Basics of Differential Geometry & Group Theory

Francesco Benini

SISSA – PhD course Fall 2019

Contents

Last update: October 25, 2019

1 Introduction

This course is divided into two parts. The first part is about differential geometry and fibre bundles. The material is very standard, and is mainly taken from the book of M. Nakahara. The second part is about Lie algebras, Lie groups and their representations. A good concise exposition is in Chapter 13 of the book of P. Di Francesco. Another good and simple reference is the book of R. N. Cahn. More technical details and proofs can be found in the lecture notes by V. Kac. For spinors in various dimensions, a good reference is Appendix B of the book of J. Polchinski, Volume 2.

We will first introduce Lie groups in terms of their action on differentiable manifolds, which is a concrete way they appear in physics, and then move on to study their formal properties.

Suggested readings:

- M. Nakahara, "*Geometry, Topology and Physics*," Taylor & Francis (2003).
- C. J. Isham, "*Modern Differential Geometry for Physicists*," World Scientific (1999).
- C. Nash, S. Sen, "*Topology and Geometry for Physicists*," Elsevier (1988).
- P. Di Francesco, P. Mathieu, D. Senechal, "*Conformal Field Theory*," Springer (1996). Charpter 13
- V. Kac, "*Introduction to Lie Algebras*," lecture notes (2010) available online: `http://math.mit.edu/classes/18.745/index`
- R. N. Cahn, "*Semi-Simple Lie Algebras and Their Representations*," Benjamin/Cummings Publishing Company (1984). (Available online)
- J. Polchinski, "*String Theory Volume II*," Cambridge University Press (1998). Appendix B

2 Topology

One of the simplest structures we can define on a set X is a *topology*.

Definition 2.1. Let X be a set and $\mathcal{I} = \{U_i \,|\, i \in I,\, U_i \subset X\}$ a certain collection of subsets of X, called **open sets**. Then $(X, \mathcal{I})$ is called a **topological space** if $\mathcal{I}$ satisfies the following properties:

(i) $\varnothing, X \in \mathcal{I}$.

(ii) If $\{U_j \,|\, j \in J\}$ is any (maybe infinite) subcollection in $\mathcal{I}$, then $\bigcup_{j\in J} U_j \in \mathcal{I}$.

(iii) If $\{U_k \,|\, k \in K\}$ is any *finite* subcollection in $\mathcal{I}$, then $\bigcap_{k\in K} U_k \in \mathcal{I}$.

The collection $\mathcal{I}$ defines a **topology** on X.

Example 2.2.

(a) If X is a set and $\mathcal{I}$ is the collection of all subsets in X, then (i)–(iii) are satisfied. This topology is called the **discrete topology**.[1]

(b) If X is a set and $\mathcal{I} = \{\varnothing, X\}$, then (i)–(iii) are satisfied. This topology is called the **trivial topology**.[2]

(c) Let X be the real line $\mathbb{R}$, and $\mathcal{I}$ the set of all open intervals (a, b) and their unions, where a can be $-\infty$ and b can be $+\infty$. This topology is called the **usual topology**. The usual topology on $\mathbb{R}^n$ is constructed in the same way from products of open sets $(a_1, b_1) \times \ldots \times (a_n, b_n)$.

Let $(X, \mathcal{I})$ be a topological space and A a subset of X. Then $\mathcal{I}$ induces the **relative topology** on A given by $\mathcal{I}' = \{U_i \cap A \,|\, U_i \in \mathcal{I}\}$.

A topology on sets is useful because it allows us to define *continuity*.

Definition 2.3. Let X and Y be topological spaces. A function $f : X \to Y$ is **continuous** if the *inverse* image of an open set in Y is an open set in X.

The inverse image of a subset $C \subset Y$ is given by

$$f^{-1}(C) = \{x \in X \,|\, f(x) \in C\} \ .$$

¶ *Exercise* 1. Show that a discontinuous function, for instance

$$f(x) = \begin{cases} x & \text{for } x \le 0 \\ x + 1 & \text{for } x > 0 \ , \end{cases}$$

fails to be continuous by that definition. On the other hand, show that a continuous function can send an open set to a set which is not open—for instance $f(x) = x^2$.

[1] In this topology each point is an open set, thus is its own neighbourhood. Hence points are disconnected.

[2] In this topology the only neighbourhood of a point is the whole X, therefore all points are "infinitely close" to each other. If X contains more than one point, $(X, \mathcal{I})$ is not Hausdorff (see footnote 3).

A special class of topological spaces is given by *metric spaces* on which there is an extra structure—a metric—which induces a topology.

Definition 2.4. A **metric** $d : X \times X \to \mathbb{R}$ is a function that satisfies the conditions:

(i) It is symmetric, *i.e.* $d(x, y) = d(y, x)$

(ii) $d(x, y) \geq 0$ and the equality holds if and only if $x = y$

(iii) $d(x, y) + d(y, z) \geq d(x, z)$.

If X is endowed with a metric d, there is an induced **metric topology** on X in which the open sets are the "open discs"

$$U_\epsilon(x) = \{y \in X \,|\, d(x, y) < \epsilon\}$$

and all their unions. The topological space $(X, \mathcal{I})$ is called a **metric space**.[3]

A subset $C \subset X$ is **closed** if its complement in X is open, *i.e.* if $X \backslash C \in \mathcal{I}$. From Definition 2.1, $\varnothing$ and X are *both* open and closed. Moreover the intersection of any (possibly infinite) collection of closed sets is closed, and the union of a finite number of closed sets is closed. The **closure** of A is the smallest closed set that contains A, and is denoted $\overline{A}$. The **interior** of A is the largest open subset of A, and is denoted A°. Notice that A is closed if and only if $\overline{A} = A$, and is open if and only if $A^\circ = A$. The **boundary** of A is the complement of A° in $\overline{A}$, and is denoted $\partial A = \overline{A} \setminus A^\circ$.

A family $\{A_i\}$ of subsets of X if called a **covering** of X if

$$\bigcup\nolimits_{i \in I} A_i = X \ .$$

If all A_i are open, the covering is called an **open covering**.

Definition 2.5. A set X is **compact** if, for every open covering $\{U_i \,|\, i \in I\}$, there exists a *finite* subset J of I such that $\{U_j \,|\, j \in J\}$ is also a covering of X.

Theorem 2.6. *A subset A of $\mathbb{R}^n$ is* compact *(with respect to the usual topology) if and only if it is* closed *and* bounded.

Definition 2.7.

(a) A topological space X is **connected** if it cannot be written as $X = X_1 \cup X_2$ where $X_{1,2}$ are both open, non-empty, and $X_1 \cap X_2 = \varnothing$. Otherwise X is **disconnected**.

(b) A topological space X is **arcwise connected** if, for any points $x, y \in X$, there exists a continuous function $f : [0, 1] \to X$ such that $f(0) = x$ and $f(1) = y$. With a few pathological exceptions, arcwise connectedness is equivalent to connectedness.

[3] All metric spaces are Hausdorff. A topological space $(X, \mathcal{I})$ is **Hausdorff**, or separable, if for any two distinct points x, y, there exist a neighbourhood U of x and a neighbourhood V of y that are disjoint.

(c) A loop is a continuous map $f : [0,1] \to X$ such that $f(0) = f(1)$. If any loop in X can be continuously shrunk to a point,[4] X is called **simply connected**.

Example 2.8. $\mathbb{R} \setminus \{0\}$ is disconnected, $\mathbb{R}^2 \setminus \{0\}$ is connected but not simply-connected, $\mathbb{R}^{n\geq 3} \setminus \{0\}$ is connected and simply-connected.

Definition 2.9. Let X_1 and X_2 be topological spaces. A map $f : X_1 \to X_2$ is a **homeomorphism** if it is continuous and it has an inverse $f^{-1} : X_2 \to X_1$ which is also continuous. If there exists a homeomorphism between X_1 and X_2, they are said to be **homeomorphic**.

Quantities that are preserved under homeomorphisms are called *topological invariants*. Clearly if two spaces have at least one topological invariant that differ, they cannot be homeomorphic. Unfortunately we do not know a complete list of topological invariants that can characterize different topologies.

A coarser equivalence class than homeomorphism, is *homotopy type*.

Definition 2.10. Two continuous functions $f, g : X \to Y$ from one topological space to another are **homotopic** if one can be continuously deformed into the other, namely if there exists a continuous function $H : X \times [0,1] \to Y$ such that $H(\,\cdot\,, 0) = f$ and $H(\,\cdot\,, 1) = g$.
The two spaces X, Y are **homotopy equivalent**, or of the same **homotopy type**, if there exist continuous maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f$ is homotopic to id_X and $f \circ g$ is homotopic to id_Y.

Two spaces that are homeomorphic are also of the same homotopy type, but the converse is not true.

¶ *Exercise* 2. Show that the circle S^1 and the infinite cylinder $S^1 \times \mathbb{R}$ are of the same homotopy type.

Spaces that are homotopy equivalent to a point are called **contractible**.

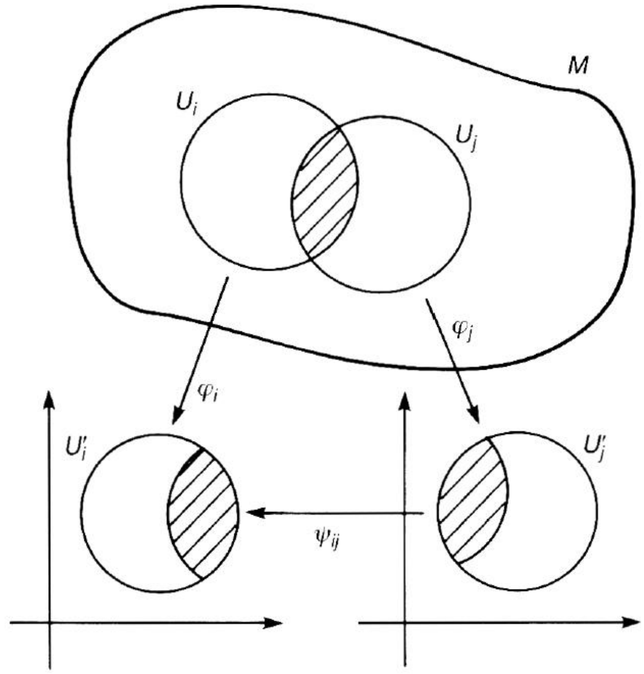
3 Differentiable manifolds

Adding "topology" to a space X allows us to talk about continuous functions, *i.e.* it introduces a concept of continuity. A much richer structure is the concept of differentiability, what is called a "smooth (C^∞) structure". A *manifold* is an object that is locally homeomorphic to $\mathbb{R}^m$, and thus inherits a differentiable structure.

Definition 3.1. M is an m-dimensional **differentiable manifold** if:

(i) M is a topological space

[4] This can be stated as the existence of a continuous map $g : [0,1] \times [0,1] \to X$ such that $g(x,0) = f(x)$ and $g(x,1) = y$ for some point $y \in X$. Alternatively, we demand the existence of a disk in X whose boundary is the loop.



(ii) M is provided with a collection of pairs $\{(U_i, \varphi_i)\}$.

$\{U_i\}$ is a collection of open sets which covers M. φ_i is a homeomorphism from U_i onto an open subset U_i' in $\mathbb{R}^m$

(iii) Given U_i and U_j with $U_i \cap U_j \neq \varnothing$, the map $\psi_{ij} = \varphi_i \circ \varphi_j^{-1}$ from $\varphi_j(U_i \cap U_j)$ to $\varphi_i(U_i \cap U_j)$ is C^∞ (we call it *smooth*).

Each pair (U_i, φ_i) is called a **chart** (see Figure), while the whole family $\{(U_i, \varphi_i)\}$ is called an **atlas**. The homeomorphism φ_i is represented by m functions

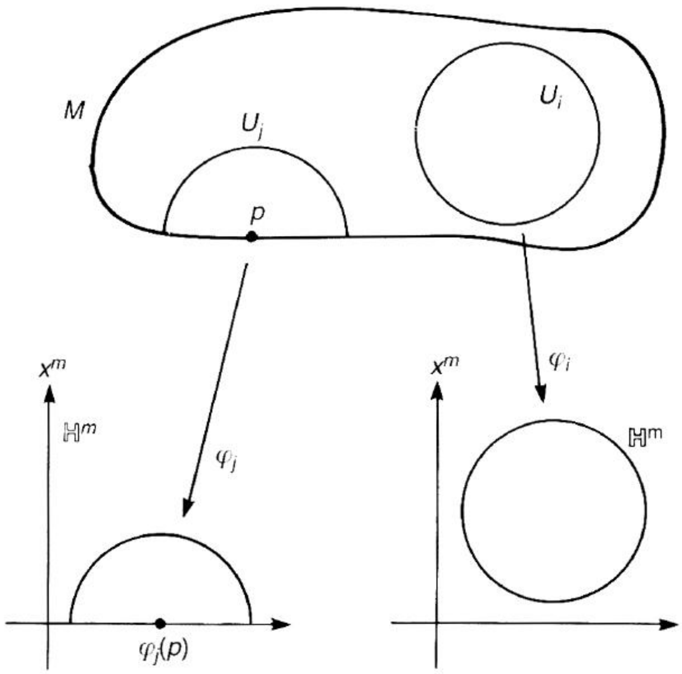
$$\varphi_i(p) = \{x^1(p), \ldots, x^m(p)\} = \{x^\mu(p)\}$$

called **coordinates**. If U_i and U_j overlap, two coordinate systems are assigned to a point. The functions $\psi_{ij} = \varphi_i \circ \varphi_j^{-1}$, called **coordinate transition functions**, are from an open set in $\mathbb{R}^m$ to another open set in $\mathbb{R}^m$. We write them as

$$x^\mu(y^\nu)$$

and are required to be C^∞ with respect to the standard definition in $\mathbb{R}^m$. If the union of two atlases $\{(U_i, \varphi_i)\}$, $\{(V_j, \psi_j)\}$ is again an atlas, the two atlases are said to be **compatible**. Compatibility is an equivalence relation, and each possible equivalence class is called a **differentiable structure** on M.

Manifolds with boundaries can be defined in a similar way. The topological space M must be covered by a family of open sets $\{U_i\}$ each homeomorphic to an open set in



$\mathbb{H}^m = \{(x^1, \ldots, x^m) \,|\, x^m \geq 0\}$. The set of points that are mapped to points with $x^m = 0$ is the **boundary** of M, denoted by ∂M (see Figure).[5]

The boundary ∂M is itself a manifold, with an atlas induced from M.

Example 3.2. These are examples of manifolds.

(a) The Euclidean space $\mathbb{R}^m$ is covered by a single chart and φ is the identity map.

(b) The n-dimensional sphere S^n, realized in $\mathbb{R}^{n+1}$ as

$$\sum_{i=0}^{n} (x^i)^2 = 1 \; .$$

It can be covered with $2(n+1)$ patches:[6]

$$U_{i\pm} = \{(x^0, \ldots, x^n) \in S^n \,|\, x^i \gtrless 0\} \; .$$

The coordinate maps $\varphi_{i\pm} : U_{i\pm} \to \mathbb{R}^n$ are given by

$$\varphi_{i\pm}(x^0, \ldots, x^n) = (x^0, \ldots, x^{i-1}, x^{i+1}, \ldots, x^n)$$

where we omit x^i. The coordinate transition functions are

$$\psi_{is,jt} = \varphi_{is} \circ \varphi_{jt}^{-1} = \Big(x^0, \ldots,\, x^j = t\sqrt{1 - \textstyle\sum_{k(\neq j)} (x^k)^2},\, \ldots \widehat{x^i} \ldots, x^n\Big)$$

[5] The boundary of a manifold M is *not* the same as the boundary of M in the topological sense. In fact, if X is a topological space then $\overline{X} = X^\circ = X$ and thus $\partial X = \varnothing$.

[6] Two patches are enough to cover S^n, just remove a point of S^n in each case. However writing down the coordinates in terms of those in $\mathbb{R}^{n+1}$ requires the stereographic projection, and it is not completely trivial. Our parametrization here is simpler.

where $s, t \in \{\pm 1\}$ and we omit x^i from the list.

In fact, since $S^n \setminus \{\text{pt}\}$ is homeomorphic to $\mathbb{R}^n$, we can cover the sphere we two patches obtained by removing antipodal points.

(c) The real projective space $\mathbb{RP}^n$ is the set of lines through the origin in $\mathbb{R}^{n+1}$. If $x = (x^0, \ldots, x^n) \neq 0$ is a point in $\mathbb{R}^{n+1}$, it defines a line through the origin. Another non-zero point y defines the same line if and only if there exists $a \in \mathbb{R}_{\neq 0}$ such that $y = ax$. This defines an equivalence relation $x \sim y$, and we define

$$\mathbb{RP}^n = \big(\mathbb{R}^{n+1} \setminus \{0\}\big)/\sim \ .$$

The $n+1$ numbers $x^0, \ldots, x^n$ are called **homogeneous coordinates**, but they are redundant. Instead we define patches

$$U_i = \{\text{lines in } \mathbb{RP}^n \,|\, x^i \neq 0\} \ .$$

and **inhomogeneous coordinates** on U_i given by

$$\xi^j_{(i)} = x^j / x^i \ ,$$

where the entry $\xi^i_{(i)} = 1$ is omitted. For $x \in U_i \cap U_j$ the coordinate transition functions are

$$\psi_{ij} : \xi^k_{(j)} \to \xi^k_{(i)} = x^k/x^i = \frac{x^k}{x^j}\frac{x^j}{x^i} = \xi^k_{(j)}/\xi^i_{(j)} \ .$$

Let $f : M \to N$ be a map from an m-dimensional manifold M to an n-dimensional manifold N. A point $p \in M$ is mapped to a point $f(p) \in N$. Take charts (U, φ) on M and (V, ψ) on N such that $p \in U$ and $f(p) \in V$. Then f has the following coordinate presentation:

$$\psi \circ f \circ \varphi^{-1} : \mathbb{R}^m \to \mathbb{R}^n$$

(the function is defined on $\varphi(U) \subset \mathbb{R}^m$). If we write $\varphi(p) = \{x^\mu\}$ and $\psi(f(p)) = \{y^\alpha\}$, then $\psi \circ f \circ \varphi^{-1}$ is given by n functions of m variables, and with some abuse of notation we write

$$y^\alpha = f^\alpha(x^\mu)$$

(really we should write $y^\alpha = (\psi f \varphi^{-1})^\alpha(x^\mu)$). We say that f is **smooth** (C^∞) if $\psi \circ f \circ \varphi^{-1}$ is, as a function from $\mathbb{R}^m$ to $\mathbb{R}^n$. The property of being C^∞ does not depend on the coordinates used, since coordinate transformations are C^∞.

Definition 3.3. If f is invertible and both $y = \psi f \varphi^{-1}(x)$ and $x = \varphi f^{-1} \psi^{-1}(y)$ are C^∞ for all charts, f is called a **diffeomorphism** and M is **diffeomorphic** to N, denoted $M \equiv N$.

Clearly, if $M \equiv N$ then $\dim M = \dim N$.[7] Two spaces that are diffeomorphic are also homeomorphic, however the opposite is not true. There can be multiple differentiable structures on the same topological space. For instance, S^7 admits 28 differentiable structures (Milnor, 1956).[8]

[7] If f is a diffeomorphism, then by definition $y(x) : \mathbb{R}^m \to \mathbb{R}^n$ is smooth and with smooth inverse. In particular, at any point, $\frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial x^\mu}{\partial y^\beta} = \delta^\alpha_\beta$ and $\frac{\partial x^\mu}{\partial y^\alpha}\frac{\partial y^\alpha}{\partial x^\nu} = \delta^\mu_\nu$. This implies that $m = n$.

[8] The 28 exotic 7-spheres can be constructed as follows: intersect in $\mathbb{C}^5$ the zero-locus of the complex

3.1 Vectors

An open **curve** on a manifold M is a map

$$c : (a, b) \to M \ ,$$

where (a, b) is an open interval and we take $a < 0 < b$. We also assume that the curve does not intersect itself. On a chart (U, φ), the curve $c(t)$ has presentation $x = \varphi\, c : \mathbb{R} \to \mathbb{R}^m$.

A **function** on M is a smooth map $f : M \to \mathbb{R}$. On a chart (U, φ), the function has presentation $f\,\varphi^{-1} : \mathbb{R}^m \to \mathbb{R}$ which is a real function of m variables. We denote the set of smooth functions on M by $\mathscr{F}(M)$.

Given a curve $c : (a, b) \to M$, we define the tangent vector to c at $c(0)$ as a *directional derivative* of functions $f : M \to \mathbb{R}$ along the curve $c(t)$ at $t = 0$. Such a derivative is

$$\left.\frac{\mathrm{d}f(c(t))}{\mathrm{d}t}\right|_{t=0} .$$

In terms of local coordinates, this becomes

$$(\partial f / \partial x^\mu) \left(\mathrm{d}x^\mu(c(t))/\mathrm{d}t\right)\big|_{t=0} .$$

We use the convention that repeated indices are summed over. [Note the abuse of notation! $\partial f / \partial x^\mu$ means $\partial(f\varphi^{-1}(x))/\partial x^\mu$.] Thus, $\mathrm{d}f(c(t))/\mathrm{d}t$ at $t = 0$ is obtained by applying a first-order linear differential operator X to f, where

$$X \equiv X^\mu \frac{\partial}{\partial x^\mu} \qquad \text{and} \qquad X^\mu = \left.\frac{\mathrm{d}x^\mu(c(t))}{\mathrm{d}t}\right|_{t=0} ,$$

namely

$$\left.\frac{\mathrm{d}f(c(t))}{\mathrm{d}t}\right|_{t=0} = X^\mu \frac{\partial}{\partial x^\mu} f = X[f] .$$

We define $X = X^\mu \partial / \partial x^\mu = X^\mu \partial_\mu$ as the **tangent vector** to M at $c(0)$ along the direction given by the curve $c(t)$.

We could introduce an equivalence relation between curves in M. If two curves $c_1(t)$ and $c_2(t)$ satisfy

(i) $c_1(0) = c_2(0) = p$

(ii) $\mathrm{d}x^\mu(c_1(t))/\mathrm{d}t\big|_{t=0} = \mathrm{d}x^\mu(c_2(t))/\mathrm{d}t\big|_{t=0}$

then they yield the same differential operator X at p. We declare $c_1(t) \sim c_2(t)$ and identify the *tangent vector* X with the *equivalence class of curves*.

equation $a^2 + b^2 + c^2 + d^3 + e^{6k-1} = 0$ with a small sphere around the origin. For $k = 1, \ldots, 28$ one obtains all differentiable structures on S^7 (Brieskorn, 1966). These are called Brieskorn spheres.

The set of linear first-order differential operators at $p \in M$ is a vector space called the **tangent space** of M at p, denoted by T_pM. The structure of vector space is natural in terms of linear differential operators.[9] Evidently,

$$e_\mu = \frac{\partial}{\partial x^\mu} \equiv \partial_\mu \qquad\qquad \mu = 1, \dots, m$$

are basis vectors of T_pM and $\dim T_pM = m$. The basis $\{e_\mu\}$ is called the **coordinate basis**. If a vector $V \in T_pM$ is written as

$$V = V^\mu e_\mu \ ,$$

the numbers V^μ are called the components of V with respect to the basis $\{e_\mu\}$. The transformation law of the components under change of coordinates follows from the fact that a vector V exists independently of any choice of coordinates:

$$V = V^\mu \frac{\partial}{\partial x^\mu} = \widetilde{V}^\nu \frac{\partial}{\partial \widetilde{x}^\nu} \qquad \Rightarrow \qquad \widetilde{V}^\nu = V^\mu \frac{\partial \widetilde{x}^\nu}{\partial x^\mu} \ .$$

Notice that indices are contracted in the natural way.

3.2 One-forms

Since T_pM is a vector space, there exists a *dual vector space* T_p^*M given by linear functions $\omega : T_pM \to \mathbb{R}$ and called the **cotangent space** at p. The elements of T_p^*M are called **one-forms**.

A simple example of one-form is the differential $\mathrm{d}f$ of a function $f \in \mathscr{F}(M)$. The action of $\mathrm{d}f \in T_p^*M$ on $V \in T_pM$ is defined as

$$\langle \mathrm{d}f, V\rangle \equiv V[f] = V^\mu \partial_\mu f \in \mathbb{R} \ .$$

Since $\mathrm{d}f = (\partial f/\partial x^\mu)\, \mathrm{d}x^\mu$, we regard $\{\mathrm{d}x^\mu\}$ as a basis of T_p^*M. Notice that, indeed,

$$\left\langle \mathrm{d}x^\mu, \frac{\partial}{\partial x^\nu} \right\rangle = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu \ .$$

Thus $\{\mathrm{d}x^\mu\}$ is the dual basis to $\{\partial_\mu\}$. An arbitrary one-form is written as $\omega = \omega_\mu \mathrm{d}x^\mu$ and the action on a vector $V = V^\mu \partial_\mu$ is

$$\langle \omega, V\rangle = \omega_\mu V^\mu \ .$$

The transformation law of the components under change of coordinates is easily obtained:

$$\widetilde{\omega}_\nu = \omega_\mu \frac{\partial x^\mu}{\partial \widetilde{x}^\nu} \ .$$

It is such that intrinsic objects are coordinate-independent.

[9] We can define a sum of curves through p as follows: use coordinates where p is the origin of $\mathbb{R}^m$, then sum the coordinates of the two curves. We define multiplication of a curve by a real number in a similar way. These definitions depend on the coordinates chosen, however the induced operations on the equivalence classes (namely on tangent vectors) do not.

3.3 Tensors

A **tensor** of type (q, r) can be defined as a multilinear object

$$T \in \mathrm{Hom}\left(T_pM^{\otimes r}, T_pM^{\otimes q}\right) = T_p^*M^{\otimes r} \otimes T_pM^{\otimes q} \; .$$

In components it is written as

$$T = T^{\mu_1 \dots \mu_q}_{\nu_1 \dots \nu_r} \, \partial_{\mu_1} \dots \partial_{\mu_q} \, \mathrm{d}x^{\nu_1} \dots \mathrm{d}x^{\nu_r} \; .$$

We denote the vector space of (q, r) tensors at p by $\mathcal{T}^q_{r,p}M$.

3.4 Tensor fields

If a vector is assigned *smoothly* to each point of M, it is called a **vector field** on M. A vector field is a map

$$V : \mathscr{F}(M) \to \mathscr{F}(M) \; .$$

The set of vector fields on M is denoted as $\mathcal{X}(M)$. A vector field V at p is denoted by $V|_p \in T_pM$.

Similarly, we define a **tensor field** of type (q, r) by a smooth assignment of an element of $\mathcal{T}^q_{r,p}M$ at each point $p \in M$, and we denote the set of tensor fields by $\mathcal{T}^q_r(M)$. Notice that

$$\mathcal{T}^0_0(M) = \mathscr{F}(M) = \Omega^0(M) \; , \qquad \mathcal{T}^1_0(M) = \mathcal{X}(M) \; , \qquad \mathcal{T}^0_1(M) = \Omega^1(M) \; .$$

3.5 Pull-back and push-forward

Consider a smooth map $f : M \to N$ between manifolds. If we have a function $g : N \to \mathbb{R}$ in $\mathscr{F}(N)$, then f can be used to pull it back to M:

$$f^*g \equiv g \circ f : M \to \mathbb{R}$$

is in $\mathscr{F}(M)$.

The map f also induces a map f_* called the **differential map**, or **push-forward** of a vector,

$$f_* : T_pM \to T_{f(p)}N \; .$$

Since a vector in $f(p)$ is a derivative, we define

$$(f_*V)[g] = V[f^*g]$$

where $V \in T_pM$. Writing such a definition in coordinates and setting $V = V^\mu \, \partial/\partial x^\mu$, $f_*V = W^\nu \partial/\partial y^\nu$ and $y^\nu(x^\mu)$ representing f in a chart, we find[10]

$$W^\nu = (f^*V)^\nu = V^\mu \, \frac{\partial y^\nu}{\partial x^\mu} \; .$$

[10]Indeed $(f_*V)[g] = W^\nu \dfrac{\partial g}{\partial y^\nu} = V[f^*g] = V^\mu \dfrac{\partial}{\partial x^\mu}(g \circ f) = V^\mu \dfrac{\partial g}{\partial y^\nu} \dfrac{\partial y^\nu}{\partial x^\mu}$.

This is the same transformation law as for a coordinate transformation, but here it is more general. The map f_* is linear, and equal to the Jacobian $(\partial y^\nu/\partial x^\mu)$ of f at p. The push forward is naturally extended to tensors of type $(q, 0)$.

In general the push-forward cannot be promoted to a map between vector fields: different points on M can be mapped by f to the same point on N, and f could be not surjective. However if f is a diffeomorphism[11] then $f_* : \mathcal{X}(M) \to \mathcal{X}(N)$. [More generally, f could be a diffeomorphism from an open subset of M to an open subset of N.]

The map f also induces a map f^* called the **pull-back**,

$$f^* : T^*_{f(p)}N \to T^*_p M \ .$$

Since a one-form is a linear function on the tangent space, we define

$$\langle f^*\omega, V\rangle = \langle \omega, f_* V\rangle \ .$$

In components

$$(f^*\omega)_\mu = \omega_\nu \, \frac{\partial y^\nu}{\partial x^\mu} \ .$$

If ω is a one-form *field* on N, then $f^*\omega$ is well-defined at all points on M. Therefore the pull-back f^* can be promoted to a map of one-form fields:

$$f^* : \Omega^1(N) \to \Omega^1(M) \ .$$

The pull-back naturally extends to tensor fields of type $(0, r)$.

3.6 Submanifolds

Definition 3.4. Let $f : M \to N$ be a smooth map and $\dim M \leq \dim N$.

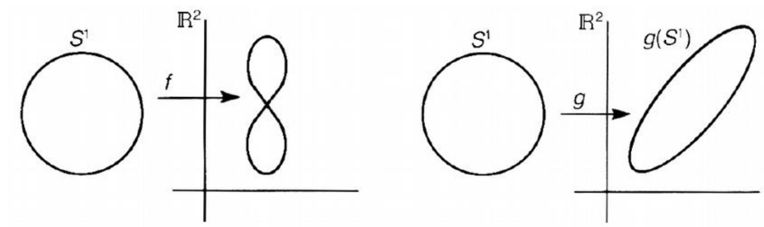
(a) f is an **immersion** of M into N if $f_* : T_pM \to T_{f(p)}N$ has $\operatorname{rank} f_* = \dim M$ (f_* is injective) for all $p \in M$.

(b) f is an **embedding** if f is an immersion and is an homeomorphism (in particular f is injective).

In this case $f(M)$ is a **submanifold** of N. [In practice, $f(M)$ is diffeomorphic to M.]

See Figure.

[11] This condition cannot be weakened. For example, consider the vector field $V = \partial_x$ on $\mathbb{R}$. Consider the map $f : M = \mathbb{R} \to N = \mathbb{R}$ given by $y = f(x) = x^3$. This map is C^∞, invertible, $f^{-1} \in C^0$ but is not smooth. The push-forward of the vector field is $f_*V = 1 \cdot 3x^2 \cdot \partial_y = 3y^{2/3}\partial_y$. The vector f_*V is defined at all points of N, but does not form a vector field because it is not smooth at $y = 0$ (the component is only C^0).



3.7 Lie derivative

Let X be a vector field in M. An **integral curve** $x(t)$ of X is a curve in M, whose tangent vector at $x(t)$ is $X|_{x(t)}$ for all t. Given a chart, the condition can be written in components as[12]

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = X^\mu\big(x(t)\big) \; .$$

This is an ODE, therefore given an initial condition $x_0^\mu = x^\mu(0)$, the solution exists and is unique.

Let $\sigma(t, x_0)$ be an integral curve of X passing through x_0 at $t = 0$ and denote its coordinates by $\sigma^\mu(t, x_0)$. It satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\sigma^\mu(t, x_0) = X^\mu\big(\sigma(t, x_0)\big) \qquad \text{with} \qquad \sigma^\mu(0, x_0) = x_0^\mu \; .$$

The map

$$\sigma : \mathbb{R} \times M \to M$$

is called a **flow** generated by $X \in \mathcal{X}(M)$. Each point of M is "evolved" along the vector field X, and in particular

$$\sigma\big(t, \sigma(s, x_0)\big) = \sigma(t + s, x_0) \; .$$

[¶ The two sides of the equation solve the same ODE.]

If we fix $t \in \mathbb{R}$, the flow $\sigma_t : M \to M$ is a diffeomorphism. In fact, we have a one-parameter commutative group of diffeomorphisms, satisfying

(i) $\sigma_t \circ \sigma_s = \sigma_{t+s}$

(ii) $\sigma_0 = \mathrm{id}$

(iii) $\sigma_{-t} = (\sigma_t)^{-1}$.

The last property guarantees that they are all diffeomorphisms.

[12]This definition does not depend on the coordinates on M. It can be written in a more invariant way as

$$x_* \frac{\mathrm{d}}{\mathrm{d}t} = X\big|_{x(t)} \; .$$

However it does depend on the parametrization of the curve. This is because an integral curve has more structure on $\mathbb{R}$: one obtains a group structure.

We would like to define derivatives, in the direction of a vector X, of the various objects we have constructed on a manifold M. We have already defined the derivative of a function: let us denote it by

$$\mathcal{L}_X f = X[f] = \frac{\mathrm{d}}{\mathrm{d}t}\, f(\sigma_t)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\, \sigma_t^* f\,\Big|_{t=0}\ ,$$

where we have rewritten the definition of a vector X in terms of the flow σ_t it generates.

