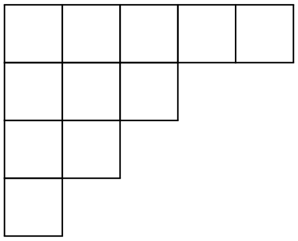
Algebraic and Geometric Methods in Engineering and Physics

José Mourão, José Natário and João P. Nunes[1]



January 9, 2020

[1] Department of Mathematics and Center for Mathematical Analysis, Geometry and Dynamical Systems, Instituto Superior Técnico, University of Lisbon.

Contents

Chapter 1

Introduction and Motivation

The goal of the present course is to introduce the student to selected topics of the mathematical methods necessary to understand aspects of some of the most striking applications of algebra, topology and geometry to modern physics and engineering.

The target audience includes physics and engineering advanced BSc or MSc students and therefore we will assume only knowledge of single and multivariable calculus, linear algebra, elementary complex analysis and differential equations.

On the other hand, since the audience includes also math students we will give detailed descriptions of the mathematical models behind each of the physics and engineering applications discussed in the course.

In particular we will study aspects of the following:

- Application of number theory to criptography;
- Methods of topology and algebraic topology to gain insight about large data;
- How the representation theory of finite groups can help us to build more stable buildings;
- Use Lie groups to study the kinematics of robots;
- The same Lie groups turn out to play a crucial role in modern cosmology, the theory that studies the evolution of the Universe in large scales from the Big Bang some 15 billion years ago all the way to the Universe today. In fact there is the following at first sight mysterious bijection of sets

  $$\{\text{Spacial homogeneous non-isotropic Cosmological models}\}$$
  $$\cong$$
  $$\{\text{Three dimensional Lie groups}\}$$

  One of our goals in the course will be to explain how come the physics problem of classifying homogeneous non-isotropic cosmological models is equivalent to the pure mathematics algebraic problem of classifying three dimensional Lie groups.
- Lie groups and their representation theory appear also in the attempts (partly successful and partly not) to unify the fundamental interactions of Nature, with the exception of gravity.

A disclaimer is in order. Given the number and variety of topics that we will address we will not manage to go very deep at them. Rather, our main goal will be to enable the interested student to go deeper by himself into the subtopics that interest him most.

Chapter 2

Topics of Algebra and Applications

The main reference for this chapter are [Ko, La, St].

2.1 Relations and Examples

2.1.1 Relations

This is a set theory topic but it will be very useful throughout the course.

Definition 2.1.1 (relations) *Let S be a set. A relation on S is a subset $\mathcal{R}$ of the Cartesian product $S \times S$,*

$$\mathcal{R} \subset S \times S = \{(s, s'),\ s, s' \in S\} \ .$$

If $(s_1, s_2) \in \mathcal{R}$ we say that s_1 is $\mathcal{R}$–related with s_2 and we write $s_1 \mathcal{R} s_2$.

With this generality the concept of relation is not very useful. We will consider two types of relations that satisfy additional properties that make them more useful.

Definition 2.1.2 (equivalence and partial order relations) *Let $\mathcal{R} \subset S \times S$ be a relation on S. Then:*

(i) $\mathcal{R}$ is said to be transitive if $s_1 \mathcal{R} s_2$ and $s_2 \mathcal{R} s_3$ implies that $s_1 \mathcal{R} s_3$ for all $s_1, s_2, s_3 \in S$.

(ii) $\mathcal{R}$ is said to be reflexive if for all $s \in S$, $(s, s) \in \mathcal{R}$, i.e. for every $s \in S$, $s \mathcal{R} s$.

(iii) $\mathcal{R}$ is called symmetric if $(s_1, s_2) \in \mathcal{R}$ implies that $(s_2, s_1) \in \mathcal{R}$, for all $s_1, s_2 \in S$.

(iv) R is called antisymmetric if for every $(s_1, s_2) \in \mathcal{R}$ such that $(s_2, s_1) \in \mathcal{R}$ then $s_1 = s_2$.

- *A relation $\mathcal{R}$ that is transitive, reflexive and symmetric (i.e. satisfies (i), (ii), (iii)) is called an equivalence relation and if $(s_1, s_2) \in \mathcal{R}$ one also writes $s_1 \sim s_2$ as equal to $s_1 \mathcal{R} s_2$.*

- *A relation $\mathcal{R}$ that is transitive, reflexive and antisymmetric (i.e. satisfies (i), (ii), (iv)) is called a partial order. If $(s_1, s_2) \in \mathcal{R}$ one also writes $s_1 \preceq s_2$ as equal to $s_1 \mathcal{R} s_2$.*

2.1.2 Equivalence Relations

Let $\mathcal{R}$ be an equivalence relation on the set S. For any $t \in S$ we define its equivalence class $[t]$ as the following subset of S:

$$[t] = \{s \in S \, : \, s \sim t\} \subset S \, .$$

Notice that equivalence classes are never empty as reflexivity implies that $[t]$ contains always, at least, the element t.

Proposition 2.1.3 *Let $\mathcal{R}$ be an equivalence relation on S and $s_1, s_2 \in S$. Then either $[s_1] = [s_2]$ or $[s_1] \cap [s_2] = \emptyset$.*

Proof. Let us show first that $[s_1] = [s_2]$ if and only if $s_1 \sim s_2$. Suppose that $[s_1] = [s_2]$. Reflexivity implies that $s_1 \in [s_1]$. Then $s_1 \in [s_2]$ and therefore $s_1 \sim s_2$. Suppose now that $s_1 \sim s_2$ and let $s \in [s_1]$. Then $s \sim s_1$ and since $s_1 \sim s_2$, transitivity implies that $s \sim s_2$ and therefore $s \in [s_2]$ and $[s_1] \subset [s_2]$. One shows analogously that $[s_2] \subset [s_1]$ so that $[s_1] = [s_2]$.

To conclude the proof we just have to show that if s_1, s_2 are not equivalent, $(s_1, s_2) \notin \mathcal{R}$, then $[s_1] \cap [s_2] = \emptyset$. Suppose that there exists an element $s \in [s_1] \cap [s_2]$. But in that case $s \in s_1$ and $s \in s_2$ and therefore $s_1 \sim s_2$, which contradicts the assumption that $(s_1, s_2) \notin \mathcal{R}$. ■

Definition 2.1.4 (partition) *A partition of a set S is a collection $(S_j)_{j \in J}$ of subsets of S such that the following two conditions are verified*

(i) $\cup_{j \in J} S_j = S$,

(ii) $S_j \cap S_k = \emptyset$ *if* $j \neq k$.

The main result in the theory of equivalence relations is then the following.

Theorem 2.1.5 (Equivalence relations) *Let $\mathcal{R}$ be an equivalence relation on the set S. Then the set of equivalence classes (called quotient set or quotient space of S with respect to the relation $\mathcal{R}$),*

$$S/\mathcal{R} \equiv S/\sim = \{[t] \, , \, t \in S\} \; ,$$

defines a partition of S. Reciprocally, a partition $(S_j)_{j \in J}$ of the set S, defines an equivalence relation $\mathcal{R}$ on S for which the equivalence classes are the sets $S_j, j \in J$, and

$$S/\sim = \{S_j \; j \in J\} \; .$$

Proof. The first part follows from Proposition 2.1.3 with $J = S/\sim$ and the fact that all elements of S are in equivalence classes, due to the reflexive property of $\mathcal{R}$. For the second part let $(S_j)_{j \in J}$ be a partition of S and define the following relation on S,

$$\mathcal{R} = \{(s,t) \, , \, s,t \in S_j , \; j \in J\} \; .$$

We see that $\mathcal{R}$ is obviously reflexive, symmetric. Suppose that $s \sim t$ and $t \sim u$. This means that exist $j, k \in J$ such that $s, t \in S_j$ and $t, u \in S_k$. But then $t \in S_j \cap S_k$ and therefore $j = k$ so that $s \sim u$ so that $\mathcal{R}$ is an equivalence relation. Finally if $t \in S_j$ then $[t] = S_j$ so that the $\mathcal{R}$–equivalence classes are the sets S_j. ■

Given an equivalence relation on S the map

$$\begin{array}{rcl} \pi : S & \longrightarrow & S/\sim \\ x & \mapsto & [x]\,, \end{array}$$

is called canonical projection or just canonical map.

Example 2.1.6 *Let S be a set and $f : S \longrightarrow M$ be a map with domain S. The map f defines on S the equivalence relation $\mathcal{R}_f$,*

$$x \sim y \Leftrightarrow f(x) = f(y)\,.$$

Exercise 2.1.1 *Consider the relation $\mathcal{R}_f$ in example 2.1.6.*

a) Show that $\mathcal{R}_f$ is an equivalence relation.

b) Show that the map

$$\begin{array}{rcl} \tilde{f} : S/\sim & \longrightarrow & M \\ [x] & \mapsto & f(x)\,, \end{array}$$

is well defined and is injective (so that the quotient $S \to S/\sim$ "cures" the lack of injectivity of f).

2.1.3 Partial Orderings

A set with a partial order relation is called a partially ordered set or, for short, a poset.

Example 2.1.7 *Examples of posets:*

- *The set of real numbers $\mathbb{R}$ with the usual partial order relation, $x \preceq y \Leftrightarrow x \leq y$. The real numbers are totally ordered in the sense that for any $x, y \in \mathbb{R}$ either $x \leq y$ or $y \leq x$.*

- *If $(S, \mathcal{R})$ is a poset then the associated lexicographic order on $S^n = S \times \cdots \times S$ is the following partial order:*
  $x = (x_1, \ldots, x_n) \preceq y = (y_1, \ldots, y_n)$ if for the first entries (from the left) that are different, $x_{j_0} \neq y_{j_0}$, one has $x_{j_0} \preceq y_{j_0}$. In particular one gets (total) lexicographic order on $\mathbb{R}^n$ and (induced from that on $\mathbb{R}^n$) on $\mathbb{Z}^n$.

- *The set, $S = 2^X$, of subsets of a set X is a poset for the partial (non total if X has more than one element, $|X| > 1$) order of set inclusion: Let $A, B \subset X$. Then*

$$A \preceq B \Leftrightarrow A \subset B\,.$$

- *The set of natural numbers $\mathbb{N}$ with partial order given by divisibility, $n \preceq m \Leftrightarrow n|m$ (i.e. n divides m), is another example of a poset with a non total order.*

- *Let P be the set of people. This set is partially ordered by descendancy.*

2.2 Groups

2.2.1 Groups and Examples

A binary operation f on a set S is a map,

$$f \,:\, S \times S \longrightarrow S\,.$$

In group theory the binary operation is called composition and, for $x, y \in S$, one frequently writes,

$$f(x,y) = x \circ y = x\,y\,,$$

and say that this element is the composition of x with y.

Definition 2.2.1 (group) *A pair $(G, \circ)$, where G is a set and $\circ \,:\, G \times G \longrightarrow G$ is a composition, is called a group if the composition satisfies the following three properties:*

(i) Is associative, i.e.

$$g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3\,, \qquad \forall g_1, g_2, g_3 \in G\,.$$

(ii) There exists a neutral element $e \in G$ such that

$$e \circ g = g \circ e = g, \qquad \forall g \in G\,.$$

(iii) For every $g \in G$ there exists $h \in G$, called inverse of g, such that,

$$g \circ h = h \circ g = e\,.$$

We write $h = g^{-1}$.

A group G is called abelian if the composition is commutative, i.e.

$$g \circ h = h \circ g\,, \qquad \forall g, h \in G\,.$$

The number of elements in a group G is called the order of the group and denoted by $|G|$.

Exercise 2.2.1 *Let G be a group. Prove that the neutral element e and the inverse g^{-1} of an element g are unique.*

Example 2.2.2 *The following are examples of groups:*

- *The sets $\mathbb{R}$ and $\mathbb{C}$ (or, as we will see, any field $\mathbb{K}$) and vector spaces V over $\mathbb{K}$ are abelian groups with composition given by addition, $\circ = +$. The neutral element is the zero, $0 \in V$, and the (additive) inverse of $v \in V$ is denoted by $-v$.*

- *The sets $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ and $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$ (or, as we will see, any field $\mathbb{K}^* := \mathbb{K} \setminus \{0\}$) are abelian groups with respect to the composition given by multiplication.*

- *The set $GL(n,\mathbb{K})$ of nonsingular square $n\times n$ matrices with entries in $\mathbb{K}$,*

$$GL(n,\mathbb{K}) = \{A \in \mathrm{Mat}_n(\mathbb{K}) \,:\, \det(A) \neq 0\,\}\ ,$$

is a group with composition given by matrix multiplication. For $n > 1$ this group is nonabelian.

Very important groups are permutation groups of sets. Let M be a nonempty set and $\mathrm{Map}(M,M)$ be the set of maps from M to M. This set has a natural composition given by composition of maps. This composition is associative and has a neutral element given by the identity map, $id_M \,:\, id_M(x) = x\,,\quad \forall x \in M$. However, for sets M with more than one element, not all maps from $\mathrm{Map}(M,M)$ have inverses. Those with inverses are the bijective maps that we will also call permutations of M. The set of of all permutations of the set M is a group and is called the symmetric group of M,

$$\mathrm{Sym}(M) = \left\{\varphi \in \mathrm{Map}(M,M) \,:\, \exists \varphi^{-1}\right\}\ .$$

Indeed, its easy to see that the composition of maps defines a composition on $\mathrm{Sym}(M)$, which satisfies all the three properties of group composition. For $M = \{1,\ldots,n\}$ we write $\mathrm{Sym}(\{1,\ldots,n\}) = S_n$. S_n is nonabelian for $n \geq 3$ and $|S_n| = n!$ (while for $\mathrm{Map}(M,M)$ we have, $|\mathrm{Map}(M,M)| = |M|^{|M|} = n^n$).

For finite groups the composition can be conveniently represented with a table. Let us illustrate this for S_3, $|S_3| = 3! = 6$.

Example 2.2.3 (S_3) *Consider the notation*

$$\begin{pmatrix} 1 & 2 & 3 \\ \varphi(1) & \varphi(2) & \varphi(3) \end{pmatrix}\ ,$$

to represent the bijective permutation mapping j to $\varphi(j), j = 1,2,3$ and

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}\ , \quad a = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\ , \quad b = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$
$$c = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}\ , \quad d = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}\ , \quad f = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

We find that the composition table for S_3 is

Table 2.1: S_3

$\cdot$	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	e	d	f	b	c
b	b	f	e	d	c	a
c	c	d	f	e	a	b
d	d	c	a	b	f	e
f	f	b	c	a	e	d

Definition 2.2.4 (subgroup) *Let G be a group. A nonempty subset $H \subset G$ is called a subgroup of G if the composition of G defines on H the structure of a group, or, equivalently, if the following conditions are satisfied:*

(i) $e \in H$.

(ii) $h^{-1} \in H$, *for every* $h \in H$.

(iii) $h_1 h_2 \in H$ *for all* $h_1, h_2 \in H$.

Examples of subgroups of the multiplicative group of nonzero complex numbers are $\mathbb{R}^*, \mathbb{Q}^*$, S^1,

$$S^1 = \{z \in \mathbb{C} \, : \, |z| = 1\} ,$$

and finite subgroups of S^1, *corresponding to the nth roots of unity,*

$$R(n) = \{z \in \mathbb{C} \, : \, z^n = 1\} \; .$$

Let ζ_n denote the following primitive nth root of unity, $\zeta_n = e^{i\frac{2\pi}{n}}$. For e.g. $R(3)$ the composition table is

Table 2.2: $R(3)$

$\cdot$	1	χ_3	χ_3^2
1	1	χ_3	χ_3^2
χ_3	χ_3	χ_3^2	1
χ_3^2	χ_3^2	1	χ_3

It is easy to see that the only subgroups of $R(3)$ are the trivial subgroups, $\{1\}$ and $R(3)$. On the other hand S_3 has one (trivial) subgroup of order one

$$\{e\},$$

three subgroups of order two

$$\{e, a\}, \{e, b\}, \{e, c\},$$

one of order three

$$\{e, d, f\},$$

none of order four or five and one (trivial) of order six, S_3. As we will see shortly it is not a coincidence that the order of subgroups of S_3 divides the order $|S_3| = 6$ of S_3. This is in fact always the case, which explains why $R(3)$ (and indeed $R(p)$ for any prime number p) has only the trivial subgroups.

Exercise 2.2.2 *Show that all subgroups of* $\mathbb{Z}$ *have the form* $d\mathbb{Z}$ *for some* $d \in \mathbb{N}_0$.

Let H be a subgroup of the group G and g be any element of G. Then the subsets gH,

$$gH = \{gh, \; h \in H\}$$

and Hg,

$$Hg = \{hg, \; h \in H\}$$

are called left and right cosets of H, respectively. The set of all left cosets of H is denoted G/H and the set of all right cosets is denoted $H \setminus G$.

Exercise 2.2.3 *Let H be a subgroup of the group G.*

a) Let $gH, g'H$ be any two left cosets of H. Show that by multiplying elements of gH on the left by $g'\,g^{-1}$ we obtain a bijective map to $g'H$. In particular, if H is finite, then all cosets (left and right) have the same number of elements.

b) Show that the relation $g \sim g'$ if and only if $g'^{-1}g \in H$ is an equivalence relation in G.

c) Show that the equivalence classes for $\sim$ are left cosets of H so that $G/H = G/\sim$.

d) Show that analogous statements are valid for right cosets.

Let $|A|$ denote the number of elements of the set A. As a consequence of the properties of cosets we obtain Lagrange's theorem.

Theorem 2.2.5 (Lagrange) *Let G be a finite group and $H \subset G$ a subgroup. Then*

$$|G| = |H|\,|G/H| = |H|\,|H \setminus G|\,. \tag{2.2.1}$$

Proof. Let us show the theorem for left cosets. For right cosets the proof is analogous.

From the Exercise 2.2.3 we know that the cosets define a partition of G on $|G/H|$ subsets and all those subsets (the cosets) are bijective and thus have the same number of elements, equal to the order $|H|$ of the subgroup H. This proves the first equality in (2.2.1). ■

Definition 2.2.6 (index of a subgroup) *The number $|G/H| = |G \setminus H|$ of cosets is called index of H in G and denoted by $[G : H]$.*

We see from (2.2.1) that if G is finite then $[G : H] = \frac{|G|}{|H|}$. But if both H and G are infinite the index of H in G can still be finite. Examples are the subgroups $d\mathbb{Z}$ of $\mathbb{Z}$, with $d > 0$.

Exercise 2.2.4 *Show, using Exercise 2.2.2, that all subgroups of $\mathbb{Z}$, except $\{0\}$, have finite index.*

It is natural to wonder when does the set G/H,

$$G/H = \{gH\,,\, g \in G\}\ ,$$

have structure of group induced by that on G. In other words, when does

$$g_1H\, g_2H = g_1g_2H\ ,\ \ g_1, g_2 \in G\,, \tag{2.2.2}$$

define a composition in G/H? First of all for (2.2.2) to define a composition on G/H the coset on the rhs should not depend on the choice of representatives g_1, g_2 of the cosets g_1H, g_2H on the left. Let g_1', g_2' be two other representatives of the same cosets. In that case there exist $h_1, h_2 \in H$ such that, $g_1' = g_1h_1$ and $g_2' = g_2h_2$ and we want to have

$$g_1'g_2'H = g_1g_2H\,, \tag{2.2.3}$$

for every $g_1, g_2 \in G,\ h_1, h_2 \in H$. The equality (2.2.3) is equivalent to

$$\begin{aligned}(g_1'g_2')^{-1}\, g_1g_2 \in H \ &\Leftrightarrow\ (g_2')^{-1}(g_1')^{-1}\, g_1g_2 \in H \\ &\Leftrightarrow\ h_2^{-1}g_2^{-1}h_1^{-1}g_1^{-1}\, g_1g_2 = h_2^{-1}g_2^{-1}h_1^{-1}g_2 \in H\,.\end{aligned}$$

We see that for this to hold we must have that, for every $g_2 \in G$ and $h_1 \in H$, there must be $h_3 \in H$ such that,

$$h_1^{-1}g_2 = g_2h_3\,,$$

or, equivalently,

$$Hg = gH \Leftrightarrow gHg^{-1} = H\,, \tag{2.2.4}$$

for every $g \in G$.

Definition 2.2.7 (normal subgroup) *A subgroup H of the group G is called a normal subgroup of G if (2.2.4) is verified for all $g \in G$, i.e. if all left cosets of H are equal to the right cosets.*

Corollary 2.2.8 *Let H be a normal subgroup of the group G. Then (2.2.2) for all $g_1, g_2 \in G$ defines a composition on the set of (left=right) cosets G/H, with respect to which G/H becomes a group, called the quotient group of G by H.*

Proof.

■

Remark 2.2.9 Obviously, if G is abelian, then (2.2.4) holds for all subgroups H and all $g \in G$ so that all subgroups of an abelian group are normal subgroups. ◊

Example 2.2.10 *As we saw in the Exercise 2.2.4 the possible subgroups of $\mathbb{Z}$ are $n\mathbb{Z}$, for some natural number or zero, $n \in \mathbb{N}_0$. The corresponding quotient groups, $\mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z}$, have order n and are called groups of $\mathbb{Z}$ mod n,*

$$\mathbb{Z}_n = \{[0] = n\mathbb{Z}, [1] = 1 + n\mathbb{Z}, \ldots, [n-1] = n - 1 + n\mathbb{Z}\}\ .$$

Exercise 2.2.5 *In the notation of Example 2.2.3 consider the subgroups $H = \{e, a\}$ and $A_3 = \{e, d, f\}$. A_3 is called alternating group of degree 3.*

a) Show that H is not a normal subgroup.

b) Show that A_3 is a normal subgroup and find the composition table for S_3/A_3.

Definition 2.2.11 (homomorphism, isomorphism) *Let G, H be groups. A map $\varphi : G \longrightarrow H$ such that,*

$$\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2) \qquad \forall g_1, g_2 \in G\,,$$

is called a (group) homomorphism. If φ is bijective than φ is called an isomorphism. Two groups are said to be isomorphic if there is an isomorphism from one to the other.

Exercise 2.2.6 *Let $\varphi : G \longrightarrow H$ be a group homomorphism. Show the following:*

a) $\varphi(e_G) = e_H$.

b) $\varphi(g^{-1}) = \varphi(g)^{-1}$, $\quad \forall g \in G$.

c) Let $\psi : H \longrightarrow K$ *be another (group) homomorphism. Then* $\psi \circ \varphi$ *is a group homomorphism from* G *to* K.

Solution. TO BE ADDED ■

Being isomorphic defines an equivalence relation in the set of all groups. From the point of view of algebra one is mostly interested in groups up to isomorphism.

Example 2.2.12 *The groups* $(\mathbb{Z}_n, +)$ *and* $R(n)$ *are isomorphic with isomorphism given by,*

$$\begin{array}{rcl}\tilde{\varphi} : \mathbb{Z}_n & \longrightarrow & R(n) \\ {[k]} & \mapsto & (\chi_1)^k = e^{2\pi i \frac{k}{n}} .\end{array}$$

We will return to this example below in Exercise 2.2.7.

Example 2.2.13 *Let* $a \in \mathbb{R}^* := \mathbb{R} \setminus \{0\}$*. A one parameter family of isomorphisms,* φ_a*, from the additive group of real numbers to the multiplicative group of positive real numbers is given by*

$$\begin{array}{rcl}\varphi_a : \mathbb{R} & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & e^{ax} .\end{array}$$

Example 2.2.14 *Let* G *be a subgroup of the group of real nonsingular square matrices,*

$$G \subset GL_n(\mathbb{R}) .$$

The map det *from* G *to* $\mathbb{R}^*$,

$$A \longmapsto \det(A) ,$$

is a group homomorphism.

Example 2.2.15 *Let* N *be a normal subgroup of the group* G*. From (2.2.2) and Corollary 2.2.8 it follows that the canonical projection* π *from* G *to* G/N *is a (surjective) homomorphism,*

$$\pi(g) = gN .$$

Definition 2.2.16 *The kernel of a group homomorphism,* $\varphi : G \longrightarrow H$*, is the pre-image of the identity,*

$$\ker(\varphi) = \varphi^{-1}(\{e\}) = \{g \in G : \varphi(g) = e \in H\} ,$$

A very important theorem is the isomorphism theorem.

Theorem 2.2.17 (isomorphism theorem) *Let* $\varphi : G \longrightarrow H$ *be a group homomorphism. Then,*

(a) $Im(\varphi)$ *is a subgroup of* H.

(b) $\ker(\varphi)$ *is a normal subgroup of* G.

(c) The map $\tilde{\varphi} : G/\ker(\varphi) \longrightarrow Im(\varphi)$,

$$\tilde{\varphi}(g\ker(\varphi)) = \varphi(g)\,,$$

induced by φ*, is an isomorphism.*

Proof.

(a) Let us prove that $Im(\varphi)$ is a subgroup of H.

(i) We saw in Exercise 2.2.6 that $e_H = \varphi(e_G) \in Im(\varphi)$.

(ii)

$$\begin{aligned} h \in Im(\varphi) &\Leftrightarrow \exists\, g \in G : h = \varphi(g) \\ &\Leftrightarrow \exists\, g \in G : h^{-1} = \varphi(g)^{-1} = \varphi(g^{-1}) \\ &\Leftrightarrow h^{-1} \in Im(\varphi). \end{aligned}$$

(iii)

$$\begin{aligned} \text{Let } h_1, h_2 \in Im(\varphi) &\Leftrightarrow \exists g_1, g_2 \in G : h_1 = \varphi(g_1)\,, h_2 = \varphi(g_2) \\ &\Rightarrow h_1h_2 = \varphi(g_1)\varphi(g_2) = \varphi(g_1g_2) \in Im(\varphi). \end{aligned}$$

(b) We now prove that $\ker(\varphi)$ is a normal subgroup of G.

(i) We have $\varphi(e_G) = e_H \Rightarrow e_G \in \ker(\varphi)$.

(ii) Let $g \in \ker(\varphi)$,

$$\begin{aligned} &\Leftrightarrow \varphi(g) = e_H \\ &\Leftrightarrow \varphi(g)^{-1} = \varphi(g^{-1}) = e_H \\ &\Leftrightarrow g^{-1} \in \ker(\varphi)\,. \end{aligned}$$

(iii) Let $g_1, g_2 \in \ker(\varphi)$,

$$\begin{aligned} &\Leftrightarrow \varphi(g_1) = \varphi(g_2) = e_H. \\ &\Rightarrow \varphi(g_1g_2) = \varphi(g_1)\varphi(g_2) = e_H \Leftrightarrow g_1g_2 \in \ker(\varphi). \end{aligned}$$

To show that $\ker(\varphi)$ is a normal subgroup we have to show that

$$g\ker(\varphi)g^{-1} = \ker(\varphi)\,, \forall g \in G. \tag{2.2.5}$$

Let us show first that $g\ker(\varphi)g^{-1} \subset \ker(\varphi), \forall g \in G$. Let $g' \in g\ker(\varphi)g^{-1} \Leftrightarrow \exists n \in \ker(\varphi) : g' = gng^{-1}$. Then,

$$\varphi(g') = \varphi(gng^{-1}) = \varphi(g)\varphi(n)\varphi(g^{-1}) = \varphi(g)e_H\varphi(g)^{-1} = e_H,$$

and therefore $g' \in \ker(\varphi)\,, \forall g \in G$.
Let us now show that $\ker(\varphi) \subset g\ker(\varphi)g^{-1}, \forall g \in G$. Let $n \in \ker(\varphi)$. Then,

$$n = (gg^{-1})\,n\,(gg^{-1}) = g\left(g^{-1}ng\right)g^{-1}.$$

Since we have shown that $g^{-1}ng \in \ker(\varphi)$ we conclude that $n \in g\ker(\varphi)g^{-1}\,, \forall g \in G$ and therefore we have proved (2.2.5).

(c) To show that φ induces a bijective map

$$\begin{array}{rcl}\tilde{\varphi}\,:\,G/\ker(\varphi) & \longrightarrow & Im(\varphi)\\ \tilde{\varphi}(g\ker(\varphi)) & = & \varphi(g)\,,\end{array} \tag{2.2.6}$$

and following exercise 2.1.1 we have to show that the equivalence relation $\mathcal{R}_\varphi$ defined on G by φ coincides with the equivalence relation associated with $\ker(\varphi)$. Indeed,

$$\varphi(g_1)=\varphi(g_2)\Leftrightarrow\varphi(g_2)^{-1}\varphi(g_1)=\varphi(g_2^{-1}g_1)=e_H\Leftrightarrow g_2^{-1}g_1\in\ker(\varphi)\ \ \forall\, g_1,g_2\in G\,.$$

So the equivalence classes of these equivalence relations coincide and therefore the $\ker(\varphi)$ cosets coincide with the points in G with the same image under φ. Then, from exercise 2.1.1, we know that $\tilde{\varphi}$ in (2.2.6) is well defined and bijective. We now show that $\tilde{\varphi}$ is an homomorphism.

$$\tilde{\varphi}([g_1][g_2])=\tilde{\varphi}([g_1g_2])=\varphi(g_1g_2)=\varphi(g_1)\varphi(g_2)=\tilde{\varphi}([g_1])\tilde{\varphi}([g_2])\ ,\ \forall[g_1]\,,\,[g_2]\in G/\ker(\varphi)\,.$$

■

Obviously if N is a normal subgroup of G then it is easy to construct a homomorphism, $\varphi\,:\,G\longrightarrow H$, with kernel, $\ker(\varphi)=N$: just choose $H=G/N$ and $\varphi=\pi$, where π denotes the canonical projection.

Remark 2.2.18 From *(c)* in the isomorphism theorem we obtain a convenient way of understanding quotient groups, G/N. It is sufficient to find a group H and a surjective homomorphism $\varphi\,:\,G\longrightarrow H$ with kernel, $\ker(\varphi)=N$. Then we know from Theorem 2.2.17 that G/N is isomorphic to H with isomorphism given by $\tilde{\varphi}$.

♢

Exercise 2.2.7 *Let $n\in\mathbb{N}$. Show that the additive group $\mathbb{Z}/n\mathbb{Z}$ is isomorphic to the multiplicative group $R(n)$ of the nth roots of unity.*

Solution. We will solve this exercise by using (c) in the Isomorphism Theorem 2.2.17. So we need to find a surjective homomorphism, $\varphi\,:\,\mathbb{Z}\longrightarrow R(n)$ such that $\ker(\varphi)=n\mathbb{Z}$. Let us show that indeed the map φ corresponding to $\tilde{\varphi}$ in Example 2.2.12 above satisfies these conditions,

$$\begin{array}{rcl}\varphi\,:\,\mathbb{Z} & \longrightarrow & R(n)\\ m & \longmapsto & e^{2\pi i\frac{m}{n}}\,.\end{array}$$

We see that φ is surjective. Let us show that it is a group homomorphism.

$$\varphi(m_1+m_2)=e^{2\pi i\frac{m_1+m_2}{n}}=e^{2\pi i\frac{m_1}{n}}\,e^{2\pi i\frac{m_2}{n}}=\varphi(m_1)\,\varphi(m_2)\,,\qquad\forall m_1,m_2\in\mathbb{Z}.$$

For the kernel we find,

$$\ker\varphi=\left\{m\in\mathbb{Z}\,:\,e^{2\pi i\frac{m}{n}}=1\right\}\,=\,n\mathbb{Z}\,.$$

We see that indeed we are in the conditions of (c) in Theorem 2.2.17 and therefore the map $\tilde{\varphi}$ in the Example 2.2.12 is a group isomorphism.

■

To any element $g \in G$ we can associate an homomorphism from (the additive group of integers) $\mathbb{Z}$ to G,

$$\begin{array}{rcl} \varphi_g : \mathbb{Z} & \longrightarrow & G \\ n & \longmapsto & g^n = g \cdots g \,, \end{array}$$

where $g^0 := e$. The subgroup $\varphi_g(\mathbb{Z}) = Im(\varphi_g) \subset G$ is called subgroup generated by g and denoted by $\langle g \rangle$. It is easy to see that $Im(\varphi_g)$ is the minimal subgroup of G that contains the element g. From Exercise 2.2.2 we know that $\ker(\varphi_g) = \mathbb{Z}_d$ for $d \in \mathbb{N}$, where d is the minimal natural number for which $g^d = e$. This number is called order of the element g, $\mathrm{ord}(g)$, and we see from the Isomorphism Theorem 2.2.17 that $Im(\varphi_g)$ is isomorphic to $\mathbb{Z}_d = \mathbb{Z}/d\mathbb{Z} \cong R(d)$.

Remark 2.2.19 From Lagrange Theorem 2.2.5, if $|G| < \infty$ then

$$d = \mathrm{ord}(g) = \mathrm{ord}(< g >) \,|\, |G| \,,$$

i.e. the order of every element $g \in G$ divides the order of the group G.

If G has prime order, $|G| = p$, then all its elements different from e, $g \neq e$, must have order p. Indeed, $\mathrm{ord}(g) \geq 2$ and $\mathrm{ord}(g)|p = |G|$ so that $\mathrm{ord}(g) = p$ and therefore G is cyclic with cyclic generator any element g different from e,

$$G = \langle g \rangle \,.$$

◊

Exercise 2.2.8 *Construct a surjective homomorphism φ from S_3 to $\mathbb{Z}_2$. Which quotient group is described by this homomorphism?*

Solution. Let $\varphi : S_3 \longrightarrow \mathbb{Z}_2$, as follows,

$$\varphi(g) = \left\{ \begin{array}{ll} [0] & \text{if}\, g \in A_3 \\ [1] & \text{if}\, g \notin A_3 \end{array} \right. .$$

φ has to be surjective, constant on the two A_3 cosets and $\ker(\varphi) = A_3$. These conditions fix φ uniquely. The homomorphism property can be easily verified with the composition tables of S_3 and $\mathbb{Z}_2$ (the details are left as an exercise to the reader).

■

Definition 2.2.20 *A group G is called cyclic if it is generated by one of its elements, i.e. if there exists $g \in G$ such that $G = \langle g \rangle$.*

2.2.2 Applications to Number Theory

We will be assuming throughout the present subsection some results of elementary number theory like the fundamental theorem of arithmetic (or prime factorization theorem):

Theorem 2.2.21 (prime factorization theorem) *(see [La, Theorem 1.8.5]) Let* $n \in \mathbb{N}, n > 1$. *Then* n *can be factored uniquely into a product of prime numbers*

$$n = p_1 \cdots p_r \, ,$$

with $r \geq 1$ *and* $p_1 \leq p_2 \leq \cdots \leq p_r$.

Recall that if $a = bc$, $a, b, c \in \mathbb{Z}$, we say that b (and c) divides a and we write $b|c$.

Let us now recall the Euclidean algorithm to find the greatest common divisor of two integers m, n, $\gcd(m,n)$. We will need the well known result:

Theorem 2.2.22 (division with remainder) *(see [La, Theorem 1.2.1]) Let* $d \in \mathbb{N}$. *Then for every* $n \in \mathbb{Z}$ *there exist unique* $q \in \mathbb{Z}$ *and* $r \in \{0, 1, \ldots, d-1\}$ *such that*

$$n = qd + r \, .$$

Crucial to the Euclidean algorithm is the following result.

Proposition 2.2.23 *(see [La, Theorem 1.5.1]) Let* $m, n \in \mathbb{Z}$. *Then,*

(i) $\gcd(m, 0) = m$, *for* $m \in \mathbb{N}$.

(ii) $\gcd(m,n) = \gcd(m - qn, n) \ , \forall q \in \mathbb{Z}$.

By using this proposition and Theorem 2.2.22 we can find $\gcd(m,n)$ for $m \geq n$ as follows. Applying division with remainder,

$$m = qn + r$$

we obtain, $\gcd(m,n) = \gcd(n,r)$. Applying again division with remainder to n, r we obtain, $n = q_1 r + r_1$ and therefore, $\gcd(n,r) = \gcd(r, r_1)$, with $n > r > r_1 \geq 0$. This process, called Euclidean algorithm, ends after a finite number of steps when we get to $\gcd(m,n) = \gcd(\ell, 0) = \ell$ (see [La, p. 10]). A very important corollary of this algorithm is that it can be used to show that for $m, n \in \mathbb{Z}$ there exist $\lambda, \mu \in \mathbb{Z}$ such that the $\gcd(m,n)$ can be represented as a linear combination with integer coefficients of m, n,

$$\gcd(m,n) = \lambda m + \mu n \, . \tag{2.2.7}$$

This way of finding λ, μ from the Euclidean algorithm is called *extended Euclidean algorithm.* Let us illustrate this in the following exercise.

Exercise 2.2.9 (extended Euclidean algorithm) *Let* $m = 35$ *and* $n = 18$.

a) Find $\gcd(35, 18)$, *using the Euclidean algorithm.*

b) Find λ *and* μ *as in (2.2.7), using the extended Euclidean algorithm.*

Solution.

a) We have,

$$\begin{aligned} 35 &= 18 \times 1 + 17 \\ 18 &= 17 \times 1 + 1 \end{aligned}$$

and therefore $\gcd(35, 18) = \gcd(18, 17) = \gcd(17, 1) = \gcd(1, 0) = 1$.

b) From the steps of the Euclidean algorithm we obtain

$$\begin{aligned} 17 &= 35 - 18 \times 1 \\ 1 &= -17 + 18 = \\ &= -(35 - 18) + 18 = \\ &= -35 + 18 \times 2\,, \end{aligned}$$

so that, in this example, $\lambda = -1$ and $\mu = 2$.

