Mastermath course Differential Geometry 2015/2016

Marius Crainic

[1]Department of Mathematics, Utrecht University, 3508 TA Utrecht, The Netherlands.
E-mail address: `m.crainic@uu.nl`.
Received by the editors September, 2015.

Contents

Contents- to be continued

Contents- to be continued

CHAPTER 1

Vector bundles and connections

1.1. Vector bundles

In this section we recall some basics on vector bundles.

1.1.1. The definition/terminology

Definition 1.1. *A (real) vector bundle of rank r over a manifold M consists of:*

- *a manifold E,*
- *a surjective map $\pi : E \longrightarrow M$; for $x \in M$, the fiber $\pi^{-1}(x)$ is denoted E_x*
- *for each $x \in M$, a structure of r-dimensional (real) vector space on E_x*

satisfying the following local triviality condition: for each $x_0 \in M$, there exists an open neighborhood U of x_0 and a diffeomorphism

$$h : E|_U := \pi^{-1}(U) \xrightarrow{\sim} U \times \mathbb{R}^r$$

with the property that it sends each fiber E_x isomorphically to $\{x\} \times \mathbb{R}^r$, where "isomorphically" means by a vector space isomorphism and where we identify $\{x\} \times \mathbb{R}^r$ with $\mathbb{R}^r$. Complex vector bundles are defined similarly, replacing $\mathbb{R}$ by $\mathbb{C}$.

Strictly speaking, a vector bundle is an entire triple (E, π, M) as above; accordingly, one sometimes uses notations of type

$$\xi = (E, \pi, M) \tag{1.2}$$

and one refers to the triple ξ as being a vector bundle; with this terminology, E is usually denoted $E(\xi)$ and is called the total space of the vector bundle ξ. However, we will adopt a simpler terminology (hopefully not too confusing): we will just mention E (i.e. we just say that E is a vector bundle over M), by which it is self understood that E comes with a map onto M, generically denoted π_E, and called the projection of the vector bundle E. Intuitively, one should think about E as the collection

$$\{E_x\}_{x \in M}$$

of vector spaces (of rank r), "smoothly parametrized by $x \in M$".

Note that, as hinted by the notation $E|_U$, a vector bundle E over M can be restricted to an arbitrary open $U \subset M$. More precisely,

$$E|_U := \pi^{-1}(U)$$

together with the restriction of π gives a vector bundle $\pi_U : E|_U \longrightarrow U$ over U.

Here are some basic concepts/constructions regarding vector bundles.

1.1.2. Morphisms

Given two vector bundles E and F over M, a morphism from E to F (of vector bundles over M) is a smooth map $u : E \longrightarrow F$ with the property that, for each $x \in M$, u sends E_x to F_x and

$$u_x := u|_{E_x} : E_x \longrightarrow F_x$$

is linear. We say that u is an isomorphism (of vector bundles over M) if each u_x is an isomorphism (or, equivalently, if u is also a diffeomorphism). Again, one should think of a morphism u as a collection $\{u_x\}_{x\in M}$ of linear maps between the fibers, "smoothly parametrized by x".

1.1.3. Trivial vector bundles; trivializations

The trivial vector bundle of rank r over M is the product $M \times \mathbb{R}^r$ together with the first projection

$$\mathrm{pr}_1 : M \times \mathbb{R}^r \longrightarrow M$$

and the usual vector space structure on each fiber $\{x\} \times \mathbb{R}^r$. When using the notation (1.2), the trivial vector bundle of rank r is usually denoted

$$\epsilon^r = (M \times \mathbb{R}^r, \mathrm{pr}_1, M) \tag{1.3}$$

(or, if one wants to be more precise about the base, then one uses the notation ϵ^r_M).

We say that a vector bundle E (of rank r) over M is trivializable if E is isomorphic to $M \times \mathbb{R}^r$. A trivialization of E is the choice of such an isomorphism.

With these in mind, we see that the local triviality condition from the definition of vector bundles says that E is locally trivializable, i.e. each point in M admits an open neighborhood U such that the restriction $E|_U$ is trivializable.

1.1.4. Sections

One of the main objects associated to vector bundles are their (local) sections. Given a vector bundle $\pi : E \longrightarrow M$, a section of E is a smooth map $s : M \longrightarrow E$ satisfying $\pi \circ s = \mathrm{Id}$, i.e. with the property that

$$s(x) \in E_x \quad \forall\, x \in M.$$

We denote by

$$\Gamma(M, E) = \Gamma(E)$$

the space of all smooth sections. For $U \subset M$ open, the space of local sections of E defined over U is

$$\Gamma(U, E) := \Gamma(E|_U).$$

Sections can be added pointwise:

$$(s + s')(x) := s(x) + s'(x)$$

and, similarly, can be multiplied by scalars $\lambda \in \mathbb{R}$:

$$(\lambda s)(x) := \lambda s(x).$$

With these, $\Gamma(E)$ becomes a vector space. Furthermore, any section $s \in \Gamma(E)$ can be multiplied pointwise by any real-valued smooth function $f \in C^{\infty}(M)$ giving rise to another section $fs \in \Gamma(E)$:

$$(fs)(x) := f(x)s(x).$$

The resulting operation

$$C^\infty(M) \times \Gamma(E) \longrightarrow \Gamma(E), \quad (f, s) \mapsto fs$$

makes $\Gamma(E)$ into a module over the algebra $C^\infty(M)$. Actually the entire vector bundle E is fully encoded in the space of sections $\Gamma(E)$ together with this module structure.

For the curious reader. The precise formulation of the last statement is given by Swan's theorem which says that the construction $E \mapsto \Gamma(E)$ gives a 1-1 correspondence between vector bundles over M and finitely generated projective modules over $C^\infty(M)$. Recall here that, for a ring R, an R-module $\mathcal{E}$ is said to be finitely generated and projective if there exists another R-module $\mathcal{F}$ such that the direct sum R-module $\mathcal{E} \oplus \mathcal{F}$ is isomorphic to the free R-module R^k for some k. This corresponds to a basic property of vector bundles: for any vector bundle E over M, one can always find another vector bundle F over M such that the direct sum vector bundle $E \otimes F$ (see below) is isomorphic to the a trivial vector bundle.

Note that this discussion is very much related, at least in spirit, with the Gelfand-Naimark theorem which says that (compact) topological spaces X can be recovered from the algebra $C(X)$ of continuous functions on X (see Chapter 8, section 3, from our bachelor course "Inleiding Topologie", available at: http://www.staff.science.uu.nl/∼crain101/topologie2014/).

A simpler illustration of the previous principle is the following:

Lemma 1.4. *Let E and F be two vector bundles over M. Then there is a bijection between':*

- *morphisms $u : E \longrightarrow F$ of vector bundles over M.*
- *morphisms $u_* : \Gamma(E) \longrightarrow \Gamma(F)$ of $C^\infty(M)$-modules.*

Explicitely, given u, the associated u_* is given by

$$u_*(s)(x) = u_x(s(x)).$$

1.1.5. Frames

Let $\pi : E \longrightarrow M$ be a vector bundle over M. A frame of E is a collection

$$s = (s^1, \ldots, s^r)$$

consisting of sections s^i of E with the property that, for each $x \in M$,

$$(s^1(x), \ldots, s^r(x))$$

is a frame of E_x (i.e. a basis of the vector space E_x). A local frame of E is a frame s of $E|_U$ for some open $U \subset M$; we also say that s is a local frame over U.

Remark 1.5. Choosing a frame s of E is equivalent to choosing a trivialization $u : M \times \mathbb{R}^r \longrightarrow E$ of E. Hence, the local triviality condition from the definition of vector bundles can be phrased in terms of local frames as follows: around any point of M one can find a local frame of E.

1.1.6. Remark on the construction of vector bundles

Often the vector bundles that one encounters do not arise right away as in the original definition of vector bundles. Instead, one has just a collection $E = \{E_x\}_{x \in M}$ of vector spaces indexed by $x \in M$ and certain "smooth sections". Let us formalize this a bit. We will use the name "discrete vector bundle over M (of rank r)" for any collection $\{E_x\}_{x \in M}$ of (r-dimensional) vector spaces indexed by $x \in M$. We will identify such a collection with the resulting disjoint union and the associated projection

$$E := \{(x, v_x) : x \in M, v_x \in E_x\}, \ \pi : E \longrightarrow M, \quad (x, v_x) \mapsto x$$

(so that each E_x is identified with the fiber $\pi^{-1}(x)$).

For such a discrete vector bundle E we can talk about discrete sections, which are simply functions s as above,

$$M \ni x \mapsto s(x) \in E_x$$

(but without any smoothness condition). Denote by $\Gamma_{\text{discr}}(E)$ the set of such sections. Similarly we can talk about discrete local sections, frames and local frames. As in the case of charts of manifolds, there is a natural notion of "smooth compatibility" of local frames. To be more precise, we assume that

$$s = (s^1, \ldots, s^r), \ \tilde{s} = (\tilde{s}^1, \ldots, \tilde{s}^r)$$

are two local frames defined over U and $\tilde{U}$, respectively. Then, over $U \cap \tilde{U}$, one can write

$$\tilde{s}^i(x) = \sum_{j=1}^{r} g_j^i(x) s^j(x),$$

giving rise to functions

$$g_j^i : U \cap \tilde{U} \longrightarrow \mathbb{R} \ \ (1 \leq i, j \leq r).$$

We say that s and $\tilde{s}$ are smoothly compatible if all the functions g_j^i are smooth. The following is an instructive exercise.

Exercise 1. Let $E = \{E_x\}_{x \in M}$ be a discrete vector bundle over M of rank r. Assume that we are given an open cover $\mathcal{U}$ of M and, for each open $U \in \mathcal{U}$, a discrete local frame s_U of E over U. Assume that, for any $U, V \in \mathcal{U}$, s_U and s_V are smoothly compatible. Then E admits a unique smooth structure which makes it into a vector bundle over M with the property that all the s_U become (smooth) local frames.

Moreover, the (smooth) sections of E can be recognized as those discrete sections s with the property that they are smooth with respect to the given data $\{s_U\}_{U \in \mathcal{U}}$ in the following sense: for any $U \in \mathcal{U}$, writing

$$u(x) = f_1(x) s_U^1(x) + \ldots + f_r(x) s_U^r(x) \quad (x \in U),$$

all the functions f_i are smooth on E.

Example 1.6. For a manifold M one consider all the tangent spaces $TM = \{T_x M\}_x \in M$ and view it as a discrete vector bundle. Given a chart $\chi : U_\chi \longrightarrow \mathbb{R}^n$ for M, then the associated tangent vectors

$$(\frac{\partial}{\partial \chi_1}(x), \ldots, \frac{\partial}{\partial \chi_n}(x))$$

can be seen as a discrete local frame of TM over U_χ. Starting with an atlas $\mathcal{A}$ of M, we obtain in this way precisely the data that we need in order to apply the previous exercise; this makes TM into a vector bundle over M in the sense of the original definition.

Example 1.7. Here is another interesting vector bundle- the so called tautological line bundle over the projective space $\mathbb{R}P^n$. In the notation (1.2), this is usually denoted by γ_1. We describe the total space $E(\gamma^1)$. Recalling the a point $l \in \mathbb{R}P^n$ is a line $l \subset \mathbb{R}^{n+1}$ through the origin, the fiber of $E(\gamma^1)$ above such a point is precisely the line l, interpreted as a 1-dimensional vector space. Equivalently:

$$E(\gamma^1) = \{(l, v) \in \mathbb{R}P^n \times \mathbb{R}^{n+1} : v \in l\},$$

or, denoting by $[x_0 : \ldots : x_n]$ the line through $(x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \setminus \{0\}$,

$$E(\gamma^1) = \{([x_0 : \ldots : x_n], \lambda \cdot x_0, \ldots, \lambda \cdot x_n)) : (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \setminus \{0\}, \lambda \in \mathbb{R}\}.$$

With these, γ_1 is a vector bundle of rank one (a line bundle) over $\mathbb{R}P^n$.

Such vector bundles exist for arbitrary ranks: one replaces $\mathbb{R}P^n$ with the Grassmannian $\mathrm{Gr}_k(\mathbb{R}^{n+k})$ (whose points are the k-dimensional vector subspaces of $\mathbb{R}^{n+k}$) and γ_1 by γ_k whose fiber above some $V \in \mathrm{Gr}_k(\mathbb{R}^{n+k})$ is the vector space V itself.

Of course, these examples can be adapted to the complex setting, by working with complex lines (and complex vector subspaces) to define $\mathbb{C}P^n$ (and $\mathrm{Gr}_k(\mathbb{C}^{n+k})$), and the tautological complex vector bundles on them.

1.1.7. Operations with vector bundles

The principle is very simple: natural operations with vector spaces, applied fiberwise, extend to vector bundles.

Direct sums: Let us start with the direct sum operation. Given two vector spaces V and W we consider their direct sum vector space $V \oplus W$. Assume now that $p_E : E \longrightarrow M$ and $p_F : F \longrightarrow M$ are vector bundles over M. Then the direct sum $E \oplus F$ is another vector bundle over M, with fibers

$$(E \oplus F)_x := E_x \oplus F_x. \tag{1.8}$$

These equations force the definition of the total space $E \oplus F$ and of the projection into M. To exhibit the smooth structure of $E \oplus F$ one can e.g. use Exercise 1. Indeed, choosing opens $U \subset M$ over which we can find (smooth) local frames $e = (e_1, \ldots, e_p)$ of E and $f = (f_1, \ldots, f_q)$ of F, one can form the direct sum local frame

$$e \oplus f = (e_1, \ldots, e_p, f_1, \ldots, f_q)$$

and we consider the smooth structure on $E \oplus F$ which makes all the local frames of type $e \oplus f$ smooth.

This procedure of extending operations between vector spaces to operations between vector bundles is rather general. In some cases however, one can further take advantage of the actual operation one deals with and obtain "more concrete" descriptions. This is the case also with the direct sum operation. Indeed, recall that for any two vector spaces V and W, their direct sum $V \oplus W$ can be described as the set-theoretical product $V \times W$ with the vector space operations

$$(v, w) + (v', w') = (v + v', w + w'), \lambda \cdot (v, w) = (\lambda \cdot v, \lambda \cdot w)$$

(the passing to the notation $V \oplus W$ indicates that we identify the elements $v \in V$ with $(v, 0) \in V \times W$, $w \in W$ with $(0, w) \in V \times W$, so that an arbitrary element (v, w) can be written uniquely as $v + w$ with $v \in V$, $w \in W$). Hence one can just define $E \oplus F$ as the submanifold of $E \times F$

$$E \times_M F := \{(e, f) \in E \times F : p_E(e) = p_F(f)\}.$$

The condition (1.8) is clearly satisfied (and specify the vector space structure on the fibers) and is not difficult to see that the resulting $E \oplus F$ is a vector bundle over M. Note that the space of sections of $E \oplus F$ coincides with the direct sum $\Gamma(E) \oplus \Gamma(F)$.

Duals: Let us now look at the operation that associates to a vector space V its dual V^*. Starting with a vector bundle E over M, its dual E^* is another vector bundle over M with the property that

$$(E^*)_x = (E_x)^*$$

for all $x \in M$. Again, this determines E^* as a set and its projection into M. Moreover, using dual basis, we see that any smooth local frame $e = (e_1, \ldots, e_r)$ of E induces a local frame e^* for E^* and we can invoke again Exercise 1 to obtain the smooth structure of E^*.

Hom-bundles: Next we look at the operation that associates to two vector spaces V and W the vector space $\mathrm{Hom}(V, W)$ consisting of all linear maps from V to W. Given now two vector bundles E and F over M, we form the new vector bundle $\mathrm{Hom}(E, F)$ over M with fibers

$$\mathrm{Hom}(E, F)_x = \mathrm{Hom}(E_x, F_x).$$

And, again, we see that local frames of E and F induce a local frame of $\mathrm{Hom}(E, F)$, and then we obtain a canonical smooth structure on the hom-vector bundle. Note that a section of $\mathrm{Hom}(E, F)$ is the same thing as a morphism $u : E \longrightarrow F$ of vector bundles over M. Hence Lemma 1.4 identifies sections of $\mathrm{Hom}(E, F)$ with $C^\infty(M)$-linear maps $\Gamma(E) \longrightarrow \Gamma(F)$.

Of course, when F is the trivial vector vector bundle of rank 1 ($F = M \times \mathbb{R}$), we recover the dual of E:

$$E^* = \mathrm{Hom}(E, M \times \mathbb{R}).$$

Hence Lemma 1.4 identifies the sections of E^* with the dual of $\Gamma(E)$ as an $C^\infty(M)$-module.

Quotients: Next we look at the operation that associates to a vector subspace W of a vector space V the quotient vector space V/W. Note that the notion of subspace (of a vector space) gives rise to a notion of vector sub-bundle: given a vector bundle E over M, a (vector) sub-bundle of E is any vector bundle F over M with the property that each F_x is a vector subspace of E_x and the inclusion $F \hookrightarrow E$ is smooth.

Exercise: Show that the last condition(smoothness) is equivalent to the condition that for any smooth (local) section s of F, interpreting s as a section of E using the inclusion $F \subset E$, s is a smooth section of E.

Now, given a vector sub-bundle $F \subset E$ it should be clear how to proceed to define E/F: it is a new vector bundle over M whose fiber above $x \in M$ is the quotient E_x/F_x and with the smooth structure uniquely characterised by the fact that the quotient map from E to E/F is smooth (hence a morphism of vector bundles).

Tensor products: One proceeds similarly for the tensor product operation on vector spaces. Since we work with finite dimensional vector spaces, this operation can be expressed using duals and homs:

$$V \otimes W = \mathrm{Hom}(V^*, W).$$

(where for $v \in V$, $w \in W$, the tensor $v \otimes w$ is identified with (or stands for) the linear map

$$V^* \longrightarrow W, \quad \xi \mapsto \xi(v)w.)$$

Other operations: Similar to taking the dual of a vector space one can consider operations of type

$$V \mapsto S^k(V^*), \text{ (or: } V \mapsto \Lambda^k(V^*))$$

which associate to a vector space V the space of all k-multilinear symmetric (or: anti-symmetric) maps $V \times \ldots \times V \longrightarrow \mathbb{R}$. Again, one has to remember/notice that any frame of V induces a frame of S^kV^* (or: Λ^kV^*). Slightly more generally, one can consider operations of type

$$(V, W) \longrightarrow S^k(V^*) \otimes W \text{ (or: } (V, W) \mapsto \Lambda^k(V^*) \otimes W)$$

which associate to a pair (V, W) of vector spaces the space of k-multilinear symmetric (or: anti-symmetric) maps on V with values in W and then one obtains similar operations on vector bundles. Note the following generalization of Exercise 1:

Exercise 2. Show that, for any two vector bundles E and F over M and $k \geq 1$ integer, there is a 1-1 correspondence between:

- sections u of $S^kE^* \otimes F$.
- symmetric maps

$$u_* : \underbrace{\Gamma(E) \times \ldots \times \Gamma(E)}_{k-\text{times}} \longrightarrow \Gamma(F)$$

  which is $C^\infty(M)$-linear in each argument.

Similarly for sections of $\Lambda^kE^* \otimes F$ and antisymmetric maps as above.

Pull-backs: Another important operation with vector bundles, but which does not fit in the previous framework, is the operation of taking pull-backs. More precisely, given a smooth map

$$f : M \longrightarrow N,$$

starting with any vector bundle E over N, one can pull-it back via f to a vector bundle f^*E over M. Fiberwise,

$$(f^*E)_x) = E_{f(x)}$$

for all $x \in M$. One can use again Exercise 1 to make f^*E into a vector bundle; the key remark is that any section s of E induces a section f^*s of F^*E by

$$(f^*s)(x) := s(f(x))$$

and similarly for local sections and local frames.

Note that, when $f = i : M \hookrightarrow N$ is an inclusion of a submanifold M of N, then i^*E is also denoted $E|_N$ and is called the restriction of E to M.

Real versus complex vector bundles: Of course, any complex vector bundle F can be seen as a real vector bundle; when we want to emphasize that we look at F as being a real vector bundle, we use the notation $F_{\mathbb{R}}$. Note that the rank (over $\mathbb{R}$) of $F_{\mathbb{R}}$ is twice the rank (over $F_{\mathbb{C}}$) of F. A natural question is: when can a real vector bundle can be made into a comple one.

Exercise 3. For a real vector bundle E over M, there is a 1-1 correspondence between complex vector bundles F over M such that $F_{\mathbb{R}} = E$ and vector bundle morphisms $J : E \to E$ such that $J^2 = -\mathrm{Id}$.

One can also proceed the other way: starting with a real vector bundle E one can complexify it, i.e. consider the complex vector bundle

$$E_{\mathbb{C}} := E \otimes \mathbb{C}$$

whose fiber above an arbitrary point $x \in M$ is

$$E_x \otimes \mathbb{C} = \{v \otimes 1 + w \otimes i : u, v \in E_x\}.$$

One can also use the direct sum of real vector bundles and define

$$E_{\mathbb{C}} = E \oplus E,$$

in which the multiplication by complex numbers given by

$$(a + bi) \cdot (u, v) = (au - bv, av + bu) \quad \text{for } a + ib \in \mathbb{C}.$$

Or, using the previous exercise, we deal with the vector bundle $E \oplus E$ with the complex structure induced by

$$J(u, v) = (-v, u).$$

Conjugation: For any complex vector bundle F one can define a new one, denoted $\overline{F}$ and called the conjugate of F, whose underlying total space is the same as that of F, with the only difference that we change the multiplication by complex numbers on the fiber to:

$$z \cdot v := \overline{z}v$$

where $\overline{z}$ is the complex conjugate of z. Note that, in terms of the previous exercise, what we do is to change J by $-J$.

Note that, in general, $\overline{F}$ is not isomorphic to F (as complex vector bundles). For instance, this happens for the tautological line bundle γ^1 from Example 1.7 (however, proving this is not so easy at this point!).

Exercise 4. Consider the tangent bundle TS^n of the n-dimensional sphere. Show that the direct sum of TS^n with the trivial bundle of rank one is isomorphic to the trivial vector bundle of rank $n+1$ (note: this does not imply that TS^n is isomorphic to a trivial bundle, and it is not!).

Exercise 5. Show that the tangent bundle of S^3 is trivializable. (1st hint: quaternions; 2nd hint: first do it for S^1; then write your proof using complex numbers; then go to the previous hint).

For the curious reader. Manifolds M with the property that TM is trivializable are called "parallelizable". This happens rather rarely. The situation is interesting even for the spheres; we discuss it here in a bit more detail. Note that, the question of whether S^n is parallelizable or not can be rewritten in very down to earth terms: one is looking for $n + 1$ functions

$$F^1, \ldots, F^n \longrightarrow \mathbb{R}^{n+1}$$

with the property that, for each $x \in S^n \subset \mathbb{R}^{n+1}$, vectors $F^1(x), \ldots, F^n(x)$ are linearly independent and take value in the hyperplane P_x orthogonal to x:

$$F^i(x) \in P_x := \{v \in \mathbb{R}^{n+1} : \langle v, x\rangle = 0\}.$$

This comes from the standard identification of the tangent spaces T_xS^n with $P_x \subset \mathbb{R}^{n+1}$. In the case $n = 1$ one can take

$$F^1(x_1, x_2) = (x_2, -x_1)$$

or, interpreting $\mathbb{R}^2$ as $\mathbb{C}$ ($(x_1, x_2) \mapsto x_1 + ix_2$),

$$F^1(z) = iz.$$

It works quite similarly for S^3, but using quaternions (previous exercise). The 2-sphere however is not parallelizable. Actually, on S^2 any vector field must vanish at at least one point. This is popularly known as the "hairy ball theorem", and stated as: "you can't comb a hairy ball flat without creating a cowlick"(see also the "Cyclone consequences" on the wikipedia). The mathematical reason that S^2 is not parallelizable (and actually that it does not admit any no-where vanishing vector field) is that its "Euler characteristic" does not vanish. Actually, a nice (but non-trivial) theorem in differential topology says that the Euler characteristic of a compact manifold vanishes if and only if it admits a no-where vanishing vector field. By the way, for the same reason, all the even dimensional spheres S^{2n} are not parallelizable.

Are the other (odd dimensional) spheres parallelizable? Yes: there is also S^7. This uses octonions (also called the Cayley algebra) instead of quaternions (S^3) and complex numbers (S^1). How general is this? In what dimensions does it work? Looking at the proofs (e.g. for S^1 and S^3) we see that what we need is a "normed division algebra" structure on $\mathbb{R}^{n+1}$ (in order to handle the n-sphere S^n). By that we mean a multiplication "$\cdot$" on $\mathbb{R}^{n+1}$, with unit element $1 := (1, 0, \ldots, 0)$ (but not necessarily associative- as in the case of the octonions) and satisfying the norm condition

$$|x \cdot y| = |x||y| \quad \forall\, x, y \in \mathbb{R}^{n+1},$$

where $|\cdot|$ is the standard norm. The term "division algebra" comes from the fact that the norm condition implies that the product has the "division property": $xy = 0$ implies that either $x = 0$ or $y = 0$ (if we keep this condition, but we give up on the norm condition, we talk about division algebras). Indeed, any such operation induces

$$F^i(x) = e_i \cdot x$$

proving that S^n is parallelizable. But on which $\mathbb{R}^{n+1}$ do there exist such operations? Well: only on $\mathbb{R}$, $\mathbb{R}^2$ (complex), $\mathbb{R}^4$ (quaternions) and $\mathbb{R}^8$ (octonions)! This was known since the 19th century. It is also interesting to know why was there interest on such operations already in the 19th century: number theory and the question of which numbers can be written as sums of two, three, etc squares. For sum of two squares, the central formula which shows that a product of two numbers that can be written as a sum of two squares can itself be written as a sum of two squares is:

$$(x^2 + y^2)(a^2 + b^2) = (xa - yb)^2 + (xb + ya)^2.$$

Or, in terms of the complex numbers $z_1 = x + iy$, $z_2 = a_ib$:

$$|z_1 z_2| = |z_1||z_2|.$$

The search for similar "magic formulas" for sum of three squares never worked, but it did for four:

$$(x^2 + y^2 + z^2 + t^2)(a^2 + b^2 + c^2 + d^2) =$$

$$(xa + yb + zc + td)^2 + (xb - ya - zd + tc)^2 +$$

$$+(xc + yd - za - tb)^2 + (xd - yc + zb - ta)^2.$$

This is governed by the quaternions and its norm equation.

Any way, returning to the spheres, we see that the trick with the multiplication can only work for S^1, S^3 and S^7. And, indeed, one can prove that there are the only parallelizable spheres! Well, S^0 as well if you insist. The proof is highly non-trivial and makes use of the machinery of Algebraic Topology.

There are a few more things that are worth mentioning here. One is that probably the largest class of manifolds that are automatically parallelizable (because of their structure) are the Lie groups (see also later); this applies in particular to the closed subgroups of GL_n. This is one of the reasons that S^1 and S^3 are parallelizable. The circle is clearly a Lie group (with complex multiplication), which can be identified with the rotation group

$$\left\{ \begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix} \right\} \subset GL_2(\mathbb{R}).$$

The fact that S^3 can be made into a Lie group follows by characterizing S^3 as the set of unit quaternionic vectors (i.e. similar to S^1, but by using quaternions instead of complex numbers). The similar argument for S^7 (but using octonions instead of quaternions) does not work because the multiplication of octonions is not associative (to ensure that S^7 is parallelizable, the division property is enough, but for making it into a Lie group one would need associativity). Can other spheres be made into Lie groups (besides S^1 and S^3)? No!

What else? Note that the negative answer in the case of even dimensional spheres - due to the fact that such spheres do not even admit nowhere vanishing vector fields- comes with a very natural question: ok, but, looking at S^n, which is the largest number of linearly independent vector fields that it admits? Again, this is a simple but very non-trivial problem, whose solution (half way in the 20th century) requires again the machinery of Algebraic Topology. But here is the answer: write $n + 1 = 2^{4a+r}m$ with m odd, $a \geq 0$ integer, $r \in \{0, 1, 2, 3\}$. Then the maximal number of linearly independent vector fields on S^n is

$$r(n) = 8a + 2^r - 1.$$

Note that $r(n) = 0$ if n is even (no nowhere vanishing vector fields on the even dimensional spheres), while the parallelizability of S^n is equivalent to $r(n) = n$, i.e. $8a + 2r = 2^{4a+r}m$ which is easily seen to have the only solutions

$$m = 1, a = 0, r \in \{0, 1, 2, 3\}$$

giving $n = 0, 2, 4, 8$ hence again the spheres S^0, S^1, S^3 and S^7.

Exercise 6. Show that $S^2 \times S^1$ is parallelizable.

Exercise 7. Let E be a vector bundle over M, and assume that it sits inside a trivial vector bundle $M \times \mathbb{R}^k$ of some rank k; in other words E_x is a vector subspace of $\mathbb{R}^k$ for each x and E is a sub-manifold of $M \times \mathbb{R}^k$. For each $x \in M$, we denote by $E_x^{\perp}$ the orthogonal complement of E_x in $\mathbb{R}^k$. Show that these define a new vector bundle $E^{\perp}$ over M and $E \oplus E^{\perp}$ is isomorphic to $M \times \mathbb{R}^k$.

Exercise 8. Consider the tautological line bundle over $\mathbb{R}P^n$ (see Example 1.7), denoted here by γ, and we also consider the vector bundle $\gamma^{\perp}$ whose fiber over $l \in \mathbb{R}P^n$ is the orthogonal complement of l in $\mathbb{R}^{n+1}$ (see the previous Exercise). Show that one has an isomorphism of vector bundles:

$$T(\mathbb{R}P^n) \cong \mathrm{Hom}(\gamma, \gamma^{\perp}),$$

where $T(\mathbb{R}P^n)$ is the tangent bundle of $\mathbb{R}P^n$. Then do the same for the complex projective space.

1.1.8. Differential forms with coefficients in vector bundles

Vector bundles also allow us to talk about more general differential forms: with coefficients. The standard differential forms are those with coefficients in the trivial vector bundle of rank 1. Recall here that the space of (standard) differential forms of degree p on a manifold M, $\Omega^p(M)$, is defined as the space of sections of the bundle $\Lambda^p T^*M$. Equivalently, a p-form on M is the same thing as a $C^{\infty}(M)$-multilinear, antisymmetric map

$$\omega : \underbrace{\mathcal{X}(M) \times \ldots \times \mathcal{X}(M)}_{p \text{ times}} \longrightarrow C^{\infty}(M), \tag{1.9}$$

where $\mathcal{X}(M)$ is the space of vector fields on M. Such p-forms can be written locally, over the domain U of a coordinate chart $(U, \chi_1, \ldots, \chi_n)$ as:

$$\omega = \sum_{i_1,\ldots,i_p} f^{i_1,\ldots,i_p} d\chi_{i_1} \ldots d\chi_{i_p}, \tag{1.10}$$

with $f^{i_1,\ldots,i_p}$-smooth functions on U.

Assume now that E is a vector bundle over M. We define the space of E-valued p-differential forms on M

$$\Omega^p(M; E) = \Gamma(\Lambda^p T^*M \otimes E).$$

As before, an element $\omega \in \Omega^p(M; E)$ can be thought of as a $C^{\infty}(M)$-multilinear antisymmetric map

$$\omega : \underbrace{\mathcal{X}(M) \times \ldots \times \mathcal{X}(M)}_{p \text{ times}} \longrightarrow \Gamma(E). \tag{1.11}$$

Also, locally, with respect to a coordinate chart $(U, \chi_1, \ldots, \chi_n)$, one can write

$$\omega = \sum_{i_1,\ldots,i_p} d\chi_{i_1} \ldots d\chi_{i_p} \otimes e^{i_1,\ldots,i_p}. \tag{1.12}$$

with $e^{i_1,\ldots,i_p}$ local sections of E (defined on U). Using also a local frame $e = \{e_1, \ldots, e_r\}$ for E, we obtain expressions of type

$$\sum_{i_1,\ldots,i_p,i} f_i^{i_1,\ldots,i_p} dx_{i_1} \ldots dx_{i_p} \otimes e_i.$$

Recall also that

$$\Omega(M) = \bigoplus_p \Omega^p(M)$$

is an algebra with respect to the wedge product: given $\omega \in \Omega^p(M)$, $\eta \in \Omega^q(M)$, their wedge product $\omega \wedge \eta \in \Omega^{p+q}(M)$, also denoted $\omega\eta$, is given by
(1.13)
$$(\omega \wedge \eta)(X_1, \ldots, X_{p+q}) = \sum_{\sigma} sign(\sigma)\omega(X_{\sigma(1)}, \ldots, X_{\sigma(p)}) \cdot \eta(X_{\sigma(p+1)}, \ldots, X_{\sigma(p+q)}),$$

where the sum is over all (p, q)-shuffles σ, i.e. all permutations σ with $\sigma(1) < \ldots < \sigma(p)$ and $\sigma(p+1) < \ldots < \sigma(p+q)$. Although this formula no longer makes sense when ω and η are both E-valued differential forms, it does make sense when one of them is E-valued and the other one is a usual form. The resulting operation makes

$$\Omega(M, E) = \bigoplus_{p} \Omega^p(M, E)$$

into a (left and right) module over $\Omega(M)$. Keeping in mind the fact that the spaces Ω are graded (i.e are direct sums indexed by integers) and the fact that the wedge products involved are compatible with the grading (i.e. $\Omega^p \wedge \Omega^q \subset \Omega^{p+q}$), we say that $\Omega(M)$ is a graded algebra and $\Omega(M, E)$ is a graded bimodule over $\Omega(M)$. As for the usual wedge product of forms, the left and right actions are related by[1]

$$\omega \wedge \eta = (-1)^{pq}\eta \wedge \omega \quad \forall\, \omega \in \Omega^p(M), \eta \in \Omega^q(M, E).$$

1.2. Connections on vector bundles

1.2.1. The definition

Throughout this section E is a vector bundle over a manifold M. Unlike the case of smooth functions on manifolds (which are sections of the trivial line bundle!), there is no canonical way of taking derivatives of sections of (an arbitrary) E along vector fields. That is where connections come in.

Definition 1.14. *A connection on E is a bilinear map ∇*

$$\mathcal{X}(M) \times \Gamma(E) \longrightarrow \Gamma(E), \quad (X, s) \mapsto \nabla_X(s),$$

satisfying

$$\nabla_{fX}(s) = f\nabla_X(s), \;\; \nabla_X(fs) = f\nabla_X(s) + L_X(f)s,$$

for all $f \in C^\infty(M)$, $X \in \mathcal{X}(M)$, $s \in \Gamma(E)$.

Remark 1.15. In the case when E is the trivial vector bundle of rank r, $M \times \mathbb{R}^r$ or $M \times \mathbb{C}^r$, one has to so-called canonical (flat) connection on the trivial bundle, denoted ∇^{can}, uniquely characterized by

$$\nabla_X(e_j) = 0,$$

where

$$e = \{e_1, \ldots, e_r\},$$

is its canonical frame (if E is not already trivialized, this says that any frame gives rise to a canonical connection). Actually, giving a connection on E is the same thing as giving an r by r matrix whose entries are 1-forms on M:

$$\omega := (\omega_i^j)_{i,j} \in M_r(\Omega^1(M)).$$

[1]Important: this is the first manifestation of what is known as the "graded sign rule": in an formula that involves graded elements, if two elements a and b of degrees p and q are interchanged, then the sign $(-1)^{pq}$ is introduced

This 1-1 correspondence is uniquely characterized by

$$\nabla_X(e_j) = \sum_{i=1}^{r} \omega_j^i(X) e_i.$$

(work out the details!). The canonical connection corresponds to the zero matrix.

Please be aware of our conventions: for the matrix $\omega = \{\omega_j^i\}_{i,j}$, the upper indices i count the rows, while the lower ones the columns:

$$\omega = \begin{pmatrix} \omega_1^1 & \dots & \omega_r^1 \\ \dots & \dots & \dots \\ \omega_1^r & \dots & \omega_r^r \end{pmatrix}$$

The other convention (switching the rows and the columns) would correspond to considering the transpose matrix ${}^t(\omega)$; that convention is taken in some text-books and accounts for the sign changes between our formulas involving ω and the ones in those text-books.

Exercise 9. Let E be a vector bundle over M, and assume that it sits inside a trivial vector bundle $M \times \mathbb{R}^k$ of some rank k; hence E_x is a vector sub-space of $\mathbb{R}^k$ for each x. We denote by

$$\mathrm{pr}_x : \mathbb{R}^k \to E_x$$

the orthogonal projection into E_x and we use the canonical connection ∇^{can} on the trivial bundle. Show that

$$\nabla_X(s)(x) := \mathrm{pr}_x(\nabla_X^{\mathrm{can}}(s)(x))$$

defines a connection on E. The same for complex vector bundles. Use this to exhibit a connection on the tautological line bundle γ^1 (see Example 1.7).

Exercise 10. Let E and E' be two vector bundles over M endowed with connections ∇ and ∇', respectively. We look at E^*, $E \oplus E'$ and $E \otimes E'$. Show that one has:

1. an induced connection ∇^* on E^*, defined by the Leibniz-type equation
$$L_X(\xi(s)) = \nabla_X^*(\xi)(s) + \xi(\nabla_X(s))$$
for all $s \in \Gamma(E)$, $\xi \in \Gamma(E^*)$ and $X \in \mathcal{X}(M)$.
2. an induced connection $\nabla \oplus \nabla'$ on $E \oplus E'$ given by
$$(\nabla \oplus \nabla')_X(s, s') = (\nabla_X(s), \nabla'_X(s')).$$
3. an induced connection $\nabla \otimes \nabla'$ on $E \oplus E'$ given by
$$(\nabla \otimes \nabla')_X(s \otimes s') = \nabla_X(s) \otimes s' + s \otimes \nabla'_X(s').$$

Exercise 11. Prove that any convex linear combination of two connections is again a connection, i.e., given ∇^1 and ∇^2 connections on E and ρ_1, ρ_2 smooth functions on M (the base manifold) such that $\rho_1 + \rho_2 = 1$, then

$$\nabla = \rho_1 \nabla^1 + \rho_2 \nabla^2$$

is also a connection.

1.2.2. Locality; connection matrices

Connections are local in the sense that, for a connection ∇ and $x \in M$,

$$\nabla_X(s)(x) = 0$$

for any $X \in \mathcal{X}(M)$, $s \in \Gamma(E)$ such that $X = 0$ or $s = 0$ in a neighborhood U of x. This can be checked directly, or can be derived from the remark that ∇ is a differential operator of order one in X and of order zero in f.

Locality implies that, for $U \subset M$ open, ∇ induces a connection ∇^U on the vector bundle $E|_U$ over U, uniquely determined by the condition

$$\nabla_X(s)|_U = \nabla^U_{X|_U}(s_U).$$

Choosing U the domain of a trivialization of E, with corresponding local frame $e = \{e_1, \ldots, e_r\}$, the previous remark shows that, over U, ∇ is uniquely determined by a matrix

$$\omega := (\omega^i_j)_{i,j} \in M_r(\Omega^1(U)).$$

This matrix is called the connection matrix of ∇ with respect to the local frame e and; when we want to emphasize the dependence on ∇ and e, aspect, we use the notation

$$\omega = \omega(\nabla, e) \in M_r(\Omega^1(U)). \tag{1.16}$$

Proposition 1.17. *Any vector bundle E admits a connection.*

Proof. Start with a partition of unity η_i subordinated to an open cover $\{U_i\}$ such that $E|_{U_i}$ is trivializable. On each $E|_{U_i}$ we consider a connection ∇^i (e.g., in the previous remark consider the zero matrix). Define ∇ by

$$\nabla_X(s) := \sum_i (\nabla_{X|_{U_i}})(\eta_i s).$$

□

1.2.3. More than locality: derivatives of paths

We have seen that ∇ is local: if we want to know $\nabla_X(s)$ at the point $x \in X$, then it suffices to know X and s in a neighborhood of x. However, much more is true.

Lemma 1.18. *Let ∇ be a connection on E and look at*

$$\nabla_X(s)(x_0),$$

with $X \in \mathcal{X}(M)$, $s \in \Gamma(E)$, $x_0 \in M$. This expression vanishes in each of the cases:

1. *s-arbitrary but $X(x_0) = 0$.*
2. *X-arbitrary but there exists*

$$\gamma : (-\epsilon, \epsilon) \to M \quad \text{with } \gamma(0) = x_0, \dot{\gamma}(0) = X_{x_0}$$

such that $s(\gamma(t)) = 0$ for all t near 0.

Proof. We deal with a local problem and we can concentrate on an open U containing x on which we have a given local frame e of E; ∇ will then be specified by its connection matrix. An arbitrary section s can now be written on U as

$$s = \sum_{i=1}^r f^i e_i$$

with $f^i \in C^\infty(U)$; on such a section we find using the Leibniz identity and then, using the connection matrix:

$$\nabla_X(s)(x) = \sum_i (df^i)(X_x)e_i(x) + \sum_{i,j} f^j(x)\omega_j^i(X_x)e_i(x).$$

It is clear that this is zero when $X(x) = 0$. In the second case we find

$$\nabla_X(s)(x) = \sum_i \frac{df^i \circ \gamma}{dt}(0)e_i(x) + \sum_{i,j} f^j(\gamma(0))\omega_j^i(X_x)e_i(x) \tag{1.19}$$

which clearly vanishes under the condition that $f^i(\gamma(t)) = 0$ for t near 0. □

The first type of condition in the lemma tells us that, given $s \in \Gamma(E)$, it makes sense to talk about

$$\nabla_{X_x}(s) \in E_x$$

for all $X_x \in T_xM$. In other words, ∇ can be reinterpreted as an operator

$$d_\nabla : \Gamma(E) \longrightarrow \Omega^1(M, E), \quad d_\nabla(s)(X) := \nabla_X(s).$$

The properties of ∇ immediately imply that d_∇ is linear (with respect to the multiplication by scalars) and satisfies the Leibniz identity, i.e.

$$d_\nabla(fs) = fd_\nabla(s) + df \otimes s$$

for all $f \in C^\infty(M)$ and $s \in \Gamma(E)$. Of course, ∇ can be recovered from d_∇ and this gives rise to another way of looking at connections.