To compute the derivative of a vector field $Y \in \mathcal{X}(M)$ we encounter a problem: how do we compare vectors at different points of M, since they live in different vector spaces T_pM? Since the points are connected by the flow σ_t generated by X, we can use the push-forward (or differential map) $(\sigma_t)_*$ to map—in a natural way—the vector space at a point into the one at another (since σ_t is a diffeomorphism, the push-forward is a map between vector fields). More precisely we use

$$(\sigma_{-t})_* : T_{\sigma_t(p)}M \to T_pM\ .$$

Thus we define the **Lie derivative** of the vector field Y along the vector field X as

$$\mathcal{L}_X Y = \lim_{t\to 0}\, \frac{1}{t}\Big[(\sigma_{-t})_* Y|_{\sigma_t(p)} - Y|_p \Big] = \frac{\mathrm{d}}{\mathrm{d}t}\, (\sigma_{-t})_* Y|_{\sigma_t(p)}\,\Big|_{t=0}\ .$$

¶ *Exercise* 3. Compute the Lie derivative in components on a chart with coordinates x. Let $X = X^\mu \partial_\mu$ and $Y = Y^\mu \partial_\mu$. The flow of X is

$$\big(\sigma_t(x)\big)^\mu = x^\mu + tX^\mu + \mathcal{O}(t^2)$$

and we can work at first order in t. We have

$$Y|_{\sigma_t(x)} \;\simeq\; Y^\mu\big(\sigma_t(x)\big)\, \partial_\mu\big|_{x+tX} \;\simeq\; \big[Y^\mu(x) + \partial_\rho Y^\mu(x)\cdot tX^\rho \big]\, \partial_\mu\big|_{x+tX}\ .$$

Then we use the formula for the push-forward:

$$(\sigma_{-t})_* Y|_{\sigma_t(x)} \;\simeq\; \big(Y|_{\sigma_t(x)}\big)^\mu\, \partial_\mu(x^\nu - tX^\nu)\, \partial_\nu\big|_x$$

Putting the pieces together, we find

$$\mathcal{L}_X Y = \big(X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu \big)\, \partial_\nu\ .$$

¶ *Exercise* 4. Define the **Lie bracket** $[X,Y]$ of two vector fields $X, Y \in \mathcal{X}(M)$ by

$$[X,Y]f = X[Y[f]] - Y[X[f]]\ ,$$

in terms of $f \in \mathscr{F}(M)$. Show that $[X,Y]$ is a vector field (a field of linear first-order differential operators) given by

$$[X,Y] = \big(X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu \big)\, \partial_\nu\ .$$

These two exercises show that the Lie derivative of Y along X is given by

$$\mathcal{L}_X Y = [X, Y] \ .$$

Proposition 3.5. *The Lie bracket satisfies the following properties:*

(a) Skew-symmetry

$$[X, Y] = -[Y, X] \ .$$

(b) Bilinearity

$$[X, c_1 Y_1 + c_2 Y_2] = c_1 [X, Y_1] + c_2 [X, Y_2]$$

for constants c_1, c_2.

(c) Jacobi identity

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \ .$$

(d) Liebnitz rule

$$[X, fY] = X[f]\, Y + f\, [X, Y] \ .$$

They are evident from the definition.

Remark.

(a) The properties above imply that the Lie derivative satisfies the Liebnitz rule:

$$\mathcal{L}_X\, fY = (\mathcal{L}_X f)\, Y + f\, \mathcal{L}_X Y$$
$$\mathcal{L}_X [Y, Z] = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z] \ .$$

We also have

$$\mathcal{L}_{fX} Y = f\, \mathcal{L}_X Y - Y[f]\, X \ .$$

(b) Given $X, Y \in \mathcal{X}(M)$ and a diffeomorphism $f : M \to N$, one can show that

$$f_*[X, Y] = [f_* X, f_* Y] \ .$$

In a similar way, we can define the Lie derivative of a one-form $\omega \in \Omega^1(M)$ along the vector field X using the pull-back:

$$\mathcal{L}_X \omega = \lim_{t \to 0} \frac{1}{t} \left[(\sigma_t)^* \omega|_{\sigma_t(p)} - \omega|_p \right] = \frac{\mathrm{d}}{\mathrm{d}t} \, (\sigma_t)^* \omega|_{\sigma_t(p)} \Big|_{t=0} \ .$$

¶ *Exercise* 5. With $\omega = \omega_\mu \mathrm{d}x^\mu$, we compute the Lie derivative in components. At first order in t we find

$$(\sigma_t)^* \omega \big|_{\sigma_t(x)} \simeq \left[\omega_\nu(x) + t X^\rho \partial_\rho \omega_\nu(x) \right] \partial_\mu (x^\nu + t X^\nu) \, \mathrm{d}x^\mu \big|_x$$

which leads to

$$\mathcal{L}_X \omega = \left(X^\nu \partial_\nu \omega_\mu + \partial_\mu X^\nu \omega_\nu \right) \mathrm{d}x^\mu \ .$$

The Lie derivative of a tensor field is defined in a similar way. Given a tensor of type (q, r), we can map $\mathcal{T}^q_{r,\sigma_t(p)}M \to \mathcal{T}^q_{r,p}M$ using $(\sigma_{-t})_*^{\otimes q} \otimes (\sigma_t)^{*\,\otimes r}$. However it is more conveniently defined by the following proposition.

Proposition 3.6. *The Lie derivative is completely specified by the following properties:*

(a) For a function and a vector field,

$$\mathcal{L}_X f = X[f] \ , \qquad \mathcal{L}_X Y = [X, Y] \ .$$

(b) For two tensors $t_{1,2}$ of the same type,

$$\mathcal{L}_X(t_1 + t_2) = \mathcal{L}_X t_1 + \mathcal{L}_X t_2 \ .$$

(c) For arbitrary tensors $t_{1,2}$ the Liebnitz rule holds:

$$\mathcal{L}_X(t_1 \otimes t_2) = (\mathcal{L}_X t_1) \otimes t_2 + t_1 \otimes (\mathcal{L}_X t_2) \ .$$

This holds even when the tensors are (partially) contracted.

Proof. It is enough to notice that $(\sigma_{-t})_*^{\otimes q} \otimes (\sigma_t)^{*\,\otimes r}\, T|_{\sigma_t(p)}$, when expanded at first order, is a sum of terms each from the action of a single factor. If $T = t_1 \otimes t_2$, we get the sum of terms from t_1 and the sum of terms from t_2. □

The formula for the Lie derivative of a one-form follows from contraction with an arbitrary vector field Y:

$$\langle \mathcal{L}_X \omega, Y\rangle = \mathcal{L}_X \langle \omega, Y\rangle - \langle \omega, \mathcal{L}_X Y\rangle = X\big[\langle \omega, Y\rangle\big] - \langle \omega, [X, Y]\rangle \ .$$

For an arbitrary tensor field, we write $T = T^{\mu_1 \dots \mu_q}_{\nu_1 \dots \nu_r}\, \partial_{\mu_1} \dots \partial_{\mu_q} \mathrm{d}x^{\nu_1} \dots \mathrm{d}x^{\nu_r}$ and apply the Lie derivative to each factor separately. We find

$$(\mathcal{L}_X T)^{\mu_1 \dots \mu_q}_{\nu_1 \dots \nu_r} = X^\lambda \partial_\lambda T^{\mu_1 \dots \mu_q}_{\nu_1 \dots \nu_r} - \sum_{s=1}^{q} \partial_\lambda X^{\mu_s}\, T^{\mu_1 \dots \lambda \dots \mu_q}_{\nu_1 \dots \nu_r} + \sum_{s=1}^{r} \partial_{\nu_s} X^\lambda\, T^{\mu_1 \dots \mu_q}_{\nu_1 \dots \lambda \dots \nu_r} \ .$$

This reproduces the special cases above.

Proposition 3.7. *For an arbitrary tensor field t:*

$$\mathcal{L}_{[X,Y]} t = \mathcal{L}_X \mathcal{L}_Y t - \mathcal{L}_Y \mathcal{L}_X t \ .$$

Proof. The equality is valid when t is a function or a vector field. Moreover both sides are linear in t and satisfy the Liebnitz rule. It follows that they are equal. □

3.8 Differential forms

Definition 3.8. A **differential form** of order r, or **r-form**, is a *totally antisymmetric* tensor of type $(0, r)$.

A $(0, r)$-tensor ω is a multi-linear function on $(T_pM)^{\otimes r}$. In the coordinate basis,

$$\omega(\partial_{\mu_1}, \dots, \partial_{\mu_r}) = \omega_{\mu_1 \dots \mu_r}$$

gives its components. A totally antisymmetric tensor is such that

$$\omega(V_{P(1)}, \dots, V_{P(r)}) = \operatorname{sgn}(P)\, \omega(V_1, \dots, V_r)$$

where P is a permutation of r elements and $\operatorname{sgn}(P) = +1$ for even permutations and -1 for odd permutations.[13]

Definition 3.9. The **wedge product** $\wedge$ of r one-forms is the totally antisymmetric tensor product

$$\mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} = \sum_{P \in S_r} \operatorname{sgn}(P)\, \mathrm{d}x^{\mu_{P(1)}} \otimes \ldots \otimes \mathrm{d}x^{\mu_{P(r)}} \;. \tag{3.1}$$

For example,

$$\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu - \mathrm{d}x^\nu \otimes \mathrm{d}x^\mu \;.$$

Clearly $\mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} = 0$ if some index μ appears at least twice.

We denote the vector space of r-forms at $p \in M$ by $\Omega^r_p(M)$. A basis is given by the r-forms (3.1), and a general element is expanded as

$$\omega = \frac{1}{r!}\, \omega_{\mu_1 \dots \mu_r}\, \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}$$

where the components $\omega_{\mu_1 \dots \mu_r}$ are totally antisymmetric. We can introduce the *antisymmetrization of indices*:

$$T_{[\mu_1 \dots \mu_r]} = \frac{1}{r!} \sum_{P \in S_r} \operatorname{sgn}(P)\, T_{\mu_{P(1)} \cdots \mu_{P(r)}} \;.$$

Then we can write

$$\omega_{\mu_1 \dots \mu_r} = \omega_{[\mu_1 \dots \mu_r]} \;.$$

The dimension of $\Omega^r_p(M)$ is

$$\binom{m}{r} = \frac{m!}{r!\,(m-r)!} \;.$$

In particular $\Omega^0_p(M) = \mathbb{R}$, $\Omega^1_p(M) = T^*_pM$, and there are no r-forms for $r > m$.

[13] The sign of a permutation P can be defined as the determinant of the matrix that represents the permutation.

Definition 3.10. The **exterior product** of forms,

$$\wedge : \Omega_p^q(M) \times \Omega_p^r(M) \to \Omega_p^{q+r}(M) ,$$

is defined as the linear extension of the same object acting on products of one-forms.

It immediately follows that, if $\xi \in \Omega_p^q(M)$ and $\eta \in \Omega_p^r(M)$,

$$\xi \wedge \eta = (-1)^{qr} \, \eta \wedge \xi$$

and in particular $\xi \wedge \xi = 0$ if q is odd. With this product, we define an *algebra*

$$\Omega_p^\bullet(M) = \Omega_p^0(M) \oplus \Omega_p^1(M) \oplus \ldots \oplus \Omega_p^m(M)$$

of all differential forms at p. The spaces of smooth r-forms on M are denoted as $\Omega^r(M)$, and in particular $\Omega^0(M) = \mathscr{F}(M)$ is the space of functions.

Definition 3.11. The **exterior derivative** is a map (more precisely, a collection of maps) $\mathrm{d} : \Omega^r(M) \to \Omega^{r+1}(M)$ whose action on a form

$$\omega = \frac{1}{r!} \, \omega_{\mu_1 \ldots \mu_r} \, \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}$$

is

$$\mathrm{d}\omega = \frac{1}{r!} \left(\frac{\partial}{\partial x^\nu} \omega_{\mu_1 \ldots \mu_r} \right) \mathrm{d}x^\nu \wedge \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} .$$

Notice that only the antisymmetric part $\partial_{[\nu} \omega_{\mu_1 \ldots \mu_r]}$ contributes to the sum.

Proposition 3.12. *The exterior derivative satisfies the two following important properties:*

(a) Liebnitz rule: let $\xi \in \Omega^q(M)$ *and* $\omega \in \Omega^r(M)$*, then*

$$\mathrm{d}(\xi \wedge \omega) = \mathrm{d}\xi \wedge \omega + (-1)^q \, \xi \wedge \mathrm{d}\omega .$$

(b) Nilpotency:

$$\mathrm{d}^2 = 0 .$$

In fact, these two properties together with $\mathrm{d}f = (\partial f / \partial x^\mu) \, \mathrm{d}x^\mu$ *are an alternative definition of* d.

Proof. (a) can be proved as an exercise. (b) is proven by computing $\mathrm{d}^2\omega$:

$$\mathrm{d}^2\omega = \frac{1}{r!} \, \frac{\partial^2 \omega_{\mu_1 \ldots \mu_r}}{\partial x^\lambda \, \partial x^\nu} \, \mathrm{d}x^\lambda \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} .$$

This vanishes because the derivative is symmetric in $\lambda \leftrightarrow \nu$ while the differential is antisymmetric. Our previous definition of d immediately follows from the properties. □

One can write coordinate-free expressions for the exterior derivative of forms. For instance for one-forms:

$$\mathrm{d}\omega(X, Y) = X[\omega(Y)] - Y[\omega(X)] - \omega([X, Y]) .$$

(¶ It can be proven in components.)

¶ *Exercise* 6. The formula can be generalized to n-forms.

Recall that a map $f : M \to N$ induces the pull-back $f^* : T^*_{f(p)}N \to T^*_p(M)$, which is naturally extended to tensor fields of type $(0,r)$ and thus also to forms.

¶ *Exercise* 7. Let $\xi, \omega \in \Omega^\bullet(M)$ and let $f : M \to N$. Then verify that

$$\begin{aligned} \mathrm{d}(f^*\omega) &= f^*(\mathrm{d}\omega) \\ f^*(\xi \wedge \omega) &= (f^*\xi) \wedge (f^*\omega) \ . \end{aligned}$$

The exterior derivative induces the sequence

$$0 \xrightarrow{i} \Omega^0(M) \xrightarrow{\mathrm{d}} \Omega^1(M) \xrightarrow{\mathrm{d}} \ldots \xrightarrow{\mathrm{d}} \Omega^{m-1}(M) \xrightarrow{\mathrm{d}} \Omega^m(M) \xrightarrow{\mathrm{d}} 0 \ .$$

This is called the **de Rham complex**. Since $\mathrm{d}^2 = 0$, we have

$$\operatorname{im} d \subseteq \ker \mathrm{d} \ .$$

An element of $\ker \mathrm{d}$ is called a **closed form**, namely $\mathrm{d}\omega = 0$. An element of $\operatorname{im} \mathrm{d}$ is called an **exact form**, namely $\omega = \mathrm{d}\psi$ for some ψ. The quotient space $\ker \mathrm{d}_r / \operatorname{im} d_{r-1}$ is called the r-th **de Rham cohomology group**.

Definition 3.13. The **interior product** $\mathrm{i}_X : \Omega^r(M) \to \Omega^{r-1}(M)$, where $X \in \mathcal{X}(M)$ is a vector field, is defined as

$$\mathrm{i}_X\omega(X_1, \ldots, X_{r-1}) = \omega(X, X_1, \ldots, X_{r-1})$$

when $\omega \in \Omega^r(M)$. In components it takes the following form:

$$\mathrm{i}_X\omega = \frac{1}{(r-1)!} X^\nu \, \omega_{\nu\mu_2\ldots\mu_r} \, \mathrm{d}x^{\mu_2} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} \ .$$

Proposition 3.14. *The interior product* i_X *is an* anti-derivation*, meaning that is satisfies the Liebnitz rule*

$$\mathrm{i}_X(\omega \wedge \eta) = \mathrm{i}_X\omega \wedge \eta + (-1)^r \omega \wedge \mathrm{i}_X\eta$$

when $\omega \in \Omega^r(M)$*, and it is nilpotent:*

$$\mathrm{i}_X^2 = 0 \ ,$$

while mapping forms to forms of lower degree.

Proof. (¶) Show as an exercise.

Proposition 3.15. *The Lie derivative of a form can be expressed as*

$$\mathcal{L}_X\omega = (\mathrm{d}\,\mathrm{i}_X + \mathrm{i}_X\mathrm{d})\,\omega$$

where $\omega \in \Omega^r(M)$*. In other words* $\mathcal{L}_X = \{\mathrm{d}, \mathrm{i}_X\}$*.*

Proof. A formal proof is to show that both $\mathcal{L}_X$ and $\{\mathrm{d}, \mathrm{i}_X\}$ satisfy the four defining properties. First, they agree on 0-forms (functions):

$$\mathcal{L}_X f = X[f] = \mathrm{i}_X \mathrm{d} f = (\mathrm{i}_X \mathrm{d} + \mathrm{d}\mathrm{i}_X) f \ .$$

Second, they agree on a coordinate basis of 1-forms:

$$\mathcal{L}_X \mathrm{d}x^\mu = \partial_\nu X^\mu \, \mathrm{d}x^\nu = \mathrm{d}(X^\mu) = (\mathrm{d}\, \mathrm{i}_X + \mathrm{i}_X \mathrm{d}) \mathrm{d}x^\mu \ .$$

Third, both $\mathcal{L}_X$ and $\{\mathrm{d}, \mathrm{i}_X\}$ are linear maps, because both d and i_X are. Finally, one verifies that $\{\mathrm{d}, \mathrm{i}_X\}$ satisfies the (ungraded) Liebnitz rule.

One could check agreement on general one-forms ω in the following way. We evaluate the result on a generic vector Y:

$$\begin{aligned} \big\langle (\mathrm{d}\mathrm{i}_X + \mathrm{i}_X \mathrm{d})\omega, Y \big\rangle &= \big\langle d\langle \omega, X\rangle + \mathrm{i}_X d\omega, Y \big\rangle = Y\big[\langle \omega, X\rangle\big] + \mathrm{i}_Y \mathrm{i}_X d\omega = d\omega(X,Y) + Y\big[\omega(X)\big] \\ &= X\big[\omega(Y)\big] - \omega\big([X,Y]\big) = \mathcal{L}_X \langle \omega, Y\rangle - \langle \omega, \mathcal{L}_X Y\rangle = \langle \mathcal{L}_X \omega, Y\rangle \ . \end{aligned}$$

To go from the first to the second line, we used the coordinate-free expression for the differential of a one-form that we found before.

A more practical proof is in components. Consider the case of a 1-form $\omega = \omega_\mu \mathrm{d}x^\mu$. One computes

$$\begin{aligned} (\mathrm{d}\mathrm{i}_X + \mathrm{i}_X \mathrm{d})\omega &= \mathrm{d}(X^\mu \omega_\mu) + \mathrm{i}_X \big[\tfrac{1}{2} (\partial_\mu \omega_\nu - \partial_\nu \omega_\mu) \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \big] \\ &= \big(\partial_\nu X^\mu \omega_\mu + X^\mu \partial_\nu \omega_\mu\big) \mathrm{d}x^\nu + X^\mu \big(\partial_\mu \omega_\nu - \partial_\nu \omega_\mu\big) \mathrm{d}x^\nu \\ &= (\partial_\nu X^\mu \omega_\mu + X^\mu \partial_\mu \omega_\nu) \mathrm{d}x^\nu = \mathcal{L}_X \omega \end{aligned}$$

comparing with our previous expression for $\mathcal{L}_X \omega$. □

¶ *Exercise* 8. Verify the formula in the general case of an r-form.

¶ *Exercise* 9. Let $X, Y \in \mathcal{X}(M)$. Show that

$$[\mathcal{L}_X, \mathrm{i}_Y] = \mathrm{i}_{[X,Y]} \ .$$

The nilpotency of d and i_X can be used to easily prove

$$[\mathcal{L}_X, \mathrm{i}_X] = 0 \qquad \text{and} \qquad [\mathcal{L}_X, \mathrm{d}] = 0 \ .$$

Proof. We can show that $[\mathcal{L}_X, \mathrm{i}_Y]$ and $\mathrm{i}_{[X,Y]}$ agree on functions and one-forms, and are both anti-derivations. When acting on functions, both operators give zero. When acting of one-forms:

$$\big(\mathcal{L}_X \mathrm{i}_Y - \mathrm{i}_Y \mathcal{L}_X\big)\omega = \mathcal{L}_X \langle \omega, Y\rangle - \langle \mathcal{L}_X \omega, Y\rangle = \langle \omega, \mathcal{L}_X Y\rangle = \mathrm{i}_{[X,Y]}\omega \ .$$

The (graded) Liebnitz rule can be proven easily. □

3.9 Integration of forms

An integration of a differential m-form (top form) over an m-dimensional manifold M is defined only when M is *orientable*. So we first define an **orientation**.

At a point $p \in M$, the tangent space T_pM is spanned by the basis $\{\partial_\mu \equiv \partial/\partial x^\mu\}$ in terms of the local coordinates x^μ on the chart U_i to which p belongs. Let U_j be another chart with coordinates $\{\widetilde{x}^\nu\}$, such that $p \in U_i \cap U_j$. Then T_pM is also spanned by $\{\widetilde{\partial}_\nu = \partial/\partial\widetilde{x}^\nu\}$. The change of basis is

$$\widetilde{\partial}_\nu = \frac{\partial x^\mu}{\partial \widetilde{x}^\nu}\,\partial_\mu \;.$$

If $J = \det\big(\partial x^\mu/\partial\widetilde{x}^\nu\big) > 0$ on $U_i \cap U_j$, then $\{\partial_\mu\}$ and $\{\widetilde{\partial}_\nu\}$ are said to define the same *orientation* on $U_i \cap U_j$, otherwise if $J < 0$ they define opposite orientations.

Definition 3.16. Let M be a connected manifold covered by $\{U_i\}$. Then M is **orientable** if there exist local coordinates $\{x^\mu_{(i)}\}$ on U_i such that, for any overlapping charts U_i and U_j, one has $J = \det\big(\partial x^\mu_{(i)}/\partial x^\nu_{(j)}\big) > 0$.

Definition 3.17. An m-form (top form) η which nowhere vanishes on a manifold M is called a **volume form**.

Proposition 3.18. *A volume form η exists on a manifold M if and only if M is orientable, and η defines an orientation.*

Proof. Suppose M is orientable. Start with a chart U and take an m-form

$$\eta = h(p)\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^m$$

with $h(p)$ a positive-definite function on U. Then extend η to the whole M, in such a way that it never vanishes. When you transit from U_i to U_j, choose coordinates that preserve the orientation. In the new coordinates

$$\eta = h(p)\,\det\Big(\frac{\partial x^\mu}{\partial y^\nu}\Big)\,\mathrm{d}y^1 \wedge \ldots \wedge \mathrm{d}y^m \;,$$

namely the component of η transforms with the Jacobian J. The component is still positive-definite on $U_i \cap U_j$ thanks to orientability, $J > 0$, and can be extended in a positive-definite way on U_j. Proceeding in this way, one can consistently construct a volume form on M.

Conversely, suppose a volume form η on M is given. Since it is nowhere vanishing, its component on any chart U_i is either positive or negative. On the charts where it is negative, perform the change of coordinates $x^1 \to -x^1$ which makes it positive. The resulting set of charts is then oriented. □

Recall that a manifold M with boundary is such that the open sets $\{U_i\}$ are mapped to open sets in $\mathbb{H}^m = \big\{(x^1, \ldots, x^m)\,\big|\,(-1)^m x^m \geq 0\big\}$, [This choice avoids unpleasant signs in Stokes' theorem] and the boundary ∂M consists of the points mapped to points with $x^m = 0$.

This automatically provides an atlas of the boundary ∂M, which is itself a manifold. If M is oriented, the atlas for ∂M is also oriented. This is because the Jacobian evaluated at the boundary is

$$\frac{\partial x^\mu}{\partial \widetilde{x}^\nu}\Big|_{\widetilde{x}^m=0} = \left(\begin{array}{c|c} \dfrac{\partial x^i}{\partial \widetilde{x}^j} & \dfrac{\partial x^m}{\partial \widetilde{x}^j} \\ \hline 0 & \dfrac{\partial x^m}{\partial \widetilde{x}^m} > 0 \end{array}\right) > 0 \qquad\qquad \text{with } i,j = 1,\ldots,m-1 \;.$$

This leads to

Proposition 3.19. *Given an oriented manifold M with boundary ∂M, the boundary inherits an orientation from M.*

Consider an oriented manifold M. The integral of a top form ω on a coordinate neighborhood U_i with coordinates x^μ is defined by

$$\int_{U_i} \omega \;\equiv\; \int_{\varphi_i(U_i)} \omega_{1\ldots m}\big(\varphi_i^{-1}(x)\big)\, \mathrm{d}x^1 \ldots \mathrm{d}x^m \;.$$

This definition is coordinate-invariant, as long as we consider oriented changes of coordinates, because $\omega_{1\ldots m}$ transforms with the Jacobian while $\mathrm{d}x^1 \ldots \mathrm{d}x^m$ transforms with the absolute value of the Jacobian.

Then we want to define the integral on M of a top-form ω with compact support. The last assumption ensures that we can cover the support with a finite number of open patches.

Definition 3.20. Take an open covering $\{U_i\}$ of M. A **partition of unity** subordinate to the covering $\{U_i\}$ is a family of differentiable functions $\{\varepsilon_i(p)\}$ satisfying

(i) $0 \leq \varepsilon_i(p) \leq 1$

(ii) $\varepsilon_i(p) = 0$ if $p \notin U_i$

(iii) $\varepsilon_1(p) + \varepsilon_2(p) + \ldots = 1$ for any point $p \in M$.

It follows that

$$\omega(p) = \sum\nolimits_i \varepsilon_i(p)\, \omega(p) \;,$$

and each term in the sum can be integrated on U_i because it vanishes outside. Thus we can define

$$\int_M \omega \;\equiv\; \sum_i \int_{U_i} \varepsilon_i\, \omega \;.$$

Notice that, since the support of ω is compact, we can cover it with a finite number of U_i and thus the sum is finite.

Theorem 3.21. (Stokes' theorem) *Let $\omega \in \Omega^{m-1}(M)$. Then*

$$\int_M \mathrm{d}\omega = \int_{\partial M} \omega \;.$$

If M has no boundary, the integral vanishes.

Proof. Using a partition of unity, we can reduce the formula to coordinate patches:

$$\sum_i \int_{U_i} \mathrm{d}(\varepsilon_i\,\omega) = \sum_i \int_{U_i\cap\partial M} \varepsilon_i\,\omega \ .$$

Given $\varepsilon_i\omega = \frac{1}{(m-1)!}\varepsilon_i\omega_{\mu_1\ldots\mu_{m-1}}\mathrm{d}x^{\mu_1}\wedge\ldots\wedge\mathrm{d}x^{\mu_{m-1}}$, its differential can be written as

$$\mathrm{d}(\varepsilon_i\omega) = \sum_{\mu=1}^{m} (-1)^{\mu-1}\,\partial_\mu\big(\varepsilon_i\,\omega_{1\ldots\hat{\mu}\ldots m}\big)\ \mathrm{d}x^1\wedge\ldots\wedge\mathrm{d}x^m\ ,$$

where $\hat{\mu}$ means that we omit that index. Recalling that coordinate neighborhoods U_i are mapped by φ_i to open subsets in $\mathbb{H}^m = \{x\in\mathbb{R}^m\,|\,(-1)^m x^m\geq 0\}$, we integrate by parts:

$$\begin{aligned}\int_{U_i}\mathrm{d}\big(\varepsilon_i\,\omega\big) &= \sum_{\mu=1}^{m}(-1)^{\mu-1}\int_{\{(-1)^m x^m\geq 0\}}\partial_\mu\big(\varepsilon_i\,\omega_{1\ldots\hat{\mu}\ldots m}\big)\ \mathrm{d}x^1\ldots\mathrm{d}x^m\\ &= (-1)^{m-1}\int_{\{(-1)^m x^m\geq 0\}}\partial_m\big(\varepsilon_i\,\omega_{1\ldots m-1}\big)\ \mathrm{d}x^1\ldots\mathrm{d}x^m\\ &= \int_{\{x^m=0\}}\varepsilon_i\ \omega_{1\ldots m-1}\ \mathrm{d}x^1\ldots\mathrm{d}x^{m-1} = \int_{U_i\cap\partial M}\varepsilon_i\,\omega\ .\end{aligned}$$

In the second equality we have picked the only term that, after integration by parts, has a chance to give a non-vanishing result. □

4 Lie groups and Lie algebras

We will give a first look at Lie groups in the context of differential geometry.

4.1 Groups

A **group** is a set G with a binary operation (product)

$$\cdot\ : G \times G \to G$$

that combines two elements a and b into another element, ab. The product must satisfy:

(i) **Associativity**: for all $a, b, c \in G$:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c)\ .$$

(ii) **Identity element**: there exists an element $\mathbb{1} \in G$ such that

$$\mathbb{1} \cdot a = a \cdot \mathbb{1} = a$$

for all $a \in G$.

(iii) **Inverse element**: for each $a \in G$, there exists an element a^{-1} such that

$$a \cdot a^{-1} = a^{-1} \cdot a = \mathbb{1}\ .$$

Notice that both the identity element and the inverse of a are unique. Suppose $\mathbb{1}$ and $\mathbb{1}'$ are identify elements, then $\mathbb{1} = \mathbb{1} \cdot \mathbb{1}' = \mathbb{1}'$. Suppose that b and b' are inverses of a, then

$$b = b \cdot \mathbb{1} = b \cdot (a \cdot b') = (b \cdot a) \cdot b' = \mathbb{1} \cdot b' = b'\ .$$

In general the order of the two factors in the product matters, namely $a \cdot b \neq b \cdot a$. If

$$a \cdot b = b \cdot a \qquad\qquad \forall\, a, b \in G$$

then G is an **Abelian group** (and the group operation can be represented as $+$). Groups, as sets, can be either finite, countably infinite or continuous.

Example 4.1. Consider the following examples of groups:

(a) The integers $(\mathbb{Z}, +)$ with addition (Abelian group).

(b) The non-zero rationals $(\mathcal{Q} \setminus \{0\}, \cdot)$ with multiplication (Abelian group).

(c) The positive real numbers $(\mathbb{R}_+, \cdot)$ with multiplication (Abelian).

(d) The integers modulo n $(\mathbb{Z}_n \equiv \mathbb{Z}/n\mathbb{Z}, +)$ with addition (Abelian).

(e) The group of permutations of n objects $(S_n, \circ)$ with composition.

(f) The group of phases $e^{i\alpha}$ with $\alpha \in [0, 2\pi)$, $(S^1 \equiv U(1), \cdot)$ with multiplication (Abelian).

(g) The group of $n \times n$ matrices with non-vanishing determinant, $(GL(n), \cdot)$ with matrix multiplication.

4.2 Lie groups

Definition 4.2. A **Lie group** G is a differentiable manifold which is endowed with a group structure such that the group operations

(i) $\cdot : G \times G \to G$

(ii) ${}^{-1} : G \to G$

are differentiable.

The unit element is written as e or $\mathbb{1}$. The dimension of a Lie group is the dimension of G as a manifold.

Example 4.3. The following are examples of Lie groups.

(a) Let S^1 be the unit circle on the complex plane,

$$S^1 = \left\{ e^{i\theta} \,\middle|\, \theta \in \mathbb{R} \quad (\text{mod } 2\pi) \right\} ,$$

and take the group operation $e^{i\theta} e^{i\varphi} = e^{i(\theta+\varphi)}$ and $(e^{i\theta})^{-1} = e^{-i\theta}$, which are differentiable. This makes S^1 into a Lie group, called $U(1)$.

(b) The **general linear group** $GL(n, \mathbb{R})$ of $n \times n$ real matrices with non-vanishing determinant is a Lie group, with the operations of matrix multiplication and inverse. Its dimension is n^2 and it is non-compact. Interesting Lie subgroups are:

orthogonal	$O(n) = \left\{ M \in GL(n, \mathbb{R}) \,\middle|\, M M^\mathsf{T} = \mathbb{1}_n \right\}$
special linear	$SL(n, \mathbb{R}) = \left\{ M \in GL(n, \mathbb{R}) \,\middle|\, \det M = 1 \right\}$
special orthogonal	$SO(n) = O(n) \cap SL(n, \mathbb{R})$
(real) symplectic	$Sp(2n, \mathbb{R}) = \left\{ M \in GL(2n, \mathbb{R}) \,\middle|\, M\Omega M^\mathsf{T} = \Omega \right\} \subset SL(2n, \mathbb{R})$

where the symplectic unit is $\Omega = \left(\begin{smallmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{smallmatrix}\right)$.[14]

An interesting fact is $Sp(2n, \mathbb{R}) \cap SO(2n) \cong U(n)$.

(c) The **Lorentz group** is

$$O(3,1) = \left\{ M \in GL(4, \mathbb{R}) \,\middle|\, M\eta M^\mathsf{T} = \eta \right\}$$

where $\eta = \text{diag}(-1, 1, 1, 1)$ is the Minkowski metric. This is non-compact and has 4 connected components, distinguished by the sign of the determinant and the sign of the $(0, 0)$ entry.[15] ¶

[14] The fact that a symplectic matrix has determinant 1 can be shown as follows. The Pfaffian of an antisymmetric matrix $A = -A^\mathsf{T}$ is defined as $\text{Pf}(A) = \frac{1}{2^n n!} A_{a_1 a_2} \ldots A_{a_{n-1} a_n} \epsilon^{a_1 \ldots a_n}$. It follows that

$$\text{Pf}\big(BAB^\mathsf{T}\big) = \tfrac{1}{2^n n!} B_{a_1 j_1} A_{j_1 j_2} B_{a_2 j_2} \ldots B_{a_{n-1} j_{n-1}} A_{j_{n-1} j_n} B_{a_n j_n} \epsilon^{a_1 \ldots a_n} = \det(B)\, \text{Pf}(A) \ .$$

Therefore the defining equation implies $\det M = 1$.

[15] From the equation it follows $\det M = \pm 1$ and $M_{00} \neq 0$.

(d) The **general linear group** $GL(n,\mathbb{C})$ of $n \times n$ complex matrices with non-vanishing determinant has (real) dimension $2n^2$. Interesting subgroups are:

unitary	$U(n) = \{M \in GL(n,\mathbb{C}) \,\big|\, MM^\dagger = \mathbb{1}_n\}$
special linear	$SL(n,\mathbb{C}) = \{M \in GL(n,\mathbb{C}) \,\big|\, \det M = 1\}$
special unitary	$SU(n) = U(n) \cap SL(n,\mathbb{C})$
(complex) symplectic	$Sp(2n,\mathbb{C}) = \{M \in GL(2n,\mathbb{C}) \,\big|\, M\Omega M^\mathsf{T} = \Omega\} \subset SL(2n,\mathbb{C})$
compact symplectic	$USp(2n) = Sp(2n,\mathbb{C}) \cap U(2n)$.