■

In order to bring in group theory in its full force we will need the second, besides addition, composition on the integers and on the the Abelian (additive) groups $\mathbb{Z}_n$: the multiplication. The triple $(\mathbb{Z}, +, \cdot)$ defines on $\mathbb{Z}$ the structure of a commutative ring.

Definition 2.2.24 (ring and field) *A ring is a triple $(R, +, \cdot)$, where $(R, +)$ is an Abelian group and the second composition, called multiplication of the ring, satisfies the following properties:*

(i) [associativity] $a \cdot (b \cdot c) = (a \cdot b) \cdot c\,, \quad \forall a, b, c \in R.$

(ii) [neutral element] The multiplication, $\cdot$, has a neutral element, $1 \in R$.

(iii) [distributivity] $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c\ , \quad \forall a, b, c \in R.$

If the multiplication is commutative (like in the case of $\mathbb{Z}$) the ring is called commutative. If the multiplication is commutative and if all elements of $R \setminus \{0\}$ have multiplicative inverses, R is called a field.

An element $a \in R$ with multiplicative inverse, i.e. $\exists\, b \in R$ such that $a \cdot b = b \cdot a = 1$ is called a unit of R. We write $b = a^{-1}$. It is easy to see that the set of all units in a ring, denoted by R^*, is a group, called group of units of R. For fields $R^* = R \setminus \{0\}$. An element $a \in R$ is called a zero divisor if there is a nonzero element, $b \neq 0$, of the ring such that $a \cdot b = 0$ or $b \cdot a = 0$. A zero divisor can not be a unit. Indeed, suppose that a is a zero divisor and a unit. Let $b \neq 0, b \in R$ be such that

$$b \cdot a = 0 \quad \text{(the case with } a \cdot b = 0 \text{ is analogous)}\,.$$

By multiplying both terms by a^{-1} on the right we obtain,

$$(b \cdot a) \cdot a^{-1} = 0 \cdot a^{-1} \Leftrightarrow b \cdot (a \cdot a^{-1}) = 0 \Leftrightarrow b = 0, \tag{2.2.8}$$

which contradicts the hypothesis that $b \neq 0$.

Example 2.2.25 (rings) *Some examples of rings/fields are the following.*

1. *Examples of fields are $\mathbb{C}, \mathbb{R}, \mathbb{Q}$ and the finite fields $\mathbb{Z}_p$, for p prime.*

2. *$\mathbb{Z}$ is a commutative ring but is not a field as its group of units is $\mathbb{Z}^* = \{1, -1\} \neq \mathbb{Z} \setminus \{0\}$.*

3. *The set of continuous (real valued) functions on $\mathbb{R}$, $C(\mathbb{R})$ (or on any manifold), is a commutative ring. Its group of units, $C(\mathbb{R})^*$, is the set (i.e. group) of functions that are different from zero everywhere. So $f_1(x) = e^x$ and $f_2(x) = x^2 + 1$ are units in $C(\mathbb{R})$, but $f_3(x) = x$ and $f_4(x) = \sin(x)$, are not.*

4. *The set of all real square matrices $n \times n$, $Mat_n(\mathbb{R})$, is a ring that is noncommutative for $n > 1$. Its group of units coincides with the general linear group of all real invertible matrices $n \times n$,*
$$Mat_n(\mathbb{R})^* = GL_n(\mathbb{R})\,.$$

It is easy to show that the multiplication on $\mathbb{Z}$ defines a multiplication on $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ i.e.

$$[m] \cdot [k] = [mk]$$

does not depend on the class representatives.

The group $\mathbb{Z}_n^*$ is much less trivial than the group $\mathbb{Z}_n$: in particular it is not easy to calculate its order and it is not always cyclic (unlike $\mathbb{Z}_n$).

Define the Euler φ function, $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$, as follows:

$$\varphi(n) = \text{number of natural numbers smaller than } n \text{ that are coprime to n,}$$

We have:

Theorem 2.2.26 *Let $n \in \mathbb{N}$. Then,*

$$\mathbb{Z}_n^* = \{[m] \in \mathbb{Z}_n \,:\, gcd(m,n) = 1\}\ ,$$

so that, in particular, $|\mathbb{Z}_n^*| = \varphi(n)$.

Proof. Let $a \in \mathbb{Z}$, $0 \le a < n$. Let us first show that if a is not coprime with n then $[a]$ is not a unit, $[a] \notin \mathbb{Z}_n^*$. Suppose $\gcd(a,n) = b > 1$. Then,

$$\frac{n}{b}\,a = n\,\frac{a}{b} \equiv 0 \pmod{n}\,,$$

so that $[a]$ is a zero divisor and thus can not be a unit (see (2.2.8)).

Let us now show that if a is relatively prime to n then $[a]$ is a unit in $\mathbb{Z}_n$. From the extended Euclidean algorithm we know that $\exists\lambda, \mu \in \mathbb{Z}$ such that

$$\begin{aligned} \lambda\, a + \mu\, n &= 1 \Leftrightarrow \\ \lambda\, a &\equiv 1 \pmod{n} \Leftrightarrow \\ [\lambda]\,[a] &= [1] \Leftrightarrow [\lambda] = [a]^{-1}\,. \end{aligned}$$

■

We conclude that if p is prime, then $\mathbb{Z}_p$ is a field and $\varphi(p) = p - 1$. In general (see [La, p. 24]) one shows that if $\gcd(n,m) = 1$ then $\varphi(nm) = \varphi(n)\,\varphi(m)$ and, given the prime factorization of n, $n = p_1^{r_1} \cdots p_s^{r_s}$, one finds,

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_s}\right)\,.$$

Exercise 2.2.10 (see Example 1.5.3 of [La]) *Notice that* $\gcd(34 = 17 \times 2, 13) = 1$.

(a) *Find* x, y *such that* $13\,x + 34\,y = 1$.

(b) *Find* $[13]^{-1}$ *in* $\mathbb{Z}_{34}^*$.

Solution.

(a) Let us apply the extended Euclidean algorithm to find x and y.

$$\begin{aligned} 34 &= 2 \cdot 13 + 8 \\ 13 &= 8 \cdot 1 + 5 \\ 8 &= 5 \cdot 1 + 3 \\ 5 &= 3 \cdot 1 + 2 \\ 3 &= 2 \cdot 1 + 1 \end{aligned}$$

Then,

$$\begin{aligned} 8 &= -2 \cdot 13 + 34 \\ 5 &= 13 - 8 = 13 - (-2 \cdot 13 + 34) = 3 \cdot 13 - 34 \\ 3 &= 8 - 5 = (-2 \cdot 13 + 34) - (3 \cdot 13 - 34) = -5 \cdot 13 + 2 \cdot 34 \\ 2 &= 5 - 3 = (3 \cdot 13 - 34) - (-5 \cdot 13 + 2 \cdot 34) = 8 \cdot 13 - 3 \cdot 34 \\ 1 &= 3 - 2 = (-5 \cdot 13 + 2 \cdot 34) - (8 \cdot 13 - 3 \cdot 34) = -13 \cdot 13 + 5 \cdot 34\,, \end{aligned}$$

so that $x = -13$ and $y = 5$.

(b) We see that

$$\begin{aligned} [-13] \cdot [\ 13] &= [1] \ \Leftrightarrow \\ [21] \cdot [13] &= [1] \ \Leftrightarrow \\ [13]^{-1} &= [21]\,. \end{aligned}$$

■

Example 2.2.27 *The multiplication table for* $\mathbb{Z}_5^*$

Table 2.3: $\mathbb{Z}_5^*$

×	[1]	[2]	[3]	[4]
[1]	[1]	[2]	[3]	[4]
[2]	[2]	[4]	[1]	[3]
[3]	[3]	[1]	[4]	[2]
[4]	[4]	[3]	[2]	[1]

Exercise 2.2.11 *Prove that* $\mathbb{Z}_{15}^*$ *is not cyclic by using the map* $f\,:\,[m] \mapsto [m]^2$.

Solution. Since $15 = 3 \times 5$ the order of $\mathbb{Z}_{15}^*$ is $\varphi(15) = (3-1)(5-1) = 8$. So $\mathbb{Z}_{15}^*$ is cyclic if and only if it has elements of order 8.

Explicitly,

$$\mathbb{Z}_{15}^* = \{[1], [2], [4], [7], [8], [11], [13], [14]\} \ .$$

Let us now write the table for f (notice that e.g. $[13]^2 = [-2]^2 = [4]$).

Table 2.4: $[m] \mapsto [m]^2$

$[m]$	$[1]$	$[2]$	$[4]$	$[7]$	$[8]$	$[11]$	$[13]$	$[14]$
$[m]^2$	$[1]$	$[4]$	$[1]$	$[4]$	$[4]$	$[1]$	$[4]$	$[1]$

Since the order of the units in $\mathbb{Z}_{15}$ divides 8, from Table 2.4 we see that $[4], [11], [14]$ have order 2 and $[2], [7], [8], [13]$ have order 4. No element has order 8. ■

Let us now study two typical applications of group theory to number theory by proving the Euler theorem and the Chinese remainder theorem.

Theorem 2.2.28 *Let a and n be relatively prime natural numbers. Then*

$$a^{\varphi(n)} \equiv 1 \pmod{n} . \tag{2.2.9}$$

Proof. Since $\gcd(a, n) = 1$, $[a]$ is a unit, $[a] \in \mathbb{Z}_n^*$, and therefore the order of $[a]$ in $\mathbb{Z}_n^*$ divides the order of the group (see remark 2.2.19), i.e. divides $\varphi(n)$, $ord([a]) \,|\, \varphi(n)$. This implies that

$$[a]^{\varphi(n)} = [1] \qquad (\text{in } \mathbb{Z}_n),$$

which is equivalent to (2.2.9). ■

Let us now prove the Chinese remainder theorem with the help of the isomorphism theorem.

Theorem 2.2.29 *Let $n_1, \ldots, n_r \in \mathbb{N}$ be relatively prime, $N = n_1 \ldots n_r$, and let*

$$\psi_i \ : \ \mathbb{Z} \longrightarrow \mathbb{Z}/n_i\mathbb{Z} \, ,$$

denote the canonical surjective homomorphisms. The following map

$$\begin{array}{rcl} \tilde{\Psi} \ : \ \mathbb{Z}/N\mathbb{Z} & \longrightarrow & \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z} \\ \tilde{\Psi}(m + N\mathbb{Z}) & = & (\psi_1(m), \ldots, \psi_r(m)) \\ & = & (m + n_1\mathbb{Z}, \ldots, m + n_r\mathbb{Z}) \ , \end{array} \tag{2.2.10}$$

is well defined and an isomorphism.

Proof. Consider the homomorphism

$$\begin{array}{rcl} \Psi \ : \ \mathbb{Z} & \longrightarrow & \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_r} \\ \Psi(m) & = & (\psi_1(m), \ldots, \psi_r(m)) \ . \end{array}$$

Let us first show that $\ker(\Psi) = N\mathbb{Z}$. The inclusion $N\mathbb{Z} \subset \ker(\Psi)$ is easy. Indeed,

$$\begin{array}{rcl} \Psi(Nk) & = & (Nk + n_1\mathbb{Z}, \ldots Nk + n_r\mathbb{Z}) \\ & = & (n_1\mathbb{Z}, \ldots, n_r\mathbb{Z}) \\ & = & ([0], \ldots, [0]) \ . \end{array}$$

Let us now show that,

$$\ker(\Psi) \subset N\mathbb{Z}\,.$$

Let $m \in \ker(\Psi)$. Then we have,

$$\begin{aligned}\Psi(m) &= (m + n_1\mathbb{Z}, \ldots, m + n_r\mathbb{Z}) \\ &= (n_1\mathbb{Z}, \ldots, n_r\mathbb{Z})\ ,\end{aligned}$$

which implies that $n_1|m, \ldots, n_r|m$. But, since $n_1, \ldots, n_r$ are relatively prime, this in turn implies that $N = n_1 \cdots n_r$ divides m and therefore $m \in N\mathbb{Z}$, which concludes the proof that $N\mathbb{Z} = \ker(\Psi)$. Therefore, from the isomorphism theorem, Ψ induces an injective homomorphism $\tilde{\Psi}$ defined by (2.2.10). The surjectivity of $\tilde{\Psi}$ follows from its injectivity and the fact that

$$|\mathbb{Z}/N\mathbb{Z}| = N = n_1 \cdots n_r = |\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_r}|\,.$$

■

Exercise 2.2.12 *Construct the isomorphism* $\tilde{\Psi} : \mathbb{Z}_6 \longrightarrow \mathbb{Z}_3 \times \mathbb{Z}_2$.

Solution.

$$\begin{aligned}\tilde{\Psi}(6\mathbb{Z}) &= ([0], [0]) \\ \tilde{\Psi}(1 + 6\mathbb{Z}) &= (1 + 3\mathbb{Z}, 1 + 2\mathbb{Z}) \\ &= ([1], [1]) \\ \tilde{\Psi}(2 + 6\mathbb{Z}) &= (2 + 3\mathbb{Z}, 2 + 2\mathbb{Z}) = (2 + 3\mathbb{Z}, 2\mathbb{Z}) \\ &= ([2], [0]) \\ \tilde{\Psi}(3 + 6\mathbb{Z}) &= (3 + 3\mathbb{Z}, 3 + 2\mathbb{Z}) = (3\mathbb{Z}, 1 + 2\mathbb{Z}) \\ &= ([0], [1]) \\ \tilde{\Psi}(4 + 6\mathbb{Z}) &= (4 + 3\mathbb{Z}, 4 + 2\mathbb{Z}) = (1 + 3\mathbb{Z}, 2\mathbb{Z}) \\ &= ([1], [0]) \\ \tilde{\Psi}(5 + 6\mathbb{Z}) &= (5 + 3\mathbb{Z}, 5 + 2\mathbb{Z}) = (2 + 3\mathbb{Z}, 1 + 2\mathbb{Z}) \\ &= ([2], [1])\ .\end{aligned}$$

■

The Chinese remainder theorem can be viewed as a partial structure theorem for finite abelian groups. More generally we have the following theorem.

Theorem 2.2.30 *Let G be a finite abelian group of order n with prime factorization given by $n = p_1^{r_1} \cdots p_s^{r_s}$. Then,*

$$G \cong G_1 \times \cdots \times G_s\,,$$

where G_j have order, $|G_j| = p_j^{r_j}$. For any such group G_j there is a representation of r_j as a sum of nonincreasing natural numbers, a_j^i, $r_j = a_j^1 + \cdots + a_j^\ell$, such that

$$G_j \cong \mathbb{Z}_{p_j^{a_j^1}} \times \cdots \times \mathbb{Z}_{p_j^{a_j^\ell}}\,.$$

We conclude that every finite Abelian group is isomorphic to the direct product of cyclic groups of power of a prime order. For example of order 8 there are three nonisomorphic (classes of) abelian groups (there are also two nonisomorphic nonabelian groups):

$$\mathbb{Z}_8, \mathbb{Z}_4 \times \mathbb{Z}_2, \mathbb{Z}_2^3\,,$$

while of order 12 there are only two nonisomorphic classes of abelian groups:

$$\mathbb{Z}_3 \times \mathbb{Z}_4 \cong \mathbb{Z}_{12},\ \mathbb{Z}_3 \times \mathbb{Z}_2^2\,.$$

2.2.3 Application to Criptography

Public key criptography was invented in the 70's by Rivert, Shamir and Adelman (RSA), who proposed the first RSA cryptosystem based on the so called "one way" functions. These functions are provided naturally by number theory and are based on the difficulty (at least for the known algorithms) of factoring large natural numbers into primes. The one way functions proposed initially by RSA are obtained by fixing two very large distinct primes, p, q, and considering on $\mathbb{Z}_N$, $N = pq$, the function

$$F([X]) = [X]^e\,,$$

where e is called the encryption exponent, with e chosen such that,

$$\gcd(e, \varphi(N)) = \gcd(e, (p-1)(q-1)) = 1\,.$$

Then the decryption exponent, d, is chosen so that

$$([X]^e)^d = [X]^{ed} = [X] \text{ in } \mathbb{Z}_N.$$

The issue is that to find d, knowing only N and e, is equivalent to factoring N and thus a very difficult task for large p and q.

The idea of RSA is that Alice has a form to have Bob sending her encrypted messages through unsecure channels. She displays in a public page the two numbers N and e that Bob will use to encrypt his messages, $[X] \in \mathbb{Z}_N$, by taking them to the encryption power: $[X] \mapsto [X]^e$. Then Bob sends $[X]^e$ to Alice and only she knows the private key i.e. the decryption power d, which allows her to get back $[X]$. The key result, justifying the above, is the following.

Theorem 2.2.31 *Let $N = pq$, with p, q distinct prime numbers and $e \in \mathbb{N}$ be such that* $\gcd(e, \varphi(N)) = 1$. *Then, there exists $d \in \mathbb{N}$ such that*

$$[X]^{ed} = ([X]^e)^d = [X]\,, \qquad \forall [X] \in \mathbb{Z}_N\,. \tag{2.2.11}$$

Proof. Since, due to the Chinese remainder theorem, $\mathbb{Z}_N \cong \mathbb{Z}_p \times \mathbb{Z}_q$, the equation (2.2.11) is equivalent to the system

$$\left\{ \begin{array}{lll} X^{ed} & \equiv & X \pmod{p} \\ X^{ed} & \equiv & X \pmod{q} \end{array} \right. . \tag{2.2.12}$$

Let us show (2.2.12) only $\bmod p$ as $\bmod q$ is analogous. Suppose first that $p|X$. Then $X \equiv 0 \pmod{p}$ and therefore

$$\begin{aligned} X^{ed} &\equiv 0 \pmod{p} \\ &\equiv X \pmod{p}\,. \end{aligned}$$

Suppose now that X is coprime with p. Then,

$$X^{\varphi(p)} = X^{p-1} \equiv 1 \pmod{p}\,.$$

On the other hand, since $\gcd(e, (p-1)(q-1)) = 1$, there exist $\lambda, \mu \in \mathbb{Z}$ such that

$$\lambda(p-1)(q-1) + \mu e = 1\,. \tag{2.2.13}$$

By choosing $d = \mu$ we obtain,

$$\begin{aligned} X^{ed} &= X^{1-\lambda(p-1)(q-1)} \\ &= X\left(X^{p-1}\right)^{-\lambda(q-1)} \equiv X \pmod{p}\,. \end{aligned}$$

$\blacksquare$

Remark 2.2.32 We see that finding d in (2.2.13) is equivalent to finding the prime factors p, q of N. $\Diamond$

Exercise 2.2.13 *For $N = 15 = 3 \times 5$ and $e = 3$ ($\gcd(e, 2 \times 4) = \gcd(3, 8) = 1$) find d such that Alice can decode the message Bob encrypted as $[X]^3$ in $\mathbb{Z}_{15}$.*

Solution. We have to find $\lambda, \mu \in \mathbb{Z}$ such that, $8 \cdot \lambda + 3 \cdot \mu = 1$. In this case we can just guess the solution, $\lambda = -1, \mu = 3$. Then we know from Theorem 2.2.31 that the decryption exponent can be chosen, $d = \mu = 3$, and therefore,

$$\left([X]^3\right)^3 = [X]^9 = [X]\,, \quad \forall [X] \in \mathbb{Z}_{15}.$$

$\blacksquare$

2.2.4 Actions of Groups

Definition 2.2.33 (action of a group) *Let G be a group and M a set. An action of $G \overset{\varphi}{\curvearrowright} M$ of G on M is an homomorphism*

$$\begin{aligned} \varphi\,:\, G &\longrightarrow \mathrm{Sym}(M) \\ g &\longmapsto \varphi_g\,. \end{aligned}$$

The action φ is called effective if $\ker(\varphi) = \{e\}$.

Remark 2.2.34 A non effective action $G \overset{\varphi}{\curvearrowright} M$ defines, via the isomorphism theorem, an effective action of $\tilde{G} = G/\ker(\varphi)$ on M. $\Diamond$

Lauritzen [La, Def. 2.10.1] gives a different definition of action: an action of G on M is a map

$$\hat{\varphi} : G \times M \longrightarrow M \tag{2.2.14}$$

satisfying the following properties

(i) $\hat{\varphi}(e, x) = x$,

(ii) $\hat{\varphi}(g_1, \hat{\varphi}(g_2, x)) = \hat{\varphi}(g_1\, g_2, x)$, $\quad \forall g_1, g_2 \in G, \forall x \in M$.

Exercise 2.2.14 *Show that that the two definitions of group action are equivalent by relating φ in the Definition 2.2.33 with $\hat{\varphi}$ in (2.2.14), through*

$$\hat{\varphi}(g, x) = \varphi_g(x) \,. \tag{2.2.15}$$

Solution.

⇒ We assume that $\varphi : G \longrightarrow \mathrm{Sym}(M)$ is an homomorphism and that $\hat{\varphi}$ is given by (2.2.15). Let us prove that $\hat{\varphi}$ satisfies the properties *(i)* and *(ii)* of the second definition.

(i)

$$\hat{\varphi}(e, x) = \varphi_e(x) = \mathrm{id}_M(x) = x \quad .$$

(ii)

$$\hat{\varphi}(g_1, \hat{\varphi}(g_2, x)) = \varphi_{g_1}(\varphi_{g_2}(x)) = \varphi_{g_1 g_2}(x) = \hat{\varphi}(g_1 g_2, x) \,, \quad \forall g_1, g_2 \in G \,, \forall x \in M \,.$$

⇐ We now assume that $\hat{\varphi}$ satisfies the properties *(i)* and *(ii)* above and need to show that φ, defined by (2.2.15), is an homomorphism to $\mathrm{Sym}(M)$. First we have to show that, for every $g \in G$, the map φ_g is a permutation of M, i.e. is bijective or, equivalently, has inverse. Let us show that the inverse is given by $\varphi_{g^{-1}}$.

$$\begin{aligned}(\varphi_{g^{-1}} \circ \varphi_g)(x) &= \varphi_{g^{-1}}(\varphi_g(x)) = \hat{\varphi}(g^{-1}, \hat{\varphi}(g, x)) \\ &= \hat{\varphi}(g^{-1}g, x) = \hat{\varphi}(e, x) = x \,, \quad \forall g \in G, \forall x \in M \,.\end{aligned}$$

Analogously,

$$\begin{aligned}(\varphi_g \circ \varphi_{g^{-1}})(x) &= \varphi_g(\varphi_{g^{-1}}(x)) = \hat{\varphi}(g, \hat{\varphi}(g^{-1}, x)) \\ &= \hat{\varphi}(e, x) = x \,, \quad \forall g \in G, \forall x \in M \,,\end{aligned}$$

so that indeed $\varphi_g \in \mathrm{Sym}(M)$, $\forall g \in G$. Let us now show that φ is a group homomorphism.

$$\begin{aligned}\varphi_{g_1 g_2}(x) &= \hat{\varphi}(g_1 g_2, x) = \hat{\varphi}(g_1, \hat{\varphi}(g_2, x)) \\ &= \varphi_{g_1}(\varphi_{g_2}(x)) = (\varphi_{g_1} \circ \varphi_{g_2})(x) \,, \quad \forall g_1, g_2 \in G, \forall x \in M \,.\end{aligned}$$

■

Example 2.2.35 *Of course the most natural actions are by subgroups H of* $\mathrm{Sym}(M)$*,* $H \subset \mathrm{Sym}(M)$*, on M given by,*

$$\begin{aligned} \varphi &= \left(\mathrm{id}_{\mathrm{Sym}(M)}\right)_{|H} \\ \varphi_h &= h \\ \varphi_h(x) &= h(x)\,, \quad \forall x \in M\,. \end{aligned}$$

Subgroups of $\mathrm{Sym}(M)$ *were the first to be studied and are sometimes called "concrete" groups. On the other hand effective actions, φ of G, correspond to injective homomorphisms so that G is isomorphic to its image* $\mathrm{Im}(\varphi)$ *in* $\mathrm{Sym}(M)$*.*

Definition 2.2.36 (orbits, stabilizers and fixed points) *Consider the action $G \overset{\varphi}{\curvearrowright} M$.*

(a) [orbits and orbit space] Let $x_0 \in M$. The set

$$\varphi_G\, x_0 := \{\varphi_g(x_0)\,, g \in G\}\ ,$$

is called orbit of G through x_0. The set of orbits or orbit space defines a partition of M and is denoted by

$$M/\varphi_G \quad (= M/G)\,.$$

(b) [stabilizers] Let,

$$\varphi_g\, X := \{\varphi_g\, x\,,\, x \in X\}\ .$$

The subgroup of G:

$$G_X = \{g \in G\ :\ \varphi_g\, X = X\}\ ,$$

is called stabilizer of the set $X \subset M$. The stabilizer of a point $\{x_0\}$

$$G_{\{x_0\}} = G_{x_0} = \{g \in G\ :\ \varphi_g\, x_0 = x_0\}\ ,$$

is also called isotropy subgroup of x_0.

(c) [fixed points] The points,

$$x \in M\ :\ G_x = G\,,$$

are called fixed points of the G action. The set of all G–fixed points is denoted by,

$$M^G = \{x \in M\ :\ \varphi_g\, x = x\,,\ \forall g \in G\}\ .$$

Exercise 2.2.15 (HW) *Given an action $G \overset{\varphi}{\curvearrowright} M$ show that the following relation on M,*

$$x_1 \sim x_2 \Leftrightarrow \exists\, g \in G\ :\ \varphi_g\, x_1 = x_2\,,$$

is an equivalence relation and that the equivalence classes are the G–orbits,

$$[x] = \varphi_G\, x = \{\varphi_g\, x\,,\ g\ \in G\}\ .$$

Example 2.2.37 *Consider S_3 and the standard actions on $M = \{x_1, x_2, x_3\}$ by its subgroups $H = \{e, a\}$ and $A_3 = \{e, d, f\}$ (see example 2.2.3 and exercise 2.2.5).*

(a) $H \overset{\varphi}{\curvearrowright} \{x_1, x_2, x_3\}$.
Since the orbits are the sets, $\varphi_H(x) = \{x, a(x)\}$, *there are two orbits,*

$$\begin{aligned} \varphi_H(1) &= \{1\} \\ \varphi_H(2) &= \{2, a(2) = 3\} = \{2, 3\} \ , \end{aligned}$$

where we used the fact that $a(1) = 1$*. The orbit space has therefore two points,*

$$\{x_1, x_2, x_3\} / H = \{[1], [2]\} = \{\{1\}, \{2, 3\}\} \ .$$

The only fixed point, corresponding to the only orbit with just one point, is $x = 1$,

$$\{1, 2, 3\}^H = \{1\} \, .$$

(b) *Transitive action:* $A_3 \overset{\varphi}{\curvearrowright} \{x_1, x_2, x_3\}$.
There is only one orbit of A_3,

$$\varphi_{A_3}(1) = \{1, d(1) = 2, f(1) = 3\} = M \, .$$

Such actions, with only one orbit, or equivalently, such that for any pair of points, $x, y \in M$, *there exists a group element* $g \in G$ *taking* x *to* $y = \varphi_g(x)$, *are called transitive actions.*

Example 2.2.38 *Consider the standard action of* $GL_n(\mathbb{R})$,

$$GL_n(\mathbb{R}) = \{A \in \mathrm{Mat}(\mathbb{R}^n) \, : \, \det(A) \neq 0\} \ ,$$

and its subgroups on $\mathbb{R}^n$ *by matrix multiplication,*

$$\varphi_A(x) = Ax \, ,$$

with

$$x = (x_1, \ldots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1 \ \ldots \ x_n)^T$$

where, for simplicity of the notation, we identify $\mathbb{R}^n$ *with the space of column* $(n \times 1)$ *matrices.*

(a) $GL_n(\mathbb{R}) \overset{\varphi}{\curvearrowright} \mathbb{R}^n$.
This action has only two orbits: The fixed point $\{0\} = (\mathbb{R}^n)^{GL_n(\mathbb{R})}$ *and its complement,*

$$\varphi_{GL_n(\mathbb{R})}(x_0) = \mathbb{R}^n \setminus \{0\} \ , \tag{2.2.16}$$

where $x_0 \neq 0$*. To prove (2.2.16), i.e. that the action of* $GL_n(\mathbb{R})$ *is transitive on* $\mathbb{R}^n \setminus \{0\}$, *consider two arbitrary vectors,* $x', x'' \in \mathbb{R}^n \setminus \{0\}$ *and two ordered basis,*

$$\begin{aligned} B' &= \{v'_1, \ldots, v'_n\} \ , \\ B'' &= \{v''_1, \ldots, v''_n\} \end{aligned}$$

of $\mathbb{R}^n$, chosen such that their first vectors are equal to x', x'', i.e. $v_1' = x'$ and $v_1'' = x''$. Let us find a nonsingular matrix C taking x' to x''. Then, let e_1 be the first vector of the canonical basis of $\mathbb{R}^n$,

$$e_1 = (1, 0, \ldots, 0) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} ,$$

and A', A'' be the two matrices in $GL_n(\mathbb{R})$ having v_j' and v_j'' as their j-th column, respectively, $j = 1, \ldots, n$.

we have that

$$\begin{aligned} A'e_1 &= v_1' = x' \\ A''e_1 &= v_1'' = x'' . \end{aligned}$$

Then we can choose $C = A''(A')^{-1}$. Indeed,

$$C\, x' = A''(A')^{-1}\, x' = x'' ,$$

which shows transitivity.

(b) $SO_n(\mathbb{R}) \overset{\varphi}{\curvearrowright} \mathbb{R}^n$.
The group of special (i.e. determinant one), orthogonal matrices $n \times n$ is

$$SO_n(\mathbb{R}) = \left\{ A \in \mathrm{Mat}_n(\mathbb{R}) \,:\, AA^T = I_n \,, \det(A) = 1 \right\} .$$

Consider the usual inner product on $\mathbb{R}^n$,

$$\langle y, x \rangle = y^T x = (y_1 \, \cdots \, y_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j=1}^{n} y_j x_j \,,$$

and the associated norm,

$$||x|| = \sqrt{\langle x, x \rangle} \,.$$

Let us show that, by restricting from $GL_n(\mathbb{R})$ to $SO_n(\mathbb{R})$, the orbits in $\mathbb{R}^n \setminus \{0\}$, become the $n-1$–dimensional spheres with center in 0,

$$S_r^{n-1} = \left\{ x \in \mathbb{R}^n \,:\, ||x||^2 = \sum_{j=1}^{n} x_j^2 = r^2 \right\} .$$

Orthogonal matrices preserve inner products,

$$\begin{aligned} \langle Ay, Ax \rangle &= (Ay)^T\, Ax = y^T A^T Ax \\ &= y^T I_n x = y^T x = \langle y, x \rangle \,, \quad \forall y, x \in \mathbb{R}^n,\ A \in SO_n(\mathbb{R}) \,, \end{aligned}$$

and thus also norms, $||Ax|| = ||x||$. This implies that the $SO_n(\mathbb{R})$–orbits are contained in the spheres S_r^{n-1}. Indeed, let x_0 be a norm r vector. Then,

$$\varphi_{SO_n(\mathbb{R})}\, x_0 = \{Ax_0\,,\ A \in SO_n(\mathbb{R})\} \subset S_r^{n-1}\,.$$

Let us show that the inclusion above is in fact an equality, or, equivalently, the spheres S_r^{n-1} are $SO_n(\mathbb{R})$–orbits or also $SO_n(\mathbb{R})$ acts transitively on S_r^{n-1}. Consider two vectors x', x'' in S_r^{n-1} and suppose that they both have their first component different from zero. [1] *Complement x', x'' with the basis vectors of the canonical basis, $e_2, \ldots, e_n$ to get two bases,*

$$\begin{aligned} \tilde{B}' &= \left\{\frac{x'}{r}, e_2, \ldots, e_n\right\} \\ \tilde{B}'' &= \left\{\frac{x''}{r}, e_2, \ldots, e_n\right\}\,, \end{aligned}$$

and apply Gram-Schmidt orthogonalization to both these bases to get two orthonormal bases

$$\begin{aligned} B' &= \{u'_1, \ldots, u'_n\} \\ B'' &= \{u''_1, \ldots, u''_n\}\,, \end{aligned}$$

with $u'_1 = \frac{x'}{r}$ and $u''_1 = \frac{x''}{r}$. Consider now the matrices A', A'' having u'_j and u''_j as j–th columns, respectively. If one or both of these matrices has determinant equal to -1 just replace its last column by its symmetric, i.e. u'_n and/or u''_n replaced by $-u'_n$ and/or $-u''_n$. We have thus now two special orthogonal matrices, $A', A'' \in SO_n(\mathbb{R})$, such that

$$\begin{aligned} A'e_1 &= u'_1 = \frac{x'}{r} \\ A''e_1 &= u''_1 = \frac{x''}{r}\,. \end{aligned}$$

or, equivalently,

$$\begin{aligned} A'r\, e_1 &= r\, u'_1 = x' \\ A''r\, e_1 &= r\, u''_1 = x''\,. \end{aligned} \qquad (2.2.17)$$

Then,

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

so that the action of $SO_n(\mathbb{R})$ on spheres is indeed transitive.

The orbit space, $\mathbb{R}^n/\varphi_{SO_n(\mathbb{R})}$, is naturally bijective to $\mathbb{R}_{\geq 0}$, with bijection given by,

$$\begin{aligned} \mathbb{R}^n/\varphi_{SO_n(\mathbb{R})} &\longrightarrow \mathbb{R}_{\geq 0} \\ [x] = SO_n(\mathbb{R})\, x &\longmapsto ||x||\,. \end{aligned}$$

[1] if one or both the vectors x', x'' have their first component equal to zero just adapt the construction by complementing one or both of these vectors with $n-1$ vectors of the canonical basis that form a basis with them.

Exercise 2.2.16 *Consider the standard action* $SO_n(\mathbb{R}) \overset{\varphi}{\curvearrowright} \mathbb{R}^n$*. Show that the stabilizer of every* $x_0 \neq 0$*,* $(SO_n(\mathbb{R}))_{x_0}$*, is isomorphic to* $SO_{n-1}(\mathbb{R})$*.*

Before solving this exercise let us prove a relevant (to the exercise) general result.

Proposition 2.2.39 *Consider the action* $G \overset{\varphi}{\curvearrowright} M$ *and let* $x_1, x_2 \in M, g \in G$ *be such that*

$$x_2 = \varphi_g(x_1), \tag{2.2.18}$$

i.e. x_1, x_2 *are in the same* G*–orbit. Then, their stabilizers are conjugate subgroups of* G*,*

$$G_{x_2} = g\, G_{x_1}\, g^{-1}\,. \tag{2.2.19}$$

Proof.

It will be sufficient to show that

$$g\, G_{x_1}\, g^{-1} \subset G_{x_2} \tag{2.2.20}$$

for all triples x_1, x_2, g satisfying (2.2.18). Indeed, the opposite (to (2.2.20)) inclusion,

$$g\, G_{x_1}\, g^{-1} \supset G_{x_2}\,, \tag{2.2.21}$$

is equivalent, by conjugating with g^{-1} both sides (i.e. by multiplying on the left by g^{-1} and on the right by g), to the inclusion

$$g^{-1}\, G_{x_2}\, g \subset G_{x_1}\,,$$

which is equivalent to (2.2.20) with the triple (x_1, x_2, g) replaced by (x_2, x_1, g^{-1}). So we are left with proving (2.2.20). Let $\tilde{h} \in g\, G_{x_1}\, g^{-1}$, i.e. exists $h \in G_{x_1}$ such that $\tilde{h} = ghg^{-1}$. Then

$$\begin{aligned} \varphi_{\tilde{h}}(x_2) &= \varphi_{ghg^{-1}}(x_2) = \left(\varphi_g \circ \varphi_h \circ \varphi_{g^{-1}}\right)(x_2) \\ &= \left(\varphi_g \circ \varphi_h\right)(x_1) = \varphi_g(\varphi_h(x_1)) = x_2\,, \end{aligned}$$

which implies that $\tilde{h} \in G_{x_2}$.

■

We are now ready to solve the exercise 2.2.16.