Exercise 12. Show that $\nabla \longleftrightarrow d_\nabla$ gives a 1-1 correspondence between connections ∇ on E and linear operators from $\Gamma(E)$ to $\Omega^1(M, E)$ satisfying the Leibniz identity.

This will give rise to the re-interpretation of connections as DeRham-like operators, discussed a bit later. We now concentrate on the second type of condition in the lemma, and we re-interpret it more conceptually. Given a path

$$\gamma : I \longrightarrow M$$

(i.e. a smooth map, defined on some interval I, typically $[0, 1]$ or of type $(-\epsilon, \epsilon)$), by a path in E above γ we mean any path $u : I \longrightarrow E$ with the property that

$$u(t) \in E_{\gamma(t)} \quad \forall\ t \in I.$$

One way to produce such paths above γ is by using sections of E: any section $s \in \Gamma(E)$ induces the path

$$s \circ \gamma : I \longrightarrow E$$

above γ. The previous lemma implies that the expression

$$\nabla_{\dot\gamma}(s)(\gamma(t))$$

makes sense, depends on the path $s \circ \gamma$ and defines is a path above γ. It is denoted

$$\frac{\nabla(s \circ \gamma)}{dt}.$$

Slightly more generally, for any path $u : I \longrightarrow E$ above γ one can define the new path above γ

$$\frac{\nabla u}{dt} : I \longrightarrow E.$$

Locally with respect to a frame e, writing $u(t) = \sum_j u^j(t)e_j(\gamma(t))$, the formula is just the obvious version of (1.19) and we find the components

$$\left(\frac{\nabla u}{dt}\right)^i = \frac{du^i}{dt}(t) + \sum_j u^j(t)\omega^i_j(\dot\gamma(t)). \tag{1.20}$$

Exercise 13. On the tangent bundle of $\mathbb{R}^n$ consider the connection

$$\nabla_X Y = \sum_i X(Y^i)\frac{\partial}{\partial x_i}.$$

Let γ be a curve in $\mathbb{R}^n$ and let $\frac{\nabla}{dt}$ be the derivative induced along γ by the connection. What is $\frac{\nabla\dot\gamma}{dt}$?

1.2.4. Parallel transport

One of the main use of connections comes from the fact that a connection ∇ on E can be used to move from one fiber of E to another, along paths in the base. This is the so called parallel transport. To explain this, let us return to paths $u : I \longrightarrow E$. We say that u is parallel (with respect to ∇) if

$$\frac{\nabla u}{dt} = 0 \quad \forall\, t \in I.$$

Lemma 1.21. *Let ∇ be a connection on the vector bundle E and $\gamma : I \longrightarrow M$ a curve in M, $t_0 \in I$. Then for any $u_0 \in E_{\gamma(t_0)}$ there exists and is unique a parallel path above γ, $u : I \longrightarrow E$, with $u(t_0) = u_0$.*

Proof. We can proceed locally (also because the uniqueness locally implies that the local pieces can be glued), on the domain of a local frame e. By formula (1.20), we have to find

$$u = (u^1, \ldots, u^r) : I \longrightarrow \mathbb{R}^r$$

satisfying

$$\frac{du^i}{dt}(t) = -\sum_j u^j(t)\omega^i_j(\dot\gamma(t)),\ u(0) = u_0.$$

In a matricial form (with u viewed as a column matrix), writing $A(t)$ for the matrix $-\omega(\dot\gamma(t))$, we deal with the equation

$$\dot u(t) = A(t)u(t), \quad u(t_0) = u_0$$

and the existence and uniqueness is a standard result about first order linear ODE's. □

Definition 1.22. *Given a connection ∇ on E and a curve $\gamma : I \longrightarrow M$, $t_0, t_1 \in I$, the parallel transport along γ (with respect to ∇) from time t_0 to time t_1 is the map*

$$T^{t_0,t_1}_\gamma : E_{\gamma(t_0)} \longrightarrow E_{\gamma(t_1)}$$

which associates to $u_0 \in E_{\gamma(t_0)}$ the vector $u(t_1) \in E_{\gamma(t_1)}$, where u is the unique parallel curve above γ with $u(t_0) = u_0$.

Exercise 14. Show that
1. each T^{t_0,t_1}_γ is a linear isomorphism.
2. $T^{t_1,t_2}_\gamma \circ T^{t_0,t_1}_\gamma = T^{t_0,t_2}_\gamma$ for all $t_0, t_1, t_2 \in I$.

Then try to guess how ∇ can be recovered from all these parallel transports (there is a natural guess!). Then prove it.

Exercise 15. Show that any vector bundle on a contractible manifold is trivializable.

1.2.5. Connections as DeRham-like operators

Next, we point out a slightly different way of looking at connections, in terms of differential forms on M. Recall that the standard DeRham differential d acts on the space $\Omega(M)$ of differential forms on M, increasing the degree by one

$$d : \Omega^*(M) \longrightarrow \Omega^{*+1}(M),$$

and satisfying the Leibniz identity:

$$d(\omega \wedge \eta) = d(\omega) \wedge \eta + (-1)^{|\omega|}\omega \wedge d(\eta),$$

where $|\omega|$ is the degree of ω[2], and is a differential (i.e. $d \circ d = 0$). Locally, writing ω as in (1.10) in 1.1.8, we have

$$d\omega = \sum_{i} \sum_{i_1,\ldots,i_p} \frac{\partial f^{i_1,\ldots,i_p}}{\partial \chi_k} d\chi_i d\chi_{i_1} \ldots d\chi_{i_p}.$$

Globally, thinking of ω as a $C^\infty(M)$-linear map as in (1.9), one has

$$\begin{aligned} d(\omega)(X_1, \ldots, X_{p+1}) &= \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots X_{p+1})) \\ &+ \sum_{i=1}^{p+1} (-1)^{i+1} L_{X_i}(\omega(X_1, \ldots, \hat{X}_i, \ldots, X_{p+1})). \end{aligned} \tag{1.23}$$

where L_X denotes the Lie derivative along the vector field X.

Let us now pass to differential forms with coefficients in a vector bundle E (see 1.1.8). The key remark here is that, while there is no canonical (i.e. free of choices) analogue of DeRham differential on $\Omega(M, E)$, connections are precisely the piece that is needed in order to define such operators. Indeed, assuming that ∇ is a connection on E, and thinking of forms $\omega \in \Omega^p(M, E)$ as $C^\infty(M)$-multilinear maps as in 1.11, we see that the previous formula for the DeRham differential does makes sense if we replace the Lie derivatives L_{X_i} by ∇_{X_i}. Hence one has an induced operator

$$d_\nabla : \Omega^\bullet(M, E) \longrightarrow \Omega^{\bullet+1}(M, E).$$

As in the case of DeRham operator, d_∇ satisfies the Leibniz identity

$$d_\nabla(\omega \wedge \eta) = d(\omega) \wedge \eta + (-1)^{|\omega|}\omega \wedge d_\nabla(\eta)$$

for all $\omega \in \Omega(M)$, $\eta \in \Omega(M, E)$.

Note that, in degree zero, d_∇ acts on $\Omega^0(M, E) = \Gamma(E)$ and it coincides with the operator d_∇ previously discussed. Moreover, d_∇ defined on $\Omega^\bullet(M, E)$ is unquely determined by what it does on the degree zero part and the fact that d_∇ satisfies the Leibniz identity.

[2]Note: the sign in the formula agrees with the graded sign rule: we interchange d which has degree 1 and ω

Exercise 16. As a continuation of Exercise 12 conclude that one has a 1-1 correspondence between connections ∇ on E and operators d_∇ as above i.e. defined on $\Omega^\bullet(M, E)$, increasing the degree by one and satisfying the Leibniz identity).

Note also that, as for the DeRham operator, d_∇ can be described locally, using the connection matrices. First of all, if U is the domain of a local frame $e = \{e_1, \ldots, e_r\}$ with connection matrix ω, then one can write

$$d_\nabla(e_j) = \sum_{i=1}^{r} \omega_j^i e_j, \tag{1.24}$$

Assume now that U is also the domain of a coordinate chart $(U, \chi_1, \ldots, \chi_n)$. Representing $\omega \in \Omega^p(M, E)$ locally as in (1.12), the Leibniz identity gives the formula

$$d_\nabla(\omega) = \sum_{i_1,\ldots,i_p} (-1)^p dx_{i_1} \ldots dx_{i_p} \otimes d_\nabla(e^{i_1,\ldots,i_p}).$$

hence it suffices to describe d_∇ on sections of E. The same Leibniz formula implies that it suffices to describe d_∇ on the frame e- and that what (1.24) does.

Exercise 17. Consider the tautological line bundle γ^1 over $\mathbb{C}P^1$ (see Example 1.7) and consider the connection that arise from Exercise 9. Consider the chart of $\mathbb{C}P^1$ given by

$$U = \{[1 : z] : z \in \mathbb{C}\} \subset \mathbb{C}P^1$$

over which we consider the local frame of γ^1 (i.e. just a nowhere vanishing local section of γ^1, since γ^1 is of rank one) given by

$$e([1 : z]) = e_1 + ze_2$$

where $\{e_1, e_2\}$ is the canonical basis of $\mathbb{C}^2$. Show that

$$d_\nabla(e) = \frac{\overline{z} \cdot dz}{1 + |z|^2} e.$$

1.3. Curvature

1.3.1. The definition

Recall that, for the standard Lie derivatives of functions along vector fields,

$$L_{[X,Y]} = L_X L_Y(f) - L_Y L_X(f).$$

Of course, this can be seen just as the definition of the Lie bracket $[X, Y]$ of vector fields but, even so, it still says something: the right hand side is a derivation on f (i.e., indeed, it comes from a vector field). The similar formula for connections fails dramatically (i.e. there are few vector bundles which admit a connection for which the analogue of this formula holds). The failure is measured by the curvature of the connection.

Proposition 1.25. *For any connection ∇, the expression*

$$k_\nabla(X, Y)s = \nabla_X \nabla_Y(s) - \nabla_Y \nabla_X(s) - \nabla_{[X,Y]}(s), \tag{1.26}$$

is $C^\infty(M)$-linear in the entries $X, Y \in \mathcal{X}(M)$, $s \in \Gamma(E)$. Hence it defines an element

$$k_\nabla \in \Gamma(\Lambda^2 T^*M \otimes End(E)) = \Omega^2(M; End(E)), \tag{1.27}$$

called the curvature of ∇.

Proof. It follows from the properties of ∇. For instance, we have

$$\begin{aligned}\nabla_X\nabla_Y(fs) &= \nabla_X(f\nabla_Y(s)+L_Y(f)s)\\ &= f\nabla_X\nabla_Y(s)+L_X(f)\nabla_Y(s)+L_X(f)\nabla_Y(s)+L_XL_Y(f)s,\end{aligned}$$

and the similar formula for $\nabla_X\nabla_Y(fs)$, while

$$\nabla_{[X,Y]}(fs) = f\nabla_{[X,Y]}(s)+L_{[X,Y]}(f)s.$$

Hence, using $L_{[X,Y]} = L_XL_Y - L_YL_X$, we deduce that

$$k_\nabla(X,Y)(fs) = fk_\nabla(X,Y)(s),$$

and similarly the others. $\square$

Exercise 18. This is a continuation of Exercise 10, hence we assume the same notations. For two finite dimensional vector spaces V and V' and $A \in \mathrm{End}(V)$ and $A' \in \mathrm{End}(V')$, we consider

$$A^* \in \mathrm{End}(V^*) \quad \text{given by } A^*(\xi)(v) = \xi(A(v)),$$

$$A \oplus A' \in \mathrm{End}(V\oplus V') \quad \text{given by } (A\oplus A')(v,v') = (A(v),A'(v')),$$

$$A \otimes A' \in \mathrm{End}(V\otimes V') \quad \text{given by } (A\otimes A')(v\otimes v') = A(v)\otimes A'(v').$$

We keep the same notations for endomorphisms of vector bundles. Show that the curvatures of ∇^*, $\nabla\oplus\nabla'$ and $\nabla\otimes\nabla'$ are given by

$$k_{\nabla^*}(X,Y) = k_\nabla(X,Y)^*,$$

$$k_{\nabla\oplus\nabla'}(X,Y) = k_\nabla(X,Y)\oplus k_{\nabla'}(X,Y),$$

$$k_{\nabla\otimes\nabla'}(X,Y) = \mathrm{Id}_E\otimes k_{\nabla'}(X,Y) + k_\nabla(X,Y)\otimes \mathrm{Id}_{E'}.$$

Remark 1.28. One can express the curvature locally, with respect to a local frame $e = \{e_1,\ldots,e_r\}$ of E over an open U, as

$$k_\nabla(X,Y)e_j = \sum_{j=1}^{r} k^i_j(X,Y)e_i,$$

where $k^i_j(X,Y) \in C^\infty(U)$ are smooth functions on U depending on $X, Y \in \mathcal{X}(M)$. The previous proposition implies that each k^j_i is a differential form (of degree two). Hence k_∇ is locally determined by a matrix

$$k = k(\nabla, e) := (k^i_j)_{i,j} \in M_r(\Omega^2(U)),$$

called the curvature matrix of ∇ over U, with respect to the local frame e. A simple computation (exercise!) gives the explicit formula for k in terms of the connection matrix ω:

$$k = d\omega + \omega\wedge\omega,$$

where $\omega\wedge\omega$ is the matrix of 2-forms given by

$$(\omega\wedge\omega)^i_j = \sum_k \omega^i_k\wedge\omega^k_j.$$

1.3.2. The curvature as failure of "$\frac{d^2}{dsdt} = \frac{d^2}{dtds}$"

The fact that the curvature measure the failure of the connections to satisfy "usual formulas", which is clear already from the definition, can be illustrated in several other ways. For instance, while for functions $\gamma(t,s)$ of two variables, taking values in some $\mathbb{R}^n$, the order of taking derivatives does not matter, i.e.

$$\frac{d^2\gamma}{ds\,dt} = \frac{d^2\gamma}{dt\,ds},$$

the curvature of a connection shows up as the failure of such an identity for the operator $\frac{\nabla}{dt}$ induced by ∇. To explain this, assume that ∇ is a connection on a vector bundle E over M,

$$\gamma = \gamma(t,s) : I_1 \times I_2 \to M$$

is a smooth map defined on the product of two intervals $I_1, I_2 \subset \mathbb{R}$ and

$$u = u(t,s) : I_1 \times I_2 \to E$$

smooth covering γ. Applying $\frac{\nabla}{dt}$ and then$\frac{\nabla}{ds}$ obe obtains a new

$$\frac{\nabla^2 u}{ds\,dt} : I_1 \times I_2 \to E$$

covering u.

Exercise 19. In the setting described above, show that

$$\frac{\nabla^2 u}{ds\,dt} - \frac{\nabla^2 u}{dt\,ds} = k_\nabla(\frac{d\gamma}{ds}, \frac{d\gamma}{dt}).$$

Corollary 1.29. *Assume that ∇ is a flat connection on a vector E, i.e. with the property that $k_\nabla = 0$. Then, for $x, y \in M$, looking at paths γ in M from x to y, the induced parallel transport*

$$T_\gamma^{0,1} : E_x \to E_y$$

only depends on the path-homotopy class of gamma.

This means that, if γ_0, γ_1 are two paths which can be joined by a smooth family of paths $\{\gamma_s\}_{s\in[0,1]}$ starting at x and ending at y, then $T_{\gamma_0}^{0,1} = T_{\gamma_1}^{0,1}$. Of ocurse, the smoothness of the family means that

$$\gamma(t,s) := \gamma_s(t)$$

is a smooth map from $[0,1] \times [0,1]$ to M.

Proof. Consider a smooth family $\{\gamma_s\}_{s\in[0,1]}$ as above, encoded in $\gamma(t,s) = \gamma_s(t)$. The conditions on γ are:

$$\gamma(0,s) = x, \gamma(1,s) = y, \quad (\gamma_s \text{ starts at } x \text{ and ends at } y),$$
$$\gamma(t,0) = \gamma_0(t), \gamma(t,1) = \gamma_1(t).$$

We fix $u_0 \in E_x$ and we are going to prove that $T_{\gamma_0}^{0,1}(u_0) = T_{\gamma_1}^{0,1}(u_0)$. For that, consider $u(t,s) = T_{\gamma_s}^{0,t}(u)$. By the definition of parallel transport, $\frac{\nabla}{u}dt = 0$. Using the flatness of ∇ and the previous exercise we see that $v := \frac{\nabla u}{ds}$ must satisfy

$$\frac{\nabla v}{dt} = 0.$$

But, for $t = 0$ and any s one has $u(0,s) = u_0$ hence $v(0,s) = 0$. In other words, is we fix s, $v(\cdot, s)$ is a path in E that is parallel w.r.t. ∇ and starts at 0; by uniqueness of the parallel transport we deduce that $v = 0$, i.e. $\frac{\nabla u}{ds} = 0$ at all (t, s). We use this at $t = 1$, where the situation is more special bacause all $u(1, s)$ take value in one single fiber- namely E_y; moreover, from the definition of $\frac{\nabla}{ds}$, we see that $\frac{\nabla u}{ds}(1, s)$ is the usual derivative of the resulting path

$$[0, 1] \ni s \mapsto u(1, s) \in E_y.$$

We deduce that this last path is constant in s; in particular, $u(1, 0) = u(1, 1)$, i.e. $T^{0,1}_{\gamma_0}(u_0) = T^{0,1}_{\gamma_1}(u_0)$. $\square$

Recall that, fixing a point $x \in M$, the path-homotopy classes of paths in M that start and end at x (loops at x) form a group- the fundamental group of M with base point x, denoted $\pi(M, x)$ (the group operation is given by the concatenation of paths). The previous corollary implies that the parallel transport defines a map

$$\rho : \pi(M, x) \to GL(E_x), \quad [\gamma] \mapsto T^{0,1}_{\gamma},$$

where $GL(E_x)$ is the group of all linear isomorphisms from E_x to itself. Moreover, it is not difficult to see that this is actually a group homomorphism. In other words, ρ is a representation of the group $\pi(M, x)$ on the vector space E_x. This is called the monodromy representation of the flat connection (at the base point x).

Exercise 20. Show that a vector bundle is trivializable if and only if it admits a flat connection ∇ whose monodromy representation (at some base point) is trivial.

1.3.3. The curvature as failure of $d^2_{\nabla} = 0$

There is another interpretation of the curvature, in terms of forms with values in $End(E)$. While ∇ defines the operator d_{∇} which is a generalization of the DeRham operator d, it is very rarely that it squares to zero (as d does). Again, k_{∇} measure this failure. To explain this, we first look more closely to elements

$$K \in \Omega^p(M, End(E)).$$

The wedge product formula (1.13) has a version when $\omega = K$ and $\eta \in \Omega^q(M, E)$:

$$(K \wedge \eta)(X_1, \ldots, X_{p+q}) = \sum_{\sigma} sign(\sigma) K(X_{\sigma(1)}, \ldots, X_{\sigma(p)})(\eta(X_{\sigma(p+1)}, \ldots, X_{\sigma(p+q)})),$$

Any such K induces a linear map

$$\hat{K} : \Omega^{\bullet}(M, E) \longrightarrow \Omega^{\bullet+p}(M, E), \ \hat{K}(\eta) = K \wedge \eta.$$

For the later use not also that the same formula for the wedge product has an obvious version also when applied to elements $K \in \Omega^p(M, End(E))$ and $K' \in \Omega^q(M, End(E))$, giving rise to operations

$$\wedge : \Omega^p(M, End(E)) \times \Omega^q(M, End(E)) \longrightarrow \Omega^{p+q}(M, End(E)) \tag{1.30}$$

which make $\Omega(M, End(E))$ into a (graded) algebra.

Exercise 21. Show that $\hat{K}$ is an endomorphism of the graded (left) $\Omega(M)$-module $\Omega(M, E)$ i.e., according to the graded sign rule (see the previous footnotes):

$$\hat{K}(\omega \wedge \eta) = (-1)^{pq} \omega \wedge K(\eta),$$

for all $\omega \in \Omega^q(M)$. Moreover, the correspondence $K \mapsto \hat{K}$ defines a bijection

$$\Omega^p(M, End(E)) \cong End^p_{\Omega(M)}(\Omega(M, E))$$

between $\Omega^p(M, End(E))$ and the space of all endomorphisms of the graded (left) $\Omega(M)$-module $\Omega(M, E)$ which rise the degree by p.

Finally, via this bijection, the wedge operation (1.30) becomes the composition of operators, i.e.

$$\widehat{K \wedge K'} = \hat{K} \circ \hat{K}'$$

for all $K, K' \in \Omega(M, End(E))$.

Due to the previous exercise, we will tacitly identify the element K with the induced operator m_K. For curvature of connections we have

Proposition 1.31. *If ∇ is a connection on E, then*

$$d_\nabla^2 = d_\nabla \circ d_\nabla : \Omega^\bullet(M, E) \longrightarrow \Omega^{\bullet+2}(M, E)$$

is given by

$$d_\nabla^2(\eta) = k_\nabla \wedge \eta$$

for all $\eta \in \Omega^*(M; E)$, and this determines k_∇ uniquely.

Proof. First of all, d_∇ is $\Omega(M)$-linear: for $\omega \in \Omega^p(M)$ and $\eta \in \Omega^p(M, E)$,

$$\begin{aligned} d_\nabla^2(\omega \wedge \eta) &= d_\nabla(d(\omega) \wedge \eta + (-1)^p \omega \wedge d_\nabla(\eta)) \\ &= [d^2(\omega) \wedge \eta + (-1)^{p+1} d(\omega) \wedge d_\nabla(\eta)] + (-1)^p [d(\omega) \wedge d_\nabla(\eta) + (-1)^p \omega \wedge d_\nabla^2(\eta)) \\ &= \omega \wedge d_\nabla(\eta). \end{aligned}$$

Hence, by the previous exercise, it comes from multiplication by an element $k \in \Omega^2(M)$. Using the explicit Koszul-formula for d_∇ to compute d_∇^2 on $\Gamma(E)$, we see that $d_\nabla^2(s) = k_\nabla \wedge s$ for all $s \in \Gamma(E)$. We deduce that $k = k_\nabla$. □

1.3.4. More on connection and curvature matrices; the first Chern class

Recall that, given a connection ∇ on a vector bundle E, ∇ is locally determined by its conection matrices. More precisely, for any local frame

$$e = \{e_1, \ldots, e_r\}$$

over some open $U \subset M$, one can write

$$\nabla_X(e_j) = \sum_{i=1}^{r} \omega_j^i(X) e_i$$

and ∇ is uniquely determined by the matrix

$$\omega = \omega(\nabla, e) := (\omega_i^j)_{i,j} \in M_r(\Omega^1(U)).$$

Similarly, one can talk about the connection matrix

$$k = k(\nabla, e) := (k_j^i)_{i,j} \in M_r(\Omega^2(U)),$$

which describes the curvature of ∇ w.r.t. the local frame e by

$$k(X, Y)(e_j) = \sum_{i=1}^{r} k_j^i(X, Y) e_i.$$

As we have already remarked, the curvature matrix can be obtained from the connection matrix by the formula

$$k = d\omega + \omega \wedge \omega,$$

where $\omega \wedge \omega$ is the matrix of 2-forms given by

$$(\omega \wedge \omega)^i_j = \sum_k \omega^i_k \wedge \omega^k_j.$$

Here we would like to see how these matrices change when one changes the (local) frames.

Lemma 1.32. *Let ∇ be a connection on E. Let $e = \{e_1, \ldots, e_r\}$ be a local frame of E over an open U and let ω and k be the associated connection matrix and curvature matrix, respectively. Let $e' = \{e'_1, \ldots, e'_r\}$ be another local frame of E over some open U' and let ω' and k' be the associated connection and curvature matrix of ∇. Let*

$$g = (g^j_i) \in M_n(C^\infty(U \cap U'))$$

be the matrix of coordinate changes from e to e', i.e. defined by:

$$e'_i = \sum_{j=1}^{r} g^j_i e_j$$

over $U \cap U'$. Then, on $U \cap U'$,

$$\boxed{\omega' = g^{-1} \cdot (dg) + g^{-1} \cdot \omega \cdot g.}$$

$$\boxed{k' = g^{-1} \cdot k \cdot g.}$$

Proof. Using formula (1.24) for d_∇ we have:

$$\begin{aligned} d_\nabla(e'_i) &= d_\nabla(\sum_l g^l_i e_l) \\ &= \sum_l d(g^l_i)e_l + \sum_{l,m} g^l_i \theta^m_l e_m, \end{aligned}$$

where for the last equality we have used the Leibniz rule and the formulas defining θ. Using the inverse matrix $g^{-1} = (\overline{g}^i_j)_{i,j}$ we change back from the frame e to e' by $e_j = \sum_i \overline{g}^i_j \omega_i$ and we obtain

$$d_\nabla(e'_i) = \sum_{l,j} d(g^l_i)\overline{g}^j_l e'_j + \sum_{l,m,j} g^l_i \theta^m_l \overline{g}^j_m e'_j.$$

Hence

$$(\theta')^j_i = \sum_l d(g^l_i)\overline{g}^j_l + \sum_{l,m} g^l_i \theta^m_l \overline{g}^j_m,$$

i.e. the first formula in the statement. To prove the second equation, we will use the formula (refinvariance) which expresses k in terms of θ. We have

$$d\theta' = d(g^{-1}dg + g^{-1}\theta g) = d(g^{-1})d(g) + d(g^{-1})\theta g + g^{-1}d(\theta)g - g^{-1}\theta d(g). \tag{1.33}$$

For $\theta' \wedge \theta'$ we find

$$g^{-1}dg \wedge g^{-1}d(g) + g^{-1}dg \wedge g^{-1}\theta g + g^{-1}\theta g \wedge g^{-1}dg + g^{-1}\theta g \wedge g^{-1}\theta g.$$

Since

$$g^{-1}dg = d(g^{-1}g) - d(g^{-1})g = -d(g^{-1})g,$$

the expression above equals to

$$-d(g^{-1}) \wedge d(g) - d(g^{-1}) \wedge \theta g + g^{-1}\theta \wedge d(g) + g^{-1}\omega \wedge \omega g.$$

Adding up with (1.33), the first three term cancel and we are left with:

$$k' = d\omega' + \omega' \wedge \omega' = g^{-1}(d\omega + \omega \wedge \omega)g = g^{-1}kg.$$

$\square$

Exercise 22. Show that the formula $k' = g^{-1}kg$ is also a direct consequence of the fact that k_∇ is defined as a global section of $\Lambda^2 T^*M \otimes End(E)$.

Since the curvature matrix does not depend on the local frame "up to conjugation", it follows that any expression that is invariant under conjugation will produce a globally defined form on M. The simplest such expression is obtained by applying the trace:

$$Tr(k) = \sum_i k_i^i \in \Omega^2(U);$$

indeed, it does have the fundamental property that

$$Tr(gkg^{-1}) = Tr(k).$$

It follows immediately that, if k' corresponds to another local frame e' over U', then $Tr(k) = Tr(k')$ on the overlap $U \cap U'$. Hence all these pieces glue to a global 2-form on M, denoted

$$Tr(k_\nabla) \in \Omega^2(M). \tag{1.34}$$

This construction can be looked at a bit differently: for any finite dimensional vector space V one can talk about the trace map

$$Tr : \mathrm{End}(V) \to \mathbb{R} \text{ (or } \mathbb{C} \text{ in the complex case)}$$

defined on the space $\mathrm{End}(V)$ of all linear maps from V to itself; applying this to the fibers of E we find a trace map

$$Tr : \Omega^2(M, \mathrm{End}(E)) \to \Omega^2(M),$$

so that (1.34) is indeed obtained from k_∇ by applying Tr. Of course, $\Omega^\bullet(M)$ denotes the space of differential forms with coefficients in the field of scalars ($\mathbb{R}$ or $\mathbb{C}$).

Theorem 1.35. *For any vector bundle E over M,*

1. *$Tr(k_\nabla) \in \Omega^2(M)$ is a closed differential form, whose cohomology class only depends on the vector bundle E and not on the connection ∇ .*
2. *In the real case, the cohomology class of $Tr(k_\nabla)$ is actually zero.*
3. *in the complex case,*

$$c_1(E) := \left[\frac{1}{2\pi i} Tr(k_\nabla)\right] \in H^2(M, \mathbb{C})$$

   is actually a real cohomology class. It is called the first Chern class of E.
4. *For any two complex vector bundles E and E' one has*

$$c_1(\overline{E}) = -c_1(E),$$

$$c_1(E \oplus E') = c_1(E) + c_1(E'),$$

$$c_1(E \otimes E') = rank(E) \cdot c_1(E') + rank(E') \cdot c_1(E).$$

Proof. We will use the following fundamental property of the trace map:

$$Tr([A, B]) = 0$$

for any two matrices with scalar entires, $A, B \in \nabla \times \nabla$, where $[A, B] = AB - BA$. The same formula continues to hold for matrices with entries differential forms on M,

$$A \in \mathcal{M}_{r\times r}(\Omega^p(M)), B \in \mathcal{M}_{r\times r}(\Omega^q(M))$$

provided we use the graded commutators:

$$[A, B] = AB - (-1)^{pq}BA$$

(check this!). The fact that $Tr(k_\nabla)$ is closed can be checked locally, for which we can use local frames and the associated connection and curvature matrices. Then, using $k = d\omega + \omega \wedge \omega$ and $d^2 = 0$ we find

$$dk = d(\omega) \wedge \omega - \omega \wedge d(\omega);$$

writing $d\omega = k - \omega \wedge \omega$ we find

$$dk = k \wedge \omega - \omega \wedge k = [k, \omega],$$

therefore, using the fundamental property of the trace,

$$dTr(k) = Tr(dk) = Tr([k, \omega]) = 0.$$

To check that the class does not depend on ∇, we assume that we have another connection ∇'. With respect to a (any) local frame over some U we obtain two connection matrices ω and ω' and similarly two curvature matrices. Consider

$$\alpha := \omega' - \omega \in \mathcal{M}_{r\times r}(\Omega^1(U)).$$

Note that the previous lemma implies that, if one changes the local frame by another one then α changes in a similar way as k: by conjugation by g. In particular, the resulting 1-forms $Tr(\alpha) \in \Omega^1(U)$ will glue to a globally defined 1-form

$$Tr(\alpha) \in \Omega^1(M).$$

(as for the curvature, this form can be interpreted more directly, but a bit more abstractly: note that the expression $\nabla'_X(s) - \nabla_X(s) \in \Gamma(E)$ is $C^\infty(M)$-linear in all entries, hence it defines a tensor, usually denoted

$$\nabla' - \nabla \in \Gamma(T^*M \otimes \mathrm{End}(E)) = \Omega^1(M, \mathrm{End}(E)).$$

With this, the previous form is just $Tr(\nabla' - \nabla)$).

We claim that $Tr(k'_{\nabla'}) - Tr(k_\nabla) = d\alpha$. Again, this can be checked locally, using a local frame and the resulting matrices. Using $\omega' = \omega + \alpha$ we compute k' and we find

$$k' = k + d\alpha + \alpha \wedge \omega + \omega \wedge \alpha + \alpha \wedge \alpha.$$

This can be re-written using the graded commutators as

$$k' = k + d\alpha + [\alpha, \omega] + \frac{1}{2}[\alpha, \alpha].$$

Applying the trace and its fundamental property we find that, indeed,

$$Tr(k') = Tr(k) + d\alpha.$$

This closes the proof of 1. The proofs of 2 and 3 are postponed till after the discussion of metrics (see Exercise 30). For the first equality in 4 note that a connection ∇ on E serves also as one on $\overline{E}$, hence the only difference in computing $c_1(\overline{E})$ comes from the way one computes the trace $Tr_V : \mathrm{End}(V) \to \mathbb{C}$ for a finite dimensional vector space, when one replaces V by its conjugate $\overline{V}$; what happens is that

$$Tr_{\overline{V}}(A) = \overline{Tr_V(A)}.$$

Due also to the presence of the $2\pi i$ we find that

$$c_1(\overline{E}) = [\frac{1}{2\pi i}\overline{Tr(k_\nabla)}] = -\overline{[\frac{1}{2\pi i}Tr(k_\nabla)]} = -\overline{c_1(E)}.$$

Since $c_1(E)$ is a real class, we deduce that $c_1(\overline{E}) = -c_1(E)$.

For the last two equalities in 4 we use the connections $\nabla \oplus \nabla'$ and $\nabla \otimes \nabla'$ from Exercise 10 and 18. Choosing local frames e and e' for E and E' (over some U) and computing the connection matrix corresponding to the resulting local frame $e \oplus e'$ we find the direct sum of the connection matrices of ∇ and ∇':

$$(\omega \oplus \omega') = \begin{pmatrix} \omega & 0 \\ 0 & \omega' \end{pmatrix} \in \mathcal{M}_{r+r',r+r'}(\Omega^1(U))$$

and similarly for the curvature matrices; applying the trace, we find that

$$Tr(k_{\nabla\oplus\nabla'}) = Tr(k_\nabla + Tr(k_{\nabla'}).$$

Of course, one could have also proceeded more globally, using the second equality from Exercise 18 to which we apply the global trace Tr. The main point of this argument was that, for two finite dimensional vector spaces V and V',

$$Tr : \mathrm{End}(V \oplus V') \to \mathbb{C}$$

has the property that $Tr(A \oplus A') = Tr(A) + Tr(A')$ for any $A \in \mathrm{End}(V)$, and $A' \in \mathrm{End}(V')$. The computation for the tensor product connection is completely similar. The main point is that for any finite dimensional vector spaces V and V' one has $Tr(Id_V) = \dim(V)$ and

$$Tr : \mathrm{End}(V \otimes V') \to \mathbb{C}$$

satisfies

$$Tr(A \otimes B) = Tr(A) \cdot Tr(B).$$

Applying this to the last equality from Exercise 18, one obtains the desired equality. □

Exercise 23. Consider the tautological line bundle γ^1 over $\mathbb{C}P^1$. Continue the computations from Exercise 17 with the computation of the curvature matrix (just a 2-form in this case) and then show that

$$\int_{\mathbb{C}P^1} c_1(\gamma^1) = -1.$$

Deduce that γ^1 is not trivializable, is not isomorphic to its conjugate and is not the complexification of a real vector bundle.

For the curious reader. The first Chern class $c_1(E)$ is just one of the interesting invariants that one can associate to vector bundles and which can be used to distinguish vector bundle. At the heart of its construction was the trace map and its fundamental property that it is invariant under conjugation: $Tr(g \cdot k \cdot g^{-1}) = Tr(k)$ for any invertible matrix g. Generalizations of this can be obtained replacing the trace map $A \mapsto Tr(A)$ by $A \mapsto Tr(A^k)$ (for $k \geq 0$ any integer). The most general possibility is to work with arbitrary "invariant polynomials". Let us work over the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. We denote by $I_r(\mathbb{F})$ the space of all functions

$$P : M_r(\mathbb{F}) \longrightarrow \mathbb{F}$$

which are polynomial (in the sense that $P(A)$ is a polynomial in the entries of A), and which are invariant under the conjugation, i.e.

$$P(gAg^{-1}) = P(A)$$

for all $A \in M_r(\mathbb{F})$, $g \in Gl_r(\mathbb{F})$. In complete analogy with the trace (which is an such an invariant polynomial, of degree 1!), for any $P \in I_r(\mathbb{F})$, an invariant polynomial of degree d, and any vector bundle over $\mathbb{F}$ of rank r, one can choose a connection ∇ on E and one can apply P to its curvature to obtain a differential form

$$P(k_\nabla) \in \Omega^2 d(M).$$

Moreover, this form is closed and its cohomology class does not depend on the choice of ∇; it is denoted by

$$P(E) \in H^{2d}(M)$$

(the P-characteristic class of E).

Which are the possible/interesting choices for P? It is worth to first remark that $I_r(\mathbb{F})$ is an algebra (the product of two invariant polynomials is invariant). Moreover, for two invariant polynomials P and Q, it is clear that

$$(P \cdot Q)(E) = P(E) \cdot Q(E)$$

for any vector bundle E of rank r; hence, in some sense, it suffices to apply the previous construction to Ps that form a basis of the algebra $I_r(\mathbb{F})$. There are several interesting bases that one can use.

Example 1.36. One has the polynomial functions

$$\sigma_p : M_r(\mathbb{F}) \longrightarrow \mathbb{F}$$

defined by the equation

$$det(I + tA) = \sum_{d=0}^{r} \sigma_d(A)t^d.$$

For instance, $\sigma_1 = \Sigma_1$ is just the trace while $\sigma_r(A) = det(A)$. One can prove that

$$I_r(\mathbb{F}) = \mathbb{F}[\sigma_1, \ldots, \sigma_r].$$

Over $\mathbb{F} = \mathbb{C}$, one usually rescales these functions to

$$c_d = \left(\frac{1}{2\pi i}\right)^d \sigma_d \in I_r(\mathbb{C})$$

(so that the resulting characteristic classes are real, or even integral) or, equivalently:

$$det(I + \frac{t}{2\pi i}A) = \sum_{p=0}^{r} c_d(A)t^d.$$

The resulting classes,

$$c_d(E) \in H^{2d}(M)$$

are called the Chern classes of the (complex) vector bundle E. Putting them together, one obtains the total Chern class of E,

$$c(E) = \sum_{d=0}^{r} c_d(E) \in H^{\text{even}}(M).$$

Applying the same construction to real vector bundles, it so happens that the first, the third, etc classes vanish; hence the interesting classes are obtained using only the even-degree polynomials σ_{2k},

$$p_k := \left(\frac{1}{2\pi}\right)^{2k} \sigma_{2k} \in I_r(\mathbb{R}).$$

The resulting classes,

$$p_k(E) \in H^{4k}(M),$$

are called the Pontrjagin classes of the (real) vector bundle E.

Example 1.37. Another set of generators are obtained using the invariant polynomial functions

$$\Sigma_d : M_r(\mathbb{F}) \longrightarrow \mathbb{F},\ \Sigma_d(A) = Tr(A^d)$$

(and one can actually show that these elements with $0 \leq d \leq r$ generate the entire algebra $I_r(\mathbb{F})$). Again, it is customary to work over $\mathbb{C}$ and rescale these functions to

$$Ch_d = \frac{1}{d!}\left(-\frac{1}{2\pi i}\right)^d \Sigma_d,$$

so that one can write, formally,

$$Tr(e^{-\frac{t}{2\pi i}A}) = \sum_d Ch_d(A)t^d.$$

Applied to complex vector bundles E we obtain the so called Chern-character

$$Ch(E) = \sum_{d\geq 0} Ch_d(E) \in H^{\text{even}}(M).$$

Of course, this does not contain more (or less) information than the Chern class, but it has a better behavior with respect to the standard operations of vector bundles: it satisfies

$$Ch(E \oplus F) = Ch(E) + Ch(F), \quad Ch(E \otimes F) = Ch(E)Ch(F).$$

Remark 1.38. And here is yet another way to think about the invariant polynomials: interpret them as symmetric polynomials, over the base field $\mathbb{F}$, in r variables $x_1, \ldots x_r$ which play the role of the eigenvalues of a generic matrix A. More precisely, one has an isomorphism of algebras

$$I_r(\mathbb{F}) \cong Sym_{\mathbb{F}}[x_1, \ldots, x_r]$$

which associates to a symmetric polynomial S the invariant function (still denoted by S) given by

$$S(A) = S(x_1(A), \ldots, x_r(A)),$$

where $x_i(A)$ are the eigenvalues of A. Conversely, any $P \in I_r(\mathbb{F})$ can be viewed as a symmetric polynomial by evaluating it on diagonal matrices:

$$P(x_1, \ldots, x_r) := P(diag(x_1, \ldots, x_r)).$$

For instance, via this bijection, the Σ_p's correspond to the polynomials

$$\Sigma_p(x_1, \ldots, x_r) = \sum_i (x_i)^p,$$

while the σ_p's correspond to

$$\sigma_p(x_1, \ldots, x_r) = \sum_{i_1 < \ldots < i_p} x_{i_1} \ldots x_{i_p},$$

With this it is now easier to express the Σ's in term of the σ's and the other way around (using "Newton's formulas": $\Sigma_1 = \sigma_1$, $\Sigma_2 = (\sigma_1)^2 - 2\sigma_2$, $\Sigma_3 = (\sigma_1)^3 - 3\sigma_1\sigma_2 + 3\sigma_3$, etc.).

For more on characteristic classes, including full proofs, other Characteristic classes, more applications, you can have a look at lecture 10-11-12 of our more advanced course "Analysis on Manifolds" (available at http://www.staff.science.uu.nl/~crain101/AS-2013/).

1.4. Connections compatible with a metric

Very often the vector bundles one deals with comes with extra-structure and one is interested in connections that preserve the extra-structure. A good illustration of this principle is that of vector bundles endowed with metrics.

1.4.1. Metrics on vector bundles

Here, by a metric on a real vector space V we mean an inner product on V, i.e. a bilinear map $g : V \times V \to \mathbb{R}$ which is bi-linear, symmetric and satisfies $g(u, u) \geq 0$, with equality only when $u = 0$. Similalrly, by a metric on a complex vector space V we mean a hermitian inner-product $g : V \times V \to \mathbb{C}$- the difference with the real case being the conjugated symmetry requirement:

$$g(v, u) = \overline{g(u, v)} \quad \forall\ u, v \in V,$$

and (therefore) the conjugated linearity in the second argument

$$g(u, \lambda \cdot v) = \overline{\lambda} g(u, v) \quad \text{for } \lambda \in \mathbb{C}.$$

Given a (real or complex) vector bundle $\pi : E \to M$, a metric on E is a family

$$g = \{g_x\}_{x \in M}$$

of metrics g_x on the vector spaces E_x (one for each $x \in M$), which vary smoothly with respect to x in the sense that, for any two smooth local sections s and s' of E, the function $g(s, s')$ defined by

$$g(s, s')(x) := g_x(s(x), s'(x))$$

is smooth.

Exercise 24. For a metric g as above and a local frame $e = \{e_1, \ldots, e_r\}$ of E over some $U \subset M$, we define the coefficients of g w.r.t. e as the functions $g_{i,j} := g(e_i, e_j)$ (defined on U). Show that the smoothness of g is equivalent to the condition that all these coefficients (for all local frames) are smooth.

