in[t_{j-1},t_j]} \|\phi'(t)\|\,\|t_j - t_{j-1}\| \\
&\leqslant \sum_{j\in J} \sup_{t\in[t_{j-1},t_j]} \left[f\phi(t) \sup_{t\in[t_{j-1},t_j]} \|\phi'(t)\| \right] \|t_j - t_{j-1}\| \\
&\leqslant \sum_{j\in J} \sup_{t\in[t_{j-1},t_j]} f\phi(t)\left(\|\phi'(t)\| + \varepsilon\right)\|t_j - t_{j-1}\| \\
&\leqslant \sum_{j\in J} \sup_{t\in[t_{j-1},t_j]} f\phi(t)\|\phi'(t)\|\,\|t_j - t_{j-1}\| + \varepsilon \sum_{j\in J} \sup_{t\in[t_{j-1},t_j]} f\phi(t)\|t_j - t_{j-1}\| \\
&\leqslant \sum_{j\in J} \sup_{t\in[t_{j-1},t_j]} f\phi(t)\|\phi'(t)\|\,\|t_j - t_{j-1}\| + \varepsilon M \operatorname{vol} I_t
\end{aligned}$$

where $\operatorname{vol} I_t$ happens to be the length of the interval, since $I_t \subseteq \mathbf{R}$. Since ϕ is uniformly continuous, we find, from the last inequalities, that in the limit $\operatorname{mesh} P_t \to 0$ (which implies in $\operatorname{mesh} P_x \to 0$):

$$\overline{\int_{I_x}} f(x)\,\mathrm{d}x \leqslant \overline{\int_{I_t}} f\phi(t)\|\phi'(t)\|\,\mathrm{d}t.$$

Notice that we can use the same analogous proof for the C^1-isomorphism ϕ^{-1} and the map $I_t \ni t \mapsto f\phi(t)\|\phi'(t)\|$, we find that

$$\overline{\int_{I_t}} f\phi(t)\|\phi'(t)\|\,\mathrm{d}t \leqslant \overline{\int_{I_x}} f(x)\,\mathrm{d}x,$$

thus the equality for the first relation has been established for non-negative f. Notice, however, that $f = \max(f, 0) - \max(-f, 0)$, for any map f, thus the equality holds for the general case. ♮

Change of Variables on Elementary C^1-Isomorphisms

Definition A.9.8 (Elementary C^1-isomorphism in $\mathbf{R}^n$)**.** Let $\{t_j\}_{j=1}^n$ and $\{x_j\}_{j=1}^n$ be basis for the euclidean space $\mathbf{R}^n$. Given any $1 \leqslant k \leqslant n$, a C^1-isomorphism $\phi\colon D_t \xrightarrow{\simeq} D_x$ such that $\phi_j(t_j) = t_j$ for every $j \neq k$, and $\phi_k(t_k) = x_k$, is said to be a k-elementary C^1-isomorphism — that is, the only coordinate of $\mathbf{R}^n$ changed under the mapping of ϕ is the k-th.

Lemma A.9.9. In the context of Theorem A.9.3, if ϕ is an elementary C^1-isomorphism, then the proposition is valid.

Proof. Let ϕ be an elementary C^1-isomorphism on the k-th coordinate. Given points $x, t \in \mathbf{R}^n$, we define $x' := (x_j)_{j=1, j\neq k}^n \in \mathbf{R}^{n-1}$ and $t' := (t_j)_{j=1, j\neq k}^n \in \mathbf{R}^{n-1}$. Moreover, for every $x', t' \in \mathbf{R}^{n-1}$, we define 1-dimensional sections of the domains D_x and D_t, respectively, as

$$\begin{aligned} D_{x_k}(x') &:= \{p \in D_x : p_k = x_k \text{ and } p_j = x'_j \text{ for } j \neq k\}, \\ D_{t_k}(t') &:= \{p \in D_t : p_k = t_k \text{ and } t_j = t'_j \text{ for } j \neq k\}. \end{aligned}$$

Let $I_x \subseteq \mathbf{R}^n$ be a closed interval containing D_x and we let $I^0_{x'} \times I_{x_k} \times I^1_{x'} := I_x$ be the representation of I_x by closed intervals $I^0_{x'} \subseteq \mathbf{R}^{k-1}$, $I_{x_k} \subseteq \mathbf{R}$ and $I^1_{x'} \subseteq \mathbf{R}^{n-k}$. Analogously, let $I_t \subseteq \mathbf{R}^n$ be a closed interval containing D_t and let $I^0_{t'} \times I_{t_k} \times I^1_{t'} := I_t$ where, as before,

$I^0_{t'} \subseteq \mathbf{R}^{k-1}$, $I_{t_k} \subseteq \mathbf{R}$ and $I^1_{t'} \subseteq \mathbf{R}^{n-k}$ are all closed intervals. Thus, we have

$$\begin{aligned}
\int_{D_x} f(x)\,\mathrm{d}x &= \int_{I_x} f_{\chi D_x}(x)\,\mathrm{d}x \\
&= \int_{I^0_{x'}\times I^1_{x'}} \left(\int_{I_{x_k}} f_{\chi D_x}(x')\,\mathrm{d}x_k\right)\mathrm{d}x' && \text{(A.46)}\\
&= \int_{I^0_{x'}\times I^1_{x'}} \left(\int_{D_{x_k}(x')} f(x)\,\mathrm{d}x\right)\mathrm{d}x'_0 && \text{(A.47)}\\
&= \int_{I^0_{t'}\times I^1_{t'}} \left(\int_{I_{t_k}} f\phi_{\chi D_t}(t'_1,\dots,t_k,\dots,t'_n)\|\det\left(\mathrm{Jac}\,\phi_{\chi D_t}(t'_1,\dots,t_k,\dots,t'_n)\right)\|\,\mathrm{d}t_k\right)\mathrm{d}t' && \text{(A.48)}\\
&= \int_{I^0_{t'}\times I_{t_k}\times I^1_{t'}} f\phi_{\chi D_t}(t)\|\det(\mathrm{Jac}\,\phi_{\chi D_t}(t))\|\,\mathrm{d}t = \int_{I_t} f\phi_{\chi D_t}(t)\|\det(\mathrm{Jac}\,\phi_{\chi D_t}(t))\|\,\mathrm{d}t && \text{(A.49)}\\
&= \int_{D_t} f\phi(t)\|\det(\mathrm{Jac}\,\phi(t))\|\,\mathrm{d}t && \text{(A.50)}
\end{aligned}$$

where we have the following use of theorems and definitions for each of the equations: Eq. (A.46) used Theorem A.8.4, Eq. (A.47) used Definition A.7.20, Eq. (A.48) used Lemma A.9.6, Eq. (A.49) used Theorem A.8.4, and, finally, Eq. (A.50) used Definition A.7.20. This proves the lemma. ♮

Proposition A.9.10. Let $f\colon E \to \mathbf{R}^n$ be a C^1 map on the open set $E \subseteq \mathbf{R}^n$, with $0 \in E$ — moreover, we impose that $f(0) := 0$ and $\mathrm{d}f(0)\colon T_0\mathbf{R}^n \to T_0\mathbf{R}^n$ is an $\mathbf{R}$-linear isomorphism. Then there exists a neighbourhood $U \subseteq E$ of $0 \in \mathbf{R}^n$, with mappings:

- $\{\phi_j\colon U \to \mathbf{R}^n\}_{j=1}^n$ of elementary C^1-isomorphisms, such that $\phi_j(0) = 0$ and $\mathrm{d}\phi(0)$ is an $\mathbf{R}$-linear isomorphism.
- We define, for each $1 \leqslant k \leqslant n$, the map $\tau_k\colon \mathbf{R}^n \to \mathbf{R}^n$ as $\tau_k(\sum_{j=1}^n x_j e_j) := \sum_{j=1}^n x_{\tau(j)} e_j$, where $\tau \in \mathrm{Aut}_{\mathtt{Grp}}([n])$ is a transposition.

These maps allow us to write $f|_U$ as the composition of elementary C^1-isomorphisms and transpositions

$$f(x) = \tau_1 \dots \tau_{n-1} g_n \dots g_1(x), \text{ for all } x \in U.$$

Proof. Define, for every $0 \leqslant k \leqslant n$, the map $p_k\colon \mathbf{R}^n \to \mathbf{R}^n$ by $p_k(\sum_{j=1}^n x_j e_j) := \sum_{j=1}^k x_j e_j$.
We do induction in the following proposition, where $1 \leqslant m \leqslant n-1$

- There exists a neighbourhood $V_m \subseteq \mathbf{R}^n$ of zero, a map $f_m \in C^1(V_m)$ — with $f_m(0) = 0$ and $\mathrm{d}f_m(0)$ linear isomorphism (we define $f_1 := f$)— for which

$$p_{m-1} f_m(x) = p_{m-1}(x), \text{ for all } x \in V_m.$$

Notice that the proposition is clearly true for $m = 1$, now, we assume it's true for $1 < m < n-1$ as our hypothesis of induction.