Solution. The best way to find the stabilizer of a point $x_0 \in M$ is usually to find the stabilizer of a simpler point in the same orbit as x_0. We saw in example 2.2.38 (b) that the orbits are spheres. So as a simpler vector in the same orbit as x_0 let us choose $||x_0||\, e_1$. For the stabilizer of $||x_0||\, e_1$ we find,

$$\begin{aligned} (SO_n(\mathbb{R}))_{||x_0||e_1} &= (SO_n(\mathbb{R}))_{e_1} = \{A \in SO_n(\mathbb{R}) \,:\, A\, e_1 = e_1\} \\ &= \left\{ A = \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{A} \end{pmatrix} \,,\ \tilde{A} \in SO_{n-1}(\mathbb{R}) \right\} \cong SO_{n-1}(\mathbb{R})\,, \end{aligned}$$

where, in the first column of A,

$$0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathrm{Mat}_{(n-1)\times 1}(\mathbb{R}) = \mathbb{R}^{n-1}$$

and we used the fact that a matrix A is orthogonal if and only if its columns form an orthonormal basis. Therefore

$$(SO_n(\mathbb{R}))_{x_0} \cong (SO_n(\mathbb{R}))_{||x_0||e_1} = (SO_n(\mathbb{R}))_{e_1} \cong SO_{n-1}(\mathbb{R})\,,$$

where the first isomorphism is a conjugation isomorphism by an orthogonal matrix A_0 rotating $||x_0||e_1$ to x_0, as in (2.2.17),

$$(SO_n(\mathbb{R}))_{x_0} = A_0\,(SO_n(\mathbb{R}))_{||x_0||e_1}\,A_0^{-1}\,.$$

■

Definition 2.2.40 *An action $G \overset{\varphi}{\curvearrowright} M$ is called free if the stabilizers of all points are trivial, $G_x = \{e\}\,,\ \forall x \in M$.*

Example 2.2.41 *Very natural examples of free actions are those of subgroups $H \subset G$ on G by left and right translations. Left and right translations of G by a group element, h, are the following bijective maps,*

$$\begin{array}{rcl} L_h \,:\, G & \longrightarrow & G \\ L_h(g) & = & hg\,, \end{array}$$

and

$$\begin{array}{rcl} R_h \,:\, G & \longrightarrow & G \\ R_h(g) & = & gh\,. \end{array}$$

Then the following maps,

$$\begin{array}{rcl} L \,:\, H & \longrightarrow & \mathrm{Sym}(G) \\ h & \longmapsto & L_h\,, \end{array}$$

and

$$\begin{array}{rcl} \tilde{R} \,:\, H & \longrightarrow & \mathrm{Sym}(G) \\ h & \longmapsto & R_{h^{-1}}\,, \end{array}$$

are well defined and define injective group homomorphisms (check this!) and therefore define effective actions of H on G, for any subgroup $H \subset G$, which furthermore are free.

The left H orbits,

$$L_H\, g = H\, g\,,$$

are right H cosets and the right H orbits,

$$\tilde{R}_H\, g = g\, H\,,$$

are left H cosets, so that we obtain,

$$\begin{array}{rcl} G/H & = & G/\tilde{R}_H \\ H \setminus G & = & G/L_H\,. \end{array}$$

Example 2.2.42 *The left action of G on itself descends to a non free but transitive action of G on $G/H = G/\tilde{R}_H$, that we call canonical action, $G \overset{\varphi^{\mathrm{can}}}{\curvearrowright} G/H$,*

$$\varphi^{\mathrm{can}}_{g_1}[g] = \varphi^{\mathrm{can}}_{g_1}\, gH = g_1 g\, H = [g_1 g]\,.$$

The transitivity of this action follows from the transitivity of the left action on G. Indeed, let $[g_1], [g_2] \in G/H$. Then,

$$\varphi^{\mathrm{can}}_{g_2 g_1^{-1}}([g_1]) = [g_2 g_1^{-1} g_1] = [g_2]\,.$$

The class of the neutral element $[e]$ is called origin of the coset space G/H. The stabilizer of the origin of G/H is H,

$$G_{[e]} = \left\{ g \in G \,:\, \varphi^{\mathrm{can}}_g\,[e] = [g] = [e] = H \right\} = H\,.$$

Notice that descending the right action of G on itself to $H \setminus G$ leads to analogous results.

Remark 2.2.43 We already saw that G/N has the structure of group, induced from that of G, if N is a normal subgroup. What we saw in the example 2.2.42 is that the coset spaces G/H (and $H \setminus G$) have another structure (for any subgroup H!): that of spaces of transitive actions of G.

We will see soon that much more than simple examples of transitive actions, the coset spaces, $G \overset{\varphi^{\mathrm{can}}}{\curvearrowright} G/H$, provide, up to equivalence of actions, *all* transitive actions. ♢

Definition 2.2.44 (equivalence of actions) *Let $G \overset{\varphi}{\curvearrowright} M$ and $G' \overset{\varphi'}{\curvearrowright} M'$ be two G–actions. Then,*

(i) A map, $T \,:\, M \longrightarrow M'$, is called G–equivariant if

$$T \circ \varphi_g = \varphi'_g \circ T\ , \quad \forall g \in G\,.$$

(ii) An equivariant map, $T \,:\, M \to M'$, is called an equivalence of G–actions if it is bijective. Two actions are said to be equivalent if there exists an equivalence of G–actions, $T \,:\, M \longrightarrow M'$.

Exercise 2.2.17 (HW) *Let $T \,:\, M \longrightarrow M'$ be a G–equivariant map. Show that T maps G–orbits to G–orbits, and therefore induces a map, $\widetilde{T}$, from M/φ_G to M'/φ'_G. Describe $\widetilde{T}$.*

Theorem 2.2.45 (transitive actions) *Let G act transitively on M, $G \overset{\varphi}{\curvearrowright} M$. Then*

$$G \overset{\varphi}{\curvearrowright} M \;\cong\; G \overset{\varphi^{\mathrm{can}}}{\curvearrowright} G/G_{x_0}\,,$$

where G_{x_0} is the stabilizer of any point, $x_0 \in M$. The equivalence is given by the map,

$$\begin{array}{rcl} T_{x_0} \,:\, G/G_{x_0} & \longrightarrow & M \\ g\,G_{x_0} = [g] & \longmapsto & \varphi_g(x_0)\,. \end{array} \tag{2.2.22}$$

Proof. Let us first show that T_{x_0} is well defined. Consider the map

$$\begin{array}{rcl} \widehat{T}_{x_0} \,:\, G & \longrightarrow & M \\ g & \longmapsto & \varphi_g(x_0)\,. \end{array}$$

Due to the transitivity of the action $G \overset{\varphi}{\curvearrowright} M$, the map $\widehat{T}_{x_0}$ is surjective and the equivalence relation it defines on G,

$$\begin{array}{rcl} g_1 \sim g_2 & \Leftrightarrow & \widehat{T}_{x_0}(g_1) = \widehat{T}_{x_0}(g_2) \Leftrightarrow \varphi_{g_1}(x_0) = \varphi_{g_2}(x_0) \\ & \Leftrightarrow & \varphi_{g_2^{-1}g_1}(x_0) = x_0 \,\Leftrightarrow\, g_2^{-1}g_1 \in G_{x_0} \\ & \Leftrightarrow & g_1 G_{x_0} = g_2 G_{x_0}\,, \end{array}$$

coincides with that corresponding to the right action of G_{x_0} on G, whose orbits are the left G_{x_0}–cosets. Therefore $G/G_{x_0} = G/\sim$ and, as in exercise 2.1.1, $\widehat{T}_{x_0}$ defines the injective map $T_{x_0} \,:\, G/G_{x_0} \longrightarrow M$, as the unique map satisfying,

$$T_{x_0} \circ \pi_{x_0} = \widehat{T}_{x_0}\,,$$

where π_{x_0} denotes the canonical projection. This is the map in (2.2.22). Since $\widehat{T}_{x_0}$ is surjective, T_{x_0} is also surjective and thus bijective.

We are only left with proving that T_{x_0} is G–equivariant.

$$\begin{array}{rcl} \left(T_{x_0} \circ \varphi^{\mathrm{can}}_{g_1}\right)([g]) & = & T_{x_0}([g_1 g]) = \varphi_{g_1 g}(x_0) \\ & = & \varphi_{g_1}(\varphi_g(x_0)) = \varphi_{g_1}\left(T_{x_0}([g])\right) \\ & = & \left(\varphi_{g_1} \circ T_{x_0}\right)([g])\,, \qquad \forall [g] \in G/G_{x_0}\,, \forall g_1 \in G\,. \end{array}$$

■

Remark 2.2.46 Like the isomorphism theorem was usefull to find groups H isomorphic to quotient groups G/N with N a normal subgroup (see remark 2.2.18) the above theorem is also useful to find spaces of transitive actions that are equivariantly bijective to G/H for any H: find a transitive action $G \overset{\varphi}{\curvearrowright} M$ such that G_{x_0} is conjugated to H, i.e. exists $g \in G$ such that,

$$G_{x_0} = gHg^{-1}\,.$$

◊

Exercise 2.2.18 *Show that* $S^{n-1} \cong SO_n(\mathbb{R})/\psi(SO_{n-1}(\mathbb{R}))$*, where*

$$\begin{array}{rcl} \psi \,:\, SO_{n-1}(\mathbb{R})) & \longrightarrow & SO_n(\mathbb{R}) \\ \tilde{A} & \longmapsto & \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{A} \end{pmatrix} \end{array} \quad ,$$

and $S^{n-1} = S^{n-1}_1 = \{x \in \mathbb{R}^n \,:\, ||x|| = 1\}$*.*

Solution. This is a direct consequence of the Theorem 2.2.45. Indeed, we have already seen in exercise 2.2.16 that $(SO_n(\mathbb{R}))_{e_1} = \psi(SO_{n-1}(\mathbb{R}))$ so that S^{n-1} is bijective to $SO_n(\mathbb{R})/\psi(SO_{n-1}(\mathbb{R}))$ with (inverse) G–equivariant bijection, T_{e_1} given by,

$$\begin{array}{rcl} T_{e_1} \,:\, SO_n(\mathbb{R})/\psi(SO_{n-1}(\mathbb{R})) & \longrightarrow & S^n \\ [A] & \longmapsto & A\,e_1\,. \end{array}$$

■

Conjugation action

Of big importance to the representation theory of finite (and not only) groups is the conjugation action of G on itself.

$$\begin{array}{rcl} G & \longrightarrow & \mathrm{Sym}(G) \\ g_1 & \longmapsto & \varphi^c_{g_1} \\ \varphi^c_{g_1}(g) & = & g_1 g g_1^{-1} \,. \end{array}$$

The orbits of the conjugation action

$$\varphi^c_G(g) = \left\{\varphi^c_{g_1}(g) = g_1 g g_1^{-1}\,,\ \ g_1 \in G\right\} =: C(g)\,,$$

are called conjugacy classes of G or also conjugacy classes of $g \in G$. The stabilizer of an element $g \in G$,

$$G_g = \left\{g_1 \in G\ :\ g_1 g g_1^{-1} = g \Leftrightarrow g_1 g = g g_1\right\} =: Z(g)\,,$$

is called centralizer of g.

The stabilizer of a subgroup $H \subset G$,

$$G_H = \left\{g \in G\ :\ gHg^{-1} = H \Leftrightarrow gH = Hg\right\}\ =: N_G(H)\,,$$

is called normalizer of H in G. It is the maximal subgroup of G containing H as a normal subgroup.

The fixed point set under the conjugation action is

$$\begin{array}{rcl} G^G & = & \left\{g \in G\ :\ \varphi^c_{g_1}(g) = g_1 g g_1^{-1} = g\,,\quad \forall g_1 \in G\right\} \\ & = & \left\{g \in G\ :\ g_1 g = g g_1\,,\quad \forall g_1 \in G\right\}\ =: Z(G) \end{array}$$

is called center of G.

Symmetric group

Let us study the conjugacy classes of the symmetric group S_n. But first let us simplify notation and formulate some results. Let $j_1, \dots, j_k \in \{1, 2, \dots, n\}$ be k distinct elements and let us represent by $(j_1\ \dots\ j_k)$ the following permutation $\sigma \in S_n$:

$$\begin{array}{rcl} \sigma(j_l) & = & j_{l+1} \\ \sigma(x) & = & x\,,\ \text{if}\ x \notin \{j_1, \dots, j_k\} \end{array}$$

where $j_{k+1} = j_1$. We call this permutation a cycle of length k.

Notice that the cycles of length one are all equal to the neutral element, $(j) = e$, for every j. Cycles of length two, (ij), are called transpositions. It is easy to see that transpositions are their own inverses,

$$\begin{array}{rcl} ((i\,j) \circ (i\,j))\,(i) & = & (i\,j)\,((i\,j)(i)) = (ij)(j) = i \\ ((i\,j) \circ (i\,j))\,(j) & = & (i\,j)\,((i\,j)(j)) = (i\,j)(i) = j \\ ((i\,j) \circ (i\,j))\,(x) & = & (i\,j)\,((i\,j)(x)) = (i\,j)(x) = x\,,\quad \text{for every}\ x \notin \{i, j\}\,, \end{array}$$

so that indeed,

$$(i\,j)\circ(i\,j) = (i\,j)^2 = e\,.$$

A transposition (ij) is called simple if $|i-j|=1$. There are $n-1$ simple transpositions in S_n, $s_1=(1\,2), s_2=(2\,3),\dots,s_{n-1}=((n-1)\,n)$.

An important structure result is the following.

Theorem 2.2.47 *The symmetric group is isomorphic to the group with $n-1$ generators subject to the following relations,*

$$\begin{array}{rcl} s_i^2 &=& e \quad \forall i \\ s_i\circ s_j &=& s_j\circ s_i \quad \text{if } |i-j|>1 \\ s_i\circ s_{i+1} && \textit{is of order } 3\,, \quad i\le n-1\,. \end{array}$$

In the cycle notation for S_3 we have (see notation in example 2.2.3).

$$\begin{array}{rcl} a &=& (23) = s_2 \\ b &=& (13) \\ c &=& (12) = s_1 \\ d &=& (123) \\ f &=& (132)\,. \end{array}$$

The cycles have cyclic symmetry

$$(j_1\,j_2\dots\,j_{k-1}\,j_k) = (j_k\,j_1\,j_2\dots\,j_{k-2}\,j_{k-1})\,,$$

have order equal to their length,

$$(j_1\,\dots\,j_k)^k = e\,,$$

and are equal to products of overlapping cycles,

$$(j_1\,\dots j_{r-1}\,j_r)\circ(j_r\,j_{r+1}\dots\,j_k) = (j_1\,\dots\,j_{r-1}\,j_r\,j_{r+1}\dots\,j_k)\,.$$

So, for example, in S_3 we get

$$\begin{array}{rcl} d &=& (1\,2\,3) = (1\,2)\circ(2\,3) = s_1\circ s_2 \\ f &=& (1\,3\,2) = (1\,3)\circ(3\,2) = (2\,1\,3) = (2\,1)\circ(1\,3) \end{array}$$

Let us represent $(1\,3)$ as a product of the two simple transpositions. Let us do that first for f,

$$f = (1\,3\,2) = (3\,2\,1) = (3\,2)\circ(2\,1) = (2\,3)\circ(1\,2) = s_2\circ s_1\,.$$

Now, since also,

$$f = (2\,1)\circ(1\,3)$$

we conclude that

$$(1\,3) = (2\,1)\circ f = s_1\circ s_2\circ s_1\,.$$

Proposition 2.2.48 (see Proposition 2.9.6 of [La]) *Every permutation $\sigma\in S_n$ is a product of unique disjoint cycles (i.e. cycles that involve disjoint subsets of $\{1,2,\dots,n\}$).*

Exercise 2.2.19 *Represent the permutation*

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 2 & 1 & 7 & 6 \end{pmatrix}$$

as a product of disjoint cycles.

Solution. We have,

$$\sigma = (1\,3\,5) \circ (2\,4) \circ (6\,7) = (1\,3\,5)(2\,4)(6\,7)\,.$$

■

Definition 2.2.49 (sign of a permutation) *The sign of a permutation, $\sigma \in S_n$, is*

$$\operatorname{sgn}(\sigma) = (-1)^{n(\sigma)}\,, \tag{2.2.23}$$

where $n(\sigma)$ is the minimal number of simple transpositions in the representation of σ as a product of simple transpositions. Permutations with sign $+1$ are called even and permutations with sign -1 are called odd.

Remark 2.2.50 The representation of σ as a product of simple transpositions is not unique but the number of transpositions in the representation of σ as a product of (not necessarily simple) transpositions is well defined mod 2. Thus, in the exponent in (2.2.23), we could as well put the number of transpositions in *any* representation of σ as a product of transpositions. ♢

Proposition 2.2.51 (see Proposition 2.9.6 of [La]) *The map* sgn *is a group homomorphism*

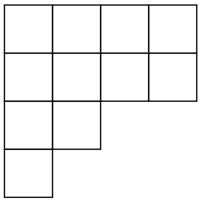
$$\begin{array}{rcl} \operatorname{sgn}\,:\, S_n & \longrightarrow & \{1, -1\} \\ \sigma & \longmapsto & \operatorname{sgn}(\sigma)\,. \end{array}$$

The kernel of ker(sgn) is a normal subgroup of S_n called alternating subgroup (of even permutations), A_n. Since, for $n > 1$, sgn is surjective, we find from Lagrange Theorem that

$$|A_n| = |S_n/2| = \frac{n!}{2}\,.$$

Definition 2.2.52 (partition and Young diagram) *A partition λ of $n \in \mathbb{N}$ is a tuple of natural numbers, $\lambda = (i_1, \ldots, i_k)$, such that, $i_1 \geq i_2 \geq \cdots \geq i_k$, and $i_1 + \cdots + i_k = n$. To a partition λ of n we associate a diagram with i_1 boxes in the first row, i_2 in the second, $\ldots$, i_k in the last row. This diagram is called the Young diagram of the partition λ.*

The Young diagram of the partition $\lambda = (4, 4, 2, 1)$ of 11 is:



Very important is the following result on the conjugacy classes of the symmetric group.

Theorem 2.2.53 (conjugacy classes of S_n) *Let $\sigma \in S_n$ and*

$$\sigma = \sigma_1 \sigma_2 \dots \sigma_k \tag{2.2.24}$$

be the representation of σ as a product of disjoint cycles of decreasing lengths, i_j, corresponding to a partition of n,

$$\lambda = (i_1, \dots, i_k) .$$

The conjugacy class of σ is given by all permutations with cycle decomposition having the same sequence of lengths as σ, i.e.

$$C(\sigma) = D_{i_1,\dots,i_k} ,$$

where

$$D_{i_1,\dots,i_k} := \{\tilde{\sigma} \in S_n \, : \, \tilde{\sigma} = \tilde{\sigma}_1 \dots \tilde{\sigma}_k \, , \textit{the cycles } \tilde{\sigma}_j \textit{ are disjoint and } \tilde{\sigma}_j \textit{ has length } i_j\} .$$

Proof. We will need the following Lemma.

Lemma 2.2.54 *Let σ and $(j_1 \dots j_k)$ be permutations of S_n. Then,*

$$\sigma \, (j_1 \dots j_k) \, \sigma^{-1} = (\sigma(j_1) \dots \sigma(j_k)) . \tag{2.2.25}$$

Proof of the Lemma. Let us show the equality of the permutations in (2.2.25) by showing that they are equal as maps from the set $\{1, 2, \dots, , n\}$ to itself. The rhs of (2.2.25) maps $\sigma(j_r)$ to $\sigma(j_{r+1})$ (with $\sigma(j_{k+1}) = \sigma(j_1)$). Let us now apply the lhs to $\sigma(j_r)$:

$$\begin{aligned} \left(\sigma \circ (j_1 \dots j_k) \circ \sigma^{-1}\right)(\sigma(j_r)) &= (\sigma \circ (j_1 \dots j_k))\,(j_r) \\ &= \sigma(j_{r+1}) . \end{aligned}$$

Let now $i \notin \{\sigma(j_1), \dots, \sigma(j_k)\}$. The rhs of (2.2.25) maps i to itself. The lhs maps it first (with σ^{-1}) to $\sigma^{-1}(i) \notin \{j_1, \dots, j_k\}$. Then,

$$\left(\sigma \circ (j_1 \dots j_k) \circ \sigma^{-1}\right)(i) = (\sigma \circ (j_1 \dots j_k))\,(\sigma^{-1}(i)) = \sigma\left(\sigma^{-1}(i)\right) = i ,$$

which concludes the proof of the Lemma. ■

Returning to the proof of the theorem, we find, for the conjugacy class of σ in (2.2.24),

$$\begin{aligned} C(\sigma) &= \left\{\tau \, \sigma \, \tau^{-1} \, , \, \tau \in S_n\right\} \\ &= \left\{\tau \, \sigma_1 \, \tau^{-1} \dots \tau \, \sigma_k \, \tau^{-1}, \, \tau \in S_n\right\} \end{aligned}$$

so that from the Lemma we see that

$$C(\sigma) \subset D_{i_1,\dots,i_k} .$$

Let us now show that $D_{i_1,\dots,i_k} \subset C(\sigma)$. Let $\tilde{\sigma} \in D_{i_1,\dots,i_k}$ and let

$$\tilde{\sigma} = \tilde{\sigma}_1 \dots \tilde{\sigma}_k .$$

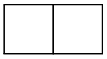
To find a permutation τ mapping σ_j to $\tilde{\sigma}_j$ for all $j = 1, \ldots, k$ just order all the cycles and the elements inside each cycle and define τ to map the cycle entries in every σ_j to the corresponding entries in $\tilde{\sigma}_j$. Then we see that

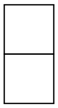
$$\tilde{\sigma} = \tau\sigma\tau^{-1} \in C(\sigma)\,.$$

■

Therefore there is a natural bijection from the conjugation orbit space of S_n and the set of all partitions of n or, equivalently, the set of all Young diagrams with n boxes.

Example 2.2.55 *Let us list the conjugacy classes of the first symmetric groups.*

(a) For S_2 we have two partitions of 2. First $\lambda_1 = (2)$ corresponds to the conjugacy class with only one element $C((12))$ and to the Young diagram  *. Then $\lambda_2 = (1,1)$ so that the Young tableau reads,*

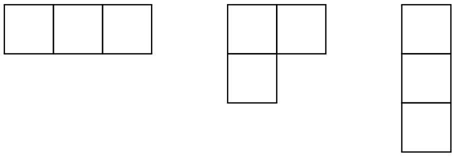


.

and the conjugacy class is the class of the identity,

$$C(e) = \{e\}\,.$$

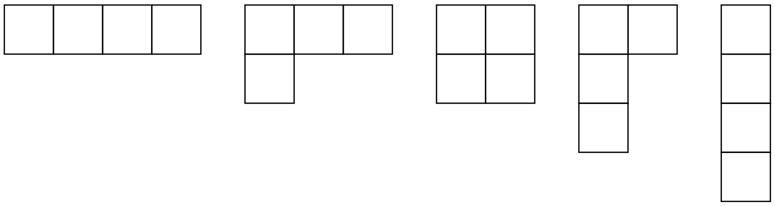
(b) For S_3 there are three partitions of 3, $\lambda_1 = (3), \lambda_2 = (2,1)$ and $\lambda_3 = (1,1,1)$, with Young diagrams given by



and corresponding conjugacy classes

$$\begin{array}{rcl} C((123)) & = & \{(1\,2\,3),\ (1\,3\,2)\} \\ C((12)(3)) & = & \{(1\,2),(1\,3),(2\,3)\} \\ C(e) & = & \{e\}\,. \end{array}$$

(c) There are five partitions of 4: $\lambda_1 = (4), \lambda_2 = (3,1), \lambda_3 = (2,2), \lambda_4 = (2,1,1)$ and $\lambda_5 = (1,1,1,1)$, with Young diagrams given by,



and the corresponding conjugacy classes

$$C((1234))\,,\ C((123))\,,\ C((12)(34))\,,\ C((12))\,,\ C(e) = \{e\}\,.$$

2.2.5 Representations of groups

In the present section we will be studying linear actions of a group G on linear spaces, called representations of G. We will restrict ourselves to complex representations of finite degree, i.e. to representations on finite dimensional complex vector spaces.

Given an ordered basis $B = (v_1, \ldots, v_n)$ on a vector space we obtain an isomorphism of vector spaces,

$$\begin{aligned} T_B : V \quad &\longrightarrow \quad \mathbb{C}^n \\ v = a_1 v_1 + \cdots + a_n v_n \quad &\longmapsto \quad (a_1, \ldots, a_n)\,. \end{aligned} \tag{2.2.26}$$

To simplify notation we will identify $\mathbb{C}^n$ with the (vector) space of $n \times 1$ column matrices, $\mathbb{C}^n = \mathrm{Mat}_{n\times 1}(\mathbb{C})$,

$$(a_1, \ldots, a_n) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

An inner product on V is a function

$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{C}\,,$$

satisfying the following properties.

(i) Linearity in the first argument:

$$\langle \alpha_1 v_1 + \alpha_2 v_2, w\rangle = \alpha_1 \langle v_1, w\rangle + \alpha_2 \langle v_2, w\rangle\,, \qquad \forall\, \alpha_1, \alpha_2 \in \mathbb{C}\,, v_1, v_2, w \in V\,.$$

(ii) Conjugate symmetry:

$$\langle v, w\rangle = \overline{\langle w, v\rangle}\,, \qquad \forall\, v, w \in V\,.$$

(iii) Positive definiteness:

$$\langle v, v\rangle > 0\,, \quad \forall\, v \in V \setminus \{0\}\,.$$

Example 2.2.56 *The standard inner product on $\mathbb{C}^n$ is,*

$$\langle v, w\rangle = v^T \cdot \bar{w} = (a_1 \ldots a_n) \begin{pmatrix} \bar{b}_1 \\ \vdots \\ \bar{b}_n \end{pmatrix} = a_1 \bar{b}_1 + \cdots + a_n \bar{b}_n\,.$$

Example 2.2.57 *Let M be a finite set. The set of all complex valued functions on M,*

$$\mathbb{C}^M = \mathrm{Map}(M, \mathbb{C})\,,$$

has a natural structure of complex vector space inherited from that of $\mathbb{C}$ and has a standard inner product,

$$\langle f_1\,,\, f_2\rangle = \sum_{x\in M} f_1(x)\, \overline{f_2(x)}\,. \tag{2.2.27}$$

Exercise 2.2.20 *Show that a basis of* $\mathbb{C}^M$ *is given by*

$$\{\delta_x \, , \, x \in M\} \, , \tag{2.2.28}$$

where

$$\delta_x(y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{else} \, . \end{cases}$$

Solution. Let us first show that,

$$\text{span}_{\mathbb{C}} \{\delta_x \, , \, x \in M\} = \mathbb{C}^M \, ,$$

by showing that, for any $f \in \mathbb{C}^M$, we have,

$$f = \sum_{x \in M} f(x) \, \delta_x \, . \tag{2.2.29}$$

Indeed, the value of the function on the rhs of (2.2.29), at the point $y \in M$, is

$$\left(\sum_{x \in M} f(x) \, \delta_x \right) (y) = \sum_{x \in M} f(x) \, \delta_x(y) = f(y) \, ,$$

which is of course equal to the value of the function f, on the lhs of the same equation, at the same point y, for every $y \in M$, which shows that the two functions are equal. Let us now show the linear independence of the set in (2.2.28).

$$\begin{aligned} \sum_{x \in M} a_x \, \delta_x &= 0 \\ \Leftrightarrow \sum_{x \in M} a_x \, \delta_x(y) &= 0 \\ \Leftrightarrow a_y &= 0 \, , \qquad \forall \, y \in M \, . \end{aligned}$$

■

We call this basis the canonical basis of $\mathbb{C}^M$,

$$B_{\text{can}}(\mathbb{C}^M) = \{\delta_x \, , \, x \in M\} \, .$$

The canonical basis is orthonormal for the inner product (2.2.27).

Definition 2.2.58 *Let* $(V, \langle \ , \ \rangle)$ *be a complex inner product space,* $\dim V < \infty$*. Then* $U \in \text{Hom}(V, V)$ *is called a unitary linear transformation or unitary operator if*

$$\langle U(v), U(w) \rangle = \langle v, w \rangle \, , \qquad \forall v, w \in V \, .$$

Exercise 2.2.21 *Show that a unitary operator* U *is necessarily invertible,* $U \in GL(V)$*.*

Solution. We have that,

$$v \in \ker(U) \Rightarrow \langle U(v), U(v) \rangle = \langle v, v \rangle = 0 \Leftrightarrow v = 0 \, ,$$

so that $\ker(U) = \{0\}$, i.e. $U \in \text{Hom}(V, V)$ is injective and therefore, since $\dim(V) < \infty$, U is bijective. ■

Example 2.2.59 *We will identify square $n \times n$ matrices A with with the linear transformations on $\mathbb{C}^n$, having the matrix A in the canonical basis of $\mathbb{C}^n$ or, equivalently,*

$$\begin{array}{rcl} \mathbb{C}^n & \longrightarrow & \mathbb{C}^n \\ \\ v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} & \longmapsto & A\,v \qquad , \end{array}$$

The matrix A is unitary, for the standard inner product on $\mathbb{C}^n$, if and only if,

$$\begin{array}{rcl} \langle Av, Aw \rangle & = & (Av)^T\,\overline{Av} \\ & = & v^T\,A^T\,\overline{A}\,\overline{w} \\ & = & v^T\,\overline{w} = \langle v, w \rangle\,, \qquad \forall v, w \in \mathbb{C}^n \\ \Leftrightarrow A^T\overline{A} & = & I_n \Leftrightarrow A^\dagger\,A = I_n \\ \Leftrightarrow \langle \mathrm{Col}_j(A), \mathrm{Col}_k(A) \rangle & = & \delta_{jk}\,, \qquad 1 \leq j, k \leq n\ , \end{array}$$

where $\mathrm{Col}_k(A)$ denotes the kth column of the matrix A and

$$\delta_{jk} = \left\{ \begin{array}{ll} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{array} \right. \qquad .$$

Thus A is unitary if and only if its columns form an orthonormal basis of $\mathbb{C}^n$. The set of all unitary $n \times n$ matrices is a subgroup of $GL_n(\mathbb{C})$ denoted by,

$$U_n = \left\{ A \in GL_n(\mathbb{C})\ :\ A^\dagger\,A = I_n \right\}\ .$$

A normal subgroup of U_n is the group of special unitary matrices,

$$SU_n = \left\{ A \in GL_n(\mathbb{C})\ :\ A^\dagger\,A = I_n\,,\ \det(A) = 1 \right\}\ .$$

The fact that SU_n is a normal subgroup of U_n follows from the fact that it coincides with the kernel of the homomorphism,

$$\begin{array}{rcl} \det\ :\ U_n & \longrightarrow & S^1 \subset \mathbb{C}^* \\ A & \longmapsto & \det(A) \qquad . \end{array}$$

For analogous reasons $SL_n(\mathbb{K})$ and $SO_n(\mathbb{K})$ are normal subgroups of $GL_n(\mathbb{K})$ and $O_n(\mathbb{K})$, respectively, for any field $\mathbb{K}$.

We are now ready to introduce representations of groups as special actions on vector spaces.

Definition 2.2.60 (representation) *Let G be a group and V be a vector space.*

(a) A representation of G is a linear action of G on V, i.e. an action $G \overset{\varphi}{\curvearrowright} V$, such that

$$\mathrm{Im}(\varphi) \subset GL(V) \subset \mathrm{Sym}(V)\,,$$

or, equivalently, is an homomorphism (compare with the definition 2.2.33),

$$\varphi\ :\ G \longrightarrow GL(V)\,.$$

The dimension of V is called degree of the representation,

$$\deg(\varphi) = \dim(V)\,.$$

(b) Let V be a finite dimensional inner product space. A representation, $G \overset{\varphi}{\curvearrowright} V$, is called unitary if φ_g are unitary operators, for every $g \in G$.

Remark 2.2.61 If $\dim(V) = \infty$, a representation $G \overset{\varphi}{\curvearrowright} V$ is called unitary if V is an Hilbert space, i.e. an inner product space, complete (see next chapter) with respect to the norm defined by the inner product and the linear transformations φ_g, besides preserving the inner product, are required to be surjective (and thus invertible) for every $g \in G$. $\Diamond$

Definition 2.2.62 (intertwining maps and equivalent representations) *Let $G \overset{\varphi}{\curvearrowright} V$ and $G \overset{\varphi'}{\curvearrowright} V'$ be two representations of G.*

(i) A map, $T : V \longrightarrow V'$, is called an intertwining map between the two G–representations if it is a G–equivariant linear map (see definition 2.2.44) i.e. is linear and such that,

$$T \circ \varphi_g = \varphi'_g \circ T \,, \qquad \forall\, g \in G \,.$$

(ii) If $T : V \longrightarrow V'$ is an intertwining map and a vector space isomorphism (i.e. is bijective) then T is called an equivalence of representations. In that case we have,

$$\varphi'_g = T \circ \varphi_g \circ T^{-1} \,, \qquad \forall\, g \in G \,.$$

Definition 2.2.63 (invariant subspace) *Let $G \overset{\varphi}{\curvearrowright} V$ be a representation of the group G. A subspace $W \subset V$ is called (G–)invariant if*

$$G_W = G \;\Leftrightarrow\; \varphi_g(W) \subset W \,, \quad \forall g \in G.$$

Remark 2.2.64 If a G–invariant subspace W is finite dimensional then $\varphi_g(W) \subset W \;\Leftrightarrow\; \varphi_g(W) = W$. $\Diamond$

Exercise 2.2.22 *Let W be a G–invariant subspace of a unitary representation φ of G. Show that the orthogonal complement $W^\perp$,*

$$W^\perp = \{v \in V \,:\, \langle v, w\rangle = 0 \;, \quad \forall\, w \in W\} \;,$$

is also G–invariant.

Solution. Let $v \in W^\perp$. Then, for every $w \in W$, we have

$$\begin{aligned} \langle \varphi_g v, w\rangle &= \langle \varphi_{g^{-1}}\, \varphi_g v, \varphi_{g^{-1}} w\rangle \\ &= \langle v, \varphi_{g^{-1}} w\rangle = 0 \quad , \forall\, g \in G \\ &\Leftrightarrow\; W^\perp \text{ is } G\text{–invariant.} \end{aligned}$$

■

Definition 2.2.65 (direct sum) *Let $G \overset{\varphi}{\curvearrowright} V$, $G \overset{\varphi'}{\curvearrowright} V'$ be two representations. Define $\varphi \oplus \varphi'$ as the following representation of G on $V \oplus V'$.*

$$\left(\varphi \oplus \varphi'\right)_g (v, v') = \left(\varphi_g(v), \varphi'_g(v')\right) \,, \qquad g \in G \,.$$

Remark 2.2.66 Consider the representation $G \overset{\varphi}{\curvearrowright} V$ and let $B = (v_1, \ldots, v_n)$ be an ordered basis on V. Then, with the isomorphism $T_B : V \longrightarrow \mathbb{C}^n$ of (2.2.26), we define an equivalent matrix representation of G,

$$\begin{array}{rcl} T_B : V & \longrightarrow & \mathbb{C}^n \\ \widetilde{\varphi} : G & \longrightarrow & GL_n(\mathbb{C}) \\ \widetilde{\varphi}_g & = & T_B \circ \varphi_g \circ T_B^{-1} \,, \qquad g \in G \,. \end{array}$$

Given ordered bases B, B',

$$B = (v_1, \ldots, v_n) \ , \ B' = (v'_1, \ldots, v'_{n'})$$

of V, V' and two representations, $G \overset{\varphi}{\curvearrowright} V$ and $G \overset{\varphi'}{\curvearrowright} V'$, consider the ordered basis

$$B \oplus B' := ((v_1, 0), \ldots, (v_n, 0), (0, v'_1), \ldots, (0, v'_{n'}))$$

of $V \oplus V'$. Then, by considering the isomorphism,

$$T_{B \oplus B'} : V \oplus V' \longrightarrow \mathbb{C}^n \oplus \mathbb{C}^{n'} \,,$$

we get an equivalent matrix representation, with matrices in a block diagonal form,

$$(\varphi \oplus \varphi')_g \longmapsto \widetilde{(\varphi \oplus \varphi')}_g = \left(\begin{array}{cc} \widetilde{\varphi}_g & 0 \\ 0 & \widetilde{\varphi}'_g \end{array} \right) .$$

$\Diamond$

Example 2.2.67 *For S_n there is a natural representation on $\mathbb{C}^n$ by permuting the canonical basis vectors, that we call the* <u>*standard*</u> *representation*

$$\varphi^{\mathrm{st}}_\sigma(e_j) = e_{\sigma(j)} \,, \quad 1 \leq j \leq n \,.$$

For S_3 we find,

$$\begin{array}{rcl} \varphi^{\mathrm{st}}_{(12)} & = & \left(\begin{array}{ccc} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right) \ , \ \varphi^{\mathrm{st}}_{(23)} = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array} \right) \ , \ \varphi^{\mathrm{st}}_{(13)} = \left(\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{array} \right) \\ \varphi^{\mathrm{st}}_{(123)} & = & \left(\begin{array}{ccc} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right) \ , \ \varphi^{\mathrm{st}}_{(132)} = \left(\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{array} \right) . \end{array}$$

Exercise 2.2.23 (from [St]) *Verify that the following matrices*

$$\rho_{s_1} = \left(\begin{array}{cc} -1 & -1 \\ 0 & 1 \end{array} \right) \ , \ \rho_{(123)} = \left(\begin{array}{cc} -1 & -1 \\ 1 & 0 \end{array} \right) ,$$

where s_1 is the simple transposition, $s_1 = (12)$, define a (degree two, nonunitary) representation of S_3 on $\mathbb{C}^2$.