Exercise 25. A local frame e over U is called orthonormal w.r.t. g if

$$g(e_i, e_j) = \delta_{i,j} \quad \forall\ i, j \in \{1, \ldots, r\}.$$

Show that for any $x \in M$ there exists an orthonormal local frame defined on an open neighborhood of x.

Proposition 1.39. *Any vector bundle E admits a metric g.*

Proof. We consider a locally finite open cover of $M, \mathcal{U} = \{U_i\}_{i \in I}$ (where, this time, I is just a set of indices) together with a partition of unity $\{\eta_i\}$ subordinated to it such that, over each U_i, one finds a local frame. Each such local frame gives rise to a metric on $E|_{U_i}$; denote it by g^i, We then define

$$g = \sum_i \eta_i g^i,$$

or, more explicitly,

$$g_x(v, w) = \sum_i \eta_i(x) g^i_x(v, w).$$

From the definition of partitions of unity, it follows that g is smooth metric; for instance, the fact that $g_x(v, v) = 0$ happens only when $v = 0$ follows from the fact that the same is true for each g^i_x, while the coefficients $\eta_i(x)$ sum up to 1 (hence at least one of them is non-zero). $\square$

Exercise 26. Deduce that, for any real vector bundle E, its dual E^* is isomorphic to E. Similarly, from any complex vector bundle F, it dual F^* is isomorphic to its conjugate $\overline{F}$.

And here is another interesting use of metrics on a vector bundle E: for vector sub-bundles $F \subset E$ it allows us to produce a concrete model for the abstractly defined quotient E/F (see subsection 1.1.7):

Exercise 27. Let g be a metric on a vector bundle E over M and let F be a vector sub-bundle of E. For $x \in M$ we denote by $(F_x)^{\perp}$ the orthogonal of F_x in E_x, with respect to g_x. Show that:

1. the orthogonals $(F_x)^{\perp}$ together form a vector sub-bundle $F^{\perp}$ of E.
2. the quotient vector bundle E/F is isomorphic to $F^{\perp}$.

Exercise 28. Deduce that, for any vector bundle E over M, the following are equivalent:

1. E can be embedded in a trivial bundle (of some rank, usually very large).
2. one can find another vector bundle F such that $E \oplus F$ is isomorphic to the trivial bundle.

Actually, one can show that these items always hold true (for any vector bundle E); prove this when M is compact.

Exercise 29. Show that, if E is a real vector bundle of rank k over the compact manifold M then there exists a large enough natural number N and a map

$$f : M \to \mathrm{Gr}_k(\mathbb{R}^{N+k})$$

such that E is isomorphic to $f^*\gamma_k$- the pull-back by f of the tautological rank k-bundle (see Example 1.7). Similarly in the complex case.

1.4.2. Connections compatible with a metric

Fixing now a vector bundle E together with a metric g, there are various ways to make sense of the fact that a connection is compatible with the metric.

Proposition 1.40. *Given a vector bundle E endowed with a metric g, for any connection ∇ on E the following are equivalent:*

1. *Global: for any path $\gamma : [0,1] \to M$ between x and y, the induced parallel transport $T_\gamma^{0,1}$ is an isometry between (E_x, g_x) and (E_y, g_y).*
2. *Infinitesimal: the following Leibniz identity holds for any $s, s' \in \Gamma(E)$ and $X \in \mathcal{X}(M)$:*
$$L_X(g(s,s')) = g(\nabla_X(s), s') + g(s, \nabla_X(s')).$$
3. *Local: for any local orthonormal frame e, the connection matrix $\omega = \omega(\nabla, e)$ is anti self-adjoint:*
$$\overline{\omega_i^j} = -\omega_j^i,$$
where the over-line denotes complex conjugation (hence, in the real case, this just means that the matrix is antisymmetric).

A connection is said to be compatible with g if it satisfies one of these equivalent conditions.

Proof. Since condition 2. is local and since one can always choose local orthonormal frames (see the exercise above), it is enough to require 2. on members of local orthonormal frames- and then one arives at the equivalence of 2. with 3. For the equivalence with 1, note first that the condition in 2. can be re-written in terms of the derivatives $\frac{\nabla}{dt}$ induced by ∇ as follows: for any path γ in M and for any two paths $u, u' : I \to M$ above γ,
$$\frac{d}{dt} g(u,u') = g(\frac{\nabla u}{dt}, u') + g(u, \frac{\nabla u'}{dt}).$$
Choosing $u(t) = T_\gamma^{0,t}(v)$, $u'(t) = T_\gamma^{0,t}(v')$, with $v, v' \in E_x$ ($x = \gamma(0)$), we find that $g(u,u')$ is constant hence, in particular, it takes the same values at 1 and 0:
$$g(T_\gamma^{0,1}(v), T_\gamma^{0,1}(v')) = g(v, v');$$
hence $T_\gamma^{0,1}$ is an isometry. Moreover, one can check that the last argument can be reversed to prove that 1. implies 2. (exercise!). □

Proposition 1.41. *Any vector bundle E admits a metric g; for any g there exists a connection ∇ on E compatible with g.*

Proof. The existence of metrics was the subject of Proposition 1.39. Assume now that we have an arbitrary metric g. Choose again a partition of unity but, this time, with the property that, over each U_i, one finds a local frame that is orthonormal with respect to g. Each such local frame gives rise to a connection on U_i, denoted ∇^i, which is clearly compatible with g. Define then the global connection ∇ given by
$$\nabla_X(s)(x) := \sum_i \eta_i(x) \nabla^i_{X|_{U_i}}(s|_{U_i})(x).$$
It is not difficult to see that this defines indeed a connection and a computation of $g(\nabla_X(s), s') + g(s, \nabla_X(s'))$ using the fact that each ∇^i is compatible with g (over

U_i) and using that $\sum \eta_i = 1$ leads to $L_X g(s, s')$, i.e. the compatibility of ∇ with g. □

Exercise 30. Return to Theorem 1.35. Using the fact that the cohomology class of $[Tr(k_\nabla)]$ is independent of ∇ and the existence of metrics and compatible connections, prove now parts 2 and 3 of the theorem.

Exercise 31. Let F be a complex vector bundle over M. Show that if F can be written as the complexification of a real vector bundle over M, then $c_1(F) = 0$.

1.4.3. Riemannian manifolds I: The Levi-Civita connection

A Riemannian structure on a manifold M is a metric g on its tangent bundle TM. One also says that (M, g) is a Riemannian manifold. From the previous subsection we know that any manifold can be made into a Riemannian manifold and, given a Riemannian manifold (M, g), one can choose connections on TM compatible with g. However, working with TM instead of general vector bundles is a bit more special. First of all, $\Gamma(TM) = \mathcal{X}(M)$, so that a connection on TM is an operator

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M).$$

Because of this, one can talk about the torsion of a connection on TM:

Proposition 1.42. *For any connection ∇ on TM, the expression*

$$T_\nabla(X, Y) = \nabla_X(Y) - \nabla_Y(X) - [X, Y]$$

is $C^\infty(M)$-linear in the entries $X, Y \in \mathcal{X}(M)$. Hence it defines an element

$$T_\nabla \in \Gamma(\Lambda^2 T^*M \otimes TM) = \Omega^2(M; TM),$$

called the torsion of ∇.

This follows by a computation completely similar to (but simpler than) the one from the proof of Proposition 1.25; we leave it as an exercise. A connection on TM will be called torsion-free if $T_\nabla = 0$.

Theorem 1.43. *For any Riemannian manifold (M, g) there exists a unique connection ∇ on TM which is compatible with g and which is torsion free. It is called the Levi-Civita connection of (M, g).*

Proof. This proof here (which is the standard one) is not very enlightening (hopefully we will see a much more transparent/geometric argument a bit later in the course). One just plays with the compatibility equations applied to (X, Y, Z) and their cyclic permutations, combined in such a way that most of the appearances of the ∇ disappear by using the torsion free condition. One ends up with the following identity (which, once written down, can also be checked directly):

$$\begin{aligned} 2g(\nabla_X(Y), Z) &= L_X(g(Y, Z)) + L_Y(g(X, Z)) - L_Z(g(X, Y)) + \\ &+ g([X, Y], Z) + g([Z, X], Y) + g([Z, Y], X). \end{aligned} \tag{1.44}$$

Hence, fixing X and Y, we see that for the expression we are looking for, $\nabla_X(Y)$, the product with respect to any Z is predetermined. Since g is a metric, this forces the definition of $\nabla_X(Y)$ pointwise. It remains to check the identities that ensure that ∇ is a torsion free metric connection; each such identity can be written as $E = 0$ for some expression E; to prove such an identity one proves instead that $g(E, V) = 0$ for all vector fields V; with this trick one can use the definition of ∇ and the computations become straightforward. □

Remark 1.45 (The Christoffel symbols)**.** While, locally, connections ∇ on a vector bundle E are determined by connection matrices,

$$\nabla_X(e_j) = \sum_i \omega^i_j(X), \quad (e = \{e_1, \ldots, e_r\} \text{ – local frame of } E)$$

when looking at connections on $E = TM$, it is natural to use local frames

$$\frac{\partial}{\partial \chi^1}, \ldots, \frac{\partial}{\partial \chi^n}$$

induced by coordinate charts

$$(U, \chi = (\chi^1, \ldots, \chi^n))$$

of M. Then the resulting 1-forms can be further expressed in terms of functions:

$$\omega^i_j = \sum_p \Gamma^i_{pj} d\chi^p,$$

where Γ^i_{pj} are smooth functions on U. Equivalently:

$$\nabla_{\frac{\partial}{\partial \chi^p}}\left(\frac{\partial}{\partial \chi^j}\right) = \sum_i \Gamma^i_{pj} \frac{\partial}{\partial \chi^i}.$$

Locally, the fact that ∇ is torsion free is equivalent to

$$\Gamma^i_{pj} = -\Gamma^i_{jp}$$

for all i, j and p. While a metric g on TM is locally determined by the functions

$$g_{ij} := g\left(\frac{\partial}{\partial \chi^i}, \frac{\partial}{\partial \chi^j}\right),$$

the equation (1.44) that expresses the compatibility of ∇ with g becomes:

$$2 \sum_i \Gamma^i_{pj} g_{ik} = \frac{\partial g_{jk}}{\partial x_p} + \frac{\partial g_{pk}}{\partial x_j} - \frac{\partial g_{pj}}{\partial x_k}.$$

Using the inverse (g^{ij}) of the matrix (g_{ij}) we find that

$$\Gamma^i_{pj} = \frac{1}{2} \sum_k \left\{ \frac{\partial g_{jk}}{\partial x_p} + \frac{\partial g_{pk}}{\partial x_j} - \frac{\partial g_{pj}}{\partial x_k} \right\} g^{ki}.$$

These are called the Christoffel symbols of g with respect to the local chart (U, χ).

1.4.4. Riemannian manifolds II: geodesics and the exponential map

We have seen that the case of connections on the tangent bundle TM of a manifold M is more special: unlike the case of connections on general vector bundles, we could talk about the torsion of a connection; the vanishing of the torsion, together with the compatibility with a given metric g on TM (i.e. a Riemannian structure on M) were conditions that implied that a Riemannian manifold (M, g) comes with a canonical connection on TM.

Similarly, while a connection ∇ on a vector bundle π over M allows us to talk about the derivative

$$\frac{\nabla u}{dt} : I \to E$$

of a path $u : I \to E$, sitting above some path $\gamma : I \to M$, in the case when $E = TM$ there is a canonical path $u_\gamma : I \to TM$ sitting above any γ: its usual derivative.

In particular, one can talk about the second order derivative of $\gamma : I \to M$ (with respect to a connection ∇ on TM) defined as

$$\frac{\nabla(\dot{\gamma})}{dt} : I \to TM.$$

In particular, for a Riemannian manifold (M, g) and the associated Levi-Civita connection, one uses the notation:

$$\frac{D^2\gamma}{dt^2} : I \to M$$

for the resulting second derivative of paths $\gamma : I \to M$.

Definition 1.46. *Given a Riemannian manifold (M, g), a geodesic of (M, g) is any path $\gamma : I \to M$ with the property that $\frac{D^2\gamma}{dt^2} = 0$.*

The interesting existence question for geodesics is: given any $x \in M$ (read: initial point) and $v \in T_xM$ (read: initial speed), can one find a geodesic $\gamma : I \to M$ such that $\gamma(0) = x$, $\dot{\gamma}(0) = v$ (where I is an interval containing 0)? Is it unique? Is there a maximal one? I.e. the same questions as for flows as vector fields; and the answers will be very similar as well. Actually, to obtain the desired statements, we just have to realize that the geodesic equations can be written as flow equations.

Proposition 1.47. *For any Riemanninan manifold (M, g), there exists a unique vector field $\mathcal{X}$ on TM such that the integral curves of $\mathcal{X}$ are precisely those of type*

$$u_\gamma(t) = (\gamma(t), \dot{\gamma}(t)),$$

with γ a geodesic of (M, g); in other words,

$$\gamma \longleftrightarrow u_\gamma$$

defines a bijection between the geodesics of (M, g) and the integral curves of $\mathcal{X}$.

Proof. We first look at the geodesic equations locally, in a coordinate chart (U, χ); using the description of ∇ in terms of the Christoffel symbols, we see that $\gamma = (\gamma^1, \ldots, \gamma^n)$ is a geodesic if and only if

$$\frac{d^2\gamma^k}{dt^2} = -\sum_{i,j} \Gamma^k_{ij}(\gamma(t)) \frac{d\gamma^i}{dt}\frac{d\gamma^j}{dt}.$$

The idea is to write this system as a first order system:

$$\begin{cases} \frac{d\gamma^k}{dt} = & y^k(t), \\ \frac{dy^k}{dt} = & -\sum_{i,j} \Gamma^k_{ij}(\gamma(t)) y^i(t) y^j(t) \end{cases} \tag{1.48}$$

Let us formalize this. First, (U, χ) induces the chart $(TU, (\chi, y_\chi))$ for TM:

$$TU \ni (x, y^1 \frac{\partial}{\partial \chi^1} + \ldots + y^n \frac{\partial}{\partial \chi^n}) \mapsto (\chi, y_\chi)(\chi(x), y^1, \ldots, y^n) \in \mathbb{R}^{2n}.$$

It is now clear that, in these coordinates,

$$\mathcal{X} := \sum_k y^k_\chi \frac{\partial}{\partial \chi^k} - \sum \Gamma^k_{ij} y^i_\chi y^j_\chi \frac{\partial}{\partial y^k_\chi}$$

will have the desired properties. Since the geodesic equation is global and the local integral curves of a vector field determines the vector field, it follows that different charts (U', χ') give rise to vector fields $\mathcal{X}'$ on TU' which agree with $\mathcal{X}$ on $TU \cap TU'$. Therefore $\mathcal{X}$ is globally defined and has the desired properties. □

The vector field that arises from the previous proposition is called the geodesic vector field of (M, g). Using the properties of flows of vector fields applied to $\mathcal{X}$, we find an open subset

$$\mathcal{D} \subset \mathbb{R} \times TM \quad (\text{containing } \{0\} \times TM)$$

(the domain of the geodesic flow) and a smooth map (the geodesic flow)

$$\gamma : \mathcal{D} \to M$$

so that, for any $(x, v) \in TM$,

$$\gamma_{x,v}(t) := \gamma(t, x, v),$$

as a curve $\gamma_{x,v}$ defined on

$$I_{x,v} = \{t \in \mathbb{R} : (t, x, v) \in \mathcal{D}\}$$

(an open interval!), is the maximal geodesic with

$$\gamma_{x,v}(0) = x, \ \dot{\gamma}_{x,v}(0) = v.$$

Exercise 32. With the notations above show that, for any $c > 0$, one has

$$I_{x,cv} = \frac{1}{c} I_{x,v} \ \text{ and } \ \gamma(\frac{1}{c}t, x, cv) = \gamma(t, x, v) \ \text{ for all } t \in I_{x,v}.$$

The previous exercise indicates that the (local) information encoded in the geodesic flow is contained already in what it does at time $t = 1$. More precisely, one defines the exponential map of (M, g)

$$\exp(x, v) := \gamma(1, x, v),$$

which makes sense when (x, v) belongs to the open

$$\mathcal{U} := \{(x, v) \in TM : (1, x, v) \in \mathcal{D}\},$$

so that exp becomes a smooth map

$$TM \supset \mathcal{U} \mapsto \exp M.$$

With this, for all $(t, x, v) \in \mathcal{D}$, one has that $(x, tv) \in \mathcal{U}$ and

$$\gamma(t, x, v) = \exp(x, tv).$$

When one looks around a given point $x \in M$ it is common to consider

$$\exp_x : \mathcal{U}_x \to M,$$

the restriction of the exponential map to $\mathcal{U}_x = \mathcal{U} \cap TM$. And here is one of the fundamental properties of the exponential map:

Proposition 1.49. *For any $x \in M$,*

$$T_xM \supset \mathcal{U}_x \mapsto exp_x M$$

sends some open neighborhood of $0_x \in T_xM$ diffeomorphically into an open neighborhood of x in M. Even more: the differential at 0 of exp_x, combined with the canonical identification of $T_{0_x}(U_x)$ with T_xM, is the identity map.

Proof. By the inverse function theorem, it suffices to prove the last statement. Let $v \in T_xM$. Recall that the standard identification of T_xM with $T_{0_x}(U_x)$ interprets v as the speed at $t = 0$ of the curve $t \mapsto tv$. Therefore, it is sent by $(d\exp_x)$ to

$$\frac{d}{dt}_{|t=0} \exp_x(tv) = \frac{d}{dt}_{|t=0} \gamma(1, x, tv) = \frac{d}{dt}_{|t=0} \gamma(t, x, v) = v.$$

□

1.4.5. Riemannian manifolds III: application to tubular neighborhoods

And here is an important application of Riemannian metrics and of the exponential maps, application that is of interest outside Riemannian geometry: the existence of tubular neighborhoods. The concept of tubular neighborhood makes sense whenever we fix a submanifold N of a manifold M; then one can talk about tubular neighborhoods of N in M which, roughly speaking, are "linear approximations of M around N".

Remark 1.50 (One explanation)**.** To understand this, and the use of Riemannian metrics in this context, let us first consider the case when $N = \{x\}$ and return to the intuitive meaning of the tangent space T_xM at the point $x \in M$: it is the linear (i.e. vector space) approximation of the manifold M around the point x. The fact that M, around x, looks like T_xM is obvious: a coordinate chart $(U, \chi_1, \ldots, \chi_n)$ around x provides:

- an identification of a neighborhood of x in M with $\mathbb{R}^r$.
- a basis of T_xM, hence also an identification of T_xM with $\mathbb{R}^r$.

However, this way of relating M near x with T_xM is not so intrinsic: it depends on the choice of a chart around x- and this is a real problem if we want to pass from points x to general submanifolds $N \subset M$ (e.g. because we cannot find coordinate charts that contain an arbitrary N). The exponential map associated to a metric allows us to remove this problem since it relates M near x with T_xM (cf. Proposition 1.49) without a reference to coordinates- hence it has a better chance to work more generally.

The general idea is that, while the linear approximation of M around a point is a vector space, the linear approximation around a submanifold N will be a vector bundle over N.

Definition 1.51. *Given a submanifold N of a manifold M, a tubular neighborhood of N inside M is an open $U \subset M$ containing N, together with a diffeomorhpism*

$$\phi : E \to U$$

defined on a vector bundle E over N, such that $\phi(0_x) = x$ for all $x \in N$.

We now make some comments on the definition. In general, for a vector bundle E over a manifold N, one sees N as a submanifold of E "as the zero-section" by identifying each $x \in N$ with $0_x \in E$ (the zero vector of the fiber E_x).

Exercise 33. Show that, indeed, the zero section $\mathbb{O} : N \to E$, $\mathbb{O}(x) = 0_x$, is an embedding of N in E (where we view E as a manifold).

With this identification in mind, the condition $\phi(0_x) = x$ in the definition of tubular neighborhoods reads: ϕ is the identity on N. The next interesting remark is that the vector bundle E from the definition is actually determined by the way

N sits inside M. To make this precise, we consider the tangent bundle TN of N, sitting inside the restriction $T_N M = (TM)|_N$ of the tangent bundle of M to N, and we consider the resulting quotient

$$\nu_M(N) := T_N M/TN$$

(a vector bundle over N!). This is called the normal bundle of N in M, For the general construction of quotients of vector bundles, please see subsection 1.1.7. However, as in Exercise 27, metrics are helpful to to realize such quotients in more concrete terms:

Exercise 34. Go again through the argument, with the conclusion: if (M, g) is a Riemannian manifold and $N \subset M$ is a submanifold then, as vector bundles over N, $\nu_M(N)$ is isomorphic to the orthogonal of TN inside TM with respect to the Riemannian metric.

However, in some situations, normal bundles can also be "computed" without the choice of a metric:

Exercise 35. If $\pi : E \to N$ is a vector bundle show that the normal bundle of N inside the manifold E (where we view N as a submanifold of E as explained above) is isomorphic to the vector bundle $\pi : E \to N$.

In particular, this exercise tells us that any vector bundle over N can be realized as the normal bundle of N inside a bigger manifold.

Remark 1.52 (from "linear approximations" to the normal bundle)**.** It is interesting to go back to the original question: what are "linear approximations" of a manifold M around a submanifold N? Intuitively, one may think of opens in M containing N which are "as linear as possible around N". Denoting by m and n the dimensions of M, and N respectively, one may think that the most "linear" m-dimensional manifold containing N is

$$N \times \mathbb{R}^{m-n},$$

which contains $N \times \{0\}$ as a copy of N. However, hoping that (in general) a neighborhood of N inside M is diffeomorphic to such a "linear model" is too optimistic: indeed, it would imply that the normal bundle of N inside M was be trivial(izable). Which, by the previous exercise, cannot happen in general. However, the same exercise also indicates what to do: replace the too optimistic linear models $N \times \mathbb{R}^{m-n}$ by vector bundles over N. The same exercise tells us even more- namely what the vector bundle must be: the normal bundle of N in M.

Because of the last remark, it is customary that the definition of tubular neighborhoods uses right from the beginning $E = \nu_M(N)$. This closes our comments on the definition of tubular neighborhoods. And here is the existence result:

Theorem 1.53. *Any embedded submanifold N of a manifold M admits a tubular neighborhood in M.*

Proof. For simplicity, we restrict to the case when N is compact. Consider a Riemannian metric g on M and let E be the vector bundle over N which is the orthogonal complement of TN inside TM with respect to g (a more concrete realization of the normal bundle). We look at the exponential map restricted to E,

$$\exp_E : E \to M.$$

Strictly speaking, this is defined only on an open neighborhood of the zero section but, since all the arguments below are just around such "small enough" opens, we may assume that exp is defined on the entire E (just to simplify notations). We claim that, similar to Proposition 1.49, the differential of $\exp_E$ at any point of type $0_x = (x, 0) \in E$ is an isomorphism. To see this, note first that there is a canonical identification (independent of the metric and valid for any vector bundle E over N):

$$T_{0_x}E \cong T_xN \oplus E_x = T_xM. \tag{1.54}$$

Actually, the first identification is valid for any vector bundle E over a manifold N; let us explain this. First of all, for each $v \in E_x$ one has a short exact sequence

$$0 \to E_x \stackrel{i_v}{\to} T_vE \stackrel{d\pi}{\to} T_xN \to 0$$

(i.e. the first map is injective, the last one is surjective, and the image of the first coincides with the kernel of the second) where $\pi : E \to N$ is the projection of E and

$$i_v(w) = \frac{d}{dt}_{|t=0}(v + t \cdot w) \in T_vE.$$

The situation at $v = 0_x$ is more special- the previous sequence is canonically split (see the next exercise): the differential of the zero section $\mathbb{O} : N \to E$, $(d\mathbb{O})_x : T_xN \to T_{0_x}E$, is a right inverse of $d\pi$; and such a splitting of a short exact sequence induces an isomorphism between the middle space and the direct sum of the extreme ones,

$$E_x \oplus T_xN \cong T_{0_x}E,$$

by the map which is i_0 on E_x and is $(d\mathbb{O})_x$ on T_xN.

We now return to our Riemannian manifold and compute the differential

$$(d\exp_E)_{0_x} : T_{0_x}E \to T_xM$$

using the identification (1.54). For a vector coming from $w \in E_x$ we find, like in the proof of Proposition 1.49,

$$\frac{d}{dt}_{|t=0}\exp_E(x, tw) = \frac{d}{dt}_{|t=0}\gamma(1, x, tw) = w.$$

And, for a vector coming from $u \in T_xN$, say $u = \dot{\alpha}(0)$ for some curve α, we find

$$\frac{d}{dt}_{|t=0}\exp_E(0_{\alpha(t)}) = \frac{d}{dt}_{|t=0}\alpha(t) = u.$$

Therefore, with respect to (1.54), $(d\exp_E)_{0_x}$ is an isomorphism. Hence, for each $x \in N$, one finds an open neighborhood of 0_x in E on which $\exp_E$ becomes a diffeomorphism onto an open inside M. We claim that it continues to have the same property when restricted to

$$B_\epsilon := \{v \in E : ||v|| \leq \epsilon\},$$

for some $\epsilon > 0$ small enough. The only problem is to ensure that, for some ϵ, the exponential becomes injective on B_ϵ. We prove the claim by contradiction. If such ϵ does not exists, we find

$$v_n, w_n \in B_{\frac{1}{n}} \text{ with } v_n \neq w_n,\ \exp_E(v_n) = \exp_E(w_n).$$

Since N is compact, so are the closed balls $\overline{B}_\epsilon$ so, after eventually passing to convergent subsequences, we may assume that $v_n \to v$ and $w_n \to w$ with $v, w \in E$. Since $||v_n|| < \frac{1}{n}$, we find that $v = 0_x$ for some $x \in N$ and, similarly, $w = 0_y$. Passing to limit in $\exp_E(v_n) = \exp_E(w_n)$ we find that $x = y$. Therefore, both v_n as well as

w_n converge to 0_x. But this is a contradiction with the fact that $\exp_E$ is a local diffeomorphism around 0_x.

Therefore the restriction of the exponential to some $B_\epsilon \subset E$ is a diffeomorphism into an open inside M. To find a diffeomorphism defined on the entire E, one chooses any diffeomorphism $\phi : [0, \infty) \to [0, \epsilon)$ which is the identity near $t - 0$, and one composes the exponential with the diffeomorhism

$$\hat{\phi} : E \to B_\epsilon, \ \hat{\phi}(v) = \frac{\phi(||v||)}{||v||} \cdot v.$$

□

Exercise 36. A short exact sequence of (finite dimensional) vector spaces is a sequence

$$0 \to U \xrightarrow{i} V \xrightarrow{\pi} W \to 0 \tag{1.55}$$

consisting of vector spaces U, V and W, and linear maps i and π between them, satisfying the conditions: i-injective, π-surjective, $\mathrm{Im}(i) = \mathrm{Ker}(\pi)$.

A right splitting of (1.55) is any linear map $\sigma : W \to V$ which is a right inverse of π (i.e. $\pi \circ \sigma = \mathrm{Id}_W$); similarly, a left splitting is a linear map $p : V \to U$ which is a left inverse of i (i.e. $p \circ i = \mathrm{Id}_U$). Show that:

- there is a 1-1 correspondence between left splittings p and right splittings σ, uniquely characterized by the condition that
$$\sigma \circ \pi + i \circ p = \mathrm{Id}_V.$$
- the choice of a right splitting σ (or, equivalently, of a left splitting p) induces an isomorphism
$$U \oplus W \xrightarrow{\sim} V, \ (u, w) \mapsto (i(u), \sigma(w)),$$
with inverse given by $v \mapsto (p(v), \pi(v))$.

CHAPTER 2

Principal bundles

2.1. Digression: Lie groups and actions

2.1.1. The basic definitions

One of many good references for general Lie group theory is F.W. Warner, *Foundations of Differentiable manifolds and Lie groups.*

Definition 2.1. *A Lie group G is a group which is also a manifold, such that the two structures are compatible in the sense that the multiplication and the inversion operations,*

$$m : G \times G \to G, m(g, h) = gh, \quad \iota : G \to G, \iota(g) = g^{-1},$$

are smooth.

Given two Lie groups G and H, a Lie group homomorphism from G to H is any group homomorphism $f : G \to H$ which is also smooth. When f is also a diffeomorphism, we say that f is an isomorphism of Lie groups.

Example 2.2. The unit circle S^1, identified with the space of complex numbers of norm one,

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

is a Lie group with respect to the usual multiplication of complex numbers. Similarly, the 3-sphere S^3 can be made into a (non-commutative, this time) Lie group. For that we replace $\mathbb{C}$ by the space of quaternions:

$$\mathbb{H} = \{x + iy + jz + kt : x, y, z, t \in \mathbb{R}\}$$

where we recall that the product in $\mathbb{H}$ is uniquely determined by the fact that it is $\mathbb{R}$-bilinear and $i^2 = j^2 = k^2 = -1, ij = k, jk = i, ki = j$. Recall also that for

$$u = x + iy + jz + kt \in \mathbb{H}$$

one defines

$$u^* = x - iy - jz - kt \in \mathbb{H}, \quad |u| = \sqrt{uu^*} = \sqrt{x^2 + y^2 + z^2 + t^2} \in \mathbb{R}.$$

Then, the basic property $|u \cdot v| = |u| \cdot |v|$ still holds and we see that, identifying S^3 with the space of quaternionic numbers of norm 1, S^3 becomes a Lie group.

Example 2.3. The group GL_n of $n \times n$ invertible matrices (with the notation $GL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$ when we want to be more precise about the base field) is a Lie group. Indeed, denoting by

$$gl_n := \mathcal{M}_{n,n}$$

the space of $n \times n$ matrices, gl_n is just an Euclidean space hence has a canonical smooth structure, hence

$$GL_n = \{A \in gl_n : \det(A) \neq 0\},$$

being open in gl_n (since set : $gl_n \to \mathbb{R}$ is continuous, being a polynomial function), has an induced smooth structure as well. The product of matrices is a polynomial function in the entires, hence it is smooth. Similalry, the components of the map $A \mapsto A^{-1}$ are rational functions, hence smooth. Inside GL_n one can find several interesting Lie groups, such as

$$O(n) = \{A \in GL_n(\mathbb{R}) : A \cdot A^T = I\} \text{ (the orthogonal group)}$$

$$O(n) = \{A \in O(n) : \det(A) = 1\} \text{ (the special orthogonal group)}$$

$$SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) : \det(A) = 1\} \text{ (the special linear group)}$$

$$U(n) = \{A \in GL_n(\mathbb{C}) : A \cdot A^* = I\} \text{ (the unitary group)}$$

$$SU(n) = \{A \in GL_n(\mathbb{C}) : A \cdot A^* = I, \det(A) = 1\} \text{ (the special unitary group)}$$

$$Sp_k(\mathbb{R}) := \{A \in GL_{2k}(\mathbb{R}) : A^T J_{can} A = J_{can}\} \text{ (the symplectic group)}$$

where A^T denotes the transpose of A, A^* the conjugate transpose, and where

$$J_{\text{can}} = \begin{pmatrix} 0 & -I_k \\ I_k & 0 \end{pmatrix}. \tag{2.4}$$

It is straightforward to check that all these are subgroups of the general linear groups. To see that they are actually Lie groups, since GL_n is, it suffices to check that these subgroups are actually submanifolds of GL_n. This can be proven in each case, e.g. by using the regular value theorem[1]. However, a lot more is true (to be proven soon): any closed subgroup of GL_n is automatically a submanifold (hence a Lie group on its own); and it is clear that the previous subgroups are closed. Nevertheless, let us give the "direct argument" at least for one of the examples: $O(n)$. The main remark is that $A \cdot A^T$ is always a symmetric matrix and, denoting by Sym_n the space of such matrices (again an Euclidean space, of dimension $\frac{n(n+1)}{2}$), the map

$$f : GL_n \to Sym_n, \quad f(A) = A \cdot A^T$$

(actually defined on the entire Euclidean space gl_n) is a submersion. The last part follows by computing

$$(df)_A : gl_n \to Sym_n, \; X \mapsto \frac{d}{dt}_{t=0} f(A + tX) = A \cdot X^T + A^T \cdot X. \tag{2.5}$$

It follows that

$$O(n) = f^{-1}(\{I\})$$

is a smooth sub manifold of GL_n of dimension $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

Exercise 37. Work out a similar argument for $SL_n(\mathbb{R})$.

[1]for a smooth map $f : M \to N$ between manifolds, if $y \in N$ is a regular value, i.e. if $(df)_x : T_xM \to T_yM$ is surjective at all $x \in f^{-1}(y)$ (in particular if f is a submersion), then $f^{-1}(y)$ is either empty or a smooth submanifold of M of dimension $\dim(M) - \dim(N)$; moreover, for each $x \in f^{-1}(y)$, $T_x f^{-1}(y) \subset T_xM$ is precisely the kernel of $(df)_x$.

Finally, here is another interesting use of J_{can}: it characterizes a copy of $GL_k(\mathbb{C})$ inside $GL_{2k}(\mathbb{R})$:

$$j : GL_k(\mathbb{C}) \to GL_{2k}(\mathbb{R}), A + iB \mapsto \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$$

is a morphism of Lie groups which maps $GL_k(\mathbb{C})$ diffeomorphically into the subgroup of $GL_{2k}(\mathbb{R})$ given by

$$\{M \in GL_{2k}(\mathbb{R}) : J_{\text{can}}M = MJ_{\text{can}}\}.$$

Exercise 38. Show that

$$f : S^1 \to SO(2), \quad f(x, y) = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$

is an isomorphism of Lie groups. Similarly for

$$F : S^3 \to SU(2), \quad F(\alpha, \beta) = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}$$

where we interpret S^3 as $\{(\alpha, \beta) \in \mathbb{C}^2 : |\alpha|^2 + |\beta|^2 = 1\}$ and we use the group structure from Example 2.2.

Example 2.6. For any vector space V (over $\mathbb{R}$ or $\mathbb{C}$) one can talk about the general linear group associated to V,

$$GL(V) = \{A : V \to V : A = \text{a linear isomorphism}\}.$$

When V is finite dimensional then, as above, $GL(V)$ is a Lie group (sitting openly inside the finite dimensional vector space $gl(V)$ of all linear maps from V to itself). Actually, $GL(V)$ is isomorphic to GL_n (where n is the dimension of V), but writing down an isomorphism depends on the choice of a basis of V; because of this (and especially when we do not have canonical bases around- e.g. when passing from vector spaces to vector bundles), it is often useful to treat $GL(V)$ intrinsically, without identifying it with GL_n.

Definition 2.7. *Given a Lie group G and a manifold M, a left action of the Lie group G on M is a smooth map*

$$m : G \times M \to M, \quad \textit{also denoted } m(g, x) = g \cdot x$$

which satisfy the axioms for an action:

$$g \cdot (h \cdot (x)) = (gh) \cdot x, \ e \cdot x = x$$

for all $g, h \in G$, $x \in M$, where $e \in G$ is the unit of the group. Similarly one defines the notion of right action. If $M = V$ is a finite dimensional vector space, then such an action is said to be linear if, for each $a \in G$,

$$m(a, \cdot) : V \to V$$

is linear. In this case one usually re-writes the action as the map

$$r : G \to GL(V), \quad a \mapsto r_a := m(a, \cdot)$$

and one says that (V, r) (or simply V if r is clear) is a representation of G. In other words, a representation of G is a pair (V, r) with V a vector space and $r : V \to GL(V)$ a morphism of Lie groups (check it!).

Example 2.8. Of course, any of the subgroups of $GL_n(\mathbb{R})$ mentioned above comes with a linear action on $\mathbb{R}^n$ (and similarly over $\mathbb{C}$). Actually, each one of them can be characterized as the group of transformations of the Euclidean space (real or complex) preserving some structure. E.g., for $O(n)$ (and $U(n)$), it is about the preservation of the standard inner product on $\mathbb{R}^n$ (or $\mathbb{C}^n$, respectively); this implies in particular that the action of $O(n)$ preserves lengths hence it restricts to an action on the $n-1$-dimensional sphere

$$O(n) \times S^{n-1} \to S^{n-1}.$$

Given a Lie group G, the multiplication itself can be seen as both a left, as well as a right, action of G on itself. Depending on the point of view, it gives rise to:

- left translations: any $a \in G$ induces the left translation
$$L_a : G \to G, L_a(g) = a \cdot g.$$
- right translations, $R_a : G \to G$, $R_a(g) = g \cdot a$.

Another interesting action of G on itself is the so called (left) adjoint action

$$Ad : G \times G \to G,\ Ad(a, g) = a \cdot g \cdot a^{-1}. \tag{2.9}$$

Fixing $a \in G$, one has

$$Ad_a = L_a R_{a^{-1}} : G \to G,\ Ad_a(g) = a \cdot g \cdot a^{-1};$$

it is called the adjoint action of a on G.

2.1.2. The Lie algebra

The left/right translations allow one to move from different points in the group; or, even better, to move from any point to its unit element $e \in G$. For instance this philosophy can be applied when looking at tangent vectors, and move everything to the vector space:

$$\mathfrak{g} := T_e G.$$

More precisely, for any $a \in G$, the left translation L_a is a diffeomorphism (with inverse $L_{a^{-1}}$), hence its differential induces an isomorphism

$$(dL_a)_e : \mathfrak{g} \xrightarrow{\sim} T_a G. \tag{2.10}$$

Exercise 39. Deduce that, for any Lie group G, the tangent bundle TG is trivializable (i.e. any Lie group is parallelizable- cf. the discussion after Exercise 5).

Using the left translations (2.10), we see that any vector

$$\alpha \in \mathfrak{g}$$

induces, by left translating them, tangent vectors at all points in G; in other words, it induces a vector field

$$\alpha^L \in \mathcal{X}(G),$$

defined by

$$\alpha^L(a) = (dL_a)_e(\alpha) \in T_a G.$$

These vector fields can be characterized slightly differently, using the notion of left invariance. We say that a vector field $X \in \mathcal{X}(G)$ is left invariant if

$$X(ag) = (dL_a)_g(X(g))\ \ \forall\, a, g \in G. \tag{2.11}$$

We denote by $\mathcal{X}^{\mathrm{inv}}(G)$ the subspace of $\mathcal{X}(G)$ consisting of left-invariant vector fields (it is clearly a vector subspace).

Exercise 40. Recall that, for a diffeomorphism $\phi : M \to M$, one has an induced push-forward operation $\phi_* : \mathcal{X}(M) \to \mathcal{X}(M)$ given by

$$\phi_*(X)(x) = (d\phi)_{\phi^{-1}(x)}(X(\phi^{-1}(x)).$$

1. Show that $X \in \mathcal{X}(G)$ is left invariant iff $(L_a)_*(X) = X$ for all $a \in G$.
2. Given a left action of a Lie group G on a manifold M, define the notion of left-invariance of vector fields on M so that, when applied to the left action of G on itself, one recovers the notion of left invariance discussed above.

Next, we discuss the Lie algebra of a Lie group.

Proposition 2.12. *For any Lie group G,*

1. *The construction $\alpha \mapsto \alpha^L$ defines a bijection between $\mathfrak{g}$ and $\mathcal{X}^{inv}(G)$.*
2. *If two vector fields $X, Y \in G$ are left invariant, then so is their Lie bracket $[X, Y]$. In particular, there is an operation*
$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, (\alpha, \beta) \mapsto [\alpha, \beta],$$
unique with the property that, for any $\alpha, \beta \in \mathfrak{g}$,
$$[\alpha, \beta]^L = [\alpha^L, \beta^L].$$
3. *The previous operation on $\mathfrak{g}$ is bi-linear, antisymmetric and satisfies the so-called Jacobi identity:*
$$[[\alpha, \beta], \gamma] + [[\beta, \gamma], \alpha] + [[\gamma, \alpha], \beta] = 0.$$

Proof. If $X \in \mathcal{X}^{\mathrm{inv}}(G)$, then the left invariance condition (2.11) at $g = e$ implies that $X = \alpha^L$ where $\alpha = X(e)$. Conversely, it is straightforward to check that α^L is left invariance. The second part follows from the previous proposition and the fact that the push-forward operation ϕ_* is compatible with the brackets: $\phi_*([X, Y]) = [\phi_*(X), \phi_*(Y)]$. For the third part, it suffices to remember that the Jacobi identity is satisfied by the Lie bracket of vector fields (these general facts about vector fields are actually straightforward to check using the definition/characterization of the Lie bracket of vector fields: $[X, Y](f) = X(Y(f)) - Y(X(f))$ for all smooth functions $f : M \to \mathbb{R}$). ☐

Definition 2.13. *A Lie algebra is a vector space $\mathfrak{a}$ endowed with an operation*

$$[\cdot, \cdot] : \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$$

which is bi-linear, antisymmetric and which satisfies the Jacobi identity.

Hence the last part of the proposition says that, for a Lie group G, $\mathfrak{g}$ endowed with the bracket $[\cdot, \cdot]$ is a Lie algebra; it is called the Lie algebra of the Lie group G.

Example 2.14 (GL_n and gl_n)**.** Let us show how the Lie algebra of GL_n can be identified with gl_n in such a way that the Lie bracket becomes the usual commutator bracket

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

For simplicity, we restrict to the real case. As we have already mentioned, gl_n is seen as an Euclidean space; we denote coordinates functions by

$$u^i_j : gl_n \to \mathbb{R}$$

(it picks the element on the position (i, j)). While GL_n sits openly inside gl_n, its tangent space at the identity matrix I (and similarly at any point) is canonically identified with gl_n: any $X \in gl_n$ is identified with the speed at $t = 0$ of $t \mapsto (I + tX)$. After left translating, we find that that left invariant vector field X^L that it defines is

$$X^L(A) = \frac{d}{dt}_{|t=0} A \cdot (I + tX) \in T_A GL_n \tag{2.15}$$

for all $A \in GL_n$. Let $X, Y \in gl_n$ and let $[X, Y] \in gl_n$ be the resulting bracket coming from GL_n (defined by $[X, Y]^L = [X^L, Y^L]$, where the last bracket is the Lie bracket of vector fields on GL_n). To compute it, we use the defining equation for the Lie bracket of two vector fields:

$$L_{[X^L, Y^L]}(F) = L_{X^L} L_{Y^L}(F) - L_{Y^L} L_{X^L}(F) \tag{2.16}$$

for all smooth functions

$$F : GL_n \to \mathbb{R}.$$

From the previous description of X^L we deduce that

$$L_{X^L}(F)(A) = \frac{d}{dt}_{|t=0} F(A \cdot (I + tX)) = \sum_{i,j,k} \frac{\partial F}{\partial u^i_j}(A)\, A^i_k X^k_j.$$

In particular, applied to a coordinate function $F = u^i_j$, we find

$$L_{X^L}(u^i_j)(A) = \sum_k A^i_k X^k_j$$

or, equivalently, $L_{X^L}(u^i_j) = \sum_k u^i_k X^k_j$. Therefore

$$L_{Y^L}(L_{X^L}(u^i_j)) = \sum_k L_{Y^L}(u^i_k) X^k_j = \sum_{k,l} u^i_l Y^l_k X^k_j,$$

which is precisely $L_{(Y \cdot X)^L}(u^i_j)$. Since this holds for all coordinate functions, we have

$$L_{Y^L} \circ L_{X^L} = L_{(Y \cdot X)^L};$$

we immediately deduce that $[X, Y]$ is $X \cdot Y - Y \cdot X$, i.e. the usual commutator of matrices.