Let G be a Lie group and $H \subset G$ a Lie subgroup of G. Define an equivalence relation $g \sim g'$ if there exists an element $h \in H$ such that $g' = gh$. An equivalence class $[g]$ is a set $\{gh \,|\, h \in H\}$. The **coset space**

$$G/H = \{[g] \,\big|\, g \in G\}$$

is the set of equivalence classes, and it is a manifold with $\dim G/H = \dim G - \dim H$. H is said to be a **normal subgroup** of G if

$$ghg^{-1} \in H \qquad \text{for any } g \in G \text{ and } h \in H \;,$$

i.e. if H is preserved under the adjoint action of G. When H is a normal subgroup of G, then G/H is a (Lie) group. The group operations are simply defined as

$$[g] \cdot [g'] = [gg'] \;, \qquad\qquad [g]^{-1} = [g^{-1}] \;.$$

Let us check that they are well-defined. If gh and $g'h'$ are representatives of $[g]$ and $[g']$, respectively, then

$$ghg'h' = gg'g'^{-1}hg'h' = gg'h''h'$$

in the same class $[gg']$. Similarly

$$(gh)^{-1} = h^{-1}g^{-1} = g^{-1}gh^{-1}g^{-1} = g^{-1}h''$$

in the same class $[g^{-1}]$.

4.3 Action of Lie groups on manifolds

Lie groups often appear as set of transformations acting on manifolds. We already discussed one example: a vector field X defines a *flow* on M, which is a map $\sigma : \mathbb{R} \times M \to M$ in which $\mathbb{R}$ acts as an additive group. We can generalize the idea.

Definition 4.4. Let G be a Lie group and M be a manifold. The **action** of G on M is a differentiable map $\sigma : G \times M \to M$ such that

(i) $\sigma(\mathbb{1}, p) = p \qquad$ for all $p \in M$

(ii) $\sigma\big(g_1,\, \sigma(g_2, p)\big) = \sigma(g_1 g_2,\, p)$.

In other words the action respects the group structure. We use the notation gp for $\sigma(g, p)$.

The action is said to be

(a) **transitive** if, for any $p_1, p_2 \in M$, there exists an element $g \in G$ such that $g\, p_1 = p_2$;

(b) **free** if every $g \neq \mathbb{1}$ of G has no fixed points on M, namely if the existence of a point $p \in M$ such that $g\, p = p$ necessarily implies $g = \mathbb{1}$;

(c) **faithful** or **effective** if the unit element $\mathbb{1}$ is the only element that acts trivially on M.

4.4 Lie algebras

Definition 4.5. Let a, g be elements of a Lie group G. The **left-translation** $L_a : G \to G$ and the **right-translation** $R_a : G \to G$ are defined by

$$L_a g = ag \ , \qquad\qquad R_a g = ga \ .$$

Clearly L_a, R_a are diffeomorphisms from G to G, hence they induce differential maps $L_{a*} : T_g G \to T_{ag} G$ and $R_{a*} : T_g G \to T_{ga} G$. Since the two cases are equivalent, in the following we discuss left-translations.

Given a Lie group G, there exists a special class of vector fields characterized by the invariance under group action.

Definition 4.6. Let X be a vector field on a Lie group G. Then X is said to be a **left-invariant vector field** if $L_{a*} X|_g = X|_{ag}$.

A vector $V \in T_{\mathbb{1}} G$ defines a *unique* left-invariant vector field X_V on G by

$$X_V\Big|_g = L_{g*} V \qquad \text{for } g \in G \ .$$

In fact $X_V|_{ag} = L_{ag*} V = L_{a*} L_{g*} V = L_{a*} X_V|_g$ thus X_V is left-invariant. Conversely, a left-invariant vector field X defines a unique vector $V = X|_{\mathbb{1}} \in T_{\mathbb{1}} G$. We denote the set of left-invariant vector fields on G by $\mathfrak{g}$. The map $T_{\mathbb{1}} G \leftrightarrow \mathfrak{g}$ defined by $V \leftrightarrow X_V$ is an isomorphism, in particular $\dim \mathfrak{g} = \dim G$.

Since $\mathfrak{g}$ is a set of vector fields, it is a subset of $\mathcal{X}(G)$ and the Lie bracket is defined on $\mathfrak{g}$. We show that $\mathfrak{g}$ is closed under the Lie bracket. Let $X, Y \in \mathfrak{g}$, then

$$[X, Y]\Big|_{ag} = [X|_{ag}, Y|_{ag}] = [L_{a*} X|_g, L_{a*} Y|_g] = L_{a*}[X, Y]\Big|_g \ ,$$

implying that $[X, Y] \in \mathfrak{g}$.

Definition 4.7. The set of left-invariant vector fields $\mathfrak{g}$ with the Lie bracket $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is called the **Lie algebra** of a Lie group G.

Let the set of n vectors $\{V_1, \ldots, V_n\}$ be a basis of $T_{\mathbb{1}}G$, where $n = \dim G$ and we consider n finite. The basis defines a set of n linearly independent left-invariant vector fields $\{X_1, \ldots, X_n\}$ on G. Since $[X_a, X_b]$ is again an element of $\mathfrak{g}$, it can be expanded in terms of $\{X_a\}$:

$$[X_a, X_b] = C^c_{ab}\, X_c \ ,$$

The coefficients C^c_{ab} are called the **structure constants** of the Lie group G.

¶ *Exercise* 10. Show that the structure constants satisfy

(a) skew-symmetry

$$C^c_{ab} = -C^c_{ba}$$

(b) Jacobi identity

$$C^f_{ab} C^g_{fc} + C^f_{ca} C^g_{fb} + C^f_{bc} C^g_{fa} = 0 \ .$$

Let us introduce a dual basis to $\{X_a\}$ and denote it by $\{\theta^a\}$, in other words

$$\langle \theta^a, X_b \rangle = \delta^a_b \qquad \text{at all points } g \in G \ .$$

Then $\{\theta^a\}$ is a basis for the left-invariant one-forms. We can show that the dual basis satisfies **Maurer–Cartan's structure equations**:

$$\mathrm{d}\theta^a = -\frac{1}{2}\, C^a_{bc}\, \theta^b \wedge \theta^c \ .$$

To prove it we use the coordinate-free expression for the differential of a one-form:[16]

$$\begin{aligned} \mathrm{d}\theta^a(X_b, X_c) &= X_b[\theta^a(X_c)] - X_c[\theta^a(X_b)] - \theta^a([X_b, X_c]) \\ &= X_b[\delta^a_c] - X_c[\delta^a_b] - \theta^a(C^d_{bc} X_d) = -C^a_{bc} \ . \end{aligned}$$

We can define a Lie-algebra valued 1-form $\theta : T_gG \to T_{\mathbb{1}}G$ as

$$\theta : X \to (L_{g^{-1}})_* X = (L_g)^{-1}_* X \qquad \text{for } X \in T_gG \ .$$

This is called the **Maurer–Cartan form** on G.

Theorem 4.8. *The Maurer-Cartan form θ is*

$$\theta = \theta^a \otimes V_a$$

where $\{\theta^a\}$ is the dual basis of left-invariant one-forms and $\{V_a\}$ is a basis of $T_{\mathbb{1}}G$. Then θ satisfies

$$\mathrm{d}\theta + \theta \wedge \theta \ \equiv \ \mathrm{d}\theta^a \otimes V_a + \frac{1}{2}\, \theta^b \wedge \theta^c \otimes [V_b, V_c] = 0 \ ,$$

where the commutator is the one in the Lie algebra.

[16] Such a formula is: $\mathrm{d}\omega(X, Y) = X\big[\omega(Y)\big] - Y\big[\omega(X)\big] - \omega\big([X, Y]\big)$.

Proof. Take a vector $Y \equiv Y^a X_a \in T_g G$, where $\{X_a\}$ are left-invariant vector fields with $X_a|_g = (L_g)_* V_a$. From the definition of θ:

$$\theta(Y) = Y^a \, \theta(X_a) = Y^a (L_g)_*^{-1} (L_g)_* V_a = Y^a \, V_a = \theta^a(Y) \otimes V_a \ ,$$

and in the last step we used that $\{\theta^a\}$ is the dual basis to $\{X_a\}$.

The Maurer–Cartan structure equation gives

$$\mathrm{d}\theta + \theta \wedge \theta = -\frac{1}{2} \, C^a_{bc} \, \theta^b \wedge \theta^c \otimes V_a + \frac{1}{2} \, \theta^b \wedge \theta^c \otimes C^a_{bc} \, V_a = 0 \ .$$

□

4.5 Exponential map

We saw that a vector field $X \in \mathcal{X}(M)$ generates a flow $\sigma(t,x)$ in M. Now we want to consider the flow $\sigma(t,g)$ in a Lie group G generated by a left-invariant vector field $X \in \mathfrak{g}$. The flow satisfies

$$\sigma(\cdot, g)_* \frac{\mathrm{d}}{\mathrm{d}t}\Big|_t = X\big|_{\sigma(t,g)} \ ,$$

where $\sigma(\cdot, g)$ is a map $\mathbb{R} \to G$. We define a **one-parameter subgroup** of G

$$\phi_X(t) \ \equiv \ \sigma(t, \mathbb{1}) \ .$$

It is not obvious that the integral curve of X through the identity forms a subgroup of G (with the product induced by G).[17] This is shown in the

Proposition 4.9. *The one-parameter subgroup of G generated by $X \in \mathfrak{g}$ satisfies*

$$\phi_X(s+t) = \phi_X(s) \, \phi_X(t) \ , \qquad\qquad \phi_X(0) = \mathbb{1} \ .$$

Proof. The second equation is obvious. We compare the two sides of the first equation as functions of t (we keep X implicit). At $t=0$ they agree. The function on the LHS satisfies

$$\phi(s + \ \cdot)_* \frac{\mathrm{d}}{\mathrm{d}t}\Big|_t = X\big|_{\phi(s+t)} \ .$$

To compute the derivative of the RHS we use that X is left-invariant. We have

$$\big(\phi(s) \, \phi(\,\cdot\,)\big)_* \frac{\mathrm{d}}{\mathrm{d}t}\Big|_t = \big(L_{\phi(s)} \, \sigma(\cdot, \mathbb{1})\big)_* \frac{\mathrm{d}}{\mathrm{d}t}\Big|_t = L_{\phi(s)*} \, \sigma(\cdot, \mathbb{1})_* \frac{\mathrm{d}}{\mathrm{d}t}\Big|_t = L_{\phi(s)*} \, X\big|_{\phi(t)} = X\big|_{\phi(s)\,\phi(t)} \ .$$

Since the two sides of the equation solve the same ODE and have the same initial condition, they agree for all t's.

Notice that the statement is different from $\sigma\big(t, \sigma(s, \mathbb{1})\big) = \sigma(t+s, \mathbb{1})$ that we already proved. □

[17] In fact, while the curve $\sigma(\mathbb{R}, \mathbb{1})$ is an immersion of $\mathbb{R}$ into G for any vector field X, it actually forms a subgroup of G only when X is a left-invariant vector field.

Definition 4.10. Let G be a Lie group and $V \in \mathfrak{g}$ (equivalently, $V \in T_{\mathbb{1}}G$). The **exponential map** $\exp : \mathfrak{g} \to G$ is defined by

$$\exp V = \phi_V(1) \ ,$$

and it satisfies

$$\exp(t\,V) = \phi_V(t) \ .$$

¶ *Exercise* 11. The last part should be proven, namely that $\phi_{tV}(1) = \phi_V(t)$. Show that both $\phi_{\lambda V}(t)$ and $\phi_V(\lambda t)$, where $\lambda \in \mathbb{R}$ is fixed, are one-parameter subgroups of G generated by λV, hence they are equal. The claim follows by setting $t = 1$ and $\lambda \to t$.

5 Riemannian geometry

Besides a smooth structure, a manifold may carry a further structure: a metric. This is an inner product between vectors in the tangent space.

Definition 5.1. A **Riemannian metric** g on a differentiable manifold M is a type $(0,2)$ tensor field on M such that, at each point $p \in M$:

(i) $g_p(U,V) = g_p(V,U)$;

(ii) $g_p(U,U) \geq 0$, where equality implies $U = 0$.

In short, g_p is a symmetric positive-definite bilinear form.

A **pseudo-Riemannian metric** is such that the bilinear form is symmetric and non-degenerate,[18] but not necessarily positive-definite. A special case is a Lorentzian metric, whose signature has only one negative entry.

In coordinates, we expand the metric as

$$g_p = g_{\mu\nu}(p)\, \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu \;\equiv\; \mathrm{d}s^2 \;.$$

The notation $\mathrm{d}s^2$ is used in physics, because the metric is an infinitesimal distance squared. We regard $g_{\mu\nu}$ at a given point as a matrix, and since it has maximal rank, it has an inverse denoted by $g^{\mu\nu}$. The determinant $\det(g_{\mu\nu})$ is sometimes denoted by g.

The metric g_p is a map $T_pM \otimes T_pM \to \mathbb{R}$, and for each vector $U \in T_pM$ it defines a linear map

$$\begin{aligned} g_p(U,\cdot) :\; T_pM &\to \mathbb{R} \\ V &\to g_p(U,V) \;, \end{aligned}$$

which is a one-form $\omega_U \in T_p^*M$. Since g_p is non-degenerate, this gives an isomorphism between T_pM and T_p^*M. In components

$$\omega_\mu = g_{\mu\nu} U^\nu \;, \qquad\qquad U^\mu = g^{\mu\nu}\omega_\nu \;.$$

If a smooth manifold M is equipped with a Riemannian metric g, the pair (M,g) is called a **Riemannian manifold**. In fact, all manifolds admit a Riemannian metric.[19] This is not so for Lorentzian or other pseudo-Riemannian metrics.

If $f : M \to N$ is an *immersion* of an m-dimensional manifold M into an n-dimensional manifold N with Riemannian metric g_N, then the pullback map f^* induces the natural metric $g_M = f^* g_N$ on M, called the **induced metric**. In components:

$$g_{M\mu\nu}(x) = g_{N\alpha\beta}\big(f(x)\big)\, \frac{\partial f^\alpha}{\partial x^\mu}\, \frac{\partial f^\beta}{\partial x^\nu} \;.$$

[18] The metric is non-degenerate if $g_p(U,V) = 0$ for every $V \in T_pM$ implies that $U = 0$.

[19] Simply, on every patch U_i introduce a Riemannian metric $g_p^{(i)}$ with compact support. Then sum all the metrics: $g_p = \sum_i g_p^{(i)}$ (one should use a locally-finite covering to avoid divergences). One uses that the sum of Riemannian metrics is a Riemannian metric g_p.

If N is pseudo-Riemannian, f^*g_N is not guaranteed to be a metric.

¶ *Exercise* 12. Can you give an example?

¶ *Exercise* 13. Consider the unit sphere S^2 with coordinates (θ, ϕ) embedded into $\mathbb{R}^3$ as

$$f : (\theta, \phi) \rightarrow \big(\sin\theta \, \cos\phi, \sin\theta \, \sin\phi, \cos\theta\big) .$$

Taking the Euclidean metric on $\mathbb{R}^3$, compute the induced metric on $\mathbb{R}^2$.

5.1 Connections

A vector X is a "directional derivative" acting on functions $f \in \mathscr{F}(M)$ as $X : f \rightarrow X[f]$. There is no natural directional derivative acting on tensor fields of type (p, q): the Lie derivative $\mathcal{L}_X Y = [X, Y]$ is not a directional derivative, because it requires X to be a vector field, not just a vector (it depends on the *derivative* of X). What we need is an extra structure, a *connection*, which specifies how tensors are transported along a curve.

Definition 5.2. An **affine connection** ∇ is a map $\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M)$, in short $\nabla_X Y$, such that:

$$\begin{aligned} \nabla_X(Y + Z) &= \nabla_X Y + \nabla_X Z \\ \nabla_{(X+Y)} Z &= \nabla_X Z + \nabla_Y Z \\ \nabla_{fX} Y &= f \, \nabla_X Y \\ \nabla_X(fY) &= X[f] \, Y + f \, \nabla_X Y . \end{aligned}$$

The third property distinguishes the Lie derivative from an affine connection.[20] ∇ is also called a **covariant derivative**.

Given a coordinate basis $\{e_\mu\} = \{\partial/\partial x^\mu\}$ on a chart, one can define the **connection coefficients** $\Gamma^\lambda_{\mu\nu}$ as

$$\nabla_\mu e_\nu \equiv \nabla_{e_\mu} e_\nu = \Gamma^\lambda_{\mu\nu} \, e_\lambda .$$

They characterize the action of the connection ∇ on any vector field:

$$\nabla_V W = V^\mu \, \nabla_{e_\mu}\big(W^\nu e_\nu\big) = V^\mu \left(\frac{\partial W^\lambda}{\partial x^\mu} + W^\nu \, \Gamma^\lambda_{\mu\nu} \right) e_\lambda .$$

Let $c : (a, b) \rightarrow M$ be a curve in M, with tangent vector $V = \mathrm{d}x^\mu\big(c(t)\big)/\mathrm{d}t \, e_\mu|_{c(t)}$. Let X be a vector field, defined at least along $c(t)$. If X satisfies the condition

$$\nabla_V X = 0 \qquad\qquad \forall t \in (a, b) ,$$

X is said to be **parallel transported** along $c(t)$. If the tangent vector $V(t)$ itself is parallel transported along $c(t)$, namely if

$$\nabla_V V = 0 ,$$

[20] Recall that $\mathcal{L}_{fX} Y = f \mathcal{L}_X Y - Y[f] \, X$.

the curve $c(t)$ is called a **geodesic**. In components the geodesic equation is

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}t^2} + \Gamma^\mu_{\nu\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}t} \frac{\mathrm{d}x^\lambda}{\mathrm{d}t} = 0 \ .$$

Notice that this equation is not invariant under reparametrizations.

¶ *Exercise* 14. Show that if a curve $c(t)$ satisfies the weaker condition

$$\nabla_V V = f\, V \qquad\qquad \text{with } f \in \mathscr{F}(M) \ ,$$

it is always possible to change parametrization $t \to t'$ such that $c(t')$ is a geodesic.

It is natural do define the covariant derivative of a function $f \in \mathscr{F}(M)$ as the ordinary directional derivative:

$$\nabla_X f = X[f] \ .$$

Then the fourth property above looks like the Liebnitz rule. We require that this be true for any product of tensors,

$$\nabla_X (T_1 \otimes T_2) = \nabla_X T_1 \otimes T_2 + T_1 \otimes \nabla_X T_2 \ ,$$

even when some of the indices are contracted. This fixes the action of ∇ on tensors of general type. For instance, using

$$X\big[\langle \omega, Y\rangle\big] = \nabla_X \langle \omega, Y\rangle = \langle \nabla_X \omega, Y\rangle + \langle \omega, \nabla_X Y\rangle$$

where $\omega \in \Omega^1(M)$ is a one-form, one obtains the action in components:

$$(\nabla_\mu \omega)_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu} \omega_\lambda \ .$$

The action on tensors of general type is:

$$(\nabla_\mu T)^{\nu_1 \ldots \nu_p}_{\lambda_1 \ldots \lambda_q} = \partial_\mu T^{\nu_1 \ldots \nu_p}_{\lambda_1 \ldots \lambda_q} + \sum_{s=1}^p \Gamma^{\nu_s}_{\mu\rho} T^{\nu_1 \ldots \rho \ldots \nu_p}_{\lambda_1 \ldots \lambda_q} - \sum_{s=1}^q \Gamma^\rho_{\mu\lambda_s} T^{\nu_1 \ldots \nu_p}_{\lambda_1 \ldots \rho \ldots \lambda_q} \ .$$

Consider two overlapping charts U, V with $U \cap V \neq \varnothing$, and with coordinates x^μ and $\widetilde{x}^\mu$, respectively. Imposing that $\nabla_X Y$ transforms as a tensor, imposes that the connection coefficients transform as

$$\widetilde{\Gamma}^\gamma_{\alpha\beta} = \frac{\partial x^\mu}{\partial \widetilde{x}^\alpha} \frac{\partial x^\nu}{\partial \widetilde{x}^\beta} \frac{\partial \widetilde{x}^\gamma}{\partial x^\lambda} \Gamma^\lambda_{\mu\nu} + \frac{\partial^2 x^\mu}{\partial \widetilde{x}^\alpha \partial \widetilde{x}^\beta} \frac{\partial \widetilde{x}^\gamma}{\partial x^\mu} \ .$$

Because of the second inhomogeneous term, they do not transform as a tensor. However, the difference of two connections $\Gamma^\lambda_{\mu\nu} - \overline{\Gamma}^\lambda_{\mu\nu}$ is a tensor of type $(1, 2)$.

5.2 Torsion and curvature

From the connection ∇ we can construct two intrinsic geometric objects: the **torsion tensor** $T : \mathcal{X}(M)^2 \to \mathcal{X}(M)$ and the **Riemann tensor** $R : \mathcal{X}(M)^3 \to \mathcal{X}(M)$, as follows.

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$$
$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \ .$$

They are both antisymmetric in $X \leftrightarrow Y$. Although not manifest, with a bit of algebra one can show that these objects are *tensors*—as opposed to differential operators—namely they only depend on X, Y, Z at the point p and not on their derivatives.

¶ *Exercise* 15. Verify that $T(fX, gY) = fg\, T(X,Y)$ and $R(fX, gY)hZ = fgh\, R(X,Y)Z$.

Acting on a coordinate basis, we define the tensors in components:

$$T(e_\mu, e_\nu) = T^\lambda_{\mu\nu}\, e_\lambda \ , \qquad\qquad R(e_\mu, e_\nu)e_\lambda = R^\sigma{}_{\lambda\mu\nu}\, e_\sigma$$

(notice the order of indices in R). They can be explicitly expressed in terms of the connection coefficients:

$$T^\lambda_{\mu\nu} = 2\Gamma^\lambda_{[\mu\nu]}$$
$$R^\sigma{}_{\lambda\mu\nu} = 2\partial_{[\mu}\Gamma^\sigma_{\nu]\lambda} + 2\Gamma^\sigma_{[\mu|\rho}\Gamma^\rho_{\nu]\lambda} \ ,$$

where $[\mu, \nu]$ means antisymmetrization of indices.

The curvature tensor $R(X,Y)Z$ measures the difference between parallel transport of Z along infinitesimal X_p and then Y_p, as opposed to Y_p and then X_p. The torsion tensor measures the difference between the point indicated by X parallel transported along infinitesimal Y_p, as opposed to Y parallel transported along infinitesimal X_p.

5.3 The Levi-Civita connection

So far we have left the connection Γ arbitrary. When the manifold M is endowed with a metric, we can demand that $g_{\mu\nu}$ be *covariantly constant*. This means that if two vectors X and Y are parallel transported along a curve, then their inner product remains constant.

Definition 5.3. An affine connection ∇ is said to be **metric compatible**, or simply a **metric connection**, if

$$\nabla_V g = 0$$

for any vector field V and at every point on M. In components:

$$(\nabla_\lambda g)_{\mu\nu} = 0 \ .$$

Writing this condition in terms of the connection coefficients, one finds an expression for the symmetric part $\Gamma^{\lambda}_{(\mu\nu)}$ in terms of the metric and the torsion tensor (¶). Recalling that the antisymmetric part is itself the torsion tensor, one finds

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}\Big(2\partial_{(\mu} g_{\nu)\rho} - \partial_{\rho} g_{\mu\nu} + 2T^{\sigma}_{\rho(\mu} g_{\nu)\sigma}\Big) + \frac{1}{2} T^{\lambda}_{\mu\nu}\ .$$

Definition 5.4. A connection ∇ is said **symmetric** if the torsion tensor vanishes, namely if $\Gamma^{\lambda}_{[\mu\nu]} = 0$.

¶ *Exercise* 16. Let ∇ be a symmetry connection.

(a) Let $f \in \mathscr{F}(M)$. Show that $\quad \nabla_{\mu}\nabla_{\nu} f = \nabla_{\nu}\nabla_{\mu} f$.

(b) Let $\omega \in \Omega^1(M)$. Show that $\quad \mathrm{d}\omega = (\nabla_{\mu}\omega)_{\nu}\, \mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu}$.

Proposition 5.5. *On a (pseudo-) Riemannian manifold* (M, g) *there exists a unique* symmetric *connection* compatible *with the metric* g*. This connection is called the* **Levi-Civita connection**.

Proof. The constructive proof is the explicit expression for the Levi-Civita connection:

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}\left(\partial_{\mu} g_{\nu\rho} + \partial_{\nu} g_{\mu\rho} - \partial_{\rho} g_{\mu\nu}\right) .$$

□

¶ *Exercise* 17. Geodesics and minimal length

¶ *Exercise* 18. Geodesics on the upper half plane with Poincaré metric

Definition 5.6. The **Ricci tensor** is a $(0, 2)$ tensor defined by

$$\mathrm{Ric}(X, Y) = \langle \mathrm{d}x^{\mu},\, R(e_{\mu}, Y)X\rangle\ , \qquad\qquad \mathrm{Ric}_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}\ .$$

This can be defined for any connection. In the presence of a metric, one can also define

$$\mathscr{R} = g^{\mu\nu}\mathrm{Ric}_{\mu\nu}$$

called the **scalar curvature**

The Riemann tensor defined with respect to the Levi-Civita connection satisfies some special properties.

(i) Symmetry. The object $R_{\lambda\rho\mu\nu} = g_{\lambda\sigma} R^{\sigma}{}_{\rho\mu\nu}$ satisfies[21]

$$R_{(\lambda\rho)\mu\nu} = 0 \qquad\qquad R_{\lambda\rho\mu\nu} = R_{\mu\nu\lambda\rho}$$
$$R_{\lambda\rho(\mu\nu)} = 0 \qquad\qquad \mathrm{Ric}_{\mu\nu} = \mathrm{Ric}_{\nu\mu}\ .$$

[21] To prove these symmetry properties, one can first write the Riemann tensor as

$$R_{\lambda\rho\mu\nu} = \frac{1}{2}\left(\frac{\partial^2 g_{\lambda\mu}}{\partial x^{\rho}\,\partial x^{\nu}} - \frac{\partial^2 g_{\rho\mu}}{\partial x^{\lambda}\,\partial x^{\nu}} - \frac{\partial^2 g_{\lambda\nu}}{\partial x^{\rho}\,\partial x^{\mu}} + \frac{\partial^2 g_{\rho\nu}}{\partial x^{\lambda}\,\partial x^{\mu}}\right) + g_{\zeta\eta}\Big(\Gamma^{\zeta}_{\lambda\mu}\Gamma^{\eta}_{\rho\nu} - \Gamma^{\zeta}_{\lambda\nu}\Gamma^{\eta}_{\rho\mu}\Big)\ .$$

(ii) Bianchi identities:

$$\begin{aligned} 0 &= R(X,Y)Z + R(Z,X)Y + R(Y,Z)X \\ 0 &= (\nabla_X R)(Y,Z)V + (\nabla_Z R)(X,Y)V + (\nabla_Y R)(Z,X)V \ . \end{aligned}$$

In components they take the simple form

$$R^\lambda{}_{[\rho\mu\nu]} = 0 \ , \qquad\qquad \nabla_{[\sigma} R^\lambda{}_{\rho|\mu\nu]} = 0$$

in terms of antisymmetrization.

(iii) Conservation of the Einstein tensor:[22]

$$\nabla_\mu \left(\mathrm{Ric}^{\mu\nu} - \frac{1}{2} g^{\mu\nu} \mathscr{R} \right) \equiv \nabla_\mu G^{\mu\nu} = 0 \ .$$

Here $G_{\mu\nu}$ is called the *Einstein tensor*. This relation shows that it is automatically conserved.

Taking into account the symmetry properties and the first Bianchi identity, one finds[23] that in m dimensions the number of algebraically independent components of the Riemann tensor is $F(m) = m^2(m^2-1)/12$. We find $F(1) = 0$, indeed every one-dimensional manifold is flat. Then $F(2) = 1$, indeed in two dimensions the curvature is fixed by the scalar curvature. Finally $F(3) = 6$, indeed in three dimensions the curvature is fixed by the Ricci tensor.

5.4 Isometries and conformal transformations

Definition 5.7. Let (M, g) be a (pseudo-) Riemannian manifold. We say that a diffeomorphism $f : M \to M$ is an **isometry** if it preserves the metric:

$$f^* g_{f(p)} = g_p \ .$$

In components:

$$\frac{\partial f^\alpha}{\partial x^\mu} \frac{\partial f^\beta}{\partial x^\nu} \, g_{\alpha\beta}\big(f(x)\big) = g_{\mu\nu}(x) \ .$$

The identity map, the inverse of an isometry and the composition of isometries are isometries, therefore they form a *group*.

[22]The relation simply follows from a double contraction of the second Bianchi identity with the compatible metric.

[23]Each antisymmetric pair $[\mu\nu]$ has $N = \binom{m}{2}$ components. The symmetrization of two pairs gives $\binom{N+1}{2}$ components. The first Bianchi identity, $R_{\lambda\rho\mu\nu} + R_{\lambda\mu\nu\rho} + R_{\lambda\nu\rho\mu} = 0$ is a totally antisymmetric tensor in four indices, and imposes $\binom{m}{4}$ constraints. The formula for $F(m)$ follows.

In 2 dimensions: $R_{\lambda\rho\mu\nu} = \mathscr{R}\, g_{\lambda[\mu} g_{\nu]\rho}$. In 3 dimensions: $R_{\lambda\rho\mu\nu} = 2\big(g_{\lambda[\mu} \mathrm{Ric}_{\nu]\rho} - g_{\rho[\mu} \mathrm{Ric}_{\nu]\lambda}\big) - \mathscr{R}\, g_{\lambda[\mu} g_{\nu]\rho}$.

Definition 5.8. Let (M, g) be a Riemannian manifold and $X \in \mathcal{X}(M)$ a vector field. If X generates a one-parameter family of isometries, it is called a **Killing vector field**. The infinitesimal isometries are $f : x^\mu \to x^\mu + \varepsilon X^\mu$ with ε infinitesimal. This leads to

$$X^\lambda \partial_\lambda g_{\mu\nu} + 2\partial_{(\mu} X^\lambda g_{\nu)\lambda} = (\mathcal{L}_X g)_{\mu\nu} = 0 ,$$

called **Killing equation**.

¶ *Exercise* 19. Show that the Killing equation can be rewritten in terms of the Levi-Civita connection ∇ as[24]

$$(\nabla_\mu X)_\nu + (\nabla_\nu X)_\mu = 0 .$$

Let X, Y be two Killing vector fields. We easily verify:

(a) any linear combination $aX + bY$ $(a, b \in \mathbb{R})$ is a Killing vector field;

(b) the Lie bracket $[X, Y]$ is a Killing vector field.

Thus Killing vector fields form a Lie algebra of symmetries (isometries) of the manifold M.

Definition 5.9. Let (M, g) be a (pseudo-) Riemannian manifold. We say that a diffeomorphism $f : M \to M$ is a **conformal transformation** if it preserves the metric *up to a scale*:

$$f^* g_{f(p)} = e^{2\sigma(p)}\, g_p , \qquad\qquad \sigma \in \mathscr{F}(M) .$$

In components:

$$\frac{\partial f^\alpha}{\partial x^\mu} \frac{\partial f^\beta}{\partial x^\nu}\, g_{\alpha\beta}\big(f(x)\big) = e^{2\sigma(x)}\, g_{\mu\nu}(x) .$$

Let us define the angle θ between two vectors $X, Y \in T_p(M)$ as

$$\cos\theta = g_p(X, Y) / \big[g_p(X, X)\, g_p(Y, Y)\big]^{1/2} .$$

Conformal transformations change lengths but preserve angles. The set of conformal transformations on M is a *group*, the **conformal group** $\mathrm{Conf}(M)$.

Definition 5.10. Let (M, g) be a Riemannian manifold and $X \in \mathcal{X}(M)$ a vector field. If X generates a one-parameter family of conformal transformations, it is called a **conformal Killing vector field** (CKV). The infinitesimal isometries are $f : x^\mu \to x^\mu + \varepsilon X^\mu$ with ε infinitesimal. The scale factor must be proportional to ε, thus set $\sigma = \varepsilon\psi/2$. This leads to

$$(\mathcal{L}_X g)_{\mu\nu} = \psi\, g_{\mu\nu} , \qquad\qquad \psi \in \mathscr{F}(M) .$$

It is easy to find $\psi = m^{-1}\big(X^\lambda g^{\mu\nu}\partial_\lambda g_{\mu\nu} + 2\partial_\lambda X^\lambda\big)$ where $m = \dim M$.

CKVs form a Lie algebra of conformal transformations, closed under linear combinations and Lie bracket (¶).

[24]Notice that the Killing condition is weaker than the covariant constancy condition. A covariantly constant vector field, $\nabla_\mu X_\nu = 0$, is both Killing and with constant norm, $\partial_\mu(V_\nu V^\nu) = 0$.

A concept related to conformal transformations is Weyl rescalings. Let $g, \overline{g}$ be metrics on a manifold M. $\overline{g}$ is said to be **conformally equivalent** to g if

$$\overline{g}_p = e^{2\sigma(p)}\, g_p \ .$$

This is an equivalence relation between metrics on M, and each equivalence class is called a **conformal structure**. The transformations $g \to e^{2\sigma} g$, called **Weyl rescalings**, form an infinite-dimensional group denoted by $\mathrm{Weyl}(M)$.