Solution. From Theorem 2.2.47 we just have to check that ρ_{s_1} and ρ_{s_2} have order two and $\rho_{s_1 s_2} = \rho_{(123)}$ has order three. We have that $s_2 = s_1\,(123)$ and therefore

$$\rho_{s_2} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

We check now the relations given by Theorem 2.2.47.

$$\begin{aligned} \rho_{s_1}^2 &= \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} = I_2 \\ \rho_{s_2}^2 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = I_2 \\ \rho_{s_1 s_2}^3 &= \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} = I_2 \ . \end{aligned}$$

■

Definition 2.2.68 (irreducible representation) *A representation $G \overset{\varphi}{\curvearrowright} V$ is called irreducible if the only invariant subspaces are $\{0\}$ and V.*

Exercise 2.2.24 *Verify that ρ in the exercise 2.2.23, $\rho : S_3 \longrightarrow GL_2(\mathbb{C})$, is irreducible.*

Solution. Suppose that ρ is not irreducible. Since the only nontrivial (i.e. different from $\{0\}$ and V) invariant subspaces are of dimension one, this means that there is a vector $v_0 \in \mathbb{C}^2 \setminus \{0\}$ such that the line it generates, span$\{v_0\}$, is an S_3–invariant subspace. Then, for every $\sigma \in S_3$ there exists $\alpha_\sigma \in \mathbb{C}$, such that,

$$\rho_\sigma(v_0) = \alpha_\sigma\, v_0 \, ,$$

i.e. v_0 is a joint eigenvector of ρ_σ for every $\sigma \in S_3$. Let us find the eigenvalues and eigenvectors of ρ_{s_1}.

$$\det\left(\rho_{s_1} - \lambda\, I_2\right) = \begin{vmatrix} -1-\lambda & -1 \\ 0 & 1-\lambda \end{vmatrix} = -(1+\lambda)(1-\lambda) \, .$$

We see from the matrix ρ_{s_1} that e_1 is eigenvector with eigenvalue -1. For the eigenspace with eigenvalue 1 we find

$$\ker\left(\rho_{s_1} - I_2\right) = \ker \begin{pmatrix} -2 & -1 \\ 0 & 0 \end{pmatrix} = \text{span}\{(1,-2)\} \, .$$

It is now easy to verify that neither $e_1 = (1,0)$ nor $(1,-2)$ are eigenvectors of ρ_{s_2} and therefore there is not a one–dimensional invariant subspace of $\mathbb{C}^2$. ■

We see from the solution of the exercise 2.2.24 that the following result is true.

Proposition 2.2.69 *Let $G \overset{\varphi}{\curvearrowright} V$ be a representation of degree 2. Then φ is irreducible if and only if there is not a common eigenvector of φ_g, for every $g \in G$.*

The following result is very important.

Proposition 2.2.70 (see [St] Proposition 3.2.4) *Every finite–dimensional representation of a finite group is equivalent to a unitary representation.*

As a consequence of exercise 2.2.22 and proposition 2.2.70 one obtains the following crucial result.

Theorem 2.2.71 (Maschke on complete reducibility) *Every representation, $G \overset{\varphi}{\curvearrowright} V$, of a finite group is completely reducible, i.e. it is equivalent to a direct sum of irreducible representations.*

Proof. From proposition 2.2.70 we know that every finite–dimensional representation is equivalent to to a unitary representation. Let us assume then that φ is unitary. From exercise 2.2.22, we know that if W is a G–invariant subspace of V then its orthogonal complement $W^\perp$ is also invariant and $V = W \oplus W^\perp$. We can then complete the argument by (strong) induction on the dimension of V. Namely, if $\dim V = 1$, then φ is irreducible (and thus completely reducible also). Let us now assume that every m–dimensional representation with $m \leq n$ is completely reducible and prove that the same is true for representations of dimension $n+1$. Let $\dim V = n+1$ and let us assume that it is not completely reducible. In particular φ is not irreducible so that it has a G–invariant subspace W, with $0 < \dim W \leq n$. But then $W^\perp$ is also G–invariant and $0 < \dim(W^\perp) \leq n$ so that, the induction hypothesis, implies that both W and $W^\perp$ are completely reducible and therefore so is $V = W \oplus W^\perp$. ■

Let us denote the set of all intertwining operators, $T : G \overset{\varphi}{\curvearrowright} V \longrightarrow G \overset{\psi}{\curvearrowright} W$ by $\mathrm{Hom}_G(\varphi, \psi) \subset \mathrm{Hom}(V, W)$. We are now ready for the crucial Schur's Lemma.

Lemma 2.2.72 (Schur's Lemma) *Let φ, ψ be two irreducible representations of G on V and W and $T \in \mathrm{Hom}_G(\varphi, \psi)$. Then either $T = 0$ or T is invertible and*

(a) If $\varphi \nsim \psi$ then $\mathrm{Hom}_G(\varphi, \psi) = 0$.

(b) If $\varphi = \psi$ then $\exists\, \lambda \in \mathbb{C}$ such that $T = \lambda \,\mathrm{Id}_V$, or, equivalently,

$$\mathrm{Hom}_G(\varphi, \varphi) = \{\lambda \,\mathrm{Id}_V \,, \lambda \in \mathbb{C}\} \cong \mathbb{C}\,.$$

Proof.

Let $T \in \mathrm{Hom}_G(\varphi, \psi)$. If $T = 0$ we are done. Suppose now that $T \neq 0$ and let us prove that T is invertible. We show first the triviality of the kernel of T by showing that it is invariant. Let $v \in \ker(T)$.

$$T(\varphi_g(v)) = \psi_g(T(v)) = 0 \Leftrightarrow \varphi_g(v) \in \ker(T),\ \forall v \in \ker(T),\ g \in G.$$

So $\ker T$ is invariant and therefore $\ker T = 0$ or $\ker T = V$. But since $T \neq 0$ then $\ker(T) = 0$ and T is injective. Let us now show the surjectivity of T. Let $w \in Im(T)$, i.e. exists $v \in V$ such that $w = T(v)$. Then,

$$\psi_g(w) = \psi_g(T(v)) = T(\varphi_g(v)) \in Im(T) \qquad \forall g \in G.$$

So $Im(T)$ is G–invariant and thus either $Im(T) = 0$ or $Im(T) = W$. But $Im(T)$ can not be zero since by supposition $T \neq 0$ and therefore $Im(T) = W$ and T is surjective.

(a) This part is obvious.

(b) Let now $\psi = \varphi$ and $T \in \mathrm{Hom}_G(\varphi, \varphi) \subset \mathrm{Hom}(V, V)$ and λ_0 be an eigenvalue of T. Then,

$$T - \lambda_0 \,\mathrm{Id}_V \;\in\; \mathrm{Hom}_G(\varphi, \varphi)\,,$$

and $T - \lambda_0 \,\mathrm{Id}_V$ can not be invertible by definition of eigenvalue. It has then to be equal to zero and therefore,

$$T = \lambda_0 \,\mathrm{Id}_V\,.$$

■

Corollary 2.2.73 *Let G be an abelian group. Then every irreducible representation of G has degree one.*

Proof. Let $\varphi : G \longrightarrow GL(V)$ be an irreducible representation and let $g_0 \in G$. Since G is abelian,

$$\varphi_{g_0} \circ \varphi_g = \varphi_{g_0\,g} = \varphi_{g\,g_0} = \varphi_g \circ \varphi_{g_0}\,,$$

so that $\varphi_{g_0} \in \mathrm{Hom}_G(\varphi, \varphi)$. But then from Schur's Lemma we conclude that, for every $g_0 \in G$, there exists $\alpha(g_0) \in \mathbb{C}^*$ such that,

$$\varphi_{g_0} = \alpha(g_0)\,\mathrm{Id}_V\,.$$

Then every subspace of V is a G–invariant subspace, which, since φ was assumed to be irreducible, implies that $\dim V = 1$.

■

Remark 2.2.74 If $\dim V = 1$, then $V = \mathrm{span}\{v_0\}$, for any $v_0 \neq 0$ and $\mathrm{Hom}(V, V)$ is canonically isomorphic to $\mathbb{C}$. Indeed, if $T \in \mathrm{Hom}(V, V)$, then

$$T(v_0) = \lambda_T\, v_0$$

and the canonical isomorphism makes correspond to T its eigenvalue.

$$\begin{array}{rcl} \mathrm{Hom}(V, V) & \longrightarrow & \mathbb{C} \\ T & \longmapsto & \lambda_T\,. \end{array}$$

◊

Corollary 2.2.75 *If G is a finite abelian group any representation of G is equivalent to a direct sum of 1–dimensional representations.*

Proposition 2.2.76 *Let $G \overset{\varphi}{\curvearrowright} V$ be a representation of G. Then, for every $g \in G$, φ_g is diagonalizable with $d = ord(g)$–roots of unity as eigenvalues.*

Proof. Since $g^d = e$ we conclude that

$$(\varphi_g)^d = \mathrm{Id}_V \,,$$

We have that the subgroup of G generated by g is

$$\langle g\rangle \cong \mathbb{Z}_d \,,$$

and therefore,

$$\begin{array}{rcl} \widehat{\varphi} : \mathbb{Z}_d & \longrightarrow & G \longrightarrow GL(V) \\ {[m]} & \longmapsto & g^m \longmapsto (\varphi_g)^m \,, \end{array}$$

is a representation of the abelian group $\mathbb{Z}_d$, so that it is equivalent to the direct sum of 1–dimensional representations, corresponding to the eigenspaces of φ_g,

$$V = V_1 \oplus \cdots \oplus V_n \,, \tag{2.2.30}$$

with $\dim(V_j) = 1$.

In a basis B adapted to (2.2.30),

$$T_B \circ \widehat{\varphi}_{[1]} \circ T_B^{-1} = T_B \circ \varphi_g \circ T_B^{-1} = \begin{pmatrix} \alpha_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \alpha_n \end{pmatrix} \,,$$

with $\alpha_j^d = 1$ so that

$$(\varphi_g)^d = \mathrm{Id}_V \,.$$

■

Definition 2.2.77 (regular representation) *Let G be finite and*

$$L(G) = \mathbb{C}^G = \mathrm{Map}(G, \mathbb{C}) \,.$$

The regular representation, $G \overset{\varphi^r}{\curvearrowright} L(G)$, of G is defined by

$$\left(\varphi^r_g f\right)(h) = f(g^{-1}h) \,, \qquad \forall g, h \in G.$$

The inner product on $L(G)$ is the inner product (2.2.27) divided by the order of the group, $|G|$,

$$\langle f_1, f_2\rangle = \frac{1}{|G|} \sum_{g\in G} f_1(g)\,\overline{f_2(g)} \,. \tag{2.2.31}$$

Proposition 2.2.78 *The regular representation, $G \overset{\varphi^r}{\curvearrowright} L(G)$, is unitary.*

Before proving the proposition let us prove a lemma.

Lemma 2.2.79 *Let $(V, \langle\ ,\ \rangle)$ be an inner product space. The linear transformation, $U : V \longrightarrow V$, is unitary if and only if maps an orthonormal basis to another orthonormal basis.*

Proof. Let $B = (u_1, \dots, u_n)$ be an orthonormal basis of V and $U(B) = (U(u_1), \dots, U(u_n))$ its image under U. If U is unitary it is obvious that $U(B)$ is an orthonormal basis. Let us now prove that if the basis $U(B)$ is orthonormal, then U is unitary. Indeed, for $v = \sum_{j=1}^{n} a_j\, u_j$ and $w = \sum_{j=1}^{n} b_j\, u_j$ arbitrary vectors, we find

$$\langle U(v), U(w)\rangle = \sum_{j,k=1}^{n} a_j \bar{b}_k \,\langle U(u_j), U(u_k)\rangle \; = \sum_{j=1}^{n} a_j \bar{b}_j = \langle v, w\rangle \,.$$

■

Let us now prove proposition 2.2.78.

Proof. Let us show that, for every $h \in G$, φ^r_h acts on the canonical orthonormal basis of $L(G)$,

$$B_{\mathrm{can}}(L(G)) = \left\{ \sqrt{|G|}\; \delta_g \;,\; g \in G \right\} \,,$$

by just permuting its elements,

$$\varphi^r_h(\delta_g) = \delta_{hg} \;, \qquad h, g \in G \,, \tag{2.2.32}$$

where, for simplicity, we have omitted the common factor $\sqrt{|G|}$. Indeed,

$$\begin{aligned} \varphi^r_h(\delta_g)(g_1) = \delta_g(h^{-1}g_1) \quad &= \quad \begin{cases} 1 & \text{if } h^{-1}g_1 = g \\ 0 & \text{if } h^{-1}g_1 \neq g \end{cases} \\ &= \quad \begin{cases} 1 & \text{if } g_1 = hg \\ 0 & \text{if } g_1 \neq hg \end{cases} \\ &= \quad \delta_{hg}(g_1) \,, \qquad \forall h, g, g_1 \in G \,. \end{aligned}$$

Since, for every $h \in G$, φ^r_h maps the canonical orthonormal basis to itself it is unitary. ■

Let φ be a matrix representation of $G \overset{\varphi}{\curvearrowright} \mathbb{C}^n$, $\varphi : G \longrightarrow GL_n(\mathbb{C})$. Then the entries of the representation matrices define functions on $L(G)$,

$$\varphi_{jk}(g) := (\varphi_g)_{jk} \,.$$

In fact, as we will see now, if the representation is irreducible and unitary of dimension n, these functions are all orthogonal to each other and thus span a n^2–dimensional subspace of $L(G)$.

Theorem 2.2.80 (Schur's orthogonality relations) *(See [St], Theorem 4.2.8)*

Let $\varphi : G \longrightarrow U_n$, $\rho : G \longrightarrow U_m$, be any two inequivalent irreducible unitary representations of G. Then,

1. *Matrix entries of inequivalent representations are orthogonal, i.e.*

$$\langle \varphi_{ij}, \rho_{kl}\rangle = 0 \;, \quad \forall\; i.j, k, l \;.$$

2. *Matrix entries of an unitary irreducible representation form an orthogonal system in $L(G)$,*

$$\langle \varphi_{ij}, \varphi_{kl}\rangle = \begin{cases} \frac{1}{n} & \text{if } i = k, j = l \\ 0 & \text{else}\,. \end{cases} \tag{2.2.33}$$

Corollary 2.2.81 *Exist finitely many equivalence classes of irreducible representations,*

$$\begin{aligned} \widehat{G} &= \{\textit{irreducible representations of } G\} / \sim \\ &= \{[\varphi^{(1)}], \ldots, [\varphi^{(s)}]\} \ . \end{aligned}$$

satisfying the inequalities,

$$s \leq d_1^2 + \cdots + d_s^2 \leq |G| \, , \tag{2.2.34}$$

where $d_j = \deg(\varphi^{(j)}) = \dim(V_j)$*.*

Proof. The inequalities (2.2.33) are a direct consequence of the Schur's orthogonality relations as for every equivalence class $[\varphi^{(j)}]$ of irreducible representations we obtain subspaces of $L(G)$ of dimension d_j^2. On the other hand the left inequality is due to the fact that all classes $[\varphi^{(j)}]$ define subspaces of $L(G)$, which are orthogonal for inequivalent representations. ■

Remark 2.2.82 From Theorem 2.2.80 and Corollary 2.2.73 it follows that equality in the left inequality in (2.2.34) takes place for abelian groups. In fact only for them since a nonabelian group has at least one irreducible representation of dimension bigger than 1.

On the other hand, we will show below in Corollary 2.2.93 that the right inequality in (2.2.34) is in fact always an equality.

$\Diamond$

Characters and class functions

Definition 2.2.83 *Let* $G \overset{\varphi}{\curvearrowright} V$ *be a representation. The character* χ_φ *of* φ *is the following function* $\chi_\varphi \in L(G) = \mathbb{C}^G$*,*

$$\chi_\varphi(g) = \operatorname{tr}(\varphi_g) = \sum_{j=1}^{d} (\varphi_g)_{jj} = \sum_{j=1}^{d} \varphi_{jj}(g) \ .$$

Remark 2.2.84 In accordance with remark 2.2.74, if $\deg(\varphi) = \dim(V) = 1$, then $\varphi : G \longrightarrow \mathbb{C}^*$ (in fact it takes values in S^1) and,

$$\chi_\varphi = \varphi \, .$$

In particular, for degree one representations, χ_φ is an homomorphism.

$\Diamond$

Remark 2.2.85 Let $G \overset{\varphi}{\curvearrowright} V$ be a representation of G. At the identity e we have,

$$\chi_\varphi(e) = \operatorname{tr}(\varphi_e) = \operatorname{tr}(\mathrm{Id}_V) = \dim V = \deg(\varphi) \ .$$

$\Diamond$

Proposition 2.2.86 *If* φ *and* ψ *are equivalent representations their characters are equal,*

$$\chi_\varphi = \chi_\psi \ .$$

Proof. Since they are equivalent there is an invertible linear transformation T such that,

$$\psi_g = T \circ \varphi_g \circ T^{-1}\,, \qquad g \in G.$$

Then,

$$\begin{aligned} \chi_\psi(g) &= \operatorname{tr}(\psi_g) = \operatorname{tr}\left(T \circ \varphi_g \circ T^{-1}\right) = \operatorname{tr}\left(T^{-1} \circ T \circ \varphi_g\right) \\ &= \operatorname{tr}(\varphi_g) = \chi_\varphi(g)\ , \quad \forall\, g \in G. \end{aligned}$$

■

Proposition 2.2.87 *Let $G \overset{\varphi}{\curvearrowright} V$ be a representation. Its character, χ_φ, is constant on conjugacy classes, i.e.,*

$$\chi_\varphi(hgh^{-1}) = \chi_\varphi(g)\ , \qquad \forall g, h \in G\ .$$

Proof. The proof is very similar to the proof of the previous proposition as it also uses the cyclic property of trace.

$$\begin{aligned} \chi_\varphi(hgh^{-1}) &= \operatorname{tr}\left(\varphi_{hgh^{-1}}\right) = \operatorname{tr}\left(\varphi_h\, \varphi_g\, \varphi_{h^{-1}}\right) \\ &= \operatorname{tr}\left(\varphi_h^{-1}\, \varphi_h\, \varphi_g\right) = \operatorname{tr}(\varphi_h) = \chi_\varphi(h)\ , \quad \forall \varphi \in \widehat{G}\,,\ g, h \in G\ . \end{aligned}$$

■

Definition 2.2.88 *A function $f \in L(G)$ is called a class function if it is constant on conjugacy classes,*

$$f(g) = f(hgh^{-1})\ , \qquad \forall g, h \in G\,.$$

The vector subspace of $L(G)$ of all class functions on G is denoted by $Z(L(G))$.

The vector space $Z(L(G))$ of class functions on G is naturally isomorphic to the space of functions on the set of conjugacy classes of G, $Cl(G) = G/\varphi^c_G$, with isomorphism given by,

$$\begin{aligned} Z(L(G)) &\longrightarrow \mathbb{C}^{Cl(G)} \\ f &\longmapsto \widetilde{f} \\ \widetilde{f}([g]) &= f(g)\,. \end{aligned}$$

In particular, as in exercise 2.2.20, we have,

$$\dim(Z(L(G))) = |Cl(G)|\,. \tag{2.2.35}$$

Theorem 2.2.89 *Let φ and ρ be irreducible representations of G. Then,*

$$\langle \chi_\varphi, \chi_\rho \rangle = \begin{cases} 1 & \text{if } \rho \sim \varphi \\ 0 & \text{if } \rho \nsim \varphi\ . \end{cases}$$

Proof. From propositions 2.2.70 and 2.2.86 we can assume that both representations φ and ρ are unitary matrix representations, $\varphi : G \longrightarrow U_n$ and $\rho : G \longrightarrow U_m$,

$$\begin{aligned}
\langle \chi_\varphi, \chi_\rho \rangle &= \frac{1}{|G|} \sum_{g\in G} \chi_\varphi(g)\, \overline{\chi_\rho(g)} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{m} \frac{1}{|G|} \sum_{g\in G} \varphi_{jj}(g)\, \overline{\rho_{kk}(g)} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{m} \langle \varphi_{jj}, \rho_{kk} \rangle = \left\{ \begin{array}{ll} 0 & \text{if } \varphi \nsim \rho \\ 1 & \text{if } \varphi = \rho \ . \end{array} \right.
\end{aligned}$$

The general case $\varphi \sim \rho$ is already included in the proof since in that case, $\chi_\varphi = \chi_\rho$ (see proposition 2.2.86), and therefore,

$$\langle \chi_\varphi, \chi_\rho \rangle = \langle \chi_\varphi, \chi_\varphi \rangle = 1 \ .$$

■

Corollary 2.2.90 *The number of classes of inequivalent irreducible representations of G is smaller or equal than the number of conjugacy classes of G,*

$$|\widehat{G}| \leq |Cl(G)|\,.$$

Proof. The characters of irreducible representations are class functions so that, from Theorem 2.2.89, the set $\{\chi_\varphi\,,\, \varphi \in \widehat{G}\}$ is an orthonormal set in $Z(L(G))$ and therefore its cardinality ($= |\widehat{G}|$) is smaller or equal than $\dim(Z(L(G)))$ ($= |Cl(G)|$, see (2.2.35)).

■

Remark 2.2.91 We will see below that the inequality in the Corollary 2.2.90 is in fact always equality. ♢

Theorem 2.2.92 *Let the space of equivalence classes of irreducible representations of the group G be*

$$\widehat{G} = \left\{ [\varphi^{(1)}], \ldots, [\varphi^{(s)}] \right\} \,,$$

and ρ a representation of G,

$$\rho \sim m_1 \varphi^{(1)} \oplus \cdots \oplus m_s \varphi^{(s)} \,,$$

where $m\,\varphi$ denotes m copies of the representation φ,

$$m\,\varphi = \underbrace{\varphi \oplus \cdots \oplus \varphi}_{m} \ . \tag{2.2.36}$$

Then,

(a) The multiplicities m_j are given by,

$$m_j = \langle \chi_\rho, \chi_{\varphi^{(j)}} \rangle \,. \tag{2.2.37}$$

(b) The norm square of the character of ρ is equal to the sums of squares of the multiplicities,

$$||\chi_\rho||^2 = \langle \chi_\rho, \chi_\rho \rangle = m_1^2 + \cdots + m_s^2 \,. \tag{2.2.38}$$

Proof.

(a) From (2.2.36) we see that,

$$\chi_{m\varphi} = m\, \chi_\varphi \,,$$

and therefore,

$$\chi_\rho = m_1\, \chi_{\varphi^{(1)}} + \cdots + m_s\, \chi_{\varphi^{(s)}} \,. \tag{2.2.39}$$

By taking the inner product of both sides of (2.2.39) with $\chi_{\varphi^{(j)}}$ we obtain (2.2.37).

(b) The equality (2.2.38) is obtained by taking the norm square of both sides in (2.2.39).

■

Corollary 2.2.93 *Let G be a finite group and*

$$\widehat{G} = \left\{ [\varphi^{(1)}], \ldots, [\varphi^{(s)}] \right\} \,.$$

Then,

(a)

$$|G| = d_1^2 + \cdots + d_s^2 \,, \tag{2.2.40}$$

where

$$d_j = \deg(\varphi^{(j)}) \,.$$

(b) A representation ρ is irreducible if and only if

$$\langle \chi_\rho, \chi_\rho \rangle = 1 \,. \tag{2.2.41}$$

Proof.

(a) We know from Corollary 2.2.81 that, $d_1^2 + \cdots + d_s^2 \leq |G|$. Let us find the character of the regular representation, χ_{φ^r}, which will allow us to prove (2.2.40) easily. Let us order the elements in the group G to have the canonical basis of $L(G)$ also ordered,

$$\begin{aligned} G &= \left\{ g_1 = e, g_2, \ldots, g_{|G|} \right\} \\ B_{\text{can}}(L(G)) &= \left\{ \delta_{g_k} \,, \quad 1 \leq k \leq |G| \right\} \,. \end{aligned}$$

We have seen in (2.2.32) that,

$$\varphi_h^r(\delta_{g_k}) = \delta_{hg_k}.$$

But if $h \neq e$, then $\delta_{hg_k} \neq \delta_{g_k}$ for every $k = 1, \ldots, |G|$, which implies that the matrices of φ_h^r, in the canonical basis of $L(G)$, have one entry per column (outside the main diagonal) equal to 1 and all other entries equal to zero. Then, we have

$$\chi_{\varphi^r}(h) = \operatorname{tr}(\varphi_h^r) = \begin{cases} 0 & \text{if } h \neq e \\ |G| & \text{if } h = e \,. \end{cases}$$

We find,

$$\begin{aligned}\langle \chi_{\varphi^r}, \chi_{\varphi^{(j)}} \rangle &= \frac{1}{|G|} \sum_{g\in G} \chi_{\varphi^r}(g)\, \overline{\chi_{\varphi^{(j)}}(g)} \\ &= \frac{1}{|G|}\,(|G|\, d_j + 0) = d_j = \deg(\varphi^{(j)})\,, \qquad 1 \le j \le s = |\widehat{G}|\,.\end{aligned}$$

From Theorem 2.2.92 we now conclude that

$$\varphi^r \sim d_1\, \varphi^{(1)} \oplus \cdots \oplus d_s\, \varphi^{(s)}\,,$$

and therefore,

$$||\chi_{\varphi^r}||^2 = \frac{1}{|G|}\,|G|^2 = |G| = d_1^2 + \cdots + d_s^2\,.$$

(b) Follows immediately from (2.2.38).

■

Exercise 2.2.25 *Complete the character table of S_3, knowing that $|Cl(S_3)| = 3 \ge |\widehat{G}|$ and that the two 1–dimensional (thus irreducible) representations (see Proposition 2.2.51 and remark 2.2.84) are given by*

$$\begin{aligned}\chi_{\varphi^{(1)}}(\sigma) &= \varphi^{(1)}_\sigma = 1\,, \qquad \sigma \in S_3 \\ \chi_{\varphi^{(2)}}(\sigma) &= \varphi^{(2)}_\sigma = \mathrm{sgn}(\sigma) = \begin{cases} 1 & \text{if } \sigma \in \{e, (123), (132)\} \\ -1 & \text{else}\,, \end{cases}\end{aligned}$$

but without using the representation ρ in the exercise 2.2.23.

Solution. Let $d_j = \deg(\varphi^{(j)})$. We have that $d_1 = d_2 = 1$ so that there has to be at least one more irreducible representation of S_3 in order to satisfy (2.2.40). Since, on the other hand, $|\widehat{G}|$ is bound above by the number of conjugacy classes ($= 3$) we must have $|\widehat{G}| = 3$ and the degree d_3 of $\varphi^{(3)}$ is found from,

$$6 = |S_3| = 1 + 1 + d_3^2 \Leftrightarrow d_3 = 2\,.$$

We have then, for the character table of S_3,

Table 2.5

	e	$C(12)$	$C(123)$
$\chi_{\varphi^{(1)}}$	1	1	1
$\chi_{\varphi^{(2)}}$	1	-1	1
$\chi_{\varphi^{(3)}}$	$d_3 = 2$	a_2	a_3

We find a_2 and a_3 from the fact that the characters form an orthonormal system (see Theorem 2.2.89).

$$\begin{aligned}\langle \chi_{\varphi^{(3)}}, \chi_{\varphi^{(1)}} \rangle &= \frac{1}{6}\,(2 + 3a_2 + 2a_3) = 0 \\ \langle \chi_{\varphi^{(3)}}, \chi_{\varphi^{(2)}} \rangle &= \frac{1}{6}\,(2 - 3a_2 + 2a_3) = 0\,,\end{aligned}$$

which gives, $a_2 = 0$ and $a_3 = -1$, so that the complete character table of S_3 reads,

Table 2.6: Character table of S_3

	e	$C(12)$	$C(123)$
$\chi_{\varphi^{(1)}}$	1	1	1
$\chi_{\varphi^{(2)}}$	1	-1	1
$\chi_{\varphi^{(3)}}$	2	0	-1

■

The last line of the table 2.6 is of course in agreement with the exercise 2.2.23 and $\rho \sim \varphi^{(3)}$. Notice however that ρ is not unitary for the standard inner product on $\mathbb{C}^2$.

Exercise 2.2.26 *Consider the standard representation of S_3 (see example 2.2.67) and, in the notation of the exercise 2.2.4, the decomposition,*

$$\varphi^{\text{st}} \sim m_1\,\varphi^{(1)} \oplus m_2\,\varphi^{(2)} \oplus m_3\,\varphi^{(3)}\ .$$

Find $m_j\,,\ j = 1, 2, 3$.

Solution. From the example 2.2.67 we find the character of φ^{st},

Table 2.7: Character of φ^{st}

	e	$C(12)$	$C(123)$
$\chi_{\varphi^{\text{st}}}$	3	1	0

We now use Theorem 2.2.92 to obtain the multiplicities of the irreducible representations in the standard representation,

$$\begin{aligned} m_1 &= \langle \chi^{\text{st}}, \chi_{\varphi^{(1)}} \rangle = \frac{1}{6}\,(3 \times 1 \times 1 + 1 \times 1 \times 3 + 0 \times 1 \times 2) = 1 \\ m_2 &= \langle \chi^{\text{st}}, \chi_{\varphi^{(2)}} \rangle = \frac{1}{6}\,(3 - 3 + 0) = 0 \\ m_3 &= \langle \chi^{\text{st}}, \chi_{\varphi^{(3)}} \rangle = \frac{1}{6}\,(6 + 0 + 0) = 1\,. \end{aligned}$$

We then have,

$$\varphi^{\text{st}} \sim \varphi^{(1)} \oplus \varphi^{(3)}\ .$$

■

Let us now consider the following extension of a representation of a group G, $\varphi : G \longrightarrow GL(V)$, to the following linear map from $L(G)$ to $\text{Hom}(V, V)$,

$$L(G) \ni f \longmapsto \varphi_f := \sum_{g \in G} f(g)\,\varphi_g \in \text{Hom}(V, V)\,. \qquad (2.2.42)$$

This map can be thought of as an extension of the domain of the homomorphism φ from G to $L(G)$ because for the functions δ_h, supported at $h \in G$, we have

$$\varphi_{\delta_h} = \varphi_h\,.$$

Indeed,

$$\varphi_{\delta_h} = \sum_{g \in G} \delta_h(g)\,\varphi_g = \varphi_h\ , \qquad \forall h \in G\,.$$

Remark 2.2.94 [group ring $(L(G),\cdot)$] By defining on $L(G)$ a composition that extends the natural product,

$$\delta_{g_1}\cdot\delta_{g_2}=\delta_{g_1 g_2}\,,$$

we obtain a ring that is called the group ring of G, with convolution product,

$$(f_1\cdot f_2)\,(g)=\sum_{h\in G} f_1(gh^{-1})\,f_2(h)\,.$$

The map in (2.2.42) defines an homomorphism of the group ring to $\mathrm{Hom}(V,V)$, i.e., one can verify that

$$\varphi_{f_1\cdot f_2}=\varphi_{f_1}\circ\varphi_{f_2}\,.$$

One can show that the center of the group ring (i.e. the functions commuting with all the others) coincides with the space (subring) of class functions $Z(L(G))$, which explains the notation we are using. For more details see the Section 5.2 of [St]. $\Diamond$

Lemma 2.2.95 *Let $G \overset{\varphi}{\curvearrowright} V$ be a representation and consider the map in (2.2.42).*

(a) If f is a class function then φ_f is an intertwining map or a G–morphism, $\varphi_f\in\mathrm{Hom}_G(\varphi,\varphi)$ (see definition 2.2.62 of intertwining maps).

(b) If f is a class function and φ is irreducible, then φ_f is proportional to the identity,

$$\varphi_f=\frac{|G|}{\deg(\varphi)}\,\langle f,\overline{\chi_\varphi}\rangle\,\mathrm{Id}_V\ . \tag{2.2.43}$$

Proof. Let f be a class function, $f\in Z(L(G))$.

(a)

$$\begin{aligned}\varphi_h\circ\varphi_f &= \varphi_h\circ\sum_{g\in G} f(g)\,\varphi_g=\sum_{g\in G} f(g)\,\varphi_{hg}\\ &\underset{k=hg}{=} \sum_{k\in G} f(h^{-1}k)\,\varphi_k \underset{k=\tilde g h}{=} \sum_{\tilde g\in G} f(h^{-1}\tilde g h)\,\delta_{\tilde g h}\\ &= \sum_{\tilde g\in G} f(\tilde g)\,\delta_{\tilde g}\circ\delta_h=\varphi_f\circ\varphi_h\ ,\qquad \forall h\in G\ ,\end{aligned}$$

so that φ_f commutes with the action of G and therefore is an intertwining map, $\varphi_f\in\mathrm{Hom}_G(\varphi,\varphi)$.

(b) If φ is irreducible than from Schur's Lemma we conclude that there exists a complex number, $c=c(f,\varphi)$, such that,

$$\varphi_f=c\,\mathrm{Id}_V\,.$$

To find c let us take trace of both sides. Then,

$$\begin{aligned}|G|\frac{1}{|G|}\sum_{g\in G} f(g)\,\chi_\varphi(g) &= c\,\deg(\varphi)\\ \Leftrightarrow c &= \frac{|G|}{\deg(\varphi)}\,\langle f,\overline{\chi_\varphi}\rangle\,,\end{aligned}$$

and we obtain (2.2.43).

■

Theorem 2.2.96 *The characters of G form an orthonormal basis of $Z(L(G))$.*

Remark 2.2.97 Notice that, in particular, this theorem implies that

$$|\widehat{G}| = |Cl(G)|\,.$$

$\Diamond$

Proof. Let us prove the theorem. We already know from Theorem 2.2.89 that the characters form an orthonormal system in $Z(L(G))$. We only have to show that the characters generate $Z(L(G))$ or, equivalently, that a class function orthogonal to all characters must be equal to zero. Suppose then that $f \in Z(L(G))$ and $\langle \chi_{\varphi^{(j)}}, f \rangle = 0\,,\ 1 \leq j \leq s = |\widehat{G}|$. This, together with (2.2.43), implies that

$$\big(\varphi^{(j)}\big)_{\overline{f}} = \sum_{g \in G} \overline{f(g)}\,\varphi^{(j)}_g = \frac{|G|}{\deg(\varphi^{(j)})}\,\langle \overline{f},\, \overline{\chi_{\varphi^{(j)}}} \rangle\,\mathrm{Id}_V = 0\,.$$

But if $\varphi_{\overline{f}} = 0$ for all irreducible representations then it is zero for all (not necessarily irreducible) representations. In particular, by considering the regular representation, we obtain

$$(\varphi^r)_{\overline{f}} = \sum_{g \in G} \overline{f(g)}\,\varphi^r_g = 0\,,$$

and therefore, from (2.2.32), we obtain

$$\varphi^r_{\overline{f}}(\delta_h) = \sum_{g \in G} \overline{f(g)}\,\varphi^r_g(\delta_h) = \sum_{g \in G} \overline{f(g)}\,\delta_{gh} = 0\,,$$

for every $h \in G$. In particular, by taking $h = e$, we conclude that (see (2.2.29))

$$\sum_{g \in G} \overline{f(g)}\,\delta_g = \overline{f} = 0\,,$$

and therefore $f = 0$. ■

Projection operators

Let us use the operators φ_f for class functions to construct projection operators into isotypical components of representations.

[TO BE CONTINUED]

2.2.6 Applications to Civil Engineering

Chapter 3

Topics of Geometry & Topology and Applications

The main references for this chapter are [GN, Is, Ne].

3.1 Topological and metric spaces

3.1.1 Introduction

The sets that are used to describe physical and engeneering systems are not always open subsets of $\mathbb{R}^n$ for which we have natural notions like distance between points, convergence of sequences, continuity, differentiability and integrability of functions (i.e. usual calculus) well defined. Yet, many of those sets do have natural generalizations of the above notions. In the present section we will study first metric spaces with a notion of distance between points well defined and then the more general notion of topological spaces based just on open sets.

3.1.2 Metric spaces

Definition 3.1.1 *Let M be a set. A distance function (or a metric) on M is a map*

$$d\,:\,M \times M \longrightarrow \mathbb{R}_+\ ,$$

satisfying the following properties,

1. *Symmetry:*

$$d(x,y) = d(y,x)\,, \qquad \forall x, y \in M\,.$$

2. *Triangle inequality:*

$$d(x,y) \leq d(x,z) + d(z,y)\ , \qquad \forall x, y, z \in M\,.$$

3. *Positive definiteness:*

$$d(x,y) \geq 0 \text{ and } d(x,y) = 0 \Leftrightarrow x = y\ , \qquad \forall x, y \in M\,.$$

A pair (M, d), where M is a set and d is a distance function on M is called a metric space.

Example 3.1.2 *An inner product space $(V, \langle\ ,\ \rangle)$ is a metric space with distance function given by,*

$$d(v_1, v_2) = ||v_1 - v_2|| = \sqrt{\langle v_1 - v_2, v_1 - v_2\rangle}\ .$$

Example 3.1.3 *Let $M \neq \emptyset$. Then*

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

defines a distance function called discrete distance function. The pair (M, d^{dis}) is called discrete metric space.

Example 3.1.4 *Consider the sphere $S^2 = S^2_1$ of radius 1 and center in the origin of $\mathbb{R}^3$. Given $u, v \in S^2, u \neq v$, intersect the sphere with the plane H containing u, v and the origin. Let*

$$d_{S^2}(u, v) = \mathrm{length}(C_{u,v})\,,$$

where $C_{u,v}$ denotes the shortest arc from u to v in $H \cap S^2$.