Definition 2.17. *A morphism between two Lie algebras* $(\mathfrak{g}, [\cdot, \cdot])$ *and* $(\mathfrak{g}', [\cdot, \cdot]')$ *is any linear map* $f : \mathfrak{g} \to \mathfrak{g}'$ *which is compatible with the brackets in the sense that*

$$f([\alpha, \beta]) = [f(\alpha), f(\beta)]' \quad \forall\ \alpha, \beta \in \mathfrak{g}.$$

Exercise 41. Show that if $F : H \to G$ is a morphism between two Lie groups with Lie algebras denoted by $\mathfrak{h}$ and $\mathfrak{g}$, respectively, then the differential of F of the identity element, $f = (dF)_e : \mathfrak{h} \to \mathfrak{g}$ is a morphism of Lie algebras.

The previous exercise shows that, if G is a Lie group with Lie algebra $\mathfrak{g}$ then for any Lie subgroup H of G the Lie algebra $\mathfrak{h}$ of H is a Lie sub-algebra of $\mathfrak{g}$; i.e., while $\mathfrak{h}$ is a vector subspace of $\mathfrak{g}$ (as a tangent space of a submanifold), the exercise tells us that the Lie bracket of $\mathfrak{h}$ can be computed using the Lie bracket of $\mathfrak{g}$.

Example 2.18. For any of the subgroups of GL_n from Example 2.3, their Lie algebras can be described explicitly as Lie sub-algebras of gl_n. For instance, for $O(n)$, using the smooth structure as discussed in Example 2.3, we see that $T_I O(n)$ is the kernel of the map (2.5) at $A = I$; hence we find that the Lie algebra of $O(n)$ is

$$o(n) = \{X \in gl_n(\mathbb{R}) : X + X^T = 0\}.$$

For the other groups from Example 2.3 one finds

$$\begin{aligned} so(n) &= o(n), \\ sl_n(\mathbb{R}) &= \{A \in gl_n(\mathbb{R}) : Tr(X) = 0\}, \\ u(n) &= \{X \in gl_n(\mathbb{C}) : X + X^* = 0\}, \\ su(n) &= \{X \in gl_n(\mathbb{C}) : X + X^* = 0, Tr(X) = 0\}, \\ sp_n(\mathbb{R}) &= \{X \in gl_{2n}(\mathbb{C}) : J_{\text{can}} \cdot X + X^T \cdot J_{\text{can}} = 0\}. \end{aligned}$$

Example 2.19. While all these examples are finite dimensional, Lie algebras can also be infinite dimensional. For instance, already implicit in our previous discussions is the fact that for any manifold M the space $\mathcal{X}(M)$ of vector fields on M, endowed with the usual Lie bracket of vector fields, is a Lie algebra. Also, as a generalization of gl_n, for any vector space V one can talk about the Lie algebra

$$gl(V) = \{A; V \to V : A - \text{linear map}\},$$

endowed with usual commutator bracket $[A, B] = A \circ B - B \circ A$. This is the infinitesimal analogue of the general linear group $GL(V)$ from Example 2.6. As there, if V is finite dimensional, then $gl(V)$ is isomorphic to gl_n but it is often useful to think about $gl(V)$ without making this identification (which depends on the choice of a basis of V). For instance, while Example 2.14 does imply that the Lie algebra of $GL(V)$ is isomorphic to $gl(V)$, the formula (2.15) describing the identification still makes sense in this setting (and it is a nice exercise to write the entire argument without choosing a basis).

While representations of Lie groups were defined using $GL(V)$, similarly:

Definition 2.20. *A representation of a Lie algebra $\mathfrak{g}$ on a vector space V is any Lie algebra morphism*

$$\rho : \mathfrak{g} \to gl(V).$$

Note that, by the previous discussion, any representation $r : G \to GL(V)$ of a Lie group G induces, after differentiating at the unit, a representation $\rho : \mathfrak{g} \to gl(V)$ of its Lie algebra.

Example 2.21 (The adjoint action(s))**.** Of particular interests are the actions obtained differentiating the adjoint action of G on itself (see (2.9)). More precisely, since Ad_a sends the unit to the unit, its differential at the unit,

$$(d\,Ad)_e : \mathfrak{g} \to \mathfrak{g},$$

still denoted by Ad_a, defines an element of $GL(\mathfrak{g})$. This gives rise to the adjoint action of G on its Lie algebra,

$$\mathrm{Ad} : G \to GL(\mathfrak{g}),\ a \mapsto \mathrm{Ad}_a.$$

The corresponding infinitesimal action (obtained by differentiating Ad at the unit) is denoted

$$\mathrm{ad} : \mathfrak{g} \to GL(\mathfrak{g}),\ \alpha \mapsto \mathrm{ad}_\alpha.$$

and is called the (infinitesimal) adjoint action of $\mathfrak{g}$ on itself. The exponential map that will be discussed in the next subsection will allows us to write down explicit formulas for these representations. In particular, the next exercise will become easier after the next subsection.

Exercise 42. If $\mathfrak{g}$ is the Lie algebra of a Lie group G show that

$$\mathrm{ad}_\alpha(\beta) = [\alpha, \beta].$$

(Note 1: hence the Lie bracket of $\mathfrak{g}$ can be recovered via the adjoint action.
Note 2: the fact that ad is a representation is equivalent to the Jacobi identity.)

2.1.3. The exponential map

And here is another central object in the theory of Lie groups: the exponential map.

Proposition 2.22. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. For any $\alpha \in \mathfrak{g}$, the vector field α^L is complete and its flow $\phi^t_{\alpha^L}$ satisfies*

$$\phi^t_{\alpha^L}(ag) = a\phi^t_{\alpha^L}(g).$$

In particular, the flow can be reconstructed from what it does at time $t = 1$, at the identity element of G, i.e. from

$$exp : \mathfrak{g} \to G, \ exp(\alpha) := \phi^1_{\alpha^L}(e),$$

by the formula

$$\phi^t_{\alpha^L}(a) = a\, exp(t\alpha).$$

Moreover, the exponential is a local diffeomorphism around the origin: it sends some open neighborhood of the origin in $\mathfrak{g}$ diffeomorphically into an open neighborhood of the identity matrix in G.

Proof. For the first part the key remark is that if $\gamma : I \to G$ is an integral curve of α^L defined on some interval I, i.e. if

$$\frac{g\gamma}{dt}(t) = \alpha^L(\gamma(t)) \quad \forall\ t \in I$$

then, for any $a \in G$, the left translate $a \cdot \gamma : t \mapsto a \cdot \gamma(t)$ is again an integral curve (do the computation!). If γ is the flow-line through some point $g \in G$ (i.e. $\gamma(0) = g_0$) then $a \cdot \gamma$ will be the one through $a \cdot g$- and that proves the first identity (for as long as both terms are defined). For the completeness, by the same argument, if the integral curve γ that starts at some point $g \in G$ is defined for all $t \in [0, \epsilon]$ the, for any $a \in G$,

$$\gamma'(t) := a \cdot \gamma(t - \epsilon)$$

will be an integral curve that is defined at least on $[\epsilon, 2\epsilon]$; moreover, we can arrange that $\gamma'(\epsilon) = \gamma(\epsilon)$ by using $a = \gamma(\epsilon)\gamma(0)^{-1}$; hence, from the uniqueness of integral curves passing through a point, we see that γ is actually defined on the entire $[0, 2\epsilon]$ hence, repeating the argument, on the entire $[0, \infty)$; and similarly on the negative line.

We are left with the last part. For that it suffices to show that $(d\exp)_0$ is an isomorphism. But

$$(d\exp)_0 : \mathfrak{g} \to T_eG = \mathfrak{g}$$

sends $\alpha \in \mathfrak{g}$ to

$$\frac{d}{dt}_{|t=0} \exp(t\alpha) = \frac{d}{dt}_{|t=0} \phi^t_{\alpha^L}(e) = \alpha^L(e) = \alpha,$$

i.e. the differential is actually the identity map. □

Exercise 43. Recall (from the basic differential geometry course ...) that the Lie bracket of two vector fields X and Y on a manifold M can also be described using flows as

$$[X,Y](x) = \frac{d}{dt}_{|t=0}(\phi_X^{-t})_*(Y)(x) = \frac{d}{dt}_{|t=0}(d\phi_X^{-t})(Y(\phi_X^t(x))). \tag{2.23}$$

For a Lie group G with Lie algebra $\mathfrak{g}$, deduce that

$$[\alpha,\beta] = \frac{d}{dt}_{|t=0} Ad_{\exp(t\alpha)}(\beta).$$

Look again at Exercise 42 .

Exercise 44. Consider a linear action $G \times V \to V$, interpreted as a representation $r : G \to GL(V)$, and let $\rho : \mathfrak{g} \to gl(V)$ be the induced representation of its Lie algebra (obtained by differentiating r at the identity element of G). Show that

$$\rho_\alpha(v) = \frac{d}{dt}_{|t=0} \exp(t\alpha)\cdot v = \frac{d}{dt}_{|t=0} r_{\exp(t\alpha)}(v) \quad (\alpha \in \mathfrak{g}, v \in V).$$

Using the basic formula (2.23) from the previous exercise, check directly that ρ is a Lie algebra homomorphism.

2.1.4. Closed subgroups of GL_n

In this subsection, as an application of the exponential map, we will show that any subgroup of GL_n which is closed is automatically a Lie subgroup. Let us starts by looking closer at the exponential map for GL_n.

Example 2.24. We continue Example 2.14 and show that, for GL_n, the exponential map becomes the usual exponential map for matrices:

$$e : gl_n \to GL_n,\ X \mapsto e^X := \sum_{k=0}^{\infty} \frac{X^k}{k!}$$

(where we work with the conventions that $X^0 = I$ the identity matrix and $0! = 1$). Let us compute $\exp(X)$ for an arbitrary $X \in gl_n = T_I GL_n$; according to the definition, it is $\gamma(1)$, where γ is the unique path in GL_n satisfying

$$\gamma(0) = I,\ \frac{d\gamma}{dt}(t) = X^L(\gamma(t)).$$

The last equality takes place in $T_{\gamma(t)}GL_n$, which is canonically identified with gl_n (as in the previous example, $Y \in gl_n$ is identified with the speed at $\epsilon = 0$ of $\epsilon \mapsto \gamma(t) + \epsilon Y$). By this identification, $\frac{d\gamma}{dt}(t)$ goes to the usual derivative of γ (as a path in the Euclidean space gl_n) and $X^L(\gamma(t))$ goes to $\gamma(t)\cdot X$. Hence γ is the solution of

$$\gamma(0) = I,\ \gamma'(t) = X \cdot \gamma(t),$$

which is precisely what $t \mapsto e^{tX}$ does. We deduce that $\exp(X) = e^X$.

Proposition 2.25. *Any closed subgroup G of GL_n is a Lie subgroup of GL_n with Lie algebra*

$$\mathfrak{g} := \{X \in gl_n : e^{tX} \in G \ \ \forall \ \ t \in \mathbb{R}\}.$$

Proof. For the first part we have to prove that G is an embedded submanifold of GL_n; for simplicity, assume we work over $\mathbb{R}$. We will show that $\mathfrak{g}$ is a linear subspace of gl_n; then we will choose a complement $\mathfrak{g}'$ of $\mathfrak{g}$ in gl_n and we show that one can find an open neighborhood V of the origin in $\mathfrak{g}$, and a similar one V' in $\mathfrak{g}'$, so that

$$\phi : V \times V' \to GL_n, \phi(X, X') = e^X \cdot e^{X'}$$

is a local diffeomorphism onto an open neighborhood of the identity matrix and so that

$$G \cap \phi(V \times V') = \{\phi(v, 0) : v \in V\}. \tag{2.26}$$

This means that the submanifold condition is verified around the identity matrix and then, using left translations, it will hold at all points of G. Note also that the differential of ϕ at 0 (with ϕ viewed as a map defined on the entire $\mathfrak{g} \times \mathfrak{g}'$) is the identity map:

$$(d\phi)_0 : \mathfrak{g} \times \mathfrak{g}' \to T_I GL_n = gl_n, (X, X') \mapsto \frac{d}{dt}_{t=0} e^{tX} e^{tX'} = X + X'.$$

In particular, ϕ is a indeed a local diffeomorphism around the origin. Hence the main condition that we have to take care of is (2.26), i.e. that if $X \in V$, $X' \in V'$ satisfy $e^X \cdot e^{X'} \in G$, then $X' = 0$. Hence it suffices to show that one can find V' so that

$$X' \in V', e^{X'} \in G \Longrightarrow X' = 0$$

(then just choose any V small enough such that ϕ is a diffeomorphism from $V \times V'$ into an open neighborhood of the identity in GL_n). For that we proceed by contradiction. If such V' would not exist, we would find a sequence $Y_n \to 0$ in $\mathfrak{g}'$ such that $e^{Y_n} \in G$. Since Y_n converges to 0, we find a sequence of integers $k_n \to \infty$ such that $X_n := k_n Y_n$ stay in a closed bounded region of $\mathfrak{g}'$ not containing the origin (e.g. on $\{X \in \mathfrak{g}' : 1 \leq ||X|| \leq 2\}$, for some norm on $\mathfrak{g}'$); then, after eventually passing to a subsequence, we may then assume that $X_n \to X \in \mathfrak{g}'$ non-zero. In particular, X cannot belong to $\mathfrak{g}$. Setting $t_n = \frac{1}{k_n}$, we see that $e^{t_n X_n} = e^{Y_n} \in G$, and then the desired contradiction will follow from:

Lemma 2.27. *Let $X_n \to X$ be a sequence of elements in $\mathfrak{gl}_n$ and $t_n \to 0$ a sequence of non-zero real numbers. Suppose $e^{t_n X_n} \in G$ for all n: then $e^{tX} \in G$ for all t, i.e., $X \in \mathfrak{g}$.*

Proof. See the book by Sternberg (Lemma 4.2, Chapter V). □

The same Lemma can also be used to prove the claim we made (and used) at the start of the argument: that $\mathfrak{g}$ is a linear subspace of gl_n. The $\mathbb{R}$-linearity follows immediately from the definition. Assume now that $X, Y \in \mathfrak{g}$ and we show that $X + Y \in \mathfrak{g}$. Write now

$$e^{tX} e^{tY} = e^{f(t)}, \quad \text{with} \quad \lim_{t \to 0} \frac{f(t)}{t} = X + Y$$

(exercise: why is this possible?). Taking t_n a sequence of real numbers converging to 0 and $X_n = \frac{f(t_n)}{t_n}$ in the lemma, to conclude that $X + Y \in \mathfrak{g}$.

This closes the proof of the fact that G is a submanifold of GL_n (hence also a Lie group). It should now also be clear that $\mathfrak{g}$ is the Lie algebra of G since $\phi|_{V \times \{0\}}$ is precisely the restoration of the exponential map e to the open neighborhood V of the origin in $\mathfrak{g}$. □

Exercise 45. In particular, it follows that $\mathfrak{g}$ is closed under the commutator bracket of matrices:

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

However, prove this directly, using the previous Lemma, by an argument similar to the one we used to prove that $\mathfrak{g}$ is a linear subspace of gl_n.

2.1.5. Infinitesimal actions

Finally, a few more words about actions. Recall that, besides the finite dimensional Lie algebras arising from Lie groups, another important Lie algebra that appeared in our discussion was the Lie algebra $\mathcal{X}(M)$ of vector fields on a manifold M. The concept of infinitesimal action combines the two:

Definition 2.28. *An infinitesimal action of a Lie algebra $\mathfrak{g}$ on a manifold M is any Lie algebra morphism*

$$a : \mathfrak{g} \to \mathcal{X}(M).$$

Of course, this definition arises when looking at the infinitesimal counterpart of the notion of group actions (i.e. what is left after differentiation). So, in some sense, the next proposition actually comes before the previous proposition:

Proposition 2.29. *Let $\mu : M \times G \to M$ be a smooth right action of the Lie group G on a manifold M. For $\alpha \in \mathfrak{g}$, one defines the infinitesimal generator induced by α as the $\alpha_M \in \mathcal{X}(M)$ whose flow is given by*

$$\phi^t_{\alpha_M}(x) = x \cdot exp(t\alpha).$$

Then $\alpha \mapsto \alpha_M$ defines an infinitesimal action of $\mathfrak{g}$ on M.

Note that, explicitly,

$$\alpha_M(x) = \frac{d}{dt}_{|t=0} x \cdot exp(t\alpha);$$

equivalently, $\alpha_M(x)$ is obtained by differentiating $\mu^x : G \to M$, $\mu^x(a) = a \cdot x$:

$$\alpha_M(x) = (d\mu^x)_e(\alpha) \in T_x M. \tag{2.30}$$

Note that, in the previous proposition, we have chosen to use right actions in order to obtain a nicer statement. Indeed, the similar proposition for left actions $\mu : G \times M \to M$ will require the use of an extra-sign: the vectors α_M will then be defined by adding a minus sign to (2.30) or, using the exponential map, as

$$\alpha_M(x) = \frac{d}{dt}_{|t=0} exp(-t\alpha) \cdot x. \tag{2.31}$$

This is due to the fact that passing between left and right actions is given by

$$g \cdot x = x \cdot g^{-1}$$

and the differential of the map $g \mapsto g^{-1}$, at the identity elements of G, is $\alpha \mapsto -\alpha$.

Corollary 2.32. *For any finite dimensional vector space V,*

$$i : gl(V) \to \mathcal{X}(V), \quad i(A)(v) := \frac{d}{dt}_{|t=0} exp(-tA)(v)$$

is an injective morphism of Lie algebras.

Note that this identifies $gl(V)$ with the sub Lie algebra of $\mathcal{X}(V)$ of "linear vector fields" (i.e. vector fields on V whose coefficients, with respect to a/any basis of V, are linear functions).

The next exercise explains the note in the statement and also provides a short, direct way to see the need of the minus sign.

Exercise 46. Take $G = GL_n$ with the usual left action on $M = \mathbb{R}^n$:

$$GL_n \times \mathbb{R}^n \to \mathbb{R}^n, (A, v) \mapsto A(v) = \sum_{i,j} A^i_j v^j e_i$$

(i.e. so that $\begin{pmatrix} A(v)^1 \\ \dots \\ A(v)^n \end{pmatrix} = \begin{pmatrix} A^1_1 & \dots & A^1_n \\ \dots & \dots & \dots \\ A^n_1 & \dots & A^n_n \end{pmatrix} \cdot \begin{pmatrix} v^1 \\ \dots \\ v^n \end{pmatrix}$) where e_i are the components of the standard basis of $\mathbb{R}^n$ and v^j are the components of v with respect to this basis. Check that the previous proposition/discussions gives rise to the infinitesimal action

$$\rho : gl_n \to \mathcal{X}(\mathbb{R}^n),\ \rho(A) = -\sum_{i,p} A^i_p x^p \frac{\partial}{\partial x_i}$$

Then check by a direct computation that, indeed, ρ is a Lie algebra map (and the minus sign above is needed!).

Note that, in the case of a linear action of a Lie group G on a vector space V, $G \times V \to V$, we now have two ways of passing to "infinitesimal data":

1. The induced representation of the Lie algebra $\mathfrak{g}$ of G (see e.g. Exercise 44):
$$\rho : \mathfrak{g} \to gl(V),\ \rho_\alpha(v) = \frac{d}{dt}_{|t=0} \exp(tv) \cdot v.$$
2. The induced infinitesimal action of $\mathfrak{g}$ on V (from the previous proposition),
$$a : \mathfrak{g} \to \mathcal{X}(V),\ \alpha \mapsto \alpha_V.$$

Exercise 47. With the previous notations, show that ρ and a are the same, modulo the inclusion i from the previous corollary; more precisely, $a = i \circ \rho$ or, as a commutative diagram,

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\rho"] \arrow[dr, "a"'] & gl(V) \arrow[d, "i"] \\ & \mathcal{X}(V) \end{tikzcd}$$

2.1.6. Free and proper actions

Group actions give rise to quotient spaces M/G whose elements are the orbits $G \cdot x$ of the action; in general, one endows M/G with the quotient topology, i.e. the largest topology with the property that the canonical projection

$$\pi : M \to M/G$$

(sending a point $x \in M$ to the orbit $G \cdot x$ through x) is continuous. However, M/G is often rather pathological; in particular, it is rare that this quotient can be made into a manifold in a natural way (i.e. such that π becomes a submersion). We now look at restricted class of actions when this problem can be overcome.

Definition 2.33. *A (left, say) Lie group action of G on a M is called:*

1. *free: if $g \cdot x = h \cdot x$ (for some $x \in M$) implies $g = h$.*
2. *proper: if the map $G \times M \to M \times M$, $(g, x) \mapsto (gx, x)$ is proper.*

(a map $f : M \to N$ is proper if $f^{-1}(K)$ is compact for any $K \subset N$ compact).

Before the main theorem, here is a lemma that we will need:

Lemma 2.34. *If the action is free then, for all $x \in M$,*

$$a_x : \mathfrak{g} \to T_x M,\ \ \alpha \mapsto \alpha_M(x) \ \ (\textit{see Proposition 2.29})$$

is injective.

Because of this lemma, (infinitesimal) actions with the previous property are called infinitesimally free.

Proof. From the definition of $a(\alpha)$, its flow is given by

$$\phi^t_{a(\alpha)}(x) = \exp(-t\,\alpha)\cdot x.$$

Now, if $a_x(\alpha) = 0$ at some point x, i.e. if the vector field $a(\alpha)$ vanishes at x, then the flow of this vector fields through the point x is constant:

$$\exp(-t\,\alpha)\cdot x = x$$

for all t. The freeness of the action implies that $\exp(-t\,\alpha) = 0$ for all t, hence $\alpha = 0$. □

Theorem 2.35. *If a Lie group G acts freely and properly on a manifold M, then the quotient M/G admits a unique smooth structure with the property that the quotient map $\pi : M \to M/G$ is a submersion.*

Proof. Remark first that, if $\pi : M \to N$ is a surjective map from a manifold M to a set N, then N admits at most one smooth structure with the property that π is a submersion. This follows from the local form of submersions (fill in the details!). This implies not only the uniqueness part of the theorem but also allows us to proceed locally: we will show that for any $x_0 \in M$, one finds an invariant open neighborhood U of x_0 and a smooth structure on $\pi(U)$ such that $\pi|_U : U \to \pi(U)$ is a submersion; then $\pi(U)$s can be used as charts and the (smooth) compatibility of the charts will follow using again the uniqueness property. So, fix $x_0 \in M$ and look for U. Choose first a submanifold S of M containing x_0 such that

$$a_{x_0}(\mathfrak{g}) \oplus T_{x_0}S = T_{x_0}M,$$

where $a_{x_0} : \mathfrak{g} \to T_{x_0}M$ is the infinitesimal action at x_0. S will be important only in a neighborhood of x_0. Consider the map

$$f : G \times S \to M,\ \ f(g,x) = g\cdot x.$$

Computing the differential at the point (e, x_0) we find

$$(df)_{e,x_0} : \mathfrak{g} \times T_{x_0}S \to T_{x_0}M, (\alpha, v) \mapsto a_{x_0}(\alpha) + v.$$

Since this is an isomorphism, $(df)_{e,x}$ continues to be an isomorphism for x close to x_0. Hence, after eventually shrinking S, we may assume that this happens at all $x \in G$. Hence f is a local diffeomorphism around all points (e, x) with $x \in S$. Furthermore, since f is G-equivariant, using left translations it follows that the same is true at all points (g, x) with $g \in G$, $x \in S$. We claim that, after eventually shrinking S even more, one may assume that the map f is injective on $G \times S$. We do that by assuming the contrary: then we find $(g_k, x_k) \neq (g'_k, x'_k)$ in $G \times S$ such that $g_k x_k = g'_k x'_k$; from the freeness, it follows that $g_k \neq g'_k$. Set $a_k = (g'_k)^{-1} g_k$. Then

$$x'_k = a_k x_k,\ \ \ a_k \neq e.$$

Now, since $(a_k x_k, x_k) \to (x_0, x_0)$, this sequence will stay inside a compact region hence, using the properness of the action (and the fact that compacts are sequentially compact) it follows, eventually after passing to a subsequence of a_k, that (a_k) is convergent to some $a \in G$. Since both x_k and $a_k x_k = x_k^{'}$ converge to x_0, we find $a\, x_0 = x_0$ hence, by freeness, $a = e$. Now, we have

$$f(e, x_k) = f(a_k^{-1}, a_k x_k).$$

Now, both (e, x_k) and $(a_k^{-1} a_k x_k)$ belong to $G \times S$ and converge to (e, x_0); moreover, we also know that f is injective (even a diffeomorphism) in a neighborhood V of (e, x_0) in $G \times S$. Hence, for k large enough, $(e, x_k) = (a_k^{-1} a_k x_k)$; in particular, $a_k = e$, which gives the desired contradiction.

Since we now arranged that $f : G \times S \to M$ is an injective local diffeomorphism; it will be a diffeomorphism into its image; denoting by U the image (necessarily an open in M!), f will identify U equivarianty with $G \times S$ and the restriction of π to U with the projection $G \times S \to S$, proving the desired existence of the smooth structure.

□

Exercise 48. Consider the space of k-tuples of orthonormal vectors in $\mathbb{R}^{n+k}$ (known under the name of Stiefel manifold):

$$V_k(\mathbb{R}^{n+k}) = \{(v_1, \ldots, v_k) : e_i \in \mathbb{R}^{n+k} \ \langle v_i, v_j \rangle = \delta_{i,j}\}.$$

Consider

$$M = O(n + k), \quad G = O(n)$$

and the right action of G on M arising by interpreting an $n \times n$ matrix as an $(n+k) \times (n+k)$ one and by using the right action of $O(n+k)$ on itself. Show that for any $A \in O(n+k)$, denoting by A^i the ith row of A intepreted as a vector,

$$A^i = (A^i_1, \ldots, A^i_{n+k}) \in \mathbb{R}^{n+k},$$

then

$$\pi(A) := (A^1, \ldots, A^k) \in V_k(\mathbb{R}^{n+k}).$$

You now have all ingredients to put a smooth structure on $V_k(\mathbb{R}^{n+k})$. Do it!

What do you get when $k = 1$?

Exercise 49. Use now

$$M = V_k(\mathbb{R}^{n+k}),$$

the Stiefel manifold from the previous exercise, to exhibit an interesting smooth structure on the Grasmanian $\mathrm{Gr}_k(\mathbb{R}^{n+k})$ (Example 1.7): uniquely determined by the condition that

$$\pi : V_k(\mathbb{R}^{n+k}) \to \mathrm{Gr}_k(\mathbb{R}^{n+k}), \ (v_1, \ldots, v_k) \mapsto \mathrm{span}_{\mathbb{R}}(v_1, \ldots, v_k)$$

becomes a submersion.

(Hint: $G = O(k)$).

2.2. Principal bundles

2.2.1. The definition/terminology

Here are some basics on principal bundles.

Definition 2.36. *Let M be a manifold and G a Lie group. A principal G-bundle over M consists of a manifold P together with*

- *a right action of G on P, $P \times G \longrightarrow P$, $(p, g) \mapsto pg$,*
- *a surjective map*
$$\pi : P \longrightarrow M$$
which is G-invariant (i.e. $\pi_P(pg) = \pi_P(p)$ for all p and g).

satisfying the following local triviality condition: for each $x_0 \in M$, there exists an open neighborhood U of x_0 and a diffeomorphism

$$\Psi : P_U := \pi^{-1}(U) \longrightarrow U \times G$$

which maps each fiber $\pi^{-1}(x)$ to the fiber $\{x\} \times G$ and which is G-equivariant; here, the right action of G on $U \times G$ is the one on the last factor:

$$(x, a)g = (x, ag). \tag{2.37}$$

As for vector bundles, the local triviality conditions says that any point in M has a neighborhood U such that $P|_U$ is isomorphic to the trivial principal G-bundle where we use the following two notions:

Trivial principal bundles: the trivial principal G-bundle over a manifold M is $M \times G$, endowed with the second projection and the action of G given by (2.37).

Isomorphisms: given two principal G-bundles $\pi_i : P_i \to M$, $i \in \{1, 2\}$, an isomorphism between them is a diffeomorphisms $F : P_1 \to P_2$ which commutes with the projections ($\pi_s \circ F = \pi_1$) and which is G-equivariant ($F(pg) = F(p)g$ for all $p \in P_1$, $g \in G$).

Exercise 50. Define the notion of morphism between two principal G-bundles over a manifold M. Show that any morphism is an isomorphism.

Exercise 51. Show that for any principal G-bundle P the action of G on P is free.

Given a principal G-bundle $\pi : P \to M$ one can still talk about:

Fibers: One has the fiber

$$P_x := \pi^{-1}(x)$$

above any point $x \in M$. This time, this is no longer a vector-space but a G-a manifold, where the (right) action of G is obtained by restriction the one on P. Even more, while the fibers of a rank r vector bundle are vector spaces isomorphic (non-canonically!) to the Euclidean model $\mathbb{R}^r$, the fibers of P are G-spaces (but not groups!) diffeomorphic (as G-spaces!) to G itself (where G is viewed as a G-space with the right action given by right translations).

Exercise 52. Fix $x \in M$. Show that any $p \in P_x$ induces an equivariant diffeomorphism between P_x and G. How does it depend on the choice of p?

As a more intrinsic formulation of this property of the fibers, we can say that for any two points $p, q \in P$ in the same fiber there is a unique element

$$[p : q] \in G \quad \text{(read: } p \text{ divided by } q\text{)}$$

such that

$$q \cdot [p : q] = p. \tag{2.38}$$

Sections: A section of $\pi : P \to M$ is a smooth map $\sigma : M \longrightarrow P$ which sends each $x \in M$ into an element $\sigma(x)$ in the fiber P_x (equation-wise: $\pi_P \circ \sigma = \mathrm{Id}_M$). Similarly one can talk about local sections over some open $U \subset M$: sections of $P|_U$.

Exercise 53. Show that a principal bundle admits a (global) section if and only if it is trivializable. In more detail: show that if σ is a section of the principal G-bundle $\pi : P \to M$ then

$$F_\sigma : M \times G \to P, \; F_\sigma(x, g) = \sigma(x) \cdot g$$

is an isomorphism of principal G-bundles; and then show that $\sigma \longleftrightarrow F_\sigma$ defines a 1-1 correspondence between sections of P and isomorphisms between the trivial principal G-bundle and P.

Hence the local triviality condition can be rephrased as saying that for any point in M one can find a local section defined on a neighborhood of the point.

Exercise 54. Compare the last sentence with the fact that, for vector bundles, the local triviality condition is equivalent to the fact that for any point in M one can find a local frame defined on a neighborhood of the point; a bit more precisely, using this analogy as a hint, try to associate to any vector bundle over M a principal bundle over M. What is the relevant Lie group?

2.2.2. Remarks on the definition

Here are some of the consequences of the axioms for a principal G-bundle $\pi : P \longrightarrow M$ (please prove the ones that we did not do yet!):

- π is a submersion.
- π is G-invariant or, equivalently, the action of G preserves the fibers of π.
- the action of G on P is free and proper.
- The map

$$P \times G \longrightarrow P \times_M P, \quad (p, g) \mapsto (p, pg). \tag{2.39}$$

  is a diffeomorphism; its inverse is given by $(p, q) \longrightarrow (p, [p : q])$.

Moreover, one can slightly change the point of view and take some of these consequences as axioms- giving rise to slightly different (but equivalent) definitions of the notion of principal G-bundle. For instance, one can realize that the entire information is contained in the manifold P and the action of G (without any reference to M or π):

Proposition 2.40. *An action of a Lie group G on a manifold P is part of the structure of a principal G-bundle if and only if the action is free and proper.*

Moreover, in this case, the base manifold must be (diffeomorphic to) $M = P/G$ endowed with the unique smooth structure that makes the quotient map $\pi_{can} : P \to M$ into a submersion, and the principal bundle projection must be π_{can} itself.

Proof. Theorem 2.35. □

Before we mention other ways of encoding principal bundles, let us first take advantage of the previous proposition and provide a few examples of principal bundles. We start with some concrete ones.

Exercise 55. Show that the operation which associates to a matrix $A \in SO(3)$ its first column defines a map

$$\pi_0 : SO(3) \to S^2$$

and, together with the right action of S^1 on $SO(3)$ given by

$$A \cdot \lambda = A \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{for } \lambda = \cos(\alpha) + i\sin(\alpha) \in S^1, \tag{2.41}$$

$SO(3)$ becomes a principal S^1-bundle over S^2. Then generalize to higher dimensions.

Exercise 56 (The Hopf fibration)**.** Consider the S^1-action on S^3 given by

$$(x, y, z, t) \cdot (a, b) = (ax - by, ay + bx, az - bt, at + bz)$$

and

$$\pi : S^3 \longrightarrow S^2, \quad \pi(x, y, z, t) = (x^2 + y^2 - z^2 - t^2, 2(yz - tx), 2(ty + xz)).$$

Check that, in complex coordinates, writing

$$S^3 = \{(z_0, z_1) \in \mathbb{C}^2 : |z_0|^2 + |z_1|^2 = 1\},$$

we have

$$(z_0, z_1) \cdot \lambda := (z_0\lambda, z_1\lambda), \ \pi(z_0, z_1) = (|z_0|^2 - |z_1|^2, i\overline{z}_0 z_1). \tag{2.42}$$

Then show that $\pi : S^3 \to S^2$ is a principal S^1-bundle.

Exercise 57. Here is the relationship between the Hopf fibration and the (simpler) fibration from Exercise 55. For that we use the space of quaternions from Example 2.2 to identify S^3 with the space of quaternionic numbers of norm 1,

$$S^3 \xrightarrow{\sim} \{u \in \mathbb{H} : |u| = 1\}, \ (x, y, z, t) \mapsto x + iy + jz + kt$$

and $\mathbb{R}^3$ with the space of pure quaternions

$$\mathbb{R}^3 \xrightarrow{\sim} \{v \in \mathbb{H} : v + v^* = 0\}, \ (a, b, c) \mapsto ai + bj + ck.$$

Show that

$$A_u(v) := u^* v u$$

defines, for each $u \in S^3$, a linear map $A_u : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ which, as a matrix, gives

$$A_u \in SO(3).$$

Consider the induced map

$$\phi : S^3 \longrightarrow SO(3), \ u \mapsto A_u.$$

Show that:

1. The projections of the principal bundles from the previous two exercises are related via ϕ, i.e.:
$$\pi = \pi_0 \circ \phi.$$

2. However, ϕ is not S^1-equivariant. Instead,
$$\phi(u\lambda) = \phi(u) \cdot \lambda^2 \ \ \forall\, u \in S^3,\ \lambda \in S^1,$$
where $u \cdot \lambda$ uses to the action of S^1 on S^3 from Exercise 56 (given by (2.42)) and $\phi(u)\cdot\lambda^2$ uses the action of S^1 on $SO(3)$ from Exercise 55 (given by (2.41)).
3. Deduce also that $SO(3)$ can be identified with (i.e. it is diffeomorphic to) the real projective space $\mathbb{P}^3$

$$\begin{tikzcd} S^3 \arrow[rr, "\phi"] \arrow[dr, "\pi"'] & & SO(3) \arrow[dl, "\pi_0"] \\ & S^2 & \end{tikzcd}$$

Exercise 58 (The generalized Hopf fibrations)**.** Taking in the previous proposition
$$P = S^{2n+1} = \{z = (z_0, \ldots, z_n) \in \mathbb{C}^n : |z_0|^2 + \ldots + |z_n|^2 = 1\}, \quad G = S^1,$$
with the action given by complex multiplication, deduce that $\mathbb{C}P^n$ is a smooth manifold and S^{2n+1} is a principal S^1-bundle over $\mathbb{C}P^n$, with projection
$$\pi_1 : S^{2n+1} \to \mathbb{C}P^n, \pi_1(z_1, \ldots, z_n) = [z_0 : \ldots : z_n],$$
where $[z_0 : \ldots : z_n]$ is the complex line through the origin and the point $(z_0 : \ldots : z_n)$.

Then, in the case $n = 1$, show that we recover the Hopf fibration. More precisely, show that there exists a diffeomorphism
$$\Psi : S^2 \xrightarrow{\sim} \mathbb{CP}^1 \tag{2.43}$$
uniquely determined by $\pi_1 = \Psi \circ \pi$.

And here is a diagram that puts together (parts of) the previous two exercises:

$$\begin{tikzcd} S^3 \arrow[rr, "\phi"] \arrow[dr, "\pi"'] & & SO(3) \arrow[dl, "\pi_0"] \arrow[dr, "\pi_1"] & \\ & S^2 \arrow[rr, "\Psi", "\sim"'] & & \mathbb{C}P^1 \end{tikzcd}$$

Exercise 59. As a generalization of the previous exercise, and as a continuation of exercise 49, show that $V_k(\mathbb{R}^{n+k})$ can be made into a principal $O(k)$-bundle over $\mathrm{Gr}_k(\mathbb{R}^{n+k})$.

And here are some more applications of Proposition 2.40, this time to "general constructions".

Exercise 60 (Pull-backs)**.** Given $f : N \longrightarrow M$ smooth and a principal G-bundle P over M, one forms
$$f^*(P) := N \times_M P := \{(x, p) \in N \times P : f(x) = \pi(p)\}$$
with projection the standard projection on the first coordinate (built so that its fiber at $x \in M$ is just a copy of $P_{f(x)}$), with the right action of G:
$$(x, p)g = (x, pg).$$
Show that f^*P is a principal G-bundle over P.

While there is no direct analogue of the other operations on vector bundles such as direct sums, duals, etc, there is a (related) notion of push-forward along morphisms of Lie groups:

Exercise 61. Morphisms $\rho : H \longrightarrow G$ of Lie groups induce operations $P \mapsto \rho_*(P)$ which associate to a principal H-bundle P a principal G-bundle $\rho_*(P)$.

To be more precise, let $\pi : P \to M$ be a principal H-bundle. Consider

$$\rho_*(P) := (P \times G)/H,$$

the quotient of $P \times G$ modulo the action of H given by $(p, g)h = (ph, g\rho(h))$. Show that, denoting by $[p, g] \in \rho_*(P)$ the class of $(p, g) \in P \times G$, the "projection"

$$\tilde{\pi} : \rho_*(P) \longrightarrow M, \ [p, g] \mapsto \pi(p)$$

and the G-action $[p, g] \cdot g' = [p, gg']$ make $\rho_*(P)$ into a principal G-bundle over M.

Exercise 62. Consider the following two principal S^1-bundles over S^2: $SO(3)$ as in Exercise 55 and S^3 as in Exercise 56. Consider the group homomorphism

$$\rho : S^1 \to S^1, \ \rho(z) = z^2.$$

Using the map ϕ from Exercise 57 show that $\rho_*(S^3)$ is isomorphic to $SO(3)$, as principal S^1-bundles.

We now return to the general remarks on the principal bundle axioms; here is another way:

Exercise 63. Assume that G acts on the manifold P from the right and that $\pi : P \longrightarrow M$ is a surjective submersion. Show that P is a principal G-bundle if and only if (2.39) is well-defined and it is a diffeomorphism.

Yet another (equivalent) way to look at/define/produce principal bundles is to start with a setting in which P is only a set (endowed with an action of G) and the rest of the data is used to put the smooth structure on P.

Exercise 64. Formulate and prove an analogue of Exercise 1 for principal bundles.

And, in the same direction, one has the description of principal bundles in terms of transition functions that we now outline. Let $\pi : P \to M$ be a principal G-bundle in the sense of our original definition. In particular, we find a collection of local sections

$$\sigma_i \in \Gamma(P|_{U_i}), \quad i \in I \quad (I - \text{some indexing set})$$

defined over opens $U_i \subset M$ such that $\mathcal{U} = \{U_i\}_{i \in I}$ is an open cover of M. On the overlaps $U_i \cap U_j$ we can consider the quotient of σ_j and σ_i:

$$g_{i,j} : U_i \cap U_j \to G, \quad g_{i,j}(x) = [\sigma_j(x) : \sigma_i(x)].$$

(see (2.38)). These functions are called the transition functions of P with respect to the given collection of local sections. As it can readily be seen, they satisfy the so called "cocycle relations": for $x \in U_i \cap U_j \cap U_k$, one has

$$g_{i,k}(x) = g_{i,j}(x) g_{j,k}(x).$$

One also has $g_{i,i}(x) = e$ for all $x \in U_i$ and $g_{j,i}(x) = (g_{i,j}(x))^{-1}$ for all $x \in U_i \cap U_j$, but these are just formal consequences of the previous equations (check that!).

Exercise 65. Assume that M is a manifold, G is a Lie group, $\mathcal{U} = \{U_i\}_{i \in I}$ is an open cover of M and $g_{i,j} : U_i \cap U_j \to G$ is a collection of smooth functions satisfying the previous cocycle relations.

Construct a principal G-bundle $\pi : P \to M$ with the property that it admits a collection of local sections as above, so that the associated transition functions are precisely $g_{i,j}$. Then show that any two principal bundles with this property are isomorphic.