The Riemann tensor of $\overline{g}$ is different from the Riemann tensor of g. However, it turns out that the *traceless* part of the Riemann tensor is invariant under Weyl rescalings:

$$W_{\lambda\rho\mu\nu} = R_{\lambda\rho\mu\nu} - \frac{2}{m-2}\Big(g_{\lambda[\mu}\mathrm{Ric}_{\nu]\rho} - g_{\rho[\mu}\mathrm{Ric}_{\nu]\lambda}\Big) + \frac{2}{(m-1)(m-2)}\, \mathscr{R}\, g_{\lambda[\mu}g_{\nu]\rho} \ .$$

This is called the **Weyl tensor**. It vanishes identically for $m = 3$.[25]

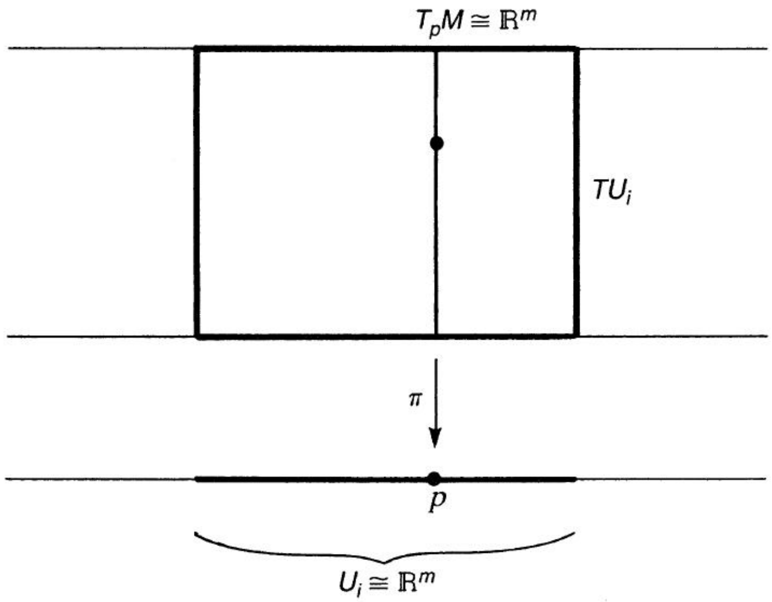
A (pseudo-) Riemannian manifold (M, g) such that in each patch $g_{\mu\nu} = e^{2\sigma}\eta_{\mu\nu}$ (with the suitable signature) is said to be **conformally flat**. Since the Weyl tensor vanishes for a flat metric, it also vanishes for a conformally flat metric. If $m \geq 4$, also the converse is true:

Theorem 5.11. *If* $\dim M \geq 4$*, a (pseudo-) Riemannian manifold is conformally flat if and only if its Weyl tensor vanishes.*

If $\dim M = 3$*, a (pseudo-) Riemannian manifold is conformally flat if and only if its Cotton tensor vanishes.*

If $\dim M = 2$*, every (pseudo-) Riemannian manifold is conformally flat.*

[25] The expression given is not valid for $m = 2$. However in two dimensions the Riemann tensor is pure trace, therefore its traceless part vanishes.



6 Fibre bundles

A manifold is a topological space that locally looks like $\mathbb{R}^n$. A *fibre bundle* is a topological space which locally looks like a direct product of $\mathbb{R}^n$ and another space.

Definition 6.1. A (differentiable) **fibre bundle** (E, π, M, F, G) consists of the following elements:

(i) A manifold E called the **total space**.

(ii) A manifold M called the **base space**.

(iii) A manifold F called the **fibre**.

(iv) A surjective map

$$\pi : E \to M$$

called the **projection**. The inverse image of any point, $\pi^{-1}(p) = F_p$, must be isomorphic to F and is the fibre at p. (See Figure.)

(v) A Lie group G called the **structure group**, acting on F from the left.

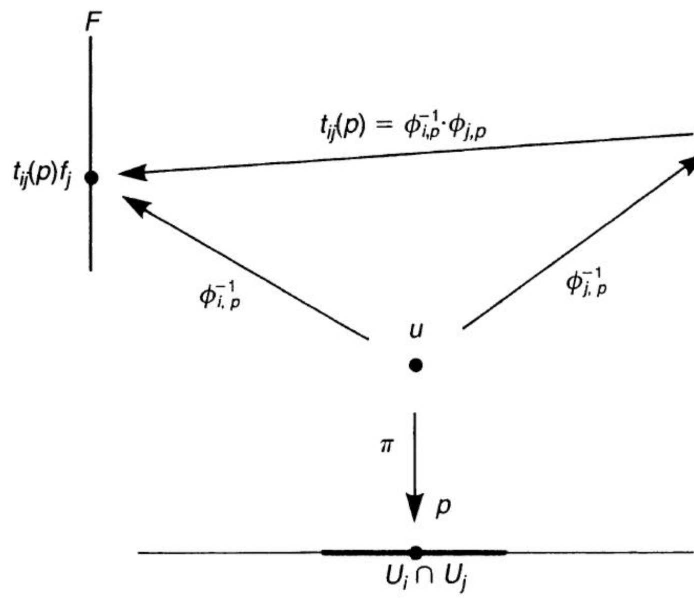
(vi) A set of open coverings $\{U_i\}$ of M with diffeomorphisms $\phi_i : U_i \times F \to \pi^{-1}(U_i)$ such that $\pi \circ \phi_i(p, f) = p$. The maps ϕ_i are called **local trivializations**. (See Figure.)

(vii) Writing $\phi_i(p, f) \equiv \phi_{i,p}(f)$, the maps $\phi_{i,p} : F \to F_p$ are diffeomorphisms. On $U_i \cap U_j \neq \varnothing$ we require that

$$t_{ij}(p) \equiv \phi_{i,p}^{-1} \circ \phi_{j,p} : F \to F$$

be an element of G. Then ϕ_i and ϕ_j are related by a smooth map $t_{ij} : U_i \cap U_j \to G$ as

$$\phi_j(p, f) = \phi_i\big(p, t_{ij}(p) f\big) \ .$$



The maps t_{ij} are called **transition functions**.

The effect of the transition functions can be written in an alternative way. Consider the diffeomorphisms $\phi_i^{-1} : \pi^{-1}(U_i) \to U_i \times F$. Take a point u such that $\pi(u) = p \in U_i \cap U_j$. Then

$$\phi_i^{-1}(u)\big|_F = t_{ij}(p) \ \ \phi_j^{-1}(u)\big|_F \ .$$

On a well-defined bundle, the transition functions satisfy the following consistency conditions:

$$\begin{aligned} t_{ii}(p) &= \text{id} && (p \in U_i) \\ t_{ij}(p) &= t_{ji}(p)^{-1} && (p \in U_i \cap U_j) \\ t_{ij}(p)\cdot t_{jk}(p) &= t_{ik}(p) && (p \in U_i \cap U_j \cap U_k)\ . \end{aligned}$$

If all transition functions can be taken to be the identity map, then the bundle is called a **trivial bundle** and it is a direct product $M \times F$.

For a given bundle $E \stackrel{\pi}{\to} M$, the set of possible transition functions is not unique. We can change the local trivializations $\{\phi_i\}$, without changing the bundle, by choosing maps $g_i(p) : F \to F$ at each point $p \in M$, required to be diffeomorphisms that belong to G, and then defining

$$\widetilde{\phi}_{i,p} = \phi_{i,p} \circ g_i(p) \ .$$

The transition functions for the new trivializations are

$$\widetilde{t}_{ij}(p) = g_i(p)^{-1} \circ t_{ij}(p) \circ g_j(p) \ .$$

Physically, the choice of trivializations $\{\phi_i\}$ is a choice of gauge and the maps $\{g_i\}$ are *gauge transformations*, one on each covering patch U_i.

Definition 6.2. Let $E \xrightarrow{\pi} M$ be a fibre bundle. A **section** $s : M \to E$ is a smooth map which satisfies $\pi \circ s = \mathrm{id}_M$. Clearly $s(p) = s|_p$ is an element of $F_p = \pi^{-1}(p)$. The set of sections on M is denoted by $\Gamma(M, E)$.

If $U \subset M$, we may talk of a **local section** which is defined only on U, and $\Gamma(U, F)$ denotes the set of local sections on U. Notice that not all fibre bundles admit global sections.

It turns out that a fibre bundle (E, π, M, F, G) can be *reconstructed* from the data

$$M,\ \{U_i\},\ t_{ij}(p),\ F,\ G\ .$$

This amounts to finding E and π from the given data. Construct

$$X = \bigsqcup_i U_i \times F\ .$$

Introduce an equivalence relation $\sim$ between $(p, f) \in U_i \times F$ and $(q, f') \in U_j \times F$ if $p = q$ and $f' = t_{ij}(p)f$ (this is possible if the transition functions satisfy the consistency relations). A fibre bundle E is then defined as

$$E = X/\sim\ .$$

Denote an element of E as $[(p, f)]$. The projection is given by

$$\pi : [(p, f)] \mapsto p\ .$$

The local trivializations $\phi_i : U_i \times F \to \pi^{-1}(U_i)$ are given by

$$\phi_i : (p, f) \mapsto [(p, f)]\ .$$

Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be fibre bundles. A *smooth* map $\bar{f} : E \to E'$ si called a **bundle map** if it maps each fibre F_p of E to a fibre F'_q of E'. Then $\bar{f}$ naturally induces a smooth map $f : M \to M'$ such that $f(p) = q$. Then the diagram

$$\begin{array}{ccc} E & \xrightarrow{\bar{f}} & E' \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi'} \\ M & \xrightarrow{f} & M' \end{array}$$

commutes.

Let $E \xrightarrow{\pi} M$ be a fibre bundle with typical fibre F. Given a map $f : N \to M$, we can define a new fibre bundle f^*E over N with the same fibre F, called the **pulled-back bundle**. Consider the diagram

$$\begin{array}{ccc} f^*E & \xrightarrow{\pi_2} & E \\ \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \pi} \\ N & \xrightarrow{f} & M \end{array} \qquad \left(\begin{array}{ccc} (p, u) & \xrightarrow{\pi_2} & u \\ \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \pi} \\ p & \xrightarrow{f} & f(p) \end{array} \right) .$$

We define f^*E as the following subspace of $N \times E$:

$$f^*E = \big\{(p,u) \in N \times E \,\big|\, f(p) = \pi(u)\big\} \ .$$

We define the projections $\pi_1 : (p,u) \mapsto p$ and $\pi_2 : (p,u) \mapsto u$. This makes the diagram above commuting. The fibre $\widetilde{F}_p$ of f^*E at p is equal to $F_{f(p)}$. Then the transition functions of f^*E are the pull-back of those of E:

$$\widetilde{t}_{ij}(p) = t_{ij}(f(p)) = f^*t_{ij}(p)$$

(¶ check as an exercise).

6.1 Vector bundles

A **vector bundle** $E \xrightarrow{\pi} M$ is a fibre bundle whose fibre is a vector space. Let F be $\mathbb{R}^k$ and M an m-dimensional manifold. We call k the **fibre dimension**. The transition functions belong to $GL(k,\mathbb{R})$, since those map the vector space into itself isomorphically. If F is a complex vector space $\mathbb{C}^k$, the structure group is $GL(k,\mathbb{C})$.

Example 6.3. The **tangent bundle** TM over an m-dimensional manifold M is a vector bundle whose fibre is $\mathbb{R}^m$. Let u be a point in TM such that $\pi(u) = p \in U_i \cap U_j$. Let $x^\mu = \varphi_i(p)$ and $\widetilde{x}^\nu = \varphi_j(p)$ be coordinate systems on U_i and U_j, respectively. The vector V corresponding to u is expressed as

$$V = V^\mu \partial_\mu\big|_p = \widetilde{V}^\nu \widetilde{\partial}_\nu\big|_p \ .$$

The local trivializations are

$$\phi_i^{-1}(u) = \big(p, \{V^\mu\}\big) \ , \qquad \phi_j^{-1}(u) = \big(p, \{\widetilde{V}^\nu\}\big) \ .$$

The fibre coordinates $\{V^\mu\}$ and $\{\widetilde{V}^\nu\}$ are related as

$$V^\mu = \left(\frac{\partial x^\mu}{\partial \widetilde{x}^\nu}\right)_p \widetilde{V}^\nu \ ,$$

therefore the transition function ${G^\mu}_\nu(p) = (\partial x^\mu/\partial \widetilde{x}^\nu)_p \in GL(m,\mathbb{R})$. Hence, a tangent bundle is $\big(TM, \pi, M, \mathbb{R}^m, GL(m,\mathbb{R})\big)$.

The sections of TM are vector fields on M, namely $\Gamma(M,TM) = \mathcal{X}(M)$.

Example 6.4. The **cotangent bundle** $T^*M = \bigcup_{p\in M} T_p^*M$ is defined similarly to the tangent bundle. On a chart U_i whose coordinates are x^μ, a basis of T_p^*M can be taken to be $\{\mathrm{d}x^\mu\}$, which is dual to $\{\partial/\partial x^\mu\}$. If $p \in U_i \cap U_j$ a one-form ω is represented as

$$\omega = \omega_\mu \, \mathrm{d}x^\mu = \widetilde{\omega}_\nu \, \mathrm{d}\widetilde{x}^\nu \ .$$

Therefore the fibre coordinates $\{\omega_\mu\}$ and $\{\widetilde{\omega}_\nu\}$ are related as

$$\omega_\mu = \Big(\frac{\partial \widetilde{x}^\nu}{\partial x^\mu}\Big)\,\widetilde{\omega}_\nu = \Big(\frac{\partial x}{\partial \widetilde{x}}\Big)^{-1}_{\mu\nu}\,\widetilde{\omega}_\nu$$

and the transition functions are $G(p)^{-1}$ in terms of the transition functions $G(p)$ of the tangent bundle, and the structure group is still $GL(m,\mathbb{R})$. Notice that in the contraction

$$\langle \omega, V\rangle = \omega_\mu V^\mu$$

the transition functions cancel, namely the contraction takes values in the trivial bundle with fibre $\mathbb{R}$.

The sections of T^*M are the one-forms on M, namely $\Omega^1(M) = \Gamma(M, T^*M)$.

The construction is more general. Given a vector bundle $E \xrightarrow{\pi} M$ with fibre F, we may define its **dual bundle** $E^* \xrightarrow{\pi} M$. The fibre F^* of E^* is the vector space of linear maps from F to $\mathbb{R}$ (or $\mathbb{C}$). Given a basis $\{e_a(p)\}$ of F_p, we define the dual basis $\{\theta^a(p)\}$ of F^*_p such that $\langle \theta^a(p), e_a(p)\rangle = \delta^a_b$. The transition functions of E^* are the inverse of those of E.

The set of **sections** of a vector bundle form an infinite-dimensional vector space. Addition and scalar multiplication are defined pointwisely:

$$(s+s')(p) = s(p) + s'(p) \qquad\qquad (fs)(p) = f(p)\,s(p)\ .$$

We see that the field of coefficients can be taken as the field $\mathscr{F}(M)$ of functions on M, more general than just $\mathbb{R}$. Vector bundles always admit a special section called the **null section** $s_0 \in \Gamma(M,E)$, such that $\phi_i^{-1}\big(s_0(p)\big) = (p,0)$ in any local trivialization. This is also the origin of the vector space of sections.

Given a metric $h_{\mu\nu}$ on the fibre, this defines an inner product

$$(s,s')_p = h_{\mu\nu}(p)\,s^\mu(p)\,s'^\nu(p)\ .$$

This is a function $f \in \mathscr{F}(M)$, not a number, as that is the field of scalars in the vector space.

6.2 Principal bundles

A **principle bundle**, or ***G*-bundle**, $P \xrightarrow{\pi} M$ or $P(M,G)$ is a fibre bundle whose fibre F is equal to the structure (Lie) group G.

The transition functions act on the fibres from the left as before. In addition, we can define an action of G on F from the right. Let $\phi_i : U_i \times G \to \pi^{-1}(U_i)$ be a local trivialization given by $\phi_i^{-1}(u) = (p, g_i)$ where $p = \pi(u)$. Then

$$\phi_i^{-1}(ua) = (p, g_i a) \qquad \forall\, a \in G\ .$$

Since transition functions act from the left, and the left and right actions commute, this definition is independent of any local trivialization: if $p \in U_i \cap U_j$,

$$ua = \phi_j(p, g_j a) = \phi_i\big(p,\, t_{ij}(p)\, g_j a\big) = \phi_i(p,\, g_i a) \;.$$

Thus the right multiplication is denoted as $P \times G \to P$, or $(u, a) \mapsto ua$. Notice that $\pi(ua) = \pi(u)$.

Moreover, the right multiplication on $\pi^{-1}(p) = F_p \cong G$ is:

(a) *Transitive*, *i.e.* for any $u_1, u_2 \in \pi^{-1}(p)$ there exists an element $a \in G$ such that $u_1 = u_2 a$. This means that we can reconstruct the whole fibre as $\pi^{-1}(p) = \{ua \,|\, a \in G\}$ starting from u.

(b) *Free*, *i.e.* if $ua = u$ for some $u \in P$ then $a = \mathbb{1}$.

Given a local section $s_i(p)$ over U_i, it selects a preferred local trivialization ϕ_i as follows. For $u \in \pi^{-1}(p)$ and $p \in U_i$, there is a *unique* element $g_u \in G$ such that $u = s_i(p) g_u$. Then we define

$$\phi_i^{-1}(u) = (p, g_u) \;.$$

In other words, this is the local trivialization in which the local section appears as

$$s_i(p) = \phi_i(p, \mathbb{1}) \;.$$

The transition functions relate the various local sections:

$$s_j(p) = s_i(p)\, t_{ij}(p) \;.$$

This follows from $s_j(p) = \phi_j(p, \mathbb{1}) = \phi_i\big(p, t_{ij}(p)\mathbb{1}\big) = \phi_i\big(p, \mathbb{1}\, t_{ij}(p)\big) = s_i(p)\, t_{ij}(p)$.

Proposition 6.5. *A principal bundle is* trivial*, or* parallelizable*, if and only if it admits a globally defined section $s(p)$.*

Proof. If there is a globally defined section, we can use it to define local trivializations on all patches U_i. The transition functions between those trivializations are trivial, $t_{ij}(p) = \mathbb{1}$, proving parallelizability. On the contrary, a trivial bundle is $P = M \times G$ and any constant section is globally defined. □

Proposition 6.6. *Let G be a Lie group and H a closed Lie subgroup. Then G is a principal bundle over $M = G/H$ with fibre H.*

The projection $\pi : G \to M = G/H$ is given by $\pi : g \mapsto [g] = \{gh \,|\, h \in H\}$. By definition, $\pi^{-1}(p) = F_p \cong H$. To construct local trivializations, we cover M with patches U_i and then construct local sections $s_i(p)$.

Example 6.7. **Hopf map**. A nice example is $SU(2)$, which is a $U(1)$ bundle over $SU(2)/U(1)$.

First, $SU(2) \cong S^3$. Let $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $a, b, c, d \in \mathbb{C}$ be a complex matrix. Then M is in $SU(2)$ if $M^{-1} = \frac{1}{\det M}\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$ is equal to $M^\dagger$ and $\det M = 1$. We find that the matrices of $SU(2)$ are $M = \left(\begin{smallmatrix} a & b \\ -b^* & a^* \end{smallmatrix}\right)$ with $|a|^2 + |b|^2 = 1$. This is S^3.

Second, $SU(2)/U(1) \cong S^2$. The subgroup $U(1)$ is given by matrices $U = \left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{smallmatrix}\right)$ which rotate the phase of a and b, and we should identify points related by a phase rotation. For $a \neq 0$ we can use the phase rotation to set $a \in \mathbb{R}_+$, and we are left with $a^2 + b_1^2 + b_2^2 = 1$ (where $b = b_1 + ib_2$) which is a semi-sphere B_2. However for $a = 0$ the phase rotation identifies all points at the boundary of the semi-sphere, and we get S^2.

We obtain that S^3 is a $U(1)$ bundle over S^2. This is called the Hopf fibration.

6.3 Associated bundles

Given a principal bundle $P(M, G)$ and a manifold F on which G acts from the left, we can construct an **associated bundle**. We define an action of G on $P \times F$ as[26]

$$g : (u, f) \mapsto (ug, g^{-1} f) .$$

The associated bundle (E, π, M, F, G) is

$$E = \big(P \times F\big)/G \qquad \text{where} \qquad (u, f) \sim (ug, g^{-1} f) .$$

In other words, the associated bundle has the same base space M, it has fibre F and its transition functions are the same as those of the principal bundle. The projection is $\pi_E(u, f) = \pi(u)$. The local trivializations are

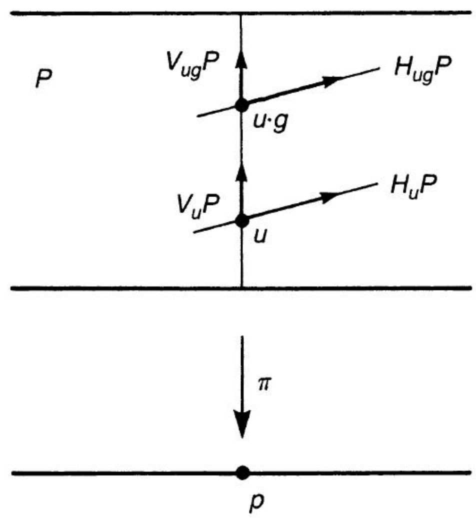
$$\psi_i : U_i \times F \to \pi_E^{-1}(U_i) \qquad \text{such that} \qquad (p, f) \mapsto \big(\phi_i(p, \mathbb{1}),\, f\big) .$$

For example, we can take F to be a k-dimensional vector space V which provides a k-dimensional representation ρ of G. Thus from a principal G-bundle we can construct an associated $\mathbb{R}^k$ vector bundle. The transition functions are $\rho\big(t_{ij}(p)\big)$.

7 Connections on fibre bundles

Manifold and bundles only have a topological and differentiable structure. We would like to introduce a new structure: the "parallel transport" of elements of the fibre F as we move along the base M. Such a structure is called a *connection*. We start defining a connection on *principal bundles*. We use a geometric and coordinate-invariant definition.

[26]This definition guarantees that G acts with a group action, because $g_2^{-1} g_1^{-1} = (g_1 g_2)^{-1}$.



Definition 7.1. Let u be an element of a principal bundle $P(M, G)$ and let G_p be the fibre at $p = \pi(u)$. The **vertical subspace** V_uP is the subspace of T_uP that is tangent to G_p at u. [*Warning*: T_uP is the tangent space of P, not just of M.] That is

$$V_uP = \ker \pi_* \ .$$

To construct V_uP, take $A \in \mathfrak{g}$. The right-action

$$R_{\exp(tA)}\, u = u \exp(tA)$$

gives a curve through u in P and within G_p. Define a vector $A^\sharp \in T_uP$ by

$$A^\sharp[f] = \frac{\mathrm{d}}{\mathrm{d}t} f\big(u \exp(tA)\big)\Big|_{t=0}$$

where $f \in \mathscr{F}(P)$. As we vary $u \in P$, this gives a vector field $A^\sharp \in \mathcal{X}(P)$.

The map $\sharp : \mathfrak{g} \to V_uP$ given by $A \mapsto A^\sharp$ is an isomorphism of vector spaces.

¶ *Exercise* 20. We can explicitly check that $\pi_* A^\sharp = 0$:

$$(\pi_* A^\sharp)[f] = A^\sharp[f \circ \pi] = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \pi\big(u \exp(tA)\big)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \pi(u)\Big|_{t=0} = 0 \ ,$$

where $f \in \mathscr{F}(M)$.

Definition 7.2. Let $P(M, G)$ be a principal bundle. A **connection** on P is a separation of the tangent space T_uP into the vertical subspace V_uP and a **horizontal subspace** H_uP such that

(i) $T_uP = H_uP \oplus V_uP$

(ii) The separation is smooth, meaning that any (smooth) vector field X on P is separated into (smooth) vector fields $X^H \in H_uP$ and $X^V \in V_uP$ as $X = X^H + X^V$

(iii) $H_{ug}P = R_{g*}H_uP$,
i.e. the separation at a point on the fibre fixes the separation on the whole fibre. (See Figure)

The separation between vertical and horizontal subspaces is specified by the the connection one-form.

Definition 7.3. A **connection one-form** $\omega \in T^*P \otimes \mathfrak{g}$ is a projection of T_uP onto the vertical subspace $V_uP \cong \mathfrak{g}$:[27]

$$\begin{aligned} &\text{(i)} \qquad \omega(A^\sharp) = A \qquad \forall\, A \in \mathfrak{g} \\ &\text{(ii)} \qquad R_g^*\omega = g^{-1}\omega g \ . \end{aligned}$$

Then ω defines the horizontal subspace H_uP as its kernel:

$$H_uP \;\equiv\; \left\{ X \in T_uP \,\middle|\, \omega(X) = 0 \right\} .$$

Remark. We can check that this definition is consistent with the definition of H_uP. We have

$$R_{g*}H_uP = R_{g*}\left\{ X \in T_uP \,\middle|\, \omega(X) = 0 \right\} .$$

Using the definition of ω we rewrite

$$0 = \omega(X) = g\, R_g^*\omega(X)\, g^{-1} = g\,\omega(R_{g*}X)\, g^{-1} \ .$$

Since the adjoint action is invertible, we conclude

$$R_{g*}H_uP = \left\{ R_{g*}X \in T_{ug}P \,\middle|\, \omega(R_{g*}X) = 0 \right\} = H_{ug}P \ .$$

Let $\{U_i\}$ be an open covering of M and let $\sigma_i : M \to P$ be local sections defined on each U_i. Then we can represent ω by Lie-algebra-valued one-forms $\mathcal{A}_i$ on U_i:

$$\mathcal{A}_i \equiv \sigma_i^*\omega \quad \in \ \Omega^1(U_i) \otimes \mathfrak{g} \ .$$

[It turns out that one can reconstruct ω from $\mathcal{A}_i$.][28] In components we write

$$\mathcal{A}_i = (\mathcal{A}_i)^a_\mu \,\mathrm{d}x^\mu \, T_a \ ,$$

where $\{T_a\}$ is a basis of $\mathfrak{g}$.

[27] Equation (ii) can be written as $(R_g^*\omega)_u(X) \equiv \omega_{ug}(R_{g*}X) = g^{-1}\omega_u(X)g$. On the RHS is the adjoint action of G on $\mathfrak{g}$. To define it, construct the map $\mathrm{ad}_g : G \to G$ that maps $x \mapsto gxg^{-1}$ for any $g \in G$. In particular it maps $\mathbb{1} \mapsto \mathbb{1}$. Its differential map at $\mathbb{1}$ is $\mathrm{ad}_{g*} \equiv \mathrm{Ad}_g : T_{\mathbb{1}}G \to T_{\mathbb{1}}G$. Identifying $T_{\mathbb{1}}G \cong \mathfrak{g}$, we obtain the adjoint action on the algebra.

[28] One reconstructs $\mathfrak{g}$-valued one-forms ω_i on $\pi^{-1}(U_i) \subset P$ by

$$\omega_i = g_i^{-1}\pi^*\mathcal{A}_i g_i + g_i^{-1}\mathrm{d}_P g_i \ .$$

If $\mathcal{A}_i$ transform as a gauge potential from patch to patch, then the ω_i's agree on intersections and give ω.

Lemma 7.4. *Let σ_i and σ_j be local sections of a principal bundle $P(M,G)$ on $U_i \cap U_j$, related by $\sigma_j(p) = \sigma_i(p)\, t_{ij}(p)$. Then for $X \in T_pM$ (and $p \in U_i \cap U_j$):*

$$\sigma_{j*}X = R_{t_{ij}*}(\sigma_{i*}X) + \big(t_{ij}^{-1}\mathrm{d}t_{ij}(X)\big)^\sharp$$

which is a vector in $T_{\sigma_j(p)}P$.

Proof. We take a curve $\gamma : [-1,1] \to M$ such that $\gamma(0) = p$ and $\dot\gamma(p) = X$. Recall that, in components, $(\sigma_* X)^\nu = \frac{\mathrm{d}}{\mathrm{d}t}\sigma^\nu\big(\gamma(t)\big)\big|_{t=0}$. Keeping the index ν implicit and using the shorthand notation $\sigma(t)$ for $\sigma(\gamma(t))$ and $t_{ij}(t)$ for $t_{ij}(\gamma(t))$, we find:

$$\begin{aligned}
\sigma_{j*}X &= \frac{\mathrm{d}}{\mathrm{d}t}\,\sigma_j(t)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\big[\sigma_i(t)\, t_{ij}(t)\big]_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\sigma_i(t)\cdot t_{ij}(p) + \sigma_i(p)\cdot \frac{\mathrm{d}}{\mathrm{d}t} t_{ij}(t)\Big|_{t=0} \\
&= R_{t_{ij}*}(\sigma_{i*}X) + \sigma_j(p)\, t_{ij}(p)^{-1}\frac{\mathrm{d}}{\mathrm{d}t}t_{ij}(t)\Big|_{t=0} \,.
\end{aligned}$$

We should interpret the last term. Notice that

$$t_{ij}(p)^{-1}\,\mathrm{d}t_{ij}(X) = t_{ij}(p)^{-1}\,\frac{\mathrm{d}}{\mathrm{d}t}\,t_{ij}(t)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\big[t_{ij}(p)^{-1}\,t_{ij}(t)\big]_{t=0} \;\in\; T_{\mathbb{1}}G \cong \mathfrak{g}\,.$$

If we compare with the definition of $A^\sharp$ at $u \in P$:

$$(A^\sharp_u)^\nu = \frac{\mathrm{d}}{\mathrm{d}t}\Big(u\,(\mathbb{1} + tA + \dots)\Big)^\nu\Big|_{t=0}\,,$$

we see that the last term is $\big(t_{ij}^{-1}(p)\,\mathrm{d}t_{ij}(X)\big)^\sharp$ at $\sigma_j(p)$. □

We use the lemma to compare pull-backs of the connection one-form:

$$\begin{aligned}
\sigma_j^*\omega(X) &= \omega(\sigma_{j*}X) = \omega(R_{t_{ij}*}\sigma_{i*}X) + \omega\big((t_{ij}^{-1}\mathrm{d}t_{ij}(X))^\sharp\big) \\
&= R^*_{t_{ij}}\omega(\sigma_{i*}X) + t_{ij}^{-1}\,\mathrm{d}t_{ij}(X) = t_{ij}^{-1}\,\sigma_i^*\omega(X)\,t_{ij} + t_{ij}^{-1}\,\mathrm{d}t_{ij}(X)\,.
\end{aligned}$$

We used the properties of the connection one-form. Since this relation is valid for any X, we conclude

$$\mathcal{A}_j = t_{ij}^{-1}\,\mathcal{A}_i\,t_{ij} + t_{ij}^{-1}\,\mathrm{d}t_{ij}\,.$$

This is precisely the way in which a *gauge potential* transforms under gauge transformations.

7.1 Curvature

Definition 7.5. The **curvature two-form** $\Omega \in \Omega^2(P)\otimes\mathfrak{g}$ is defined as

$$\Omega = \mathrm{d}_P\omega + \omega\wedge\omega\,,$$

in terms of the connection one-form $\omega \in \Omega^1(P)\otimes\mathfrak{g}$.

Here d_P is the differential on P. Let us define the last term. Let ζ, η be $\mathfrak{g}$-valued forms: $\zeta \in \Omega^p(M) \otimes \mathfrak{g}$ and $\eta \in \Omega^q \otimes \mathfrak{g}$. This means that we can decompose

$$\zeta = \zeta^a \otimes T_a \,, \qquad \eta = \eta^b \otimes T_b$$

where $\{T_a\}$ is a basis of $\mathfrak{g}$, while $\zeta^a \in \Omega^p(M)$ and $\eta^b \in \Omega^q(M)$. Then

$$\begin{aligned}[\zeta, \eta] &\equiv \zeta \wedge \eta - (-1)^{pq} \eta \wedge \zeta \\ &= \zeta^a \wedge \eta^b \, T_a T_b - (-1)^{pq} \eta^b \wedge \zeta^a \, T_b T_a \\ &= \zeta^a \wedge \eta^b \otimes [T_a, T_b] = C^c_{ab} \, \zeta^a \wedge \eta^b \otimes T_c \,.\end{aligned}$$

In the special case that $\zeta = \eta$ and $p = q$ is odd:

$$\zeta \wedge \zeta = \frac{1}{2} [\zeta, \zeta] = \frac{1}{2} \zeta^a \wedge \zeta^b \otimes [T_a, T_b] \,.$$

Lemma 7.6. *The curvature two-form Ω satisfies*

$$R_g^* \Omega = g^{-1} \Omega g \qquad \textit{for } g \in G \,.$$

Proof. It is enough to expand

$$R_g^* \Omega = R_g^* (\mathrm{d}_P \omega + \omega \wedge \omega) = \mathrm{d}_P R_g^* \omega + R_g^* \omega \wedge R_g^* \omega = \mathrm{d}_P (g^{-1} \omega g) + g^{-1} \omega g \wedge g^{-1} \omega g$$

and use that g is a constant. □

Proposition 7.7. *Defining the local form $\mathcal{F}$ of the curvature Ω as*

$$\mathcal{F} \equiv \sigma^* \Omega \,,$$

where $\sigma : U \to P$ is a local section of the principal bundle $P(M, G)$ on a chart U, it is expressed in terms of the gauge potential $\mathcal{A} = \sigma^* \omega$ as

$$\mathcal{F} = \mathrm{d}\mathcal{A} + \mathcal{A} \wedge \mathcal{A} \,.$$

Proof. We have $\mathcal{F} = \sigma^* (\mathrm{d}_P \omega + \omega \wedge \omega) = \mathrm{d} \, \sigma^* \omega + \sigma^* \omega \wedge \sigma^* \omega$. □

In components we can write

$$\mathcal{F} = \frac{1}{2} \mathcal{F}_{\mu\nu} \, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = \frac{1}{2} \mathcal{F}^a_{\mu\nu} \, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \, T_a = \mathcal{F}^a \, T_a \,,$$

where we have expanded the form and/or Algebra-valued part. Then

$$\begin{aligned}\mathcal{F}_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \\ \mathcal{F}^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + C^a_{bc} A^a_\mu A^b_\nu \,.\end{aligned}$$

The first expression is still $\mathfrak{g}$-valued.

Proposition 7.8. *Let U_i and U_j be overlapping charts of M with local forms $\mathcal{F}_i$ and $\mathcal{F}_j$ of the curvature. On $U_i \cap U_j$ they satisfy*

$$\mathcal{F}_j = t_{ij}^{-1} \mathcal{F}_i t_{ij}$$

where t_{ij} is the transition function.

Proof. The simplest proof is to start from $\mathcal{F}_j = \mathrm{d}\mathcal{A}_j + \mathcal{A}_j \wedge \mathcal{A}_j$ and substitute the relation $\mathcal{A}_j = t_{ij}^{-1} \mathcal{A}_i t_{ij} + t_{ij}^{-1} \mathrm{d}t_{ij}$. One makes use of $t_{ij}^{-1} \mathrm{d}t_{ij}\, t_{ij}^{-1} = -\mathrm{d}t_{ij}^{-1}$. The details are left as an exercise ¶. □

The curvature two-form satisfies a constraint known as the **Bianchi identity**. In a coordinate-invariant form this is

$$\mathrm{d}_P \Omega(X^H, Y^H, Z^H) = 0$$

for any $X^H, Y^H, Z^H \in H_u P$. This follows from $\mathrm{d}_P \Omega = \mathrm{d}_P \omega \wedge \omega - \omega \wedge \mathrm{d}_P \omega$ and the fact that ω is a projection to the vertical subspace.