Example 3.1.5 *Consider the torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$. The following function*

$$\begin{aligned} d_T([(x_1, y_1)], [(x_2, y_2)]) &= \min_{(k,m)\in\mathbb{Z}^2} ||(x_1, y_1) - (x_2 + k, y_2 + m)|| \\ &= \min_{(k,m)\in\mathbb{Z}^2} \sqrt{(x_1 - (x_2 + k))^2 + (y_1 - (y_2 + m))^2}\,, \end{aligned}$$

is a distance function.

Example 3.1.6 (product distance) *If (M, d_M) and $(M', d_{M'})$ are metric spaces then a distance function on $M \times M'$ can be defined by*

$$d_{M\times M'}((x, x'), (y, y')) = d_M(x, y) + d_{M'}(x', y')\,.$$

Example 3.1.7 *If (M, d) is a metric space and $N \subset M$ is a subset, we can define a distance function on N by restricting the distance function on M,*

$$\begin{aligned} d_N &= d|_{N\times N} \\ d_N(x, y) &= d(x, y)\,, \qquad \forall x, y \in N\,. \end{aligned}$$

The metric d_N is called subspace metric.

Remark 3.1.8 Notice that for the radius one sphere S^2 in $\mathbb{R}^3$ the distance function d_{S^2} introduced in the example 3.1.4 does not coincide with the subspace metric,d'_{S^2} , induced on S^2 by the standard distance function on $\mathbb{R}^3$,

$$d_{\mathbb{R}^3}(u, v) = ||u - v|| = \sqrt{\langle u - v, u - v\rangle} = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}\,.$$

We have,

$$d'_{S^2}(u, v) = (d_{R^3})\,|_{S^2\times S^2}(u, v) < d_{S^2}(u, v)\,, \qquad \forall\, u, v \in S^2\,, u \neq v\,.$$

◊

Definition 3.1.9 (open metric ball) *Let (M, d) be a metric space. The open (metric) ball with radius $r > 0$ and center $x \in M$ is*

$$B_r(x) = \{y \in M \, : \, d(y, x) < r\} \, .$$

Exercise 3.1.1 *Find the open balls of the discrete metric space, (M, d^{dis}) (see example 3.1.3).*

Solution. From the definition of the discrete distance function we find that

$$B_r(x) = \begin{cases} \{x\} & \text{if } r \leq 1 \\ M & \text{if } r > 1 \, . \end{cases}$$

■

Definition 3.1.10 (open sets) *Let (M, d) be a metric space.*

(i) A subset $U \subset M$ is called open if for every $x \in U$, there exists a $r(x) > 0$ such that, $B_{r(x)}(x) \subset U$. The collection of all open sets is denoted by $\mathcal{T}_d$.

(ii) A subset $C \subset M$ is called closed if its complement, $C^c = M \setminus C$, is open.

Proposition 3.1.11 *[Ne, Proposition 2.1.2] Let (M, d) be a metric space and (N, d_N) a subset with subspace metric. The collection of open sets, $\mathcal{T}_d$, satisfies the following properties:*

1. *M and $\emptyset$ are open, i.e. $M, \emptyset \in \mathcal{T}_d$.*
2. *For any subcollection $\mathcal{U} \subset \mathcal{T}_d$,*
$$\bigcup_{U \in \mathcal{U}} U \in \mathcal{T}_d \, .$$
3. *For any finite subcollection, $\mathcal{U} = \{U_1, \ldots, U_m\} \subset \mathcal{T}_d$,*
$$\bigcap_{j=1}^{m} U_j \in \mathcal{T}_d \; ,$$
so that $\mathcal{T}_d$ is closed under arbitrary unions and finite intersections.

Exercise 3.1.2 *Give an example which shows that $\mathcal{T}_d$ is (in general) not closed under countable intersections.*

Solution. Consider the metric space $(\mathbb{R}, d_{\mathbb{R}})$ and the open sets,

$$U_n = (-\frac{1}{n}, \frac{1}{n}) \, , \quad n \in \mathbb{N} \, .$$

The intersection of these open sets is not open,

$$\bigcap_{n=1}^{\infty} (-\frac{1}{n}, \frac{1}{n}) = \{0\} \, .$$

■

Proposition 3.1.12 *[Ne, Proposition 2.1.4] Let (M, d) be a metric space and (N, d_N) a subspace with subspace metric. Then the open subsets of (N, d_N) are given by $U \cap N$, where U is open in M, i.e.,*

$$\mathcal{T}_{d_N} = \{U \cap N \, , \quad U \in \mathcal{T}_d\} \; .$$

3.1.3 Topological spaces

Distance functions in sets are very usefull. We have seen that they lead to an associated collection of open sets with which one can (as we will see) define important notions like continuity of maps, compactness and connectedness of metric spaces, etc. There are however sets with natural families of open sets which cannot come from a metric. This leads one naturally to try to use the proposition 3.1.11 to abstract the properties of open sets to the definition of topological spaces, which, as we will see are indeed more general than metric spaces.

Definition 3.1.13 (topology and topological space) *Let M be a set. A topology $\mathcal{T}$ on M is a collection of subsets of M, satisfying the following properties,*

1. *$M, \emptyset \in \mathcal{T}$.*

2. *For any subcollection $\mathcal{U} \subset \mathcal{T}$,*
$$\bigcup_{U \in \mathcal{U}} U \in \mathcal{T} .$$

3. *For any finite subcollection, $\mathcal{U} = \{U_1, \ldots, U_m\} \subset \mathcal{T}$,*
$$\bigcap_{j=1}^{m} U_j \in \mathcal{T} .$$

Sets $U \in \mathcal{T}$ are called open sets and their complements, $C = U^c = M \setminus U$, are called closed sets.

A pair $(M, \mathcal{T})$, where M is a set and $\mathcal{T}$ is a topology on M is called a topological space. A neighborhood of a point in a topological space is any open set containing the point.

Example 3.1.14 *A metric space is an example of a topological space with open sets given as defined in definition 3.1.10,*
$$\mathcal{T} = \mathcal{T}_d .$$
The topology $\mathcal{T}_d$ is called metric topology.

Remark 3.1.15 Not all topologies are metric topologies in the sense that, for a given topology $\mathcal{T}$ on a set M, a distance function d such that

$$\mathcal{T} = \mathcal{T}_d,$$

may not exist.

$\diamondsuit$

Exercise 3.1.3 *Consider the set, $M = \{a, b, c\}$ and the collection of subsets,*

$$\mathcal{T} = \{\emptyset, \{b\}, \{a, b\}, \{c, b\}, M\} .$$

Verify that $\mathcal{T}$ is a topology.

Definition 3.1.16 *A topology $\mathcal{T}$ on the set M is called Hausdorff if for every $x, y \in M$, $x \neq y$, there exist neighborhoods U, V of x, y separating these points i.e. such that $U \cap V = \emptyset$.*

A topological space $(M, \mathcal{T})$ is called (non)Hausdorff if its topology is (non)Hausdorff.

Example 3.1.17 *The topology in exercise 3.1.3 is non–Hausdorff as every neighborhood of a (and of c) contains b.*

Exercise 3.1.4 *Show that a metric topology is always Hausdorff.*

Solution. Let (M, d) be a metric space, $x, y \in M$ be arbitrary distinct points of M and $\delta = d(x, y) > 0$. Choose the following neighborhoods $U_x = B_{\delta/2}(x)$ of x and $U_y = B_{\delta/2}(y)$ of y. Suppose that $\exists\, z \in U_x \cap U_y$. But then,

$$\delta = d(x, y) \leq d(x, z) + d(z, y) < \frac{\delta}{2} + \frac{\delta}{2} = \delta \ .$$

This contradiction implies that

$$U_x \cap U_y = \emptyset\,,$$

and therefore $(M, \mathcal{T}_d)$ is an Hausdorff topological space. ■

Exercise 3.1.5 *If M is finite and d is a distance function show that $\mathcal{T}_d = \mathcal{T}^{\text{dis}} = 2^M$.*

Solution. Let,

$$0 < \delta = \min\{d(x, y),\, x, y \in M, x \neq y\} \ .$$

Then, all one point sets are open,

$$B_\delta(x) = \{y \in M \ : \ d(y, x) < \delta\} = \{x\} \in \mathcal{T}_d\,, \qquad \forall x \in M\,,$$

and therefore all subsets, $A \subset M$, are open,

$$A = \bigcup_{x \in A} \{x\} \in \mathcal{T}_d\,.$$

■

Proposition 3.1.18 *Let $(M, \mathcal{T})$ be a topological space and $\mathcal{V} \subset \mathcal{T}$. Then the following properties of the collection of open sets $\mathcal{V}$ are equivalent.*

P1. Every open set $U \in \mathcal{T}$ can be expressed as a union,

$$U = \bigcup_{\lambda \in \Lambda_U} V_\lambda\,,$$

over some subcollection of sets from $\mathcal{V}$,

$$\{V_\lambda,\ \lambda \in \Lambda_U\} \subset \mathcal{V}\,.$$

P2. For each $U \in \mathcal{T}$ and each $x \in U$ there exists $V \in \mathcal{V}$ such that

$$x \in V \subset U\,.$$

Proof. $\underline{P1 \implies P2}$:
Let $U \in \mathcal{T}$ and

$$U = \bigcup_{\lambda \in \Lambda_U} V_\lambda \, ,$$

be as in $P1$. Then for every $x \in U$, exists $\lambda_x \in \Lambda_U$ such that $x \in V_{\lambda_x} \subset U$.
$\underline{P2 \implies P1}$:
For every $U \in \mathcal{T}$ there exists $V_x \in \mathcal{V}$ such that $x \in V_x \subset U$. Then,

$$U = \bigcup_{x \in U} V_x \, .$$

■

Definition 3.1.19 *If the collection of open sets $\mathcal{V} \subset \mathcal{T}$ satisfies any (thus both) of the properties $P1$ or $P2$ in the proposition 3.1.18 then it is called a basis of the topology $\mathcal{T}$.*

Example 3.1.20 *Let (M,d) be a metric space. From the definition 3.1.10 of metric topology $\mathcal{T}_d$ it follows that a basis for this topology is given by the open balls $B_r(x)$ for all $r > 0$ and $x \in M$.*

Exercise 3.1.6 *Let $(M,\mathcal{T})$ be a topological space and $N \subset M$ a subset. Show that the following collection of sets,*

$$\mathcal{T}_N = \{U \cap N, \quad U \in \mathcal{T}\} \ ,$$

is a topology on N.

Definition 3.1.21 (subspace topology) *The topology $\mathcal{T}_N$ defined in exercise 3.1.6 on the subset $N \subset M$ of the topological space $(M,\mathcal{T})$ is called subspace topology.*

Exercise 3.1.7 *Consider $[a,b) \subset \mathbb{R}$, $(a < b)$, with subspace topology. Find a neighborhood of a different from $[a,b)$.*

Solution. From the definition we see that to obtain a neighborhood of a point in $[a,b)$ for the subspace topology we need (any) neighborhood of the point in $\mathbb{R}$ and then intersect it with $[a,b)$. Let $c \in (a,b)$ and consider the following neighborhood of a in $\mathbb{R}$,

$$U_a = (a-1,c) \, .$$

Then the set,

$$[a,c) = [a,b) \cap (a-1,c) \subsetneq [a,b) \, ,$$

is a neighborhood of a in $[a,b)$ (distinct from $[a,b)$).

■

Definition 3.1.22 (product topology) *Let $(M,\mathcal{T})$ and $(M',\mathcal{T}')$ be two topological spaces. Then the product topology is the topology generated by the following basis,*

$$\mathcal{B} = \{U \times U' \, , \quad U \in \mathcal{T}, U' \in \mathcal{T}'\} \ .$$

Definition 3.1.23 (closure of a set and denseness) *Let $(M,\mathcal{T})$ be a topological space and $N \subset M$. The closure $\overline{N}$ of N in M is the minimal closed set containing N,*

$$\overline{N} = \bigcap_{F^c \in \mathcal{T},\, F \supset N} F \qquad .$$

The subset $N \subset M$ is said to be dense in M if $\overline{N} = M$.

Proposition 3.1.24 *Let $(M,\mathcal{T})$ be a topological space and $N \subset M$. A point $x \in \overline{N}$ if and only if every neighborhood U of x contains points from N, $U \cap N \neq \emptyset$.*

Proof. TO BE ADDED ■

3.1.4 Continuous maps

The definition of continuity that corresponds to the usual one of calculus is the following.

Definition 3.1.25 (continuity I) *Let $(M,\mathcal{T})$ and $(M',\mathcal{T}')$ be topological spaces. A map $f : M \longrightarrow M'$ is called continuous at the point $x \in M$ if for every neighborhood U' of $f(x)$, exists a neighborhood U of x such that*

$$f(U) \subset U'.$$

The map f is called continuous if it is continuous at every point $x \in M$.

It turns out that continuity of a map can be equivalently defined in a much more concise and elegant way.

Definition 3.1.26 (continuity II) *Let $(M,\mathcal{T})$ and $(M',\mathcal{T}')$ be topological spaces. A map $f : M \longrightarrow M'$ is called continuous if for every $U' \in \mathcal{T}'$ its pre-image is open, $f^{-1}(U') \in \mathcal{T}$, i.e. the pre-image of every open set in M' is open in M.*

Theorem 3.1.27 *The definitions of continuity 3.1.25 and 3.1.26 are equivalent.*

Proof. TO BE ADDED ■

Exercise 3.1.8 *Let $f : (M,\mathcal{T}) \longrightarrow (M',\mathcal{T}')$.*

(a) Show that if $\mathcal{T} = \mathcal{T}^{\text{dis}}$, where $\mathcal{T}^{\text{dis}} = 2^M$ is the discrete topology all maps f are continuous, for any topology $\mathcal{T}'$ on M'.

(b) Show that if $\mathcal{T}' = \mathcal{T}^{\text{und}}$, where $\mathcal{T}^{\text{und}} = \{\emptyset, M'\}$ is the undiscrete topology, all maps f are continuous, for any topology $\mathcal{T}$ on M.

Solution.

(a) Let $U' \in \mathcal{T}'$ be any open set. Its pre-image,

$$f^{-1}(U) = \{x \in M \,:\, f(x) \in U'\} \subset M\,,$$

is a subset of M and therefore $f^{-1}(U) \in \mathcal{T}^{\text{dis}}$ so that f is continuous.

(b) Since for any map f and any topology $\mathcal{T}$ on M, we have $f^{-1}(\emptyset) = \emptyset \in \mathcal{T}$ and $f^{-1}(M') = M \in \mathcal{T}$ we conclude that f is continuous.

■

Theorem 3.1.28 *If $(M_1, \mathcal{T}_1), (M_2, \mathcal{T}_2), (M_3, \mathcal{T}_3)$ are topological spaces and $f : M_1 \longrightarrow M_2$, $h : M_2 \longrightarrow M_3$ are continuous, then $h \circ f : M_1 \longrightarrow M_3$ is also continuous.*

Proof. Notice that, for any set $A \subset M_3$, $(h \circ f)^{-1}(A) = f^{-1}(h^{-1}(A))$. Let U_3 be any open subset of M_3, $U_3 \in \mathcal{T}_3$. Then we have that $h^{-1}(U_3) \in \mathcal{T}_2$ because h is continuous and

$$(h \circ f)^{-1}(U_3) = f^{-1}(h^{-1}(U_3)) \in \mathcal{T}_1$$

because f is continuous. ■

Definition 3.1.29 (quotient topology) *Let $(M, \mathcal{T})$ be a topological space and $\sim$ an equivalence relation. The quotient topology on the set of equivalence classes, $M/\sim$, is the maximal (or finest) topology for which π is continuous,*

$$\mathcal{T}_\pi = \left\{ U \subset M/\sim :\ \pi^{-1}(U) \in \mathcal{T} \right\} .$$

Definition 3.1.30 (homeomorphism) *Let $(M_1, \mathcal{T}_1)$ and $(M_2, \mathcal{T}_2)$ be topological spaces. A function,*

$$f : M_1 \longrightarrow M_2 ,$$

is called a homeomorphism if is bijective and both f and f^{-1} are continuous. If there exists a homeomorphism f the topological spaces $(M_1, \mathcal{T}_1)$ and $(M_2, \mathcal{T}_2)$ are said to be homeomorphic.

Exercise 3.1.9 *Consider on $\mathbb{R}$ the subgroup $\mathbb{Z}$ and on the quotient group the quotient topology. Consider the bijective map,*

$$\begin{array}{rcl} f : \mathbb{R}/\mathbb{Z} & \longrightarrow & [0,1) \\ [x] = x + \mathbb{Z} & \longmapsto & f([x]) = x - \lfloor x \rfloor \end{array}$$

and the topology $\mathcal{T}_{\text{per}}$ on $[0,1)$ for which f is an homeomorphism. Give an example of a neighborhood of $0 \in [0,1)$ different from the whole set $[0,1)$.

Solution.

We have that $U \subset [0,1)$ is open if and only if $f^{-1}(U)$ is open in $\mathbb{R}/\mathbb{Z}$ and this in turn is equivalent to $\pi^{-1}(f^{-1}(U)) = (f \circ \pi)^{-1}(U) \in \mathcal{T}_{\mathbb{R}}$, where

$$(f \circ \pi)(x) = x - \lfloor x \rfloor .$$

Let $\epsilon \in (0, \frac{1}{2})$. As we saw in the exercise 3.1.7, if we had the subspace topology on $[0,1)$ (from $\mathbb{R}$), a neighborhood of 0 would be given by $[0, \epsilon)$. In the present example however, the set

$$(f \circ \pi)^{-1}([0, \epsilon)) = \bigcup_{n \in \mathbb{Z}} [n, n + \epsilon) ,$$

is not open in $\mathbb{R}$ and therefore $[0, \epsilon) \notin \mathcal{T}_{\text{per}}$. A neighborhood of 0 is given by

$$[0, \epsilon) \cup (1 - \epsilon, 1) \,.$$

Indeed,

$$(f \circ \pi)^{-1} \left([0, \epsilon) \cup (1 - \epsilon, 1)\right) = \bigcup_{n \in \mathbb{Z}} (n - \epsilon, n + \epsilon) \in \mathcal{T}_{\mathbb{R}} \,,$$

and therefore $[0, \epsilon) \cup (1 - \epsilon, 1) \in \mathcal{T}_{\text{per}}$.

■

[TO BE CONTINUED]

3.2 Manifolds

3.2.1 Introduction and definition

Differentiable manifolds of dimension m are topological spaces that are locally like open subsets of $\mathbb{R}^m$ on which we can do all operations of usual calculus on $\mathbb{R}^m$ and have additionally fixed a rule on how to glue operations performed on different neighborhoods.

Definition 3.2.1 *Let $(M, \mathcal{T})$ be a topological space, $M \neq \emptyset$.*

(i) (coordinate chart)
A m–dimensional coordinate chart on M is a pair (U, ϕ), where U is an open subset of M and φ is a homeomorphism from U onto an open subset $\varphi(U)$ of $\mathbb{R}^m$, [1]

$$\varphi \,:\, U \xrightarrow{\sim} \varphi(U) \subset \mathbb{R}^m \,.$$

(ii) (transition functions)
Two (m–dimensional coordinate charts (U, φ), (V, ψ)) are said to be C^∞–related or C^∞–compatible if the maps

$$\psi \circ \varphi^{-1} \,:\, \mathbb{R}^m \supset \varphi(U \cap V) \xrightarrow{\psi \circ \varphi^{-1}} \psi(U \cap V) \subset \mathbb{R}^m$$

and

$$\varphi \circ \psi^{-1} \,:\, \mathbb{R}^m \supset \psi(U \cap V) \xrightarrow{\varphi \circ \psi^{-1}} \varphi(U \cap V) \subset \mathbb{R}^m$$

are C^∞–maps.

(iii) (atlas)
A m–dimensional atlas

$$\mathcal{A} = \{(U_i, \varphi_i), i \in I\} \,,$$

on $(M, \mathcal{T})$ is a collection of coordinate charts such that

$$\mathcal{U} = \{U_i \,,\, i \in I\} \,,$$

is an open cover of M and (U_i, φ_i), (U_j, φ_j) are (C^∞)–compatible for all $i, j \in I$.

[1] The existence of such a chart tells us that, at least on points of U, the topological space is homeomorphic to (an open subset of) $\mathbb{R}^m$. In case there is one such chart, with the same m, in a neighborhood of every point this fixes the dimension of M as no open subset of $\mathbb{R}^m$ (except $\emptyset$) is homeomorphic to an open subset of $\mathbb{R}^n$, with $n \neq m$.

(iv) (compatibility of atlases)
Two atlases $\mathcal{A}_1, \mathcal{A}_2$ are said to be compatible if $\mathcal{A}_1 \cup \mathcal{A}_2$ is an atlas.

Remark 3.2.2 One can show that compatibility of atlases is an equivalence relation on the set of m–dimensional atlases on a topological manifold. The equivalence classes of atlases are called differentiable or smooth structures on $(M, \mathcal{T})$,

$$\mathcal{D}(\mathcal{A}) = \left\{ \widetilde{\mathcal{A}} - \text{atlas on } (M, \mathcal{T}) \,:\, \widetilde{\mathcal{A}} \sim \mathcal{A} \right\} .$$

On every equivalence class exists a maximal atlas,

$$\begin{aligned} \mathcal{A}_{\max} &\in \mathcal{D}(\mathcal{A}) \\ \mathcal{A}_{\max} &= \bigcup_{\widetilde{\mathcal{A}} \in \mathcal{D}(\mathcal{A})} \widetilde{\mathcal{A}} . \end{aligned}$$

◊

Definition 3.2.3 (differentiable manifold) *A m–dimensional differentiable manifold is a triple $(M, \mathcal{T}, \mathcal{D})$, where $(M, \mathcal{T})$ is a Hausdorff topological space and $\mathcal{D}$ is a m–dimensional differentiable structure on $(M, \mathcal{T})$.*

Remark 3.2.4 (a) Since a differentiable structure $\mathcal{D}$ is completely defined by any atlas $\mathcal{A} \in \mathcal{D}$, one can define the manifold structure just by exhibiting one atlas

$$\mathcal{A} = \{(U_i, \varphi_i),\ i \in I\} .$$

(b) One defines C^k structures and C^ω structures analogously by requiring that the transition functions, $\psi \circ \varphi^{-1}$, are of class C^k or C^ω.

◊

Example 3.2.5 *Consider the sphere, S^n,*

$$S^n = \left\{ x \in \mathbb{R}^{n+1} \,:\, ||x||^2 = x_1^2 + \cdots + x_{n+1}^2 = 1 \right\} \subset \mathbb{R}^{n+1} .$$

We could introduce the differentiable structure on S^n with an atlas using $2(n+1)$ charts corresponding to projections from $\{x_j > 0\} \cap S^n$ and $\{x_j < 0\} \cap S^n$ to the coordinate hyperplanes $\{x_j = 0\}$. Instead we will use the atlas corresponding to stereographic projections, which uses only two charts. Denote by N the point $(0, \ldots, 0, 1) \in S^n$ and by S the point $(0, \ldots, 0, -1) \in S^n$. Then let $U_N = S^n \setminus \{N\}$ and φ_N be the stereographic projection from N, mapping homeomorphically U_N onto $\mathbb{R}^n$, with S being mapped to $0 \in \mathbb{R}^n$. Denote $\varphi_N(x) = u^N$ and $x = (v, x_{n+1})$, with v denoting the projection of $x \in U_N$ to the hyperplane $\{x_{n+1} = 0\}$, $v = (x_1, \ldots, x_n)$. Then we find,

$$\begin{aligned} ||x||^2 &= ||v||^2 + x_{n+1}^2 = 1 \\ \frac{1}{||u_N||} &= \frac{1 - x_{n+1}}{||v||} . \end{aligned}$$

The vectors $u_N, v \in \mathbb{R}^n$ are parallel and therefore,

$$u_N = \varphi_N(x) = \frac{v}{1-x_{n+1}} = \frac{1}{1-x_{n+1}}\,(x_1,\ldots,x_n)\,.$$

Analogously, one shows that on $U_S = S^n \setminus \{S\}$ the stereographic projection from S gives,

$$u_S = \varphi_S(x) = \frac{v}{1+x_{n+1}} = \frac{1}{1+x_{n+1}}\,(x_1,\ldots,x_n)\,.$$

Let us find the transition functions,

$$\varphi_S \circ \varphi_N^{-1} \,:\, \mathbb{R}^n \setminus \{0\} \longrightarrow \mathbb{R}^n \setminus \{0\}\,.$$

We see that,

$$\begin{aligned} \varphi_N^{-1}(u^N) &= ((1-x_{n+1})\,u^N, x_{n+1}) \\ (1-x_{n+1})^2||u^N||^2 + x_{n+1}^2 &= 1\,. \end{aligned}$$

Then,

$$\begin{aligned} \varphi_S \circ \varphi_N^{-1} \,:\, \varphi_N(U_N \cap U_M) = \mathbb{R}^n \setminus \{0\} &\longrightarrow \varphi_S(U_N \cap U_M) = \mathbb{R}^n \setminus \{0\} \\ \left(\varphi_S \circ \varphi_N^{-1}\right)(u^N) &= \frac{1-x_{n+1}}{1+x_{n+1}}\,u^N = \frac{u^N}{||u^N||^2} = u^S\,. \end{aligned}$$

This function is of class C^∞ and is a homeomorphism from $\mathbb{R}^n \setminus \{0\}$ onto $\mathbb{R}^n \setminus \{0\}$. The inverse of this map is also of class C^∞,

$$\left(\varphi_N \circ \varphi_S^{-1}\right)(u^S) = u^N = \frac{u^S}{||u^S||^2}\,,$$

which confirms that these two coordinate charts form an atlas,

$$\mathcal{A} = \{(U_N, \varphi_N), (U_S, \varphi_S)\}\,,$$

defining a differentiable structure on S^n.

Definition 3.2.6 (differentiable maps) *Let M, N be differentiable manifolds of dimensions m, n respectively and (U, φ), (V, ψ) be coordinate charts such that and $f(U) \subset V$ and $p \in U$.*

1. *The map $f \,:\, M \longrightarrow N$ is called differentiable ($\equiv$ smooth) at the point $p \in M$ if the local representative of f,*

$$\begin{aligned} \widehat{f} &= \psi \circ f \circ \varphi^{-1} \,:\, \mathbb{R}^m \supset \varphi(U) \longrightarrow \psi(V) \subset \mathbb{R}^n \\ \widehat{f} \,:\, &\quad \begin{cases} y_1 = \widehat{f}_1(x_1,\ldots,x_m) \\ \quad\vdots \\ y_n = \widehat{f}_n(x_1,\ldots,x_m)\,, \end{cases} \end{aligned}$$

is of class C^∞.

2. *The map $f : M \longrightarrow N$ is called a differentiable map if it is differentiable at every point $p \in M$.*

Definition 3.2.7 *(a) A map $f : M \longrightarrow N$ is called a diffeomorphism if it is bijective and both f and f^{-1} are differentiable.*

(b) A map $f : M \longrightarrow N$ is called a local diffeomorphism at $p \in M$ if exist neighborhoods U of p and V of $f(p)$ such that

$$f|_U : U \subset M \longrightarrow V \subset N$$

is a diffeomorphism.

Example 3.2.8 *The function,*

$$\begin{array}{rcl} f : \mathbb{R} & \longrightarrow & \mathbb{R}_+ = (0,\infty) \\ x & \longmapsto & e^x \end{array}$$

is a diffeomorphism. Indeed it is a homeomorphism and both f and f^{-1},

$$\begin{array}{rcl} f^{-1} : \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \log(y) \ , \end{array}$$

are differentiable, i.e. of class C^∞.

2. *The function,*

$$\begin{array}{rcl} g : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^3 \end{array}$$

is a differentiable and a homeomorphism but not a diffeomorphism because the inverse function,

$$\begin{array}{rcl} g^{-1} : \mathbb{R} & \longrightarrow & \mathbb{R} \\ y & \longmapsto & (y)^{\frac{1}{3}} \ , \end{array}$$

is not differentiable.

Remark 3.2.9 A map, $f : M \longrightarrow N$, may be a local diffeomorphism at every point of $p \in M$ (see definition 3.2.7, (b)) but not a (global) diffeomorphism to the image. An example is given by the map

$$\begin{array}{rcl} f : \ \mathbb{R} & \longrightarrow & S^1 \\ x & \longmapsto & e^{ix} \ . \end{array}$$

◊

3.2.2 Tangent bundle, vector fields and flows

The concept of tangent vector to a (differentiable) manifold at a given point $p \in M$ is one of the most important concepts in differential geometry. For submanifolds in $\mathbb{R}^n$ it is easy to define and "visualize" tangent vectors. Let S be a smooth surface (embedded) in $\mathbb{R}^3$. Then we say that $v \in \mathbb{R}^3$ is tangent to S at p_0 if there exists a (parametrized) differentiable curve

$$c \, : \, (-\epsilon, \epsilon) \longrightarrow S \subset \mathbb{R}^3 \, ,$$

such that $c(0) = p_0 = (x_0, y_0, z_0) \in S$ and,

$$\begin{aligned} v &= \frac{dc}{dt}(0) = c'(0) = \lim_{t \to 0} \frac{c(t) - c(0)}{t} \\ &= (c_1'(0), c_2'(0), c_3'(0)) \\ &= ((x \circ c)'(0), (y \circ c)'(0), (z \circ c)'(0)) \, , \end{aligned}$$

where here, x, y, z, represent the coordinate functions corresponding to projections to the coordinate axis, e.g.

$$\begin{aligned} x = pr_1 \, : \, \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x_1, y_1, z_1) &\longmapsto x_1 \, . \end{aligned}$$

The set of all tangent vectors to S at p_0, denoted by $\widetilde{T_{p_0} S}$, is a two dimensional subspace of $\mathbb{R}^3$,

$$\widetilde{T_{p_0} S} \subset \mathbb{R}^3 \, . \tag{3.2.1}$$

We call $\widetilde{T_{p_0} S}$ the naive tangent space to S at p_0. To have a definition of the tangent space to S at p_0 that does not use the embedding, $S \subset \mathbb{R}^3$, and is naturally isomorphic to the naive tangent space, let us consider the map from $\widetilde{T_{p_0} S}$ to the set of directional derivatives of differentiable functions at $p_0 \in S$ that we denote by $T_{p_0} S$,

$$\begin{aligned} T_{p_0} S &= \left\{ v \, : \, C^\infty(p) \ni f \mapsto \frac{d}{dt}(f \circ c)(0) \in \mathbb{R} \right\} \\ &= \{\text{directional derivatives of smooth functions at } p_0\} \, . \end{aligned}$$

From the chain rule we know that

$$\begin{aligned} \frac{d}{dt}(f \circ c)(0) &= \frac{d}{dt}_{|t=0} f(x_1(c(t)), x_2(c(t)), x_3(c(t))) \\ &= \sum_{j=1}^{3} \frac{\partial f}{\partial x_j}(c(0)) \, \frac{d}{dt}(x_j \circ c)(0) \\ &= \sum_{j=1}^{3} v_j \, \frac{\partial f}{\partial x_j}(p_0) \, . \end{aligned}$$

We see that we have a map,

$$\begin{aligned} \widetilde{T_{p_0} S} &\longrightarrow T_{p_0} S \\ \widetilde{v} = (v_1, v_2, v_3) &\longmapsto \left(v \, : \, f \longmapsto v(f) = \sum_{j=1}^{3} v_j \frac{\partial f}{\partial x_j}(p_0) \right) \, , \end{aligned} \tag{3.2.2}$$

that one can easily show, using the methods below, is bijective and an isomorphism of vector spaces for the natural structure of vector space on $T_{p_0}S$. So we can use the operators (or linear functionals) of directional derivatives of smooth functions at p_0 as an alternative definition of tangent vectors at p_0. The big advantage of this definition is that it does not envolve the embedding $S \subset \mathbb{R}^3$ and can be easily generalized to any differentiable manifold!

Definition 3.2.10 (tangent vector and tangent space) *Let M be a m–dimensional manifold.*

1. *Let $c : (-\epsilon, \epsilon) \longrightarrow M$ be a parametrized smooth curve on M such that $c(0) = p$. The velocity vector of this curve at p (= tangent vector $v = \frac{dc}{dt}(0)$) is defined to be the directional derivative of functions smooth at p in the direction of c, i.e.*

$$\begin{aligned}\frac{dc}{dt}(0) : C^\infty(p) &\longrightarrow \mathbb{R} \\ f &\longmapsto \left(\frac{dc}{dt}(0)\right)(f) := \frac{d}{dt}(f \circ c)(0) \\ &= \lim_{t\to 0} \frac{f(c(t)) - f(c(0))}{t}\,.\end{aligned}$$

2. *The tangent space, T_pM, of M at the point p is the set of all tangent vectors at p i.e. all operators of directional derivatives along curves at p,*

$$T_pM = \left\{\frac{dc}{dt}(0) : C^\infty(p) \longrightarrow \mathbb{R}\,, c\, \textit{is a smooth curve such that } c(0) = p\right\}.$$

The structure of vector space on T_pM and natural coordinate basis can be introduced with the help of coordinate charts,

$$B_{(U,\varphi,p)} = \left(\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_m}\right)_p\right). \tag{3.2.3}$$

Let us define these coordinate tangent vectors. We call

$$\hat{c} = \varphi \circ c : (-\epsilon, \epsilon) \longrightarrow \mathbb{R}^m$$

and

$$\hat{f} = f \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^m \longrightarrow \mathbb{R}$$

local coordinate representatives of the curve c and of the function f, respectively. The chain rule brings calculus on $\mathbb{R}^m$ into the picture. Indeed, since,

$$f \circ c = f \circ \varphi^{-1} \circ \varphi \circ c = \hat{f} \circ \hat{c}\,,$$

we get that the directional derivative of f along c is equal to the directional derivative of the usual ($\mathbb{R}^m$) function $\hat{f}$ on $\mathbb{R}^m$ along the curve $\hat{c}$ on $\mathbb{R}^m$ and therefore,

$$\begin{aligned} v(f) = \frac{d}{dt}(\hat{f} \circ \hat{c})(0) &= \sum_{j=1}^m \frac{\partial \hat{f}}{\partial x_j}(a)\, \frac{d}{dt}(x_j \circ \hat{c})(0) \\ &= \sum_{j=1}^m v_j\, \frac{\partial \hat{f}}{\partial x_j}(a)\,, \end{aligned} \tag{3.2.4}$$

where $a = \varphi(p) \in \mathbb{R}^m$.

Associated with the chart (U, φ) there are, through every $p \in U \subset M$, m coordinate curves

$$\begin{array}{rcl} c^{(j)} \,:\, (-\epsilon, \epsilon) & \longrightarrow U & \subset M \\ c^{(j)} & = & \varphi^{-1} \circ \hat{c}^{(j)}\,, \end{array}$$

where $\hat{c}^{(j)}$ is the j–th coordinate line on $\mathbb{R}^m$,

$$\hat{c}^{(j)}(t) = (a_1, \ldots, a_j + t, \ldots, a_m)\,,$$

and therefore,

$$c^{(j)}(t) = \varphi^{-1}(a_1, \ldots, a_j + t, \ldots, a_m)\,.$$

The tangent vector

$$\left(\frac{\partial}{\partial x_j}\right)_a \in T_a\mathbb{R}^m\,,$$

as in usual calculus, is the partial derivative, i.e. the directional derivative along the coordinate line $\hat{c}^{(j)}$ at the point a,

$$\begin{aligned} \left(\frac{\partial}{\partial x_j}\right)_a \,:\, \hat{f} \longmapsto \left(\frac{\partial}{\partial x_j}\right)(\hat{f})(a) &= \frac{d}{dt}\left(\hat{f} \circ \hat{c}^{(j)}\right)(0) \\ &= \lim_{t \to 0} \frac{\hat{f}(a_1, \ldots, a_j + t, \ldots, a_m) - \hat{f}(a_1, \ldots, a_j, \ldots, a_m)}{t}\,. \end{aligned}$$

It is then natural to define,

$$\left(\frac{\partial}{\partial x_j}\right)_p \in T_pM\,,$$

as the directional derivative along the coordinate line $c^{(j)}$ at the point p,

$$\begin{aligned} \left(\frac{\partial}{\partial x_j}\right)_p \,:\, f \longmapsto \left(\frac{\partial}{\partial x_j}\right)_p (f) &:= \frac{d}{dt}\left(f \circ c^{(j)}\right)(0) \\ &= \lim_{t \to 0} \frac{f(c^{(j)}(t)) - f(p)}{t} \\ &= \frac{d}{dt}\left(f \circ \varphi^{-1} \circ \varphi \circ c^{(j)}\right)(0) \qquad (3.2.5) \\ &= \frac{d}{dt}\left(\hat{f} \circ \hat{c}^{(j)}\right)(0) \\ &=: \left(\frac{\partial}{\partial x_j}\right)_a (\hat{f})\,. \end{aligned}$$

Comparing (3.2.5) with (3.2.4) we see that

$$T_pM \subset \mathrm{Span}_{\mathbb{R}} \left\{ \left(\frac{\partial}{\partial x_j}\right)_p \,, j = 1, \ldots, m \right\}\,. \qquad (3.2.6)$$

Let us show that in fact we have equality, i.e. for every $v = \sum_{j=1}^{m} v_j \left(\frac{\partial}{\partial x_j}\right)_p$, there exists a curve

$$c^{(v)} : (-\epsilon, \epsilon) \longrightarrow M,\ c^{(v)}(0) = p,$$

such that

$$v_j = \frac{d}{dt}\left(x_j \circ c^{(v)}\right)(0)\,.$$

Indeed, we can choose $c^{(v)} = \varphi^{-1} \circ \hat{c}^{(v)}$, where

$$\begin{array}{rcl} c^{(v)} : (-\epsilon, \epsilon) & \longrightarrow & \mathbb{R}^m \\ t & \longmapsto & c^{(v)}(t) = a + tv = (a_1 + tv_1, \ldots, a_m + tv_m)\,, \end{array}$$

and $\epsilon > 0$ is such that $c^{(v)}((-\epsilon, \epsilon)) \subset \varphi(U)$. Therefore we have equality in (3.2.6),

$$T_pM = \mathrm{Span}_{\mathbb{R}} \left\{ \left(\frac{\partial}{\partial x_j}\right)_p, j = 1, \ldots, m \right\}.$$

In fact this set of coordinate vectors is linearly independent and therefore is a basis of T_pM.