2.2.3. From vector bundles to principal bundles (the frame bundle)

The main aim of this and the next subsection is to prove (and understand) in detail the intepretation of vector bundles via principal bundles; as a simplified version, we state here:

Theorem 2.44. *For any manifold M and any $r \geq 0$ integer, there is a 1-1 correspondence between (isomorphism classes of) vector bundles of rank r over M and principal GL_r-bundles over M.*

In this subsection we go in one direction. Actually, with Exercise 54 in mind, it should be rather clear how to go from vector bundles to principal bundles: starting with a vector bundle$E$ of rank r over M (real or complex), one forms the frame bundle associated to E

$$\mathrm{Fr}(E) = \{(x, u) : x \in M, u - \text{a frame of } E_x\}.$$

It comes with an action of GL_r from the right, that we could just write down explicitly right away; but this is a good moment to discuss a bit more the notion of frame and be clearer about our conventions. For simplicity, we work over $\mathbb{R}$.

First of all, we write a matrix $A \in gl_r$ as

$$A = (A^i_j)_{1 \leq i, j \leq r}$$

by which we mean that

$$A = \begin{pmatrix} A^1_1 & \dots & A^1_r \\ \dots & \dots & \dots \\ A^r_1 & \dots & A^r_r \end{pmatrix}$$

Therefore, the usual product of two matrices is given by

$$(A \cdot B)^i_j = \sum_k A^i_k B^k_j.$$

Any matrix A gives rise to a linear map

$$\hat{A} : \mathbb{R}^r \to \mathbb{R}^r$$

defined on the canonical basis $\{e_1, \dots, e_r\}$ of $\mathbb{R}^r$ by

$$\hat{A}(e_j) = \sum_{i=1}^{r} A^i_j e_i;$$

however, we often omit the hat from the notation, i.e. we will identify the matrix A with the linear map that it induces. Note that, for two matrices A and B, one has

$$\widehat{AB} = \hat{A} \circ \hat{B},$$

i.e. the notations are chosen so that, when writing AB, it does not matter whether we interpret it as matrix multiplication or as composition of linear maps.

Keeping for the moment the notation $\hat{A}$ note that, on a general vector $v = v^1 e_1 + \dots + v^r e_r$, one has

$$\hat{A}(v) = \sum_i (\sum_j A^i_j v^j) e_i.$$

To write also this formula in terms of matrix multiplication, one associates to any $v \in \mathbb{R}^r$ the column matrix

$$[v] = \begin{pmatrix} v^1 \\ \dots \\ v^r \end{pmatrix}$$

With this, the previous formula becomes

$$[\hat{A}(v)] = A \cdot [v]$$

where "$\cdot$" refers to the multiplication of matrices. While we identify matrices A with the associated linear maps $\hat{A}$, we will also identify the vectors $v \in \mathbb{R}^r$ with the associated column matrices $[v]$; so, we will just omit the brackets in the notation $[\cdot]$.

We now pass to an arbitrary r-dimensional vector space V and we consider the set $\mathrm{Fr}(V)$ of all frames of V, i.e ordered collections

$$\phi = (\phi_1, \dots, \phi_r)$$

of vectors $\phi_i \in V$ that form a basis of V. While any r-dimensional vector space V can be identified with (i.e. it is isomorphic to) the standard model $\mathbb{R}^r$, there is no such canonical identification. Choosing a frame should be seen as choosing such an identification; indeed, one has the following simple:

Exercise 66. For $\phi = (\phi_1, \dots, \phi_r) \in \mathrm{Fr}(V)$, consider the linear map

$$\hat{\phi} : \mathbb{R}^r \to V$$

that sends the component e^i of the standard basis of $\mathbb{R}^r$ to ϕ^i, for each i. Show that $\phi \longleftrightarrow \hat{\phi}$ defines a 1-1 correspondence between frames of V and linear isomorphism from $\mathbb{R}^r$ to V.

It should now be clear how GL_r acts on $\mathrm{Fr}(V)$ (from the right!):

$$\mathrm{Fr}(V) \times GL_r \to \mathrm{Fr}(V), \quad (\phi, A) \mapsto \phi \cdot A$$

so that

$$\widehat{\phi \cdot A} = \hat{\phi} \circ \hat{A}.$$

(and, again, we just omit the hat from the notation $\hat{\phi}$). We find the explicit formula:

$$(\phi \cdot A)_i = \sum^{i} A_i^j \phi_j.$$

Returning to a general rank r vector bundle $E \to M$, the frame bundle $\mathrm{Fr}(E)$ just puts together the frame spaces $\mathrm{Fr}(E_x)$ of all the fibers; the action of GL_r on these spaces defines a (free, right) action of GL_r on $\mathrm{Fr}(E)$. There are various ways to exhibit the smooth structure on $\mathrm{Fr}(M)$; one of the most intuitive ones comes from the fact that we know which local sections of $\mathrm{Fr}(E)$ we want to be smooth: any local frame of the vector bundle E. This (together with Exercise 64) specifies the smooth structure on $\mathrm{Fr}(E)$ and it is not difficult to see that it does, indeed, a principal GL_r-bundle over M.

Exercise 67. Provide the details.

Exercise 68. Describe an embedding $i : SO(3) \hookrightarrow \mathrm{Fr}(TS^2)$ such that the composition of the projection $\mathrm{Fr}(TS^2) \to S^2$ with i is the projection π_0 from Exercise 55. Why is i an embedding? Make sense of the statement: "$SO(3)$ is the oriented orthonormal frame bundle associated to TS^2".

Here are some exercises which show that, if one looks for an analogue of the various standard operations on vector bundles (such as taking duals, direct sums etc) that applies to principal bundles, then one has to look at morphisms of Lie groups and the corresponding push-forward operations (Exercise 61):

Exercise 69. Consider the group homomorphism

$$\rho : GL_r \longrightarrow GL_r, \quad \rho(A) = {}^t(A^{-1}).$$

Show that for any vector bundle E, the frame bundles associated to E and to its dual are related by

$$\mathrm{Fr}(E^*) \cong \rho_*(\mathrm{Fr}(E)).$$

Exercise 70. Show that if $\pi_i : P_i \longrightarrow M$ are principal G_i-bundles, for $i \in \{1, 2\}$, then "the fibered product"

$$P_1 \times_M P_2 := \{(p_1, p_2) \in P_1 \times P_2 : \pi_1(p_1) = \pi_2(p_2)\}$$

carries a natural structure of principal $G_1 \times G_2$-bundle.

Assume now that $P_i = \mathrm{Fr}(E_i)$ is the frame bundles associated to a vector bundles E_i of rank r_i (hence also $G_i = GL_{r_i}$), $i \in \{1, 2\}$. Consider $r = r_1 + r_2$ and the group homomorphism

$$\rho : GL_{r_1} \times GL_{r_2} \longrightarrow GL_r, \quad (A, B) \mapsto A \oplus B := \mathrm{diag}(A, B).$$

Show that the frame bundle associated to $E_1 \oplus E_2$ is isomorphic to the push-forward of $P_1 \times_M P_2$ via ρ.

2.2.4. From principal bundles to vector bundles

We now describe a reverse construction, which associates to principal GL_r-bundles vector bundles. Actually, we proceed in a slightly more general setting, for a general Lie group G.

What we need is a principal G-bundle $\pi : P \longrightarrow M$ as well as a representation $\rho : G \longrightarrow GL(V)$ of G on a finite dimensional vector space (again, over $\mathbb{R}$ or $\mathbb{C}$). To such data we will associate a vector bundle

$$E(P, V)$$

over M, of rank equal to the dimension of V, called the vector bundle obtained by attaching to P the representation (V, ρ) (a more faithful notation would be $E(P, G, V, \rho)$, but G and ρ are usually clear from the context). Of course, when returning to the 1-1 correspondence between vector bundles and principal bundles, we will apply this construction to $G = GL_r$, and ρ the canonical representation of matrices as linear maps on $V = \mathbb{R}^r$.

In this general farmework, the representation ρ encodes a linear action of G on V from the left which, combined with the right action of G on P, induces a (right) action of G on $P \times V$:

$$(p, v) \cdot g := (pg, g^{-1}v).$$

We then define $E(P, V)$ as the resulting quotient

$$E(P, V) := (P \times V)/G$$

and we denote by $[p, v] \in E(P, V)$ the element in $E(P, V)$ represented by $(p, v) \in P \times G$; hence, in the quotient,

$$[pg, v] = [p, gv]$$

for all $p \in P$, $v \in V$, $g \in G$. We endow $E(P, V)$ with the projection

$$\tilde{\pi} : E(P, V) \longrightarrow M, \quad \tilde{\pi}(p, v) = \pi(v).$$

Note that the fiber of $\tilde{\pi}$ above an arbitrary point $x \in M$,

$$E(P, V)_x = \{[p, v] : p \in P_x, v \in V\}.$$

The following is a simple exercise:

Exercise 71. Show that there exists a unique vector space structure on the fibers $E(P, V)_x$ such that, for each $p \in P$ with $\pi(p) = x$,

$$\phi_p : V \to E(P, V)_x, \quad v \mapsto [p, v],$$

is an isomorphism of vector spaces.

To describe the smooth structure on $E(P, V)$ one can proceed in several ways. One is to use again the theorem on the smoothness of quotients (Theorem 2.35), i.e. apply Prroposition 2.40; for that one just needs the simple remark that, since the action of G on P is free and proper, so is the action on $P \times V$. Another way to see that $E(P, V)$ is a smooth vector bundle is to use local trivializations of P:

$$\Psi : P|_U \longrightarrow U \times G$$

and note that any such trivialization induces a trivialization of $E(P, V)$:

$$\tilde{\Psi} : \tilde{\pi}^{-1}(U) \longrightarrow U \times V, \; p \mapsto (\pi(p), g_{\Psi(p)} v),$$

where $g_{\Psi(p)} \in G$ is the second component of $\Psi(p)$. Of course, the local trivializations $\tilde{\Psi}$ will serve as charts of a smooth atlas for $E(P, V)$.

Corollary 2.45. *For any principal G-bundle P over M and any representation V of G, $E(P, V)$ is a vector bundle over M.*

As we mentioned, this construction is particularly interesting in the case when

$$G = GL_r, \; V = \mathbb{R}^r, \; \rho = \mathrm{Id},$$

when it associates to a principal GL_r bundle P the vector bundle $E(P, \mathbb{R}^r)$. Let us apply this to the frame bundle

$$P = \mathrm{Fr}(E)$$

of a vector bundle E of rank r over M. In this case, since an element $p \in P_x$ is the same thing as the choice of an isomorphism $i_p : \mathbb{R}^r \longrightarrow E_x$, we see that that a pair $[p, v] \in E(P, \mathbb{R}^r)_x$ can be identified with $i_p(v) \in E_x$. In other words, one has an isomorphism of vector bundles over M:

$$E(P, \mathbb{R}^r) \xrightarrow{\sim} E, \quad [p, v] \mapsto i_p(v).$$

Putting everything together, we find:

Theorem 2.46. *The constructions which associate:*

- *to a vector bundle E its frame bundle $Fr(E)$*
- *to a principal GL_r-bundle P the vector bundle $E(P, \mathbb{R}^r)$*

defines a 1-1 correspondence between (isomorphism classes of) vector bundles of rank r over M and principal GL_r-bundles over M.

Here are some concrete illustrations of the general construction from this subsection.

Exercise 72. Consider the principal S^1-bundle $\pi_0 : SO(3) \longrightarrow S^2$ from Exercise 55. Consider also the representation of S^1 on $\mathbb{R}^2$ given by

$$r : S^1 \to GL_2(\mathbb{R}) = GL(\mathbb{R}^2),\ \cos(\alpha) + i\sin(\alpha) \mapsto \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Show that the associated vector bundle $E(SO(3), \mathbb{R}^2, r)$ is isomorphic to the tangent bundle of S^2.

Exercise 73. Consider the Hopf fibrations, i.e. the principal S^1-bundle $\pi : S^3 \longrightarrow S^2$ from Exercise 55. Consider also the square r^2 of the representation from the previous exercise, i.e.

$$r^2 : S^1 \to GL_2(\mathbb{R}) = GL(\mathbb{R}^2),\ \cos(\alpha) + i\sin(\alpha) \mapsto \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ -\sin(2\alpha) & \cos(2\alpha) \end{pmatrix}.$$

Show that the associated vector bundle $E(S^3, \mathbb{R}^2, r^2)$ is isomorphic to the tangent bundle of S^2.

Exercise 74. Consider again the Hopf fibration, but use now the representation r to build the new vector bundle $E = E(S^3, \mathbb{R}^2, r)$ over S^2. Show that this corresponds to the tautological line bundle over $\mathbb{C}P^1$. More precisely:

- recall that we have the tautological complex line bundle L over $\mathbb{C}P^1$ defined as the following sub-bundle of the trivial bundle $\mathbb{C}P^1 \times \mathbb{C}^2$:
$$E := \{(l, v) : l \in \mathbb{CP}^1, v \in l\} \subset \mathbb{CP}^1 \times \mathbb{C}^2,$$
- recall that there we have a diffeomorphism $\Psi : S^2 \to \mathbb{C}P^1$ constructed in Exercise 58 (see (2.43)),

and show that E is isomorphic to $\Psi^*(L)$, as vector bundles over S^2.

2.2.5. Sections/forms for vector bundles associated to principal ones

Due to the importance of sections of a vector bundle E as well as of the space of E-valued differential forms (e.g. when discussing connections), when

$$E = E(P, V)$$

is the vector bundle associated to a principal G-bundle $\pi : P \to M$ and a representation $\rho : G \to GL(V)$, it is useful to rewrite those spaces more directly in terms of the starting data: the principal bundle and the representation. That is what we do in this subsection; hence, we place ourselves in the previous setting.

For sections of $E(P, V)$, the important space is $C^\infty(P, V)$- the space of smooth functions on P with values in the vector space V, and the fact that this space comes enowed with a canonical (linear) action of G:

$$G \times C^\infty(P, V) \to C^\infty(P, V), \quad (g, f) \mapsto g \cdot f,$$

induced by the actions of G on P and V by:

$$(g \cdot f)(p) = \rho_g(f(pg)).$$

Exercise 75. Check that this defines, indeed, a left action.

As for any action of a group, we can talk about the space of G-invariant elements:

$$C^\infty(P, V)^G := \{f \in C^\infty(P, V) : g \cdot f = f \ \ \forall\, g \in G\}.$$

The key-point is that any such G-equivariant function f induces a section of $E(P,V)$: first of all it induces

$$(\mathrm{id}, f) : P \longrightarrow P \times V,\ p \mapsto (p, f(p))$$

which is equivariant (where the right action of G on $P \times V$ is precisely the one used to define $E(P,V)$); hence, passing to the quotient modulo G, it induces a map

$$s_f : M \longrightarrow E(P,V).$$

Lemma 2.47. *If $E = E(P,V)$ is the associated vector bundle, then one has a bijection*

$$C^\infty(P,V)^G \xrightarrow{\sim} \Gamma(E), \quad f \mapsto s_f.$$

Proof. Smooth functions $f : P \to V$ are in 1-1 correspondence with sections of the trivial vector bundle $P \times V$ over P: to f one associates

$$\tilde{s}_f : P \to P \times V,\ \tilde{s}_f(p,v) = (p, f(v)).$$

Moreover, the equivariance of f is equivalent to s_f being equivariant, where $P \times V$ is endowed with the action of G used to define E. This implies that $\tilde{s}_f$ descends to a map

$$s_f : M \to (P \times V)/G = E,$$

which is a section of E. A minute of thinking reveals that this discussion can easily be reversed. □

Let us now generalize this construction to E-valued differential forms on M of arbitrary degreee k (i.e. going from $k = 0$ to arbitrary k). Of course, instead of $C^\infty(P,V)$ we now consider the space $\Omega^k(P,V)$ of forms on P with values in the vector space V. As above, one has an action of G on this space:

$$G \times \Omega^k(P,V) \to P, (g, \omega) \mapsto g \cdot \omega,$$

built out from the action of G on P and V; explicitely,

$$(g \cdot \omega)(-) := \rho_g(R_g^*(\omega)(-)),$$

where R_g^* is the pull-back of forms via the right multiplication by g, $R_g : P \longrightarrow P$, $R_g(p) = pg$. Hence, on vectors $X_p^i \in T_pP$, $1 \le i \le k$:

$$(g \cdot \omega)_p(X^1, \ldots, X^k) = \rho(g)(\omega_{pg}((dR_g)_p(X^1), \ldots, (dR_g)_p(X^k))).$$

This allows us to talk about G-invariant V-valued forms on P. However, this is not quite the space we are interested in. We also have to make use of the infinitesimal action of $\mathfrak{g}$ (the Lie algebra of G) on P induced by the action of G on P (cf. Proposition 2.29):

$$a : \mathfrak{g} \longrightarrow \mathcal{X}(P),\ a(v)_p = a_p(v) := \left.\frac{d}{dt}\right|_{t=0} p \cdot \exp(tv).$$

Definition 2.48. *We say that a differential form $\omega \in \Omega^k(P,V)$ is:*

- *equivariant: if $g \cdot \omega = \omega$ for all $g \in G$.*
- *horizontal: if $i_{a(v)}(\omega) = 0$ for all $v \in \mathfrak{g}$.*
- *basic: if it is both equivariant as well as horizontal.*

We denote by $\Omega^k(P,V)_{hor}$ the space of horizontal differential forms and by $\Omega^k(P,V)_{bas}$ the space of basic forms.

Here is an exercise explaining the term "horizontal":

Exercise 76. For $X_p \in T_pP$ ($p \in P$, $x = \pi(p)$), the following are equivalent:

1. X_p is tangent to the fiber $P_p = \pi^{-1}(x)$.
2. X_p is killed by $(d\pi)_p : T_pP \to T_xM$.
3. $X_p = a_p(v)$ for some $v \in \mathfrak{g}$ (necessarily unique).

When this happens we say that X is a vertical tangent vector (at p); the space of such vectors is denoted

$$T_p^{\text{vert}}P \subset T_pP \quad (\text{hence } T_p^{\text{vert}}P = \text{Ker}(d\pi)_p = T_p(P_p) = \text{Im}(a_p)).$$

A vector field X on P is called vertical if X_p is vertical for all $p \in P$. Deduce that a form $\omega \in \Omega^k(P, V)$ is horizontal if and only if it vanishes on all vertical vector fields.

Proposition 2.49. *If $E = E(P, V)$ is the associated vector bundle, then one has a linear isomorphism*

$$\pi^{\bullet} : \Omega^k(M, E) \xrightarrow{\sim} \Omega^k(P, V)_{bas}.$$

Proof. We start with two remarks. The first one is that that the pull-back π^*E of the vector bundle E via the projection $\pi : P \longrightarrow M$ is, as a vector bundle over P, isomorphic to the trivial vector bundle $P \times V$, by a canonical isomorphism

$$i : P \times V \longrightarrow \pi^*(E).$$

At the fiber at $p \in P$, this is simply

$$i_p : V \longrightarrow E_{\pi(p)}, \;\; v \mapsto [p, v].$$

The second remark is that the usual pull-back of forms along $\pi : P \longrightarrow M$ makes sense for forms with coefficients, giving rise to

$$\pi^* : \Omega^k(M, E) \longrightarrow \Omega^k(P, \pi^*E).$$

Explicitely, for $X^1, \ldots, X^k \in T_pP$,

$$\pi^*(\omega)(X^1, \ldots, X^k) = \omega((d\pi)_p(X^1), \ldots, (d\pi)_p(X^k)) \in E_{\pi(p)} = (\pi^*E)_p. \tag{2.50}$$

Combining the isomorphism from the previous remark, we consider

$$\pi^{\bullet} := i^{-1} \circ \pi^* : \Omega^k(M, E) \longrightarrow \Omega^k(P, V).$$

The rest of the proof consists of the following rather straightforward steps (details are left as small exercises):

1. Check directly that $\pi^{\bullet}$ actually takes values in $\Omega^k(P, V)_{\text{bas}}$ (for instance, since $(d\pi)$ kills the vertical vector fields, all forms of type $\pi^*\omega$ are horizontal).
2. The map $\pi^{\bullet}$ is injective (since the differential of π is surjective).
3. One still has to show that each $\eta \in \Omega^k(P, V)_{\text{bas}}$ can be written as $\pi^*\omega$ for some ω. Explicitely, one wants

$$\omega_{\pi(p)}((d\pi)_p(X^1), \ldots, (d\pi)_p(X^k)) = i_p(\eta_p(X^1, \ldots, X^k)).$$

for all $p \in P$ $X^i \in T_pP$. But this forces the definition of ω: for $x \in M$ and $V^i \in T_xM$, choosing $p \in P$ with $\pi(p) = x$ and choosing $X^i \in T_pP$ with $(d\pi)_p(X^i) = V^i$, we must have

$$\omega_x(V^1, \ldots, V^k) = i_p(\eta_p(X^1, \ldots, X^k)).$$

4. What is left to check is that this definition does not depend on the choices we made. First of all, once $p \in \pi^{-1}(x)$ is chosen, the formula does not depend on the choice of the X^i's- this follows from the fact that η is horizontal. The independence of the choice of the point p in $\pi^{-1}(x)$ follows from the G-invariance of η and the fact that any other $q \in \pi^{-1}(x)$ can be written as $q = pg$ for some $g \in G$.

$\square$

2.3. Connections on principal bundles

Throughout this subsection G is a Lie group and $\pi : P \longrightarrow M$ is a principal G-bundle. There are several different ways of looking at connections on P. We start with the more intuitive one.

2.3.1. Connections as horizontal distributions:

A horizontal subspace of TP at $p \in P$ is a subspace

$$\mathcal{H}_p \subset T_p P$$

with the property that

$$(d\pi)_p|_{\mathcal{H}_p} : \mathcal{H}_p \longrightarrow T_{\pi(p)}M \tag{2.51}$$

is an isomorphism.

Exercise 77. Show that, for a vector sub-space $\mathcal{H}_p \subset T_p P$, denoting $x = \pi(p)$ and using the vertical space $T_p^{\mathrm{vert}}P$ from Exercise 76, the following are equivalent:

1. $\mathcal{H}_p$ is horizontal.
2. $\mathcal{H}_p$ is a complement of T_p^{vert} inside T_pP.
3. $(d\pi)_p : T_pP \to T_xM$ restricts to an isomorphism between $\mathcal{H}_p$ and T_xM.
4. There exists a splitting of $(d\pi)_p$, i.e. a linear map

$$h_p : T_xM \to T_pP$$

satisfying $(d\pi)_p \circ h_p = \mathrm{Id}$, such that $\mathcal{H}_p$ is the image of h_p.
(Read: h_p is the operation of "horizontal lifting" associated to $\mathcal{H}_p$).

Note that if $\mathcal{H}_p$ is a horizontal space at $p \in P$ then, for any $g \in G$,

$$R_g(\mathcal{H}_p) := \{(dR_g)_p(X_p) : X_p \in \mathcal{H}_p\} \subset T_{gp}P$$

is another horizontal subspace at gp. By distribution on P we mean a vector sub-bundle $\mathcal{H}$ of TP (hence a collection $\{\mathcal{H}_p\}_{p\in P}$ of vector sub-spaces $\mathcal{H}_p$ of T_pP, all of the same dimension, which fits together into a smooth submanifold $\mathcal{H}$ of TP).

Definition 2.52. *A connection on P is a distribution $\mathcal{H}$ on P with the property that each $\mathcal{H}_p$ is a horizontal subspace of P and*

$$\mathcal{H}_{pg} = R_g(\mathcal{H}_p) \quad \forall\ p \in P, g \in G.$$

While the notion of vertical vector fields on P is intrinsic (it does not depend on any choices), a connection allows us to talk about horizontal vectors on P: the ones that lie in $\mathcal{H}$ (hence one should really make a reference to $\mathcal{H}$ in the terminology- e.g. talk about $\mathcal{H}$-horizontal). Note that any vector field on M has a lift to a horizontal vector field on P, i.e. there is an operation

$$h : \mathcal{X}(M) \longrightarrow \mathcal{X}(P)$$

(the horizontal lift operation associated to the connection) which associates to $X \in \mathcal{X}(M)$ the unique horizontal vector field $h(X)$ on P satisfying

$$(d\pi)_p(h(X)_p) = X_{\pi(p)}$$

for all $p \in P$. With the notations from the previous exercise, $h(X)_p := h_p(X_{\pi(p)})$.

For the curious reader. Putting all the possible horizontal liftings together we obtain the so-called first order jet of P,

$$J^1(P) := \{(p, h_p) : p \in P, h_p \text{ is a horizontal lifting of } P \text{ at } p\}.$$

This can be seen as a (discrete for now) bundle over P (using the projection on the first coordinate) and also as a bundle over M by further composing with the projection $\pi : P \longrightarrow M$. The terminology comes from the fact that this set encodes all the possible first order data (jets) associated to (local) sections of P. More precisely, given $x \in M$, any local section σ of P defined around x, induces:

- the 0-th order data at x: just the value $\sigma(x) \in P$.
- the 1-st order data at x: the value $\sigma(x)$ together with the differential of σ at p (which encodes the first order derivatives of σ at x),

$$(d\sigma)_x : T_xM \longrightarrow T_{\sigma(x)}P.$$

Of course, $(d\sigma)_x$ is a horizontal lifting at $\sigma(x)$ and it is not difficult to see that

$$J^1(P) = \{(\sigma(x), (d\sigma)_x) : x \in M, \sigma - \text{local section of } P \text{ defined around } p\}.$$

Note that for a section σ of P, the resulting map

$$j^1(\sigma) : M \longrightarrow J^1(P), \quad x \mapsto (\sigma(x), (d\sigma)_x))$$

is a section of $J^1(P)$ viewed as a bundle over P. Similarly for local sections. There is a natural smooth structure on $J^1(P)$ uniquely determined by the condition that all local sections of type $j^1(\sigma)$ are smooth.

2.3.2. Connection forms

Here is a slightly different point of view on connections, not so natural, but often easier to work with. Let $\mathfrak{g}$ be the Lie algebra of G. We will use the infinitesimal action of $\mathfrak{g}$ on P (see Proposition 2.29),

$$a : \mathfrak{g} \longrightarrow \mathcal{X}(P), \ a(v)_p = \left.\frac{d}{dt}\right|_{t=0} p \cdot \exp(tv),$$

and the adjoint representation of G on $\mathfrak{g}$ (see Example 2.21),

$$\mathrm{Ad} : G \longrightarrow GL(\mathfrak{g}), \quad g \mapsto \mathrm{Ad}_g.$$

The general discussion from subsection 2.2.5 gives us the notion of G-invariant forms on P with values in $\mathfrak{g}$: those $\theta \in \Omega^k(P, \mathfrak{g})$ with the property that

$$R_g^*(\omega) = \mathrm{Ad}_{g^{-1}}(\omega)$$

for all $g \in G$.

Definition 2.53. *A connection form on P is a 1-form*

$$\omega \in \Omega^1(P, \mathfrak{g})$$

with the property that it is G-invariant and satisfies

$$\omega(a(v)) = v \quad \forall \ v \in \mathfrak{g}.$$

Proposition 2.54. *There is a bijection between connection 1-forms $\omega \in \Omega^1(P, \mathfrak{g})$ and connections $\mathcal{H}$ on P. The bijection associates to ω its kernels:*

$$\mathcal{H}_p := \{X_p \in T_pP : \omega_p(X_p) = 0\}.$$

Proof. For each $p \in P$ we consider the short exact sequence of vector spaces

$$\mathfrak{g} \xrightarrow{a_p} T_pP \xrightarrow{(d\pi)_p} T_{\pi(p)}M.$$

A right splitting of this sequence is a right inverse of $(d\pi)_p$; as we have noticed, such right splittings h_p correspond to horizontal subspaces $\mathcal{H}_p$. On the other hand one can also talk about left splitting of the sequence, by which we mean a left inverse of a_p or, equivalently, linear maps $\omega_p : T_pP \longrightarrow \mathfrak{g}$ satisfying $\omega_p(a_p(v)) = v$ for all $v \in \mathfrak{g}$. What we need here is a simple linear algebra fact which says that, for a short exact sequence of vector space, there is a 1-1 correspondence between the choice of right splittings h_p and the choice of left splittings ω_p. Explicitely, this correspondence is described by the equation

$$a_p \circ \omega_p + h_p \circ (d\pi)_p = \mathrm{Id}.$$

After applying this at all $p \in P$, one obtains the 1-1 correspondence from the statement. Of course, one still has to check that the smoothness of ω is equivalent to the one of h and similarly for the G-invariance, but we leave these as an exercise. $\square$

Exercise 78. Consider the interpretation of S^3 in terms of quaternions (Example 2.2); remark that, for any $u \in S^3$,

$$iu, ju, ku \in \mathbb{H} \cong \mathbb{R}^4$$

are orthogonal to u, i.e. they define vectors tangent to S^3 at u. Hence, each of these elements defines a vector field on S^3; they are denoted X_1 (using i), X_2 (using j) and X_3 (using k). For instance,

$$X_1(x, y, z, t) = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} - t\frac{\partial}{\partial z} + z\frac{\partial}{\partial t} \in T_{x,y,z,t}S^3 \subset T_{x,y,z,t}\mathbb{R}^4.$$

We also consider the coframe $\theta^1, \theta^2, \theta^3$ associated to the frame $\{X_1, X_2, X_3\}$, i.e. the collection of 1-forms on S^3 uniquely determined by $\theta^i(X_j) = \delta^i_j$.

1. Show that the flow of X_1 induces the action of S^1 on S^3 from the Hopf fibration (Example 56).
2. Compute the differentials of each θ_i and write them in terms of wedge of θ_1, θ_2 and θ_3.
3. Show that θ^1 can be interpreted as a connection on the Hopf fibration.

2.3.3. Parallel transport

As in the case of vector bundles, connections on principal bundles have associated parallel transports. Assume we fix a connection $\mathcal{H}$ on the principal G-bundle $\pi : P \longrightarrow M$, with associated connection form ω. Then we say that a curve $u : I \longrightarrow P$ is horizontal if

$$\dot{u}(t) \in \mathcal{H}_{\gamma(t)} \quad \forall\, t \in I,$$

where $\gamma(t) = \pi(u(t))$. Then, completely similar to Lemma 1.21, we have:

Lemma 2.55. *Let $\gamma : I \longrightarrow M$ be a curve in M, $t_0 \in I$. Then for any $u_0 \in P_{\gamma(t_0)}$ there exists and is unique a horizontal path above γ, $u : I \longrightarrow P$, with $u(t_0) = u_0$.*

Proof. Again, we may work locally, i.e. we may just assume that $P = M \times G$ is the trivial principal bundle. Let us now see what a connection form on this bundle looks like. First of all, the invariance condition means that, for all $(x, a) \in M \times G$, $(X_x, V_a) \in T_{x,a}(M \times G) = T_xM \times T_aG$, and for all $g \in G$,

$$\omega_{x,ag}(X_x, R_g(V_a)) = \mathrm{Ad}_{g^{-1}}(\omega_{x,a}(X_x, V_a)),$$

where we denote by $R_g : G \longrightarrow G$ the right multiplication by g and also its differential $(dR_g)_a : T_aG \longrightarrow T_{ag}G$. Taking $a = 1$, we see that ω is determined by all the expressions of type $\omega_{x,1}(X_x, v)$ with $1 \in G$ the unit, $v \in T_1G = \mathfrak{g}$. In turn, the condition that $\omega(a(v)) = v$ for all $v \in \mathfrak{g}$ implies that ω is uniquely determined by all expressions of type

$$\omega_{x,1}(X_x, 0) \in \mathfrak{g}$$

i.e. precisely by the restriction $\eta \in \Omega^1(M, \mathfrak{g})$ of ω to M via the inclusion $M \hookrightarrow M\times G$, $x \mapsto (x, 1)$. Writing down the explicit formulas we find that the formula for ω in terms of η:

$$\omega_{x,g}(X_x, V_g) = \mathrm{Ad}_{g^{-1}}(\eta_x(X_x)) + L_{g^{-1}}(V_g),$$

where, as before, $L_{g^{-1}} : G \longrightarrow G$ stands for the left multiplication by g^{-1} and also for its differential $(dL_{g^{-1}})_g : T_gG \longrightarrow T_1G = \mathfrak{g}$.

A curve in $P = M \times G$ above γ can be written as

$$u(t) = (\gamma(t), g(t)).$$

The fact that it is horizontal means that $\dot{u}(t)$ takes values in the horizontal space or, equivalently, that it is killed by ω:

$$\omega_{\gamma(t),g(t)}(\dot{\gamma}(t), \dot{g}(t)) = 0 \quad \forall\, t \in I.$$

Using the previous formula for ω we see that this equation becomes

$$\dot{g}(t) = -R_{g(t)}(\eta_{\gamma(t)}(\dot{\gamma}(t))).$$

For fixed γ we need existence and uniqueness of g satisfying this equation, with an initial condition $g(t_0) = g_0$ given. We denote $v(t) = -\eta_{\gamma(t)}(\dot{\gamma}(t)$ (curve in $\mathfrak{g}$, so that the equation reads $\dot{g}(t) = R_{g(t)}(v(t)))$. At least for linear groups $G \subset GL_r$ we deal again (as in the case of vector bundles) with an ordinary ODE. For a general Lie group one can just interpret this equation as the flow equation of the time-dependent vector field $X(t, g)$ on G given by

$$X(t, g) = R_g(v(t))$$

and use the standard results about flows of time-dependent vector fields. However, one still has to make sure that the equation with the given initial condition $g(t_0) = g_0$ exists on the entire interval I on which v is defined; for that one uses the remark that, if $g(t)$ is an integral curve then so is $h(t) = g(t)a$ for all $a \in G$:

$$\dot{h}(t) = R_a(\dot{g}(t)) = R_a(R_{g(t)}(v(t)) = R_{h(t)}(v(t)).$$

$\square$

Finally, as in the case of vector bundles, for curves $\gamma : I \longrightarrow M$, $t_0, t_1 \in M$, one obtains the parallel transport (with respect to the connection) along γ:

$$T_{\gamma}^{t_0,t_1} : P_{\gamma(t_0)} \longrightarrow P_{\gamma(t_1)}$$

which associates to $u_0 \in P_{\gamma(t_0)}$ the element $u(t_1) \in P_{\gamma(t_1)}$, where u is the horizontal path above γ with $u(t_0) = u_0$.

Exercise 79. Show that $T_{\gamma}^{t_0,t_1}$ are diffeomorphisms and are G-equivariant.

2.3.4. Curvature

Let $\mathcal{H}$ be a connection on P, with associated horizontal lifting h and connection 1-form ω. The curvature of $\mathcal{H}$ measure the failure of $\mathcal{H}$ to be involutive or, equivalently, the failure of the horizontal lifting h to preserve the Lie bracket of vector fields. However, let us introduce it in the form that is easier to work with. For that we note that the Lie bracket $[\cdot,\cdot]$ of the Lie algebra $\mathfrak{g}$ extends to a bracket

$$[\cdot,\cdot] : \Omega^1(P,\mathfrak{g}) \times \Omega^1(P,\mathfrak{g}) \longrightarrow \Omega^2(P,\mathfrak{g}),$$

$$[\alpha,\beta](X,Y) := [\alpha(X),\beta(Y)] - [\alpha(Y),\beta(X)].$$

Note that when $\beta = \alpha \in \Omega^1(P,\mathfrak{g})$ one obtains

$$[\alpha,\alpha](X,Y) = 2[\alpha(X),\alpha(Y)].$$

Definition 2.56. *The curvature of the connection ω is defined as*

$$K := d\omega + \frac{1}{2}[\omega,\omega] \in \Omega^2(P,\mathfrak{g}).$$

Here is the main property of the curvature (for basic forms see subsection 2.2.5):

Lemma 2.57. *The curvature is a basic form:*

$$K \in \Omega^2(P,\mathfrak{g})_{bas}.$$

Proof. The G-invariance follows from the G-invariance of ω and the fact that all the operations involved in construction K (d and the bracket) are natural (functorial), hence are respected by the action of G (give the details!). The fact that K is horizontal ... □

Note that the fact that Ω is horizontal implies that, in order to know Ω, it suffices to know all the expressions of type

$$K(h(X),h(Y)) \in C^\infty(P,\mathfrak{g})$$

with $X,Y \in \mathcal{X}(M)$, where $h(X) \in \mathcal{X}(M)$ is the horizontal lift of X with respect to the connection. For the following, recall that $a : \mathfrak{g} \longrightarrow \mathcal{X}(P)$ denotes the induced infinitesimal action (see above).

Proposition 2.58. *The curvature can also be characterized as:*

1. *the horizontal part of $d\omega$, i.e. the horizontal form on P (with values in $\mathfrak{g}$) which coincides with $d\omega$ on horizontal vectors.*
2. *the unique horizontal form Ω with the property that, after applying the infinitesimal action $a : \mathfrak{g} \longrightarrow \mathcal{X}(P)$,*

$$a(K(h(X),h(Y))) = h([X,Y]) - [h(X),h(Y)]$$

for all $X,Y \in \mathcal{X}(M)$.

Proof. We have already shown that K is horizontal and then the first part is clear from the very definition of K. For the second part, by the comments before the proposition and the injectivity of a, the uniqueness follows. Hence we just have to show that the actual curvature does have this property. For that one just compute $K(h(X),h(Y))$ using the definition of K and the fact that $\omega \circ h = 0$ and we find:

$$K(h(X),h(Y)) = -\omega([h(X),h(Y)]) = \omega(h([X,Y]) - [h(X),h(Y)]).$$

On the other hand, since $h(X)$ is π-projectable to X, we have that $[h(X),h(Y)]$ is π-projectable to $[X,Y]$ and then $h([X,Y]) - [h(X),h(Y)]$ is π-projectable to $[X,Y] -$

$[X, Y] = 0$, i.e. it is vertical. Hence, at each point p, it is of type $a(v_p)$ with some $v_p \in \mathfrak{g}$. Plugging in the previous equation, we find that

$$K(h(X), h(Y))(p) = \omega(a(v_p)) = v_p,$$

and then the desired equation follows is just the defining property of v_p. $\square$

Exercise 80. As a continuation of Exercise 78: show that the curvature of the connection defined by θ^1 is a non-zero constant (that you have to compute) times the area form σ of S^2,

$$\sigma = xdy \wedge dz + ydz \wedge dx + zdx \wedge dy \in \Omega^2(S^2).$$

(be aware: the forms dx, etc make sense on the entire $\mathbb{R}^3$, but we only use their restrictions to S^3. And, while the forms dx, dy and dz are linearly independent on $\mathbb{R}^3$, their restrictions to S^2 do satisfy an extra-relation: $xdx+ydy+zdz = 0$ (why?)).

2.3.5. Principal bundle connections $\longleftrightarrow$ vector bundles ones

While we have seen that vector bundles of rank r are in 1-1 correspondence with principal GL_r-bundles, we now show that connections on vector bundles correspond (1-1) to connections on the associated principal bundles.

In one direction (from principal to vector) one can proceed, as before, slightly more generally, by starting with any Lie group G, a principal G-bundle $\pi : P \longrightarrow M$ and a representation $\rho : G \longrightarrow GL(V)$. The relevant vector bundle will be $E = E(P, V)$ previously discussed. We start with a connection on the principal G-bundle P and we would like to construct a connection ∇ on E; we will encode ∇ in the operator

$$d_\nabla : \Gamma(E) \longrightarrow \Omega^1(M, E).$$

Using Proposition 2.49 we see that it suffices to describe the

$$D_\nabla := h^\bullet d_\nabla (h^\bullet)^{-1} : C^\infty(P, V)^G \longrightarrow \Omega^1(P, V)_{\text{bas}}.$$

Proposition 2.59. *For any connection ω on P,*

$$D_\nabla : C^\infty(P, V)^G \longrightarrow \Omega^1(P, V)_{bas},\ D(f) = df + \rho(\omega(-))f$$

(where $\rho(\omega(-))f$ is the V-valued 1-form on P given by $X \mapsto \rho(\omega(X))(f)$) induces a connection ∇ on the vector bundle E.

Proof. Direct check. $\square$

Next, we consider the converse procedure: from vector bundle connections to principal ones. Start with a vector bundle E and a connection ∇ on E. We consider the associated frame bundle $\mathrm{Fr}(E)$, a principal GL_r-bundle, and we want to associated to ∇ a connection on $\mathrm{Fr}(E)$. We will do that by constructing a connection 1-form

$$\omega_\nabla \in \Omega^1(\mathrm{Fr}(E), gl_r).$$

We have to say what it does on an arbitrary vector

$$\left.\frac{d}{dt}\right|_{t=0} e(t) \in T_{e(0)}\mathrm{Fr}(E),$$

where

$$\mathbb{R} \ni t \mapsto e(t) = (e_1(t), \ldots, e_r(t)) \in \mathrm{Fr}(E)$$

is a curve in $\mathrm{Fr}(E)$. Let $\gamma : \mathbb{R} \longrightarrow M$ be the base curve of e. Consider the derivatives of the paths e_i with respect to ∇ (see 1.2.3) and decompose them with respect to the frame $e(t)$:

$$\frac{\nabla e_i}{dt}(t) = \sum_j a_i^j(t) e_j(t)$$

giving rise to the matrix

$$e(t)^{-1} \cdot \frac{\nabla e}{dt}(t) := (a_i^j(t))_{i,j} \in gl_r$$

(think that, while $e(t) : \mathbb{R}^r \xrightarrow{\sim} E_{\gamma(t)}$, $\frac{\nabla e}{dt}(t)$ is a linear map from $\mathbb{R}^r$ to $E_{\gamma(t)}$). Set now

$$\omega_\nabla(\frac{d}{dt}\Big|_{t=0} e(t)) := e(0)^{-1} \cdot \frac{\nabla e}{dt}(0) \in gl_r.$$

Theorem 2.60. *For any connection ∇ on E, ω_∇ is a connection 1-form on $Fr(E)$, and this defines a bijection between connections on E and connections on the principal GL_r-bundle $Fr(E)$.*

Proof. We first show that, for any ∇, ω_∇ is a connection 1-form. Let us first prove that $\omega(a_p(A)) = A$ for all $e \in \mathrm{Fr}(E)$, $A \in gl_r$. Recall that

$$a_p : gl_r \longrightarrow T_e\mathrm{Fr}(M),\ a_e(A) = \frac{d}{dt}\Big|_{t=0} e \cdot \exp(tA).$$

In the definition of ω above, we deal with the path $t \mapsto e \cdot \exp(tA)$ (which at $t = 0$ is e), hence

$$\omega(a_p(A)) = e^{-1}\frac{\nabla e \cdot exp(tA)}{dt}(0).$$

Using the definitions (of $\frac{\nabla}{dt}$, as well as the fact that this operation on frames is defined componentwise), the standard derivation rules apply and, since e is constant on t, we find

$$\omega(a_p(A)) = e^{-1}e \cdot \frac{dexp(tA)}{dt}(0) = A.$$

Next, we prove the invariance of ω. We have to show that for any path $e(t)$ of frames,

$$\omega((dR_g)(\frac{de}{dt}(0)) = Ad_{g^{-1}}\omega(\frac{de}{dt}(0)).$$

We start from the left hand side. We see that we deal with ω evaluated on the curve of frames $e(t)g$ hence, by the definition of ω, the expression is

$$(e(0)g)^{-1}\frac{\nabla eg}{dt}(0) = (g^{-1}e(0)^{-1})\frac{\nabla e}{dt}(0)g$$

i.e. precisely

$$g^{-1}\omega(\frac{de}{dt}(0))g = Ad_{g^{-1}}(\omega(\frac{de}{dt}(0)).$$

To show that $\nabla \mapsto \omega_\nabla$ is a bijection one either remarks that the previous argument can be read backwards, or checks directly that $\nabla \mapsto \omega_\nabla$ composed with the operation from the previous proposition is the identity. $\square$

Of course, it is not surprising that a similar story applies when looking at the notion of curvature. More precisely, one can consider:

- the curvature of the vector bundle connection ∇,

$$k_{\nabla} \in \Omega^2(M, \mathrm{End}(E)) \tag{2.61}$$

(see equation (1.27) in Proposition 1.25).