Notice that f_m is a map that does only change the last $x_m, \dots, x_n$ variables of its input, which allows us to write it as

$$f_m(x) = p_{m-1}(x) + \sum_{j=m}^{n} \alpha_j(x) e_j,$$

where $\alpha_j\colon V_m \to \mathbf{R}$ is a $C^1(V_m)$ map for all $m \leqslant j \leqslant n$, moreover, this implies in

$$\mathrm{d}f_m(0)(e_m) = \sum_{j=m}^{n} \partial_m \alpha_j(0) e_j.$$

Since $\mathrm{d}f_m(0)$ is an isomorphism, it cannot be the zero-map, thus there must exist some $m \leqslant k_0 \leqslant n$ for which $\partial_m \alpha_{k_0}(0) \neq 0$. Define τ_m to be the transposition interchanging the m-th and k_0-th values of its input. Define a map $\phi_m\colon V_m \to \mathbf{R}^n$ as

$$\phi_m(x) \coloneqq x + (\alpha_{k_0}(x) - x_m)e_m,$$

so that $\phi_m \in C^1(V_m)$, and only changes the m-th coordinate of its input. Moreover $\mathrm{d}\phi_m(0)$ is an isomorphism, thus there must exist a neighbourhood $U_m \subseteq V_m$ of zero for which the induced map $\phi_m\colon U_m \to \phi_m(U_m)$ is a C^1-isomorphism (which, in fact, is elementary). We define $V_{m+1} \coloneqq \phi_m(U_m)$, and also

$$f_{m+1}(y) \coloneqq \tau_m f_m \phi_m^{-1}(y), \text{ for all } y \in V_{m+1},$$

which makes $f_{m+1} \in C^1(V_{m+1})$, also $f_{m+1}(0) = 0$, and $\mathrm{d}f_{m+1}(0)$ is an isomorphism. Notice that, for all $x \in U_m$ we have

$$\begin{aligned}
p_m f_{m+1}(\phi_m(x)) &= p_m \tau_m f_m \phi_m^{-1}(\phi_m(x)) \\
&= p_m \tau_m f(x) \\
&= p_m \tau_m \left[p_{m-1}(x) + \sum_{j=m}^{n} \alpha_j(x) e_j \right] \\
&= p_m \left[p_{m-1}(x) + \alpha_{k_0}(x) e_m + \cdots + \alpha_m(x) e_{k_0} + \cdots + \alpha_n(x) e_n \right] \\
&= p_{m-1}(x) + \alpha_{k_0}(x) e_m \\
&= p_m \phi_m(x).
\end{aligned}$$

That is, since ϕ_m is a bijection in V_{m+1}, for any $y \in V_{m+1}$ we have the equality

$$p_m f_{m+1}(y) = p_m(y).$$

This concludes the proof by induction of our initial statement. We now must show that it, in fact, implies the proposition we set out to prove. For that we recall the definition of f_{m+1} and notice that, by taking τ_m from both sides we obtain $\tau_m f_{m+1}(y) =$

$\tau_m\tau_m f_m\phi_m^{-1}(y) = f_m\phi_m^{-1}(y)$ for every $y \in V_{m+1}$ — which allow us to write, taking $y = \phi_m(x)$,

$$f_m(x) = \tau_m f_{m+1}(\phi_m(x)), \text{ for every } x \in U_m.$$

This recursive definition implies in (recalling that we set $f_1 := f$):

$$f = \tau_1 f_2\phi_1 = \cdots = \tau_1 \dots \tau_{n-1} f_n \phi_{n-1} \dots \phi_1,$$

which is what we wanted since f_n is an elementary C^1-isomorphism for some neighbourhood of zero. ♮

Proof for the Theorem on Change of Variables

We now go for our (finally) last lemma before we can get to the proof of the main theorem of the section.

Lemma A.9.11. Let $D_s \xrightarrow{\psi} D_t \xrightarrow{\phi} D_x$ be C^1-isomorphisms. Moreover, let $f\colon D_x \to \mathbf{R}$ be a Riemann integrable map over D_x. If Theorem A.9.3 holds for both ψ, and for ϕ, then it is also valid for the composition $\phi\psi\colon D_s \xrightarrow{\simeq} D_x$.

Proof. Notice that, for every $s \in D_s$, $\operatorname{Jac}(\phi\psi)(s) = \operatorname{Jac}\phi(\psi(s)) \operatorname{Jac}\psi(s)$, then

$$\det(\operatorname{Jac}(\phi\psi)(s)) = \det(\operatorname{Jac}\phi(\psi(s))) \det(\operatorname{Jac}\psi(s)).$$

Therefore we can write

$$\begin{aligned}\int_{D_x} f(x)\,\mathrm{d}x = \int_{D_t} f\phi(t)\|\det(\operatorname{Jac}\phi(t))\|\,\mathrm{d}t &= \int_{D_s} f\phi\psi(s)\|\det(\operatorname{Jac}\phi(\psi(s)))\|\|\det(\operatorname{Jac}\psi(s))\|\,\mathrm{d}s \\ &= \int_{D_s} f\phi\psi(s)\|\det(\operatorname{Jac}(\phi\psi)(s))\|\,\mathrm{d}s,\end{aligned}$$

which proves the proposition. ♮

Now we are ready for the proof of Theorem A.9.3.

Proof. Define the compact set $K_t := \operatorname{supp}(t \mapsto f\phi(t)\|\det(\operatorname{Jac}\phi(t))\|)$, and, for every $t \in K_t$, let $U(t) \subseteq D_t$ be a neighbourhood of t with diameter less than $\delta(t)$, for some $\delta(t) > 0$ — for which $\phi|_{U(t)}$ can be decomposed into elementary C^1-isomorphisms. For each $t \in K_t$, let $U'(t) \subseteq U(t)$ be a neighbourhood of t with diameter less than $\delta(t)/2$ — with this, the collection $\{U'(t)\}_{t\in K_t}$ is a cover for K_t, more than that, since K_t is compact, there exists a finite collection of points $\{t_j\}_{j=1}^k$ such that $\{U'(t_j)\}_{j=1}^k$ is a cover for K_t. Define $\delta := \frac{1}{2}\min(\delta(t_j)_{j=1}^k)$ so that, for any set $A \subseteq D_t$ with closure of diameter less than δ and non-empty intersection with the support, $A \cap K_t \neq \varnothing$, the given set must be contained in some $U'(t_m)$ for some $1 \leqslant m \leqslant k$.

Let $I \subseteq \mathbf{R}^n$ be an interval containing D_t, and $P := \{I_s\}_{s\in S}$ a partition of I with mesh $P < \min(\delta, d)$ — where $d := \inf_{(a,b)\in\partial K_t\times\partial D_t}\|b-a\|$ is the minimal distance between the boundaries of K_t and D_t. Let S' be the index set such that $I_s \cap K_t \neq \varnothing$ if and only if $s \in S'$ — the set indices of intervals of the partition P with non-empty intersection with the support. Then

$$\begin{aligned}\int_{D_t} f\phi(t)\|\det(\operatorname{Jac}\phi(t))\|\,\mathrm{d}t &= \int_I f\phi_{\chi_{D_t}}(t)\|\det(\operatorname{Jac}\phi_{\chi_{D_t}}(t))\|\,\mathrm{d}t \\ &= \sum_{s'\in S'}\int_{I_{s'}} f\phi(t)\|\det(\operatorname{Jac}\phi(t))\|\,\mathrm{d}t.\end{aligned}$$

Since $I_s \in P$ are Jordan-measurable sets, so is their image $\phi(I_s) \subseteq D_x$ — this follows from Lemma A.9.4. Define the Jordan-measurable set $E := \bigcup_{s'\in S'} I_{s'}$ so that, by our construction, given any $x \in \operatorname{supp} f$, let $t := \phi^{-1}(x)$, then $f\phi(t)\|\det(\operatorname{Jac}\phi(t))\| \neq 0$ thus $t \in K_t$ — therefore, $t \in \bigcup_{s'\in S'} I_{s'}$, which implies in $x \in E$, that is, $\operatorname{supp} f \subseteq E$. Notice that, if $I_x \subseteq \mathbf{R}^n$ is an interval containing D_x, then

$$\begin{aligned}\int_{D_x} f(x)\,\mathrm{d}x = \int_{I_x} f_{\chi_{D_x}}(x)\,\mathrm{d}x &= \underbrace{\int_{I_x\smallsetminus E} f_{\chi_{D_x}}(x)\,\mathrm{d}x}_{0} + \int_E f_{\chi_{D_x}}(x)\,\mathrm{d}x \\ &= \int_E f(x)\,\mathrm{d}x = \sum_{s'\in S'}\int_{\phi(I_{s'})} f(x)\,\mathrm{d}x. \end{aligned} \tag{A.51}$$

Since ϕ is decomposed into elementary C^1-isomorphisms in any neighbourhood $U'(t)$, for $t \in K_t$, it follows that ϕ decomposes into elementary C^1-isomorphisms in every interval $I_{s'}$ for $s' \in S'$ — hence, by means of Lemma A.9.9 we conclude that, for every $s' \in S'$

$$\int_{\phi(I_{s'})} f(x)\,\mathrm{d}x = \int_{I_{s'}} f\phi(t)\|\det(\operatorname{Jac}\phi(t))\|\,\mathrm{d}t. \tag{A.52}$$

Thus, merging both Eq. (A.51) and Eq. (A.52), we obtain the desired equality. ♮

A.10 An Elementary Construction of Differential Forms

Definition A.10.1 (k-surface)**.** Let $E \subseteq \mathbf{R}^n$ be an open set. We define a *k-surface in E* to be a C^1 *map* $\Phi\colon D \to E$, where $D \subseteq \mathbf{R}^k$ is a *compact* set — this set is commonly referenced to as the parameter domain of Φ.