Exercise 3.2.1 *Let (U, φ) be a coordinate chart of M, $\dim(M) = m$. Show that the vectors $\left(\frac{\partial}{\partial x_j}\right)_p$, $j = 1, \ldots, m$, are linearly independent.*

Solution. Let us, for clarity, in this exercise denote with different letters the coordinate functions on $U \subset M$ ($x_j = \varphi \circ \hat{x}_j$) and on $\varphi(U) \subset \mathbb{R}^m$ ($\hat{x}_j$) (in the main text we, for simplicity, denote them with the same letters). Let us show that if

$$\sum_{j=1}^{m} v_j \left(\frac{\partial}{\partial x_j}\right)_p = 0\,,$$

then the v_j have to be all equal to zero. Let us apply both sides of this equation to $x_k \in C^\infty(U)$. Then

$$\begin{aligned} \sum_{j=1}^{m} v_j \left(\frac{\partial}{\partial x_j}\right)_p (x_k) &= 0 \\ \Leftrightarrow \sum_{j=1}^{m} v_j \left(\frac{\partial}{\partial \hat{x}_j}\right)_a (\hat{x}_k) &= 0 \\ \Leftrightarrow v_k &= 0\,, \end{aligned} \tag{3.2.7}$$

for all $k = 1, \ldots, m$, which proves linear independence so that $B_{(U,\varphi,p)}$ in (3.2.3) is indeed a basis of T_pM for every point $p \in U$.

■

A differential map

$$f : M \longrightarrow N$$

sends points of M to points of N (we say that f pushes forward points) and as a consequence, as we will see now, also pushes forward curves and vectors and pulls back functions on N to functions on M. The pushforward by f of a parametrized curve on M, $c : (-\epsilon, \epsilon) \longrightarrow M$ is just,

$$\tilde{c} = f \circ c \, : \, (-\epsilon, \epsilon) \longrightarrow N \, ,$$

and the pushforward, denoted by $(f_*)_p(v)$, of its velocity vector $v = \frac{dc}{dt}(0)$ at $p = c(0) \in M$ is the velocity vector of the pushforward of the curve,

$$(f_*)_p(v) := \frac{d}{dt}(f \circ c)(0) \, ,$$

so that we get a (linear) map

$$(f_*)_p \, : \, T_pM \longrightarrow T_{f(p)}N \, , \tag{3.2.8}$$

called pushforward induced by f or also differential of f at p and denoted alternatively by df_p.

Proposition 3.2.11 (see [GN], Prop. 4.6) *Let M, N be manifolds of dimensions m, n, respectively. The pushforward map,*

$$(f_*) \, : \, T_pM \longrightarrow T_{f(p)}N$$

is a linear transformation and

$$\begin{aligned} (f_*)_p \left(\left(\frac{\partial}{\partial x_j} \right)_p \right) &= \sum_{k=1}^{n} \frac{\partial y_k}{\partial x_j}(p) \left(\frac{\partial}{\partial y_k} \right)_{f(p)} \\ (f_*)_p \left(\sum_{j=1}^{m} v_j \left(\frac{\partial}{\partial x_j} \right)_p \right) &= \sum_{k=1}^{n} u_k \left(\frac{\partial}{\partial y_k} \right)_{f(p)} \, , \end{aligned}$$

where

$$u_k = \sum_{j=1}^{m} \frac{\partial y_k}{\partial x_j}(p) \, v_j \, , \qquad k = 1, \ldots, n \, .$$

The tangent bundle,

$$TM := \bigsqcup_{p \in M} T_pM \, ,$$

has a natural structure of $2m$–dimensional manifold with a differentiable map,

$$\begin{aligned} \pi \, : \, TM &\longrightarrow M \\ TM \supset T_pM \ni v &\longmapsto p \, , \end{aligned}$$

called canonical projection. The differentiable structure is defined as follows. For every coordinate chart (U, φ) on M we construct a coordinate chart $(\pi^{-1}(U), \widetilde{\varphi})$,

$$\begin{aligned} \widetilde{\varphi} \, : \, \pi^{-1}(U) &\longrightarrow \varphi(U) \times \mathbb{R}^m \subset \mathbb{R}^{2m}, \\ TM \supset T_pM \ni v = \sum_{j=1}^{m} v_j \left(\frac{\partial}{\partial x_j} \right)_p &\longmapsto (\varphi(p), (v_1, \ldots, v_m)) \\ &= ((x_1(p), \ldots, x_m(p)), (v_1, \ldots, v_m)) \, . \end{aligned} \tag{3.2.9}$$

It is easy to see that by taking an atlas,

$$\mathcal{A} = \{(U_\lambda, \varphi_\lambda)\,,\, \lambda \in \Lambda\}\ ,$$

on M we get a $2m$–dimensional atlas $\widetilde{\mathcal{A}}$ on TM, given by,

$$\widetilde{\mathcal{A}} = \left\{(\pi^{-1}(U_\lambda), \widetilde{\varphi}_\lambda)\,,\, \lambda \in \Lambda\right\}\ , \tag{3.2.10}$$

where the lift of the coordinate chart map φ_λ to,

$$\widetilde{\varphi}_\lambda\ :\ \pi^{-1}(U_\lambda) \subset TM \longrightarrow \mathbb{R}^{2m}\,,$$

is defined as explained in (3.2.9).

Definition 3.2.12 (vector field) *A (differentiable) vector field Y on the manifold M is a differentiable map,*

$$\begin{array}{rcl} Y\,:\,M & \longrightarrow & TM \\ p & \longmapsto & Y_p \in T_pM\,. \end{array}$$

The vector space of all vector fields on M is denoted by $\mathcal{X}(M)$.

Remark 3.2.13

1. The vector field Y is differentiable if and only if it maps differentiable functions to differentiable functions, i.e. the functions

$$Y(f)(p) = Y_p(f)\ , \qquad \forall p \in M\,,$$

are differentiable for every $f \in C^\infty(M)$, so that the vector fields define maps

$$\begin{array}{rcl} Y\,:\,C^\infty(M) & \longrightarrow & C^\infty(M) \\ f & \longmapsto & Y(f)\,. \end{array}$$

2. A definition of vector fields equivalent to definition 3.2.12 is to say that a vector field Y is a section of the tangent bundle, i.e. a differentiable map $Y\,:\,M \longrightarrow TM$ such that

$$\pi \circ Y = \mathrm{Id}_M\ . \tag{3.2.11}$$

$\Diamond$

Let us now make a small detour to introduce Lie algebras and some of their properties.

Definition 3.2.14 (Lie algebra) *A Lie algebra over the field $\mathbb{K}^2$ is a pair $(A, [\,,\,])$, where A is a vector space and $[\,,\,]$,*

$$[\,,\,]\,:\,A \times A \longrightarrow A$$

is a composition satisfying the following properties,

[2] We will consider only the cases $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$.

(i) $[\,,\,]$ *is bilinear.*

(ii) $[\,,\,]$ *is antisymmetric, i.e. for every* $X, Y \in A$,

$$[X, Y] = -[Y, X]\,.$$

(iii) $[\,,\,]$ *satisfies the Jacobi identity,*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0\,, \qquad \forall X, Y, Z \in A\,.$$

Exercise 3.2.2 *Let* $(A, \circ)$ *be an associative algebra, i.e a vector space with a bilinear composition, which is associative, i.e.*

$$(X \circ Y) \circ Z = X \circ (Y \circ Z)\ , \qquad \forall X, Y, Z \in A.$$

Show that $(A, [\,,\,])$*, with Lie bracket given by the commutator,*

$$[X, Y] = X \circ Y - Y \circ X\,, \qquad \forall X, Y \in A\,,$$

is a Lie algebra.
Solution. *We have only to prove that the associativity of* $\circ$ *implies the Jacobi identity of the commutator. Indeed, for every* $X, Y, Z \in A$*, we have*

$$\begin{aligned} & [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] \\ = \ & X \circ [Y, Z] - [Y, Z] \circ X + Y \circ [Z, X] - [Z, X] \circ Y + Z \circ [X, Y] - [X, Y] \circ Z \\ = \ & X \circ (Y \circ Z - Z \circ Y) - (Y \circ Z - Z \circ Y) \circ X \\ + \ & Y \circ (Z \circ X - X \circ Z) - \circ (Z \circ X - X \circ Z) \circ Y \\ + \ & Z \circ (X \circ Y - Y \circ X) - (X \circ Y - Y \circ X) \circ Z \\ = \ & 0\ . \end{aligned}$$

■

Due to the relation of Lie algebras with Lie groups, which we will study in the next chapter, many definitions and results in Lie algebra theory parallel those of groups.

Definition 3.2.15 (homomorphism, isomorphism) *Let* $\mathfrak{g}, \mathfrak{h}$ *be Lie algebras. A linear map* $\varphi : \mathfrak{g} \longrightarrow \mathfrak{h}$ *such that,*

$$\varphi([X, Y]) = [\varphi(X), \varphi(Y)] \qquad \forall X, Y \in \mathfrak{g}\,,$$

is called a (Lie algebra) homomorphism. If φ *is bijective than* φ *is called an isomorphism. Two Lie algebras are said to be isomorphic if there is an isomorphism from one to the other.*

The role of normal subgroups in Lie algebras is played by ideals.

Definition 3.2.16 (subalgebra and ideal) *Let* $\mathfrak{g}$ *be a Lie algebra.*

(a) *A vector subspace* $\mathfrak{h} \subset \mathfrak{g}$ *is called a Lie subalgebra if the restriction of the Lie bracket to* $\mathfrak{h}$ *defines on it a Lie algebra structure, i.e. if*

$$[X, Y] \in \mathfrak{h}\,, \qquad \forall X, Y \in \mathfrak{h}\,.$$

(b) *A Lie subalgebra* $\mathfrak{h} \subset \mathfrak{g}$ *is called an ideal of* $\mathfrak{g}$ *if,*

$$[X, Y] \in \mathfrak{h}\,, \qquad \forall X \in \mathfrak{h},\ \forall Y \in \mathfrak{g}\,.$$

As for normal subgroups (see example 2.2.8), if $\mathfrak{h}$ is an ideal of the Lie algebra $\mathfrak{g}$, the quotient space,

$$\mathfrak{g}/\mathfrak{h} = \{[X] = \pi(X) = X + \mathfrak{h}\,, \quad X \in \mathfrak{g}\}\ ,$$

as a natural structure of Lie algebra for which the canonical projection π is an homomorphism,

$$[\pi(X_1), \pi(X_2)] = \pi([X_1, X_2])\,, \qquad \forall X_1, X_2 \in \mathfrak{g}\,. \tag{3.2.12}$$

Indeed, for (3.2.12) to define a Lie bracket on $\mathfrak{g}/\mathfrak{h}$, we have to show that the formula does not depend on the representatives $X_1 \in X_1 + \mathfrak{h} = \pi(X_1)$ and $X_2 \in X_2 + \mathfrak{h} = \pi(X_2)$. Indeed let $Y_1 = X_1 + H_1$ and $Y_2 = X_2 + H_2$, where $H_1, H_2 \in \mathfrak{h}$. Then, since $\mathfrak{h}$, is an idelal we have

$$\begin{aligned} [Y_1, Y_2] &= [X_1 + H_1, X_2 + H_2] \\ &= [X_1, X_2] + [X_1, H_2] + [H_1, X_2] + [H_1, H_2] = \\ &= [X_1, X_2] + H\,, \end{aligned}$$

where $H = [X_1, H_2] + [H_1, X_2] + [H_1, H_2] \in \mathfrak{h}$.

Then we have the isomorphism theorem for Lie algebras.

Theorem 3.2.17 (isomorphism theorem) *Let* $\varphi : \mathfrak{g} \longrightarrow \mathfrak{h}$ *be a Lie algebra homomorphism. Then,*

(a) $Im(\varphi)$ *is a Lie subalgebra of* $\mathfrak{h}$.

(b) $\ker(\varphi)$ *is an ideal of* $\mathfrak{g}$.

(c) *The map* $\tilde{\varphi} : \mathfrak{g}/\ker(\varphi) \longrightarrow Im(\varphi)$,

$$\tilde{\varphi}(X + \ker(\varphi)) = \varphi(X)\,,$$

induced by φ, *is an isomorphism of Lie algebras.*

Let us now go back to vector fields on a manifold. We saw that vector fields define first order linear differential operators,

$$\mathcal{X}(M) \ni Y : C^\infty(M) \longrightarrow C^\infty(M)\,.$$

If we compose two vector fields $X, Y \in \mathcal{X}(M)$ we get a second order linear differential operator,

$$X \circ Y : C^\infty(M) \longrightarrow C^\infty(M)\,,$$

which is not a vector field. However, if we take the commutator of two vector fields the result is a vector field. To show this let $f \in C^\infty(M)$ and the vector fields $X, Y \in \mathcal{X}(M)$, on a chart (U, ψ), be given by

$$\begin{aligned} X_p &= \sum_{j=1}^n X_j(p) \left(\frac{\partial}{\partial x_j} \right)_p \\ Y_p &= \sum_{j=1}^n Y_j(p) \left(\frac{\partial}{\partial x_j} \right)_p . \end{aligned}$$

Then,

$$\begin{aligned} [X,Y](f) &= (X \circ Y - Y \circ X)(f) \\ &= \left(\sum_{j=1}^n X_j \frac{\partial}{\partial x_j} \right) \left(\sum_{i=1}^n Y_i \frac{\partial}{\partial x_i} f \right) - \left(\sum_{j=1}^n Y_j \frac{\partial}{\partial x_j} \right) \left(\sum_{i=1}^n X_i \frac{\partial}{\partial x_i} f \right) \\ &= \sum_{i=1}^n (X(Y_i) - Y(X_i)) \frac{\partial}{\partial x_i} f . \end{aligned}$$

and therefore $[X, Y]$ is indeed a vector field given in a local coordinate chart by,

$$[X,Y] = \sum_{i=1}^n (X(Y_i) - Y(X_i)) \frac{\partial}{\partial x_i} .$$

From exercise 3.2.2 it follows that this commutator satisfies the Jacobi identity and therefore defines on $\mathcal{X}(M)$ a Lie algebra structure. This is preserved under diffeomorphisms.

Proposition 3.2.18 *Let $f : M \longrightarrow N$ be a diffeomorphism. Then*

$$f_* : \mathcal{X}(M) \longrightarrow \mathcal{X}(N) ,$$

is a isomorphism of Lie algebras i.e. is a vector space isomorphism and,

$$[f_*(X), f_*(Y)] = f_*([X,Y]) , \qquad X, Y \in \mathcal{X}(M) .$$

Definition 3.2.19 (integral curve of a vector field) *Let $Y \in \mathcal{X}(M)$. A smooth curve*

$$c : (-\epsilon, \epsilon) \longrightarrow M$$

is an integral curve of Y if

$$\frac{dc}{dt}(t) = Y_{c(t)} , \qquad \forall t \in (-\epsilon, \epsilon) . \tag{3.2.13}$$

i.e., for every function $f \in C^\infty(M)$, we have

$$\frac{d(f \circ c)}{dt}(t) = Y_{c(t)}(f) , \qquad \forall t \in (-\epsilon, \epsilon) .$$

In local coordinates (U,φ) such that $c(-\epsilon,\epsilon)\subset U$ one has

$$\begin{aligned} Y_p &= \sum_{j=1}^m Y_j(p)\left(\frac{\partial}{\partial x_j}\right)_p , \\ \hat{c}(t) &= (\varphi\circ c)(t) = (\hat{c}_1(t),\ldots,\hat{c}_m(t)) , \end{aligned}$$

and therefore,

$$\frac{dc}{dt}(t) = \sum_{j=1}^m \frac{d\hat{c}_j}{dt}(t)\left(\frac{\partial}{\partial x_j}\right)_{c(t)} .$$

Therefore, the equation (3.2.13) for the integral curves is equivalent to the system

$$\frac{d\hat{c}_j}{dt}(t) = \hat{Y}_j(\hat{c}(t)) , \quad j = 1,\ldots,m,$$

or, more informally,

$$\frac{dx_j}{dt} = Y_j(x_1,\ldots,x_m) , \quad j = 1,\ldots,m.$$

From the Picard-Lindelöf theorem it follows that for $Y\in\mathcal{X}(M)$ and $p\in M$ there exists $\epsilon>0$ such that the initial value problem,

$$\left\{\begin{array}{lcl} \frac{dc}{dt}(t) &=& Y_{c(t)} \\ c(0) &=& p \end{array}\right. \tag{3.2.14}$$

has a unique solution,

$$c_p : (-\epsilon,\epsilon)\longrightarrow M .$$

By taking integral curves of Y on a neighborhood of p we get a local flow,

$$\begin{array}{rcl} \widetilde{\psi}^Y : I\times U &\longrightarrow& M \\ (t,q) &\longmapsto& \widetilde{\psi}^Y(t,q) := c_q(t) . \end{array}$$

Theorem 3.2.20 *Let $Y\in\mathcal{X}(M)$. For every $p\in M$ there exists a neighborhood U of p, $I=(-\epsilon,\epsilon)$ and a (unique) local flow of Y,*

$$\widetilde{\psi}^Y : I\times U\longrightarrow M ,$$

such that

(i) $\widetilde{\psi}^Y$ is differentiable.

(ii) For every $q\in U$ the map

$$\begin{array}{rcl} \widetilde{\psi}^Y(\cdot,q) : I &\longrightarrow& M \\ t &\longmapsto& \widetilde{\psi}^Y(t,q) , \end{array}$$

is an integral curve of Y with initial condition q i.e.

$$\left\{\begin{array}{lcl} \frac{\partial}{\partial t}\widetilde{\psi}^Y(t,q) &=& Y_{\widetilde{\psi}^Y(t,q)} \\ \widetilde{\psi}^Y(0,q) &=& q . \end{array}\right.$$

(iii) For every $t \in I$, $\widetilde{\psi}^Y(t,\cdot)$ is a diffeomorphism from U to its image in M,

$$\psi_t^Y := \widetilde{\psi}^Y(t,\cdot) \,:\, U \longrightarrow M\,.$$

(iv) For every $t, s, t+s \in I$ and $q, \widetilde{\psi}^Y(s,q) \in U$, one has

$$\widetilde{\psi}^Y(t, \widetilde{\psi}^Y(s,q)) = \widetilde{\psi}^Y(s+t,q)\,.$$

Definition 3.2.21 (complete vector field) *The vector field $Y \in \mathcal{X}(M)$ is called complete if its local flows can be extended to a global flow,*

$$\widetilde{\psi}^Y \,:\, \mathbb{R} \times M \longrightarrow M\,.$$

The subspace of all complete vector fields on M will be denoted by $\mathcal{X}_c(M)$.

Proposition 3.2.22 *All vector fields on a a compact manifold M are complete.*

Remark 3.2.23 From the Theorem 3.2.21 we conclude that global flows of complete vector fields define actions of $\mathbb{R}$ on M i.e. differentiable maps

$$\widetilde{\psi}^Y \,:\, \mathbb{R} \times M \longrightarrow M\,,$$

such that, for every $t \in \mathbb{R}$, $\psi_t^Y := \widetilde{\psi}^Y(t,\cdot) \in \mathrm{Diff}(M)$ and the map,

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathrm{Diff}(M) \\ t & \longmapsto & \psi_t^Y\,. \end{array}$$

is a group homomorphism.

Conversely, every differentiable action $\mathbb{R} \overset{\psi}{\curvearrowright} M$ corresponds to the flow of a vector field, Y,

$$Y_p = \frac{d}{dt}\,\psi_t(p)|_{t=0}\,.$$

$\Diamond$

We therefore obtain a very important equivalence

$$\{\text{infinitesimal actions of } \mathbb{R}\} := \mathcal{X}_c(M) \longleftrightarrow \{\text{Differentiable actions of } \mathbb{R} \text{ on } M\}\,.$$

Exercise 3.2.3 *Let $m \in \mathbb{N}_0$. Show that the vector field,*

$$Y_x = x^m\,\frac{\partial}{\partial x}\,,$$

is complete if and only if $m = 0, 1$.

Solution. Let us consider separately the cases with $m = 0, m = 1$ and $m \geq 1$.
Let $\underline{m = 0}$:
We have $Y_x = \frac{\partial}{\partial x}$ and therefore the initial value problem for its integral curves reads

$$\begin{cases} \dot{x} &= 1 \\ x(0) &= x_0 \,, \end{cases}$$

so that

$$x(t) = c_{x_0}(t) = x_0 + t \,, \qquad \forall t \in \mathbb{R}.$$

The global flow is then,

$$\psi_t(x) = c_x(t) = x + t \,.$$

Let $\underline{m = 1}$:
$Y_x = x\frac{\partial}{\partial x}$ and therefore

$$\begin{cases} \dot{x} &= x \\ x(0) &= x_0 \,, \end{cases}$$

so that

$$x(t) = c_{x_0}(t) = x_0 \, e^t \,, \qquad \forall t \in \mathbb{R}.$$

For the global flow we obtain,

$$\psi_t(x) = c_x(t) = x \, e^t \,.$$

Let now $\underline{m \geq 2}$:
We have now $Y_x = x^m \frac{\partial}{\partial x}$ and, as we will see, the growth of the vector field as $x \to \infty$ implies that the integral curves with initial condition $x_0 \neq 0$ reach infinity in finite time. The equation

$$\begin{cases} \dot{x} &= x^m \\ x(0) &= x_0 \,, \end{cases}$$

is equivalent to

$$\begin{aligned} \int_{x_0}^{x} \frac{du}{u^m} &= \int_0^t ds \\ \Leftrightarrow -\frac{1}{m-1}\left(\frac{1}{x^{m-1}} - \frac{1}{x_0^{m-1}}\right) &= t \\ \Leftrightarrow x(t) = c_{x_0}(t) &= \frac{x_0}{\left(1-(m-1)x_0^{m-1}\,t\right)^{\frac{1}{m-1}}} \,, \end{aligned}$$

so that, for every $x_0 \neq 0$, the solution is defined (in one of the time directions) only for $|t| < \frac{1}{(m-1)x_0^{m-1}}$ and therefore the vector field is not complete. The local flows are given by,

$$\psi_t(x) = \frac{x_0}{\left(1-(m-1)x_0^{m-1}\,t\right)^{\frac{1}{m-1}}} \,.$$

■

Exercise 3.2.4 *Consider the linear vector field on $\mathbb{R}^m$.*

$$\begin{array}{rcl} \mathrm{Mat}_m(\mathbb{R}) \ni A & \longmapsto & Y^A \in \mathcal{X}(\mathbb{R}^m) \\ Y^A_x & = & \displaystyle\sum_{j,k=1}^m a_{jk} x_k \, \frac{\partial}{\partial x_j} \, . \end{array}$$

(a) Show that the map

$$A \longmapsto Y^A \, ,$$

is an antihomomorphism of Lie algebras, i.e.

$$[Y^A, Y^B] = -Y^{[A,B]} \, , \qquad \forall A \in \mathrm{Mat}_m(\mathbb{R}) \, .$$

(b) Find, in exponential form, the flow of the vector field Y^A.

(c) Find the flows of the vector fields $X_1, X_2, X_3 \in \mathcal{X}(\mathbb{R}^3)$, given by,

$$X_1 = y \, \frac{\partial}{\partial z} - z \, \frac{\partial}{\partial y} \, , \; X_2 = z \, \frac{\partial}{\partial x} - x \, \frac{\partial}{\partial z} \, , \; X_3 = x \, \frac{\partial}{\partial y} - y \, \frac{\partial}{\partial x} \, .$$

Solution.

(a)

$$\begin{aligned} [Y^A, Y^B] &= \left[\sum_{j,k=1}^m a_{jk} x_k \, \frac{\partial}{\partial x_j}, \sum_{i,l=1}^m b_{il} x_l \, \frac{\partial}{\partial x_i} \right] \\ &= \sum_{j,k,i}^m a_{jk} b_{ij} x_k \, \frac{\partial}{\partial x_i} - \sum_{i,l,j}^m a_{ji} b_{il} x_l \, \frac{\partial}{\partial x_j} \\ &= \sum_{i,k=1}^m (BA - AB)_{i,k} \; x_k \frac{\partial}{\partial x_i} \\ &= -Y^{[A,B]} \, . \end{aligned}$$

(b) For the integral curves of Y^A we obtain the system,

$$\dot{x}_k = \sum_{j=1}^k a_{kj} \, x_j \, ,$$

and the corresponding initial value problem in matrix form,

$$\left\{ \begin{array}{rcl} \dot{x} & = & Ax \\ x(0) & = & a \, . \end{array} \right.$$

The solution, in terms of the matrix exponential, is

$$x(t) = c_a(t) = e^{At} \, a \, , \qquad \forall t \in \mathbb{R},$$

where,

$$e^{At} = \mathrm{I}_m + tA + \cdots + \frac{t^k A^k}{k!} + \cdots = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}\,.$$

The global (linear) flow is then,

$$\psi_t(x) = e^{At}\, x\,.$$

(c) These vector fields are particular cases of the vector fields studied in (b). Let us find the flow explicitly for $X_1 = Y^{T_1}$, where

$$T_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

As we saw in (b) the global flow integrating Y^{T_1} is

$$\psi_t^{X_1}(u) = e^{tT_1}\, u = e^{tT_1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}\,.$$

To find explicitly the matrix entries of the matrix exponential notice that

$$\begin{aligned}
T_1^2 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
T_1^3 &= T_1^2 T_1 = -T_1 \\
T_1^{2n} &= (-1)^n \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\,, \\
T_1^{2n+1} &= (-1)^n\, T_1\,,
\end{aligned}$$

and therefore

$$\begin{aligned}
e^{t\,T_1} &= \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!}\, T_1^{2n} + \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!}\, T_1^{2n+1} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} (-1)^n \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} (-1)^n \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(t) & -\sin(t) \\ 0 & \sin(t) & \cos(t) \end{pmatrix}\,.
\end{aligned}$$

The flow corresponds, as expected to the orthogonal rotations around the $\mathcal{O}_x$ axis,

$$\psi_t^{X_1}(x,y,z) = (x, \cos(t)y - \sin(t)z, \sin(t)y + \cos(t)z)\,.$$

The flows of X_2 and X_3 correspond to rotations around the axes $\mathcal{O}_y$ and $\mathcal{O}_z$ and read,

$$\begin{aligned}
\psi_t^{X_2}(x,y,z) &= (\cos(t)x + \sin(t)z, y, \sin(t)x + \cos(t)z) \\
\psi_t^{X_3}(x,y,z) &= (\cos(t)x - \sin(t)y, \sin(t)x + \cos(t)y, z)\,.
\end{aligned}$$

■

3.2.3 Tensor fields and their symmetries

Definition 3.2.24 *Let V be a finite dimensional vector space over $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$). The dual V^* is the space of (linear) functionals on V,*

$$V^* = \{\alpha \, : \, V \longrightarrow \mathbb{K} \, , \, \alpha \text{ is } \mathbb{K}\text{–linear}\} \; .$$

Proposition 3.2.25 *V^* is a $\mathbb{K}$ vector space and* $\dim(V^*) = \dim(V)$*.*

Proof. V^* is a vector subspace of $C^\infty(V, \mathbb{K})$ because $0 \in V^*$ and the property of being linear is closed under linear combination of functions. Let $\dim(V) = n$ and $B = (v_1, \ldots, v_n)$, be a ordered basis of V. Define n covectors, i.e. elements of V^*, v_j^* , $j = 1, \ldots n$, as follows,

$$v_j^*(v_k) = \delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{else} \, . \end{cases}$$

Let us show that,

$$B^* = (v_1^*, \ldots, v_n^*) \, ,$$

is a basis of V^*. We show first that $\text{span}_{\mathbb{R}}(B^*) = V^*$ by showing that for every $\alpha \in V^*$,

$$\alpha = \sum_{j=1}^{n} \alpha(v_j) \, v_j^* \, . \tag{3.2.15}$$

By applying the right hand side of (3.2.15) to v_k we obtain

$$\left(\sum_{j=1}^{n} \alpha(v_j) \, v_j^* \right) (v_k) = \alpha(v_k) \, ,$$

for every $k = 1, \ldots, n$. This result coincides with the result obtained by applying the left hand side of (3.2.15) to v_k, which proves (3.2.15). Let us now show that $v_1^*, \ldots, v_n^*$ are linearly independent. The same argument as above in the case of $\alpha = 0$ shows that

$$\sum_{j=1}^{n} c_j \, v_j^* = 0 \Leftrightarrow c_k = 0 \; , \quad k = 1, \ldots, n \, .$$

■

Remark 3.2.26 The ordered basis $B^* = (v_1^*, \ldots, v_n^*)$ is called dual basis of V^*. ◊

Let us now introduce the cotangent space T_p^*M to the manifold at a point $p \in M$.

Definition 3.2.27 *The $\mathbb{R}$–vector space dual to T_pM is called cotangent space to M at p,*

$$T_p^*M := (T_pM)^* \; .$$

The disjoint union of the cotangent spaces

$$T^*M = \bigsqcup_{p \in M} T_p^*M$$

is called cotangent bundle of M and the map,

$$\begin{array}{rcl} \pi_{T^*} \, : \, T^*M & \longrightarrow & M \\ T_p^*M \ni \alpha & \longmapsto & p \, , \end{array}$$

is called canonical projection.

T^*M has a differentiable structure, similar to the one on TM (see (3.2.9), (3.2.10)), making T^*M a $2m$–dimensional differentiable manifold. We will describe it below.

Let us introduce (differentiable) one forms analogously to the definition of vector fields in definition 3.2.12 and (3.2.11), as sections of the cotangent bundle,

Definition 3.2.28 (1–forms) *A (differential) 1–form β on the manifold M is a (smooth) section of the cotangent bundle i.e. a differentiable map,*

$$\beta \,:\, M \longrightarrow T^*M\,,$$

such that,

$$\pi_{T^*} \circ \beta = \mathrm{Id}_M\,. \tag{3.2.16}$$

The vector space of all differential 1–forms on M is denoted by $\Omega^1(M)$.

Remark 3.2.29

1. A section β of the cotangent bundle is differentiable if and only if it maps smooth vector fields to smooth functions,

$$\beta \,:\, \mathcal{X}(M) \ni Y \longmapsto \beta(Y) \in C^\infty(M)\,.$$

2. Given a function $f \in C^\infty(M)$ we define a 1–form, df, called differential or exterior derivative of f, and such that

$$df(Y) = Y(f)\,.$$

C^∞ functions are also called 0–forms so that the exterior derivative maps 0–forms to 1–forms,

$$\begin{array}{rcl} d \,:\, \Omega^0(M) = C^\infty(M) & \longrightarrow & \Omega^1(M) \\ f & \longmapsto & df\,. \end{array} \tag{3.2.17}$$

3. On a coordinate chart (U, φ) the exterior derivative of the coordinate functions $x_j \in C^\infty(U)$ define, at every point $p \in U$, an ordered basis of T^*_qM,

$$((dx_1)_q, \ldots, (dx_m)_q)\ ,$$

dual to the ordered basis of T_qM given by the partial derivatives,

$$\left(\left(\frac{\partial}{\partial x_1}\right)_q, \ldots, \left(\frac{\partial}{\partial x_m}\right)_q\right)\,.$$

Indeed (compare with (3.2.7)),

$$(dx_j)_q\left(\left(\frac{\partial}{\partial x_k}\right)_q\right) = \left(\frac{\partial}{\partial x_k}\right)_q(x_j) = \delta_{jk}\,.$$

4. The exterior derivative will be extended to higher degree forms below in definition 3.2.55.

◊

Let us now briefly describe the differentiable structure on T^*M. As in (3.2.9) for the tangent bundle, for every coordinate chart (U, φ) on M we construct a coordinate chart $(\pi_{T^*}^{-1}(U), \widehat{\varphi})$ on T^*M,

$$\begin{aligned} \widehat{\varphi} : \pi_{T^*}^{-1}(U) &\longrightarrow \varphi(U) \times \mathbb{R}^m \subset \mathbb{R}^{2m}, \qquad (3.2.18) \\ T^*M \supset T_q^*M \ni \alpha = \sum_{j=1}^{m} p_j \left(dx_j\right)_q &\longmapsto (\varphi(q), (p_1, \ldots, p_m)) \\ &= ((x_1(q), \ldots, x_m(q)), (p_1, \ldots, p_m)) \ . \end{aligned}$$

It is easy to see that by taking an atlas,

$$\mathcal{A} = \{(U_\lambda, \varphi_\lambda)\, ,\, \lambda \in \Lambda\} \ ,$$

on M we get a $2m$–dimensional atlas $\widehat{\mathcal{A}}$ on TM, given by,

$$\widehat{\mathcal{A}} = \left\{(\pi_{T^*}^{-1}(U_\lambda), \widehat{\varphi}_\lambda)\, ,\, \lambda \in \Lambda\right\} \ , \qquad (3.2.19)$$

where the lift of the coordinate chart map φ_λ to,

$$\widehat{\varphi}_\lambda \, : \, \pi_{T^*}^{-1}(U_\lambda) \subset T^*M \longrightarrow \mathbb{R}^{2m} \, ,$$

is defined as explained in (3.2.18).

Definition 3.2.30 (tensors on V) *Let V be a m–dimensional real vector space.*

(i) A degree r covariant tensor T on V is a r–multilinear map,

$$T \, : \, \overbrace{V \times \cdots \times V}^{r} \longrightarrow \mathbb{R} \, .$$

(ii) The vector space of all degree r covariant tensors is denoted by $T^r(V^*)$ or $T^{(r,0)}(V)$.

(iii) Given $T \in T^r(V^*)$ and $S \in T^s(V^*)$ define $T \otimes S \in T^{r+s}(V^*)$ by

$$\left(T \otimes S\right)\left(u_1, \ldots, u_r, u_{r+1}, \ldots, u_{r+s}\right) = T\left(u_1, \ldots, u_r\right)\, S\left(u_{r+1}, \ldots, u_{r+s}\right) \ ,$$

for every $u_j \in V$, $j = 1, \ldots, r+s$. This linear, associative composition of tensors is called tensor product.

Remark 3.2.31

(a) The space of covariant tensors of degree one on V coincides with the dual space,

$$T^1(V^*) = T^{(1,0)}V = V^*.$$

(b) There is a natural map,

$$\begin{array}{rcl} V & \longrightarrow & (V^*)^* \\ v & \longmapsto & i_v \,:\, (\alpha \longmapsto \alpha(v)) \,, \end{array}$$

which, for finite dimensional vector spaces V, is an isomorphism.

(c) The tensor product defines on

$$T(V^*) = \bigoplus_{r\geq 0}^{\infty} T^r(V^*) \,,$$

a structure of associative non-commutative (if $m > 1$) algebra.

◊

Example 3.2.32 *Inner products on V are examples of (symmetric, positive definite) covariant tensors of degree two on V,*

$$\begin{array}{rcl} \langle \cdot,\cdot \rangle & \in & T^2(V^*) = T^{(2,0)}(V) = V^* \otimes V^* \\ \langle \cdot,\cdot \rangle \,:\, V \times V & \longrightarrow & \mathbb{R} \end{array}$$

Definition 3.2.33 *Let V be a vector space of dimension m.*

(i) A contravariant tensor S of degree r on V is a r–multilinear map,

$$S \,:\, \overbrace{V^* \times \cdots \times V^*}^{r} \longrightarrow \mathbb{R} \,.$$

The vector space of all degree r contravariant tensors on V is denoted by $T^{(0,r)}(V)$.

(ii) A mixed tensor T of type (r, s) on V is a $r + s$–multilinear map,

$$T \,:\, \overbrace{V \times \cdots \times V}^{r} \times \overbrace{V^* \times \cdots \times V^*}^{s} \longrightarrow \mathbb{R} \,.$$

The vector space of all tensors of type (r, s) is denoted by $T^{(r,s)}(V)$.

Proposition 3.2.34 *Given an ordered basis $B_1 = (v_1, \ldots, v_m)$ of M and a basis $B_2 = (\alpha_1, \ldots, \alpha_m)$ of V^* one obtains a basis on $T^{(r,s)}(M)$ given by,*

$$\{\alpha_{j_1} \otimes \cdots \otimes \alpha_{j_r} \otimes v_{i_1} \otimes \cdots \otimes v_{i_s}\}_{j_1,\ldots,j_r,i_1,\ldots,i_s=1}^{m}$$

Remark 3.2.35 The proof of the proposition 3.2.34 is similar to the proof of the proposition 3.2.25. ◊

Corollary 3.2.36

(a) $\dim(T^{(r,s)}(M)) = m^{r+s}$.