More convenient is the local form of the Bianchi identity. We introduce a *covariant derivative* $\mathcal{D}$ on $\mathfrak{g}$-valued p-forms on M:

$$\mathcal{D}\eta \equiv d\eta + [\mathcal{A}, \eta] \ .$$

Then, expanding $\mathrm{d}\mathcal{F}$, we find

$$\mathcal{D}\mathcal{F} \equiv \mathrm{d}\mathcal{F} + [\mathcal{A}, \mathcal{F}] = 0 \ .$$

In components:

$$0 = \partial_{[\mu} \mathcal{F}_{\nu\rho]} + A_{[\mu} \mathcal{F}_{\nu\rho]} - F_{[\mu\nu} A_{\rho]} \ .$$

7.2 Parallel transport and covariant derivative

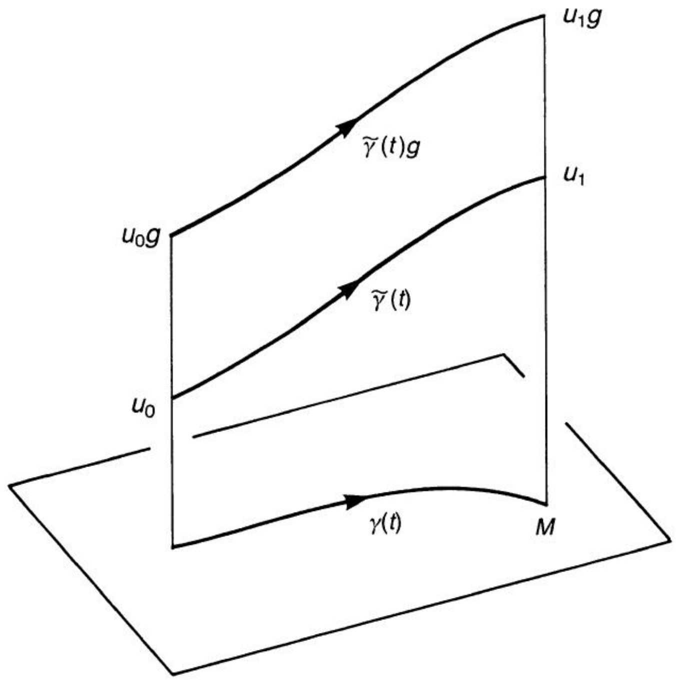
Definition 7.9. Let $P(M, G)$ be a G-bundle and let $\gamma : [0,1] \to M$ be a curve in M. A curve $\tilde{\gamma} : [0,1] \to P$ is said to be a **horizontal lift** of γ if $\pi \circ \tilde{\gamma} = \gamma$ and the tangent vector to $\tilde{\gamma}(t)$ always belongs to $H_{\tilde{\gamma}(t)} P$.

Theorem 7.10. *Let $\gamma : [0,1] \to M$ be a curve in M and let $u_0 \in \pi^{-1}\big(\gamma(0)\big)$. Then there exists a unique horizontal lift $\tilde{\gamma}(t)$ in P such that $\tilde{\gamma}(0) = u_0$.*

The proof is based on the observation that if $\widetilde{X}$ is the tangent vector to $\tilde{\gamma}$, then the curve satisfies $\omega(\widetilde{X}) = 0$. This equation reduces to an ODE, and thus the solution exists and is unique. See the Figure.

Let us write the equation. We work in a trivialization given by a local section $\sigma_i(p)$, such that $\phi_i\big(\sigma_i(p)\big) = (p, \mathbb{1})$. Then let

$$\tilde{\gamma}(t) = \sigma_i\big(\gamma(t)\big) \cdot g_i(t)$$



for some function $g_i : [0,1] \to G$ that we want to determine. With a slight modification of Lemma 7.4, we find

$$\widetilde{X} = \tilde{\gamma}_* \frac{\mathrm{d}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\gamma} = \frac{\mathrm{d}}{\mathrm{d}t}\Big[\sigma_i\big(\gamma(t)\big) \cdot g_i(t)\Big]$$
$$= R_{g_i(t)*}\,(\sigma_i \circ \gamma)_* \,\frac{\mathrm{d}}{\mathrm{d}t} + \Big(g_i^{-1}\mathrm{d}g_i\Big(\frac{\mathrm{d}}{\mathrm{d}t}\Big)\Big)^\sharp .$$

Applying ω and using its properties, we get

$$0 = \omega(\widetilde{X}) = g_i(t)^{-1}\,\omega\Big(\sigma_{i*}\gamma_*\frac{\mathrm{d}}{\mathrm{d}t}\Big)\,g_i(t) + g_i(t)^{-1}\frac{\mathrm{d}g_i(t)}{\mathrm{d}t} .$$

Multiplying by $g_i(t)$ we get

$$\frac{\mathrm{d}g_i(t)}{\mathrm{d}t} = -\omega\Big(\sigma_{i*}\gamma_*\frac{\mathrm{d}}{\mathrm{d}t}\Big) \cdot g_i(t) ,$$

which is an ODE. Noticing that $\gamma_*\frac{\mathrm{d}}{\mathrm{d}t} \equiv X$ is the vector tangent to γ in M, we can write

$$\frac{\mathrm{d}g_i(t)}{\mathrm{d}t} = -\mathcal{A}_i(X)\,g_i(t)$$

in terms of the local form of the connection.

Remark. The solution can be formally written as a path-ordered exponential:

$$g_i(t) = \mathrm{P}\exp\Big\{ - \int_{\gamma(0)}^{\gamma(t)} \mathcal{A}_{i\mu}\big(\gamma(t)\big)\, d\gamma^\mu \Big\} .$$

Corollary 7.11. *Let $\tilde{\gamma}'$ be another horizontal lift of γ, such that $\tilde{\gamma}'(0) = \tilde{\gamma}(0)\, g$. Then $\tilde{\gamma}'(t) = \tilde{\gamma}(t)\, g$ for all t.*

Proof. We have to show that if $\tilde{\gamma}(t)$ is a horizontal lift of γ, then also $\tilde{\gamma}'(t) \equiv \tilde{\gamma}(t)\, g = R_g \tilde{\gamma}(t)$ is. The horizontal lift $\tilde{\gamma}$ satisfies the equation

$$\omega\Big(\tilde{\gamma}_* \frac{\mathrm{d}}{\mathrm{d}t}\Big) = 0 \qquad\qquad \text{for all } t \in [0,1]\ .$$

Then, using the properties of the connection one-form,

$$\omega\Big(\tilde{\gamma}'_* \frac{\mathrm{d}}{\mathrm{d}t}\Big) = \omega\Big(R_{g*}\tilde{\gamma}_* \frac{\mathrm{d}}{\mathrm{d}t}\Big) = R_g^*\omega\Big(\tilde{\gamma}_* \frac{\mathrm{d}}{\mathrm{d}t}\Big) = g^{-1}\,\omega\Big(\tilde{\gamma}_* \frac{\mathrm{d}}{\mathrm{d}t}\Big)\, g = 0\ ,$$

which concludes the proof. □

Given a curve $\gamma : [0,1] \to M$ and a point $u_0 \in \pi^{-1}\big(\gamma(0)\big)$, there is a unique horizontal lift $\tilde{\gamma}$ of γ such that $\tilde{\gamma}(0) = u_0$, and hence a unique point

$$u_1 = \tilde{\gamma}(1) \in \pi^{-1}\big(\gamma(1)\big)\ .$$

The point u_1 is called the **parallel transport** of u_0 along the curve γ.

In physics, we often need to differentiate sections of a vector bundle which is associated with a certain principal bundle. For example, a charged scalar field in QED is regarded as a section of a $\mathbb{C}$ vector bundle associated to a $U(1)$ principal bundle $P(M, U(1))$. A connection one-form ω on a principal bundle enable us to define the *covariant derivative* in associated bundles to P.

Let $P(M,G)$ be a G-bundle with projection π_P, ρ a representation of G on the vector space V, and $E = P \times_\rho V$ the associated bundle whose elements are the classes

$$[(u,v)] = \big\{(ug, \rho(g)^{-1}v) \,\big|\, u \in P,\, v \in V,\, g \in G\big\}\ .$$

Given an element $[(u_0,v)] \in \pi_E^{-1}(p)$, it is natural to define its parallel transport along a curve γ in M as

$$\big[(\tilde{\gamma}(t), v)\big]$$

where $\tilde{\gamma}(t)$ is the horizontal lift of γ in P with $\tilde{\gamma}(0) = u_0$.

To define the covariant derivative, we notice that given a section $s \in \Gamma(M,E)$ and a curve $\gamma : [-1,1] \to M$, we can always represent the section along the curve as

$$s\big(\gamma(t)\big) = \big[(\tilde{\gamma}(t), \eta(t))\big]$$

for some $\eta(t) \in V$. We define the **covariant derivative** of s along $\gamma(t)$ at $p = \gamma(0)$ as

$$\nabla_X s \equiv \Big[\Big(\tilde{\gamma}(0),\, \frac{\mathrm{d}}{\mathrm{d}t}\eta(t)\big|_{t=0}\Big)\Big]\ ,$$

where X is the tangent vector to $\gamma(t)$ at p. By construction, the parallel transport is a covariantly-constant transport.

The covariant derivative can be computed at any point of M, therefore if X is a vector field on M then ∇_X is a map $\Gamma(M,E) \to \Gamma(M,E)$. Since this map turns out to be point-wise linear in X, we can think of ∇ as a map

$$\nabla : \Gamma(M,E) \to \Omega^1(M) \otimes \Gamma(M,E) \;,$$

in the sense that $\langle \nabla s, X\rangle = \nabla_X s$.

Proposition 7.12. *The covariant derivative satisfies the following properties:*

$$\begin{aligned} \nabla(a_1 s_1 + a_2 s_2) &= a_1 \nabla s_1 + a_2 \nabla s_2 \\ \nabla(fs) &= \mathrm{d}f \otimes s + f\, \nabla s \\ \nabla_{(a_1 X_1 + a_2 X_2)} s &= a_1 \nabla_{X_1} s + a_2 \nabla_{X_2} s \\ \nabla_{fX} s &= f\, \nabla_X s \;. \end{aligned}$$

Here $s \in \Gamma(M,E)$, $a_{1,2} \in \mathbb{R}$ and $f \in \mathscr{F}(M)$.

¶ *Exercise* 21. Prove them.

Proof. The first one follows from the linear structure of V. The second one, when contracted with X, follows from $\frac{\mathrm{d}}{\mathrm{d}t}\big(f(\gamma(t))\,\eta(t)\big) = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t)) \cdot \eta(t) + f\,\frac{\mathrm{d}}{\mathrm{d}t}\eta(t) = \langle df, X\rangle \eta(t) + f\,\frac{\mathrm{d}}{\mathrm{d}t}\eta(t)$.

The last two, which we already used in the definition of ∇, are more easily proven in the local expression given below. □

Take a local section $\sigma_i(p) \in \Gamma(U_i, P)$ and a local trivialization $\phi_i(p, \mathbb{1}) = \sigma_i$ on P. If e^0_α is a basis vector of V, we let $e_\alpha(p) = \big[(\sigma_i(p), e^0_\alpha)\big]$ be a local section of the associated bundle $E = P \times_\rho V$. We compute the covariant derivative $\nabla_X e_\alpha$. Let $\gamma : [-1,1] \to M$ be a curve tangent to X and

$$\tilde\gamma(t) \;\equiv\; \sigma_i(\gamma(t)) \cdot g_i(t)$$

its horizontal lift in P. Then

$$e_\alpha(t) = \big[\big(\sigma_i(t),\, e^0_\alpha\big)\big] = \big[\big(\tilde\gamma(t)\, g_i^{-1}(t),\, e^0_\alpha\big)\big] \;\sim\; \big[\big(\tilde\gamma(t),\, \rho(g_i^{-1}(t))\, e^0_\alpha\big)\big] \;.$$

The covariant derivative is

$$\begin{aligned} \nabla_X e_\alpha &= \Big[\Big(\tilde\gamma(0),\; \frac{\mathrm{d}}{\mathrm{d}t}\,\rho\big(g_i^{-1}(t)\big)\, e^0_\alpha\Big|_{t=0}\Big)\Big] \\ &= \Big[\Big(\tilde\gamma(0),\; -\rho\Big(g_i^{-1}(0)\,\frac{\mathrm{d}g_i(t)}{\mathrm{d}t}\Big|_{t=0}\, g_i^{-1}(0)\Big)\, e^0_\alpha\Big)\Big] \\ &= \Big[\Big(\sigma_i(0),\; \rho\big(\mathcal{A}_i(X)\big)\, e^0_\alpha\Big)\Big] \end{aligned}$$

In representation ρ, the one-form $\mathcal{A}_i$ is expanded into generators T_a that act as

$$T_a\, e^0_\alpha = (T_a)^\beta{}_\alpha\, e^0_\beta \;.$$

Therefore, in a local trivialization of E we have

$$\nabla_X e_\alpha = X^\mu \left(\mathcal{A}_i\right)_\mu{}^\beta{}_\alpha \, e_\beta \ .$$

The covariant derivative of a generic section $s(p) = \xi_i^\alpha(p)\, e_\alpha$ follows from the properties listed above:

$$\nabla_X s = X^\mu \left(\frac{\partial \xi_i^\alpha}{\partial x^\mu} + \left(\mathcal{A}_i\right)_\mu{}^\beta{}_\alpha \, \xi_i^\beta \right) e_\alpha \ .$$

¶ *Exercise* 22. Show that the covariant derivative is independent of the local trivialization σ_i chosen, namely that σ_i, ξ_i and $\mathcal{A}_i$ transform correctly to keep $\nabla_X s$ invariant.

¶ *Exercise* 23. Show that the last two properties in Proposition 7.12 are true.

Proposition 7.13. *We can define the action of ∇ on fibre-valued p-forms $s \otimes \eta$, where $s \in \Gamma(M, E)$ and $\eta \in \Omega^p(M)$, by*

$$\nabla(s \otimes \eta) = \nabla s \wedge \eta + s \otimes d\eta \ .$$

Then, if $s(p) = \xi_i^\alpha(p)\, e_\alpha(p)$ is a section of E, we have

$$\nabla\nabla s = e_\alpha \otimes \left(\mathcal{F}_i\right)^\alpha{}_\beta \, \xi_i^\beta = \rho(\mathcal{F}_i)\, s \ .$$

Proof. Use that, in this notation, $\nabla e_\alpha = e_\beta \otimes \left(\mathcal{A}_i\right)^\beta{}_\alpha$. □

8 Lie algebras

A Lie algebra $\mathfrak{g}$ is a vector space equipped with an antisymmetric bilinear operation

$$[\,\cdot\,,\cdot\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \;,$$

called a **Lie bracket** or **commutator**, constrained to satisfy the Jacobi identity

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \qquad\qquad \forall\, X,Y,Z \in \mathfrak{g} \;.$$

One can study *complex* or *real* Lie algebras, that are vector spaces over $\mathbb{C}$ or $\mathbb{R}$, respectively. For the purpose of studying general properties, it is convenient to work over the complex field. We will also restrict to *finite-dimensional* algebras.

A Lie algebra can be specified by a set of generators $\{J^a\}$ (in the sense that they are a basis of $\mathfrak{g}$ as a vector space) and their commutation relations

$$[J^a, J^b] = \sum_c i\, f^{ab}{}_c\, J^c \;,$$

where $f^{ab}{}_c = -f^{ba}{}_c$. The number of generators is the dimension of the algebra. We will start studying *complex* Lie algebras. The complex numbers $f^{ab}{}_c$ are the **structure constants**, and we have inserted a factor of i such that if the generators are Hermitian, $(J^a)^\dagger = J^a$, then $f^{ab}{}_c$ are real.

8.1 Representations

Definition 8.1. A **representation** ρ of a Lie algebra $\mathfrak{g}$ is a vector space V, together with a homomorphism

$$\rho : \mathfrak{g} \to \mathrm{End}(V) \;.$$

Thus, every element of $\mathfrak{g}$ is "represented" by a linear operator on V. If V is finite-dimensional, the linear operators are simply matrices. For ρ to be a homomorphism, we require that it maps the Lie bracket to a commutator:

$$\rho\big([X,Y]\big) = \rho(X)\,\rho(Y) - \rho(Y)\,\rho(X) \equiv \big[\rho(X), \rho(Y)\big] \;.$$

Notice that in a representation we can multiply operators, while in the abstract algebra this operation is not defined. The dimension of V is the **dimension** of the representation.

There is a representation that is intrinsically given once the Lie algebra $\mathfrak{g}$ is given: it is called the **adjoint** representation, and uses the algebra $\mathfrak{g}$ itself as the vector space on which the generators act:

$$\mathrm{ad}(X)\, Y \equiv [X,Y] \;.$$

The property $\mathrm{ad}\big([X,Y]\big) = \big[\mathrm{ad}(X), \mathrm{ad}(Y)\big]$ follows from the Jacobi identity.[29] In the basis $\{J^a\}$, the generators are given by the matrices $(J^a)^b{}_c = i f^{ab}{}_c$.

Two representations $\rho_{1,2}$ on V are **isomorphic** if there exist a linear map R such that $\rho_1 = R\rho_2 R^{-1}$.

Given a representation $\rho(J^a) = T^a$, the generators T^a are matrices that satisfy

$$[T^a, T^b] = i\, f^{ab}{}_c T^c \ .$$

Then the generators $\widetilde{T}^a = -T^{a\mathsf{T}}$ satisfy the same relation, and thus are also a representation. This is called the **conjugate representation** ρ^*. If ρ^* is isomorphic to ρ, we say that the representation is **self-conjugate**.

8.2 Simple and semi-simple algebras

Definition 8.2.

(a) A subspace $\mathfrak{s} \subset \mathfrak{g}$ that is closed under Lie bracket, $[\mathfrak{s},\mathfrak{s}] \subseteq \mathfrak{s}$, is called a Lie subalgebra.

(b) A subspace $\mathfrak{i} \subset \mathfrak{g}$ that is invariant under Lie bracket (this is a stronger condition),

$$[\mathfrak{i}, \mathfrak{g}] \subseteq \mathfrak{i} \ ,$$

is called an **ideal**.

If $\mathfrak{i}$ is an ideal, then the quotient $\mathfrak{g}/\mathfrak{i}$, in the sense of a quotient of vector spaces, is also a Lie algebra.[30] The elements of $\mathfrak{g}/\mathfrak{i}$ are equivalence classes of $x \sim x + a$ with $x \in \mathfrak{g}$ and $a \in \mathfrak{i}$. Then

$$[x+a, y+b] = [x,y] + [a,y] + [x,b] + [a,b] \sim [x,y]$$

because the last three terms are in $\mathfrak{i}$. Thus the Lie bracket extends to the quotient.

A canonical example of ideal is the **derived subalgebra** $\mathfrak{g}'$ of $\mathfrak{g}$:

$$\mathfrak{g}' \;\equiv\; \big\langle [\mathfrak{g},\mathfrak{g}] \big\rangle = \Big\langle \big\{ [x,y] \,\big|\, x, y \in \mathfrak{g} \big\} \Big\rangle \ ,$$

where $\langle\ \rangle$ is the linear span over the field of the algebra, say $\mathbb{C}$ or $\mathbb{R}$. It is clear that $[\mathfrak{g}', \mathfrak{g}] \subseteq \mathfrak{g}'$. The quotient

$$\mathfrak{g}/\mathfrak{g}' \qquad \text{is an Abelian Lie algebra} \ ,$$

[29] The Jacobi identity implies

$$\big(\mathrm{ad}(X)\,\mathrm{ad}(Y) - \mathrm{ad}(Y)\,\mathrm{ad}(X)\big)\, Z = [X,[Y,Z]] - [Y,[X,Z]] = [[X,Y],Z] = \mathrm{ad}\big([X,Y]\big)\, Z \ .$$

[30] This is similar to the fact that G/H is a group if H is a *normal* subgroup of G.

since any commutator is equivalent to zero.[31]

Abelian Lie algebras are not particularly interesting, as the Lie bracket is simply zero: $[\mathfrak{g}, \mathfrak{g}] = 0$. Lie algebras with a *proper* ideal, *i.e.* an ideal $\mathfrak{i} \subsetneq \mathfrak{g}$, can be understood as extensions of the quotient $\mathfrak{g}/\mathfrak{i}$ by $\mathfrak{i}$.[32] This motivates the study of *simple* Lie algebras:

Definition 8.3. A non-Abelian[33] Lie algebra that does not contain any *proper ideal* is called **simple**. A direct sum of simple algebras is called **semi-simple**.

8.3 Killing form

We introduce a symmetric bilinear form on $\mathfrak{g}$, called the **Killing form**, defined as[34]

$$k(X, Y) \equiv \frac{1}{2\tilde{g}} \operatorname{Tr}\big(\operatorname{ad}(X)\operatorname{ad}(Y)\big) .$$

The normalization constant $\tilde{g} > 0$ will be fixed in Section 10.2. The Killing form is invariant in the following sense:

$$k\big([Z, X], Y\big) + k\big(X, [Z, Y]\big) = 0 .$$

This simply follows from the definition.

Theorem 8.4. (Cartan's criterion) *A Lie algebra $\mathfrak{g}$ is semi-simple if and only if the Killing form k is non-degenerate.*

We will not prove this theorem, which requires more machinery.

We can use the Killing form to construct the totally antisymmetric tensor

$$i\, f^{abc} \equiv k\big([J^a, J^b], J^c\big) = i \sum_d f^{ab}{}_d\, k^{dc} ,$$

where

$$k^{dc} \equiv k(J^d, J^c) .$$

¶ *Exercise* 24. Show, using invariance of the Killing form, that f^{abc} is totally antisymmetric.

If the Killing form is non-degenerate, one could similarly construct f_{abc} lowering the indices with k_{ab}, which is the inverse matrix of k^{ab}.

[31] The quotient $\mathfrak{g}/\mathfrak{g}'$ is *not* the Cartan subalgebra. For instance, for $\mathfrak{g} = \mathfrak{su}(2)$ one finds $\mathfrak{g}' = \mathfrak{su}(2)$.

[32] An extension is a short exact sequence $0 \to \mathfrak{i} \to \mathfrak{g} \to \mathfrak{g}/\mathfrak{i} \to 0$. Similarly, G is an H-bundle over G/H.

[33] We should specify that the algebra is non-Abelian, otherwise the 1-dimensional algebra, which is necessarily Abelian, would turn out to be simple. Alternatively, we specify that the dimension must be ≥ 2.

[34] The trace of a linear operator A on V is defined as follows. Take a basis v^i of V, and express $Av^i = a^i{}_j v^j$. Then $\operatorname{Tr} A \equiv \operatorname{Tr}(a)$ where the last one is the trace of a matrix. One can check that this definition is independent of the basis chosen.

8.4 The example of $\mathfrak{su}(2)$

Let us consider the algebra of $SL(2,\mathbb{C})$ first. The group $SL(2,\mathbb{C})$ is the group of complex 2×2 matrices with determinant 1:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2\times 2} \qquad \text{with} \qquad \det M = 1 \ .$$

The algebra is the tangent space at $\mathbb{1}$. We use the exponential map:

$$M = \exp\left\{tA\right\} = \mathbb{1} + tA + \mathcal{O}(t^2) \ .$$

The algebra is the set $\{A \,|\, M \in SL(2,\mathbb{C})\}$. We use the formula[35]

$$1 = \det M = \exp\left\{t \operatorname{Tr} A\right\} = 1 + t \operatorname{Tr} A + \mathcal{O}(t^2) \ .$$

We conclude that

$$\mathfrak{sl}(2,\mathbb{C}) = \left\{A \in \mathbb{C}^{2\times 2} \,\middle|\, \operatorname{Tr} A = 0\right\} \ .$$

We can choose the following generators:

$$J_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \ , \qquad J_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ , \qquad J_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \ .$$

They generate $\mathfrak{sl}(2,\mathbb{C})$ as a vector space over $\mathbb{C}$.

Now consider the group $SL(2,\mathbb{R})$ of real 2×2 matrices with determinant 1. Its algebra $\mathfrak{sl}(2)$ is the set of real traceless matrices, and we can take the same generators $J_3, J_\pm$ but now in a vector space with real coefficients. In other words, with complex coefficients we get the same algebra, but there is a restriction of $\mathfrak{sl}(2,\mathbb{C})$ to real coefficients which gives the (real) algebra $\mathfrak{sl}(2)$. This restriction is called a *real form*.

As another example, consider the algebra of $SU(2)$. This is the group of unitary matrices with determinant 1:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2\times 2} \qquad \text{with} \qquad M^\dagger = M^{-1} \ , \quad \det M = 1 \ .$$

Setting

$$M = e^{itA} \ , \qquad M^\dagger = e^{-itA^\dagger} \ , \qquad M^{-1} = e^{-itA} \ ,$$

we find that the algebra is

$$\mathfrak{su}(2) = \left\{A \in \mathbb{C}^{2\times 2} \,\middle|\, A = A^\dagger \ , \quad \operatorname{Tr} A = 0\right\} \ .$$

[35]Suppose that A is diagonalizable. In a basis in which A is diagonal the formula is obvious. Since $\det M$ is a continuous function (of the entries) and since the set of diagonalizable matrices is dense in the space of all matrices, the formula follows for a generic matrix A.

A set of generators is

$$J_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} = \frac{\sigma_3}{2} \;, \qquad J_1 = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} = \frac{\sigma_1}{2} \;, \qquad J_2 = \begin{pmatrix} 0 & -\frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix} = \frac{\sigma_2}{2} \;,$$

where σ_i are the Pauli matrices, and we should use real coefficients. However over the field $\mathbb{C}$ we can also use the generators

$$J_1 \pm i J_2 = J_\pm \;.$$

Therefore, over $\mathbb{C}$ we get the same algebra as $\mathfrak{sl}(2,\mathbb{C})$, while over $\mathbb{R}$ $\mathfrak{su}(2)$ is another real form of $\mathfrak{sl}(2,\mathbb{C})$. In the following we will indicate the complex algebra as $\mathfrak{su}(2)$.

The commutation relations are

$$[J_3, J_\pm] = \pm J_\pm \;, \qquad\qquad [J_+, J_-] = 2J_3 \;.$$

Suppose we have a *finite-dimensional* representation, and let $|\lambda\rangle$ be an eigenvector of J_3:

$$J_3 \,|\lambda\rangle = \lambda \,|\lambda\rangle \;.$$

Then $J_\pm$ increase/decrease the eigenvalue,

$$J_3\big(J_\pm\,|\lambda\rangle\big) = (\lambda \pm 1)\, J_\pm\,|\lambda\rangle \;,$$

provided $J_\pm|\lambda\rangle \neq 0$. All the states we get this way are linearly independent (since they have different eigenvalue), and since the representation is finite-dimensional, the process has to stop both going up and going down.

Thus, necessarily there exists a state $|j\rangle$, with $J_3\,|j\rangle = j\,|j\rangle$, such that

$$J_+\,|j\rangle = 0 \;.$$

This is called a **highest weight state**. We then define the following other states:

$$|j-k\rangle \equiv (J_-)^k\,|j\rangle \;.$$

The action of the algebra generators on them is

$$\begin{aligned} J_3\,|j-k\rangle &= (j-k)\,|j-k\rangle \\ J_-\,|j-k\rangle &= |j-k-1\rangle \\ J_+\,|j-k\rangle &= k(2j-k+1)\,|j-k+1\rangle \;. \end{aligned}$$

The last one follows from the commutation relations (¶).[36] As we said, there must be a $k_{\max}$ such that $|j-k_{\max}\rangle \neq 0$ but $|j-k_{\max}-1\rangle = 0$. In particular $J_+|j-k_{\max}-1\rangle = 0$. Substituting above, this happens for $k_{\max} = 2j$.

[36] In fact

$$J_+|j-k\rangle = J_+J_-^k|j\rangle = \big(2J_3J_-^{k-1} + J_-J_+J_-^{k-1}\big)|j\rangle = \big(2J_3J_-^{k-1} + 2J_-J_3J_-^{k-2} + J_-^2J_+J_-^{k-2}\big)|j\rangle = \ldots$$
$$= 2\big(J_3J_-^{k-1} + J_-J_3J_-^{k-2} + \ldots + J_-^{k-1}J_3\big)|j\rangle = 2\sum_{s=0}^{k-1}(j-s)|j-k+1\rangle \;.$$

The states above form an *irreducible* representation, because they are all connected by the action of $\mathfrak{su}(2)$, and they are closed under that action. It is called a *highest weight representation*. Thus:

(i) In finite-dimensional representations, $j \in \frac{1}{2}\mathbb{Z}$.

(ii) The irreducible representation has $2j+1$ states, with $J_3 = \{-j, -j+1, \ldots, j-1, j\}$.

We could easily write the generators in this basis (essentially we already did this).

On the other hand, we can write the generators in the adjoint representation (from the commutation relations):

$$J_3 = \begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix} , \qquad J_+ = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} , \qquad J_- = \begin{pmatrix} 0 & 0 & 0 \\ -2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

We see that the adjoint representation corresponds to $j = 1$.

8.5 Cartan-Weyl basis

We are concerned with *simple Lie algebras*. A convenient choice of basis is the **Cartan-Weyl basis**. We first find a maximal set of commuting (Hermitian) generators H^i, $i = 1, \ldots, r$, where r is the **rank** of the algebra:

$$[H^i, H^j] = 0 .$$

These generators form the **Cartan subalgebra** $\mathfrak{h}$. Since the generators of the Cartan subalgebra can be simultaneously diagonalized, we choose the remaining generators to be combinations of the J^a's that are common eigenvectors:

$$[H^i, E^\alpha] = \alpha^i \, E^\alpha .$$

The vector $\alpha = (\alpha^1, \ldots, \alpha^r)$ is called a **root** and E^α its ladder operator. Because $\mathfrak{h}$ is the maximal Abelian subalgebra, the roots are non-zero. One can also prove that they are all different (see footnote 38). Since a root maps an element $H^i \in \mathfrak{h}$ to $\alpha^i \in \mathbb{R}$,

$$\alpha(H^i) = \alpha^i ,$$

roots are elements of the dual to the Cartan subalgebra: $\alpha \in \mathfrak{h}^*$.

Let us make use of the Killing form and the fact that it is non-degenerate.

- Applying invariance to $k\big([H^i, E^\alpha], H^j\big)$ we obtain $\alpha^i \, k(E^\alpha, H^j) = 0$, therefore

$$k(E^\alpha, H^j) = 0 .$$

- Applying invariance to $k\big([H^i, E^\alpha], E^\beta\big)$ we obtain $(\alpha^i + \beta^i)\, k(E^\alpha, E^\beta) = 0$, therefore

$$k(E^\alpha, E^\beta) = 0 \qquad\qquad \text{whenever } \beta \neq -\alpha \;.$$

  Since k is non-degenerate, necessarily $-\alpha$ is a root whenever α is a root.[37]

  Then necessarily $k(E^\alpha, E^{-\alpha}) \neq 0$.

  [From here we can prove that roots do not have multiplicities.][38]

- We also conclude that k restricted to the Cartan subalgebra must be non-degenerate. We can then choose H^i such that[39]

$$k(H^i, H^j) = \delta^{ij} \;.$$

We denote the **set of roots** by Δ. Consider the adjoint representation:

$$H^i \mapsto |H^i\rangle \;, \qquad\qquad E^\alpha \mapsto |E^\alpha\rangle \equiv |\alpha\rangle \;.$$

The dimension of the adjoint representation is equal to the dimension of $\mathfrak{g}$, and we see that

$$\dim \mathfrak{g} = \#\Delta + r \;.$$

We also see that the number of roots is even.

---

[37]Notice that the same conclusion is reached if we assume that H^i are Hermitian. Taking the Hermitian conjugate of the eigenvalue equation, using $H^i = (H^i)^\dagger$ and the fact that the eigenvalues are real, we find

$$[H^i, (E^\alpha)^\dagger] = -\alpha^i\, (E^\alpha)^\dagger \qquad \text{meaning} \qquad (E^\alpha)^\dagger = E^{-\alpha} \;.$$

Thus $-\alpha$ is necessarily a root, when α is.

[38]Take E^α, a vector in the root space L_α of α. Because of non-degeneracy of k, there must exist a vector $E^{-\alpha}$ in $L_{-\alpha}$ such that $k(E^\alpha, E^{-\alpha}) = C \neq 0$. Then $k\big([E^\alpha, E^{-\alpha}], H^i\big) = C\alpha^i$, and thus $[E^\alpha, E^{-\alpha}] = C\, H^\alpha$ where $H^\alpha = \alpha^i H^i$. Now consider the subspace $R = \langle E^\alpha\rangle \oplus \mathfrak{h} \oplus L_{-\alpha} \oplus \cdots \oplus L_{-p\alpha}$, where p is such that $-p\alpha$ is a root but $-(p+1)\alpha$ is not. This space is a representation of $\{E^\alpha, E^{-\alpha}, H^\alpha\}$ as it is closed under their adjoint action. We compute $\mathrm{Tr}_R\, H^\alpha = \mathrm{Tr}_R[E^\alpha, E^{-\alpha}] = 0$ because it is a commutator. On the other hand

$$\mathrm{Tr}_R\, H^\alpha = \sum_i \alpha^i \sum_{\text{spaces}} \mathrm{Tr}_{\text{spaces}}\, H^i = \sum_i \alpha^i \big(\alpha^i - \sum_\ell d_{-\ell}\, \ell \alpha^i\big) = |\alpha|^2 \Big(1 - \sum_\ell d_{-\ell}\, \ell\Big) \;,$$

where $d_{-\ell}$ is the dimension of $L_{-\ell\alpha}$. We know that $|\alpha|^2 \neq 0$ and $d_{-1} \geq 1$. We conclude that $d_{-1} = 1$ and $d_{-\ell} = 0$ for $\ell \geq 2$. In particular L_{-1} is one-dimensional.

Repeating the argument for all roots, we conclude that all L_α are one-dimensional, and moreover if α is a root, the only multiple of α which is also a root is $-\alpha$.