Definition A.10.2 (Differential k-form)**.** Let $E \subseteq \mathbf{R}^n$ be an open set. A *differential k-form* in E, for $k > 0$, is a multilinear map ω given by

$$\omega := \sum f_{j_1\dots j_k}\,\mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_k},$$

where the sum runs over the sequence of indices $(j_1, \dots, j_k)$, where $1 \leqslant j_r \leqslant n$ for each $1 \leqslant r \leqslant k$, and $f_{j_1\dots j_k}\colon E \to \mathbf{R}$ are continuous maps — that is, 0-forms. A differential k-form is said to be of class C^p if each map $f_{j_1\dots j_k}$ is of class C^p.

The form ω is said to be a *basic* k-form if we have the ordering $1 \leqslant j_1 < \cdots < j_k \leqslant n$ for its indices — in this case we commonly denote each sequence of indexes as $I := (j_1, \dots, j_k)$ and write

$$\omega = \sum_I f_I\,\mathrm{d}x_I.$$

Definition A.10.3 (Line integral)**.** The integrals of 1-forms are called *line integrals*.

Example A.10.4. Lets calculate a line integral on the 1-surface (C^1 curve) $\gamma\colon [0, 2\pi] \to \mathbf{R}^2$ defined by $t \mapsto (a\cos(t), b\sin(t))$, for constants $a, b > 0$. Let $\omega = x\,\mathrm{d}y$ be a 1-form on $\mathbf{R}^2$. Measuring this curve with respect to ω yields

$$\int_\gamma \omega = \int_0^{2\pi} a\cos(t)[b\cos(t)]\,\mathrm{d}t = ab\int_0^{2\pi} \frac{\cos(2t)-1}{2}\,\mathrm{d}t = -ab\pi.$$

Definition A.10.5 (Surface area)**.** Let $\Phi\colon D \to E$ be a k-surface on the open set $E \subseteq \mathbf{R}^n$, and $\omega := \sum_I f_{i_1\dots i_k}\,\mathrm{d}x_{i_1\dots i_k}$. We define the area of the surface Φ on ω to be given by

$$\int_\Phi \omega := \int_D \sum_I f_{i_1\dots i_k}(\Phi(t))\det\left[\partial_j \Phi_{i_r}(t)\right]_{1\leqslant r,j\leqslant k}\mathrm{d}t$$

Definition A.10.6 (Piecewise smooth)**.** We define the concept of a piecewise smooth surface in an inductive manner. A point is a zero dimensional surface of any smoothness class. A surface $S \subseteq \mathbf{R}^n$ of dimension k is piecewise smooth if, after a countable collection of at most $(k-1)$-dimensional piecewise smooth surfaces can be removed from S, the resulting surface can be decomposed into a countable collection of k-dimensional smooth surfaces.

Theorem A.10.7 (A zero form has null coefficients)**.** Let $\omega := \sum_J f_J\,\mathrm{d}x_J$ be a k-form in an open set $E \subseteq \mathbf{R}^n$, in its standard representation. If $\omega = 0$ in E (that is $\omega(\Phi) = 0$ for all k-surface Φ in E), then $f_J(x) = 0$ for all $x \in E$.

Proof. Suppose, for the sake of contradiction, that there exists an index sequence $S := (s_1, \dots, s_k)$ such that $f_S(y) \neq 0$ for some $y \in E$. Since f_S is continuous, let $\delta > 0$ be such that $f_J(x) > 0$ whenever $\|x_i - y_i\| \leqslant \delta$, for all $1 \leqslant i \leqslant n$. Let $D \subseteq \mathbf{R}^k$ be the compact set given by

$$D := \{t \in \mathbf{R}^k : \|t_j\| \leqslant \delta \text{ for all } 1 \leqslant j \leqslant k\}.$$

Define a k-surface $\Phi\colon D \to E$ to be given by

$$\Phi(t) := v + \sum_{j=1}^k t_j e_{s_j}, \text{ for all } t \in D.$$

Notice that D was chosen so that $f_S(\Phi(t)) > 0$ for all $t \in D$. Notice that

$$\int_\Phi \omega = \int_D \sum_J f_J(\Phi(t))\det[\partial_i \Phi_{j_r}(t)]_{1\leqslant i,r\leqslant k}\,\mathrm{d}t = \int_D f_S(\Phi(t))\det[\partial_i \Phi_{s_r}]_{1\leqslant i,r\leqslant k}\,\mathrm{d}t$$

since for all indexing sequence $J := (j_1, \dots, j_k) \neq S$ the matrix $[\partial_i \Phi_{j_r}(t)]_{1\leqslant i,r\leqslant k}$ has at least one column equal to zero — since there must exists at least one $j_{r_0} \in J \setminus S$, thus $\Phi_{j_0}(t) = y_{j_0}$ for all $t \in D$ and thus the column $[\partial_i \Phi_{j_0}(t)]_{1\leqslant i\leqslant k} = 0$, which implies in

$\det[\partial_i \Phi_{j_r}(t)]_{1\leqslant i,r\leqslant k} = 0$. On the other hand, notice that $\partial_i \Phi_{s_r}(t) = \delta_{ir}$ for all $1 \leqslant i, r \leqslant k$, thus $\det[\partial_i \Phi_{s_r}(t)]_{1\leqslant i,r\leqslant k} = 1$. We conclude that

$$\int_\Phi \omega = \int_D f_S(t)\,\mathrm{d}t,$$

which is strictly positive since $f_S(y) > 0$. Therefore $\omega(\Phi) \neq 0$, which is a contradiction — thus there must exist no indexing sequence S and therefore every coefficient is the zero-map. ♮

Differential Operator

Definition A.10.8. Let $f\colon E \to \mathbf{R}$ be a map of class C^1, we define an operator d which transforms any 0-form f into

$$\mathrm{d}f := \sum_{j=1}^{n} \partial_j f\,\mathrm{d}x_j.$$

Now, for any k-form $\omega := \sum_J f_J\,\mathrm{d}x_J$, where $k \geqslant 1$ and $f_J\colon E \to \mathbf{R}$ is again a map of class C^1 (a 0-form) we associate the $(k+1)$-form $\mathrm{d}\omega$ — which is defined by

$$\mathrm{d}\omega := \sum_J \mathrm{d}f_J \wedge \mathrm{d}x_J.$$

Example A.10.9. Let $E \subseteq \mathbf{R}^n$ be an open set, and $\gamma\colon [0,1] \to E$ be a 1-surface (that is, a continuous differentiable curve). If we let $f\colon E \to \mathbf{R}$ be a C^1 map, we have from the definition that the integral over the curve γ of $\mathrm{d}f$ is given by — recalling Theorem A.3.21,

$$\int_\gamma \mathrm{d}f = \int_0^1 \sum_{j=1}^{n} \partial_j f(\gamma(t))\gamma_j'(t)\,\mathrm{d}t = \int_0^1 \mathrm{d}(f\gamma)(t)\,\mathrm{d}t = f\gamma(1) - f\gamma(0).$$

since $f\gamma\colon [0,1] \to \mathbf{R}$ is a continuously differentiable map.

Theorem A.10.10. The following are properties of the differential operator on forms. Let $E \subseteq \mathbf{R}^n$ be some open set.

(a) (Skew-product rule) Let ω be a k-form and γ be a m-form, both of class $C^1(E)$. Then

$$\mathrm{d}(\omega \wedge \gamma) = (\mathrm{d}\omega) \wedge \gamma + (-1)^k \omega \wedge \mathrm{d}\gamma.$$

(b) If ω is a k-form of class $C^2(E)$, then $\mathrm{d}(\mathrm{d}\omega) = 0$.