(b) For the spaces of tensors of type (r, s) we have,

$$T^{(r,s)}(V) = \overbrace{V^* \otimes \cdots \otimes V^*}^{r} \otimes \overbrace{V \otimes \cdots \otimes V}^{s} .$$

Back to manifolds let us define mixed tensors of type (r, s) and mixed tensor fields.

Definition 3.2.37 *Let M be a m–dimensional manifold and $q \in M$.*

(i) The space of tensors of type (r, s) at the point q is

$$T_q^{(r,s)}M := T^{(r,s)}\left(T_qM\right) = \overbrace{T_q^*M \otimes \cdots \otimes T_q^*M}^{r} \otimes \overbrace{T_qM \otimes \cdots \otimes T_qM}^{s}$$

(ii) Define tensor bundle of type (r, s) on M as the disjoint union

$$T^{(r,s)}M = \bigsqcup_{q \in M} T_q^{(r,s)}M ,$$

with differentiable structure defined analogously to (3.2.9), (3.2.10) and (3.2.18), (3.2.19).

(iii) Tensor fields T of type (r, s) on M are differentiable sections of the tensor bundle $T^{(r,s)}M$, i.e. differentiable maps

$$T : M \longrightarrow T^{(r,s)}M ,$$

such that

$$\pi^{(r,s)} \circ T = \mathrm{Id}_M ,$$

where $\pi^{(r,s)}$ denotes the canonical projection from $T^{(r,s)}M$ to M. The vector space of all tensor fields on M of type (r, s) is denoted by $\mathcal{T}^{(r,s)}(M)$.

Example 3.2.38 *Particular cases of spaces of tensor fields, which we studied before are,*

$$\mathcal{T}^{(0,0)}(M) = C^\infty(M) , \mathcal{T}^{(0,1)}(M) = \mathcal{X}(M) , \ \mathcal{T}^{(1,0)}(M) = \Omega^1(M) .$$

<u>Pushforward vs pullback</u>

Given a differentiable map $f : M \longrightarrow N$, $\dim(M) = m$, $\dim(N) = n$, we have seen that, points, curves and vectors go naturally forward (i.e. in the same direction of the map f)

$$\begin{array}{ccccc} M & \ni & q \longmapsto & f(q) & \in & N \\ & & c \longmapsto & f \circ c & & \\ T_qM & \ni & v = \dfrac{dc}{dt}(0) \longmapsto & (f_*)_q(v) = \dfrac{d}{dt}\left(f \circ c\right)(0) & \in & T_{f(q)}N. \end{array}$$

while functions "go naturally" in the opposite direction i.e. are "pulled back",

$$C^\infty(N) \quad \ni \quad h \longmapsto h \circ f \quad \in \quad C^\infty(M) .$$

Like functions covariant tensors are also naturally pulled back,

$$T^{(r,0)}_{f(q)}N \;\ni S_{f(q)} \longmapsto (f^*)_q(S_{f(q)}) \;\in\; T^{(r,0)}_q M\,,$$

where, by definition, the pullback to $q \in M$ of a covariant tensor at $f(q) \in N$ acts on a r–tuple of vectors from T_qM as the original tensor at the point $f(q) \in N$ acts on the pushforward of the vectors,

$$(f^*)_q(S_{f(q)})(v_1, \ldots, v_r) := S_{f(q)}((f_*)_q(v_1), \ldots, (f_*)_q(v_r))\,, \qquad \forall v_1, \ldots, v_r \in T_qM\,.$$

Remark 3.2.39 One significative "advantage" of covariant tensor fields over contravariant tensor fields is that the pullback of a covariant tensor field on N is always a covariant tensor field on M, while for the pushforward of a contravariant tensor field on M to define a contravariant tensor field on N the map f has to be surjective (necessary condition) and usually (depends on the symmetries of the tensor field) also injective (f bijective is a sufficient condition). $\diamondsuit$

Given the above remark it is fortunate that many areas of geometry focus on studying properties of certain covariant tensor fields. The remark means that those tensor fields induce tensor fields of the same type, e.g. on submanifolds. Of course some additional properties like non-degeneracy may not be preserved under pullback.

Definition 3.2.40

(a) A Riemannian (Lorentzian) metric tensor is a tensor field, $g \in \mathcal{T}^{(2,0)}(M)$, which is symmetric,

$$g(X, Y) = g(Y, X)\,,$$

for every $X, Y \in \mathcal{X}(M)$, and defines, for every point $q \in M$, a positive definite (non-degenerate) inner product on T_qM,

$$\begin{array}{rcl} g_q(\cdot,\cdot) \,:\, T_qM \times T_qM & \longrightarrow & \mathbb{R} \\ (u, v) & \longmapsto & g_q(u, v)\,. \end{array}$$

(in the Lorentzian case the metric has signature $(-, +, \ldots, +)$).

(b) A Riemannian (Lorentzian) manifold is a pair (M, g), where M is a manifold and g is a Riemannian (Lorentzian) metric.

Remark 3.2.41 For simplicity of the notation we will introduce a symmetrized tensor product for symmetric covariant tensors and tensor fields. Let us first define a symmetrization of tensor fields, $T \in \mathcal{T}^{(r,0)}(M)$

$$\mathrm{Sym}(T) = \frac{1}{r!} \sum_{\sigma \in S_r} T \circ \sigma$$

i.e.

$$\mathrm{Sym}(T)(Y_1, \ldots, Y_r) = \frac{1}{r!} \sum_{\sigma \in S_r} T(Y_{\sigma(1)}, \ldots, Y_{\sigma(r)})\,.$$

Then the symmetrized tensor product is defined as follows. Let S, T be two symmetric covariant tensor fields. Define their symmetrized tensor product has

$$S \cdot T := \mathrm{Sym}(S \otimes T)\,.$$

For example for differential 1–forms α, β,

$$\alpha \cdot \beta = \mathrm{Sym}(\alpha \otimes \beta) = \frac{1}{2}\left(\alpha \otimes \beta + \beta \otimes \alpha\right)\,. \tag{3.2.20}$$

◊

Example 3.2.42 *Euclidean space is the Riemannian manifold* $(M, g) = (\mathbb{R}^m, g_{\mathbb{R}^m})$*, where* $g_{\mathbb{R}^m}$ *is the standard flat metric on* $\mathbb{R}^m$*,*

$$g_{\mathbb{R}^m} = dx_1^2 + \cdots + dx_m^2 = \sum_{k=1}^{m} dx_k^2\,,$$

where we have used the symmetrized tensor product introduced in the remark 3.2.41,

$$\begin{aligned} dx_j^2 &= dx_j \cdot dx_j = dx_j \otimes dx_j \\ dx_j dx_k &= \frac{1}{2}\left(dx_j \otimes dx_k + dx_k \otimes dx_j\right)\,. \end{aligned}$$

The inner product of two vectors, $u, v \in T_q\mathbb{R}^m$*,*

$$u = \sum_{j=1}^{m} u_j \left(\frac{\partial}{\partial x_j}\right)_q\,, \quad v = \sum_{i=1}^{m} v_i \left(\frac{\partial}{\partial x_i}\right)_q$$

is

$$\begin{aligned} \left(g_{\mathbb{R}^m}\right)_q(u, v) &= \sum_{k,i,j=1}^{m} u_j v_i \, (dx_k \cdot dx_k)_q \left(\left(\frac{\partial}{\partial x_j}\right)_q, \left(\frac{\partial}{\partial x_i}\right)_q\right) \\ &= \sum_{k,i,j=1}^{m} u_j v_i \, \delta_{kj}\, \delta_{ki} \\ &= \sum_{k=1}^{m} u_k v_k\,. \end{aligned}$$

Example 3.2.43 *The* m*–dimensional Minkowski space–time is the Lorentzian manifold* $(M, g) = \mathbb{R}^{1,m-1} := (\mathbb{R}^m, \eta_m)$*, where* η_m *is the flat Lorentzian metric,*

$$\eta_m = -dx_0^2 + dx_1^2 + \cdots + dx_{m-1}^2\,.$$

A very big source of interesting Riemannian metrics corresponds to taking the pullback of metrics to submanifolds of Riemannian submanifolds,

$$\begin{aligned} i : N &\hookrightarrow M \\ g_N &= i^* g\,, \end{aligned}$$

as the pullback of a Riemannian metric with respect to an injective map is always a Riemannian metric.

For Lorentzian metrics the situation is much less favourable has even for injective maps the pullback of a Lorentzian metric may be degenerate. For example the following result could be considered as somewhat of a problem for the physics theory aiming at unifying all fundamental interactions called string theory.

Theorem 3.2.44 *Let $X : \Sigma \longrightarrow M$ be an injective map from a 2–dimensional manifold without boundary Σ to a Lorentzian space–time manifold (M, g). Then, the pullback, $X^*(g)$, of g with respect to X is always degenerate if the genus of Σ is $h \geq 1$.*

Besides being always well defined on tensor fields the pullback of a tensor field is also very easy to calculate.

Proposition 3.2.45 *Let $f : M \longrightarrow N$ be a differentiable let $h \in C^\infty(f(q))$. The pullback commutes with the exterior derivative, i.e.*

$$f^*(dh) = d(f^*(h)) = d(h \circ f)\,.$$

Proof.
■

Remark 3.2.46 Let $\widehat{f}$ be the local coordinate representative of the map $f : M \longrightarrow N$ like in the definition 3.2.6,

$$\begin{aligned} \widehat{f} &= \psi \circ f \circ \varphi^{-1} : \mathbb{R}^m \supset \varphi(U) \longrightarrow \psi(V) \subset \mathbb{R}^n \\ \widehat{f} : & \quad \begin{cases} y_1 &= \widehat{f}_1(x_1, \ldots, x_m) \\ & \vdots \\ y_n &= \widehat{f}_n(x_1, \ldots, x_m)\ , \end{cases} \end{aligned}$$

By choosing in the proposition $h = y_j$ we obtain

$$f^*((dy_j)_{f(q)}) = \left(d\widehat{f}_j(x)\right)_q = \sum_{k=1}^{m} \frac{\partial \widehat{f}_j}{\partial x_k}(x)\,(dx_k)_q\ , \tag{3.2.21}$$

where $x = \varphi(q)$, $y = \psi(f(q))$. ◊

Exercise 3.2.5 *Find, using cylindrical coordinates, the metric induced on a dense open subset of the paraboloid,*

$$S = \left\{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2\right\}\,.$$

Solution. Let i denote the embedding of the parabolid,

$$\begin{aligned} i : S = M &\hookrightarrow \mathbb{R}^3 = N \\ i(q) &= q\,, \end{aligned}$$

so that $i = \mathrm{Id}_{\mathbb{R}^3|_S}$. On $\mathbb{R}^3$ we choose the standard global Cartesian coordinate chart, $(\mathbb{R}^3, \psi = \mathrm{Id}_{\mathbb{R}^3})$ and on S we use a cylindrical coordinate chart, $(U = S \setminus \{y = 0,\, x \geq 0\}, \varphi)$, where

$$\varphi^{-1} : \begin{cases} x &= \rho\cos(\theta) \\ y &= \rho\sin(\theta) \\ z &= \rho^2 , \qquad\qquad \rho > 0\,,\; 0 < \theta < 2\pi. \end{cases}$$

The local expression for i is

$$\begin{aligned} \widehat{i} &= \psi \circ i \circ \varphi^{-1} = \varphi^{-1} \\ \widehat{i}(\rho, \theta) &= \left(\varphi^{-1}\right)(\rho, \theta) \\ &= \left(\rho\cos(\theta), \rho\sin(\theta), \rho^2\right). \end{aligned}$$

From (3.2.21) we obtain

$$\begin{aligned} g_S &= i^*(g_{\mathbb{R}^3}) = (d(\rho\cos(\theta)))^2 + (d(\rho\sin(\theta)))^2 + (d(\rho^2))^2 \\ &= (\cos(\theta)d\rho - \rho\sin(\theta)d\theta)^2 + (\sin(\theta)d\rho + \rho\cos(\theta)d\theta)^2 + (2\rho d\rho)^2 \\ &= \left(1 + 4\rho^2\right)\, d\rho^2 + \rho^2\, d\theta^2 . \end{aligned}$$

■

Length of a curve:

Suppose that an unparametrized curve C is a 1–dimensional submanifold, $C \subset M$ and suppose that it has a global chart

$$\widetilde{\varphi} \,:\, C \longrightarrow \mathbb{R},$$

with inverse $c = \widetilde{\varphi}^{-1}$,

$$c \,:\, (-T, T) \longrightarrow C \subset N.$$

First of all recall from (3.2.5) that the local coordinate vector fields on a chart (U, ψ) of M have the form

$$\left(\frac{\partial}{\partial x_j}\right)_q = \frac{d}{dt} c^{(j)}(t)|_{t=0} \,,$$

where,

$$\begin{aligned} c^{(j)}(s) &= \varphi^{-1}(x_1, \ldots, x_j + s, \ldots, x_m) \\ &= \left(\varphi^{-1} \circ \hat{c}^{(j)}(s)\right)(s) \,. \end{aligned}$$

For the curve we have $q = c(t)$ and one global coordinate s,

$$\tilde{c}^{(1)}(s) = \tilde{\varphi}^{-1}(t + s) = c(t + s) \,,$$

so that c is a coordinate curve and therefore,

$$\left(\frac{\partial}{\partial t}\right)_q = \frac{d}{ds} c(t + s)|_{s=0} = \frac{d}{dt} c(t) = (c_*)_t \left(\frac{\partial}{\partial t}\right) .$$

The pullback of the metric g on M to the interval $(-T, T)$ gives,

$$c^*(g_{c(t)}) = (c^* g)_{11}(t)\, dt^2,$$

where

$$\begin{aligned}(c^*g)_{11}(t) &= (c^*g)_t\left(\frac{\partial}{\partial t},\frac{\partial}{\partial t}\right) = g_{c(t)}\left(c_*\frac{\partial}{\partial t},c_*\frac{\partial}{\partial t}\right)\\ &= g_{c(t)}\left(\frac{dc}{dt}(t),\frac{dc}{dt}(t)\right) = ||\dot{c}(t)||^2_{g_{c(t)}}\,.\end{aligned}$$

Definition 3.2.47 *In the above notation the length of the curve C is defined to be*

$$\begin{aligned}\ell_g(C) &= \int_T^T ||\dot{c}(t)||_{g_{c(t)}}\,dt\\ &= \int_T^T \sqrt{g_{c(t)}\left(\frac{dc}{dt}(t),\frac{dc}{dt}(t)\right)}\,dt\\ &= \int_T^T \sqrt{(c^*g)_t\left(\frac{\partial}{\partial t},\frac{\partial}{\partial t}\right)}\,dt\,.\end{aligned}$$

Exercise 3.2.6 *Let S be the paraboloid of exercise 3.2.5. Find the length of the circle*

$$D = S\cap\{z=2\}\,.$$

Solution. In the cylindrical coordinates of S, let $\widetilde{D} = D\setminus\left\{(\sqrt{2},0,2)\right\}$. We find,

$$\widetilde{D}\,,\; c=\widetilde{\varphi}^{-1}\,:\,\left\{\begin{array}{lcl} x &=& \sqrt{2}\,\cos(t)\\ y &=& \sqrt{2}\,\sin(t)\\ z &=& 2\,,\qquad 0<t<2\pi\,.\end{array}\right. \quad\Leftrightarrow\quad \hat{c}\,:\,\left\{\begin{array}{lcl}\rho &=& \sqrt{2}\\ \theta &=& t\,,\qquad 0<t<2\pi\,.\end{array}\right.$$

Then,

$$\begin{aligned}(c^*g)_t &= g_{c(t)}(\dot{c}(t),\dot{c}(t))\,dt^2\\ &= \hat{c}^*\left(\left(1+4\rho^2\right)\,d\rho^2+\rho^2 d\theta^2\right)\\ &= 2\,dt^2 = (c^*g)_{11}(t)\,dt^2\,,\end{aligned}$$

and

$$\ell_g(D) = \ell_g(\widetilde{D}) = \int_0^{2\pi}\sqrt{2}\,dt = 2\sqrt{2}\,\pi\,.$$

■

Symmetries of tensor fields.

The group Diff(M) acts on M,

$$\mathrm{Diff}(M)\;\curvearrowright\;M$$

and this action induces infinite dimensional representations (linear actions) on the spaces of tensor fields,

$$\begin{array}{lcl}\varphi &\longmapsto& \varphi_*\quad\in\,L(\mathcal{T}^{(0,r)}(M))\\ \varphi &\longmapsto& \varphi_* := (\varphi^{-1})^*\,\in\,L(\mathcal{T}^{(r,0)}(M))\\ \varphi &\longmapsto& \varphi_* := (\varphi^{-1})^*\otimes\varphi_*\,\in\,L(\mathcal{T}^{(r,s)}(M))\end{array}\tag{3.2.22}$$

so that indeed,

$$\mathrm{Diff}(M)\curvearrowright\mathcal{T}^{(r,s)}(M)\,.$$

Definition 3.2.48 (symmetry group of a tensor field) *Let $T \in \mathcal{T}^{(r,s)}(M)$. We say that $\varphi \in \mathrm{Diff}(M)$ is a symmetry of the tensor field T if $\varphi_* T = T$ or, equivalently, if $\varphi \in (\mathrm{Diff}(M))_T$. The stabilizer of T is the group of all symmetries of T,*

$$S(T) := (\mathrm{Diff}(M))_T \,.$$

The symmetry group of a function is always infinite dimensional and easy to find. Indeed,

$$\begin{aligned} \varphi &\in S(f) \\ \Leftrightarrow \varphi_* f &= f \Leftrightarrow \varphi^* f = f \Leftrightarrow f = f \circ \varphi \\ \Leftrightarrow f(q) &= f(\varphi(q))\,, \qquad \forall q \in M \,. \end{aligned}$$

So $\varphi \in S(T)$ if and only if φ preserves the level sets of f. For higher degree tensor fields the symmetries are usually less obvious to find. Let $T \in \mathcal{T}^{(r,s)}(M)$. Then,

$$\begin{aligned} \varphi &\in S(T) \\ \Leftrightarrow \varphi_*(T_q) &= T_{\varphi(q)} \,, \qquad \forall q \in M \,. \end{aligned}$$

Lie derivative.
Let ψ_t^Y be the local flow of a vector field Y. Then it is natural to extend the concept of directional derivative in the direction of Y to tensor fields. Recall that for functions,

$$\begin{aligned} Y_p(f) &= \frac{d}{dt}|_{t=0} \left(f \circ \psi_t^Y(q)\right) = \lim_{t\to 0} \frac{\left(\psi_t^Y\right)^* (f)(q) - f(q)}{t} \\ &= \lim_{t\to 0} \frac{f(\psi_t^Y(q)) - f(q)}{t} \,. \end{aligned}$$

Let T be a covariant tensor field of type $(r, 0)$. By analogy with the functions we define,

$$\begin{aligned} \left(\mathcal{L}_Y(T)\right)_q &= \lim_{t\to 0} \frac{\left(\left(\psi_t^Y\right)^* T\right)_q - T_q}{t} = \lim_{t\to 0} \frac{\left(\left(\psi_t^Y\right)^* T\right)_q - T_q}{t} \\ &= \lim_{t\to 0} \frac{\left(\psi_t^Y\right)^* (T_{\psi_t^Y(q)}) - T_q}{t} \,. \end{aligned} \tag{3.2.23}$$

This linear transformation is called Lie derivative of T in the direction of Y. For a general tensor field T of type (r, s) we define its Lie derivative in the direction of Y by,

$$\left(\mathcal{L}_Y(T)\right)_q = \lim_{t\to 0} \frac{\left(\psi_{-t}^Y\right)_* (T_{\psi_t^Y(q)}) - T_q}{t} \,. \tag{3.2.24}$$

Let Y be a complete vector field, $Y \in \mathcal{X}_c(M)$. From the definition of Lie derivative in (3.2.24) we see that if the diffeomorphisms from the flow of Y are symmetries of T then the Lie derivative of T in the direction of Y is zero,

$$\psi^Y = \left\{\psi_t^Y \,,\ t \in \mathbb{R}\right\} \subset S(T) \ \Rightarrow\ \mathcal{L}_Y T = 0 \,. \tag{3.2.25}$$

Complete vector fields satisfying (3.2.25) are called infinitesimal symmetries of T,

$$\mathfrak{s}(T) := \{Y \in \mathcal{X}_c(M) \,:\, \mathcal{L}_Y(T) = 0\} \,.$$

It turns out that the reverse implication in (3.2.25) is also valid, i.e. an infinitesimal symmetry generates symmetries of the tensor field.

Proposition 3.2.49 (infinitesimal symmetries are equivalent to flow symmetries) *Let $T \in \mathcal{T}^{(r,s)}$ and $Y \in \mathcal{X}_c(M)$. Then*

$$Y \in \mathfrak{s}(T) \;\Leftrightarrow\; \psi^Y = \left\{\psi_t^Y \,, t \in \mathbb{R}\right\} \subset S(T)\,.$$

Isometry groups and Killing vector fields.

In Riemannian and Lorentzian geometry the group of symmetries of the metric is called isometry group of (M, g),

$$\mathrm{Iso}(M,g) = S(g) = \left(\mathrm{Diff}(M)\right)_g = \{\varphi \in \mathrm{Diff}(M) \,:\, \varphi^*(g) = g\}\,. \tag{3.2.26}$$

Unlike the case of functions the isometry groups are always finite dimensional. In fact,

$$\dim(S(g)) = \dim(\mathfrak{s}(g)) \leq \frac{m(m+1)}{2}.$$

Infinitesimal symmetries of a metric g are called Killing vector fields of g,

$$\mathrm{Kill}(g) = \mathfrak{s}(g) = \{Y \in \mathcal{X}_c(M) \,:\, \mathcal{L}_Y(g) = 0\}\,.$$

Proposition 3.2.50 (properties of the Lie derivative) *The Lie derivative is a map from $\mathcal{X}(M)$ to linear transformations of the spaces of tensor fields,*

$$\mathcal{L} \,:\, \mathcal{X}(M) \longrightarrow \mathcal{T}^{(r,s)}(M)\,.$$

This map satisfies the following properties,

P1. The map $\mathcal{L}$ is a representation of the Lie algebra $\mathcal{X}(M)$ on the vector space $\mathcal{T}^{(r,s)}(M)$, i.e. is a Lie algebra homomorphism (see definition 3.2.15), to $L(\mathcal{T}^{(r,s)}(M))$,

$$\mathcal{L}_{[X,Y]} = \mathcal{L}_x \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X\,.$$

P2. Let $Y \in \mathcal{X}(M)$ and $f \in C^\infty(M)$. Then,

$$\mathcal{L}_X(f) = X(f)\,.$$

P3. Let $X, Y \in \mathcal{X}(M)$. Then,

$$\mathcal{L}_X Y = [X,Y]\,.$$

P4. $\mathcal{L}_Y$ is a derivation of the algebra of tensor fields, i.e. it satisfies Leibnitz identity,

$$\mathcal{L}_Y\left(T_1 \otimes T_2\right) = \mathcal{L}_Y\left(T_1\right) \otimes T_2 + T_1 \otimes \mathcal{L}_Y\left(T_2\right)\,, \qquad \forall\, T_1, T_2 \in \mathcal{T}(M)\,.$$

P5. $\mathcal{L}_Y$ commutes with all contractions, i.e. $C^\infty(M)$–linear maps

$$c \,:\, \mathcal{T}^{(r,s)}(M) \longrightarrow \mathcal{T}^{(r-p,s-p)}(M)\,,$$

$$\mathcal{L}_Y \circ c = c \circ \mathcal{L}_Y\,.$$

All contractions can be constructed from the basic one,

$$\begin{array}{rcl} c_{(1,1)} \,:\, \mathcal{T}^{(1,1)}(M) & \longrightarrow & \mathcal{T}^{(0,0)}(M) = C^\infty(M) \\ \alpha \otimes Y & \longmapsto & \alpha(Y)\,, \end{array}$$

by tensoring it with itself and with identity maps.

As we will illustrate now for 1–forms, the action of the Lie derivative on functions and vector fields (P2 and P3) plus the properties P4 and P5 completely determine the Lie derivative on tensor fields of all types.

Exercise 3.2.7 *Let $\alpha \in \Omega^1(M), X, Y \in \mathcal{X}(M)$. Prove that,*

$$\left(\mathcal{L}_X(\alpha)\right)(Y) = X\left(\alpha(Y)\right) - \alpha([X,Y])\,. \tag{3.2.27}$$

Solution. We have,

$$\begin{aligned} X\left(\alpha(Y)\right) &= \mathcal{L}_X\left(c(\alpha\otimes Y)\right) = \left(\mathcal{L}_X \circ c\right)(\alpha\otimes Y) \\ &= \left(c\circ\mathcal{L}_X\right)(\alpha\otimes Y) = c\left(\mathcal{L}_X\left(\alpha\otimes Y\right)\right) \\ &= c\left(\mathcal{L}_X\left(\alpha\right)\otimes Y + \alpha\otimes[X,Y]\right) \\ &= \left(\mathcal{L}_X\left(\alpha\right)\right)(Y) + \alpha([X,Y])\,, \end{aligned}$$

and therefore,

$$\left(\mathcal{L}_X(\alpha)\right)(Y) = X\left(\alpha(Y)\right) - \alpha([X,Y])\,.$$

■

Exercise 3.2.8 *Show that for exact 1–forms, i.e. $\alpha = df\,,\ f\in C^\infty(M)$, one has,*

$$\mathcal{L}_X(df) = d\left(X(f)\right)\,. \tag{3.2.28}$$

Solution. From (3.2.27) we have,

$$\begin{aligned} \left(\mathcal{L}_X(df)\right)(Y) &= X\left(Y(f))\right) - df([X,Y]) \\ &= X\left(Y(f)\right) - [X,Y](f) \\ &= X\left(Y(f)\right) - \left(X\left(Y(f)\right) - Y\left(X(f)\right)\right) = Y\left(X(f)\right) \\ &= d\left(X(f)\right)(Y)\,, \qquad \forall Y\in\mathcal{X}(M)\,, \end{aligned}$$

so that indeed

$$\mathcal{L}_X(df) = d\left(X(f)\right)\,.$$

■

Exercise 3.2.9 *Consider the metric g_S,*

$$g_S = (1+4\rho^2)d\rho^2 + \rho^2 d\theta^2\,,$$

on the paraboloid S of revolution around the axis $\mathcal{O}z$,

$$S = \left\{(x,y,z)\in\mathbb{R}^3\,:\, z = x^2+y^2\right\}\,,$$

studied in the exercise 3.2.5. Show that $Y = \frac{\partial}{\partial\theta}$ is a Killing vector field of this metric, i.e.,

$$\frac{\partial}{\partial\theta}\in\mathfrak{s}(g_S) \Leftrightarrow \mathcal{L}_{\frac{\partial}{\partial\theta}}\, g_S = 0\,.$$

Solution. Let $f \in C^\infty(S)$. From exercise 3.2.8 we know that

$$\mathcal{L}_{\frac{\partial}{\partial\theta}} df = d\left(\frac{\partial f}{\partial \theta}\right) .$$

From proposition 3.2.50, P4, and (3.2.20) we obtain,

$$\mathcal{L}_{\frac{\partial}{\partial\theta}}\, df\, dh = d\left(\frac{\partial f}{\partial \theta}\right) dh + df\, d\left(\frac{\partial h}{\partial \theta}\right) ,$$

for every $f, h \in C^\infty(S)$. Then, since

$$\frac{\partial}{\partial\theta} F(\rho) = 0 \,,\; \frac{\partial \theta}{\partial\theta} = 1 \,,$$

$$\mathcal{L}_{\frac{\partial}{\partial\theta}}\, d\rho^2 = 2d\left(\frac{\partial \rho}{\partial \theta}\right) d\rho = 0 \,,$$

$$\mathcal{L}_{\frac{\partial}{\partial\theta}}\, d\theta^2 = 2d\left(\frac{\partial \theta}{\partial \theta}\right) d\theta = 0 \,,$$

we conclude that

$$\begin{array}{rcl} \mathcal{L}_{\frac{\partial}{\partial\theta}}\, g_S & = & \mathcal{L}_{\frac{\partial}{\partial\theta}} \left((1+4\rho^2)d\rho^2 + \rho^2 d\theta^2\right) \\ & = & 0 \,. \end{array}$$

■

Remark 3.2.51 The vector field $\frac{\partial}{\partial\theta}$ is a Killing vector field of g_S because its flow corresponds to orthogonal rotations around the axis $\mathcal{O}z$, which are isometries of the Euclidean metric, $g_{\mathbb{R}^3}$, that act on S and therefore are isometries of $g_S = i^*(g_{\mathbb{R}^3})$.

◊

Exercise 3.2.10 *Show that if* $Y = \sum_{j=1}^m a_j(x)\frac{\partial}{\partial x_j}$, *then*

$$\mathcal{L}_Y\,(dx_k) = da_k \,.$$

Solution. From (3.2.28) and noticing that

$$Y(x_k) = a_k \,,$$

we obtain

$$\mathcal{L}_Y\,(dx_k) = d\,(Y(x_k)) = da_k \,.$$

■

A covariant tensor field $T \in \mathcal{T}^{(r,0)}(M)$ is called alternating if, as a $C^\infty(M)$ r–multilinear map,

$$\begin{array}{rcl} T \,:\, \mathcal{X}(M)\times\cdots\times\mathcal{X}(M) & \longrightarrow & C^\infty(M) \\ (Y_1,\ldots,Y_r) & \longmapsto & T(Y_1,\ldots,Y_r) \end{array}$$

satisfies the following property,

$$T\left(Y_1,\ldots,Y_i,\ldots,Y_j,\ldots,Y_r\right) = -\,T\left(Y_1,\ldots,Y_j,\ldots,Y_i,\ldots,Y_r\right)\ ,$$

for every $i,j = 1,\ldots,m$ and vector fields $Y_j\ ,\ j = 1,\ldots,m$. An alternating tensor field of degree r is also called a differential r–form. The space of all differential r–forms is denoted by $\Omega^r(M)$. The alternating property implies that $\Omega^r(M) = \{0\}$ if $r > m = \dim(M)$. A projection from $\mathcal{T}^{(r,0)}(M)$ to $\Omega^r(M)$ is given by,

$$\begin{aligned} \mathrm{Alt}\ :\ \mathcal{T}^{(r,0)}(M) &\longrightarrow \Omega^r(M) \\ T &\longmapsto \mathrm{Alt}(T) = \frac{1}{r!}\sum_{\sigma\in S_r}\mathrm{sgn}(\sigma)\,T\circ\sigma\ . \end{aligned}$$

Example 3.2.52

(i) Let $T \in \mathcal{T}^{(2,0)}(M)$*. Then,*

$$\left(\mathrm{Alt}(T)\right)(Y_1,Y_2) = \frac{1}{2}\left(T(Y_1,Y_2) - T(Y_2,Y_1)\right)\ .$$

(ii) Let $S \in \mathcal{T}^{(3,0)}(M)$*. Then,*

$$\begin{aligned} \left(\mathrm{Alt}(T)\right)(Y_1,Y_2,Y_3) &= \frac{1}{6}\left(T(Y_1,Y_2,Y_3) - T(Y_2,Y_1,Y_3) - T(Y_1,Y_3,Y_2) - T(Y_2,Y_1,Y_3)\right) \\ &+ \frac{1}{6}\left(T(Y_2,Y_3,Y_1) + T(Y_3,Y_1,Y_2)\right)\ . \end{aligned}$$

Definition 3.2.53 *Let* $\alpha \in \Omega^r(M)$ *and* $\beta \in \Omega^s(M)$*. Their wedge product is defined to be the following* $r+s$ *differential form,*

$$\alpha\wedge\beta = \frac{(r+s)!}{r!s!}\,\mathrm{Alt}\left(\alpha\otimes\beta\right)\ .$$

Example 3.2.54

(i) Let $\alpha,\beta \in \Omega^1(M)$*. Then,*

$$\alpha\wedge\beta = \frac{2}{1}\,(\mathrm{Alt}(\alpha\otimes\beta) = 2\,\frac{1}{2}\,(\alpha\otimes\beta - \beta\otimes\alpha) = \alpha\otimes\beta - \beta\otimes\alpha\,.$$

or, equivalently,

$$\left(\alpha\wedge\beta\right)(X,Y) = \alpha(X)\beta(Y) - \beta(X)\alpha(Y) = \left|\begin{array}{cc} \alpha(X) & \alpha(Y) \\ \beta(X) & \beta(Y) \end{array}\right|\ . \qquad (3.2.29)$$

(ii) Let $\alpha_j \in \Omega^1(M)\,, j = 1,\ldots,r$*. Then, one can show that the generalization of (3.2.29) for* $r > 2$ *reads*

$$\left(\alpha_1\wedge\cdots\wedge\alpha_r\right)(Y_1,\ldots,Y_r) = \det\left(\alpha_j(Y_k)\right) = \left|\begin{array}{ccc} \alpha_1(Y_1) & \ldots & \alpha_1(Y_r) \\ \vdots & \ddots & \vdots \\ \alpha_r(Y_1) & \ldots & \alpha_r(Y_r) \end{array}\right|.$$

If $\alpha \in \Omega^r(M)\,,\ \beta \in \Omega^s(M)$ then, one shows easily that,

$$\alpha \wedge \beta = (-1)^{rs}\,\beta \wedge \alpha\,.$$

In particular, for all forms of odd degree,

$$\alpha \wedge \alpha = 0\,.$$

Definition 3.2.55 *The exterior derivative of differential forms is the following linear transformation,*

$$\begin{array}{rcl} d\,:\,\Omega^r(M) & \longrightarrow & \Omega^{r+1}(M) \\ d\left(\displaystyle\sum_{j_1\ldots j_r=1}^{m} a_{j_1\ldots j_r} dx_{j_1}\wedge\cdots\wedge dx_{j_r}\right) & = & \displaystyle\sum_{j_1\ldots j_r=1}^{m} d(a_{j_1\ldots j_r})\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_r} \end{array}$$

for every $r = 0, 1, \ldots, m$.

Proposition 3.2.56 *The exterior derivative satisfies the following properties.*

1. *The exterior derivative on functions coincides with the previous definition (3.2.17).*
2. $d \circ d = 0$.
3. *If $\alpha \in \Omega^r(M)\,,\ \beta \in \Omega^s(M)$, we have*

$$d\,(\alpha \wedge \beta) = (d\alpha)\wedge\beta + (-1)^r\,\alpha\wedge d\beta\,.$$

Exercise 3.2.11 *Let $f = f_1\frac{\partial}{\partial x} + f_2\frac{\partial}{\partial y} + f_3\frac{\partial}{\partial z}$. Relate $\mathrm{rot}(f)$ and $\mathrm{div}(f)$ with the exterior derivative of differential forms on $\mathbb{R}^3$ by using the following two isomorphisms,*

$$\mathcal{X}(\mathbb{R}^3)\ \ni\ f\ \longmapsto \omega_f := f_1dx + f_2dy + f_3dz\ \in \Omega^1(M)$$

and

$$\mathcal{X}(\mathbb{R}^3)\ \ni\ f\ \longmapsto \alpha_f := f_1dy\wedge dz + f_2dz\wedge dx + f_3dx\wedge dy\ \in \Omega^2(M)$$

show that

(a) $d\omega_f = \alpha_{\mathrm{rot}(f)}$.

(b) $d\alpha_f = div(f)\,dx\wedge dy\wedge dz$.

Solution. Let us only prove that $d\omega_f = \alpha_{rotf}$. We have,

$$\begin{array}{rcl} d\omega_f & = & df_1\wedge dx + df_2\wedge dy + df_3\wedge dz \\ & = & \left(\dfrac{\partial f_1}{\partial x}dx + \dfrac{\partial f_1}{\partial y}dy + \dfrac{\partial f_1}{\partial z}dz\right)\wedge dx \\ & + & \left(\dfrac{\partial f_2}{\partial x}dx + \dfrac{\partial f_2}{\partial y}dy + \dfrac{\partial f_2}{\partial z}dz\right)\wedge dy \\ & + & \left(\dfrac{\partial f_3}{\partial x}dx + \dfrac{\partial f_3}{\partial y}dy + \dfrac{\partial f_3}{\partial z}dz\right)\wedge dz \\ & = & \left(\dfrac{\partial f_3}{\partial y} - \dfrac{\partial f_2}{\partial z}\right)dy\wedge dz + \left(\dfrac{\partial f_1}{\partial z} - \dfrac{\partial f_3}{\partial x}\right)dz\wedge dx + \left(\dfrac{\partial f_2}{\partial x} - \dfrac{\partial f_1}{\partial y}\right)dx\wedge dy \\ & = & \alpha_{\mathrm{rot}(f)}\,. \end{array}$$

■

Definition 3.2.57 *A 2–form $\omega \in \Omega^2(M)$ is called a symplectic form if it is closed, i.e. $d\omega = 0$, and non-degenerate, i.e.*

$$\omega_p(u,v) = 0\,, \qquad \forall v \in T_pM \Leftrightarrow u = 0\,.$$

A symplectic manifold is a pair (M,ω), where ω is a symplectic form.