[39]Since we are studying the algebra over $\mathbb{C}$, we can always make k positive-definite. What is non-trivial is that, in that basis, the roots are real. To show that, consider
$k(H^\alpha, H^\alpha) = \sum_i \alpha^i \alpha^i = (\alpha, \alpha) = \alpha^i \alpha^j\, \mathrm{Tr}\left[\mathrm{ad}(H^i)\mathrm{ad}(H^j)\right] = \alpha^i \alpha^j \sum_\beta \beta^i \beta^j = \sum_\beta (\alpha, \beta)^2$.
On the other hand, $(\alpha, \beta) = (\alpha, \alpha)(q_\beta - p_\beta)/2$ where q_β, p_β are the largest possible integers such that $\beta + p_\beta \alpha$ and $\beta - q_\beta \alpha$ are non-zero. We conclude that $(\alpha, \alpha) = 4/\sum_\beta (q_\beta - p_\beta)^2$ is positive, and so (α, β) is real.

Let us evaluate the restriction of the Killing form to $\mathfrak{h}$. We have $\mathrm{ad}(H^i)H^j = 0$ and $\mathrm{ad}(H^i)E^\alpha = \alpha^i E^\alpha$. Therefore, in this basis, $\mathrm{ad}(H^i)$ is a diagonal matrix with entries α^i on E^α. This implies that

$$\mathrm{Tr}\left(\mathrm{ad}(H^i)\,\mathrm{ad}(H^j)\right) = \sum\nolimits_\alpha \alpha^i \alpha^j \ .$$

Imposing $k(H^i, H^j) = \delta^{ij}$ and contracting with δ_{ij}, we find that

$$\tilde{g} = \frac{\sum_\alpha |\alpha|^2}{2r} \qquad\qquad \text{where} \qquad |\alpha|^2 \equiv \sum\nolimits_{i=1}^r \alpha^i \alpha^i \ .$$

This fixes $\tilde{g}$ in terms of the normalization of the root lengths.

The Jacobi identity implies

$$[H^i, [E^\alpha, E^\beta]] = (\alpha^i + \beta^i)\,[E^\alpha, E^\beta] \ .$$

(i) If $\alpha + \beta \in \Delta$, the commutator $[E^\alpha, E^\beta]$ must be proportional to $E^{\alpha+\beta}$ (with a proportionality constant $N_{\alpha,\beta} \neq 0$ that one can compute).

(ii) If $0 \neq \alpha + \beta \notin \Delta$, then $[E^\alpha, E^\beta] = 0$.

(iii) If $\beta = -\alpha$, then $[E^\alpha, E^{-\alpha}]$ commutes with all H^i and thus must belong to the Cartan subalgebra.

Since

$$k\big([E^\alpha, E^{-\alpha}], H^i\big) = k\big(E^{-\alpha}, [H^i, E^\alpha]\big) = \alpha^i\, k(E^\alpha, E^{-\alpha}) \ ,$$

it follows that $[E^\alpha, E^{-\alpha}] = \alpha^i H^i\, k(E^\alpha, E^{-\alpha})$. We use the symbol $\alpha \cdot H \equiv \alpha^i H^i$. We can then rescale the generators such that

$$k(E^\alpha, E^{-\alpha}) = \frac{2}{|\alpha|^2} \ .$$

We finally get

$$\boxed{\begin{aligned} &[H^i, H^j] = 0 \\ &[H^i, E^\alpha] = \alpha^i\, E^\alpha \\ &[E^\alpha, E^\beta] = N_{\alpha,\beta}\, E^{\alpha+\beta} && \text{if } \alpha + \beta \in \Delta \\ & = 2\,\frac{\alpha \cdot H}{|\alpha|^2} && \text{if } \alpha = -\beta \\ & = 0 && \text{otherwise} \ . \end{aligned}}$$

Here $N_{\alpha,\beta}$ are non-vanishing constants. This is called the **Cartan-Weyl basis**.

The fundamental role of the Killing form is to establish an isomorphism between the Cartan subalgebra $\mathfrak{h}$ and its dual $\mathfrak{h}^*$. The form $k(H^i,\,\cdot\,)$ maps $\mathfrak{h} \to \mathbb{R}$ and thus is an element of $\mathfrak{h}^*$.

To every element $\gamma \in \mathfrak{h}^*$, there corresponds an element $H^\gamma \in \mathfrak{h}$ such that

$$\gamma = k(H^\gamma,\,\cdot\,) \qquad \in \mathfrak{h}^* \,.$$

In particular $H^\alpha = \alpha \cdot H = \sum_i \alpha^i H^i$. We can use this isomorphism to induce a positive-definite scalar product on $\mathfrak{h}^*$:

$$(\gamma,\beta) \;\equiv\; k(H^\gamma, H^\beta) = \sum\nolimits_i \gamma^i \beta^i \,.$$

This is a scalar product on root space. In particular

$$(\alpha,\alpha) = |\alpha|^2 = \sum\nolimits_i \alpha^i \alpha^i \,.$$

Notice also that the non-degeneracy of k implies that the roots span $\mathfrak{h}^*$.[40]

8.6 Weights

So far we have analyzed the algebra from the point of view of the adjoint representation. But much of what we said can be repeated in other representations.

Given a representation, we consider a basis $\big\{ |\lambda\rangle \big\}$ that diagonalizes the Cartan generators:

$$H^i \,|\lambda\rangle = \lambda^i \,|\lambda\rangle \,.$$

The eigenvalues λ^i form a vector $(\lambda^1, \ldots, \lambda^r)$, called a **weight**. Weights, as the roots, are elements of $\mathfrak{h}^*$:

$$\lambda(H^i) = \lambda^i \,,$$

and the scalar product extends to them. In fact, the *roots* are the weights of the adjoint representation.

The ladder operators E^α change the eigenvalue of a weight by α:

$$H^i\big(E^\alpha \,|\lambda\rangle\big) = [H^i, E^\alpha]\,|\lambda\rangle + E^\alpha H^i \,|\lambda\rangle = (\lambda^i + \alpha^i)\, E^\alpha \,|\lambda\rangle \,.$$

[40] Suppose that the roots only span a subspace of $\mathfrak{h}^*$, then there exists r numbers $\{x_i\}$ such that $\sum_{i=1}^r x_i \alpha^i = 0$ for all $\alpha \in \Delta$. Construct $H^x \equiv \sum_{i=1}^r x_i H^i$. We can see that such an element would be orthogonal to all H^j:

$$k(H^x, H^j) = \sum_{i=1}^r x_i\, k(H^i, H^j) \;\propto\; \sum_{i=1}^r x_i \sum_\alpha \alpha^i \alpha^j = \sum_\alpha \Big(\sum\nolimits_i x_i \alpha^i \Big)\, \alpha^j = 0 \,.$$

This would contradict the fact that k is non-degenerate on $\mathfrak{h}$.

We are interested in *finite-dimensional representations*. For those, there must exist integers p, q such that

$$(E^\alpha)^{p+1}|\lambda\rangle \sim E^\alpha\,|\lambda + p\alpha\rangle = 0$$
$$(E^{-\alpha})^{q+1}|\lambda\rangle \sim E^{-\alpha}\,|\lambda - q\alpha\rangle = 0$$

for any root α. In fact, notice that E^α, $E^{-\alpha}$ and $\alpha \cdot H/|\alpha|^2$ form an $\mathfrak{su}(2)$ subalgebra of $\mathfrak{g}$. We can write

$$\Big[\frac{\alpha \cdot H}{|\alpha|^2}\,,\,E^{\pm\alpha}\Big] = \pm E^{\pm\alpha}\,, \qquad \big[E^\alpha, E^{-\alpha}\big] = 2\,\frac{\alpha \cdot H}{|\alpha|^2}\ .$$

Thus we identify $J_3 \equiv \frac{\alpha\cdot H}{|\alpha|^2}$ and $J_\pm = E^{\pm\alpha}$. The state $|\lambda\rangle$ is part of a finite-dimensional representation of $\mathfrak{su}(2)$. Let it be the state with J_3 equal to m. If the total spin is j, the highest-weight state is reached in p steps and the lowest in q, it means that $m = j-p = -j+q$. Since m is the eigenvalue of J_3 we have:

$$\frac{(\alpha,\lambda)}{|\alpha|^2}\,|\lambda\rangle = \frac{\alpha \cdot H}{|\alpha|^2}\,|\lambda\rangle = m\,|\lambda\rangle = (j-p)\,|\lambda\rangle = (-j+q)\,|\lambda\rangle\ .$$

Combining the equations we get

$$\boxed{2\frac{(\alpha,\lambda)}{|\alpha|^2} = -(p-q)\ \in \mathbb{Z}\ .}$$

This will be very important later on.

8.7 Simple roots and the Cartan matrix

The number of roots is equal to the dimension of the algebra minus the rank, and in general this is much larger than the rank. Thus the roots are linearly dependent. We fix a basis $\{\beta_1, \ldots, \beta_r\}$ in $\mathfrak{h}^*$, so that any root can be expanded as

$$\alpha = \sum_{j=1}^r n_j\,\beta_j\ .$$

In this basis we define an *ordering*:

- α is said to be *positive* if the first non-zero number in the sequence $(n_1, \ldots, n_r)$ is positive. Denote by Δ_+ the set of **positive roots**. Then Δ_- is the set of negative roots, and
$$\Delta_- = -\Delta_+\ .$$

- A positive root that cannot be written as the sum of two positive roots is called a **simple root** α_i. There are exactly[41] r simple roots $\{\alpha_1, \ldots, \alpha_r\}$, providing a convenient basis for $\mathfrak{h}^*$.

Notice that:

(i) $\alpha_i - \alpha_j \notin \Delta$.

Proof. Suppose $\alpha_i - \alpha_j = \gamma \in \Delta_+$, then $\alpha_i = \alpha_j + \gamma$. If instead $\alpha_i - \alpha_j = \gamma \in \Delta_-$, then $\alpha_j = \alpha_i + (-\gamma)$.

(ii) Every positive root is a sum of simple roots.

(In particular, every root can be written in the basis of simple roots using *integer coefficients*, either all non-negative or all non-positive.)

Proof. Every positive root is either simple, or it can be written as the sum of two positive roots. We can continue the argument, which has to stop because the number of roots is finite.

The scalar product of simple roots defines the **Cartan matrix**:

$$A_{ij} = \frac{2(\alpha_i, \alpha_j)}{|\alpha_j|^2} \in \mathbb{Z} \, .$$

(i) The matrix is not symmetric.

(ii) From the argument on the $\mathfrak{su}(2)$ subalgebra, its entries are *integers*. The diagonal elements are 2.

(iii) The Schwarz inequality[42] implies that (not summed) $A_{ij}A_{ji} < 4$ for $i \neq j$.

(iv) Since $\alpha_i - \alpha_j$ is not a root, $E^{-\alpha_j}|\alpha_i\rangle = 0$ and in our argument on the $\mathfrak{su}(2)$ subalgebra with α_i, α_j we have $q = 0$. Thus

$$(\alpha_i, \alpha_j) \leq 0 \qquad \text{for } i \neq j \, .$$

(v) We conclude that the off-diagonal elements A_{ij}, A_{ji} are either both 0, or one is -1 and the other one is -1, -2 or -3.

Notice that

$$(\alpha_i, \alpha_j) = |\alpha_i| \, |\alpha_j| \cos\theta_{ij} \, ,$$

[41] As shown below, the simple roots $\{\alpha_i\}$ span the positive roots Δ_+ over $\mathbb{Z}_{\geq 0}$ and so they span the lattice of roots over $\mathbb{Z}$. Thus they span $\mathfrak{h}^*$ over $\mathbb{R}$. Suppose they are not linearly independent, namely $\sum_i k_i \alpha_i = 0$ for some numbers k_i. The k_i cannot all be positive, thus separate the sum into a positive and a negative part, and write $\gamma = \sum a_i \alpha_i = \sum_i b_i \alpha_i$ with $a_i, b_i \geq 0$ and $a_i b_i = 0$ for all i. Notice that $\gamma \neq 0$. But then $(\gamma, \gamma) = \sum_{ij} a_i b_j (\alpha_i, \alpha_j) \leq 0$, using the fact (proven below) that $(\alpha_i, \alpha_j) \leq 0$ for $i \neq j$. The contradiction proves that the simple roots are linearly independent.

[42] Cauchy-Schwarz inequality: $\big|\langle u, v\rangle\big|^2 \leq \langle u, u\rangle \, \langle v, v\rangle$ with equality if and only if u, v are linearly dependent.

where θ_{ij} is the angle between the two roots, and is $\geq 90°$. Such an angle is expressed by the Cartan matrix:

$$\cos\theta_{ij} = -\frac{\sqrt{A_{ij}A_{ji}}}{2} \;\in \left\{0,\, -\frac{1}{2},\, -\frac{\sqrt{2}}{2},\, -\frac{\sqrt{3}}{2}\right\} .$$

The quantity A_{ij}/A_{ji} (not summed) is the ratio $|\alpha_i|^2/|\alpha_j|^2$ of lengths of the roots α_i, α_j, whenever they are not orthogonal (namely $(\alpha_i, \alpha_j) \neq 0$). This ratio can only be $1, 2, \frac{1}{2}, 3, \frac{1}{3}$, and it turns out[43] that there can only be at most two lengths in a simple Lie algebra. When all roots have the same length, the algebra is said to be **simply laced**.

It is convenient to introduce a special notation for the quantity $2\alpha_i/|\alpha_i|^2$:

$$\alpha_i^\vee = \frac{2\alpha_i}{|\alpha_i|^2} .$$

Here $\alpha_i^\vee$ is called the **coroot** associated to the root α_i. Notice in particular that

$$A_{ij} = (\alpha_i, \alpha_j^\vee) \in \mathbb{Z} .$$

A distinguished element of Δ is the **highest root** θ. It is the unique root for which, in the expansion $\sum m_i\alpha_i$, the sum $\sum m_i$ is maximized. All elements of Δ can be obtained by repeated subtraction of simple roots from θ.[44] The coefficients of the decomposition of θ in the bases $\{\alpha_i\}$ and $\{\alpha_i^\vee\}$ bear special names: the *marks* a_i and the *comarks* $a_i^\vee$:

$$\theta = \sum_{i=1}^r a_i\,\alpha_i = \frac{|\theta|^2}{2}\sum_{i=1}^r a_i^\vee\,\alpha_i^\vee \qquad\qquad a_i,\; a_i^\vee \in \mathbb{N} .$$

In terms of them one defines the **Coxeter number** g and the **dual Coxeter number** $g^\vee$:

$$g \;\equiv\; 1 + \sum_{i=1}^r a_i \;, \qquad\qquad g^\vee \;\equiv\; 1 + \sum_{i=1}^r a_i^\vee \;.$$

These definitions are independent of normalizations.

We will choose a normalization in which the longest roots have $|\alpha|^2 = 2$. It turns out that θ is always a long root, thus we fix $|\theta|^2 = 2$. As we will see (Section 10.2), the normalization of the Killing form is related to the dual Coxeter number by

$$2\tilde{g} = |\theta|^2 g^\vee \;\to\; 2g^\vee .$$

[43]It can be shown as follows. First, the classification of Dynkin diagrams implies that, among the simple roots, there can be at most two different lengths (see in particular point 7 in Chapter 9 of Cahn's book). Second, all roots can be obtained from the simple roots by applying Weyl reflections (see below), which preserve the scalar product.

[44]Therefore θ is the highest weight of the adjoint representation.

8.8 The Chevalley basis

The Cartan matrix contains all the information about the structure of a simple Lie algebra $\mathfrak{g}$. This is made manifest in the so-called Chevalley basis.

To each *simple* root we associate three generators:

$$e^i \equiv E^{\alpha_i} , \qquad f^i = E^{-\alpha_i} , \qquad h^i = 2\,\frac{\alpha_i \cdot H}{|\alpha_i|^2} .$$

Their commutation relations are fixed by the Cartan matrix:

$$\boxed{\begin{aligned} &[h^i, h^j] = 0 \\ &[h^i, e^j] = A_{ji}\, e^j \\ &[h^i, f^j] = -A_{ji}\, f^j \\ &[e^i, f^j] = \delta_{ij}\, h^j \ . \end{aligned}}$$

The remaining (ladder) generators are obtained by repeated commutations of these basic generators, subject to the *Serre relations*

$$\begin{aligned} \big[\mathrm{ad}(e^i)\big]^{1-A_{ji}} e^j &= 0 \\ \big[\mathrm{ad}(f^i)\big]^{1-A_{ji}} f^j &= 0 \ . \end{aligned}$$

For instance, $\mathrm{ad}(e^1)e^2 = [e^1, e^2]$ while $\big[\mathrm{ad}(e^1)\big]^2 e^2 = [e^1, [e^1, e^2]]$. These relations follow from the $\mathfrak{su}(2)$ subalgebra argument, applied to the adjoint representation, noticing that the difference of two simple roots is never a root. Thus in this procedure the generators e^i and f^j never mix, reflecting the separation of the roots into Δ_+ and Δ_-.

This procedure shows that the Lie algebra is reconstructed from the Cartan matrix A_{ij}.

8.9 Dynkin diagrams

All the information contained in the Cartan matrix can be encapsulated in a planar diagram: the **Dynkin diagram**:

(i) To each simple root α_i we associate a node.

(ii) We join node i and j with $A_{ij}A_{ji} \in \{0, 1, 2, 3\}$ lines.

(iii) If there is more than one line then the two roots have different length, and we add an arrow from the longer to the shorter (we can also draw short roots with shaded color).

The Cartan matrix and the Dynkin diagram should satisfy some extra properties:

(a) The Cartan matrix should be *indecomposable*, *i.e.* the Dynkin diagram should be connected (for *simple* Lie algebras).

If the diagram is disconnected, the Chevalley construction gives two or more commuting subalgebras which are ideals of the total algebra, in contradiction with simplicity.

(b) The Cartan matrix should have $\det A > 0$ and all principal (sub)minors positive. In other words (Sylvester's criterion) the matrix A should be a *positive-definite* bilinear form.

The Cartan matrix $A_{ij} = 2\frac{(\alpha_i,\alpha_j)}{(\alpha_j,\alpha_j)}$ can be written as the product of two matrices:

$$A = \widetilde{A}\, D \;, \qquad D_{jk} = \frac{2}{(\alpha_j, \alpha_j)} \delta_{jk} \;, \qquad \widetilde{A}_{ij} = (\alpha_i, \alpha_j) = \sum_k \alpha_i^k \alpha_j^k = (\alpha \alpha^\mathsf{T})_{ij} \;.$$

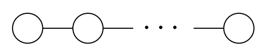
D is diagonal with positive entries, so it does not affect the conclusion. $\widetilde{A}$ is written in terms of α, which is the matrix of the roots; since the roots are linearly independent, α has maximal rank. It follows $\det \widetilde{A} = (\det \alpha)^2 > 0$. In fact, the matrix $\alpha\alpha^\mathsf{T}$ is positive definite.

It is an exercise to classify all possible Cartan matrices / Dynkin diagrams with those properties (see Chapter 9 of Cahn's book or Chapter 18 of Kac's lecture notes). This leads to the complete **classification of simple Lie algebras**. There are four infinite families and five exceptional ones:

$$A_{r\geq 1} \;, \quad B_{r\geq 2} \;, \quad C_{r\geq 1} \;, \quad D_{r\geq 2} \;, \quad E_{6,7,8} \;, \quad F_4 \;, \quad G_2 \;.$$

The subscript indicates the rank. The simply-laced algebras are A_r, D_r and $E_{6,7,8}$.

$\boldsymbol{A_{r\geq 1} \cong \mathfrak{su}(r+1)}$**.** The Dynkin diagram is

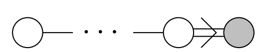


with r nodes. All roots are long—the algebra is *simply laced.* The group $SU(r+1)$ is given by $(r+1) \times (r+1)$ unitary matrices, $M^{-1} = M^\dagger$, with unit determinant, therefore

$$\mathfrak{su}(r+1) = \big\{ A \in \mathbb{C}^{(r+1)\times(r+1)} \,\big|\, A = A^\dagger, \ \operatorname{Tr} A = 0 \big\} \;.$$

The dimension is $r(r+2)$. The dual Coxeter number is $g^\vee = r+1$ (as $|\alpha|^2 = 2$ for all roots).

$\boldsymbol{B_{r\geq 2} \cong \mathfrak{so}(2r+1)}$**.** The Dynkin diagram is



with r nodes, representing $(r-1)$ long roots and 1 short root. The group $SO(N)$ is given by $N \times N$ orthogonal real matrices, $M^{-1} = M^\mathsf{T}$, with unit determinant. Writing $M = e^{tA}$ we find

$$\mathfrak{so}(N) = \big\{ A \in \mathbb{R}^{N\times N} \,\big|\, A^\mathsf{T} = -A \big\} \;.$$

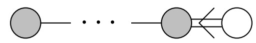
The dimension is $N(N-1)/2$. Specialized to the case $N = 2r+1$, the dimension is $r(2r+1)$.

One might be tempted to say that

$$A_1 \cong B_1 \qquad \big(\mathfrak{su}(2) \cong \mathfrak{so}(3)\big) \;.$$

This is true, but the normalizations are not the same because in B_1 the root is short while in A_1 it is long.

$\boldsymbol{C_{r\geq 1} \cong \mathfrak{sp}(2r)}$. The Dynkin diagram is



with r nodes, representing 1 long root and $(r-1)$ short roots. The group $USp(2r)$ is given by $2r \times 2r$ unitary and symplectic matrices, namely

$$M^{-1} = M^\dagger \;, \qquad M\Omega M^\mathsf{T} = \Omega \qquad \text{with} \qquad \Omega = \begin{pmatrix} 0 & \mathbb{1}_r \\ -\mathbb{1}_r & 0 \end{pmatrix} \;.$$

This group is compact. Writing $M = e^{itA}$, we find

$$\mathfrak{usp}(2r) = \left\{ A \in \mathbb{C}^{2r\times 2r} \,\middle|\, \Omega A^\mathsf{T} \Omega = A = A^\dagger \right\} \;.$$

Counting the number of independent components,[45] the dimension is $r(2r+1)$.

Alternatively, the group $Sp(2r,\mathbb{R})$ is given by $2r \times 2r$ real symplectic matrices. Writing $M = e^{tA}$ we find again the constraint $\Omega A^\mathsf{T}\Omega = A$, which however can be written as

$$\mathfrak{sp}(2r) = \left\{ A \in \mathbb{R}^{2r\times 2r} \,\middle|\, (A\Omega)^\mathsf{T} = A\Omega \right\} \;.$$

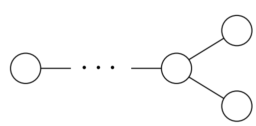
Counting symmetric matrices, the dimension is $r(2r+1)$ again. [In fact, $\mathfrak{usp}(2r)$ is the compact real form while $\mathfrak{sp}(2r)$ is a non-compact real form.]

Notice that

$$C_1 \cong A_1 \quad \big(\mathfrak{sp}(2) \cong \mathfrak{su}(2)\big) \;, \qquad\qquad C_2 \cong B_2 \quad \big(\mathfrak{sp}(4) \cong \mathfrak{so}(5)\big)$$

with the correct normalizations.

$\boldsymbol{D_{r\geq 2} \cong \mathfrak{so}(2r)}$. The Dynkin diagram is



with r nodes. All roots are long—the algebra is *simply laced*. Specializing the dimension of $\mathfrak{so}(N)$ to $N = 2r$, the dimension is $r(2r-1)$, while $g^\vee = 2r-2$.

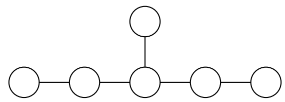
Notice that

$$D_2 \cong A_1 \times A_1 \quad \big(\mathfrak{so}(4) \cong \mathfrak{su}(2) \times \mathfrak{su}(2)\big) \;, \qquad\qquad D_3 \cong A_3 \quad \big(\mathfrak{so}(6) \cong \mathfrak{su}(4)\big)$$

with the correct normalizations.

[45] The equations $\Omega A^\mathsf{T}\Omega = A = A^\dagger$ imply that $A = \begin{pmatrix} A_1 & A_2 \\ A_2^* & -A_1^* \end{pmatrix}$ with A_1 Hermitian (r^2 components) and A_2 complex symmetric ($r(r+1)$ components).

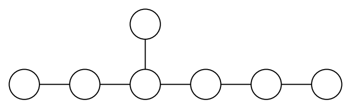
$\boldsymbol{E_6} \cong \mathfrak{e}_6$**.** The Dynkin diagram is



The algebra is simply-laced. It dimension is 78.

Exceptional Lie groups are not easy to describe. For reference, E_6 (the compact form) is the isometry group of a 32-dimensional Riemannian symmetric space known as the "bioctonionic projective plane".

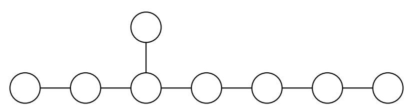
$\boldsymbol{E_7} \cong \mathfrak{e}_7$**.** The Dynkin diagram is



The algebra is simply-laced. It dimension is 133.

E_7 is the isometry group of a 64-dimensional Riemannian symmetric space known as the "quateroctonionic projective plane".

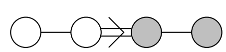
$\boldsymbol{E_8} \cong \mathfrak{e}_8$**.** The Dynkin diagram is



The algebra is simply-laced. It dimension is 248.

E_8 is the isometry group of a 128-dimensional Riemannian symmetric space known as the "octooctonionic projective plane".

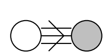
$\boldsymbol{F_4} \cong \mathfrak{f}_4$**.** The Dynkin diagram is



The dimension of the algebra is 52.

F_4 is the isometry group of a 16-dimensional Riemannian symmetric space known as the "octonionic projective plane".

$\boldsymbol{G_2} \cong \mathfrak{g}_2$**.**



The dimension of the algebra is 14.

G_2 is the automorphism group of the octonion algebra, or the little group of a Majorana spinor of $SO(7)$ (the Majorana spinor representation has real dimension 8).

8.10 Fundamental weights

Weights and roots live in the same r-dimensional vector space, since the roots are the weights of the adjoint representation. The weights could be expanded in the basis of simple roots, but the coefficients are (in general) not integers. The convenient basis $\{\omega_i\}$ is the one dual to the simple coroot basis:

$$(\omega_i, \alpha_j^\vee) = \delta_{ij} \ .$$

The ω_i are called the **fundamental weights**.

The expansion coefficients λ_i of a weight λ in the fundamental weight basis are called **Dynkin labels**:

$$\lambda = \sum_{i=1}^{r} \lambda_i \, \omega_i \qquad \Leftrightarrow \qquad \lambda_i = (\lambda, \alpha_i^\vee) \in \mathbb{Z} \ .$$

They are also the eigenvalues of the Cartan generators in the Chevalley basis:

$$\lambda_i = h^i(\lambda) \qquad\qquad \text{with} \qquad h^i = 2\, \frac{\alpha_i \cdot H}{|\alpha_i|^2} \ .$$

The fact that the Dynkin labels are integer (for finite-dimensional representations) follows from the $\mathfrak{su}(2)$ argument. We write

$$\lambda = (\lambda_1, \ldots, \lambda_r)$$

to indicate the Dynkin labels of a weight. Since we can write

$$A_{ij} = (\alpha_i, \alpha_j^\vee) \ ,$$

it follows that the entries in the rows of the Cartan matrix are the Dynkin labels of the simple roots:

$$\alpha_i = \sum_i A_{ij} \, \omega_j \ .$$

A weight of special importance, called the **Weyl vector** or *principal* vector, is the one whose Dynkin labels are all 1:

$$\rho = \sum_{i=1}^{r} \omega_i = (1, \ldots, 1) \ .$$

It has the following property: it is half the sum of all positive roots:

$$\rho = \frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha \ .$$

We will prove this formula later.

8.11 The Weyl group

Consider the projection of the adjoint representation to the $\mathfrak{su}(2)$ subalgebra associated to a root α.[46] Let m be the eigenvalue of the operator $J_3 \equiv \alpha \cdot H/|\alpha|^2$ on the state $|\beta\rangle$:

$$m\,|\beta\rangle = \frac{\alpha \cdot H}{|\alpha|^2}\,|\beta\rangle = \frac{(\alpha,\beta)}{|\alpha|^2}\,|\beta\rangle = \frac{1}{2}\,(\alpha^\vee,\beta)\,|\beta\rangle \;.$$

Then we can express

$$2m = (\alpha^\vee,\beta) \in \mathbb{Z} \;.$$

If $m \neq 0$, there must be another state in the same multiplet with J_3 eigenvalue $-m$, and this state must be $|\beta + \ell\alpha\rangle$ for some ℓ. Thus it must be

$$(\alpha^\vee, \beta + \ell\alpha) = (\alpha^\vee,\beta) + 2\ell = -(\alpha^\vee,\beta) \;,$$

which determines $\ell = -(\alpha^\vee,\beta)$. We conclude that if β is a root, then also $\beta - (\alpha^\vee,\beta)\,\alpha$ is a root.

Thus the linear operator

$$s_\alpha\,\beta \;\equiv\; \beta - (\alpha^\vee,\beta)\,\alpha \;,$$

which is a reflection with respect to the hyperplane perpendicular to α, maps roots to roots. The set of all such reflections along roots forms a group, called the **Weyl group** W of the algebra. The Weyl group is the symmetry group of the set of roots Δ.

The r elements s_i corresponding to simple roots α_i are called **simple Weyl reflections**:

$$s_i \;\equiv\; s_{\alpha_i} \;,$$

and every element $w \in W$ can be decomposed as

$$w = s_i s_j \cdots s_k \;.$$

Simple Weyl reflections have the property that they map positive roots to positive roots, with the exception of $\alpha_i \mapsto -\alpha_i$. In other words:

$$\alpha \in \Delta_+ \qquad \Rightarrow \qquad s_i\,\alpha \begin{cases} \in \Delta_+ & \text{if } \alpha \neq \alpha_i \\ = -\alpha_i & \text{if } \alpha = \alpha_i \;. \end{cases}$$

To prove it, expand $\alpha = \sum_{j=1}^r k_j\,\alpha_j$ with $k_j \geq 0$. Then $s_i\alpha = \sum_{j(\neq i)} k_j\,\alpha_j + \#\,\alpha_i$. If $\alpha \neq \alpha_i$ then some $k_{j(\neq i)} > 0$.[47] Since positive roots have non-negative coefficients in terms of simple roots, and negative roots have non-positive coefficients, we conclude that $s_i\alpha$ is positive whenever $\alpha \neq \alpha_i$.

[46] Here α can be any root, not necessarily a simple root. We define $\alpha^\vee \equiv 2\alpha/|\alpha|^2$.

[47] Here we use the fact, proved in Footnote 38, that the only multiple of a root α which is also a root, is $-\alpha$. Therefore, if α is a positive root and is not α_i, it must necessarily contain some other α_j in its expansion.

It is easy to see that the simple Weyl reflections satisfy

$$s_i^2 = \mathbb{1} \;, \qquad\qquad s_i s_j = s_j s_i \quad \text{if} \quad A_{ij} = 0 \;.$$

With a little bit of work, one can show that the full set of relations is

$$(s_i s_j)^{m_{ij}} = \mathbb{1} \qquad \text{where} \quad m_{ij} = \begin{cases} 1 & \text{if } i = j \\ 2 & \text{if } A_{ij}A_{ji} = 0 \\ 3 & \text{if } A_{ij}A_{ji} = 1 \\ 4 & \text{if } A_{ij}A_{ji} = 2 \\ 6 & \text{if } A_{ij}A_{ji} = 3 \;. \end{cases}$$

A group having such a presentation is called a *Coxeter group*.[48] On the simple roots, the simple reflections take the form

$$s_i \alpha_j = \alpha_j - A_{ji}\, \alpha_i \;.$$

Example 8.5. For $\mathfrak{su}(N)$, the Weyl group turns out to be the permutation group of N elements, $W = S_n$.

Indeed, we can interpret s_i, for $i = 1, \ldots, N-1$, as simple permutations of adjacent objects in a list of N objects $\{O_1, \ldots, O_N\}$, *e.g.* $s_1 : O_1 \leftrightarrow O_2$ and more generally

$$s_i : O_i \leftrightarrow O_{i+1} \;.$$

Each simple permutation squares to $\mathbb{1}$, non-adjacent permutations commute, while adjacent permutations satisfy

$$s_1 s_2 s_1 = s_2 s_1 s_2$$

when acting on $\{O_1, O_2, O_3\}$. This is equivalent to $(s_1 s_2)^3 = \mathbb{1}$.

We have shown that W maps Δ to itself. In fact, this is a way to generate the complete set of roots Δ from the simple roots: acting on them with the Weyl group,

$$\Delta = \big\{ w\alpha_1, \ldots, w\alpha_r \,\big|\, w \in W \big\} \;.$$

From this we see that any set $\{w\alpha_i\}$ with w fixed, provides an alternative, equally good set of simple roots. In fact, any possible set of simple roots that one obtains by choosing a different separation of $\mathfrak{h}^*$ is obtained by the action of the Weyl group on an initial set.

From the definition, it is immediate to check that the Weyl group preserves the scalar product on $\mathfrak{h}^*$:

$$\big(s_\alpha\, \lambda, s_\alpha\, \mu \big) = (\lambda, \mu) \;.$$

Thus, it is an orthogonal transformation of $\mathfrak{h}^*$.

[48] Therefore all Weyl groups of simple Lie algebras are Coxeter groups. The converse is not true: there are (an infinite number of) finite Coxeter groups that are not Weyl groups.

Proposition 8.6. *The Weyl vector takes the form*

$$\rho = \sum_{i=1}^{r} \omega_i = \frac{1}{2} \sum_{\alpha\in\Delta_+} \alpha \ .$$

Proof. Let us set $\rho = \sum_{\alpha>0} \alpha/2$. As shown above, the simple Weyl reflection s_i maps positive roots to positive roots, with the exception of $\alpha_i \mapsto -\alpha_i$. Therefore

$$s_i\rho = \tfrac{1}{2} \sum\nolimits_{\substack{\alpha>0\\ \alpha\neq\alpha_i}} \alpha - \tfrac{1}{2}\alpha_i = \rho - \alpha_i \ .$$

On the other hand s_i preserves the scalar product. Therefore

$$(s_i\rho, \alpha_i^\vee) = \begin{cases} (\rho - \alpha_i, \alpha_i^\vee) = (\rho, \alpha_i^\vee) - 2 \\ (\rho, s_i\alpha_i^\vee) = -(\rho, \alpha_i^\vee) \end{cases}$$

We conclude that $(\rho, \alpha_i^\vee) = 1$ and thus $\rho = \sum_{i=1}^{r} \omega_i$. □

8.12 Lattices and congruence classes

Given a basis $(\epsilon_1, \ldots, \epsilon_d)$ of $\mathbb{R}^d$, a *lattice* is the set of all points whose expansion in the basis has integer coefficients:

$$\mathbb{Z}\epsilon_1 + \cdots + \mathbb{Z}\epsilon_d \ .$$

In other words, it is the $\mathbb{Z}$-span of $\{\epsilon_i\}$.