Proof. (a) Let $\omega := \sum_I f_I \mathrm{d}x_I$ and $\gamma(x) := \sum_J g_J \mathrm{d}x_J$ for C^1 coefficients $f_I, g_J\colon E \rightrightarrows \mathbf{R}$ — if k or m are zero, we just omit the 1-forms from the definitions. From the wedge product we have

$$\mathrm{d}(\omega \wedge \gamma) = \sum_{I,J} \mathrm{d}\left(f_I \cdot g_J\,\mathrm{d}x_I \wedge \mathrm{d}x_J\right) = \sum_{I,J} \mathrm{d}(f_I \cdot g_J) \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J.$$

Define, for each pair I and J, the indexing sequence $(\!(I,J)\!)$ consisting of the increasing ordered union of the sequences I and J. Moreover, for each $(\!(I,J)\!)$, there will be an associated sign

$$\operatorname{sign}(I,J) := |\{j-i : j-i < 0 \text{ for } (i,j) \in I \times J\}|, \tag{A.53}$$

that is, the number of times the indices of J is greater than the ones from I. From the skew-commutativity property (Proposition 6.5.12),

$$\begin{aligned}
\mathrm{d}(\omega \wedge \gamma) &= \sum_{I,J} \operatorname{sign}(I,J)\, \mathrm{d}(f_I \cdot g_J) \wedge \mathrm{d}x_{(\!(I,J)\!)} \\
&= \sum_{(I,J)} \operatorname{sign}(I,J) \big(\mathrm{d}f_I \cdot g_J + f_I \cdot \mathrm{d}g_J\big) \wedge \mathrm{d}x_{(\!(I,J)\!)} \\
&= \sum_{I,J} \big(\mathrm{d}f_I \cdot g_J + f_I \cdot \mathrm{d}g_J\big) \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J.
\end{aligned}$$

From the distributive property, associativity and skew-commutativity we get

$$\begin{aligned}
\mathrm{d}(\omega \wedge \gamma) &= \sum_{I,J} (g_J\, \mathrm{d}f_I \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J + f_I\, \mathrm{d}g_J \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J) \\
&= \sum_{I,J} (\mathrm{d}f_I \wedge \mathrm{d}x_I) \wedge (g_J\, \mathrm{d}x_J) + (-1)^k (f_I\, \mathrm{d}x_I) \wedge (\mathrm{d}g_J \wedge \mathrm{d}x_J) \\
&= \sum_{I,J} (\mathrm{d}f_I \wedge \mathrm{d}x_I) \wedge (g_J\, \mathrm{d}x_J) + (-1)^k \sum_{I,J} (f_I\, \mathrm{d}x_I) \wedge (\mathrm{d}g_J \wedge \mathrm{d}x_J) \\
&= \Big(\sum_I \mathrm{d}f_I \wedge \mathrm{d}x_I\Big) \wedge \Big(\sum_J g_J\, \mathrm{d}x_J\Big) + (-1)^k \Big(\sum_I f_I\, \mathrm{d}x_I\Big) \wedge \Big(\sum_J \mathrm{d}g_J \wedge \mathrm{d}x_J\Big) \\
&= \mathrm{d}\omega \wedge \lambda + (-1)^k \omega \wedge \mathrm{d}\lambda.
\end{aligned}$$

(b) For the case of a zero form $f\colon E \to \mathbf{R}$, of class C^2, we have

$$\mathrm{d}(\mathrm{d}f) = \mathrm{d}\Big(\sum_{j=1}^n \partial_j f\, \mathrm{d}x_j\Big) = \sum_{j=1}^n \mathrm{d}(\partial_j f)\, \mathrm{d}x_j = \sum_{i,j=1}^n \partial_{ij} f\, \mathrm{d}x_i \wedge \mathrm{d}x_j$$

Notice however that $\partial_{ij} f = \partial_{ji} f$ for any $1 \leqslant i,j \leqslant n$, thus each pair (i,j) and (j,i) of the sum cancel with each other — since $\mathrm{d}x_i \wedge \mathrm{d}x_j = -\mathrm{d}x_j \wedge \mathrm{d}x_i$ — which implies in $\mathrm{d}(\mathrm{d}f) = 0$.

For the general case, if $\omega = \sum_I f_I\, \mathrm{d}x_I$ is any C^2 k-form, since $\mathrm{d}^2\omega = \sum_I \mathrm{d}^2 f_I\, \mathrm{d}x_I$ and $\mathrm{d}^2 f_I = 0$, for all I, it follows immediately that $\mathrm{d}^2\omega = 0$.

♮

Change of Variables — The Pullback Operation

Let $\omega := \sum_I f_I\, \mathrm{d}x_I$ be a k-form in an open set $E \subseteq \mathbf{R}^n$ and $\phi\colon V \to E$ be a C^1 map from another open set $V \subseteq \mathbf{R}^m$. Notice that there arises a *natural pullback* operation ϕ^* that allows for a *change of variables*

$$\phi^*(\omega) = \sum_I f_I \phi \, \mathrm{d}v_I, \tag{A.54}$$

where v_I represents the coordinates coming from V.

Proposition A.10.11 (Pullback properties)**.** Let ω and λ be, respectively, a k-form and an m-form in the open set $E \subseteq \mathbf{R}^n$. Let $\phi\colon V \to E$ be a C^1 map. The following properties hold:

(a) If $k = m$, then $\phi^*(\omega + \lambda) = \phi^*(\omega) + \phi^*(\lambda)$.

(b) $\phi^*(\omega \wedge \lambda) = \phi^*(\omega) \wedge \phi^*(\lambda)$.

(c) If ω is C^1 and ϕ is of class C^2, then $\mathrm{d}(\phi^*(\omega)) = \phi^*(\mathrm{d}\omega)$.

Proof. Let $\omega \coloneqq \sum f_I \, \mathrm{d}x_I$ and $\lambda \coloneqq \sum g_J \, \mathrm{d}x_J$.

(a) If $k = m$, then $\omega + \lambda$ is a k-form in E, and we have

$$\phi^*(\omega + \lambda) = \phi^*\left(\sum f_I \, \mathrm{d}x_I + \sum g_J \, \mathrm{d}x_J\right) = \sum f_I \phi \, \mathrm{d}v_I + \sum g_J \phi \, \mathrm{d}v_J = \phi^*(\omega) + \phi^*(\lambda).$$

(b) Notice that

$$\begin{aligned} \phi^*(\omega \wedge \lambda) &= \phi^*\left(\sum_{I,J} f_I \cdot g_J \, \mathrm{d}x_I \wedge \mathrm{d}x_J\right) = \sum_{I,J} (f_I \cdot g_J)\phi \, \mathrm{d}v_I \wedge \mathrm{d}v_J \\ &= \sum_{I,J} (f_I \phi) \cdot (g_J \phi) \, \mathrm{d}v_I \wedge \mathrm{d}v_J = \sum_{I,J} (f_I \phi \, \mathrm{d}v_I) \wedge (g_J \phi \, \mathrm{d}v_J) \\ &= \phi^*(\omega) \wedge \phi^*(\lambda). \end{aligned}$$

(c) We first prove the equality for the base case of a 0-form. Let $h\colon E \to \mathbf{R}$ be a 0-form, then

$$\begin{aligned} \phi^*(\mathrm{d}h) &= \phi^*\left(\sum_{i=1}^n \partial_i h \, \mathrm{d}x_i\right) = \sum_{i=1}^n (\partial_i h)\phi \, \mathrm{d}v_i \\ &= \sum_{i=1}^n (\partial_i h)\phi \left(\sum_{j=1}^n \partial_j v_i \, \mathrm{d}x_j\right) = \sum_{j=1}^n \partial_j (h\phi) \, \mathrm{d}x_j \\ &= \mathrm{d}(\phi^*(h)) \end{aligned}$$

We now turn to the general case. Notice that $\phi^*(\omega) = \sum_I \phi^*(f_I) \, \mathrm{d}v_I = \sum_I \phi^*(f_I) \, \phi^*(\mathrm{d}x_I)$, since $\phi^*(\mathrm{d}x_I) = \mathrm{d}v_I$. Therefore, assuming ϕ is of class C^2, we have

$$\begin{aligned} \mathrm{d}(\phi^*\omega) &= \sum_I \mathrm{d}(\phi^* f_I) \wedge \mathrm{d}v_I = \sum_I \phi^*(\mathrm{d}f_I) \wedge \phi^*(\mathrm{d}x_I) \\ &= \sum_I \phi^*(\mathrm{d}f_I \wedge \mathrm{d}x_I) = \phi^*\left(\sum_I \mathrm{d}f_I \wedge \mathrm{d}x_I\right) \\ &= \phi^*(\mathrm{d}\omega). \end{aligned}$$

The need for ϕ to be C^2 comes from the fact that the first equality — by Theorem A.10.10 — is obtained by noting that

$$\mathrm{d}(\phi^*(f)\,\mathrm{d}v_I) = \mathrm{d}(\phi^*(f)) \wedge \mathrm{d}v_I + \phi^*(f) \wedge \underbrace{\mathrm{d}(\mathrm{d}v_I)}_{0} = \mathrm{d}(\phi^* f) \wedge \mathrm{d}v_I.$$

♮

Lemma A.10.12 (Pullback composition)**.** Let ω be a k-form in an open set $W \subseteq \mathbf{R}^p$. Consider two composable C^1 maps $\phi \colon E \to V$ and $\psi \colon V \to W$, where $U \subseteq \mathbf{R}^n$ and $V \subseteq \mathbf{R}^m$. Then the following equality holds

$$\phi^*(\psi^*(\omega)) = (\psi\phi)^*(\omega).$$

Moreover, these are k-forms in U.