- the curvature of the principal bundle connection ω_{∇},

$$K_{\omega_{\nabla}} \in \Omega^2(\mathrm{Fr}(E), gl_r)_{\mathrm{bas}}$$

(see Lemma 2.57 for $P = \mathrm{Fr}(E)$, $G = gl_r$). Or, using the identification/isomorphism from Proposition 2.49,

$$K_{\omega_{\nabla}} \in \Omega^2(M, \mathcal{E}), \tag{2.62}$$

where $\mathcal{E}$ is obtained by the construction of Subsection 2.2.4. More precisely,

$$\mathcal{E} = E(\mathrm{Fr}(E), gl_r)$$

is the vector bundle over M obtained by attaching to the principal $G = GL_r$-bundle $P = \mathrm{Fr}(E)$ the fiber $V = gl_r$ endowed with the representation $GL_r \to GL(gl_r)$ given by the adjoint action.

Exercise 81. Show that $\mathcal{E} \cong \mathrm{End}(E)$ and, via this, the vector bundle curvature k_{∇} (see (2.61) corresponds to the principal bundle curvature $K_{\omega_{\nabla}}$ (see (2.62) .

2.3.6. Reduction of the structure group

We now return to the construction of the push-forward along group homomorphisms (Exercise 61) applied to an inclusion

$$i : H \to G$$

of a Lie subgroup of a Lie group G. Recall that the operation $P \mapsto i_*(P)$ associates to a principal H-bundle $\pi : P \to M$ the principal G-bundle

$$i_*(P) = (P \times G)/H,$$

the quotient modulo the action $(p, g) \cdot h = (ph, gh)$, endowed with the projection $[p, g] \mapsto \pi(p)$ and the G-action $[p, g] \cdot g' = [p, gg']$. With these, we say a principal G-bundle $\pi : Q \to M$ can be reduced to a Lie subgroup $H \subset G$ if there exists a principal H-bundle P (over M) such that Q is isomorphic to i_*P; since P is contained in $i_*(P)$ (with inclusion $p \mapsto [p, e]$), we can always arrange that $P \subset Q$. This leads to the following definition.

Definition 2.63. *Given a principal G-bundle $\pi : Q \to M$ and H a Lie subgroup of G, a reduction of Q to H is a choice of an H-invariant subspace $P \subset M$ with the property that $\pi|_P : P \to M$ is a principal H-bundle.*

Example 2.64. In general, saying that Q can be reduced to a subgroup H means, intuitively, that Q is simpler (and the smaller H is, the simpler Q is). For instance, in the extreme case when $H = \{e\}$ is the trivial group, choosing a reduction of Q to e is the same thing as choosing a section of Q or, equivalently, a trivialization of Q.

Proposition 2.65. *Given the principal G-bundle $\pi : Q \to M$ and $H \subset G$ a Lie subgroup, there is a 1-1 correspondence between reductions of Q to H and sections of the map $Q/H \to M$ induced by π.*

Proof. To come. $\square$

Remark 2.66. Although it is rather simple, the last part of the lemma is quite important conceptually. One reason is that it allows us to interpret reductions as (smooth) sections. This is important since functions are easier to handle. Another use of the previous corollary comes from general properties a locally trivial fiber bundles. By a locally trivial fiber bundle with fiber N (N is a manifold) we mean a submersion $\pi : Q \longrightarrow M$ with the property that any point $x \in M$ has a neighborhood U with the property that $\pi^{-1}(U)$ is diffeomorphic to $U \times N$, by a diffeomorphism that sends the fibers $\pi^{-1}(x)$ to the fibers $\{x\} \times N$. A basic but non-trivial property of locally trivial fiber bundles is that, if the fiber N is contractible, then it automatically admits global sections. Of course, this can be applied to the projection $\tilde{P}/H \longrightarrow M$ to deduce:

Corollary 2.67. *If the quotient G/H is contractible, then any principal G-bundle admits a reduction to H.*

Example 2.68. Anticipating a bit, let us mention that when $Q = \mathrm{Fr}(E)$ is the frame bundle of a rank r real vector bundle, $G = GL_r(\mathbb{R})$, then the choice of a reduction of $\mathrm{Fr}(E)$ to the orthogonal group $O(r) \subset GL_r(\mathbb{R})$ is equivalent to the choice of a metric on the vector bundle E. The existence of metrics can then be seen as a consequence of the previous corollary since $GL_r(\mathbb{R})/O(r)$ is contractible.

CHAPTER 3

G-structures by examples

3.1. Geometric structures on vector spaces

3.1.1. Linear G-structures

Many types of geometric structures on manifolds M are of infinitesimal type: they are collections of structures on the tangent spaces T_xM, "varying smoothly" with respect to $x \in M$ (think e.g. about Riemannian metrics on M). Hence, the first step is to understand (the meaning of) such structures on general finite dimensional vector spaces V.

The basic idea is to characterize the geometric structures of interest via "the frames adapted to the structure": for many types of geometric structures on vector spaces V, one can make sense of frames adapted to the given structure; for instance, for an inner-product on V, one looks at frames which are orthonormal with respect to the given inner product; or, for an orientation on V, one looks at oriented frames; etc. This process of defining the notion of frames adapted to a geometric structure depends of course on the type of structures one considers. But, for most of them, this process can be divided into several steps which we now describe. Start with a class $\mathcal{S}truct$ of geometric structures- for instance:

$$\mathcal{S}truct \in \{\text{inner products}, \text{orientations}, \text{complex structures}, \text{symplectic structures}, \text{etc}\}$$

(to be discuss in more detail later on). Hence, for any finite dimensional vector space V one has the set $\mathcal{S}truct(V)$ of $\mathcal{S}truct$-structures on V and one is interested in pairs

$$(V, \sigma)$$

consisting of a vector space V and an element $\sigma \in \mathcal{S}truct(V)$. The steps we mentioned are:

S1: One has a notion of isomorphism of such pairs (V, σ). In particular, we can talk about the automorphism group

$$gl(V, \sigma) \subset GL(V).$$

S2: One has a a standard model, which is $\mathbb{R}^n$ with "a canonical" $\mathcal{S}truct$-structure

$$\sigma_{\text{can}} \in \mathcal{S}truct(\mathbb{R}^n).$$

It is a model in the sense that any n-dimensional vector space V endowed with $\sigma \in \mathcal{S}truct(V)$ is isomorphic to $(\mathbb{R}^n, \sigma_{\text{can}})$ (just that the isomorphism is not canonical!).

S3: Of capital importance is the group of automorphisms of the standard model:

$$G := gl(\mathbb{R}^n, \sigma_{\text{can}}) \subset GL_n.$$

S4: While we identify a frame $\phi = (\phi_1, \ldots, \phi_n)$ of V with the associated linear isomorphism $\phi : \mathbb{R}^n \to V$, the condition that ϕ is adapted to a given structure $\sigma \in \mathcal{S}truct(V)$ means that ϕ is an isomorphism between $(\mathbb{R}^n, \sigma_{\text{can}})$ and (V, σ). This gives rise to the space of adapted frames

$$\mathrm{Fr(V, \sigma) \subset Fr(V)}$$

and the main idea is that σ is encoded in $\mathrm{Fr(V, \sigma)}$.

To point out the main properties of $\mathrm{Fr(V, \sigma)}$ we will use the right action of GL_n on Fr(V) (page 65) and, for two frames ϕ and ϕ' of V, we denote by

$$[\phi : \phi'] \in GL_n$$

the resulting matrix of coordinate changes- defined by $\phi = \phi' \cdot [\phi : \phi']$. With these:

- $\phi \in \mathrm{Fr(V, \sigma), A \in G} \Longrightarrow \phi \cdot \mathrm{A \in Fr(V, \sigma)}$.
- $\phi, \phi' \in \mathrm{Fr(V, \sigma)} \Longrightarrow [\phi : \phi'] \in \mathrm{G}$.

Equivalently, $\mathrm{Fr(V, \sigma)}$ is a G-invariant subspace of Fr(V) on which G acts transitively.

Note that the previous discussion can be carried out without much trouble in all the examples we will be looking at. However, the general discussion is not very precise- as it appeals to our imagination for making sense of some of the ingredients. Of course, it can be made precise by axiomatizing the notion of "class $\mathcal{S}truct$ of geometric structures on vector spaces" (e.g. using the language of functors). However, the philosophy that geometric structures are characterized by their adapted frames leads to a different approach in which the class of geometric structures $\mathcal{S}truct$ is associated to the group G and a G-structure on V is encoded in the frame bundle:

Definition 3.1. *Let G be a subgroup of $GL_n(\mathbb{R})$. A linear G-structure on an n-dimensional vector space V is a subset*

$$\mathcal{S} \subset Fr(V)$$

satisfying the axioms:

A1: *$\mathcal{S}$ is G-invariant, i.e.: $\phi \in \mathcal{S},\ A \in G \Longrightarrow \phi \cdot A \in \mathcal{S}$.*
A2: *if $\phi, \phi' \in \mathcal{S}$ then $[\phi : \phi'] \in G$.*

We denote by $\mathcal{S}truct_G(V)$ the space of linear G-structures on V.

Remark 3.2. The second axiom implies that all the frames $\phi \in \mathcal{S}$ induces the same element in the quotient,

$$\sigma_{\mathcal{S}} \in \mathrm{Fr(V)/G}.$$

Conversely, starting from $\sigma_{\mathcal{S}}$ (which can be any element in this quotient), $\mathcal{S}$ can be recovered as the set of all frames $\phi \in$ Fr that represent $\sigma_{\mathcal{S}}$. In other words,

$$\mathcal{S}truct_G(V) = \mathrm{Fr(V)/G}.$$

Remark 3.3 (isomorphisms; the standard model)**.** One can now check that the previous discussion (the steps S1-S4) can be carried out for the abstractly defined notion of linear G-structures- i.e. the $\mathcal{S}truct_G$ of the previous definition. To do that, note that one can talk about:

1. isomorphisms between two linear G-structures: given $\mathcal{S}_i$ on V_i, $i \in \{1, 2\}$, an isomorphism between the pairs $(V_i, \mathcal{S}_i)$ is any linear isomorphism

$$A : V_1 \to V_2$$

with the property that

$$(\phi_1, \ldots, \phi_n) \in \mathcal{S}_1 \Longrightarrow (A(\phi_1), \ldots, A(\phi_n)) \in \mathcal{S}_2.$$

2. the standard linear G-structure on $\mathbb{R}^n$

$$\mathcal{S}_G^{\text{can}} \subset \mathrm{Fr}(\mathbb{R}^{\mathrm{n}})$$

consisting of frames which, interpreted as matrices , are elements of G.

(We use here the conventions from pages 64-65. Hence the way that we interpret a frame $\phi = (\phi_1, \ldots, \phi_n)$ of $\mathbb{R}^n$ as a matrix is by interpreting each vector ϕ_i as a column vector:

$$[\phi_i] = \begin{pmatrix} \phi_i^1 \\ \ldots \\ \phi_i^n \end{pmatrix}$$

so that ϕ is identified with

$$[\phi] = ([\phi_1], \ldots, [\phi_n]) = \begin{pmatrix} \phi_1^1 & \ldots & \phi_n^1 \\ \ldots & \ldots & \ldots \\ \phi_1^n & \ldots & \phi_n^n \end{pmatrix}$$

In this way the right action of GL_n on $\mathrm{Fr}(\mathbb{R}^{\mathrm{n}})$ will be identified with the usual multiplication of matrices.)

With these, one can carry out steps S1-S4 for $\mathcal{S}truct_G$ and one can show that, indeed, for any $\mathcal{S} \in \mathcal{S}truct_G(V)$ (playing the role of σ in the general discussion at the start), the resulting space of adapted frames $\mathrm{Fr}(\mathrm{V}, \mathcal{S})$ is precisely $\mathcal{S}$.

3.1.2. Example: Inner products

Recall that a linear metric (inner product) on the vector space V is a symmetric bilinear map

$$g : V \times V \longrightarrow \mathbb{R}$$

with the property that $g(v, v) \geq 0$ for all $v \in V$, with equality only for $v = 0$. There is an obvious notion of frames of V adapted to g: the orthonormal ones, i.e. the ones of type $\phi = (\phi_1, \ldots, \phi_n)$ with

$$g(\phi_i, \phi_j) = \delta_{i,j} \quad \forall\ i, j. \tag{3.4}$$

However, it is instructive to find them going through stepst S1-S4 mentioned at the beginning of this chapter.

S1: Given two vector spaces endowed with inner products, (V, g), (V', g'), it is clear what an isomorphism between them should be: any linear isomorphism $A : V \longrightarrow V'$ with the property that

$$g'(A(u), A(v)) = g(u, v) \quad \forall\ u, v\ \in V.$$

One also says that A is an isometry between (V, g) and (V', g').

S2: The standard model is $\mathbb{R}^n$ with the standard inner product g_{can}:

$$g_{\text{can}}(u, v) = \langle u, v\rangle = \sum_i u^i \cdot v^i,$$

or, in the matrix notation:

$$(3.5) \qquad g_{\mathrm{can}}(u,v) = (u^1 \dots u^n) \cdot \begin{pmatrix} v^1 \\ \dots \\ v^n \end{pmatrix} = u \cdot [v].$$

(recall that $[v] = v^T$ stands for v interpreted as a column matrix- see page 65). We leave it as an exercise to show that, indeed, any vector space endowed with an inner product, (V, g), is isomorphic to $(\mathbb{R}^n, g_{\mathrm{can}})$.

S3: The symmetry group of the standard model becomes

$$G = \{A \in GL_n(\mathbb{R}) : g_{\mathrm{can}}(\hat{A}(u), \hat{A}(v)) = g_{\mathrm{can}}(u, v) \text{ for all } u, v \in \mathbb{R}^n\} \subset GL_n(\mathbb{R}).$$

Here $\hat{A}$ denotes the matrix A interpreted as a linear map (see page 64). Hence, while as column matrices one has $[\hat{A}(u)] = A \cdot [v]$ (see page 65), as a row matrix $\hat{A}(u)$ is $v \cdot A^T$. We find that

$$G = \{A \in GL_n(\mathbb{R}) : A^T \cdot A = I\} = O(n).$$

S4: By the general procedure, "the adapted frames" of (V, g) are those frames $\phi \in \mathrm{Fr}(V)$ with the property that the induced linear isomorphism $\hat{\phi} : \mathbb{R}^n \longrightarrow V$ is an isomorphism between $(\mathbb{R}^n, g_{\mathrm{can}})$ and (V, g). It is now easy to see that this happens if and only if ϕ is orthonormal with respect to g. Hence one ends up with the set of orthonormal frames of (V, g), denoted

$$\mathcal{S}_g \subset \mathrm{Fr}(V).$$

The following Proposition makes precise the fact that (and explains how) the "inner product g" is encoded by the associated set of frames $\mathcal{S}_g$.

Proposition 3.6. *Given an n-dimensional vector space V,*

$$g \longleftrightarrow \mathcal{S}_g$$

defines a 1-1 correspondence between:

1. *inner products on V.*
2. *linear $O(n)$-structures on V.*

Proof. The main point is that $\mathcal{S}_g$ determines g uniquely: using any $\phi \in \mathcal{S}_g$, and decomposing arbitrary elements of V with respect to ϕ as $u = \sum_i u^i \cdot \phi_i$, it follows that

$$g(u, v) = \sum_i u^i \cdot v^i.$$

This shows how to reconstruct g out of any linear $O(n)$-structure $\mathcal{S}$; the axioms for $\mathcal{S}$ imply the resulting g does not depend on the choice of ϕ and that $\mathcal{S}_g = \mathcal{S}$. □

3.1.3. Example: Orientations

The next example of "structure" that we consider is that of "orientation". This is very well suited to our discussion because the notion of orientation is itself defined (right from the start) using frames. More precisely, given our n-dimensional vector space V, one says that two frames $\phi, \phi' \in \mathrm{Fr}(V)$ induce the same orientation if the matrix $[\phi : \phi']$ of change of coordinates has positive determinant; this defines an equivalence relation $\sim$ on $\mathrm{Fr}(V)$ and an orientation of V is the choice of an equivalence class:

$$\mathcal{O} \in \mathrm{Orient}(V) := \mathrm{Fr}(V)/\sim .$$

Exercise 82. **S1:** Given two oriented vector spaces $(V, \mathcal{O})$, $(V', \mathcal{O}')$, what are the isomorphisms between them?

S2, S3: The local model is $\mathbb{R}^n$ with the orientation $\mathcal{O}_{\mathrm{can}}$ induced by the standard basis (frame) and computing the induced symmetry group (isomorphisms from the local model to itself) we find the subgroup of GL_n of matrices with positive determinant:

$$GL_n^+ = \{A \in GL_n : \det(A) > 0\}.$$

S4: The special frames of $(V, \mathcal{O})$ are, of course, the frames which induce the given orientation; they correspond to oriented isomorphism $(\mathbb{R}^n, \mathcal{O}_{\mathrm{can}}) \longrightarrow (V, \mathcal{O})$. Denote the set of such frames by

$$\mathcal{S}_{\mathcal{O}} \subset \mathrm{Fr}(V).$$

As for metrics, it follows that:

Proposition 3.7. *Given an n-dimensional vector space V,*

$$\mathcal{O} \longleftrightarrow \mathcal{S}_{\mathcal{O}}$$

defines a 1-1 correspondence between:

1. *orientations on V.*
2. *linear GL_n^+-structures on V.*

Note also that interpretation from Remark 3.2 applied to this context (giving an interpretation of orientations on V as elements of $\mathrm{Fr}(V)/GL_n^+$) holds by the very definition of orientations, since the equivalence relation $\sim$ discussed above is precisely the one induced by the action of GL_n^+ on $\mathrm{Fr}(V)$.

3.1.4. Example: Volume elements

Recall that a volume element on V is a non-zero element $\mu \in \Lambda^n V^*$ (where n is the dimension of V) or, equivalently, a non-zero skew-symmetric multilinear map

$$\mu : \underbrace{V \times \ldots \times V}_{n\,\mathrm{times}} \longrightarrow \mathbb{R}.$$

S1: An isomorphism between (V, μ) and $V', \mu')$ is any linear isomorphism $A : V \longrightarrow V'$ with the property that

$$\mu'(A(v_1), \ldots, A(v_n)) = \mu(v_1, \ldots, v_n)$$

for all v_i's. Equivalently, a linear map A induces $A^* : V'^* \longrightarrow V^*$ and then $A^* : \Lambda^n(V')^* \longrightarrow \Lambda^n V^*$, and we are talking about the condition $A^*\mu' = \mu$.

S2: The standard model is, again, $\mathbb{R}^n$ with the canonical volume element given μ_{can} given by (or uniquely determined by):

$$\mu_{\mathrm{can}}(e_1, \ldots, e_n) = 1.$$

S3: The associated symmetry group becomes (after the computation)

$$SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) : \det(A) = 1\}.$$

S4: Given (V, μ), its special frames will be those $\phi \in \mathrm{Fr}(V)$ satisfying:

$$\mu(\phi_1, \ldots, \phi_n) = 1.$$

Denoting by $\mathcal{S}_\mu$ the set of adapted frames, one obtains:

Proposition 3.8. *Given an n-dimensional vector space V,*

$$\mu \longleftrightarrow \mathcal{S}_\mu$$

defines a 1-1 correspondence between:

1. *volume elements on V.*
2. *linear $SL_n(\mathbb{R})$-structures on V.*

Remark 3.9. Note that a volume element on V determines, in particular, an orientation $\mathcal{O}_\mu$ on V: the one induced by frames ϕ with the property that $\mu(\phi) > 0$. This is basically due to the inclusion $SL_n(\mathbb{R}) \subset GL_n^+$ (try to make this more precise and to generalize it).

Geometrically, a volume element encodes an orientation on V together with the notion of oriented volume of any "convex body" (simplex) spanned by any n-vectors $v_1, \ldots, v_n$.

3.1.5. Example: p-directions

Another interesting "structure" one can have on a vector space V (and becomes more interesting when passing to manifolds) is that of "p-directions" in V, where p is any positive integer less or equal to the dimension n of V. By that we simply mean a p-dimensional vector subspace $W \subset V$.

The steps S1-S4 are quite clear: Given (V, W) and (V', W'), and isomorphism between them is any linear isomorphism $A : V \longrightarrow V'$ satisfying $A(W) \subset W'$. The local model is $(\mathbb{R}^n, \mathbb{R}^p)$, where we view $\mathbb{R}^p$ sitting inside $\mathbb{R}^n$ via the inclusion on the first components

$$\mathbb{R}^p \hookrightarrow \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p}, \quad u \mapsto (u, 0).$$

For the symmetry group we find

$$G = GL(p, n-p),$$

i.e. the set of invertible matrices of type

$$X = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \in GL_n(\mathbb{R})$$

with $A \in GL_p(\mathbb{R})$, $C \in GL_{n-p}(\mathbb{R})$ and B is an $p \times (n-p)$ matrix. The special frames of (V, W) will be those of type

$$\phi = (\phi_1, \ldots, \phi_n) \ \text{ with } \phi_1, \ldots, \phi_p \in W$$

and they define

$$\mathcal{S}_W \subset \mathrm{Fr}(V).$$

Proposition 3.10. *Given an n-dimensional vector space V and $p \leq n$,*

$$W \longleftrightarrow \mathcal{S}_W$$

defines a 1-1 correspondence between:

1. *p-dimensional subspaces of V.*
2. *linear $GL(p, n-p)$-structures on V.*

3.1.6. Example: Integral affine structures

By an (linear) integral affine structure on a vector space V we mean a lattice

$$\Lambda \subset V.$$

That means that Λ is a discrete subgroup of $(V, +)$. Equivalently, Λ is of type

$$\Lambda = \mathbb{Z}\phi_1 + \ldots + \mathbb{Z}\phi_n$$

for some frame $\phi = (\phi_1, \ldots, \phi_n)$. Frames of this type are called integral frames of (V, Λ) and they play the role of adapted frames; denote by $\mathcal{S}_\Lambda$ the set of such frames. The local model is $(\mathbb{R}^n, \mathbb{Z}^n)$, while the symmetry group becomes

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

One obtains:

Proposition 3.11. *Given an n-dimensional vector space V and $p \leq n$,*

$$W \longleftrightarrow \mathcal{S}_\Lambda$$

defines a 1-1 correspondence between:

1. *integral affine structures on the vector space V.*
2. *linear $GL_n(\mathbb{Z})$-structures on V.*

Remark 3.12. Note also that an integral affine structure Λ on an oriented vector space $(V, \mathcal{O})$ induces a volume element μ_Λ on V

$$\mu_\Lambda = \phi^1 \wedge \ldots \phi^n$$

where ϕ is a(ny) positively oriented integral basis (and we use the dual basis in the last equation). This comes from the fact that any matrix $A \in GL_n(\mathbb{Z})$ has determinant ± 1.

3.1.7. Example: Complex structures

A (linear) complex structure on a vector space V is a linear map

$$J : V \longrightarrow V$$

satisfying

$$J^2 = -\mathrm{Id}.$$

Such a complex structure allows us to promote the (real) vector space V to a complex vector space by defining

$$(r + is) \cdot v := rv + sJ(v).$$

When we view the vector space V as a complex vector space in this way, we will denote it by V_J. Note that this interpretation implies that complex structures can exist only on even dimensional vector spaces.

<u>S1</u>: An isomorphism between (V, J) and (V', J') is any linear isomorphism $A : V \longrightarrow V'$ satisfying

$$J' \circ A = A \circ J.$$

Equivalently, A is an isomorphism between the complex vector spaces V_J and $V'_{J'}$.

<u>S2</u>: The standard model is $\mathbb{R}^{2k} = \mathbb{R}^k \oplus \mathbb{R}^k$ with the complex structure

$$J_{\mathrm{can}}(u, v) = (-v, u)$$

(think $i \cdot (u + iv) = -v + iu$). The fact that any (V, J) is isomorphic to the standard model follows again using V_J, which must be isomorphic to $\mathbb{C}^n$ as a complex vector space. It is sometimes useful to use the matrix notations. Using the previous decomposition of $\mathbb{R}^{2k}$, we see that a linear automorphism of $\mathbb{R}^{2k}$ can be represented as a matrix

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where such a matrix encodes the map

$$\mathbb{R}^{2k} \ni (u, v) \mapsto (Au + Bv, Cu + Dv) \in \mathbb{R}^{2k}.$$

With this, one has:

$$J_{\mathrm{can}} = \begin{pmatrix} 0 & -I_k \\ I_k & 0 \end{pmatrix}. \tag{3.13}$$

S3: We find that the associated symmetry group consists of those $M \in GL_{2k}(\mathbb{R})$ satisfying $J_{\mathrm{can}}M = MJ_{\mathrm{can}}$. In the matrix notation, working out this condition, we find out that we are looking at Ms of type

$$M = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}.$$

We find that the symmetry group is a copy of $GL_k(\mathbb{C})$ embedded in $GL_{2k}(\mathbb{R})$ via

$$GL_k(\mathbb{C}) \ni A + iB \mapsto \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in GL_{2k}(\mathbb{R}).$$

S4: One can now work out the resulting notion of adapted frame for (V, J) and we arrive at frames of type

$$\phi = (\phi_1, \ldots, \phi_k, J(\phi_1), \ldots, J(\phi_k)).$$

with $k = \frac{1}{2}\mathrm{dim}(V)$, $\phi_i \in V$. These define a subspace

$$\mathcal{S}_J \subset \mathrm{Fr}(V).$$

Proposition 3.14. *Given an $n = 2k$-dimensional vector space V,*

$$J \longleftrightarrow \mathcal{S}_J$$

defines a 1-1 correspondence between:

1. *linear complex structures on the vector space V.*
2. *linear $GL_k(\mathbb{C})$-structures on V, where $GL_k(\mathbb{C})$ sits inside $GL_n = GL_n(\mathbb{R})$ as explained above.*

Remark 3.15. A complex structure J on V induces an orientation $\mathcal{O}_J$ on V: one just considers the orientation induced by the complex frames. The fact that the orientation does not depend on the frame (which is the main point of this remark) comes from the fact that, via the previous inclusion of $GL_k(\mathbb{C})$, one ends up in GL^+_{2k}.

Exercise 83. Prove the last statement. More precisely, denoting by incl the inclusion of $GL_k(\mathbb{C})$ into $GL_{2k}(\mathbb{R})$, show that

$$det(incl(Z)) = |det(Z)|^2 \quad \forall\, Z \in GL_k(\mathbb{C}).$$

(Hint: there exists a matrix $X \in GL_k(\mathbb{R})$ such that

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix} = \begin{pmatrix} X & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & X^{-1} \\ -I & X \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix})$$

For the curious reader. There is another way of encoding complex structures: via (sub)spaces instead of maps. This is especially useful when passing to manifolds. This is based on some yoga with complex vector spaces. Start with a real vector space V (no J yet). Then we introduce the complexification of V as the complex vector space

$$V_{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C} = \{u + iv : u, v \in V\}$$

with the obvious complex structure. Note that it also comes with an obvious conjugation map

$$u + iv \mapsto \overline{u + iv} := u - iv.$$

Now, if J is a complex structure on V, just view it as an $\mathbb{R}$-linear map on V and extend it to a $\mathbb{C}$-linear map on $V_{\mathbb{C}}$:

$$J(u + iv) = J(u) + iJ(v).$$

Using now that $J^2 = -\mathrm{Id}$ and that J acts now on a complex vector space (namely on $V_{\mathbb{C}}$), $V_{\mathbb{C}}$ can be split into the i and $-i$ eigenspaces of J:

$$V = V^{1,0} \oplus V^{0,1},$$

where the two summands are the complex vector spaces given by:

$$V^{1,0} = \{w \in V_{\mathbb{C}} : J(w) = iw\} = \ldots = \{u - iJ(u) : u \in V\},$$

$$V^{0,1} = \{w \in V_{\mathbb{C}} : J(w) = -iw\} = \ldots = \{u + iJ(u) : u \in V\}.$$

Of course, the two are related by conjugation:

$$V^{0,1} = \overline{V^{1,0}}.$$

The key remark here is that the complex structure J on V not only gives rise to, but it is actually determined by the resulting complex subspace of $V_{\mathbb{C}}$

$$V^{1,0} \subset V_{\mathbb{C}}.$$

The key properties of this subspace is that

$$V_{\mathbb{C}} = V^{1,0} \oplus \overline{V^{1,0}}.$$

and that the first projection $V^{1,0} \longrightarrow V$ is an isomorphism. Note that these properties are, indeed, formulated independently of J and they do determine J.

Note also the connection between the complex vector spaces that arise from the two points of view. On one hand, J gave rise to the complex vector space V_J (V viewed as a complex vector space using J). On the other hand, we had the complex subspace $V^{1,0}$ of $V_{\mathbb{C}}$. Note that the two are isomorphic as complex vector spaces, by the obvious map

$$V_J \xrightarrow{\sim} V^{1,0}, \quad u \mapsto u - iJ(u).$$

Indeed, denoting by I this map, one has:

$$i \cdot I(u) = iu + J(u) = J(u) - iJ(J(u)) = I(J(u)).$$

Similarly, the obvious map $V_J \longrightarrow V^{0,1}$ is a conjugate-linear isomorphism. One often identifies $V^{1,0}$ with V_J using I and then $V^{0,1}$ with the conjugate of V_J.

3.1.8. Example: Symplectic forms

A linear symplectic form on a vector space V is a non-degenerate antisymmetric 2-forms

$$\omega : V \times V \longrightarrow \mathbb{R}.$$

One can think of linear symplectic forms as some antisymmetric versions of inner products. Note however that, in contrast with inner products, they can only exist on even dimensional vector spaces (this will be proven below).

S1: An isomorphism between (V, ω) and (V', ω') is any linear isomorphism $A : V \longrightarrow V'$ satisfying

$$\omega'(A(u), A(v)) = \omega(u, v) \quad \forall\, u, v \in V.$$

S2: The local model is not so obvious before one thinks a bit about symplectic structures. Postponing a bit the "thinking", let us just describe the resulting local model from several points of view. First of all, it is $(\mathbb{R}^{2k}, \omega_{\mathrm{can}})$ with

$$\omega_{\mathrm{can}}((x, y), (x', y')) = \langle x', y\rangle - \langle x, y'\rangle, \tag{3.16}$$

where:

- we use $\mathbb{R}^{2k} = \mathbb{R}^k \times \mathbb{R}^k$ to represent the elements of $\mathbb{R}^{2k}$ as pairs
$$(x, y) = (x_1, \ldots, x_k, y_1, \ldots, y_k).$$
- $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^k$.

More compactly, writing the standard frame of $\mathbb{R}^{2k}$ as

$$(e_1, \ldots, e_k, f_1, \ldots, f_k),$$

and using the associated dual frame and the wedge-product operations,

$$\omega_{\text{can}} = f^1 \wedge e^1 + \ldots + f^k \wedge e^k.$$

Even more compactly, one can use the canonical inner product g_{can} and the canonical complex structure J_{can} on $\mathbb{R}^{2k}$ (see the previous subsections) to write

$$\omega_{\text{can}}(u, v) = g_{\text{can}}(u, J_{\text{can}}(v)), \quad \forall\ u, v \in \mathbb{R}^{2k}.$$

This is a fundamental relation between metrics, complex structure and symplectic structures that will be further discussed in the next section. One can go further and use the matrix expressions (3.5) and (3.13) and write:

$$\omega_{can}(u, v) = u\, J_{\text{can}}\,[v].$$

<u>S3</u>: To describe the symmetry group of $(\mathbb{R}^{2k}, \omega_{\text{can}})$ one can proceed in various ways, depending on which of the previous formulas for ω_{can} one uses. The most elegant way is to use the last formula; hence we are looking at those matrices $M \in GL_{2k}(\mathbb{R})$ for which $u \cdot J_{\text{can}} \cdot [v] = \omega_{\text{can}}(u, v)$ equals to

$$\omega_{\text{can}}(\hat{M}(u), \hat{M}(v)) = \hat{M}(u) \cdot J_{\text{can}} \cdot [\hat{M}(v)] = u \cdot M^T \cdot J_{\text{can}} \cdot M \cdot [v]$$

for all $u, v \in \mathbb{R}^{2k}$. Hence the symmetry group of interest is the symplectic group

$$Sp_k(\mathbb{R}) := \{M \in GL_{2k}(\mathbb{R}) : M^T J_{can} M = J_{can}\}$$

Note that, writing

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

the previous equations for M become

$$A^T C = C^T A, \quad D^T B = B^T D, \quad D^T A - B^T C = I_k,$$

equations at which one can arrive directly using the formula (3.16).

<u>S3</u>: The "special frames" for a pair (V, ω) (called also the symplectic frames) are somehow similar to the orthonormal frames for metrics. However, unlike there, it is not so clear "right away" how to define the notion of "symplectic frame". Of course, since we have already described the local model, we could just say that $\phi \in \mathrm{Fr}(V)$ is a symplectic frame for (V, ω) if the induced linear isomorphism $\hat{\phi}$ is an isomorphism between $(\mathbb{R}^{2k}, \omega_{\text{can}})$ and (V, ω). Explicitly, that means frames of type

$$(\phi_1, \ldots, \phi_k, \psi_1, \ldots, \psi_k) \quad \text{(in particular } n = 2k)$$

with the property that

$$\omega(\phi_i, \psi_i) = -1,\ \omega(\psi_i, \phi_i) = 1,\ \omega(\phi_i, \psi_j) = 0 \text{ for } i \neq j,\ \omega(\phi_i, \phi_j) = 0 = \omega(\psi_i, \psi_j).$$

Ok, we take this as the definition of symplectic frames, but one should keep in mind that there is still some cheating involved since, on the first place, the local model was not that obvious. Denoting by $\mathcal{S}_\omega$ the space of symplectic frames, we obtain exactly as for inner products:

Proposition 3.17. *Given an $n = 2k$-dimensional vector space V,*

$$\omega \longleftrightarrow \mathcal{S}_{\omega}$$

defines a 1-1 correspondence between:

1. *linear symplectic structures on the vector space V.*
2. *linear $Sp_k(\mathbb{R})$-structures on V.*

Remark 3.18 (from symplectic forms to volume elements)**.** Any (linear) symplectic form ω on V induces (canonical- i.e. without any further choices) a volume element

$$\mu_{\omega} := \omega^k = \underbrace{\omega \wedge \ldots \wedge \omega}_{k\text{times}} \in \Lambda^n V^*.$$

This is related to the fact that

$$Sp_k(\mathbb{R}) \subset SL_{2k}(\mathbb{R}).$$

For the curious reader. Here is more insight into the notion of symplectic vector spaces which should clarify the previous discussion/choices. So, we fix our (V, ω). The key notion will be that of Lagrangian subspace. First of all, one calls a subspace $L \subset V$ isotropic if ω vanishes on it:

$$\omega(u, v) = 0 \quad \forall\ u, v \in L.$$

Those which are maximal (of highest possible dimension) among all isotropic subspaces are called Lagrangian.

Note that for any isotropic L one has a well-defined map

$$\Phi_L : L \longrightarrow (V/L)^*, \quad \Phi_L(u)(v) = \omega(u, v)$$

(where one thinks of $(V/L)^*$ as consisting of linear maps from V to $\mathbb{R}$ that vanish on L).

Lemma 3.19. *An isotropic subspace $L \subset V$ is Lagrangian if and only if Φ_L is an isomorphism.*

Note that this implies that all Lagrangian subspaces have the same dimension, namely

$$\dim(L) = \frac{1}{2}\dim(V).$$

(in particular, the dimension of V is even).

Proof. Note that the map Φ_L is automatically injective, because ω is non-degenerate. In particular, the dimension of the domain is less than of the codomain, hence

$$\dim(L) \leq \frac{1}{2}\dim(V)$$

(... for all isotropic L). If Φ_L is an isomorphism then this becomes an equality hence L is clearly maximal among the isotropic subspaces. Assume now that L is Lagrangian. To check that Φ_L is also surjective, let $\xi \in (V/L)^*$, i.e. $\xi : V \longrightarrow \mathbb{R}$ linear satisfying $\xi|_L = 0$. Using again the nondegeneracy of ω,

$$V \longrightarrow V^*, u \mapsto \omega(u, \cdot) \tag{3.20}$$

is an isomorphism, hence we find $u_0 \in V$ such that $\xi(v) = \omega(u_0, v)$ for all $v \in V$. We have to show that $u_0 \in L$. Assume this is not the case. But then

$$L' := L + \mathbb{R}u_0$$

will be larger than L and still isotropic: indeed, since $\xi|_L = 0$, we also have $\omega(u_0, u) = 0$ for all $u \in L$. Hence the desired contradiction. ☐

Proposition 3.21. *Given (V, ω) and $L \subset V$ Lagrangian:*

1. *L admits a Lagrangian complement, i.e. $L' \subset V$ Lagrangian with*
$$V = L \oplus L'.$$
2. *for any Lagrangian complement L', one has a linear isomorphism*
$$\Phi : L' \longrightarrow L^*, \quad \Phi(l)(l') = \omega(l, l').$$
3. *The resulting linear isomorphism $V \cong L \oplus L^*$ is an isomorphism between (V, ω) and $(L \oplus L^*, \omega_L)$, where*
$$\omega_L((l, \xi), (l', \xi')) = \xi(l') - \xi'(l)).$$
(for $(l, \xi), (l', \xi') \in V_L$)

Note that the very last part tells us that (V, ω) is determined, up to isomorphism, by the bare vector space L (indeed, no extra-structure is needed to define ω_L). This provides "the local model". It is now easy to see that, for $L = \mathbb{R}^k$ one recovers, after the identifications

$$\mathbb{R}^k \oplus (\mathbb{R}^k)^* \cong \mathbb{R}^k \oplus \mathbb{R}^k \cong \mathbb{R}^{2k},$$

the local model $(\mathbb{R}^{2k}, \omega_{\text{can}})$.

Proof. Start with any complement C of L in V (Lagrangian or not), with corresponding projection $p_C : V \longrightarrow C$. Then, for any $c \in C$, the map

$$v \mapsto \frac{1}{2}\omega(c, \mathrm{pr}_C(v))$$

vanishes on L (because pr_C does) hence, using that Φ_L is an isomorphism, this map is of type $\omega(l_c, \cdot)$ for some $l_c \in L$. In particular,

$$\omega(l_c, c') = \frac{1}{2}\omega(c, c') \quad \forall\, c' \in C.$$

Replace now C by

$$L' := \{c - l_c : c \in C\}.$$

We claim that L' is the Lagrangian complement we were looking for. Indeed, using the last equation and the fact that $l_c \in L$, which is Lagrangian, we get for any two elements $c - l_c, c' - l_{c'} \in L'$

$$\begin{aligned} \omega(c - l_c, c' - l_{c'}) &= \omega(c, c') - \omega(l_c, c') + \omega(l_{c'}, c) + \omega(l_c, l_{c'}) \\ &= \omega(c, c') - \frac{1}{2}\omega(c, c') + \frac{1}{2}\omega(c', c) \\ &= 0 \end{aligned}$$

Moreover, $V = L \oplus C$ implies $V = L \oplus L'$. This proves the first part.

Assume now that L' is an arbitrary Lagrangian complement. Note that Φ is precisely $\Phi_{L'}$ after the identification $V/L' \cong L$. We deal with the resulting sequence of isomorphisms:

$$V = L \oplus L' \cong L \oplus (V/L')^* \cong L \oplus L^*.$$

For $v \in V$, denote by (l_v, ξ_v) the resulting element in $L \oplus L^*$. Going through the maps involved, we find the characterization:

$$\xi_v \circ \mathrm{pr}_L(\cdot) = \omega(v - l_v, \cdot).$$

For $v, w \in L$, compute now

$$\omega(v, w) - \omega_L((l_v, \xi_v), (l_w, \xi_w))$$

(we have to prove it is zero). Applying the definition of ω_L and the fact that $\xi_v(l_w) = \omega(v, l_w)$ we find

$$\omega(v, w) - \omega(v - l_v, l_w) + \omega(w - l_w, l_v).$$

For the middle term, using $\omega(l_v, l_w) = 0$ (L is isotropic), we find

$$\omega(v, w) - \omega(v, l_w) + \omega(w - l_w, l_v)$$

hence we find

$$\omega(v, w - l_w) + \omega(w - l_w, l_v) = \omega(v - l_v, w - l_w)$$

where we have also used the antisymmetry of ω. Since $v - l_v \in L'$ for all v and L' is isotropic, the last expression is indeed zero. □

3.1.9. Hermitian structures

Let's now briefly mention Hermitian structures (which can be thought of as analogues of inner products, but on complex vector spaces). First of all, given a complex vector space W (we reserve the letter V for real spaces), a Hermitian metric on W is an $\mathbb{R}$-bilinear map

$$h : W \times W \longrightarrow \mathbb{C}$$

which is $\mathbb{C}$-linear in the first argument (and then, because of the next axiom, $\mathbb{C}$-antilinear in the second), satisfies the conjugates symmetry

$$h(w_2, w_1) = \overline{h(w_1, w_2)} \quad \forall\, w_1, w_2 \in W$$

ans $h(w, w) \geq 0$ for all $w \in W$ with equality only for $w = 0$.