There are three important r-dimensional lattices in a Lie algebra $\mathfrak{g}$: the weight lattice

$$P = \mathbb{Z}\omega_1 + \cdots + \mathbb{Z}\omega_r \ ,$$

the root lattice

$$Q = \mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_4 \ ,$$

and the coroot lattice

$$Q^\vee = \mathbb{Z}\alpha_1^\vee + \cdots + \mathbb{Z}\alpha_r^\vee \ .$$

The weight lattice P contains all possible weights of finite-dimensional representations of $\mathfrak{g}$. Given a finite-dimensional representation, the effect of the generator E^α on a weight is to shift the weight by α (unless it gives zero), thus the weights in a representation differ by elements of the root lattice Q. Obviously $Q \subseteq P$ (as roots are the weights of the adjoint representation).

For the algebras G_2, F_4 and E_8 it turns out that $Q = P$. In all other cases, Q is a proper subset of P and P/Q is a finite group. Its order, $|P/Q|$, is equal to the determinant of the Cartan matrix (and P/Q is equal to the center of the simply-connected compact group G with algebra $\mathfrak{g}$).

The distinct elements of P/Q define **congruence classes**, and each finite-dimensional representation is in a congruence class. Then P/Q is a *conserved charge* associated to the representations, in the sense that the product of representations has a charge which is the sum of the individual charges.

Example 8.7. For $\mathfrak{su}(2)$ there two classes, given by $\lambda_1 \mod 2$: integer or half-integer spin.
Example 8.8. For $\mathfrak{su}(3)$ there are three classes, identified by the triality: $\lambda_1 + 2\lambda_2 \mod 3$.
Example 8.9. For $\mathfrak{su}(N)$ there are N classes, and the N-ality is characterized by

$$\lambda_1 + 2\lambda_2 + \cdots + (N-1)\lambda_{N-1} \mod N .$$

The congruence classes in the various cases are the following:

$\mathfrak{g}$	$\mathfrak{su}(N)$	$\mathfrak{so}(2r+1)$	$\mathfrak{sp}(2r)$	$\mathfrak{so}(2r)$	E_6	E_7	E_8	F_4	G_2
P/Q	$\mathbb{Z}_N$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2 \times \mathbb{Z}_2$ if $2r = 0 \mod 4$ $\mathbb{Z}_4$ if $2r = 2 \mod 4$	$\mathbb{Z}_3$	$\mathbb{Z}_2$	1	1	1

9 Low-rank examples

Let us discuss the algebras of rank 1 and 2.

9.1 $A_1 \cong \mathfrak{su}(2)$

There is only one simple algebra of rank 1: $A_1 \cong \mathfrak{su}(2)$. The Cartan matrix is

$$A = (2) ,$$

therefore there is only one simple root, $\alpha_1 = \theta$, obviously equal to the highest root, and related to the fundamental weight ω_1 by

$$\alpha_1 = 2\omega_1 .$$

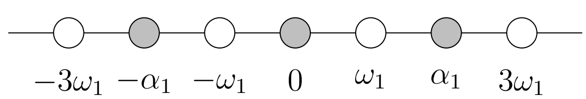
The Weyl group is generated by the simple reflection s_1, that acts on a weight $\lambda = \lambda_1\omega_1$ as

$$s_1(\lambda_1\omega_1) = \lambda_1\omega_1 - (\alpha_1^\vee, \lambda_1\omega_1)\alpha_1 = \lambda_1\omega_1 - \lambda_1\alpha_1 = -\lambda_1\omega_1 ,$$

which is a reflection with respect to the origin. Therefore $W = \{1, s_1\} = \mathbb{Z}_2$ and the roots are

$$\Delta = \{\alpha_1, -\alpha_1\} .$$

Indeed $\dim \mathfrak{su}(2) = 3$. The weight lattice is



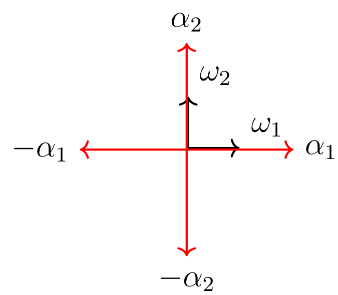
The shaded nodes form the root lattice. Then the group of congruence classes is $P/Q = \mathbb{Z}_2$, corresponding to integer or half-integer spin. The non-trivial class is generated by ω_1.

9.2 $D_2 \cong \mathfrak{so}(4) \cong A_1 \times A_1 \cong \mathfrak{su}(2) \times \mathfrak{su}(2)$

This case, which is degenerate, corresponds to Cartan matrix

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} ,$$

which is decomposable and so the algebra is semi-simple, rather than simple. The root system is



The group of congruence classes is $P/Q = \mathbb{Z}_2 \times \mathbb{Z}_2$, and the non-trivial elements are represented by ω_1, ω_2 and $\omega_1 + \omega_2$.

9.3 $A_2 \cong \mathfrak{su}(3)$

The Cartan matrix is

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

and all roots have the same length. The simple roots are α_1, α_2, with an angle of 120°, and are related to the fundamental weights by

$$\alpha_1 = \alpha_1^\vee = 2\omega_1 - \omega_2 = (2, -1)$$
$$\alpha_2 = \alpha_2^\vee = -\omega_1 + 2\omega_2 = (-1, 2) \ .$$

The Weyl group is

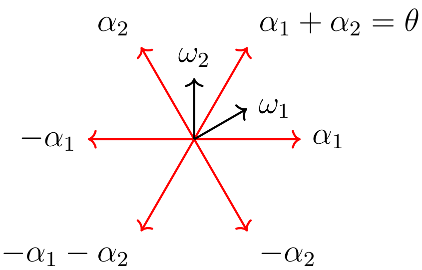
$$W = \{1, s_1, s_2, s_1 s_2, s_2 s_1, s_1 s_2 s_1\} = S_3 \ ,$$

that follows from the relation $(s_1 s_2)^3 = 1$.

The action of the Weyl group on the simple roots gives all possible roots, and one finds

$$\Delta = \{\alpha_1, \alpha_2, \alpha_1 + \alpha_2, -\alpha_1, -\alpha_2, -\alpha_1 - \alpha_2\} \ .$$

Indeed $\dim \mathfrak{su}(3) = 8$. The highest root is $\theta = \alpha_1 + \alpha_2 = (1, 1)$. The root system is



The congruence group is $P/Q = \mathbb{Z}_3$: each fundamental weight $\omega_{1,2}$ is in a different non-trivial class (while the roots are in the trivial class). This is called *triality*.

9.4 $B_2 \cong \mathfrak{so}(5) \cong C_2 \cong \mathfrak{sp}(4)$

The Cartan matrix of $B_2 \cong \mathfrak{so}(5)$ is

$$A = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} ,$$

and the algebra is not simply-laced:

$$|\alpha_1|^2 = 2 , \qquad\qquad |\alpha_2|^2 = 1 ,$$

and the angle between the two simple roots is $135°$. We have

$$\begin{aligned} \alpha_1 &= \alpha_1^\vee = 2\omega_1 - 2\omega_2 = (2, -2) \\ \alpha_2 &= \frac{1}{2}\alpha_2^\vee = -\omega_1 + 2\omega_2 = (-1, 2) . \end{aligned}$$

The structure of the Weyl group follows from $(s_1 s_2)^4 = \mathbb{1}$, so $W = D_4$ the dihedral group of 8 elements.[49]

The roots are

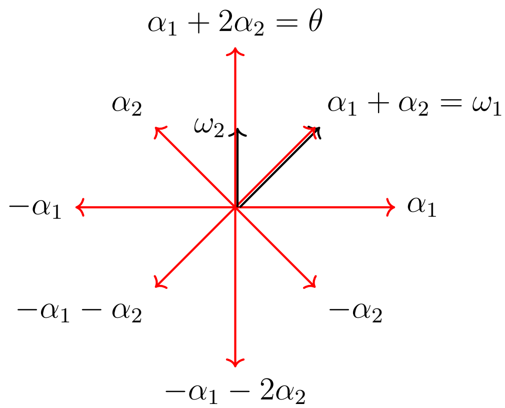
$$\Delta = \left\{ \alpha_1, \alpha_2, \alpha_1 + \alpha_2, \alpha_1 + 2\alpha_2, -\alpha_1, -\alpha_2, -\alpha_1 - \alpha_2, -\alpha_1 - 2\alpha_2 \right\} ,$$

[49]The dihedral group D_n is defined as

$$D_n = \left\{ r, s \,\middle|\, r^n = s^2 = (sr)^2 = \mathbb{1} \right\} = \mathbb{Z}_n \rtimes \mathbb{Z}_2$$

and has $2n$ elements. Setting $s_1 = s$, $s_2 = sr$, $s_1 s_2 = r$ it equals the Weyl group $s_1^2 = s_2^2 = (s_1 s_2)^n = \mathbb{1}$.

indeed $\dim \mathfrak{so}(5) = 10$, therefore the root system is



The highest root is $\theta = \alpha_1 + 2\alpha_2 = (0, 2)$.

The congruence classes are $P/Q = \mathbb{Z}_2$, corresponding to spinors and tensors. The non-trivial class is represented by ω_2.

If we exchange $\alpha_1 \leftrightarrow \alpha_2$ we obtain $C_2 \cong \mathfrak{sp}(4)$ with Cartan matrix

$$A = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} .$$

The root system is the same, but rotated by 45°. The two congruence classes correspond to tensors with odd or even rank.

9.5 G_2

The Cartan matrix of G_2 is

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} .$$

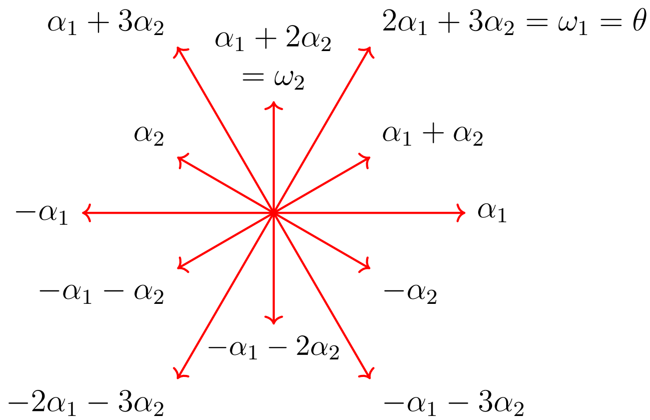
The algebra is not simply laced: $|\alpha_1|^2 = 2$, $|\alpha_2|^2 = \frac{2}{3}$ and the angle between the two is 150°. We have

$$\begin{aligned} \alpha_1 &= \alpha_1^\vee = 2\omega_1 - 3\omega_2 = (2, -3) \\ \alpha_2 &= \frac{1}{3}\alpha_2^\vee = -\omega_1 + 2\omega_2 = (-1, 2) . \end{aligned}$$

The Weyl group is given by $(s_1 s_2)^6 = 1$, and is the dihedral group D_6 with 12 elements. The positive roots are

$$\Delta = \left\{\alpha_1, \alpha_2, \alpha_1 + \alpha_2, \alpha_1 + 2\alpha_2, \alpha_1 + 3\alpha_2, 2\alpha_1 + 3\alpha_2\right\} ,$$

indeed $\dim G_2 = 14$, therefore the root system is



The highest root is $\theta = 2\alpha_1 + 3\alpha_2 = (1, 0)$. There are no congruence classes as $P/Q = 1$.

9.6 Method to construct the root system

There is a simple algorithmic method to construct the root system of any simple Lie algebra. Recall that

$$(\lambda, \alpha^\vee) = 2\frac{(\lambda, \alpha)}{|\alpha|^2} = q - p\Big|_{\text{of } \lambda \text{ along } \alpha}$$

contains information about how many times we can increase (p) or decrease (q) the weight λ by α. In particular if we shift λ by the simple roots α_i, the information is in the Dynkin labels:

$$\lambda_i = q - p\Big|_{\text{along } \alpha_i} .$$

Next, we use that the difference of two simple roots is never a root, $\alpha_i - \alpha_j \notin \Delta$. Therefore, starting with the simple roots, we know that the weight cannot be lowered and thus $q = 0$. Calling $\lambda_j^{(i)} = A_{ij}$ the j-th Dynkin label of α_i, we have

$$\lambda_j^{(i)} = -p ,$$

i.e. a negative Dynkin label tells us how many times we can shift a simple root α_i by another simple root α_j.[50]

Starting from the simple roots at the bottom, we build up all positive roots by adding simple roots; we keep track of what p and q are along the way, and every time we encounter a negative Dynkin label for a root that has $q = 0$, we infer how many times we can further add another simple root. The process stops when all Dynkin labels are positive, and that is the *highest root* θ.

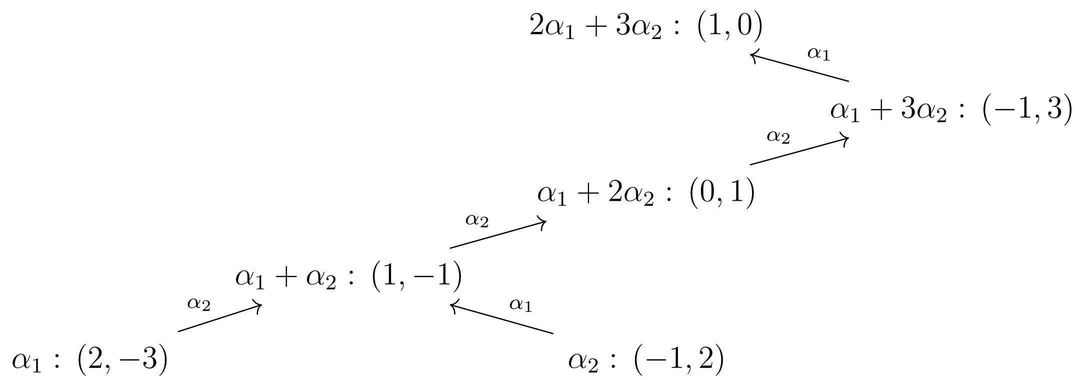
Since the Dynkin labels of the simple roots are the rows of the Cartan matrix, the process is completely specified by the Cartan matrix A.

[50] Notice that $(\alpha_i, \alpha_i^\vee) = 2$, because the only multiple of α_i that is a root is $-\alpha_i = \alpha_i - 2\alpha_i$, and of course there is the vanishing weight.

Example 9.1. Let us work out the root system of G_2. The Cartan matrix is

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} .$$

Starting from the bottom, we obtain:



This reproduces the positive roots described before. In particular the positive roots are 6, thus the dimension of the algebra is

$$\dim G_2 = 6 + 6 + 2 = 14 .$$

10 Highest-weight representations

Any *finite-dimensional irreducible representation* has a unique highest-weight state $|\lambda\rangle$, which is completely specified by its Dynkin labels λ_i.

Among all the weights in the representation, the highest weight λ is such that $\lambda + \alpha$ is not a weight for any $\alpha > 0$. That is

$$E^\alpha \, |\lambda\rangle = 0 \qquad\qquad \forall\, \alpha > 0 \;.$$

From the $\mathfrak{su}(2)$ argument,

$$2\, \frac{(\lambda, \alpha_i)}{|\alpha_i|^2} = q \geq 0 \qquad\qquad \text{because } p = 0$$

and thus the Dynkin labels are non-negative.

Moreover, to each weight λ with non-negative Dynkin labels, called **dominant weight**, corresponds a unique finite-dimensional irreducible representation V_λ (sometimes we indicate the representation as λ).

The highest root θ is the highest weight of the *adjoint representation*.

Starting from the highest-weight state $|\lambda\rangle$, all the states in the representation V_λ can be obtained by the action of the lowering operators:

$$E^{-\beta} E^{-\gamma} \dots E^{-\eta} \, |\lambda\rangle \qquad \text{for} \qquad \beta, \gamma, \dots, \eta \in \Delta_+ \;.$$

Let us call the *weight system* Ω_λ the set of weights in a representation. Any weight $\mu \in \Omega_\lambda$ is such that $\lambda - \mu \in Q$, the root lattice, thus all weights in a given representation lie in the same congruence class of P/Q.

To construct all weights μ in Ω_λ we use the $\mathfrak{su}(2)$ subalgebra:

$$(\mu, \alpha_i^\vee) = \mu_i = q_i - p_i \;, \qquad\qquad p_i, q_i \in \mathbb{Z}_+ \;.$$

Since $\mu = \lambda - \sum n_i \alpha_i$ for some $n_i \in \mathbb{Z}_+$, we can call $\sum n_i \in \mathbb{Z}$ the "level" of the weight μ and proceed level-by-level. At each step we know the value of p_i, and as long as

$$q_i = \mu_i + p_i > 0 \;,$$

we can remove α_i q_i times. When removing α_i, we reduce the Dynkin labels of μ by the Dynkin labels of α_i, which are given in the i-th row of the Cartan matrix:

$$(\alpha_i, \alpha_j^\vee) = A_{ij} \qquad \Rightarrow \qquad \alpha_i = (A_{i1}, \dots, A_{ir}) \;.$$

This process produces the full set Ω_λ.

Example 10.1. **Adjoint representation of $\mathfrak{su}(3)$.** The Cartan matrix of $\mathfrak{su}(3)$ is

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

and the adjoint representation has highest weight $\lambda = (1, 1)$. We construct the weight system Ω_λ:

$$\begin{array}{ccccc}
 & & (1,1) & & \\
 & \swarrow^{-\alpha_1} & & \searrow^{-\alpha_2} & \\
(-1,2) & & & & (2,-1) \\
 & \searrow^{-\alpha_2} & & \swarrow^{-\alpha_1} & \\
 & & (0,0) & & \\
 & \swarrow^{-\alpha_1} & & \searrow^{-\alpha_2} & \\
(-2,1) & & & & (1,-2) \\
 & \searrow^{-\alpha_2} & & \swarrow^{-\alpha_1} & \\
 & & (-1,-1) & &
\end{array}$$

Indeed, these are the roots of $\mathfrak{su}(3)$.

Example 10.2. **Fundamental representation of $\mathfrak{g}_2$.** The Cartan matrix is

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$$

and the fundamental representation has highest weight $(0, 1)$:

$$\begin{array}{ccccc}
(0,1) & & & & \\
 & \searrow^{-\alpha_2} & & & \\
 & & (1,-1) & & \\
 & \swarrow^{-\alpha_1} & & & \\
(-1,2) & & & & \\
 & \searrow^{-\alpha_2} & & & \\
 & & (0,0) & & \\
 & & & \searrow^{-\alpha_2} & \\
 & & & & (1,-2) \\
 & & & \swarrow^{-\alpha_1} & \\
 & & (-1,1) & & \\
 & & & \searrow^{-\alpha_2} & \\
 & & & & (0,-1)
\end{array}$$

This representation has dimension 7.

¶ *Exercise* 25. Compute the weight system of the fundamental representation of $\mathfrak{sp}(4)$, whose highest weight is $(1, 0)$, and verify that there are 4 weights (in fact it has dimension 4). Then consider $\mathfrak{so}(5)$ and its representation with highest weight $(0, 1)$, which is the spinorial representation.

The procedure we described does not keep track of multiplicities.

Notice that the highest weight has no multiplicity, and multiplicities can only arise when two or more arrows go into the same weight (or if an arrow starts from a weight with multiplicity).

To compute multiplicities, one can use **Freudenthal's recursion formula** (that we will prove later):

$$\text{mult}_\lambda(\mu) = \frac{\sum_{\alpha>0}\sum_{k=1}^\infty 2\,\text{mult}_\lambda(\mu+k\alpha)\;(\mu+k\alpha,\alpha)}{|\lambda+\rho|^2-|\mu+\rho|^2} \;.$$

The denominator can also be written as $(\lambda+\mu+2\rho,\lambda-\mu) = (\lambda,\lambda+2\rho)-(\mu,\mu+2\rho)$.

The formula gives the multiplicity of μ in terms of the multiplicities of the weights above it.

Example 10.3. Let us compute the multiplicity of $(0,0)$ in the adjoint representation of $\mathfrak{su}(3)$, assuming that all roots have multiplicity 1. First we use

$$\lambda=\theta=\alpha_1+\alpha_2\;,\quad \rho=\alpha_1+\alpha_2\;,\quad \mu=(0,0)\;,\quad |\lambda+\rho|^2 = \big|2(\alpha_1+\alpha_2)\big|^2 = 8\;,\quad |\mu+\rho|^2=2\;.$$

Then we see that there are three weights above $(0,0)$, and k can only be 1:

$$(\mu+k\alpha,\alpha) = (\alpha,\alpha) = 2\;.$$

We thus have

$$\text{mult}_\theta(0,0) = \frac{2(2+2+2)}{8-2} = 2\;.$$

This is correct, as the Cartan subalgebra of $\mathfrak{su}(3)$ has dimension 2.

¶ *Exercise* 26. Compute the weight system of the adjoint representation of $\mathfrak{g}_2$, whose highest weight is $(1,0)$. Verify that it has dimension 14, computing the multiplicity of the weight $(0,0)$, and that it agrees with what we described before.

Theorem 10.4. *Finite-dimensional irreducible (highest-weight) representations V_λ of a complex semi-simple Lie algebra are always* **unitary***, in the following sense. Using $(H^i)^\dagger = H^i$ and $(E^\alpha)^\dagger = E^{-\alpha}$, the norm of any state $|\lambda'\rangle$ in V_λ, computed using the commutation relations, is positive definite:*

$$|\lambda'\rangle = E^{-\beta}\dots E^{-\gamma}|\lambda\rangle \qquad\Rightarrow\qquad \langle\lambda'|\lambda'\rangle = \langle\lambda|E^\gamma\dots E^\beta E^{-\beta}\dots E^{-\gamma}|\lambda\rangle > 0$$

with $\beta,\dots,\gamma\in\Delta_+$ and taking $\langle\lambda|\lambda\rangle>0$.

This means that all finite-dimensional representations of the compact *real form are* **unitary***: the algebra can be represented by* hermitian *or anti-hermitian operators (while the complex algebra is its complexifications, and the other real forms have non-unitary finite-dimensional representations).*

10.1 Conjugate representations

Given an irreducible representation V_λ with highest weight λ, it contains a "lowest-weight state" $\widehat{\lambda} < 0$. The representation V_{λ^*} with highest weight

$$\lambda^* \equiv -\widehat{\lambda}$$

is called the **conjugate representation** of V_λ. If $\widehat{\lambda} = -\lambda$, *i.e.* $\lambda^* = \lambda$, then the representation is called **self-conjugate**.

The conjugate representation V_{λ^*} is obtained by turning "upside-down" the representation V_λ,[51] and so it has the same dimension.

In the case of $\mathfrak{su}(N)$, the conjugate representation is obtained by reversing the order of the Dynkin labels:

$$\lambda = (\lambda_1, \dots, \lambda_{N-1}) \qquad \Rightarrow \qquad \lambda^* = (\lambda_{N-1}, \dots, \lambda_1) \; .$$

Notice that this is a symmetry of the Dynkin diagram. The two representations are in opposite congruence classes. Since the Dynkin diagrams of $\mathfrak{so}(2r+1)$, $\mathfrak{sp}(2r)$, G_2, F_4, E_7 and E_8 have no symmetry, all their representations are self-conjugate.

10.2 Quadratic Casimir operator

The **Casimir operators** are operators one can construct by taking products of algebra generators, such that they commute with all elements of the algebra. The Casimir operators are not part of the algebra, but rather of the *universal enveloping algebra*, since we cannot multiply elements of the Lie algebra.[52] On the other hand, they are naturally constructed in any representation.

Since Casimir operators commute with all elements of $\mathfrak{g}$, in an irreducible representation they are proportional to the identity. The eigenvalue carries the interesting information.

The **quadratic Casimir operator** $\mathcal{Q}$ is given by

$$\mathcal{Q} = \sum_{a,b} k_{ab} J^a J^b \; , \qquad\qquad \text{where} \qquad k^{ab} = k(J^a, J^b) \; .$$

¶ *Exercise* 27. Verify that it commutes with all elements of $\mathfrak{g}$.

[51] In fact $\widehat{\lambda} = w_0\lambda$, where w_0 is the longest element of the Weyl group that maps the positive chamber to the negative chamber. It follows that $\lambda^* = -w_0\lambda$, and all weights of the conjugate representation are obtained from those of V_λ by the action of $-w_0$.

In fact w_0 is the only element of the Weyl group that maps Δ_+ to Δ_-, and thus $-w_0$ maps simple roots to simple roots. Since W preserves the scalar product of roots, conjugation $-w_0$ must implement a symmetry of the Dynkin diagram. However, notice that the Dynkin diagram could have a symmetry which is not realized by $-w_0$ (this is the case of D_{2+2k}).

[52] The universal enveloping algebra is the set of all formal power series in elements of $\mathfrak{g}$.

In the Cartan-Weyl basis, it looks like[53]

$$\mathcal{Q} = \sum_{i=1}^{r} H^i H^i + \sum_{\alpha>0} \frac{|\alpha|^2}{2} \left(E^\alpha E^{-\alpha} + E^{-\alpha} E^\alpha \right) .$$

Since $\mathcal{Q}$ has the same eigenvalue on all states of an irreducible representation, let us evaluate it on the highest weight state. First

$$\sum_i H^i H^i \left|\lambda\right\rangle = \sum_i \lambda^i \lambda^i \left|\lambda\right\rangle = (\lambda,\lambda) \left|\lambda\right\rangle .$$

Second

$$E^\alpha E^{-\alpha} \left|\lambda\right\rangle = [E^\alpha, E^{-\alpha}] \left|\lambda\right\rangle = 2\frac{\alpha\cdot H}{|\alpha|^2} \left|\lambda\right\rangle = 2\frac{(\alpha,\lambda)}{|\alpha|^2} \left|\lambda\right\rangle ,$$

where we used $E^\alpha|\lambda\rangle = 0$. Therefore

$$\mathcal{Q} \left|\lambda\right\rangle = \left[(\lambda,\lambda) + \sum_{\alpha>0} (\lambda,\alpha) \right] \left|\lambda\right\rangle = (\lambda, \lambda+2\rho) \left|\lambda\right\rangle = \Big(|\lambda+\rho|^2 - |\rho|^2 \Big) \left|\lambda\right\rangle$$

where ρ is the Weyl vector.

In the special case of the *adjoint representation*, we compute

$$(\theta,\theta+2\rho) = 2 + 2(\theta,\rho) = 2 + 2\left(\textstyle\sum_i a_i^\vee \alpha_i^\vee ,\, \sum_j \omega_j \right) = 2 + 2\sum_i a_i^\vee = 2g^\vee ,$$

where we used that $|\theta|^2 = 2$. Thus, the quadratic Casimir of the adjoint representation is twice the dual Coxeter number.

Remark. We can use the Casimir operator to fix the normalization of the Killing form. Recall that

$$k(J^a, J^b) = \frac{1}{2\tilde{g}} \operatorname{Tr}\left(\operatorname{ad}(J^a)\operatorname{ad}(J^b) \right) \equiv k^{ab} .$$

As we saw before, imposing $k(H^i,H^j) = \delta^{ij}$ fixes $\tilde{g} = \sum_\alpha |\alpha|^2/2r$. Now notice

$$\dim\mathfrak{g} = k_{ab}k^{ab} = \frac{1}{2\tilde{g}} k_{ab} \operatorname{Tr}\left(\operatorname{ad}(J^a)\, ad(J^b) \right) = \frac{1}{2\tilde{g}} \operatorname{Tr}_{\text{adj}} \mathcal{Q} = \frac{1}{2\tilde{g}} \dim\mathfrak{g}\; (\theta,\theta+2\rho) .$$

We conclude

$$2\tilde{g} = (\theta,\theta+2\rho) = |\theta|^2 g^\vee ,$$

which is independent of normalization. The choice $|\theta|^2 = 2$ leads to $\tilde{g} = g^\vee$.

[53]Recall that $k(H^i,H^j) = \delta^{ij}$, $k(H^i,E^\alpha) = 0$, $k(E^\alpha,E^\beta) = 0$ for $\beta \neq -\alpha$ while $k(E^\alpha,E^{-\alpha}) = 2/|\alpha|^2$. Therefore the matrix of the Killing form and its inverse are

$$k^{ab} = \operatorname{diag}\left(\mathbb{1}_r,\ \begin{bmatrix} 0 & 2/|\alpha|^2 \\ 2/|\alpha|^2 & 0 \end{bmatrix}, \dots \right), \qquad k_{ab} = \operatorname{diag}\left(\mathbb{1}_r,\ \begin{bmatrix} 0 & |\alpha|^2/2 \\ |\alpha|^2/2 & 0 \end{bmatrix}, \dots, \right) .$$

We have described the *quadratic* Casimir operator, that exists in all simple Lie algebras. In $\mathfrak{su}(2)$, that is the only Casimir invariant.

In general, however, there are r independent Casimir invariants of various degrees. The degrees minus 1 are called the **exponents** of the algebra.

Note. The exponents can be computed in the following way. First construct a "symmetrized Cartan matrix"

$$\widehat{A}_{ij} = 2\frac{(\alpha_i, \alpha_j)}{\sqrt{|\alpha_i|^2|\alpha_j|^2}} \ .$$

Then the exponents m_i are

$$m_i = \frac{2g}{\pi} \arcsin \sqrt{\text{Eigenvalues}\big(\widehat{A}_{ij}\big)/4} \ ,$$

where g is the *Coxeter number.*

Derivation of Freudenthal's formula. Consider an irreducible representation V_λ with highest weight λ. Let W_μ be the eigenspace related to the weight μ. The trace of the Casimir operator $\mathcal{Q}$ in W_μ is

$$\text{Tr}_{W_\mu} \mathcal{Q} = \text{mult}(\mu)\ (\lambda, \lambda + 2\rho) \ .$$

On the other hand, we can use $\mathcal{Q} = \sum_i H^i H^i + \sum_{\alpha>0} \frac{|\alpha|^2}{2}(E^\alpha E^{-\alpha} + E^{-\alpha}E^\alpha)$. The trace of the first term is

$$\text{Tr}_{W_\mu} \sum_i H^i H^i = \text{mult}(\mu)\ (\mu, \mu) \ .$$

For the second term, we consider the $\mathfrak{su}(2)$ subalgebra generated by $J_3 \equiv \frac{\alpha \cdot H}{|\alpha|^2}$, $J_\pm \equiv E^{\pm\alpha}$ for a positive root α. Each of the states in W_μ will be in an irreducible representation of this $\mathfrak{su}(2)$. Let $|\mu_j\rangle$ be a state in a representation of spin j. From our analysis of the $\mathfrak{su}(2)$ algebra we get

$$\big(J_+J_- + J_-J_+\big)|j-k\rangle = 2(j + 2jk - k^2)\,|j-k\rangle = 2(j^2 + j - m^2)\,|m\rangle \ ,$$

where we expressed as $m = j - k$ the eigenvalue of J_3. Since such an eigenvalue is, in our case,

$$m\,|\mu_j\rangle = \frac{\alpha \cdot H}{|\alpha|^2}\,|\mu_j\rangle = \frac{(\alpha, \mu)}{|\alpha|^2}\,|\mu_j\rangle \ ,$$

we conclude that

$$\big(E^\alpha E^{-\alpha} + E^{-\alpha}E^\alpha\big)\ |\mu_j\rangle = 2\bigg(j(j+1) - \frac{(\alpha, \mu)^2}{|\alpha|^4}\bigg)\ |\mu_j\rangle \ .$$

Alternatively, we can use the Casimir of $\mathfrak{su}(2)$

$$\mathcal{Q}^{\mathfrak{su}(2)} = 2J_3J_3 + J_+J_- + J_-J_+ = 2j(j+1) \ .$$

Within the full algebra it takes the form

$$\mathcal{Q}^{\mathfrak{su}(2)}\,|\mu_j\rangle = \bigg(2\,\frac{\alpha\cdot H\,\alpha\cdot H}{|\alpha|^4} + E^{\alpha}E^{-\alpha} + E^{-\alpha}E^{\alpha}\bigg)\,|\mu_j\rangle = 2j(j+1)\,|\mu_j\rangle\;.$$

We reach the same conclusion.

The state $|\mu_j\rangle$ is in a spin-j representation, so let $\mu + k\alpha$ be the highest weight for some $k \in \mathbb{Z}_{\geq 0}$, and let $|\mu + k\alpha|_j\rangle$ be its highest-weight state. Since

$$j\,|\mu + k\alpha|_j\rangle = \frac{\alpha\cdot H}{|\alpha|^2}\,|\mu + k\alpha|_j\rangle = \frac{(\alpha, \mu + k\alpha)}{|\alpha|^2}\,|\mu + k\alpha|_j\rangle$$

we get an expression for j. Substituting into the previous expression we obtain

$$\big(E^{\alpha}E^{-\alpha} + E^{-\alpha}E^{\alpha}\big)\,|\mu_j\rangle = 2\bigg(k(k+1) + (2k+1)\frac{(\alpha,\mu)}{|\alpha|^2}\bigg)\,|\mu_j\rangle\;.$$

There might be many copies of the spin-j representation that show up in W_μ: it is clear that the number of representations whose highest weight is $\mu + k\alpha$ is equal to the dimension of $W_{\mu+k\alpha}$ minus the dimension of $W_{\mu+(k+1)\alpha}$. Therefore

$$\begin{aligned}
&\operatorname{Tr}_{W_\mu} \sum_{\alpha>0} \frac{|\alpha|^2}{2}\big(E^{\alpha}E^{-\alpha} + E^{-\alpha}E^{\alpha}\big) = \\
&\quad= \sum_{\alpha>0}\sum_{k\geq 0}\Big(\operatorname{mult}(\mu + k\alpha) - \operatorname{mult}\big(\mu + (k+1)\alpha\big)\Big)\Big(k(k+1)|\alpha|^2 + (2k+1)(\alpha,\mu)\Big) \\
&\quad= \sum_{\alpha>0}\bigg[\operatorname{mult}(\mu)\;(\alpha,\mu) + \sum_{k\geq 1} 2\operatorname{mult}(\mu+k\alpha)\;(\alpha,\mu+k\alpha)\bigg]\;.
\end{aligned}$$

We conclude that

$$\begin{aligned}
\operatorname{Tr}_{W_\mu}\mathcal{Q} &= \operatorname{mult}(\mu)\;(\lambda,\lambda+2\rho) \\
&= \operatorname{mult}(\mu)\;(\mu,\mu+2\rho) + \sum_{\alpha>0}\sum_{k\geq 1} 2\operatorname{mult}(\mu+k\alpha)\;(\alpha,\mu+k\alpha)\;.
\end{aligned}$$

This leads to the multiplicity formula

$$\operatorname{mult}_\lambda(\mu) = \frac{\sum_{\alpha>0}\sum_{k\geq 1} 2\operatorname{mult}_\lambda(\mu+k\alpha)\;(\alpha,\mu+k\alpha)}{(\lambda+\mu+2\rho,\lambda-\mu)}\;.$$

The denominator can also be written as $(\lambda,\lambda+2\rho) - (\mu,\mu+2\rho) = |\lambda+\rho|^2 - |\mu+\rho|^2$.