Proof. Define $\omega \coloneqq \sum_I f_I \,\mathrm{d}w_I$. Thus

$$\phi^*(\psi^*(\omega)) = \phi^*\left(\sum_I f_I\psi \,\mathrm{d}v_I\right) = \sum_I (f_I\psi)\phi \,\mathrm{d}u_I = \sum_I f_I(\psi\phi)\,\mathrm{d}u_I = (\psi\phi)^*(\omega).$$

♮

Lemma A.10.13. Let ω be a k-form in an open set $E \subseteq \mathbf{R}^n$. We consider two k-surfaces $\Phi \colon D \to E$ and $\mathrm{id}_D \colon D \to \mathbf{R}^k$, with $\mathrm{id}_D(t) \coloneqq t$ — where $D \subseteq \mathbf{R}^k$ is a compact set. Then

$$\int_\Phi \omega = \int_{\mathrm{id}_D} \Phi^*(\omega).$$

Proof. Define $\omega \coloneqq \sum_I f_I \,\mathrm{d}x_I$, and $\Phi_j \coloneqq \pi_j \Phi$ as the j-th projection of the surface Φ — that is

$$\Phi^*(\omega) = \sum_I f_I \Phi \,\mathrm{d}\Phi_I.$$

Defining $I \coloneqq (i_p)_{p=1}^k$, and since $\mathrm{d}\Phi_{i_p} = \sum_{q=1}^k \partial_q \Phi_{i_p} \,\mathrm{d}t_q$ for all $1 \leqslant p \leqslant k$, we see that

$$\begin{aligned}
\mathrm{d}\Phi_{i_1} \wedge \cdots \wedge \mathrm{d}\Phi_{i_k} &= \left(\sum_{q=1}^k \partial_q \Phi_{i_1} \,\mathrm{d}t_q\right) \wedge \cdots \wedge \left(\sum_{q=1}^k \partial_q \Phi_{i_k} \mathrm{d}t_q\right) \\
&= \sum_{1 \leqslant i_1, \ldots, i_k \leqslant k} \left[\prod_{q=1}^k \partial_q \Phi_{i_q}\right] \mathrm{d}t_{i_1} \wedge \cdots \wedge \mathrm{d}t_{i_k} \\
&= \sum_{1 \leqslant i_1, \ldots, i_k \leqslant k} \left[\prod_{q=1}^k \partial_q \Phi_{i_q}\right] \big(\mathrm{sign}(i_1, \ldots, i_k)\,\mathrm{d}t_1 \wedge \cdots \wedge \mathrm{d}t_k\big) \\
&= \det[\partial_q \Phi_{i_p}]_{p,q=1}^n \,\mathrm{d}t_1 \wedge \cdots \wedge \mathrm{d}t_k \\
&= \det(\mathrm{Jac}\,\Phi)\,\mathrm{d}t_1 \wedge \cdots \wedge \mathrm{d}t_k.
\end{aligned}$$

Where the sign map is defined as in Eq. (A.53). Therefore, combining equations we find

$$\Phi^*(\omega) = \sum_I f_I\Phi \det(\operatorname{Jac}\Phi)\,\mathrm{d}t_1 \wedge \cdots \wedge \mathrm{d}t_k,$$

thus indeed

$$\begin{aligned}
\int_{\mathrm{id}_D} \Phi^*(\omega) &= \int_{\mathrm{id}_D} \sum_I f_I\Phi(t) \det(\operatorname{Jac}\Phi(t))\,\mathrm{d}t_1 \wedge \cdots \wedge \mathrm{d}t_k \\
&= \int_D \sum_I f_I\Phi(t) \det(\operatorname{Jac}\Phi(t))\,\mathrm{d}t \\
&= \int_\Phi \omega.
\end{aligned}$$

♮

Theorem A.10.14. Let $\psi\colon V \to E$ be a map of class C^1 on open sets $V \subseteq \mathbf{R}^m$, and $E \subseteq \mathbf{R}^n$. Consider a k-surface $\Phi\colon D \to V$, where $D \subseteq \mathbf{R}^k$ is compact, and a k-form ω in E. Then we have the following equality

$$\int_{\psi\Phi} \omega = \int_\Phi \psi^*(\omega).$$

Proof. The theorem is a direct consequence of the preceding lemmas — notice that

$$\int_{\psi\Phi} \omega = \int_{\mathrm{id}_D} (\psi\Phi)^*\omega = \int_{\mathrm{id}_D} \Phi^*(\psi^*(\omega)) = \int_\Phi \psi^*(\omega).$$

♮

Simplexes and Chains

Notation A.10.15. For the purposes of this section, we are going to define the collection

$$\Delta_{\mathrm{c}}^k := \left\{ (t_1, \ldots, t_k) \in \mathbf{R}^k : \sum_{j=1}^k t_j \leqslant 1, \text{ and } t_j \geqslant 0 \text{ for all } 1 \leqslant j \leqslant k \right\}.$$

which represents the *k-simplex* obtained by the corner of a k-dimensional unit cube.

Definition A.10.16 (Affine map)**.** Given vector spaces V and L, we say that $f\colon V \to L$ is a affine map if $f(x) = \ell + \phi(x)$ for all $x \in V$, where $\phi\colon V \to L$ is a k-linear map and $\ell = f(0) \in L$.

Definition A.10.17 (Oriented affine k-simplex)**.** Let $p_0, \ldots, p_k \in \mathbf{R}^n$ be any points. We define the oriented affine k-simplex induced by the points $p_0, \ldots, p_k$ in $\mathbf{R}^n$ as an affine map $\sigma\colon \Delta_{\mathrm{c}}^k \to \mathbf{R}^n$ given by

$$\sigma(\alpha) := p_0 + \sum_{j=1}^k \alpha_j(p_j - p_0), \text{ for all } \alpha \in \Delta_{\mathrm{c}}^k.$$

We can also denote σ by $[p_0, \dots, p_k]$ as in Definition 6.6.12. If we define the **R**-linear map $A\colon \mathbf{R}^k \to \mathbf{R}^n$ to be given by $A(e_j) \coloneqq p_j - p_0$, then we can write $\sigma(\alpha) = p_0 + A(\alpha)$.

If $\tau \in S_{k+1}$ is a permutation on $k+1$ elements, and $\mu \coloneqq [p_{\tau(0)}, p_{\tau(1)}, \dots, p_{\tau(k)}]$ is an oriented affine k-simplex, also induced by the points $p_0, \dots, p_k$, we say that

$$\tau = \operatorname{sign}(\tau)\sigma.$$

Moreover, μ is said to have the *same orientation* of σ if $\operatorname{sign}(\tau) > 0$, otherwise, if $\operatorname{sign}(\tau) < 0$, then μ is said to have the *opposite orientation* of σ.

In the special case where $k = n$ and the collection $\{p_j - p_0\}_{j=1}^k$ is linearly independent, we say that:

- σ is *positively* oriented if $\det A > 0$.
- σ is *negatively* oriented if $\det A < 0$.

In particular, the simplex $\mathrm{id}_{\mathbf{R}^k} \coloneqq [0, e_1, \dots, e_k]$ in $\mathbf{R}^k$ is positively oriented.

Example A.10.18. A special example of an oriented affine k-simplex occurs for the case $k = 0$, where we get a simplex induced by a single point $p_0 \in \mathbf{R}^k$ — in such case, two 0-simplexes are conceivable, $\sigma = p_0$ or $\sigma = -p_0$.