Given a real vector space V, a hermitian structure on V is a pair (J, h) consisting of a complex structure J on V and a Hermitian metric on the resulting complex vector space V_J,

$$h : V \times V \longrightarrow \mathbb{C}.$$

The main remark we want to make here is that such pairs can be unravelled and thought of in several equivalent ways. Here is the summary:

Lemma 3.22. *Given a real vector space V, there is a 1-1 correspondence between*

1. *Hermitian structures (h, J) on V.*

2. *pairs* (J, g) *consisting of a complex structure* J *and a metric* g *on* V *satisfying*
$$g(Ju, Jv) = g(u, v) \quad \forall\, u, v \in V,$$
3. *pairs* (J, ω) *consisting of a complex structure* J *and a symplectic structure* ω *on* V*, satisfying*
$$\omega(Ju, Jv) = \omega(u, v) \quad \forall\, u, v \in V,$$
and such that
$$\omega(u, Ju) > 0 \quad \forall\, u \in V \setminus \{0\}.$$

The defining relations between them are:
$$h(u, v) = g(u, v) - i\omega(u, v), \omega(u, v) = g(Ju, v), g(u, v) = -\omega(Ju, v).$$

Note that, using $J^2 = -\mathrm{Id}$, the equations in the statement can be written in the equivalent forms
$$g(Ju, v) = -g(u, Jv) \quad \forall\, u, v \in V,$$
$$\omega(u, Jv) = -\omega(Ju, v) \quad \forall\, u, v \in V.$$

Proof. The key remark is that, writing
$$h(u, v) = g(u, v) - i\omega(u, v)$$
(where, priory, g and ω are just $\mathbb{R}$-bilinear):

- the conjugated symmetry for h is equivalent to the fact that g is symmetric and ω is antisymmetric,
- the $\mathbb{C}$-linearity of h in the first argument of h is equivalent to
$$h(Ju, v) = ih(u, v)$$
and then to the two very last equations in the statement.
- the non-degeneracy of h is equivalent to that of g as a real bilinear form.

The rest is simple manipulations with these identities. □

Finally, there is some terminology that comes out of this lemma:

Definition 3.23. *Given a vector space* V*, a complex structure* J *on* V *and a symplectic structure* ω *on* V*, one says that* J *is* ω*-compatible if condition from point 3 of the previous proposition is satisfied.*

This terminology reflects the key idea for relating symplectic and complex structures- actually, on how to use complex structures in the study of symplectic ones. It is interesting to know that, for any symplectic ω, one can choose a ω-compatible J.

Exercise 84. Let (V, ω) be a symplectic vector space and fix a Lagrangian subspace $L \subset V$. Show that, fixing a metric g on V, there is a canonical way of constructing an ω-compatible complex structure J on V ("canonical" means that no extra-choices are necessary).

(Hint: first look at the proof of Prop 3.21, where you use as complement of L the orthogonal with respect to g. Then identify V with $L \oplus L$ where you use again the metric. On the later use $J(u, v) = (v, -u)$).

3.1.10. A few more remarks on linear G-structures

Looking back at the previous examples, there are a few remarks that we would like to make.

Example 3.24 (passing to larger groups)**.** Given subgroups

$$H \subset G \subset GL_n(\mathbb{R}),$$

any H-structure $\mathcal{S}$ on a vector space V induces a G-structure $\mathcal{S} \cdot G$ on V, where

$$\mathcal{S} \cdot G := \{\phi \cdot A : \phi \in \mathcal{S}, A \in G\}.$$

Equivalently, one can use the point of view given by Remark 3.2: while $\mathcal{S}$ is encoded by an element $\sigma_{\mathcal{S}} \in \mathrm{Fr}(V)/H$, $\sigma_{\mathcal{S} \cdot G}$ is the image of $\sigma_{\mathcal{S}}$ by the canonical projection

$$\mathrm{Fr}(V)/H \longrightarrow \mathrm{Fr}(V)/G.$$

Particular cases of this are:

1. for $SL_n(\mathbb{R}) \subset GL_n^+$ one recovers Remark 3.9: a volume element on V induces an orientation on V.
2. for $GL_k(\mathbb{C}) \subset GL_{2k}^+$ one recovers Remark 3.15: a complex structure on V induces an orientation on V.
3. for $Sp_k(\mathbb{R}) \subset SL_{2k}(\mathbb{R})$ one recovers Remark 3.18 : a symplectic form on V induces a volume element.
4. for $GL_n(\mathbb{Z}) \cap GL_n^+(\mathbb{R}) \subset SL_n(\mathbb{R})$ one recovers Remark 3.12 : an integral affine structure on an oriented vector space V induces a volume element.

Example 3.25. [e-structures] One case which looks pretty trivial in this linear discussion but which becomes important when we pass to manifolds is the case when G is the trivial group

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

In this case one talks about e-structures ("e" refers to the fact that G is trivial- fact that is often written as $G = \{e\}$ with e denoting the unit). Note that the subset

$$\mathcal{S} \subset \mathrm{Fr}(V)$$

encoding an e-structure has one element only. Hence an e-structure on V is the same thing as the a frame ϕ of V.

Example 3.26 (G-structures associated to tensors)**.** At the other extreme, many of the previous examples fit into a general type of structure: associated to various tensors. More precisely:

- for inner products, we deal with elements $g \in S^2V^*$.
- for volume elements, we deal with $\mu \in \Lambda^n V^*$.
- for symplectic forms, we deal with $\omega \in \Lambda^2 V^*$.
- for complex structures, we deal with $J \in V^* \otimes V$.

The fact that these elements were of the type we were interested in (positive definite for g, non-degenerate for ω, non-zero for μ, satisfying $J^2 = -\mathrm{Id}$) came from the fact that the standard models had these properties (actually, were determined by them, up to isomorphisms). The above examples can be generalized. In principle we can start with any

$$t_0 := \text{any canonical "tensor" on } \mathbb{R}^n.$$

To such a tensor we can associate a group $G(t_0)$ of linear isomorphisms, which preserve t_0, i.e.

$$G(t_0) := \{A \in GL_n(\mathbb{R}^n) : A^* t_0 = t_0\}.$$

Then a $G(t_0)$-structure on a vector space V corresponds to a tensor t on V, of the same type as the original t_0, with the property that (V, t) is isomorphic to $(\mathbb{R}^n, t_0)$.

3.2. Geometric structures on manifolds

3.2.1. G-structures on manifolds

Having discussed the notion of linear G-structures on vector spaces one can now pass to manifold an implement the philosophy mentioned at the very beginning of the chapter: a G-structure on a manifold M will be a collection of linear G-structures $\mathcal{S}_x$ on the tangent space T_xM, one for each $x \in M$, depending smoothly with respect to x. Such a family can be seen as a subset of the frame bundle of M,

$$\mathcal{S} := \cup_{x\in M}\mathcal{S}_x \subset \mathrm{Fr(TM)};$$

of course, the smoothness with respect to x will translate into the fact that $\mathcal{S}$ is smooth. Even more, the very notion of linear G-structure, that may seem a bit artificial at first, is very well suited to the notion of principal bundles. More precisely, we can now define:

Definition 3.27. *Let G be a Lie subgroup of $GL_n(\mathbb{R})$. A G-structure on an n-dimensional manifold M is a G-invariant submanifold*

$$\mathcal{S} \subset \mathrm{Fr(TM)}$$

with the property that, with respect to the induced action of G and with the projection $\pi : \mathcal{S} \to M$, it is a principal G-bundle.

Exercise 85. Let $\{\mathcal{S}_x\}_{x\in M}$ be a collection of G-structures on the tangent spaces T_xM and $\mathcal{S} \subset \mathrm{Fr(TM)}$ the resulting subspace. Show that the following are equivalent:

(i) $\mathcal{S}$ is a G-structure.
(ii) $\mathcal{S}_x$ carry smoothly with respect to x in the sense that: for each $x_0 \in M$ there exists a local frame ϕ of TM defined in a neighborhood U of x_0 such that $\phi(x) \in \mathcal{S}_x$ for all $x \in U$.

A central problem in the theory of G-structures is their so called integrability:

Definition 3.28. *Given a G-structure $\mathcal{S}$ on M, a coordinate chart (U, χ) of M is called adapted to the G-structure if the induced frame is adapted, i.e.:*

$$(\frac{\partial}{\partial\chi_1}(x), \ldots, \frac{\partial}{\partial\chi_n}(x)) \in \mathcal{S}_x \quad \forall\ x \in U.$$

A G-structure is called integrable if around any point $x \in M$ one can find an adapted coordinate chart.

Note that, while the definition of G-structures can be applied to any vector bundle $E \to M$, the integrability problem is specific to $E = TM$.

Remark 3.29 (isomorphism; the local model; the equivalence problem)**.** A bit similar to the discussion from the linear case, one can talk about:

S1: Isomorphism of G-structures: Given two G-structures, $\mathcal{S}$ on M and $\mathcal{S}'$ on M', an isomorphism between them is any diffeomorphism $f : M \longrightarrow M'$ with the property that

$$f_*(\mathcal{S}_x) = \mathcal{S}_{f(x)} \ \ \forall\ x \in M$$

where f_* is the map induced by $(df)_x : T_xM \longrightarrow T_{f(x)}M'$.

S2: The standard G-structure $\mathcal{S}_G^{\mathrm{can}}$ on $\mathbb{R}^n$: using the canonical identification of $T_x\mathbb{R}^n$ with $\mathbb{R}^n$ (induced by the standard basis $\partial/\partial x_i$), $(\mathcal{S}_G^{\mathrm{can}})_x$ is the standard linear G-structure on $\mathbb{R}^n$. Equivalently, the frames of $\mathbb{R}^n$ that are adapted to $\mathcal{S}_G^{\mathrm{can}}$ are precisely those of type $\phi = (\phi_1, \ldots, \phi_n)$,

$$\phi_i(x) = \sum g_{i,j}(x) \frac{\partial}{\partial x_j}(x)$$

where all the matrices $(g_{i,j}(x))_{i,j}$ belong to G.

With these, one has that G-structure $\mathcal{S}$ on M is integrable if and only if $(M, \mathcal{S})$ is locally isomorphic to $(\mathbb{R}^n, \mathcal{S}_G^{\mathrm{can}})$. In this way the integrability problem can be seen as a particular case of the so-called equivalence problem: when are two given G-structures locally equivalent?

Example 3.30 (Riemannian metrics)**.** The next subsections will be devoted to various examples. But let us mention already here the case when $G = O(n)$; the discussion from the linear case (see Proposition 3.6) shows that $O(n)$-structures on M are in 1-1 correspondence with Riemannian metrics on M. Of course, this is one of the motivating examples. However, unlike in most of the other examples that we will see, the resulting integrability problem is not so interesting. More precisely, the integrability of of Riemannian metric interpreted as an $O(n)$-structure is just too strong to impose. As we shall see, it will be equivalent to the fact that the associated Levi-Civita connection is flat.

3.2.2. $G = \{e\}$: frames and coframes

Let us briefly look at the extreme case when G is the trivial group. In that case we talk about e-structures (see also Example 3.25). We see that an e-structure on a manifold M is simply a global frame

$$\phi = (\phi_1, \ldots, \phi_n)$$

on M. The local model is, of course, $\mathbb{R}^n$ with the standard global frame $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_1}$. An isomorphism between two manifolds endowed with global frames, (M, ϕ) and (M, ϕ'), is any diffeomorphism $f : M \longrightarrow M'$ with the property that

$$(df)_x(\phi_i(x)) = \phi_i^{'}(f(x)) \;\; \forall \; c \in M.$$

The integrability of (M, ϕ) (when interpreted as an e-structure) means that: around any point $x \in M$ one can find a coordinate chart (U, χ) such that

$$\phi_i = \frac{\partial}{\partial \chi_i}$$

for all i. The integrability of a coframe can be characterized in terms of the "structure functions" of the coframes, which are the smooth functions $c_{i,j}^k$ that arise as the coefficients with respect to the frame ϕ of the Lie brackets of the vectors of ϕ:

$$[\phi_i, \phi_j] = \sum_k c_{i,j}^k \phi_k.$$

Theorem 3.31. *A frame ϕ is integrable if and only if its structure functions vanish identically.*

Proof. The condition is $[\phi_i, \phi_j] = 0$ for all i and j; this is clearly a local condition and is satisfied by the local model, so it is necessary in order to have integrability. For the sufficiency, we use the flows Φ_i^t of the vector fields ϕ_i. The condition we have implies that each two of them commute. Let $x_0 \in M$ be arbitrary and consider

$$F : \mathbb{R}^n \longrightarrow M, \quad F(t_1, \ldots, t_n) = \Phi_1^{t_1} \circ \ldots \Phi_n^{t_n}(x_0);$$

strictly speaking, this is defined on an open neighborhood of $0 \in \mathbb{R}^n$ (corresponding to the domains of the flows). Clearly $F(0) = x_0$. Computing the partial derivatives of F at an arbitrary point $x \in M$ we find the vectors $\phi_i(F(x))$. In particular, by the inverse function theorem, F is a local diffeomorphism from an open containing 0 to an open containing x_0; this defines the desired local chart. □

There is one more comment we would like to make here: very often one prefers to work with differential forms instead of vector fields. In particular, one prefers to work with coframes instead of frames, i.e. with a family of 1-forms:

$$\theta = (\theta^1, \ldots, \theta^n)$$

which, at each $x \in M$, induces a basis of T_x^*M. Of course, the two points of view are equivalent, as one has a 1-1 correspondence $\phi \longleftrightarrow \theta$ between frames and coframes given by the usual duality

$$\theta^i(\phi_j) = \delta_{i,j}.$$

However, when one goes deeper into the theory, the parallelism between the two points of view becomes a bit less obvious. One thing to keep in mind is that the role played by the Lie derivatives of vector fields, in the dual picture of forms, is taken by the DeRham differential. For instance, for a coframe θ, the resulting structure functions are characterized by

$$d\theta^k = \sum_{i<j} c_{i,j}^k \theta^i \wedge \theta^j.$$

Exercise 86. Prove that, indeed, given a frame ϕ and the induced coframe θ, the structure functions of ϕ satisfies the previous equation.

Of course, in the dual picture, the local model is $(\mathbb{R}^n, dx_1, \ldots, dx_n)$, the integrability of (M, θ) is about writing (locally) $\theta^i = d\chi_i$ and the previous proposition say that this is possible if and only if all the 1-forms are closed.

Exercise 87. Give now another proof of the previous proposition, using the fact that, locally, all closed 1-forms are exact.

3.2.3. Example: Orientations

We now look at the case of orientations, i.e. at $G = GL_n^+$. We continue the discussion from 3.1.3. According to the general philosophy, given a manifold M, we look at collections

$$\mathcal{O} = \{\mathcal{O}_x : x \in M\}$$

of orientations on the tangent spaces T_xM which vary smoothly with respect to x in the following sense: for any $x_0 \in M$ one can find a local frame ϕ of M defined over some open $U \subset M$ containing x_0, such that ϕ_x induces the orientation $\mathcal{O}_x$ for all $x \in M$ (see also Exercise 85). This is precisely what an orientation on the manifold M is. Hence an orientation on M is the same thing as an GL_n^+-structure on M.

Remark 3.32. We see that, from the point of view of G-structures, an isomorphism between two oriented manifolds $(M, \mathcal{O})$ and $(M', \mathcal{O}')$ is a diffeomorphism $f : M \longrightarrow M'$ with the property that $(df)_x : T_xM \longrightarrow T_{f(x)}M'$ sends oriented frames of M to oriented frames of M'.

Proposition 3.33. *Any orientation $\mathcal{O}$ on M, when interpreted as an GL_n^+-structure, is integrable. In other words, for any point $x_0 \in M$, one can find a coordinate chart (U, χ) around x_0 such that, for any $x \in U$, the orientation $\mathcal{O}_x$ on T_xM is induced by the frame*

$$\frac{\partial}{\partial \chi_1}(x), \ldots, \frac{\partial}{\partial \chi_n}(x).$$

Proof. We know we find a frame ϕ defined over some open U containing x_0 such that ϕ_x induces $\mathcal{O}_x$ for all $x \in U$. Choose a coordinate chart (U_0, χ) defined over some open $U_0 \subset U$ containing x_0 which we may assume to be connected, such that the corresponding frame at x_0 induces the orientation $\mathcal{O}_{x_0}$. For $x \in U_0$, the matrix $A(x)$ of coordinate changes from the frame ϕ_x to the frame induced by (U_0, χ) at x, has the entries $A_{i,j}(x)$ smooth with respect to $x \in U_0$. Hence also det(A) is smooth; the determinant is also non-zero, and we know it is strictly positive at x_0. Hence it is positive on the (connected) U_0. This implies that the frame induced by (U_0, χ) at any $x \in U_0$ induces the orientation $\mathcal{O}_x$. $\square$

The integrability of orientations gives rise to a slightly different way of looking at orientations on manifolds: via the choice of an atlas $\mathcal{A}$ for the smooth structure of M with the property that, for any (U, χ), (U', χ') in M, the change of coordinates

$$c_{\chi,\chi'} := \chi' \circ \chi^{-1}$$

(function between opens in $\mathbb{R}^n$) has positive Jacobian at all points. Such atlases are called oriented. Two oriented atlases $\mathcal{A}$ and $\mathcal{A}'$ are called oriented equivalent if $\mathcal{A} \cup \mathcal{A}'$ is an oriented atlas. We see that:

- a manifold M is orientable if and only if it admits an oriented atlas.
- the choice of an orientation is equivalent to the choice of an equivalence class of an oriented atlas.

Exercise 88. Do the same for arbitrary subgroups $G \subset GL_n(\mathbb{R})$. More precisely:

- define the notion of G-atlas and equivalence of G-atlases.
- explain how a G-atlas induces a G-structure and show that two different G-atlases induce the same G-structure if and only if they are equivalent.
- show that a G-structure comes from a G-atlas if and only if it is integrable.

Conclusion: the choice of an integrable G-structure is equivalent to the choice of an equivalence class of G-atlases (or to the choice of a maximal G-atlas).

Exercise 89. Exhibit an oriented atlas for the sphere S^n.

Of course, exhibiting oriented atlases is not quite the (practical) way to obtain orientations. Here is a general procedure (that works e.g. for S^n):

Exercise 90. A unit normal vector field on a hypersurface (or codimension one submanifold) S in $\mathbb{R}^n$ is a smooth map $\mathbf{n} : S \to \mathbb{R}^n$ with the property that $\mathbf{n}(x) \perp T_xS$ and $\|\mathbf{n}(x)\| = 1$ for all $x \in S$.

1. Prove that S is orientable if and only if S admits a unit normal vector field.

2. Let U be open in $\mathbb{R}^n$ and $\Phi : U \to \mathbb{R}$ be a smooth function. Let c be a regular value of Φ. Then the manifold $\Phi^{-1}(c)$ has a unit normal vector field and is therefore orientable.

For the curious reader. Do all manifolds admit an orientation? The answer is: no. However, most of them do.

It should be clear that any parallelizable manifold is also orientable. As pointed out in the previous exercises, the spheres are orientable. Actually, so are all the simply connected manifolds (i.e. for which the first fundamental group vanishes). What would be the non-orientable examples then? Well, probably the simplest/nicest are the even-dimensional (real) projective spaces $\mathbb{P}^{2n}$. The complex projective spaces however (as any complex manifold- see also below) are orientable.

And, there is a simple trick to replace a non-orientable manifold M by one which is orientable, denoted $\tilde{M}$, and which covers M via a projection

$$p : \tilde{M} \longrightarrow M$$

which is a 2-cover (each fiber consists of two points). Explicitly,

$$\tilde{M} = \{(x, \mathcal{O}_x) : x \in M, \mathcal{O}_x - \text{orientation on } T_xM\},\ p(x, \mathcal{O}_x) = x.$$

(I let you guess the smooth structure on $\tilde{M}$). For instance: which are the oriented covers of the (non-orientable) projective spaces $\mathbb{P}^{2n}$? (yes, it is the spheres S^{2n}).

Remark 3.34 (integration). One of the main uses of orientations comes from the fact that, once an orientation is fixed on a manifold M, one can integrate n-forms with compact supports, i.e. there is an associated integration map

$$\int_M : \Omega^n_{\text{cpt}}(M) \longrightarrow \mathbb{R},$$

where, as above, n is the dimension of M and subscript c denotes "compact supports" (i.e. forms that vanish outside a compact). Here are some details. To understand why n-forms (and not functions) and how the orientation is relevant, let us start with an n-form ω which is supported inside the domain of a coordinate chart

$$\chi : U \xrightarrow{\sim} \Omega \subset \mathbb{R}^n.$$

Using the basis induced by the coordinate chart for the tangent spaces T_xM, and then for $\Lambda^n T^*_x M$, we see we can write, on U:

$$\omega = f_\chi \circ \chi (d\chi_1) \wedge \ldots \wedge (d\chi_n)$$

for some smooth function $f_\chi : \Omega \longrightarrow \mathbb{R}$ (with compact supports). Of course, we would like to use the standard integration of functions on (opens in) $\mathbb{R}^n$ and define

$$\int_M \omega = \int_U \omega := \int_\Omega f_\chi(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

The question is: doesn't this depend on the choice of the coordinate chart on U? Of course, this should be related to the change of variable formulas for the standard integration: if $h : \Omega' \longrightarrow \Omega$ is a diffeomorphism between two opens in $\mathbb{R}^n$ then, for any $f \in C^\infty_c(\Omega)$,

$$\int_{\Omega'} |\text{Jac}(h)| f \circ h = \int_\Omega f,$$

where $|\text{Jac}(h)|$ is the absolute value of the Jacobian of h.

So, let's assume that

$$\chi' : U \longrightarrow \Omega'$$

is another coordinate chart. Then we obtain another function $f_{\chi'}$ on Ω' and the question is whether its integral coincides with that of χ. The relationship between the two functions can be expressed in terms of the coordinate change

$$h = \chi \circ (\chi')^{-1} : \Omega' \longrightarrow \Omega.$$

More precisely, writing

$$(d\chi_i)_x = \sum_j \frac{\partial h_i}{\partial x_j} (\chi'(x))(d\chi'_j)_x,$$

we find that $f_{\chi'} = \text{Jac}(h) f_\chi \circ h$. Hence we are almost there: if the Jacobian of h was everywhere positive then, by the standard change of variable formula, $\int_\Omega f_\chi = \int_{\Omega'} f_{\chi'}$ hence $\int_M \omega$ will be defined unambiguously. This is where the fixed orientation comes in: it allows us to talk about, and work only with, positive charts; and, for two such charts, the change of coordinates h will have positive Jacobian!

This explains how $\int_M \omega$ is defined if ω is supported inside a coordinate chart. Now, for an arbitrary $\omega \in \Omega^n_{\text{cpt}}(M)$ one uses a partition of unity to decompose ω as a finite sum

$$\omega = \omega_1 + \ldots + \omega_k$$

where each ω_i is supported inside some coordinate chart (with compact support there). Define then

$$\int_M \omega = \int_M \omega_1 + \ldots + \int_M \omega_k.$$

Again, this does not depend on the way we decompose ω as a sum as before; this follows basically from the additivity of the usual integral.

Note also the reason that we work with forms (and, e.g., not with functions): locally they are represented by functions f_χ (that we can integrate) and the way these functions change when we change the coordinates is compatible with the change of variables formula for the standard integration. We only had the "small problem" that the Jacobians had to be positive- which was fixed by the orientation. One can wonder here: ok, but can't one work with something else instead of forms, so that even the "small problem" disappears, so that the integration is defined even on non-orientable manifolds? The answer is yes: use "densities". We do not give further details here, but let us only mention that they are very similar to top forms: they are represented locally by functions f_χ and the formula when changing the coordinates is precisely the one coming from the change of variable formula (so that the integration can be defined without further complications). Of course, the choice of an orientation on M induces an identification between the space of densities and the one of n-forms, and then the resulting integration is the one we discussed.

Remark 3.35 (relationship with cohomology). Let $(M, \mathcal{O})$ be a compact n-dimensional oriented manifold, so that the integration

$$\int_M : \Omega^n(M) \longrightarrow \mathbb{R}$$

is well-defined. Then the Stokes formula implies that the integral vanishes on all exact forms (forms of type $d\theta$). Interesting enough (and not completely trivial) is the fact that also the converse is true if M is connected: if the integral of an n-form vanishes, then it must be exact.

This discussion fits very well with DeRham cohomology. Since $d\omega = 0$ for all n-forms ω (just because any form of degree strictly larger than the dimension of the manifold vanishes), the Stokes argument tells us that the integral descends to a linear map

$$\int_M : H^n(M) \longrightarrow \mathbb{R},$$

while the previous comment tells us that, if M is also connected, then the last map is injective (hence $H^n(M)$ is either zero or isomorphic to $\mathbb{R}$). This is an indication of the following characterization for orientability: a compact connected manifold is orientable if and only if $H^n(M)$ is isomorphic to $\mathbb{R}$ (see also the similar discussion on volume forms).

This discussion may be viewed as the starting point of an important tool of Algebraic/Differential Topology: Poincare duality. Given the compact, connected, oriented M then for any $k \in \{0, 1, \ldots, n\}$, using the wedge product of forms, passing to cohomology and then using the integration gives a pairing

$$H^k(M) \times H^{n-k}(M) \longrightarrow \mathbb{R}, \ ([\omega], [\eta]) \mapsto \int_M \omega \wedge \eta.$$

One version of the Poincare duality says that this pairing is perfect ("non-degenerate"). From this it follows for instance that, for odd dimensional M, its Euler characteristic is zero. The previous pairing is an important tool in the study of manifolds. For instance, for 4-dimensional manifolds, choosing $k = 2$, one obtains a non-degenerate bilinear form on $H^2(M)$- an algebraic invariant that tells us a lot about M. But, again, we are moving to other territories here ...

3.2.4. Example: Volume elements

We now look at the case of volume elements, i.e. at $G = SL_n(\mathbb{R})$; ee continue the discussion from 3.1.4. According to the general philosophy, given a manifold M, we look at collections

$$\mu = \{\mu_x : x \in M\}$$

of linear volume forms on the tangent spaces T_xM, varying smoothly with respect to x. I.e. at a top degree forms

$$\mu \in \Omega^n(M)$$

which do not vanish at any point $x \in M$. This is precisely what a volume element (or form) on the manifold M is. Hence a volume element on M is the same thing as an $SL_n(\mathbb{R})$-structure on M. In particular, the local model is $\mathbb{R}^n$ together with

$$\mu_{\text{can}} = (dx_1) \wedge \ldots \wedge (dx_n).$$

Remark 3.36. We see that, from the point of view of G-structures, an isomorphism between two manifolds endowed with volume forms, (M, μ) and (M', μ') is a diffeomorphism $f : M \longrightarrow M'$ with the property that $f^*(\mu) = \mu'$. These are called volume preserving diffeomorphisms.

Proposition 3.37. *Any volume form μ on a manifold M, when interpreted as a $SL_n(\mathbb{R})$-structure, is integrable. In other words, around any $x \in M$, one can find a coordinate chart (U, χ) such that, on U,*

$$\mu = d\chi_1 \wedge \ldots \wedge d\chi_n.$$

Proof. Start with an arbitrary chart (U, χ) around x. Writing μ in the resulting coordinates, it looks like

$$\mu = f d\chi_1 \wedge \ldots \wedge d\chi_n$$

for some smooth function f on U. Then just choose any function $\chi_1^{'}$ near x such that $\partial\chi_1^{'}/\partial\chi_1 = f$ and replace the old coordinates by $(\chi_1', \chi_2, \ldots, \chi_n)$. $\square$

Note that the linear story implies that the choice of a volume form on M determines the choice of an orientation on M. Actually more is true when it comes to the existence question:

Exercise 91. Show that a manifold M admits a volume form if and only if it is orientable. (Hint: time to exercise a bit with partitions of unity!)

For the curious reader. By the previous exercise, the question of existence of volume forms brings nothing new (see however below).

One of the main uses of volume forms comes from the fact that, once we fix a volume form μ on a manifold M, one can integrate compactly supported smooth functions, i.e. there is an induced integration map:

$$\int_M : C^{\infty}_{\text{cpt}}(M) \longrightarrow \mathbb{R}.$$

This goes via the integration of n-forms (see the previous subsection) and the fact that μ induces an orientation on M and an isomorphism

$$C^{\infty}(M) \xrightarrow{\sim} \Omega^n(M), \; f \mapsto f\mu.$$

In other words, the integral of a function M is defined as the integral of the n-form $f\mu$ (associated to the orientation induced by μ).

In particular, the choice of a volume form allows us to talk about the volume of M:

$$\text{Vol}_{\mu}(M) := \int_M 1$$

(where 1 is the constant function equal to 1). It is remarkable that this is invariant characterizes μ uniquely up to isomorphism:

Theorem 3.38. *[Moser] If μ and μ' are two volume elements on a compact manifold M then the following are equivalent:*

1. *there exists a diffeomorphism $h : M \longrightarrow M$ such that $\mu' = h^*\mu$.*
2. *$Vol_{\mu}(M) = Vol_{\mu'}(M)$.*

Proof. For the direct implication, note that the integration is invariant under oriented diffeomorphisms (...). Let's concentrate on the other, more difficult, implication. If you did the previous exercise, you noticed that an affine combinations of volume forms is a volume form. Hence for any $t \in [0, 1]$,

$$\mu_t := t\mu + (1-t)\mu'$$

is a volume form, and it is not difficult to see that it induces the same orientation as μ. Since $\int_M \mu = \int_M \mu'$ it follows (see the relationship of integration with DeRham cohomology) that $\mu - \mu'$ is exact. We conclude that there exists a form $\eta \in \Omega^{n-1}(M)$ such that

$$\frac{d\mu_t}{dt} = \mu - \mu' = d\eta$$

for all t. Moreover, since μ_t is a volume form, it is not difficult to see that the operation

$$\mathcal{X}(M) \longrightarrow \Omega^{n-1}(M), \quad X \mapsto i_X(\eta_t)$$

is an isomorphism (think first what happens for vector spaces) hence we find a vector field X_t such that

$$i_{X_t}(\mu_t) = -\eta.$$

All together, $\{X_t : t \in [0, 1]\}$ form a smooth family of vector fields on M (time dependent vector field). Finally, consider the flow of this family, which can be seen as a family of diffeomorphisms

$$\varphi^t : M \longrightarrow M.$$

Using the basic properties of flows, one computes:

$$\frac{d}{dt}\varphi_t^*\mu_t = \varphi_t^*(L_{X_t}(\mu_t) + \dot{\mu}_t) = \ldots = 0,$$

hence $\varphi_t^*\mu_t$ is constant with respect to t. In particular $\varphi_1^*\mu_1 = \mu_0$, i.e. $\varphi_1^*\mu = \mu'$. $\square$

Remark 3.39 (more on the relationship with integration)**.** Note that, since the integral of a volume form is always strictly positive, the resulting integration map

$$\int_M : H^n(M) \longrightarrow \mathbb{R}$$

is non-zero. Combining with the similar discussion from the orientability subsection, we deduce that : for an n-dimensional compact connected manifold, the following are equivalent:

- M is orientable.
- M admits a volume form.
- $H^n(M) \neq 0$.
- $H^n(M)$ is isomorphic to $\mathbb{R}$.

3.2.5. Example: foliations

The notion of p-directions discussed in the linear case in 3.1.5 becomes more interesting when we pass to manifolds. The resulting notion is that of p-dimensional distribution of a manifold M, by which we mean a rank p vector sub-bundle of the tangent bundle,

$$\mathcal{F} \subset TM.$$

We see that the choice of such an $\mathcal{F}$ is equivalent to the choice of an $GL(p, n-p)$-structure on M. What about integrability? In what follows, we will denote by $\Gamma(\mathcal{F})$ the space of (smooth) sections of $\mathcal{F}$, i.e. of vector fields X on M with the property that $X_x \in \mathcal{F}_x$ for all $x \in M$.

Definition 3.40. *A distribution $\mathcal{F} \subset TM$ is called involutive if:*

$$[X, Y] \in \Gamma(\mathcal{F}) \quad \forall \;\; X, Y \in \Gamma(\mathcal{F}).$$

A (p-dimensional) foliation on M is an (p-dimensional) distribution that is involutive.

Note that the local model is involutive. It is $\mathbb{R}^n$ with the distribution $\mathcal{F}_{\mathrm{can}}$ given by

$$\mathcal{F}_{\mathrm{can},x} = \mathrm{Span}\{\frac{\partial}{\partial x_1}(x), \ldots, \frac{\partial}{\partial x_p}(x)\}.$$

Theorem 3.41 (Frobenius). *For a p-dimensional distribution $\mathcal{F}$ on a manifold M, the following are equivalent:*

1. *$\mathcal{F}$ is involutive.*
2. *$\mathcal{F}$, interpreted as an $GL(p, n-p)$-structures, is integrable. Equivalently, for any $x_0 \in M$ one finds a coordinate chart*
$$(U, \chi) = (U, \chi_1, \ldots, \chi_p, \chi_{p+1}, \ldots, \chi_n)$$
around x_0 such that
$$\mathcal{F}_x = Span\{\frac{\partial}{\partial \chi_1}(x), \ldots, \frac{\partial}{\partial \chi_p}(x)\} \quad \forall \; x \in U.$$

Proof. (rather sketchy, but with all the main ingredients) The reverse direction should be clear since the bracket of two vector fields of type $\partial/\partial\chi_i$ is zero. Let us prove the direct implication. Fix $x_0 \in M$ and fix any coordinate chart (U, χ) around x_0. After eventually renumbering the coordinates, we may assume:

$$T_xM = \mathcal{F}_x \oplus \mathrm{Span}_{\mathbb{R}}\{\frac{\partial}{\partial \chi_{p+1}}(x), \ldots, \frac{\partial}{\partial \chi_n}(x)\}$$

for all x in a neighborhood $W \subset U$ of x_0 (note: for dimensional reasons the sum is any way direct; then, by assuming the previous equation to hold at x_0, it follows by a continuity argument that it holds in a neighborhood W of x_0). We may assume that W is some small ball. Consider the projection on the first p coordinates:

$$\pi : W \longrightarrow \mathbb{R}^p.$$

It follows that its differential restricted to $\mathcal{F}$ induces isomorphisms

$$(d\pi)_x : \mathcal{F}_x \xrightarrow{\sim} T_{\pi(x)}\mathbb{R}^p \quad \forall \; x \in W.$$

Hence we find

$$V_x^1, \ldots, V_x^p \in \mathcal{F}_x \quad (\text{for } x \in W)$$

which are projectable (via $(d\pi)$) to

$$\frac{\partial}{\partial x_1}(\pi(x)), \ldots, \frac{\partial}{\partial x_p}(\pi(x)) \in T_{\pi(x)}\mathbb{R}^p.$$

Moreover, each V^i is smooth w.r.t. $x \in W$ (show this using the smoothness of $\mathcal{F}$!). Hence we obtain the vector fields $V^1, \ldots, V^p$ defined on W, spanning $\mathcal{F}$ and π-projectable to $\partial/\partial x_1$, ..., $\partial/\partial x_p$. In general, the property of being π-projectable is compatible with the Lie bracket (show that!) hence the Lie brackets $[V^i, V^j]$ are π-projectable to

$$[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0.$$

Since they also belong to $\mathcal{F}$ (the involutivity condition!), from the choice of W, it follows that $[V^i, V^j] = 0$ for all i and j. Then, as in the proof of the similar result for e-structures (Theorem 3.31), it follows that the flows of the vector fields V^i combine to give F which is a diffeomorphism from a neighborhood of $0 \in \mathbb{R}^n$ to W; this provides the desired coordinates. □

Remark 3.42 (more about foliations)**.** A p-dimensional distribution $\mathcal{F}$ on M specifies certain directions. It is natural to look at curves that follow the given directions (as for flows of vector fields). Even better, one can look for integrals of $\mathcal{F}$, by which we mean submanifolds $L \subset M$ with the property that

$$T_x L = \mathcal{F}_x \quad \forall\, x \in L.$$

When is it true that through each point of M there passes an integral of $\mathcal{F}$? Well ... precisely in the involutive (integrable) case. In one direction, it is clear that the existence of integrals through each point implies involutivity, because for any two vector fields X and Y tangent to a submanifold L, $[X, Y]$ remains tangent to L. The Frobenius theorem implies the converse: around each x_0, choosing the coordinate chart as in the statement, we may assume that $\chi(x_0) = 0$ and then

$$\{x \in M : \chi^{p+1}(x) = \ldots = \chi^n(x) = 0\}$$

is an integral of $\mathcal{F}$ through x.

However, the story does not stop here: like for flows of vector fields, when one looks for *maximal* integral curves, one looks here for maximal integrals of $\mathcal{F}$. There are also called leaves of $\mathcal{F}$. These are (immersed) submanifolds $L \subset M$ which are integrals of $\mathcal{F}$ and which are maximal with this property. Starting from Frobenius theorem and proceeding as for vector fields, we see that leaves exist through each point, an the collection of all leaves of $\mathcal{F}$ gives a partition of M. The Frobenius theorem says that this partition looks locally like the partition

$$\mathbb{R}^n = \bigcup_{y \in \mathbb{R}^{n-p}} \mathbb{R}^p \times \{y\}.$$

Actually, that is precisely one should think of (and picture) a foliation: such a partition which, locally, looks trivially (the bundle $\mathcal{F}$ is handy to encode and work with the foliation). Foliation Theory study the intricate geometry of such partitions.

Exercise 92. Let $\lambda \in \mathbb{R}$ and consider the foliation on the torus $M = S^1 \longrightarrow S^1$ given by

$$\{\frac{\partial}{\partial \theta_1} + \lambda \frac{\partial}{\partial \theta_2}\}.$$

How do the leaves of $\mathcal{F}$ look like?

For the curious reader. The question of existence of foliations of a given codimension (where the codimension is the dimension of the ambient manifold minus the dimension of the foliation), on a given compact manifold, is a very interesting (and hard) question. Actually, the very birth of "Foliation Theory" is identified with the PhD thesis of George Reeb (1943), where he answered the question of existence of codimension one foliations on S^3. The answer was positive, and the examples he found is well-known under the name of "the Reeb foliation". It arises by first realizing S^3 as obtained from two copies of the solid torus $S^1 \times D^2$ (think on the picture) by gluing them along their boundary torus $S^1 \times S^1$ (indicated on picture **??**). into two solid tori.

In formulas, one thinks of the 3-sphere as

$$S^3 = \{(u, v) : u, v \in \mathbb{C} : |u|^2 + |v|^2 = 1\},$$

one consider the copy of the 2-torus inside S^3 given by

$$A = \{(u, v) \in S^3 : |v| = \frac{\sqrt{2}}{2}\}$$

and the two connected components X_1 and X_2 of $S^3 \setminus A$ (described similarly to A above, but with "$\leq$" and "$\geq$" instead of "$=$"). It is not difficult to see that X_1 and X_2 are indeed solid tori.

The idea is now to foliate (partition) each solid torus with a codimension one foliation which has, as one of the leaves, the boundary torus. This is done by wrapping around"planes" inside the solid torus, as indicated in pictures **??** and **??**.

After gluing, one gets a foliation on S^3. It is interesting to note that, although it is pretty clear that the result is smooth, the result can never be done analytically (the geometric phenomena of wrapping around is quite non-analytic). Actually, an old theorem of Haefliger says that S^3 does not admit any analytic codimension foliation. The theorem actually proves that compact manifolds that admit analytic codimension one foliation have infinite fundamental group.

Back to the existence of codimension one foliations on manifolds, it is not so difficult to see that the even dimensional spheres cannot carry such; actually, the existence of such a foliation on a compact simply connected manifold would imply the vanishing of the Euler characteristic. So one is left with the odd dimensional spheres. The case of S^5 was answered by Lawson in 1971 and then, after the work of several people, it was proved that all odd-dimensional spheres do admit codimension one foliations. This story (about codimension one foliations on compact manifold) came to an end around 1976 with the work of Thurston who showed that, on a compact orientable manifold M such a foliation exists if and only if its Euler characteristic vanishes. The story continues with the existence of foliations of other interesting dimensions/codimensions (e.g. there is a 2-dimensional one on S^7) etc etc ... and with Foliation Theory.

3.2.6. Example: Integral affine structures

We now pass to integral affine structures on manifolds- the linear version of which was discussed in 3.2.6. First of all, an almost integral affine structure on a manifold M is, by definition, a family

$$\Lambda = \{\Lambda_x\}_{x \in M}$$

of lattices Λ_x on the tangent spaces T_xM, which vary smoothly with respect to x in the sense that, for any $X_0 \in M$, there exists a local frame ϕ defined on a neighborhood U of x_0 such that

$$\Lambda_x = \mathrm{Span}_{\mathbb{Z}}\{\phi^1(x), \ldots, \phi_n(x)\} \ \ \forall\ x \in U.$$

As before, Λ can be interpreted as a subspace

$$\Lambda \subset TM.$$

We also see the we are talking about $GL_n(\mathbb{Z})$-structures on M. The adapted frames in this case are the ones whose components (vector fields) take values in Λ. Vector fields with this property are called *integral vector fields of* (M, Λ).

Definition 3.43. *We say that Λ is an integral affine structure if, when interpreted as an $GL_n(\mathbb{Z})$-structure, it is integrable.*

Note that the local model is $\mathbb{R}^n$ with the lattice

$$\Lambda_{\text{can}} = \mathbb{Z}\frac{\partial}{\partial x_1} + \ldots + \mathbb{Z}\frac{\partial}{\partial x_n}.$$

Hence the integrability of Λ means that, around any point we can find a coordinate chart (U, χ) such that

$$\Lambda_x = \mathbb{Z}\frac{\partial}{\partial \chi_1}(x) + \ldots + \mathbb{Z}\frac{\partial}{\partial \chi_n}(x)$$

for all $x \in U$. The atlas consisting of such charts have very special transition functions $F : V \longrightarrow W$ (between opens $V, W \subset \mathbb{R}^n$): they are integral affine, i.e. of type

$$F(v) = v_0 + A(v) \quad v_0 \in \mathbb{R}^n,\ A \in GL_n(\mathbb{Z}). \tag{3.44}$$

Such an atlas is called an integral affine atlas. As before, the existence of such an atlas is equivalent to the existence of an integral affine structure.