10.3 Dynkin index

The invariant bilinear forms computed as traces in different representations are all proportional to the Killing form, and the relative normalizations are called the **Dynkin index** x_λ of a representation λ. We define x_λ through

$$\mathrm{Tr}_\lambda\left(\rho(J^a)\,\rho(J^b)\right) = \frac{x_\lambda}{g^\vee}\,\mathrm{Tr}_\theta\left(\mathrm{ad}(J^a)\,\mathrm{ad}(J^b)\right) = |\theta|^2\,x_\lambda\,k^{ab}\ ,$$

where ρ is the representation λ and k^{ab} is the Killing form (we usually take $|\theta|^2 = 2$). By construction, the Dynkin index of the adjoint representation is equal to the dual Coxeter number:

$$x_\theta = g^\vee\ .$$

In all other cases, the index is computed by multiplying both sides by k_{ab}. On the LHS we get $\mathrm{Tr}_\lambda\,\mathcal{Q}$, while on the RHS we get $|\theta|^2 x_\lambda \dim\mathfrak{g}$. We thus find

$$x_\lambda = \frac{(\lambda, \lambda+2\rho)\ \dim V_\lambda}{2\dim\mathfrak{g}}\ ,$$

using $|\theta|^2 = 2$.

10.4 Tensor products of representations

Given two irreducible representations V_λ, V_μ, we can construct the *tensor product* representation

$$V_\lambda \otimes V_\mu\ .$$

In terms of the two homomorphisms ρ_λ and ρ_μ, we define

$$\rho_{\lambda\otimes\mu} = \rho_\lambda \otimes \mathbb{1} + \mathbb{1} \otimes \rho_\mu\ ,$$

which is a homomorphism.[54] This implies that the weights of $V_\lambda \otimes V_\mu$ are sums of pairs of weights of V_λ and V_μ.

However, in general, the representation $V_\lambda \otimes V_\mu$ is not irreducible, rather it is a sum—with multiplicities—of irreducible representations:

$$V_\lambda \otimes V_\mu = \bigoplus_{\nu\in P_+} \mathcal{N}_{\lambda\mu}{}^{\nu}\ V_\nu\ ,$$

where P_+ is the set of dominant weights while the integers $\mathcal{N}_{\lambda\mu}{}^{\nu}$ are the multiplicities. This is called a **Clebsh-Gordan decomposition**. The sum of highest weights $\lambda+\mu$ gives a highest weight in $V_\lambda \otimes V_\mu$, thus the representation $V_{\lambda+\mu}$ appears once in the decomposition:

$$\mathcal{N}_{\lambda\mu}{}^{\lambda+\mu} = 1\ .$$

[54] One could try to define $\rho_{\lambda\otimes\mu} = \rho_\lambda \otimes \rho_\mu$, but this would not be a linear map. For instance, one would find $\rho_{\lambda\otimes\mu}(tX) = \rho_\lambda(tX) \otimes \rho_\mu(tX) = t^2\rho_\lambda(X) \otimes \rho_\mu(X)$ which is not a homomorphism.

After removing the weights of $V_{\lambda+\mu}$, we are left with a number of other highest weights, representing other terms in the decomposition. Proceeding this way, in principle, we can work out the full decomposition.

The tensor-product coefficients should be compatible with the dimensions:

$$\dim V_\lambda \cdot \dim V_\mu = \sum_{\nu \in P_+} \mathcal{N}_{\lambda\mu}{}^{\nu} \, \dim V_\nu \ .$$

The formal set of representations form an algebra, called a **fusion algebra**. Let us call 0 the one-dimensional trivial representation, and μ^* the conjugate representation to μ. Then fusion with the trivial representation does not do anything:

$$\mathcal{N}_{\lambda 0}{}^{\nu} = \delta^\nu_\lambda \ .$$

In the tensor product of a representation λ and its conjugate λ^* there is always one and only one singlet:[55]

$$\mathcal{N}_{\lambda\lambda^*}{}^{0} = 1 \ .$$

The result is more general: the product $V_\lambda \otimes V_\mu$ contains one and only one singlet if $\mu = \lambda^*$, and no singlet if $\mu \neq \lambda^*$. In fact the singlet is a linear map $V_\lambda \to V_{\mu^*}$ which commutes with the algebra action, and then by Schur's lemma its existence implies $\mu^* \cong \lambda$ up to a change of basis.

We are familiar with the case of $\mathfrak{su}(2)$:

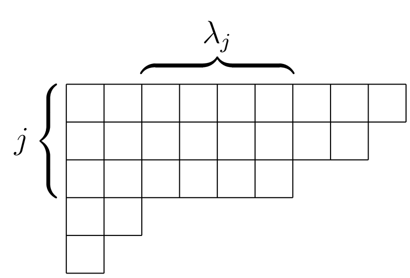
$$(2j) \otimes (2j') = \big(2(j+j')\big) \oplus \big(2(j+j'-1)\big) \oplus \cdots \oplus \big(2|j-j'|\big) \ ,$$

where we have indicated the Dynkin labels, and each representation on the RHS has multiplicity 1. The general case is complicated, and we will only explore the case of $\mathfrak{su}(N)$.

10.5 $\mathfrak{su}(N)$ and Young diagrams

There is a convenient graphical way to indicate representations of $\mathfrak{su}(N)$, called **Young diagrams**. Given a representation with Dynkin labels $(\lambda_1, \ldots, \lambda_{N-1})$, we associate a diagram containing λ_j columns of j boxes:

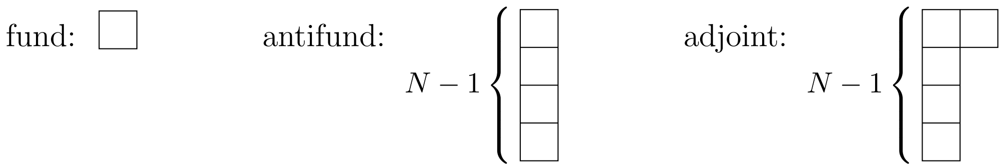
$$\big(\lambda_1, \lambda_2, \ldots, \lambda_{N-1}\big) \qquad \to$$



[55]Writing the singlet as $v_a \otimes w_i \, C^{ai}$ with $v \in V_\lambda$ and $w \in V_{\lambda^*}$, and recalling that the generators of λ^* are $-T_\alpha^{\mathsf{T}}$ in terms of those T_α of λ, we find $[T_\alpha^{\mathsf{T}}, C] = 0$ and thus, by Schur's lemma as λ is irreducible, C is the identity. We have determined the singlet, which then is unique.

The total number of boxes is $\sum_{j=1}^{r} j\lambda_j$.

For instance, the *fundamental* representation $\lambda = (1, 0, \dots)$, its conjugate the *anti-fundamental* representation $\lambda^* = (\dots, 0, 1)$ and the adjoint $\theta = (1, 0, \dots, 0, 1)$ appear as



Dimension formula. There is a simple way to compute the dimension of a representation of $\mathfrak{su}(N)$ corresponding to a Young diagram Y_λ. We need to fill the boxes with two sets of numbers, as follows.

(i) Start with the upper-left corner and assign N. Then, every time you move to the right increase the number by 1, and every time you move down you decrease the number by 1. In this way, fill the diagram with numbers n_{ij}.

(ii) Assign to each box its *hook* factor h_{ij}:

$$h_{ij} = \#\text{ boxes on the right} + \#\text{ boxes below} + 1\ .$$

The dimension of the representation is

$$\dim V_\lambda = \prod_{ij} \frac{n_{ij}}{h_{ij}}\ .$$

Example 10.5. Consider the representation $\lambda = (1, 2, 0, \dots)$ of $\mathfrak{su}(N)$. To compute its dimension, we fill in its Young diagram:

N	$N+1$	$N+2$
$N-1$	N	

4	3	1
2	1	

Then the dimension of the representation is

$$\dim \lambda = \frac{N^2(N+1)(N+2)(N-1)}{24}\ .$$

¶ *Exercise* 28. Compute the dimension of the fundamental, the anti-fundamental and the adjoint representation of $\mathfrak{su}(N)$.

Product decomposition. Given two Young diagrams Y_λ, Y_μ, the *Littlewood-Richardson rule* tells us how to decompose the product $V_\lambda \otimes V_\mu$ into irreducible representations.

(a) Fill the second diagram with numbers: 1 in all boxes of the first row, 2 in the second row, *etc.*

(b) Add all boxes with a 1 to the first diagram, keeping only the resulting diagrams that satisfy:

  (i) The resulting diagram is regular: the heights of the columns are non-increasing.

  (ii) There should not be two 1's in the same column.

  Then, from each resulting diagram remove columns with N boxes.

(c) Proceed adding all boxes with a 2 to the previous diagrams, keeping only the resulting diagrams that satisfy (i) and (ii) above, where in (ii) 1 is replaced by 2, as well as:[56]

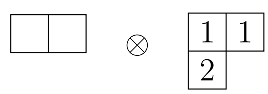
  (iii) In counting from right to left and top to bottom (*i.e.* concatenating the reversed rows into a sequence), the number of 1's must always be $\geq$ than the number of 2's, the number of 2's $\geq$ than the number of 3's, and so on.

  Then remove columns with N boxes.

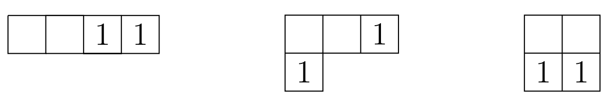
(d) Continue until all numbered boxes are used.

The resulting diagrams, with multiplicity, furnish the decomposition.

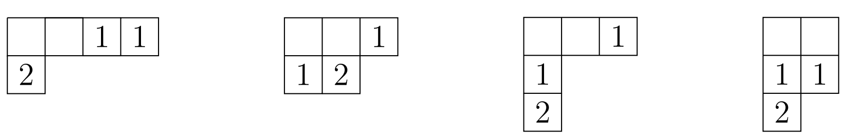
Example 10.6. We compute the decomposition of $(2,0) \otimes (1,1)$ in $\mathfrak{su}(3)$:



At the first step we add the boxes with 1, obtaining the diagrams



(a diagram with 1's on top of each other is excluded). At the second step we add the box with 2, obtaining the diagrams



After removing the columns with three boxes, we find

$$(2,0) \otimes (1,1) = (3,1) \oplus (1,2) \oplus (2,0) \oplus (0,1) \ .$$

The dimensions add up correctly: $6 \times 8 = 24 + 15 + 6 + 3$.

[56] The rules (ii) and (iii) guarantee that the procedure applied to the trivial product $1 \otimes Y_\lambda$ gives only Y_λ.

¶ *Exercise* 29. Verify the following decomposition in $\mathfrak{su}(3)$:

$$(1,1)\otimes(1,1)=(2,2)\oplus(3,0)\oplus(0,3)\oplus(1,1)\oplus(1,1)\oplus(0,0)$$

and check that the dimensions add up correctly.

¶ *Exercise* 30. Verify the following decomposition in $\mathfrak{su}(N)$:

$$(1,0\dots)\otimes(1,1,1,0\dots)=(2,1,1,0\dots)\oplus(0,2,1,0\dots)\oplus(1,0,2,0\dots)\oplus(1,1,0,1,0\dots)$$

and check that the dimensions add up correctly.

Generating representations. The fundamental representation $(1,0\dots)$ or $\Box$ has dimension N: it is a vector of $\mathfrak{su}(N)$, *i.e.* a tensor with a single index.

By taking multiple products of the fundamental representation we see that:

(a) The basic representations with a single non-vanishing Dynkin label equal to 1 are

$$\underset{j}{(\dots 0,1,0\dots)} \qquad \text{antisymmetric tensors with } j \text{ indices}$$

(b) The representations

$$(n,0\dots) \qquad \text{symmetric tensors with } n \text{ indices}$$

(c) The representations with a single non-vanishing Dynkin label are

$$\underset{j}{(\dots 0,n,0\dots)} \qquad \text{symmetric product of } n \text{ copies of } j\text{-antisymmetric tensors}$$

(d) The fact that a column with N boxes can be removed is because $\epsilon_{\mu_1\dots\mu_N}$ is a singlet of $\mathfrak{su}(N)$.

(e) A generic representation $(\lambda_1,\dots,\lambda_{N-1})$ is a tensor with $\sum_{j=1}^{N-1} j\,\lambda_j$ indices and a certain symmetry property.

Therefore all representations can be obtained from products of the fundamental representation, by applying a certain symmetry pattern.

11 Real forms and compact Lie groups

So far we have studied complex Lie algebras (because they are simple to classify), but for physical applications we are mostly interested to real Lie algebras, possibly of compact Lie groups.

In the Cartan-Weyl basis $\{H^i, E^\alpha\}$ all structure constants are real. We can then restrict to *real* linear combinations of those generators, to get a *real* Lie algebra $\mathfrak{g}_\mathbb{R}$. This real Lie algebra is called the **split real form** of $\mathfrak{g}$. For classical algebras the split real forms are

$$A_r \to \mathfrak{sl}(r+1,\mathbb{R}) \ , \qquad B_r \to \mathfrak{so}(r+1,r) \ , \qquad C_r \to \mathfrak{sp}(2r,\mathbb{R}) \ , \qquad D_r \to \mathfrak{so}(r,r) \ .$$

In the exceptional E-type cases, the split real forms are indicated as

$$\mathfrak{e}_{6(6)} \ , \qquad \mathfrak{e}_{7(7)} \ , \qquad \mathfrak{e}_{8(8)} \ ,$$

while for F_4 and G_2 there is no standard symbol.

Instead, take a basis given by

$$iH^i \ , \qquad J_1^\alpha \equiv \frac{E^\alpha + E^{-\alpha}}{i\sqrt{2}} \ , \qquad J_2^\alpha \equiv \frac{E^\alpha - E^{-\alpha}}{\sqrt{2}}$$

and consider the real algebra $\mathfrak{g}_\mathbb{R}$ of real linear combinations. It is easy to check that, again, the structure constants are real.

Since $k(H^i,H^j) = \delta^{ij}$, $k(H^i,E^\alpha) = k(E^\alpha,E^\beta) = 0$ for $\beta \neq -\alpha$ and $k(E^\alpha,E^{-\alpha}) = 2/|\alpha|^2$, in the basis above the Killing form is diagonal and *negative-definite*. This guarantees that the corresponding Lie group is compact, and therefore this is called the **compact real form**.

To show that the group is compact, one introduces an abstract notion of **Hermitian conjugation** † as follows:

$$(H^i)^\dagger = H^i \ , \qquad (E^\alpha)^\dagger = E^{-\alpha} \ ,$$

then extended to $\mathfrak{g}$ in an anti-linear way: $(cX)^\dagger = c^* X^\dagger$. Therefore

$$(iH^j)^\dagger = -iH^j \ , \qquad\qquad (J_{1,2}^\alpha)^\dagger = -J_{1,2}^\alpha \ ,$$

i.e. all elements of $\mathfrak{g}_\mathbb{R}$ are *anti-Hermitian*. Then the exponential map

$$M = e^A$$

gives a subgroup of a unitary group (since $M^{-1} = e^{-A} = M^\dagger$), which is compact. For classical algebras the compact real forms are

$$A_r \to \mathfrak{su}(r+1) \ , \qquad B_r \to \mathfrak{so}(2r+1) \ , \qquad C_r \to \mathfrak{usp}(2r) \ , \qquad D_r \to \mathfrak{so}(2r) \ .$$

while in the exceptional cases we use the same symbols as in the complex case.

Other (non-compact) real forms are $\mathfrak{su}(p,q)$ and $\mathfrak{so}(p,q)$. The full classification of real forms of $\mathfrak{g}$ boils down to classifying involutive ($\sigma^2 = 1$) automorphisms of its compact real form (due to a theorem by Cartan), and this is done with *Satake diagrams*.

A compact real form $\mathfrak{g}_{\mathbb{R}}$ is the Lie algebra of a *compact Lie group*, for instance constructed with the exponential map.[57] By taking the universal cover, to each compact real form we associate a *compact connected simply-connected Lie group* G, such that $\mathfrak{g}_{\mathbb{R}}$ is its Lie algebra:

$$\begin{aligned} &\mathfrak{su}(N) \to SU(N) \;, && \mathfrak{so}(N) \to Spin(N) \;, && \mathfrak{sp}(2r) \to USp(2r) \\ &\mathfrak{e}_r \to E_r \;, && \mathfrak{f}_4 \to F_4 \;, && \mathfrak{g}_2 \to G_2 \;. \end{aligned}$$

The group of conjugacy classes of representations P/Q is equal to the **center** Z of G:

$$Z(G) \equiv \big\{ z \in G \,\big|\, gzg^{-1} = z \;,\; \forall\, g \in G \big\} = P/Q \;.$$

Representations of the algebra are also representations of the simply-connected group G (using the exponential map). Since the center of G commutes with G, it is represented by the identity matrix multiplied by phases in each irreducible representation, and so it gives a "conserved charge" of representations. This is exactly the same as the conjugacy classes P/Q.

In fact the centers of those groups are

G	$SU(N)$	$Spin(N)$	$USp(2r)$	E_6	E_7	E_8	F_4	G_2
$Z(G)$	$\mathbb{Z}_N$	$\mathbb{Z}_2$ if $N = 1,3 \mod 4$ $\mathbb{Z}_2 \times \mathbb{Z}_2$ if $N = 0 \mod 4$ $\mathbb{Z}_4$ if $N = 2 \mod 4$	$\mathbb{Z}_2$	$\mathbb{Z}_3$	$\mathbb{Z}_2$	1	1	1

In particular

$$SO(N) = Spin(N)/\mathbb{Z}_2 \;.$$

$Spin(N)$ has all representations, including *spinor* representations; $SO(N)$ has only *tensor* representations.

From the overlap of algebras we learn that

$$\begin{aligned} &Spin(2) = U(1) \;, && Spin(3) = SU(2) \;, && Spin(4) = SU(2) \times SU(2) \\ &Spin(5) = USp(4) \;, && Spin(6) = SU(4) \;. \end{aligned}$$

11.1 Spinors in various dimensions

The basic representation of $\mathfrak{so}(N)$ and $Spin(N)$ is not the vector representation, but rather the one (for $N = 2r + 1$) or two (for $N = 2r$) spinor representations, in the sense that

[57] Starting with the algebra, one can construct the *universal enveloping algebra* which is defined as the set of all formal power series in elements of $\mathfrak{g}$. This naturally matches the product of generators in any representation. In the universal enveloping algebra, the exponential map is well-defined, and it gives a Lie group whose algebra is the Lie algebra we started with. This particular construction gives the centerless group, also called the adjoint group. Its universal cover is the simply-connected group G, and the centerless group is G/Z where Z is the center of G.

all other representations can be obtained from products of spinor representations. Let us discuss Lorentz spinors, namely spinors of $\mathfrak{so}(d-1,1)$.

We start with the **Clifford algebra**

$$\{\Gamma^\mu, \Gamma^\nu\} = 2\eta^{\mu\nu} ,$$

and take $\eta^{\mu\nu} = \mathrm{diag}(-1,+1,\ldots,+1)$. The relation between the Clifford algebra and the orthogonal algebra is that

$$\Sigma^{\mu\nu} = -\frac{i}{4}[\Gamma^\mu, \Gamma^\nu]$$

are generators of the Lorentz algebra $\mathfrak{so}(d-1,1)$, so representations of the Clifford algebra are also representations of the Lorentz group.

We start in even dimension, $d = 2k$. A faithful representation of the Clifford algebra can be constructed recursively. For $d = 2$ take

$$\Gamma^0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \qquad\qquad \Gamma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

Then from $d = 2k-2$ to $d = 2k$ take

$$\begin{aligned} \Gamma^\mu &= \gamma^\mu \otimes \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad & \mu &= 0,\ldots,d-3 \\ \Gamma^{d-2} &= \mathbb{1} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad\qquad & \Gamma^{d-1} &= \mathbb{1} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \end{aligned}$$

with γ^μ the Dirac matrices in $d = 2k-2$. This is called the **Dirac representation** and its dimension is 2^k. Notice that

$$(\Gamma^0)^\dagger = -\Gamma^0 , \qquad\qquad (\Gamma^{\mu\neq 0})^\dagger = \Gamma^\mu .$$

The Dirac representation is an irreducible representation of the Clifford algebra, but a reducible representation of the Lorentz algebra. Define

$$\Gamma = i^{k-1}\Gamma^0\Gamma^1\ldots\Gamma^{d-1} ,$$

which has the properties

$$(\Gamma)^2 = \mathbb{1} , \qquad \{\Gamma, \Gamma^\mu\} = 0 , \qquad [\Gamma, \Sigma^{\mu\nu}] = 0 .$$

Γ is called the *chirality matrix* and has eigenvalues ± 1. The 2^{k-1} states with chirality $+1$ form a **Weyl spinor representation** of the Lorentz algebra, and the 2^{k-1} states with chirality -1 form a second, inequivalent, Weyl representation.

In terms of Dynkin labels, the two Weyl spinor representations have a single 1 on one of the two small tails.

In odd dimension, $d = 2k+1$, we simply add $\Gamma^d \equiv \Gamma$ to the Dirac matrices of $d = 2k$. This is now an irreducible representation (because $\Sigma^{\mu d}$ anti-commutes with Γ^d). Thus, in $\mathfrak{so}(2k,1)$ the Dirac spinor representation has dimension 2^k, is irreducible, and in terms of Dynkin labels it has a single 1 on the shaded node (short root).

For $d = 2k$, the irreducible 2^k-dimensional representation of the Clifford algebra we constructed is unique, up to change of basis, indeed the matrices $\Gamma^{\mu *}$ and $-\Gamma^{\mu *}$ satisfy the same Clifford algebra as Γ^μ, and are related to Γ^μ by a similarity transformation.

In the basis we chose, $\Gamma^3, \Gamma^5, \ldots, \Gamma^{d-1}$ are imaginary while the other ones are real. Define

$$B_1 = \Gamma^3\Gamma^5 \ldots \Gamma^{d-1} \, , \qquad\qquad B_2 = \Gamma B_1 \, ,$$

then (using the commutation relations)

$$B_1 \Gamma^\mu B_1^{-1} = (-1)^{k-1}\, \Gamma^{\mu *} \, , \qquad\qquad B_2 \Gamma^\mu B_2^{-1} = (-1)^{k}\, \Gamma^{\mu *} \, .$$

For both matrices,

$$B \Sigma^{\mu\nu} B^{-1} = -\Sigma^{\mu\nu *} \, .$$

It follows that the spinors ζ and $B^{-1}\zeta^*$ transform in the same way under the Lorentz group, so the Dirac representation is self-conjugate (it is invariant under **charge conjugation**).

Acting on the chirality matrix:

$$B_1 \Gamma B_1^{-1} = B_2 \Gamma B_2^{-1} = (-1)^{k-1}\, \Gamma^* \, ,$$

therefore for $d = 0 \mod 4$ the two Weyl representations are conjugate to each other (charge conjugation flips the chirality), while for $d = 2 \mod 4$ each Weyl representation is self-conjugate (charge conjugation does not flip the chirality).

We can try to impose a **Majorana** (*i.e.* reality) condition on spinors, relating ζ and ζ^*. Consistency with Lorentz transformations requires

$$\zeta^* = B\zeta \qquad\qquad \text{with either } B_1 \text{ or } B_2 \, .$$

Taking the $*$ we get $\zeta = B^*\zeta^* = B^*B\zeta$, therefore such a condition is consistent if and only if $B^*B = 1$. One finds

$$B_1^* B_1 = (-1)^{k(k-1)/2} \, , \qquad\qquad B_2^* B_2 = (-1)^{(k-1)(k-2)/2} \, .$$

Therefore a Majorana condition for Dirac spinors is possible only for $d = 0, 2, 4 \mod 8$.

A Majorana condition for Weyl spinors is possible only if, moreover, the representation is self-conjugate. Therefore in $d = 2 \mod 8$ one can define **Majorana-Weyl spinors**, while in $d = 0, 4 \mod 8$ one can have either Weyl *or* Majorana spinors.

For $d = 2k + 1$, $\Gamma^d \equiv \Gamma$ and so the conjugation of Γ^d is compatible with the conjugation of Γ^μ only using B_1. Therefore a Majorana condition is possible only for $d = 1, 3 \mod 8$.

Summarizing:

d	Majorana	Weyl	Majorana-Weyl	min. rep. over $\mathbb{R}$
2	yes	self	yes	1
3	yes	–	–	2
4	yes	complex	–	4
5	–	–	–	8
6	–	self	–	8
7	–	–	–	16
8	yes	complex	–	16
9	yes	–	–	16
$10 = 8 + 2$	yes	self	yes	16
$11 = 8 + 3$	yes	–	–	32
$12 = 8 + 4$	yes	complex	–	64

Now consider spinors of $\mathfrak{so}(d)$. The discussion is similar. It turns out that the reality properties of representations of

$$\mathfrak{so}(p, q)$$

only depend on $p - q$, therefore $\mathfrak{so}(d)$ behaves like $\mathfrak{so}(d + 1, 1)$.

When the Majorana condition can be imposed, the spinor representation is **real** and the Lorentz generators can be chosen to be *imaginary*. Otherwise the representation is **pseudo-real**: the conjugate is isomorphic to itself, but the generators cannot be chosen to be imaginary and the representation is not real. A familiar example of pseudo-real representation is the **2** of $\mathfrak{so}(3)$. The product of two pseudo-real representations is a real representation.

More details can be found in Polchinski's book Volume 2, Appendix B.

12 Subalgebras

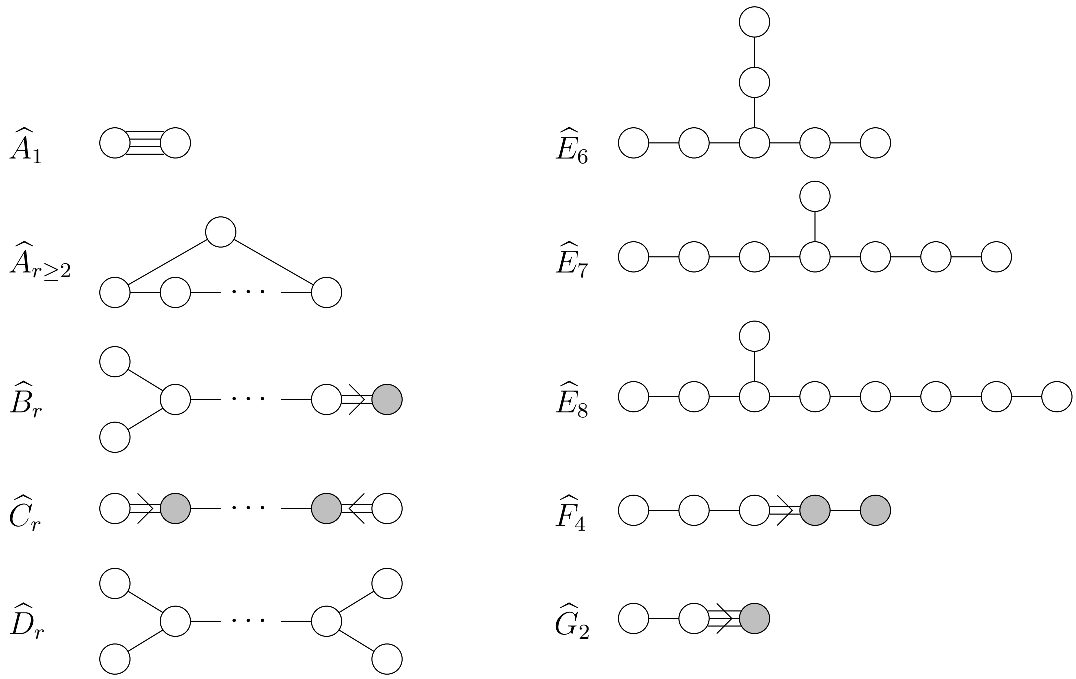
We would like to classify the possible embeddings of a semi-simple Lie algebra $\mathfrak{p}$ into a simple Lie algebra $\mathfrak{g}$. To organize the classification, one restricts to **maximal embeddings**

$$\mathfrak{p} \subset \mathfrak{g} \,,$$

for which there is no $\mathfrak{p}'$ such that $\mathfrak{p} \subset \mathfrak{p}' \subset \mathfrak{g}$. All non-maximal embeddings can be obtained from a chain of maximal ones.

We distinguish two categories: *regular subalgebras* and *special* (or non-regular) *subalgebras.*

Regular subalgebras. A regular subalgebra $\mathfrak{p}$ is one whose generators are a subset of the generators of $\mathfrak{g}$. A maximal regular subalgebra has the same rank as $\mathfrak{g}$, thus it retains all Cartan generators.



First, construct the **extended** (or **affine**) **Dynkin diagram** of $\mathfrak{g}$, by adding an extra node associated to $-\theta$ (minus the highest root).[58] Since θ is expanded in the simple-root basis in terms of *marks* a_i,

$$\theta = \sum_{i=1}^{r} a_i \, \alpha_i \ ,$$

one can compute the Cartan matrix of the extended Dynkin diagram (marks can be found in Di Francesco's book, Appendix 13.A). The extended Dynkin diagrams are in Figure.

Then, to maintain linear independence between the simple roots, we should drop (at least) one of the α_i. All *semi-simple* maximal regular subalgebras are obtained by dropping an α_i whose mark a_i is a prime number.

In the few cases in which a_i is not prime, one obtains a subalgebra of a maximal semi-simple subalgebra obtained by dropping another α_i. The exhaustive list is

$$\begin{aligned} &F_4 \supset B_4 \supset A_3 \oplus A_1 && E_7 \supset D_6 \oplus A_1 \supset A_3 \oplus A_3 \oplus A_1 \\ &E_8 \supset D_8 \supset D_5 \oplus A_3 && E_8 \supset E_6 \oplus A_2 \supset A_5 \oplus A_2 \oplus A_1 \\ &E_8 \supset E_7 \oplus A_1 \supset A_7 \oplus A_1 && E_8 \supset E_7 \oplus A_1 \supset A_5 \oplus A_2 \oplus A_1 \ . \end{aligned}$$

¶ *Exercise* 31. Check them.

[58]Promoting $-\theta$ to a "simple root" preserves the property that the difference between two simple roots is not a root (because $\alpha_i + \theta$ is not a root). In fact, $\det A = 0$ but all principal minors are positive.

Maximal regular subalgebras that are *not semi-simple* are constructed from the removal of 2 nodes with mark $a_i = 1$ and the addition of a $\mathfrak{u}(1)$ factor. For instance

$$\mathfrak{su}(p+q) \supset \mathfrak{su}(p) \oplus \mathfrak{su}(q) \oplus \mathfrak{u}(1) \qquad \text{for } p, q \geq 1 \; .$$

Special subalgebras. The maximal subalgebras of the *classical algebras* are of two types. The first type uses the fact that classical algebras are algebras of matrices. Thus[59]

$$\begin{aligned}
&\mathfrak{su}(p) \oplus \mathfrak{su}(q) && \supset \mathfrak{su}(pq) \\
&\mathfrak{so}(p) \oplus \mathfrak{so}(q) && \supset \mathfrak{so}(pq) \\
&\mathfrak{sp}(2p) \oplus \mathfrak{sp}(2q) && \supset \mathfrak{so}(4pq) \\
&\mathfrak{sp}(2p) \oplus \mathfrak{so}(q) && \supset \mathfrak{sp}(2pq) \\
&\mathfrak{so}(p) \oplus \mathfrak{so}(q) && \supset \mathfrak{so}(p+q) \qquad \text{for } p \text{ and } q \text{ odd} \; .
\end{aligned}$$

In the first four we write the indices on the right as a double index, realizing the algebras on the left. The last case is obvious, however it does not appear from manipulations of the extended Dynkin diagram.

The second type uses the fact that if $\mathfrak{p}$ has an N-dimensional representation, then it is a subalgebra of $\mathfrak{su}(N)$. Since all representations are isomorphic to unitary representations, just take the representatives as $N \times N$ matrices.

Is $\mathfrak{p}$ maximal in $\mathfrak{su}(N)$? With a few exceptions,[60] if the N-dimensional representation is real (and thus it admits a symmetric bilinear form) then $\mathfrak{p}$ is maximal in $\mathfrak{so}(N)$, if it is pseudo-real (and thus it admits an anti-symmetric bilinear form) it is maximal in $\mathfrak{sp}(N)$, otherwise it is maximal in $\mathfrak{su}(N)$.

The maximal special subalgebras of the *exceptional algebras* are listed, for instance in Table 13.1 of Di Francesco's book.

[59] The unitary, orthogonal and symplectic groups are groups of matrices that preserve a Hermitian, symmetric and antisymmetric bilinear form, respectively. Writing $\mathbb{C}^{pq} = \mathbb{C}^p \otimes \mathbb{C}^q$, the bilinear form can be written as $(\phi_1 \otimes \phi_2, \eta_1 \otimes \eta_2) = (\phi_1, \eta_1)_1 \cdot (\phi_2, \eta_2)_2$. This gives the embedding of groups.

[60] The exceptions are listed by Dynkin.