Definition A.10.19. Let $\sigma = \varepsilon p_0$ be an oriented affine 0-simplex — where $\varepsilon \in \{-1, 1\}$ and $p_0 \in \mathbf{R}^k$. If $f\colon E \to \mathbf{R}$ is a 0-form in the open set $E \subseteq \mathbf{R}^n$, we define its integral over the 0-simplex σ as

$$\int_\sigma f \coloneqq \varepsilon f(p_0).$$

Proposition A.10.20. Let $\sigma\colon \Delta^k_{\mathrm{c}} \to E$ be an oriented affine k-simplex in an open set $E \subseteq \mathbf{R}^n$, and ω be any k-form in E. If $\mu = \varepsilon\sigma$ where $\varepsilon \in \{-1, 1\}$[2], then

$$\int_\mu \omega = \varepsilon \int_\sigma \omega.$$

Proof. For the base case $k = 0$, we have Definition A.10.19. Now, let $k \geqslant 1$ and define, for each $1 \leqslant j_0 \leqslant k$, the transposition $\tau_{j_0} \in S_{k+1}$ given by $\tau_{j_0}(0) = j_0$ and $\tau_{j_0}(j_0) = 0$ — while $\tau_{j_0}(i) = i$ for all $i \neq 0, j_0$. Suppose $\sigma \coloneqq [p_0, \dots, p_k]$, then since $\operatorname{sign}(\tau_{j_0}) = -1$, the oriented affine k-simplex $\mu \coloneqq [p_{\tau_{j_0}}(0), \dots, p_{\tau_{j_0}(k)}]$ is such that $\mu = \operatorname{sign}(\tau_{j_0})\sigma = -\sigma$ — moreover,

$$\mu(u) = p_{j_0} + \sum_{j=1}^{k} u_i(p_{\tau_{j_0}(i)} - p_{j_0}) \coloneqq p_{j_0} + B(u), \text{ for all } u \in \Delta^k_{\mathrm{c}},$$

where $B\colon \mathbf{R}^k \to \mathbf{R}^n$ is a linear map with matrix representation $B = [p_{\tau_{j_0}(j)} - p_{j_0}]_{j=1}^k$ (notice we are disregarding the column with $j = 0$), where each $p_{\tau_{j_0}(j)} - p_{j_0} = [\partial_j \mu_i]_{i=1}^n$

[2] If $\sigma = [p_0, \dots, p_k]$, we say that $\mu = \operatorname{sign}(\tau)\sigma$ if $\mu = [p_{\tau(0)}, \dots, p_{\tau(k)}]$ — where $\tau \in S_{k+1}$ is a permutation on $k+1$ elements.

is the j-th column of B. Therefore, if $\omega := \sum_I f_I \,\mathrm{d}x_I$ and $I := (i_p)_{p=1}^k$ for each I, we obtain

$$\int_\mu \omega = \int_{\Delta_c^k} \sum_I f_I \mu(t) \det(B) \,\mathrm{d}t.$$

If we let $A\colon \mathbf{R}^k \to \mathbf{R}^n$ be the linear map defined by $A(u) := \sum_{i=1}^k u_i(p_i - p_0)$ for all $u \in \mathbf{R}^k$, we can write $\sigma(u) = p_0 + A(u)$. Notice however that, for each $1 \leqslant i \leqslant k$ with $i \neq j_0$, we have the relation $B(e_i) = A(e_i) - A(e_{j_0})$ (which does not affect the value of the determinant) — while $B(e_{j_0}) = -A(e_{j_0})$ (which multiplies the determinant by -1). Therefore $\det B = -\det A$. Hence,

$$\int_\mu \omega = \int_{\Delta_c^k} \sum_I f_I \mu(t)(-\det A) \,\mathrm{d}t = \int_{\Delta_c^k} \sum_I f \sigma(t) \det(A) \,\mathrm{d}t = \int_\sigma \omega. \tag{A.55}$$

For the case where we transpose a pair of indices i and j, with $0 < i < j \leqslant k$, then $\mu(u) = p_0 + C(u)$ — where C has all columns matching to those of A, except the interchange between the i-th and j-th columns, which implies in $\det C = -\det A$. Therefore we get again the same as in Eq. (A.55), which finishes the proof. ♮

Definition A.10.21 (Affine k-chain)**.** Let $E \subseteq \mathbf{R}^n$ be an open set. A finite collection of oriented affine k-simplexes in E is said to be an *affine k-chain* in E. It is to be noted that the finite collection may contain some *multiplicities* — that is, some of the k-simplexes can coincide.

If $\Gamma := (\sigma_j)_{j=1}^m$ is an affine k-chain in E, we define

$$\int_\Gamma \omega := \sum_{j=1}^m \int_{\sigma_j} \omega.$$

We'll usually denote Γ by the *formal sum* $\sum_{j=1}^m \sigma_j$ *mapping* $\omega \overset{\Gamma}{\mapsto} \sum_j \int_{\sigma_j} \omega$.

Definition A.10.22 (Boundary)**.** Let $k \geqslant 1$. Given an oriented affine k-simplex $\sigma := [p_0, \dots, p_k]$, we define the *boundary* of σ to be an affine $(k-1)$-chain defined by

$$\partial\sigma := \sum_{j=0}^k (-1)^j [p_0, \dots, p_{j-1}, p_{j+1}, \dots, p_k].$$

Definition A.10.23 (Differentiable simplexes and chains)**.** Let $E \subseteq \mathbf{R}^n$ and $V \subseteq \mathbf{R}^m$ be open sets. If $\phi\colon E \to V$ is a map of class C^2, and σ is an oriented affine k-simplex in E, then the induced map

$$\Phi := \phi\sigma\colon \Delta_c^k \longrightarrow V$$

is a k-surface in V. The surface Φ is said to be an *oriented k-simplex of class C^2* in V. Moreover, Φ has a boundary defined by

$$\partial\Phi := \phi(\partial\sigma),$$

therefore, $\partial\Phi$ is a $(k-1)$-chain of class C^2 in V.

A finite collection $\Psi := (\Phi_j)_{j=1}^r$ of oriented k-simplexes of class C^2 in V is said to be a *k-chain of class C^2 in V*. Given a k-form ω in V, we define

$$\int_\Psi \omega := \sum_{j=1}^{r} \int_{\Phi_j} \omega.$$

As expected, the boundary of Ψ is defined to be the $(k-1)$-chain of class C^2 in V given by $\partial\Psi := \sum_{j=1}^r \partial\Phi_j$.

In the context of the last definition, given an affine k-chain $\Gamma := \sum_{j=1}^r \sigma_j$ in E, the map ϕ induces a corresponding k-chain of class C^2 in V, namely, $\sum_{j=1}^r \phi\sigma_j$.

Oriented Boundaries on Sets

Definition A.10.24 (Positively oriented boundaries on a set)**.** Let $\phi\colon \Delta_c^n \rightarrowtail \mathbf{R}^n$ be an injective map of class C^2, whose Jacobian determinant is positive in the interior of Δ_c^k. Let $E := \phi(\Delta_c^n)$ — then, by the inverse map theorem (see Theorem A.5.8), E is the closure of an open set of $\mathbf{R}^n$. We define the *positively oriented boundary of the set E* to be the $(n-1)$-chain

$$\partial E := \partial\phi = \phi(\partial\,\mathrm{id}_{\mathbf{R}^n}),$$

where $\mathrm{id}_{\mathbf{R}^n} := [0, e_1, \ldots, e_n]$ — that is to say that, for any $(n-1)$-form ω in E, we have

$$\int_{\partial E} \omega := \int_{\partial\phi} \omega.$$

If $\{E_j\}_{j=1}^r$ is a collection of subsets of $\mathbf{R}^n$ with disjoint *interior*, let $\{\phi_j : \Delta_c^n \rightarrowtail \mathbf{R}^n\}_{j=1}^r$ be an associated collection of injective C^2-maps with positive Jacobian determinant in the interior of Δ_c^n, and such that $\phi_j(\Delta_c^n) = E_j$. We define the positively oriented boundary of the set $\Omega := E_1 \cup \cdots \cup E_r$ as the $(n-1)$-chain

$$\partial\Omega := \partial\phi_1 + \cdots + \partial\phi_r, \ \text{that is} \ \int_{\partial\Omega} \omega := \sum_{j=1}^{r} \int_{\partial\phi_j} \omega,$$

where ω is any $(n-1)$-form in Ω.

Proposition A.10.25. Let $\phi, \psi\colon \Delta_c^n \rightrightarrows \mathbf{R}^n$ be injective C^2-maps with positive Jacobian determinant in the interior of Δ_c^n. If $\phi(\Delta_c^n) = \psi(\Delta_c^n)$, then $\partial\phi = \partial\psi$ — that is, for every $(n-1)$-form ω in the image of the maps, we have

$$\int_{\partial\phi} \omega = \int_{\partial\psi} \omega.$$

Proof.