Example 3.2.58 *Cotangent bundles have a natural structure of symplectic manifold. Let Q be a m dimensional manifold (called configuration space in mechanics). The Liouville 1–form $\theta \in \Omega^1(T^*Q)$ is defined as follows. Let $\alpha \in T^*Q$ and $v \in T_\alpha(T^*Q)$. Then,*

$$\theta_\alpha(v) = \alpha(\pi_*(v))\,,$$

where we are denoting by π the canonical projection from T^*M to M. In the local coordinates used in (3.2.18) the Liouville form reads,

$$\theta = \sum_{j=1}^m p_j dx_j\,.$$

The canonical symplectic form on T^*Q is

$$\omega = -d\theta = \sum_{j=1}^m dx_j \wedge dp_j\,.$$

Example 3.2.59 *Let (M,ω) be a symplectic manifold. The symplectic structure leads to a very important map,*

$$\begin{array}{ccc} C^\infty(M) & \longrightarrow & \mathcal{X}(M) \\ f & \longmapsto & X_f \end{array}$$

where X_f is the vector field defined (uniquely) by

$$i_{X_f}\omega = df \Leftrightarrow \omega(X_f, Y) = df(Y) = Y(f)\,,\ \forall Y \in \mathcal{X}(M)\,.$$

This vector field is called Hamiltonian vector field corresponding to the function f. If $M = T^*Q$ with the canonical symplectic form,

$$X_h = \sum_{j=1}^m \frac{\partial h}{\partial p_j}\frac{\partial}{\partial x_j} - \frac{\partial h}{\partial x_j}\frac{\partial}{\partial p_j}\,.$$

The equations for the integral curves of X_h have, in this case, the form

$$\left\{\begin{array}{rcl} \dot{x}_j & = & \frac{\partial h}{\partial p_j} \\ \dot{p}_j & = & -\frac{\partial h}{\partial x_j}\,. \end{array}\right.$$

Definition 3.2.60 *An Hamiltonian system is a triple (M,ω,h), where (M,ω) is a symplectic manifold and h is a function, $h \in C^\infty(M)$, called Hamiltonian of the system. The Hamiltonian dynamics corresponds to the flow of the Hamiltonian vector field, X_h.*

Chapter 4

Topics of Lie Groups and Applications

4.1 Lie Groups

Definition 4.1.1 *A Lie group G is a smooth manifold and a group such that the group maps*

$$\begin{array}{rcl} G \times G & \longrightarrow & G \\ (g_1, g_2) & \longmapsto & g_1 \circ g_2 \end{array}$$

and

$$\begin{array}{rcl} G & \longrightarrow & G \\ g & \longmapsto & g^{-1} \end{array}$$

are differentiable.

A very important way of obtaining examples of Lie groups is Cartan's closed subgroup theorem.

Theorem 4.1.2 *Let G be a Lie group and $H \subset G$ a subgroup and a closed subset of G. Then H is a submanifold and a Lie group.*

Example 4.1.3

1. *The simplest m–dimensional abelian Lie group, $(\mathbb{R}^m, +)$.*

2. *The general linear group,*

$$GL_n(\mathbb{R}) = \{A \in \mathrm{Mat}_n(\mathbb{R}) \,:\, \det(A) \neq 0\}\ .$$

The vector space $\mathrm{Mat}_n(\mathbb{R})$ is isomorphic to $\mathbb{R}^{n^2}$. The group $GL_n(\mathbb{R})$ is an open subset in $\mathrm{Mat}_n(\mathbb{R})$ as is the inverse image of an open set with respect to the continuous map

$$\det \,:\, \mathrm{Mat}_n(\mathbb{R}) \longrightarrow \mathbb{R}\,.$$

Indeed we have,

$$GL_n(\mathbb{R}) = \det{}^{-1}(\mathbb{R} \setminus \{0\})\,.$$

Being an open subset in a vector space it has a natural structure of n^2–dimensional differentiable manifold (given by an atlas with a single global chart, $(U, \varphi) = (GL_n(\mathbb{R}), \mathrm{Id})$). The group operations are differentiable so that $GL_n(\mathbb{R})$ is a Lie group.

3. *The complex general linear group,*

$$GL_n(\mathbb{C}) = \{A \in \mathrm{Mat}_n(\mathbb{C}) \,:\, \det(A) \neq 0\} \subset \mathrm{Mat}_n(\mathbb{C}) \cong \mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2} \,.$$

is also the inverse image of an open set in a vector space with the associated differentiable and Lie group structure.

4. *From Cartan's theorem 4.1.2 it follows that any subgroup of the Lie groups $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$, which is a closed subset, is a Lie group.*

(4a) The real and complex orthogonal groups,

$$O_n(\mathbb{R}) = \left\{A \in GL_n(\mathbb{R}) \,:\, A^T A = \mathrm{I}_n\right\} \,,$$

and

$$O_n(\mathbb{C}) = \left\{A \in GL_n(\mathbb{C}) \,:\, A^T A = \mathrm{I}_n\right\} \,,$$

are closed subgroups of the Lie groups $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$ respectively as they are the inverse images of the closed set, $\{\mathrm{I_n}\}$, with respect to the continuous maps,

$$\begin{array}{rcl} GL_n(\mathbb{K}) & \longrightarrow & GL_n(\mathbb{K}) \\ A & \longmapsto & A^T A \,. \end{array}$$

(4b) The Lorentz group in n dimensions,

$$O(1, n-1) = \left\{A \in GL_n(\mathbb{R}) \,:\, A^T \eta_n A = \eta_n\right\} \,,$$

where η_n corresponds to the Minkowski metric and is a diagonal matrix with $(-1, 1, \ldots, 1)$ in the main diagonal. The Lorentz groups are closed subgroups of the Lie groups $GL_n(\mathbb{R})$ as they are the inverse images of the closed sets, $\{\eta_\mathrm{n}\}$, with respect to the continuous maps,

$$\begin{array}{rcl} GL_n(\mathbb{R}) & \longrightarrow & GL_n(\mathbb{R}) \\ A & \longmapsto & A^T \eta_n A \,. \end{array}$$

(4c) The unitary groups,

$$U_n = \left\{A \in GL_n(\mathbb{C}) \,:\, A^\dagger A = \mathrm{I}_n\right\} \,,$$

where † denotes Hermitian conjugation of matrices, i.e. complex conjugation and transposition, and the special unitary groups,

$$SU_n = \left\{A \in GL_n(\mathbb{C}) \,:\, A^\dagger A = \mathrm{I}_n, \ \det(A) = 1\right\} \,,$$

are closed subgroups of the Lie groups $GL_n(\mathbb{C})$. Indeed, the unitary group in n (complex) dimensions, U_n, is the inverse image of the closed set, $\{\mathrm{I_n}\}$, with respect to the continuous map,

$$\begin{array}{rcl} GL_n(\mathbb{C}) & \longrightarrow & GL_n(\mathbb{C}) \\ A & \longmapsto & A^\dagger A \,. \end{array}$$

The special unitary group in n dimensions, SU_n, is the inverse image of the closed set, $\{(\mathrm{I_n}, 1)\}$, with respect to the continuous map,

$$\begin{array}{rcl} GL_n(\mathbb{C}) & \longrightarrow & GL_n(\mathbb{C}) \times \mathbb{C}^* \\ A & \longmapsto & (A^\dagger A, \det(A))\,. \end{array}$$

(4d) The real and complex symplectic groups in $2n$ dimensions,[1]

$$Sp_n(\mathbb{K}) = \left\{ A \in GL_{2n}(\mathbb{K}) \, : \, A^T \Omega_n A = \Omega_n \right\} \, ,$$

where Ω_n is the matrix of the canonical symplectic structure in canonical coordinates,

$$\Omega_n = \left(\begin{array}{cc} 0_n & -\mathrm{I}_n \\ \mathrm{I}_n & 0_n \end{array} \right) .$$

0_n and I_n are the zero and identity $n \times n$ matrices, respectively. The symplectic groups are closed subgroups of the Lie groups $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$ respectively as they are the inverse images of the closed sets, $\{\Omega_\mathrm{n}\}$, with respect to the continuous maps,

$$\begin{array}{rcl} GL_n(\mathbb{K}) & \longrightarrow & GL_n(\mathbb{K}) \\ A & \longmapsto & A^T \Omega_n A\,. \end{array}$$

Remark 4.1.4 The dimension of the above matrix Lie groups is easier to find by finding the dimension of their Lie algebras. We will return to this, for the orthogonal and the unitary groups, in the exercise 4.1.1.

◊

Definition 4.1.5 *A (smooth) action of a Lie group G on a manifold is an action (see (2.2.14))*

$$\hat{\psi} \, : \, G \times M \longrightarrow M \, ,$$

which is a differentiable map. As in actions of groups on sets we say that G acts on M and write $G \overset{\psi}{\curvearrowright} M$.

Remark 4.1.6 For an action $G \overset{\psi}{\curvearrowright} M$ of a Lie group the bijections of M,

$$g \longmapsto \psi_g = \hat{\psi}(g, \cdot) \, ,$$

are diffeomorphisms so that the (smooth) action of G on M defines a homomorphism

$$\begin{array}{rcl} \psi \, : \, G & \longrightarrow & \mathrm{Diff}(M) \subset \mathrm{Sym}(M) \\ g & \longmapsto & \psi_g\,. \end{array}$$

◊

The theorem on transitive actions of groups on sets (Theorem 2.2.45) is strenghtened in the context of (smooth) transitive actions of Lie groups on manifolds.

[1] Some authors call this group $Sp_{2n}(\mathbb{K})$.

Theorem 4.1.7 *(see [Ar, proposition 4.2]) Let $G \overset{\psi}{\curvearrowright} M$ be a transitive action of the Lie group G on the manifold M. Then*

(a) The stabilizer G_{p_0} of (any) $p_0 \in M$ is a closed subgroup of G.

(b) The bijection (see Theorem 2.2.45),

$$\begin{array}{rcl} T_{p_0} \,:\, G/G_{p_0} & \longrightarrow & M \\ {[g]} & \longmapsto & \psi_g(p_0)\,. \end{array}$$

is a diffeomorphism for the natural differentiable structure on G/G_{p_0}.

(c) $\dim(M) = \dim(G) - dim(G_{p_0})$.

Remark 4.1.8 In particular if the action is transitive and free ($G_p = \{e\}$) then M is diffeomorphic to G, with diffeomorphism T_{p_0},

$$\begin{array}{rcl} T_{p_0} \,:\, G & \longrightarrow & M \\ g & \longmapsto & \psi_g(p_0)\,. \end{array}$$

$\Diamond$

An action of a Lie group G on the manifold M induces linear actions, i.e. representations, of G on the spaces of tensor fields $\mathcal{T}^{(r,s)}(M)$ (see (3.2.22)),

$$\begin{array}{rcl} G & \longrightarrow & L(\mathcal{T}^{(r,s)}(M)) \\ g & \longmapsto & (\psi_g)_* \end{array}$$

Very important for Lie groups are the (free and transitive) left and right actions of G on itself,

$$\begin{array}{rcl} L, R \,:\, G & \longrightarrow & \mathrm{Diff}(G) \\ L_h(g) & = & hg \\ R_h(g) & = & gh^{-1} \end{array}$$

We will denote left invariant tensor fields on G by

$$\mathcal{T}_L^{(r,s)}(G) := \left(\mathcal{T}^{(r,s)}(G)\right)^{L_G} .$$

Proposition 4.1.9 *The following equivalence is valid:*

$$T \in \mathcal{T}_L^{(r,s)}(M) \Leftrightarrow T_g = (L_g)_* (T_e) \;, \qquad \forall g \in G\,. \tag{4.1.1}$$

Proof. $\underline{\Rightarrow:}$
Let $T \in \mathcal{T}_L^{(r,s)}(G)$. We have,

$$\begin{array}{rcl} T \in \mathcal{T}_L^{(r,s)}(G) & \Leftrightarrow & (L_g)_* T = T\,, \quad \forall g \in G \\ & \Leftrightarrow & \left[(L_g)_*(T)\right]_h = T_h\,, \quad \forall g, h \in G \\ & \Leftrightarrow & (L_g)_* (T_{g^{-1}h}) = T_h\,, \quad \forall g, h \in G \\ & \overset{h=g}{\Rightarrow} & (L_g)_* (T_e) = T_g\,, \quad \forall g \in G\,. \end{array} \tag{4.1.2}$$

$\underline{\Leftarrow}$:
To complete the proof of the equivalence in (4.1.1) we only need to prove that the last implication in (4.1.2) is in fact an equivalence. So let T be such that its value at every $g \in G$ is given by,

$$T_g = (L_g)_* (T_e) \ .$$

Then,

$$(L_g)_* (T_{g^{-1}h}) = (L_g)_*(L_{g^{-1}h})_*(T_e) = (L_g \, L_{g^{-1}h})_* (T_e) = (L_h)_*(T_e) = T_h \, , \quad \forall g, h \in G \, .$$

■

We see that left invariant tensor fields on G are fully determined by their values at $e \in G$, so that the following map is a vector space isomorphism,

$$\begin{array}{rcl} \mathcal{T}_e^{(r,s)}(G) & \longrightarrow & \mathcal{T}_L^{(r,s)}(G) \\ T & \longmapsto & \widehat{T} \, , \quad \widehat{T}_g = (L_g)_*(T) \, , \quad \forall g \in G \, . \end{array} \tag{4.1.3}$$

Corollary 4.1.10 *Let G be a n–dimensional Lie group. Then,*

$$\dim \left(\mathcal{T}_L^{(r,s)}(G) \right) = n^{r+s} \, .$$

Proposition 4.1.11 *$\mathcal{X}_L(G)$ is a Lie subalgebra of $\mathcal{X}(G)$.*

Proof. Let $X, Y \in \mathcal{X}_L(G)$. Then, from proposition 3.2.18, we obtain

$$(L_g)_*[X,Y] = [(L_g)_*(X), (L_g)_*(Y)] = [X,Y] \, , \quad \forall g \in G \, ,$$

and therefore,

$$[X,Y] \in \mathcal{X}_L(G) \, .$$

■

Definition 4.1.12 (Lie algebra of a Lie group) *Let G be a Lie group and $\mathfrak{g} := T_eG$. Consider on $\mathfrak{g}$ the (unique) Lie bracket for which the vector space isomorphism*

$$\begin{array}{rcl} \mathfrak{g} & \longrightarrow & \mathcal{X}_L(G) \\ Y & \longmapsto & \widehat{Y} \, , \quad \widehat{Y}_g = (L_g)_*(Y) \, , \end{array}$$

is a Lie algebra isomorphism, i.e.,

$$[Y_1, Y_2] := [\widehat{Y}_1, \widehat{Y}_2]_e \ .$$

The pair $(\mathfrak{g}, [\ ,\ ])$ is called Lie algebra of the Lie group G.

Proposition 4.1.13 *([GN, proposition 7.3]) Let $Y \in \mathcal{X}_L(G)$. Then,*

(a) Y is complete.

(b) *Its integral curve throuph* e,

$$a_t^Y := \psi_t^{\widehat{Y}}(e)\,, \quad t \in \mathbb{R}\,,$$

is a 1*–parameter subgroup of* G*, i.e.*

$$\begin{aligned} a_s^Y \cdot a_t^Y &= \psi_s^{\widehat{Y}}(e) \cdot \psi_t^{\widehat{Y}}(e) = \psi_s^{\widehat{Y}}(e)\left(\psi_t^{\widehat{Y}}(e)\right) \\ &= \psi_{s+t}^{\widehat{Y}}(e) = a_{s+t}^Y\,, \qquad \forall s,t \in \mathbb{R}\,. \end{aligned}$$

(c) *The flow of* Y *is given by the right action of* a_t^Y*, i.e.*

$$\psi_t^{\widehat{Y}}(g) = R_{a_{-t}^Y}(g) = g\, a_t^Y\,, \quad \forall t \in \mathbb{R}, g \in G\,.$$

We define the exponential map,

$$\exp\,:\,\mathfrak{g} \longrightarrow G \tag{4.1.4}$$

as follows. First $Y \in \mathfrak{g}$ is mapped to the left invariant vector field $\widehat{Y} \in \mathcal{X}_L(G)$, which extends $\widehat{Y}_e = Y$ to G

$$\widehat{Y}_g = (L_g)_*\,(Y)\,.$$

Then, we find the integral curve of $\widehat{Y}$ that passes through e,

$$a_t^Y = \psi_t^{\widehat{Y}}(e)\,, \qquad t \in \mathbb{R}\,,$$

and follow it to time $t = 1$,

$$\exp(Y) := a_{t=1}^Y = \psi_1^{\widehat{Y}}(e)\,. \tag{4.1.5}$$

Proposition 4.1.14 *The map* $\exp\,:\,\mathfrak{g} \longrightarrow G$ *satisfies the following properties:*

P1. Is a local diffeomorphism from a neighborhood of $0 \in \mathfrak{g}$ *onto a neighborhood of* $e \in G$*.*

P2. $\exp(sY) = a_1^{sY} = a_s^Y\,, \qquad \forall s \in \mathbb{R},\, Y \in \mathfrak{g}$.

P3. Let G *be a matrix Lie group, i.e.* G *is a closed subgroup of* $GL_n(\mathbb{R})$*,*

$$G \subset GL_n(\mathbb{R})\,.$$

Then the naive tangent space to G *at* $e = \mathrm{I}_n$ *(as in (3.2.1) and (3.2.2)),*

$$\widetilde{\mathfrak{g}} := \widetilde{T_{\mathrm{I}_n} G} = \left\{ \widetilde{\dot{c}(0)} = \lim_{t\to 0} \frac{c(t) - \mathrm{I}_n}{t}\,,\ c\,:\,(-\epsilon,\epsilon) \longrightarrow G, c \,\text{is smooth},\ c(0) = I_n \right\}$$

is a Lie subalgebra of $\mathrm{Mat}_n(\mathbb{R})$ *with respect to the matrix commutator. The isomorphism of vector spaces (3.2.2), in the present case,*

$$T_{\mathrm{I}_n} G = \mathfrak{g} \longrightarrow \widetilde{\mathfrak{g}} = \widetilde{T_{\mathrm{I}_n} G} \subset \mathrm{Mat}_n(\mathbb{R})\,,$$

is a Lie algebra isomorphism and the following diagram is commutative,

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "\exp"] \arrow[d, "\sim"'] & G \\
\widetilde{\mathfrak{g}} \arrow[ur, "e^{\cdot}"] &
\end{tikzcd} \tag{4.1.6}$$

where $e^{\cdot}$ denotes the matrix exponential,

$$e^{\widetilde{Y}} = \mathrm{I}_n + \widetilde{Y} + \cdots + \frac{\widetilde{Y}^k}{k!} + \ldots \quad ,$$

i.e.

$$\begin{aligned} \exp(Y) &= e^{\widetilde{Y}} \\ &= \sum_{k=0}^{\infty} \frac{\widetilde{Y}^k}{k!}, \quad \forall Y \in \mathfrak{g}. \end{aligned}$$

We see that the commutativity of the diagram (4.1.6) means that, up to the isomorphism between $\mathfrak{g}$ and $\widetilde{\mathfrak{g}}$, the Lie group exponential, exp*, and the matrix exponential, $e^{\cdot}$ coincide.*

P4.

$$A \in \widetilde{\mathfrak{g}} \Leftrightarrow e^{tA} \in G, \qquad \forall t \in \mathbb{R}.$$

Exercise 4.1.1

(a) Find $\widetilde{Lie(O_n(\mathbb{R}))}$ and $\dim(O_n(\mathbb{R}))$.

(b) Find $\widetilde{Lie(U_n)}$ and $\dim(U_n)$.

(c) Find $\widetilde{Lie(SU_n)}$ and $\dim(SU_n)$.

Solution. Recall from 4a) and 4c) of example 4.1.3 the definition of the orthogonal, unitary and special unitary groups.

(a) From proposition 4.1.14, P4, we know that

$$\begin{aligned} \widetilde{A} \in \widetilde{Lie(O_n(\mathbb{R}))} \subset \mathrm{Mat}_n(\mathbb{R}) &\Leftrightarrow e^{s\widetilde{A}} \in O_n(\mathbb{R}), \quad \forall s \in \mathbb{R} \\ &\Leftrightarrow \left(e^{s\widetilde{A}}\right)^T e^{s\widetilde{A}} = \mathrm{I}, \quad \forall s \in \mathbb{R}, \\ &\Leftrightarrow e^{s\widetilde{A}^T} e^{s\widetilde{A}} = \mathrm{I}, \quad \forall s \in \mathbb{R}, \\ &\Leftrightarrow \mathrm{I} + s\left(\widetilde{A}^T + \widetilde{A}\right) + O(s^2) = \mathrm{I}, \quad \forall s \in \mathbb{R}, \\ &\Rightarrow \widetilde{A}^T + \widetilde{A} = 0. \end{aligned}$$

On the other hand if $\widetilde{A}^T = -\widetilde{A}$ it is easy to see that

$$e^{s\widetilde{A}^T} e^{s\widetilde{A}} = e^{-s\widetilde{A}} e^{s\widetilde{A}} = \mathrm{I}, \quad \forall s \in \mathbb{R},$$

and therefore the Lie algebra $\widetilde{Lie(O_n(\mathbb{R}))}$ is given by,

$$\widetilde{Lie(O_n(\mathbb{R}))} = \left\{ \widetilde{A} \in \mathrm{Mat}_n(\mathbb{R}) \ : \ (\widetilde{A})^T = -\widetilde{A} \right\} .$$

Then, the dimension of the Lie group $O_n(\mathbb{R})$ is equal to the dimension of its (tangent space at any point and thus equal to the dimension of its) Lie algebra, which is isomorphic to the Lie algebra of antisymmetric matrices. We therefore have,

$$\dim\left(O_n(\mathbb{R})\right) = \dim\left(Lie(O_n(\mathbb{R}))\right) = \frac{n(n-1)}{2} .$$

(b) As in (a) we have

$$\begin{aligned}
\widetilde{A} \in \widetilde{Lie(U_n)} \subset \mathrm{Mat}_n(\mathbb{C}) \ &\Leftrightarrow \ e^{s\widetilde{A}} \in U_n \,, \quad \forall s \in \mathbb{R} \\
&\Leftrightarrow \ \left(e^{s\widetilde{A}}\right)^\dagger e^{s\widetilde{A}} = \mathrm{I} \,, \quad \forall s \in \mathbb{R} \,, \\
&\Leftrightarrow \ e^{s\widetilde{A}^\dagger} e^{s\widetilde{A}} = \mathrm{I} \,, \quad \forall s \in \mathbb{R}, \\
&\Leftrightarrow \ \mathrm{I} + s\left(\widetilde{A}^\dagger + \widetilde{A}\right) + O(s^2) = \mathrm{I} \,, \quad \forall s \in \mathbb{R}, \\
&\Rightarrow \ \widetilde{A}^\dagger + \widetilde{A} = 0 \,.
\end{aligned}$$

On the other hand if $\widetilde{A}^\dagger = -\widetilde{A}$ it is easy to see that

$$\left(e^{s\widetilde{A}}\right)^\dagger e^{s\widetilde{A}} = e^{s(\widetilde{A})^\dagger} e^{s\widetilde{A}} = e^{-s\widetilde{A}} e^{s\widetilde{A}} = \mathrm{I}, \quad \forall s \in \mathbb{R} \,,$$

and therefore the Lie algebra $\widetilde{Lie(U_n)}$ is given by,

$$\widetilde{Lie(U_n)} = \left\{ \widetilde{A} \in \mathrm{Mat}_n(\mathbb{C}) \ : \ (\widetilde{A})^\dagger = -\widetilde{A} \right\} .$$

The dimension of the Lie group U_n is then equal to the dimension of its Lie algebra, which is isomorphic to the Lie algebra of antihermitian matrices, i.e. matrices with entries satisfying,

$$a_{jk} = -\overline{a_{kj}} \,. \tag{4.1.7}$$

Therefore, the entries above the main diagonal can be chosen freely as arbitrary complex numbers and on the main diagonal can be chosen as arbitrary imaginary numbers. Below the main diagonal they satisfy (4.1.7). Then, we find for the (real) dimension of U_n,

$$\dim\left(U_n\right) = \dim\left(Lie(U_n)\right) = 2\,\frac{n(n-1)}{2} + n = n^2 \,.$$

(c) We will need the following identity for the determinant of the exponential of a matrix $B \in \mathrm{Mat}_n(\mathbb{C})$,

$$\det(e^B) = e^{\mathrm{tr}(B)} \,.$$

We have

$$\begin{aligned}\widetilde{A} \in \widetilde{Lie(SU_n)} \subset \mathrm{Mat}_n(\mathbb{C}) \quad &\Leftrightarrow \quad e^{s\widetilde{A}} \in SU_n\,, \quad \forall s \in \mathbb{R} \\ &\Leftrightarrow \quad e^{s\widetilde{A}} \in U_n \text{ and } \det\left(e^{s\widetilde{A}}\right) = e^{s\,\mathrm{tr}\widetilde{A}} = 1\,, \quad \forall s \in \mathbb{R}, \\ &\Leftrightarrow \quad \widetilde{A}^\dagger + \widetilde{A} = 0 \text{ and } \mathrm{tr}\widetilde{A} = 0\,.\end{aligned}$$

Therefore the Lie algebra $\widetilde{Lie(SU_n)}$ is given by,

$$\begin{aligned}\widetilde{Lie(SU_n)} &= \left\{\widetilde{A} \in \mathrm{Mat}_n(\mathbb{C}) \,:\, (\widetilde{A})^\dagger = -\widetilde{A}\,,\ \mathrm{tr}\widetilde{A} = 0\right\} \\ &= \left\{\widetilde{A} \in \widetilde{Lie(U_n)} \,:\, \mathrm{tr}\widetilde{A} = 0\right\}\,. \end{aligned} \tag{4.1.8}$$

From (4.1.8) and (b) we find that,

$$\dim(SU_n) = \dim(\widetilde{Lie(SU_n)}) = n^2 - 1\,.$$

■

Of particular importance in quantum mechanics is the three dimensional Lie group SU_2, associated with angular momentum and spin of particles and in elementary particle physics the groups SU_2 and SU_3, the latter with dimension $3^2 - 1 = 8$.

Let G be a Lie group and let us fix a basis in its Lie algebra,

$$\begin{aligned}B = \{X_j\}_{j=1}^n \quad &\subset \quad \mathfrak{g} = Lie(G) \\ [X_j, X_k] \quad &= \quad \sum_{l=1}^n C_{jk}^l\, X_l\,,\end{aligned}$$

and the corresponding basis of left invariant vector fields (see (4.1.3)),

$$\widehat{B} = \left\{\widehat{X}_j\right\}_{j=1}^n \subset \mathcal{X}_L(G)\,.$$

The constants C_{jk}^l are called structure constants of $\mathfrak{g}$ (in the basis B).

Exercise 4.1.2 *Show that if,*

$$B^* = \{\omega_j\}_{j=1}^n \subset \Omega_L^1(G) \tag{4.1.9}$$

is a basis of left–invariant 1–forms such that their values at the identity $e \in G$ form a basis dual to the Lie algebra basis B,

$$(\omega_j)_e(X_k) = \delta_{jk}\,,$$

then B^* is a basis dual to $\widehat{B}$, i.e.

$$\omega_j(\widehat{X}_k) = \delta_{jk}\,.$$

Solution. Let us show first that $\Omega_L^1(G)$ is naturally isomorphic to the dual of $\mathcal{X}_L(G)$. This is due to the fact that if $\omega \in \Omega_L^1(G)$ and $Y \in \mathcal{X}_L(G)$, then $\omega(Y)$ is constant,

$$\omega(Y)(g) = \omega(Y)(e) = \text{const}\ \forall g \in G.$$

Indeed (see (3.2.22)),

$$\omega_g = (L_g)_*(\omega_e) = (L_{g^{-1}})^* (\omega_e),$$

and therefore,

$$\begin{aligned} \omega(Y)(g) &= \omega_g(Y_g) = [(L_{g^{-1}})^* (\omega_e)] ((L_g)_*(Y_e)) \\ &= \omega_e \left((L_{g^{-1}})_* (L_g)_* (Y_e)\right) = \omega_e(Y_e) \\ &= \omega(Y)(e)\,, \qquad \forall g \in G\,. \end{aligned}$$

The isomorphism is then given by,

$$\begin{aligned} \Omega^1_L(G) &\longrightarrow (\mathcal{X}_L(G))^* \\ \omega &\longmapsto (\mathcal{X}_L(G) \ni X \mapsto \omega(X) \in \mathbb{R})\ . \end{aligned} \tag{4.1.10}$$

Indeed, since

$$\omega_j(\widehat{X}_k)(g) = (\omega_j)_g \left((\widehat{X}_k)_g\right) = (\omega_j)_e \left((\widehat{X}_k)_e\right) = (\omega_j)_e (X_k) = \delta_{jk}\,.$$

So, the linear transformation (4.1.10) maps the basis B^* in (4.1.9) to the basis of $(\mathcal{X}_L(G))^*$ dual to the basis,

$$\widehat{B} = \left\{\widehat{X}_j\right\}_{j=1}^n \subset \mathcal{X}_L(G)\,,$$

which shows that it is indeed an isomorphism of vector spaces. ■
We will use the isomorphism (4.1.10) to identify $\Omega^1_L(G)$ with $(\mathcal{X}_L(G))^*$.

Corollary 4.1.15 *Let* $\widehat{B} = \left\{\widehat{X}_j\right\}_{j=1}^n \subset \mathcal{X}_L(G)$ *and* $\{\omega_k\}_{k=1}^n \subset \Omega^1_L(G)$ *be basis of left–invariant vector fields and* 1*–forms, respectively. Then*

$$T \in \mathcal{T}_L^{(r,s)}(G) \quad \Leftrightarrow \quad \left\{ \begin{array}{ll} T = & \sum_{i_1\ldots i_r j_1\ldots j_s} T_{i_1\ldots i_r j_1\ldots j_s}\ \omega_{i_1} \otimes \cdots \otimes \omega_{i_r} \otimes \widehat{X}_{j_1} \otimes \cdots \otimes \widehat{X}_{j_s} \\ & \\ & T_{i_1\ldots i_r j_1\ldots j_s} \textit{ are constants.} \end{array} \right.$$

Proof. Let T be left–invariant. Then,

$$\begin{aligned} T_g &= \sum_{i_1\ldots i_r j_1\ldots j_s} T_{i_1\ldots i_r j_1\ldots j_s}(g)\ \left(\omega_{i_1} \otimes \cdots \otimes \omega_{i_r} \otimes \widehat{X}_{j_1} \otimes \cdots \otimes \widehat{X}_{j_s}\right)_g \\ &= (L_g)_*(T_e) \\ &= \sum_{i_1\ldots i_r j_1\ldots j_s} T_{i_1\ldots i_r j_1\ldots j_s}(e)\ (L_g)_* \left(\omega_{i_1} \otimes \cdots \otimes \omega_{i_r} \otimes \widehat{X}_{j_1} \otimes \cdots \otimes \widehat{X}_{j_s}\right)_e \\ &= \sum_{i_1\ldots i_r j_1\ldots j_s} T_{i_1\ldots i_r j_1\ldots j_s}(e)\ \left(\omega_{i_1} \otimes \cdots \otimes \omega_{i_r} \otimes \widehat{X}_{j_1} \otimes \cdots \otimes \widehat{X}_{j_s}\right)_g\,, \end{aligned}$$

so that indeed, the tensor T is left–invariant if and only if all its components $T_{i_1\ldots i_r j_1\ldots j_s}$ in a left–invariant basis are constant.

■

4.2 Application to spatially homogeneous cosmological models

Assume that, at large distances, the Lorentzian metric γ of our space–time ($\dim(M) = 4$) is approximately invariant under the free action of a 3–dimensional Lie group G, with space–like orbits (i.e. such that the restriction of γ to their tangent space is positive definite), $\{\tau_0\} \times \Sigma$,

$$M = [0, \tau_f) \times \Sigma \,,$$

where $\tau \in [0, \tau_f)$ is the so called cosmological time. Since the action of G is transitive and free on its orbits from remark 4.1.8 it follows that the orbits $\{\tau_0\} \times \Sigma$ are diffeomorphic to G and therefore,

$$M \overset{diff}{\cong} [0, \tau_f) \times G \,.$$

We are assuming that the apace–time metric γ is, at very large distances, well approximated by a metric, γ_{sh}, which is called spatially homogeneous as it contains G in its isometry group (see (3.2.26)),

$$G \subset \mathrm{Iso}(\gamma_{\mathrm{sh}}) \,.$$

The assumption of spatial homogeneity is natural especially for $\tau \leq 380000\,\text{years}$ after the Big–Bang, before the cosmic microwave background radiation was emitted with an amazing low anisotropy as observed today at the Earth,

$$\frac{|\Delta T|}{T} \leq 10^{-5} \,.$$

Then the (approximate) metric of our space–time can be written in the form,

$$\gamma_{\mathrm{sh}} = N(t)\, dt^2 + \sum_{i=1}^{3} N_i(t)\, dt\, \omega_i + \sum_{i,j=1}^{3} h_{ij}(t)\, \omega_i\, \omega_j \,, \tag{4.2.1}$$

where $B^* = \{\omega_1, \omega_2, \omega_3\}$ is a basis of left–invariant 1–forms on G dual to a basis of left–invariant vector fields, $\widehat{B} = \left\{\widehat{X}_1, \widehat{X}_2, \widehat{X}_3\right\}$. So in spatially homogeneous cosmological models instead of ten unknown functions (the components of the metric) of four variables (the (local) coordinates of space–time) the geometry is determined by choosing just ten functions, N, N_i, h_{ij} of a single variable, the time t. By substituting (4.2.1) into the Einstein equations,

$$R_{jk}(\gamma_{\mathrm{sh}}) - \frac{1}{2}\gamma_{\mathrm{sh}\,jk} = \frac{8\pi G}{c^4}\, T_{jk} \,, \tag{4.2.2}$$

one obtains, for the geometry[2], a system of six second order nonlinear ODE for the spatial metric coefficients $h_{jk}(t)$ with constraints instead of six second order nonlinear PDE in four variables! The system of ODE for γ_{sh} depends crucially on the chosen homogeneity group G and therefore one has one physical cosmological model for each three dimensional Lie group G (see [RS, Chapter 6]).

4.3 Lie Algebras

[2] one further assumes that the distribution of matter has the same symmetries so that the symmetry group of the energy momentum tensor of matter contains G as a subgroup, $S(T) \subset G$.

Bibliography

[AF] C. Adams, R Franzosa, *Introduction to Topology Pure and Applied*, Dorling Kindersley India, 2008.

[Ar] A. Arvanitoyeorgos, *An Introduction to Lie Groups and the Geometry of Homogeneous Spaces*, AMS, 1999.

[Ca] R. Cahn, *Semi-Simple Lie Algebras and Their Representations*, Benjamin Cummings, 1984.

[CJ] R. Coquereaux, A. Jadcyzk, *Riemannian geometry, fiber bundles, Kaluza-Klein theories and all that ...*, World Scientific, 1988.

[GN] L. Godinho, J. Natario, *An Introduction to Riemannian Geometry With Applications to Mechanics and Relativity*, Springer, 2014.

[Ha] B. Hall, *Lie Groups, Lie Algebras, and Representations An Elementary Introduction,* Springer, 2015.

[Is] C. Isham, *Modern Differential Geometry for Physicists*, World Scientific Lecture Notes in Physics, **61**, 1999.

[Ka] A. Kaveh, *Optimal Analysis of Structures by Concepts of Symmetry and Regularity*, Springer, 2013.

[KN1] A. Kaveh, and M. Nikbakht, *Decomposition of symmetric mass–spring vibrating systems using groups, graphs and linear algebra*, Commun. Numer. Meth. Engng **23** (2007) 639–664.

[KN2] A. Kaveh, and M. Nikbakht, *Stability analysis of hyper symmetric skeletal structures using group theory,* Acta Mech **200** (2008) 177–197.

[Ko] Y. Kosmann-Schwarzbach, *Groups and Symmetries. From Finite Groups to Lie Groups*, Springer, 2010.

[La] N. Lauritzen, *Concrete abstract algebra. From numbers to Gröbner basis,* Cambridge University Press, 2003.

[Ma] A. Marsh, *Mathematics for Physics*, World Scientific, 2018.

[Na] M. Nakahara *Geometry, Topology and Physics*, IOP, 2003.

[Ne] T. Needham, *Introduction to Applied Algebraic Topology*, Course notes, 2019.

[Ro] G. Ross, *Grand Unified Theories*, Benjamin Cummings, 1985.

[RS] M. Ryan, L. Shepley, *Homogeneous relativistic cosmologies,* Priceton University Press, 1975.

[Sl] R. Slansky, *Group theory for unified model building,* Physics Reports **79** (1981) 1–128.

[Sp] M. Spivak A Comprehensive introduction to differential geometry bf 1, Publish or Perish, 1999.

[St] B. Steinberg, *Representation Theory of Finite Groups. An Introductory Approach,* Springer, Universitext, 2012.