Prove

♮

Example A.10.26. Let $S\colon [0,\pi]\times[0,2\pi]\to\mathbf{R}^3$ be a 2-surface in $\mathbf{R}^3$, defined by

$$S(u,v) := (\sin(u)\cos(v), \sin(u)\sin(v), \cos(u)).$$

Notice that the positively oriented boundary of S corresponds to the 4 distinct curves induced by S under the positively oriented boundary of the rectangle $[0,\pi]\times[0,2\pi]$ — that is, $\partial S = \sum_{j=1}^{4}\gamma_j$ where we have

$$\begin{aligned}
&\gamma_1{:}[0,\pi] \longrightarrow \mathbf{R}^3 \text{ mapping } u \longmapsto S(u,0) = (\sin(u),0,\cos(u));\\
&\gamma_2{:}[0,2\pi] \longrightarrow \mathbf{R}^3 \text{ mapping } v \longmapsto S(\pi,v) = (0,0,-1);\\
&\gamma_3{:}[0,\pi] \longrightarrow \mathbf{R}^3 \text{ mapping } u \longmapsto S(\pi-u,2\pi) = (\sin(u),0,-\cos(u));\\
&\gamma_4{:}[0,2\pi] \longrightarrow \mathbf{R}^3 \text{ mapping } v \longmapsto S(0,2\pi-v) = (0,0,1).
\end{aligned}$$

Now, given any 1-form $\omega := f_1\,\mathrm{d}x_1 + f_2\,\mathrm{d}x_2 + f_3\,\mathrm{d}x_3$ in $\mathbf{R}^3$, we have

$$\begin{aligned}
\int_{\partial S}\omega &= \sum_{j=1}^{4}\int_{\gamma_j}\omega\\
&= \int_0^{\pi}\sum_{i=1}^{3} f_i\gamma_1(x)\,\underbrace{\partial\gamma_1^{(i)}(x)}_{0}\,\mathrm{d}x + \int_{\gamma_2}\omega + \int_{\gamma_3}\omega + \int_0^{2\pi}\sum_{i=1}^{3} f_i\gamma_4(x)\,\underbrace{\partial\gamma_4^{(i)}(x)}_{0}\,\mathrm{d}x\\
&= \int_{\gamma_2}\omega + \int_{\gamma_3}\omega.
\end{aligned}$$

Notice however that if we make the change of variables from x to $\pi - x$ in the integral $\int_{\gamma_3}\omega$, we obtain

$$\begin{aligned}
\int_{\gamma_3}\omega &= \int_0^{\pi} -f_1\gamma_3(x)\cos(x) + f_3\gamma_3(x)\sin(x)\,\mathrm{d}x\\
&= -\int_{\pi}^{0} -f_1\gamma_3(\pi-x)\cos(x) + f_3\gamma_3(\pi-x)\sin(x)\,\mathrm{d}x\\
&= \int_0^{\pi} -f_1\gamma_2(x)\cos(x) + f_3\gamma_2(x)\sin(x)\,\mathrm{d}x\\
&= -\int_{\gamma_2}\omega
\end{aligned}$$

Where the third equality may be obtained by noting that $\gamma_3(\pi-x) = \gamma_2(x)$. We conclude that

$$\int_{\partial S}\omega = 0,$$

and, more generally, $\partial S = 0$.

Stoke's Theorem

Theorem A.10.27 (Stoke's)**.** Let Ψ be a k-chain of class C^2 in an open set $V \subseteq \mathbf{R}^n$, and let ω be a $(k-1)$-form of class C^1 in V. Then the following equality holds

$$\int_{\Psi} \mathrm{d}\omega = \int_{\partial\Psi} \omega.$$

Proof. Since $\Psi = \sum_{j=1}^{r} \Phi_j$ for some finite collection $(\Phi_j)_{j=1}^{r}$ of oriented k-simplexes of class C^2, it suffices to prove that the theorem holds for any oriented k-simplex of class C^2, since $\int_{\Psi} \gamma = \sum_{j=1}^{r} \int_{\Phi_j} \gamma$ for any $(k-1)$-form γ in V.

Let Φ be any oriented k-simplex of class C^2 in V, and consider the positively oriented affine k-simplex $\mathrm{id}_{\mathbf{R}^k} := [0, e_1, \ldots, e_k] \colon \Delta_{\mathrm{c}}^k \to \mathbf{R}^k$. From Definition A.10.23, there must exist an open set $E \subseteq \mathbf{R}^k$ with $\Delta_{\mathrm{c}}^k \subseteq E$, and a C^2-map $\phi \colon E \to V$ such that $\Phi = \phi\, \mathrm{id}_{\mathbf{R}^k}$. Therefore, given any $(k-1)$-form ω in V, we have

$$\int_{\Phi} \mathrm{d}\omega = \int_{\phi\, \mathrm{id}_{\mathbf{R}^k}} \mathrm{d}\omega = \int_{\mathrm{id}_{\mathbf{R}^k}} \phi^*(\mathrm{d}\omega) = \int_{\mathrm{id}_{\mathbf{R}^k}} \mathrm{d}(\phi^*(\omega)).$$

Moreover, we can also consider the positively oriented boundary of Φ, which yields

$$\int_{\partial\Phi} \omega = \int_{\partial(\phi\, \mathrm{id}_{\mathbf{R}^k})} \omega = \int_{\phi(\partial\, \mathrm{id}_{\mathbf{R}^k})} \omega = \int_{\partial\, \mathrm{id}_{\mathbf{R}^k}} \phi^*(\omega).$$

Hence, our goal will be to prove that the theorem is true for any $(k-1)$-form of the type $\phi^*(\omega)$, that is, any $(k-1)$-form γ in E it should be true that

$$\int_{\mathrm{id}_{\mathbf{R}^k}} \mathrm{d}\gamma = \int_{\partial\, \mathrm{id}_{\mathbf{R}^k}} \gamma. \tag{A.56}$$

Let's first deal with the base case, $k = 1$. Given any 0-form $\gamma \colon E \to \mathbf{R}$ of class C^1 in $E \subseteq \mathbf{R}$, we have — by the fundamental theorem of calculus,

$$\int_{\mathrm{id}_{\mathbf{R}}} \mathrm{d}\gamma = \int_0^1 (\mathrm{d}\gamma(x))\, \mathrm{d}x = \gamma(1) - \gamma(0) = \int_{\partial\, \mathrm{id}_{\mathbf{R}}} \gamma$$

We now assume that $k > 1$. Since any $(k-1)$-form γ in E can be written as a sum

$$\gamma = \sum_{r=1}^{k} f_r\, \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{r-1} \wedge \mathrm{d}x_{r+1} \wedge \cdots \wedge \mathrm{d}x_k := \sum_{r=1}^{k} \gamma_r,$$

we may as well simply prove Eq. (A.56) for every γ_r, with $1 \leqslant r \leqslant k$ — this is now what we'll do, for each index r. The boundary of $\mathrm{id}_{\mathbf{R}^k}$ is given by

$$\partial\, \mathrm{id}_{\mathbf{R}^k} = [e_1, \ldots, e_k] + \sum_{j=1}^{k} (-1)^j \sigma_j, \tag{A.57}$$

where we define each σ_j to be the oriented $(k-1)$-simplex $\sigma_j := [0, e_1, \dots, e_{j-1}, e_{j+1}, \dots, e_k]$ — that is, removing the j-th component of $\mathrm{id}_{\mathbf{R}^k}$, for each $1 \leqslant j \leqslant k$. Moreover, to facilitate our analysis, we'll define another oriented $(k-1)$-simplex, σ_0, defined by transposing the r-th component of $\mathrm{id}_{\mathbf{R}^k}$ an amount of $r-1$ times to the left, so that

$$\sigma_0 := [e_r, e_1, \dots, e_{r-1}, e_{r+1}, \dots, e_k] = (-1)^{r-1}[e_1, \dots, e_k].$$

Therefore, we can rewrite Eq. (A.57) as

$$\partial\, \mathrm{id}_{\mathbf{R}^k} = (-1)^{r-1}\sigma_0 + \sum_{j=1}^{k} (-1)^j \sigma_j.$$

We can analyse each of the $(k-1)$-simplexes $\sigma_j\colon \Delta_{\mathrm{c}}^{k-1} \to \mathbf{R}^k$, for $0 \leqslant j \leqslant k$:

- For the special case $j = 0$, given any $t \in \Delta_{\mathrm{c}}^{k-1}$, if $x := \sigma_0(t) \in \mathbf{R}^k$, then

$$x_i = \begin{cases} t_i, & 1 \leqslant i < r \\ 1 - \sum_{i=1}^{k-1} t_i, & i = r \\ t_{i-1}, & r < i \leqslant k \end{cases} \tag{A.58}$$

  Therefore, the matrix $[\partial_q \sigma_0^{(p)}(t)]_{i,j}$, where $1 \leqslant p, q \leqslant k-1$ and $p \neq r$, corresponds to the identity matrix on $\mathbf{R}^{k-1}$ restricted to Δ_{c}^k, therefore $\det[\partial_q \sigma_0^{(p)}] = 1$ and thus

$$\int_{\sigma_0} \gamma_r = \int_{\Delta_{\mathrm{c}}^{k-1}} f_r \sigma_r(t)\, \mathrm{d}t. \tag{A.59}$$

- For any $j \neq 0$, given any $t \in \Delta_{\mathrm{c}}^k$, if $x := \sigma_j(t) \in \mathbf{R}^k$, then

$$x_i = \begin{cases} t_i, & 1 \leqslant i < j \\ 0, & i = j \\ t_{i-1}, & j < i \leqslant k \end{cases} \tag{A.60}$$

  Moreover, since

$$\int_{\sigma_j} \gamma_r = \int_{\Delta_{\mathrm{c}}^k} f_r \sigma_j(t) \det\left[\partial_q \sigma_j^{(p)}(t)\right]_{\substack{p,q=1 \\ p \neq r}}^{k-1} \mathrm{d}t,$$

  there are two possible cases, for any $t \in \Delta_{\mathrm{c}}^k$: if $j \neq r$, then the j-th row of the Jacobian matrix is entirely composed of zeros, hence $\det[\partial_q \sigma_j^{(p)}(t)]_{p,q} = 0$, otherwise, if $j = r$, the matrix corresponds exactly to the identity matrix of $\mathbf{R}^{k-1}$ restricted to $\Delta_{\mathrm{c}}^{k-1}$ — therefore $\det[\partial_q \sigma_j^{(p)}(t)]_{p,q} = 1$ for any $t \in \Delta_{\mathrm{c}}^{k-1}$, which implies in

$$\int_{\sigma_j} \gamma_r = \begin{cases} 0, & j \neq r \\ \int_{\Delta_{\mathrm{c}}^{k-1}} f_r \sigma_r(t)\, \mathrm{d}t, & j = r \end{cases} \tag{A.61}$$

With Eq. (A.59) and Eq. (A.61) at hand, we obtain

$$
\begin{aligned}
\int_{\partial\, \mathrm{id}_{\mathbf{R}^k}} \gamma_r &= (-1)^{r-1} \int_{\sigma_0} \gamma_r + \sum_{j=1}^{r} (-1)^j \int_{\sigma_j} \gamma_r \\
&= (-1)^{r-1} \int_{\sigma_0} \gamma_r + (-1)^r \int_{\sigma_r} \gamma_r \\
&= (-1)^{r-1} \int_{\Delta_c^{k-1}} f_r \sigma_0(t) - f_r \sigma_r(t) \,\mathrm{d}t
\end{aligned}
$$

Notice, on the other hand, that

$$
\begin{aligned}
\mathrm{d}\gamma_r &= \mathrm{d}f_r \wedge \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{r-1} \wedge \mathrm{d}x_{r+1} \wedge \cdots \wedge \mathrm{d}x_k \\
&= \left(\sum_{j=1}^{k} \partial_j f_r \,\mathrm{d}x_j \right) \wedge \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{r-1} \wedge \mathrm{d}x_{r+1} \wedge \cdots \wedge \mathrm{d}x_k \\
&= \partial_r f_r \,\mathrm{d}x_r \wedge \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{r-1} \wedge \mathrm{d}x_{r+1} \wedge \cdots \wedge \mathrm{d}x_k \\
&= (-1)^{r-1} \partial_r f_r \,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k.
\end{aligned}
$$

Therefore, by Fubini's theorem (see Theorem A.8.4),

$$
\begin{aligned}
\int_{\mathrm{id}_{\mathbf{R}^k}} \mathrm{d}\gamma &= (-1)^{r-1} \int_{\Delta_c^k} \partial_r f_r(x) \,\mathrm{d}x \\
&= (-1)^{r-1} \int_{\Delta_c^{k-1}} \left[\int_0^{1-x_1+\cdots+x_{r-1}+x_{r+1}+x_{k-1}} \partial_r f_r(x) \,\mathrm{d}x_r \right] \mathrm{d}(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_k) \\
&= (-1)^{r-1} \int_{\Delta_c^{k-1}} f_r \sigma_0(t) - f_r \sigma_r(t) \,\mathrm{d}t.
\end{aligned}
$$

Where the last equality was obtained by means of Eq. (A.58) and Eq. (A.60) since the variables are independent of each other. That is, we just proved that

$$
\int_{\mathrm{id}_{\mathbf{R}^k}} \mathrm{d}\gamma_r = \int_{\partial\, \mathrm{id}_{\mathbf{R}^k}} \gamma_r,
$$

which implies in Eq. (A.56) — thus the proof is complete! ♮

Appendix B

Set Theory

B.1 ZFC

Theorem B.1.1 (Cantor)**.** Let A be a set and $f\colon A \to 2^A$ be a map, then f is not surjective and hence $|A| < |2^A|$.

Proof. Let $B := \{x \in A : x \notin f(x)\}$ be a set of 2^A and suppose, for the sake of contradiction, that f is surjective. That is, there must exist some $y \in A$ for which $f(y) = B$ — notice however that this can't be the case since by the law of excluded middle if $y \in B$ then $y \notin f(y)$ thus $f(y) \neq B$, while if $y \notin B$ then $y \in f(y)$ and thus $f(y) \neq B$. Therefore f cannot be surjective and thus $|A| < |2^A|$ since there exists an injective mapping $x \mapsto \{x\}$. ♮

Theorem B.1.2. There is no set containing all sets as members.

Proof. For the sake of contradiction, let A be a set containing every set as a member. In particular $2^A \subseteq A$ so that $|2^A| \leqslant A$ — this can't be the case by Theorem B.1.1, thus A cannot be a set. ♮

Definition B.1.3 (Cardinal numbers)**.** A *cardinal number* is an isomorphism *class* of sets, and the *cardinality* of a given set S is its isomorphism class. The following are properties pertaining to cardinal numbers:

(a) Every set has a unique cardinal number as its cardinality.

(b) Every cardinal number is the cardinality of some set.

(c) Two sets have the same cardinality if and only if they are isomorphic as sets.

Definition B.1.4 (Well ordering)**.** A *well-ordering* on a set S is a total ordering such that every non-empty subset of S has a *least* element. A set equipped with a well-order is called a *poset*.

Definition B.1.5 (Ordinal)**.** An *ordinal number* is an isomorphism *class* of *well-ordered sets*, and the *ordinal rank* of a poset S is its isomorphism class. The following are properties satisfied by the ordinal numbers in its regard to ordinal ranks:

(a) Every poset has a unique ordinal numbers as its ordinal rank.

(b) Every ordinal number is the ordinal rank of some poset.

(c) Two sets have the same ordinal rank if and only if they are isomorphic as posets.

Moreover, ordinals also have the following properties:

(d) Every ordinal α has an immediate successor $\alpha+1$ — this process entails the addition of an element to the end of a chain of a well-ordering of type α.

(e) There is a natural well-ordering on the collection of all ordinals. Given two ordinals α and β, we say that $\alpha \leqslant \beta$ if and only if there exists an initial segment of β for which α is isomorphic to.

(f) The induced well-ordering on the set $\{\beta : \beta < \alpha\}$ is the isomorphism class represented by α — therefore, one can define an ordinal as a set containing all the smaller ordinals as members.

(g) (Equivalent to the Axiom of Choice) Every set is well-orderable, hence bijective to some ordinal.

We denote by ω the *natural numbers*, so that every ordinal bijective to ω is said to be *countable*.

The cardinal numbers are well-ordered, and can be indexed by means of ordinals. We denote the α-th cardinal number by $\aleph_\alpha$ — hence $\aleph_0 = \omega$, while $\aleph_1 = \omega_1$ (the first uncountable ordinal), for instance.

Definition B.1.6 (Successor & limit ordinal)**.** Given an ordinal α, we can classify it as a *successor ordinal*, if there exists an ordinal β such that $\alpha = \beta + 1$, or as a *limit ordinal* — in particular, every cardinal is a limit ordinal.

Definition B.1.7 (Successor & limit cardinal)**.** A cardinal is said to be a *successor cardinal* if its corresponding indexing ordinal is a successor ordinal, otherwise it is a *limit cardinal*.

Bibliography

[ADH21] Araminta Amabel, Arun Debray, and Peter J. Haine. *Differential Cohomology: Categories, Characteristic Classes, and Connections*. 2021.

[Alu09] Paolo Aluffi. *Algebra: Chapter 0*. 2009.

[BBT20] Tai-Danae Bradley, Tyler Bryson, and John Terilla. *Topology: A Categorical Approach*. 2020.

[Die16] Reinhard Diestel. *Graph Theory*. 2016.

[Eng89] Ryszard Engelking. *General Topology*. 1989.

[Jos06] Jürgen Jost. *Postmodern Analysis*. 2006.

[Kim20] Dongryul Kim. *A rough guide to linear algebra*. 2020. URL: https://stanford.edu/~dkim04/assets/pdf/writings/LinAlg.pdf.

[KS06] Masaki Kashiwara and Pierre Schapira. *Categories and Sheaves*. 2006.

[Lan93] Serge Lang. *Algebra*. 1993.

[Lee11] John Lee. *Introduction to Topological Manifolds*. 2011.

[MK89] Yu I Manin and Alexei I Kostrikin. *Linear Algebra and Geometry*. 1989.

[Mun00] James Munkres. *Topology*. 2000.

[Rie16] Emily Riehl. *Category Theory in Context*. 2016.

[Rud76] Walter Rudin. *Principles of Mathematical Analysis*. 1976.

[Zor15] Vladmir A. Zorich. *Mathematical Analysis I*. 2015.

[Zor16] Vladmir A. Zorich. *Mathematical Analysis II*. 2016.