Ok, but can one characterize the integrability of Λ more directly (as we did for symplectic or complex structures?). The answer can be given in several different ways. One of them makes use of the dual

$$\Lambda^\vee \subset T^*M$$

of Λ defined by

$$\Lambda^\vee_x = \{\xi \in T^*_xM : \xi(\Lambda_x) \subset \mathbb{Z}\}.$$

Theorem 3.45. *For an almost integral affine structure Λ on a manifold M the following are equivalent:*

1. *Λ is integrable (hence an integral affine structure).*
2. *the Lie bracket of any two local vector fields which are integral (i.e. take values in Λ) vanishes.*
3. *$\Lambda^\vee$ is locally spanned by closed 1-forms, i.e. any point in M admits a neighborhood U and closed 1-forms $\theta_1, \ldots, \theta_n$ on U such that*

$$\Lambda^\vee_x = Span_{\mathbb{Z}}\{\theta_1(x), \ldots, \theta_n(x)\}.$$

Proof. That 1 implies 2 comes from the fact that the condition in 2 can be checked locally and is valid for the local model. Assume now that 2 holds. Locally, we choose any frame $\{\phi_1, \ldots, \phi_n\}$ which spans Λ (as in the description of smoothness above). The condition in 2 tells us that ϕ is integrable in the sense of e-structures (see subsection 3.2.2), hence we find charts (U, χ) such that $\phi_i = d\chi_i$, and this implies 1. The equivalence with 3 is similar, using the dual point of view (see again subsection 3.2.2). □

Exercise 93. Show that if a manifold M is orientable and admits an integral affine structure, then it admits a volume form.

3.2.7. Example: Complex structures

We now look at the case $G = GL_k(\mathbb{C})$, sitting inside $GL_{2k}(\mathbb{R})$ as explained in 3.2.7. The discussion of the linear case brings us to the following notion:

Definition 3.46. *An almost complex structure on a manifold M is a a vector bundle morphism*

$$J : TM \longrightarrow TM$$

satisfying $J^2 = -Id$.

Hence almost complex structures are the same thing as $GL_k(\mathbb{C})$-structure.

Remark 3.47 (from almost complex to orientations)**.** It should be clear from the linear story that an almost complex structure on M induces an orientation on M.

Again, the local model is $\mathbb{C}^k$, with the almost complex structure induced by the multiplication by i, after identifying the tangent spaces of $\mathbb{C}^k$ with $\mathbb{C}^k$. In other words, identifying

$$\begin{array}{ccc} \mathbb{C}^n & \xrightarrow{\sim} & \mathbb{R}^{2n} \\ (z_1 = x_1 + iy_1, \ldots, z_n = x_n + iy_n) & \mapsto & (x_1, \ldots, x_n, y_1, \ldots, y_n), \end{array}$$

we are talking about $\mathbb{R}^{2n}$ with the almost complex structure is given by

$$J_{\mathrm{can}}(\frac{\partial}{\partial x_k}) = \frac{\partial}{\partial y_k}, J_{\mathrm{can}}(\frac{\partial}{\partial y_k}) = -\frac{\partial}{\partial x_k}.$$

Definition 3.48. *The Nijenhuis tensor of an almost complex structure J is the map*

$$\mathcal{N}_J : \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M),$$
$$\mathcal{N}_J(X, Y) = [X, Y] + J([JX, Y] + [X, JY]) - [JX, JY].$$

Exercise 94. Show that $\mathcal{N}_J$ comes indeed from a tensor, i.e. it is $C^\infty(M)$-linear in its entries:

$$\mathcal{N}_J(fX, Y) = \mathcal{N}_J(X, fY) = f\mathcal{N}_J(X, Y)$$

for any two vector fields X, Y oi M and any smooth function f on M.

With these, one has the following:

Theorem 3.49 (Newlander-Nirenberg)**.** *For an almost complex structure J on M the following are equivalent:*

1. *the Nijenhuis tensor of J vanishes.*
2. *J, interpreted as a $GL_k(\mathbb{C})$-structure, is integrable. Equivalently, for any $x_0 \in M$, one finds a coordinate chart*

$$(U, \chi) = (U, x_1, \ldots, x_k, y_1, \ldots, y_k)$$

around x_0 such that, on U,

$$J(\frac{\partial}{\partial x_k}) = \frac{\partial}{\partial y_k}, \frac{\partial}{\partial y_k} = -\frac{\partial}{\partial x_k}.$$

Proof. too difficult to give here. □

Remark 3.50 (complex manifolds)**.** This implies that we are talking about complex manifolds- which are defined exactly in the same way as smooth (real) manifolds, but with charts taking values in opens in $\mathbb{C}^k$ and with transition function holomorphic: these are the charts that arise from the theorem. Indeed, if one considers a transition function corresponding to two such charts:

$$f = (f_1, \ldots, f_k, g_1, \ldots, g_k) : V \longrightarrow W \quad (V, W \subset \mathbb{R}^{2k} \text{ opens})$$

we see that

$$(df) \circ J_{\mathrm{can}} = J_{\mathrm{can}} \circ (df),$$

or, equivalently,

$$\frac{\partial f_k}{\partial x_j} = \frac{\partial g_k}{\partial y_j}, \frac{\partial f_k}{\partial y_j} = -\frac{\partial g_k}{\partial x_j},$$

which are precisely the Cauchy-Riemann equations that characterize the holomorphicity of $f_k + ig_k$. Hence F is holomorphic.

Note also that, for a complex manifold M, one can talk about its complex tangent space and, as a real vector space, it can be canonically identified with the standard tangent space of M, viewed as a smooth (real) manifold. This gives an intrinsic description of the almost complex structure on the smooth manifold underlying a complex manifold. It is also useful to think in coordinates. Then complex charts $(U, z_1, \ldots, z_k)$ induce then a basis over $\mathbb{C}$ of the tangent space

$$\frac{\partial}{\partial z_1}, \ldots, \frac{\partial}{\partial z_k}.$$

The identification of the complex tangent space (as a real vector space) with the standard tangent space of M comes from splitting th complex coordinates (of complex charts) as $z_k = x_k + iy_k$; then

$$\frac{\partial}{\partial z_k} = \frac{1}{2}(\frac{\partial}{\partial x_k} - i\frac{\partial}{\partial y_k}).$$

It is customary to consider also

$$\frac{\partial}{\partial \overline{z}_k} = \frac{1}{2}(\frac{\partial}{\partial x_k} + i\frac{\partial}{\partial y_k})$$

so that, for a smooth function f, the Cauchy-Riemann equations ensuring the holomorphicity of f take the form $\frac{\partial f}{\partial \overline{z}_k} = 0$.)

Remark 3.51 (complex foliations)**.** Almost complex structures can be interpreted (formally) like "complex foliations". This comes from the reinterpretation of a linear complex structure J on a (real) vector space V as a certain complex subspace $V^{1,0}$ of the complexification $V_{\mathbb{C}}$ of V. See Remark **??**. That discussion applied to each tangent space T_xM tells us that an almost complex structure J on a manifold M is encoded in the resulting $T^{1,0}M \subset T_{\mathbb{C}}M$ which, in terms of J is

$$T^{1,0}M = \{X - iJ(X) : X \in \mathcal{X}(M)\}$$

It is now easy to see that the condition that the Nijenhuis tensor of J vanishes is equivalent to the Frobenius-type condition:

$$[X, Y] \in \Gamma(T^{1,0}M) \quad \forall\, X, Y \in \Gamma(T^{1,0}M).$$

(where $[\cdot, \cdot]$ is just the Lie bracket of vector fields, extended by $\mathbb{C}$-linearity to the entire $\Gamma(T_{\mathbb{C}}M)$).

For the curious reader. Also in the case of complex structures, the question of existence of a complex structure on a manifold M is a rather hard one. For instance, it seems to be an open problem the question of whether S^6 admits a complex structure. On the other hand, from all the spheres, only S^2 and S^6 admit an almost complex structure; for S^2 it comes from an honest complex structure (by viewing S^2 as a complex manifold, e.g. as $\mathbb{CP}^1$); for S^6 it follows by using, again, the octonions. Actually, there is a simple trick that allows us to relate the existence of almost complex structures to parallelizability and then, using the fact that S^1, S^3 and S^7 are the only parallelizable sphere, to conclude that S^2 and S^6 are the only spheres which admit an almost complex structure. Here is the trick:

Exercise 95. Assume that $S^n \subset \mathbb{R}^{n+1}$ admits an almost complex structure J. Interpret it as a family of linear endomorphisms

$$J_y : P_y \longrightarrow P_y \quad \text{for } y \in S^n$$

of the hyperplane $P_y = \{v \in \mathbb{R}^{n+1} : \langle v, y\rangle = 0\} \subset \mathbb{R}^{n+1}$ orthogonal to y.

Pass now to $\mathbb{R}^{n+2}$; denote by $e_0, \ldots, e_{n+1}$ its standard basis and interpret $\mathbb{R}^{n+1}$ (and all its subspaces, e.g. the P_ys) as the subspace of $\mathbb{R}^{n+2}$ via the standard inclusion

$$(y_0, \ldots, y_n) \mapsto (y_0, \ldots, y_n, 0).$$

For $x = (x_0, \ldots, x_{n+1}) \in S^{n+1} \subset \mathbb{R}^{n+2}$ different from $\pm e_{n+1}$, we denote by $p(x)$ its orthogonal projection on $\mathbb{R}^{n+1}$,

$$p(x) = (x_0, \ldots, x_n) \in \mathbb{R}^{n+1}$$

and the induced element on the sphere

$$s(x) = \frac{1}{||p(x)||} p(x) \in S^n.$$

Consider the orthogonal projection onto $P_{s(x)}$,

$$\mathrm{pr}^{\perp}_{s(x)} : \mathbb{R}^{n+1} \longrightarrow P_{s(x)},\ \mathrm{pr}^{\perp}_{s(x)}(\lambda) = \lambda - \langle \lambda, s(x)\rangle s(x)$$

and define:

$$F_i(x) := x_{n+1}e_i - x_i e_{n+1} + ||p(x)|| J_{s(x)}(\mathrm{pr}^{\perp}_{s(x)}(e_i))$$

for all $x \in S^{n+1}$, where the very last term is set to be zero when it does not make sense (i.e. for $x = \pm e_{n+1}$). Show that $\{F^0, \ldots, F^n\}$ is a frame of S^{n+1}. Hence S^{n+1} is parallelizable.

Interesting enough, the product of any two odd dimensional spheres admits a complex structure.

3.2.8. Example: Symplectic forms

We now look at the case $G = Sp_k(\mathbb{R})$, sitting inside $GL_{2k}(\mathbb{R})$ as explained in 3.1.8. The discussion of the linear case brings us to the following notion:

Definition 3.52. *An almost symplectic structure on a manifold M is a 2-form*

$$\omega \in \Omega^2(M)$$

with the property that each ω_x is non-degenerate.

Hence almost symplectic structures on M are the same thing as $Sp_k(\mathbb{R})$-structures on M.

$$n = 2k.$$

Note that the local model is $\mathbb{R}^{2k}$, with

$$\omega_{\mathrm{can}} = dy_1 \wedge dx_1 + \ldots + dy_k \wedge dx_k,$$

where $(x_1, \ldots, x_k, y_1, \ldots, y_k)$ denote the coordinates of $\mathbb{R}^{2k}$. What about integrability?

Definition 3.53. *A symplectic structure on a manifold M is an almost symplectic structure $\omega \in \Omega^2(M)$ which is closed.*

Isomorphisms between symplectic manifolds (M, ω) and (M', ω') correspond to diffeomorphisms $f : M \longrightarrow M'$ satisfying $\omega = f^*\omega'$. They are also called symplectomorphisms.

Theorem 3.54 (Darboux)**.** *For an almost symplectic manifolds (M, ω), the following are equivalent:*

1. *ω is symplectic.*
2. *ω, interpreted as a $Sp_k(\mathbb{R})$-structure, is integrable. Equivalently, for any $x \in M$, one finds a coordinate chart*

   $$(U, \chi) = (U, x_1, \ldots, x_k, y_1, \ldots, y_k)$$

   around x such that

   $$\omega = dy_1 \wedge dx_1 + \ldots + dy_k \wedge dx_k.$$

Proof. This is very similar to the proof of Theorem 3.38. Let's write it in a slightly different order. The reverse implication is clear since the standard symplectic form is closed. Assume now that ω is symplectic. We may assume we are on $\mathbb{R}^{2n}$ with coordinates denoted (x, y) and we work around the origin. We may also assume that $\omega = \omega_{\mathrm{can}}$ at the origin (why???). The idea is to realize the desired diffeomorphism between open neighborhoods of the origin taking ω to ω_{can} as the flow at time $t = 1$ of a vector field. Just using vector fields does not work, but it will if we allow time-dependent vector fields. Let X_t be the time-dependent vector field we are looking for and let φ be its flow. The next idea is to connect ω to ω_{can} by a smooth family of forms, $\{\omega_t : t \in [0, 1]\}$, satisfying

$$\omega_0 = \omega_{\mathrm{can}}, \ \omega_1 = \omega$$

and make sure that $\varphi_t^*\omega_t$ stays constant or, equivalently, $\frac{d}{dt}\varphi_t^*\omega_t = 0$. Using the properties of the flow (see Exercise 96 below) we can write this condition as follows:

$$L_{X_t}(\omega_t) + \frac{d}{dt}\omega_t = 0,$$

or, if the ω_t's are all closed and using *Cartan's magic formula* $L_X = di_X + i_X d$:

$$d(i_{X_t}\omega_t) + \frac{d}{dt}\omega_t = 0.$$

That is all we have to do. Here is how one does it. First of all, one considers the most obvious family:

$$\omega_t = t\omega + (1-t)\omega_{\mathrm{can}},$$

so that the equation we have to solve (to find X_t) becomes:

$$d(i_{X_t}\omega_t) + (\omega - \omega_{\mathrm{can}}) = 0.$$

Note that, since $\omega - \omega_{\mathrm{can}}$ is closed, by working in a ball around the origin, we may assume it is exact (Poincare lemma), say of type $d\eta$ for some 1-form η. Then it is enough to solve the (simpler) equation

$$i_{X_t}(\omega_t) + \eta = 0.$$

Here we deal with the operation from vector fields to 1-forms which sends a vector V to the 1-form $i_V(\omega_t)$. This we know to be a bijection (in particular surjective) if ω_t was non-degenerate. This is certainly so for $t \in \{0,1\}$, so we just have to make sure that, after eventually shrinking the neighborhood of 0 that we are working on, all the ω_t are symplectic. Here is where the assumption that ω and ω_{can} coincide at 0 comes in. They also coincide with ω_t at 0. Since ω_t is non-degenerate at 0, it will be non-degenerate in a neighborhood W_t of 0. Since we are interested in $t \in [0,1]$ and since ω_t is a continuous family (even smooth) and $[0,1]$ is compact, it follows that one can choose the same W_t for all t; call it W. If you followed the argument, you see we are done. Otherwise, start from here and read the proof backwards. □

Exercise 96. Let α_t be a smooth family of d-forms and X_t a time-dependent vector field. Let φ_t be the flow generated by X_t. Prove that

$$\frac{d}{dt}\varphi_t^*\alpha_t = \varphi_t^*\left(\frac{d}{dt}\alpha_t + L_{X_t}\alpha_t\right).$$

(Hint: if you cannot figure it out, look at the appendix from Geiges' book.)

Remark 3.55 (from almost symplectic to volume forms and orientations)**.** Again, it should be clear by looking at the local picture that an almost symplectic structure ω on M induces a volume forms (hence also an orientation). It is customary to use a certain scaling: given ω one defines the Liouville (volume) form associated to ω as:

$$\mu_\omega := \frac{1}{k!}\omega^k = \frac{1}{k!}\underbrace{\omega \wedge \ldots \omega}_{k \text{ times}}.$$

Exercise 97. Using Riemannian metrics prove that if M admits an almost complex structure then it also admits and almost symplectic one. More precisely, show that if J is an almost complex structure and g is a Riemannian metric, then

$$\omega(X,Y) := g(X,JY) - g(JX,Y)$$

defines an almost symplectic structure on M.

Remark 3.56 (almost symplectic versus almost complex)**.** It is not true that an almost symplectic structure induces ("naturally") an almost complex one, or the other way. What is true however is that a manifold M admits an almost symplectic structure if and only if it admits an almost complex one. The reverse implication is rather easy to show once we have Riemannian metrics (Exercise 97). The direct implication is slightly more difficult, but it should be easy if you did already Exercise 84.

Exercise 98. Use Exercise 84 to show that if M admits an almost symplectic structure ω then it admits an almost complex structure J.

For the curious reader.

While the existence of an almost complex structure is equivalent to the existence of an almost symplectic structure, anything else can happen:

- There are compact manifolds which admit almost complex structures, but do not admit complex or symplectic structures. This was proven for the connected sum of three copies of $\mathbb{CP}^2$ by Taubes in 1995.
- there are compact manifolds that admit complex structures but do not admit symplectic ones. One example is $S^3 \times S^1$.
- There are compact manifolds that admit symplectic structures but not complex ones.

The key idea for relating (almost) symplectic with (almost) complex structure comes from the notion of ω-compatibility - see subsection 3.1.9 for the linear story; on manifolds, we require that condition at each point. With this, one can show that, given any almost symplectic structure ω one can find an ω-compatible almost complex structure. After that, one can proceed an use J as it was an actual complex structure, exploiting ideas from complex geometry, to derive information about ω. This is the basic idea behind the theory of "pseudo-holomorphic curves" in Symplectic Geometry.

Regarding existence results, due to the difference between symplectic and almost symplectic structures, there are really two questions to answer here. The question of which (compact) manifolds admit symplectic structure is a very hard one, especially in the case of a negative answer; this is due to the lack of (known) invariants/obstructions for symplectic structures. What one can say right away, for compact manifolds M, is that if they admit a symplectic structure, then they must be even dimensional and orientable. And a bit more: the second DeRham cohomology group $H^2(M)$ is non-zero (so this excludes all the sphere S^n with $n \neq 2$). This last remark follows using integration: if ω is a symplectic form, then one has an induced volume form and orientation and the integral of $[\omega]^k \in H^n(M)$ is non-zero, hence *omega* itself will represent a non-zero cohomology class. Here are some examples/results:

- in dimension 2, a symplectic structure is the same thing as a volume form. Hence orientable surfaces are symplectic.
- from all the spheres, only S^2 and S^6 admit an almost symplectic structure, and only S^2 admits a symplectic structure.
- the connected sum of three copies of $\mathbb{CP}^2$ admits an almost symplectic structure but neither a complex nor a symplectic one (Taubes, 1995).
- For any finitely presented group Γ one can find a compact 4-manifold M, which is symplectic, with fundamental group isomorphic to Γ (Gompf, 2000).
- The existence of an almost symplectic structure is equivalent to the existence of an almost complex structure (see also subsection **??**).
- On open manifolds (i.e. manifolds which are not compact), there is quite some flexibility- due to the so called h-principle of Gromov ("h" stands for "homotopy"). For symplectic structures it says that: starting with any almost symplectic structure on an open manifold M, one can smoothly deform it (through almost symplectic structures) into a symplectic structure.

Remark 3.57 (Hamiltonian dynamics)**.** The motion of a mass 1 particle in $\mathbb{R}^k$ in the presence of a potential force $\Phi(x) = -\frac{\partial V}{\partial x}$ is governed by Newton's second law, $\ddot{q} = \Phi(q)$. If we introduce the auxiliary variable $p = \dot{q}$, the total energy of the particle is given by

$$H = \frac{1}{2}p^2 + V(q)$$

and Newton's equation transforms into a system of first order ODE's (or, equivalently, a differential equation in the $2k$-dimensional space $\mathbb{R}^{2k}$ with coordinates (x, y)), known as a *Hamiltonian system*:

$$\frac{dy}{dt} = -\frac{\partial H}{\partial x}, \qquad \frac{dx}{dt} = \frac{\partial H}{\partial y}$$

The corresponding Hamiltonian flow φ_t sends (x_0, y_0) to the solution of the Hamiltonian system satisfying the initial condition $(x(0), y(0)) = (x_0, y_0)$. In classical mechanics, these diffeomorphisms were sometime referred to as *mechanical motions* and they have the property that they preserve the volume form

$$\mu = dy_1 \wedge dx_1 \wedge \ldots \wedge dy_k \wedge dx_k$$

The crucial observation here is that, in fact, mechanical motions preserve the canonical symplectic form on $\mathbb{R}^{2n}$, i.e. they are symplectomorphisms. Obviously this second property implies the first one, in view of

$$\mu = \mu_\omega = \frac{\omega^k}{k!}$$

and this should be more evidence that the similarity between the proofs of the Darboux theorem for symplectic forms and the Moser theorem for volume forms is not accidental (of course, the two coincide in dimension two). The volume preserving property of Hamiltonian flows already attractted a lot of attention more than a century ago. The fact that these mechanical motions preserve the symplectic form was first pointed out explicitly by Arnol'd in the 1960's and the question of whether volume-preserving diffeomorphisms were the same as symplectomorphisms is one that kept symplectic topologists busy for quite some time. The (negative) answer was provided by Gromov in the 1980's, with his famous proof of the non-squeezing theorem, and makes use of very sophisticated techniques (*J-holomorphic curves*).

So the punchline is: Hamiltonian mechanics makes sense and can be studied on any manifold endowed with a symplectic form. For any such manifold (M, ω), the symplectic form ω sets up a one-to-one correspondence between vector fields and one-forms on M (in other words, an isomorphism between the tangent and the cotangent bundle). So given a smooth function $H : M \to \mathbb{R}$ (the Hamiltonian function) there is a uniquely defined vector field X_H corresponding to the differential of H: it is called the Hamiltonian vector field and it is defined by $i_{X_H}\omega = -dH$. We are interested in integral curves of this vector field, i.e. curves γ on M that satisfy the equation:

$$\dot{\gamma}(t) = X_H(\gamma(t)),$$

which should be interpreted as the most general form of Hamilton's equations.

Exercise 99. Check that if $M = \mathbb{R}^{2n}$ with coordinates (x, y) and standard symplectic form ω_{can}, the Hamiltonian vector field is given by $J_0\nabla H = (\partial_y H, -\partial_x H)$ and we recover the classical system of Hamiltonian equations. Can you guess what the form of the Hamiltonian vector field should be for an arbitrary (M, ω)?

3.2.9. Example: Affine structures

We have discussed *integral affine structures* (where "'integral" reflects the role of $\mathbb{Z}$ in the story). What about affine structures? At least when it comes to integral affine atlases, it is clear that the role of $\mathbb{Z}$ is not important, and one can talk about affine atlases as those with the property that the transitions functions are of type (3.44) but with $A \in GL_n(\mathbb{R})$ (i.e. just affine transformations). But the rest of the story is more problematic: affine structure is a type of geometric structures which is interesting (and worth looking at) but which does not fit in the general theory of G-structures, at least not in an obvious way, and not in the way we discussed it so far. Nevertheless, a quite similar discussion can be carried out.

The central notion here is that of (affine) connection on M, by which we simply mean a connection on the vector bundle TM,

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M),\ (X, Y) \mapsto \nabla_X(Y).$$

Although we are not talking about G-structures, note that there is a natural notion of "frame adapted to ∇". More precisely, given ∇, the interesting vector fields are those $V \in \mathcal{X}(M)$ with the property that

$$\nabla_X(V) = 0 \quad \forall\ X \in \mathcal{X}(M).$$

These are called *flat vector fields* (flat with respect to ∇). With this, one can talk about flat (local) frames of (M, ∇)- and these are the adapted frames. Hence one can adapt Definition 3.28 to this situation:

Definition 3.58. *We say that an affine connection ∇ on M is integrable if around any point in M one can find a coordinate chart (U, χ) for which $\frac{\partial}{\partial \chi_i}$ are all flat.*

Continuing the analogy, note that one also has a local model: $\mathbb{R}^n$ with ∇^{can} uniquely determined by the condition that the standard vector fields $\frac{\partial}{\partial x_i}$ are flat. We see that the integrability of ∇ is equivalent to the fact that (M, ∇) is locally isomorphic to $(\mathbb{R}^n, \nabla^{\text{can}})$. And this gives rise to an atlas for which the change of coordinates are affine transformations between opens in $\mathbb{R}^n$ (hence of type $v \mapsto v_0 + A(v)$ with $v_0 \in \mathbb{R}^n$, $A \in GL_n(\mathbb{R})$). Conversely, any such atlas gives rise to an integrable ∇. Hence, with the obvious notion of equivalence of affine atlases, an integrable affine structure on M is the same things as an (equivalence class of an) affine atlas.

But can one characterize the integrability of ∇? This is where the curvature and the torsion of ∇, discussed in 1.4.3 and 1.3.1, come in:

$$T_\nabla \in \Omega^2(M, TM), \quad T_\nabla(X, Y) = \nabla_X(Y) - \nabla_Y(X) - [X, Y],$$

$$K_\nabla \in \Omega^2(M, \mathrm{End}(TM)), \quad K_\nabla(X,Y) = \nabla_X \circ \nabla_Y - \nabla_Y \circ \nabla_X - \nabla_{[X,Y]}.$$

It is clear that, for the local model, these tensors must vanish.

Theorem 3.59. *An affine connection is integrable if and only if its torsion and curvature vanish.*

Proof. The proof will be given later on. □

The local flat model of Riemannian geometry is $\mathbb{R}^n$, endowed with the standard metric given by

$$g^{\mathrm{can}}\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) = \delta_{i,j}.$$

Theorem 3.60. *A Riemannian manifold (M,g) is integrable as a $O(n)$-structure if and only if its Levi-Civita connection is flat, i.e. $K_\nabla = 0$.*

For this reason, such manifolds are also called flat Riemannian manifolds.

Proof. Follows immediately from the similar theorem for affine connections. □

CHAPTER 4

G-structures and connections

4.1. Compatible connections

4.1.1. Connections compatible with a G-structure

In the same way that we can talk about the compatibility of a connection with a metric, one can talk about the compatibility of a connection with any G-structure. Although the discussion can be carried out at the level of vector bundles E over M, we restrict here to the case $E = TM$ and to connections

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M).$$

Definition 4.1. *A connection*

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M)$$

on M is said to be compatible with the G-structure $\mathcal{S}$ if for any curve $\gamma : [0,1] \longrightarrow M$, the parallel transport along γ with respect to ∇ is an isomorphisms

$$T_\gamma^{0,1} : (E_{\gamma(0)}, \mathcal{S}_{\gamma(0)}) \longrightarrow (E_{\gamma(1)}, \mathcal{S}_{\gamma(1)})$$

between linear G-structures.

As we shall see, while connections on TM can be interpreted as principal bundle connections for Fr(TM), connections compatible with $\mathcal{S}$ can be interpreted as connections on the principal bundle $\mathcal{S}$. The first sign of this is the fact that TM can be obtained also from the principal bundle $\mathcal{S}$ by attaching as fiber a representation of G. The representation that we need here is the most natural one that we have around: $\mathbb{R}^n$ arising via the inclusion $G \subset GL_n$. As in the case $\mathcal{S} = \mathrm{Fr(TM)}$, and by the same simple arguments, we obtain:

Lemma 4.2. *For any G-structure $\mathcal{S}$ on M one has*

$$E(\mathcal{S}, \mathbb{R}^n) \cong TM,$$

as vector bundles over M.

And here are some equivalent formulations of the notion of compatible connection.

Proposition 4.3. *For a connection ∇ and a G-structure $\mathcal{S}$, the following are equivalent:*

(i) ∇ is compatible with $\mathcal{S}$.

(ii) For any local frame e over U, that belongs to $\mathcal{S}$ the resulting connection matrix (1.16) takes values in $\mathfrak{g}$:

$$\omega(\nabla, e) \in \Omega^1(U, \mathfrak{g}).$$

(iii) The induced principal bundle connection on Fr(M), represented by the connection 1-form

$$\omega \in \Omega^1(Fr(M), gl_r)$$

has the property that

$$\omega|_{\mathcal{S}} \in \Omega^1(\mathcal{S}, \mathfrak{g}).$$

Moreover, the correspondences $\nabla \mapsto \omega \mapsto \omega|_{\mathcal{S}}$ define bijections between:

- *connections ∇ satisfying (i) (or (ii)).*
- *connections on Fr(M) satisfying (iii).*
- *connections on the principal G-bundle $\mathcal{S}$.*

Proof. (a bit sketchy) It should be clear by now that the parallel transport with respect to ∇, when applied to the members of a frame, becomes the parallel transport with respect to ω. Hence, for the equivalence of (i) and (iii) it suffices to show that the parallel transport w.r.t. ω preserves $\mathcal{S}$ iff $\omega|_{\mathcal{S}}$ takes values in $\mathfrak{g}$. For the direct implication, one remarks that ω on a tangent vector $X = \dot{u}(0)$ can be computed by moving $u(t)$ to the fiber of $u(0)$ using parallel transport and then taking the derivative w.r.t. t at $t = 0$ (a tangent vector to the fiber, hence a resulting element in the Lie algebra). Hence, if the parallel transport preserves $\mathcal{S}$, if X is tangent to $\mathcal{S}$ then the entire process takes place on $\mathcal{S}$, giving rise to a vertical tangent vector to $\mathcal{S}$, hence an element in $\mathfrak{g}$. The converse is easier: $\omega|_{\mathcal{S}}$ is then a connection on $\mathcal{S}$ and for paths in $\mathcal{S} \subset$ Fr(TM), the horizontality w.r.t. $\omega|_{\mathcal{S}}$ clearlt coincides with the one w.r.t. ω. Hence the parallel transport w.r.t. ω, when applied to elements in $\mathcal{S}$, coincides with the parallel transport w.r.t. $\omega_{\mathcal{S}}$, hence they stay in $\mathcal{S}$.

The equivalence of (ii) and (iii) is rather immediate- use e.g. that $\omega(\nabla, e)$, when applied to $X_x \in T_xM$, coincides with $\omega(h_{e_x}(X_x))$.

The previous discussion and the 1-1 correspondence between connections on TM and principal bundle connections on Fr(TM) proves that $\nabla \mapsto \omega$ is a 1-1 correspondence. The main point that remains is to see that ∇ (or ω) can be reconstructed from $\omega_{\mathcal{S}}$; this follows e.g. from the previous lemma and the general construction of a connection on the associated vector bundles. □

Example 4.4 (Riemannian metrics)**.** Consider now the case $G = O(n)$ and, as in Example 3.2.8, we interpret the $O(n)$-structures on M as Riemannian metrics g on M. Given a Riemannian metric g, it follows that a connection ∇ on TM is compatible with g in the sense of Sp_k-structures if and only if, for all vector fields X, Y, Z on M,

$$L_X(g(Y, Z)) = g(\nabla_X(Y), Z) + g(Y, \nabla_X(Z))$$

i.e. if and only if ∇ is compatible with g in the sense of subsection 1.4.2. This statement follows from the interpretation of compatibility given by (ii) of the previous proposition (for the compatibility with the $O(n)$-structure) and by part 3 of Proposition 1.40 (for the compatibility with g).

Note that, for an arbitrary connection ∇ on TM, while its failure to be compatible with g is encoded by the expression

$$g(\nabla_X(Y), Z) + g(Y, \nabla_X(Z)) - L_X(g(Y, Z)),$$

this expression is $C^\infty(M)$-linear in all the entries (hence it defines a section of an appropriate vector bundle) and is symmetric in Y and Z. We see that this failure is encoded in an element denoted

$$\nabla(g) \in \Omega^1(M, S^2T^*M)$$

and called the derivative of g w.r.t ∇.

Example 4.5 (Symplectic structures)**.** Let us now pass to (almost) symplectic structures; hence we assume that $n = 2k$, $G = Sp_k$ and, as in subsection 3.2.8, we interpret the Sp_k-structures on M as presymplectic forms $\Omega \in \Omega^2(M)$.

Given an almost symplectic structure Ω, it follows that a connection ∇ on TM is compatible with Ω in the sense of Sp_k-structures if and only if, for all vector fields X, Y, Z on M,

$$L_X(\Omega(Y,Z)) = \Omega(\nabla_X Y, Z) + \Omega(Y, \nabla_X, Z). \tag{4.6}$$

This follows by exactly the same arguments as in the Riemannian case.

Note that, again, the failure of a connection ∇ to be compatible with Ω is measured by expressions that are $C^\infty(M)$-linear in the entries, but which are now antisymmetric iin Y and Z. I.e. by an element denoted

$$\nabla(\Omega) \in \Omega^1(M, \Lambda^2T^*M).$$

Example 4.7 (Complex structures)**.** Consider now the case of almost complex structures; hence we assume that $n = 2k$, $G = GL_k(\mathbb{C})$ and, as in subsection 3.2.7, we reinterpret $GL_k(\mathbb{C})$-structures on M as almost complex structures $J : TM \to TM$.

Given an almost complex structure J, it follows that a connection ∇ on TM is compatible with J in the sense of $GL_k(\mathbb{C})$-structures if and only if, for all vector fields X, Y on M,

$$\nabla_X(J(Y)) = J(\nabla_X(Y)).$$

In this case, the failure of a connection ∇ to be compatible with J is still $C^\infty(M)$-linear in the entries and defines a 1-form on M with values in $\mathrm{Hom(TM, TM)}$. Actually, we can be a bit more precise since, for each X, the map $\nabla_X \circ J - J \circ \nabla_X$ belongs to

$$\mathrm{Hom_J(TM, TM) := \{T : TM \to TM : J \circ T + T \circ J = 0.}$$

In other words, the failure of ∇ to be compatible with J is encoded into an element

$$\nabla(J) \in \Omega^1(M, \mathrm{Hom_J(TM, TM)}).$$

Example 4.8 (Foliations)**.** Consider now the case of foliations of rank p; hence we assume that $G = GL(p, n-p)$ and, as in subsection 3.2.5, we reinterpret $GL(p, n-p)$-structures on M as p-dimensional distributions $\mathcal{F} \subset TM$.

Given $\mathcal{F}$ it follows that a connection ∇ on TM is compatible with $\mathcal{F}$ in the sense of $GL(p, n-p)$-structures if and only if

$$\nabla_X(V) \in \Gamma(\mathcal{F}) \quad \forall\ X \in \mathcal{X}(M), V \in \Gamma(\mathcal{F}).$$

Again, the failure of such a compatibility condition is measured by a tensor that is obtained by looking at the class modulo $\Gamma(\mathcal{F})$, which is a section of the normal bundle $TM/\mathcal{F}$ of $\mathcal{F}$,

$$\nabla_X(V) \text{ mod } \Gamma(\mathcal{F}) \in \Gamma(\mathrm{TM}/\mathcal{F}).$$

This defines an element

$$\nabla(\mathcal{F}) \in \Omega^1(M, \mathrm{Hom}(\mathcal{F}, \mathrm{TM}/\mathcal{F}))$$

whose vanishing is equivalent to the compatibility of ∇ with $\mathcal{F}$.

4.1.2. The infinitesimal automorphism bundle

We now discuss the infinitesimal automorphism bundle of a G-structure $\mathcal{S}$. We start with the linear story.

Definition 4.9. *Given a linear G-structure $\mathcal{S}$ on a vector space V we define the automorphism group of $(V, \mathcal{S})$,*

$$GL(V, \mathcal{S}) \subset GL(V),$$

as the group of all isomorphism from $(V, \mathcal{S})$ to itself. The infinitesimal automorphism algebra of $\mathcal{S}$ is defined as the Lie algebra of $GL(\mathcal{S})$, denoted

$$gl(V, \mathcal{S}) \subset gl(V).$$

Note that $GL(V, \mathcal{S})$ is a closed subgroup of $GL(V)$; hence it is a Lie group and:

$$gl(V, \mathcal{S}) = \{A \in gl(V) : \exp(tA) \in GL(V, \mathcal{S}) \quad \text{for all t near } 0\}.$$

Actually, one has isomorphisms

$$GL(V, \mathcal{S}) \cong G, \quad gl(V, \mathcal{S}) \cong \mathfrak{g},$$

just that such isomorphisms are not canonical. More precisely, choosing an element $\phi \in \mathcal{S}_V$ and interpreting it as an isomorphism $\hat{\phi}$ between $\mathbb{R}^n$ with the standard linear G-structure and $(V, \mathcal{S}_V)$, $g \mapsto \hat{\phi} \circ g \circ \hat{\phi}^{-1}$ defines the desired isomorphisms.

Passing to manifolds:

Definition 4.10. *Given a G-structure $\mathcal{S}$ on a manifold M, define the bundle of infinitesimal automorphisms of $\mathcal{S}$ as the vector sub-bundle*

$$\mathfrak{g}_{\mathcal{S}} := gl(TM, \mathcal{S}) \subset gl(TM)$$

(where $gl(TM) = \mathrm{End(TM)}$) with fibers

$$\mathfrak{g}_{\mathcal{S},x} := gl(T_xM, \mathcal{S}_x) \subset gl(T_xM).$$

Example 4.11. Consider $G = O(n)$. Then:

(i) for a linear $O(n)$-structure on a vector space V, i.e. an inner product g on V, we obtain the group of isometries of (V, g), with the Lie algebra

$$\mathrm{gl}(V, g) = \{T \in \mathrm{gl}(V) : g(Tu, v) + g(u, Tv) = 0 \quad \forall \quad u, v \in V\}.$$

(ii) for a $O(n)$-structure on M, i.e. a Riemannian structure g on M, we obtain the vector bundle over M whose fiber at $x \in M$ is $gl(T_xM, g_x)$.

The same applies also to the various other examples, such as:

- $G = Sp_k$ and almost symplectic forms when one deals with a the Lie algebra

$$\mathrm{gl}(V, \Omega) = \{T \in \mathrm{gl}(V) : \Omega(Tu, v) + \Omega(u, Tv) = 0 \quad \forall \quad u, v \in V\}$$

  (associated to a symplectic vector space (V, Ω)).
- $G = GL_k(\mathbb{C})$ and almost complex structures when one deals with

$$\mathrm{gl}(V, J) = \{T \in \mathrm{gl}(V) : J(T(u)) = T(J(u)) \ \ \forall \ u \in V\}$$

  (associated to a complex vector space (V, J)).

- $G = GL(p, n-p)$ when one deals with
$$\mathrm{gl}(V, W) = \{T \in \mathrm{gl}(V) : T(W) \subset W\}$$
(associated to an n-dimensional vector space V together with a p-dimensional subspace W).

Finally, we claim that $\mathfrak{g}_{\mathcal{S}}$ is precisely the vector bundle that one obtains by attaching the adjoint representation of G to $\mathcal{S}$:

Proposition 4.12. *For any G-structure $\mathcal{S}$ on M one has*
$$E(\mathcal{S}, \mathfrak{g}) \cong gl(\mathcal{S}).$$

Proof. Exercise. □

Remark 4.13 (new vector bundles out of $\mathfrak{g}_{\mathcal{S}}$)**.** We will encounter several quite complicated vector bundles associated to a G-structure $\mathcal{S}$. However, while the "complicated" part comes from the linear algebra behind them, the passing to vector bundles is very simple and rather general. Here is the general discussion that separates the two steps.

The starting point is the linear context consisting of:

(4.14) a finite dimensional vector space V and a linear subspace $\mathfrak{g} \subset \mathrm{gl(V)}$.

Assume that to such a data we associate some new vector space $\mathcal{C}(\mathfrak{g})$ (well, to be more precise we should really use the notation $\mathcal{C}(\mathfrak{g}, V)$. Moreover, we also assume that $\mathcal{C}$ is functorial with respect to the obvious notion of isomorphisms between such pairs. It follows that when $\mathfrak{g} \subset \mathrm{gl}(V)$ arises from a Lie group $G \subset GL(V)$, then $\mathcal{C}(\mathfrak{g})$ is a representation of G (as action of G on $(\mathfrak{g}, V)$, one just uses the adjoint action on $\mathfrak{g}$ and the linear action on V).

With this, for a G-structure $\mathcal{S}$ on M:

(i) one obtains a vector bundle $\mathcal{C}(\mathfrak{g}_{\mathcal{S}})$ obtained by applying $\mathcal{C}$ to $\mathfrak{g}_{\mathcal{S},x} \subset \mathrm{gl}(T_xM)$:
$$\mathcal{C}(\mathfrak{g}_{\mathcal{S}})_x = \mathcal{C}(\mathfrak{g}_{\mathcal{S},x}).$$

(ii) $\mathcal{C}(\mathfrak{g}_{\mathcal{S}})$ can be obtained by attaching to $\mathcal{S}$, viewed as a principal G-bundle, the representation $\mathcal{C}(\mathfrak{g})$ of G:
$$\mathcal{C}(\mathfrak{g}_{\mathcal{S}}) \cong E(\mathcal{S}, \mathcal{C}(\mathfrak{g})). \tag{4.15}$$
(this follows from the previous lemma).

4.1.3. The compatibility tensor

Next we show that, as in the particular cases described in Examples 4.4, 4.5, 4.7 and 4.8, the failure of a connection ∇ on TM to be compatible with a given G-structure will be measured by a tensor or, more precisely, by a 1-form on M with values in the vector bundle
$$\mathcal{N}(\mathfrak{g}_{\mathcal{S}}) := \mathrm{gl}(TM)/\mathfrak{g}_{\mathcal{S}}.$$
Note that the this vector bundle follows the pattern we have just described in Remark 4.13, starting with $\mathcal{N}$ which associates to a pair $(\mathfrak{g}, V)$ as in (4.14) the vector space $\mathcal{N}(\mathfrak{g}) = \mathrm{gl}(V)/\mathfrak{g}$. In particular,
$$\mathcal{N}(\mathfrak{g}_{\mathcal{S}})_x = \mathcal{N}(\mathfrak{g}_{\mathcal{S},x})$$
and one has the principal bundle description
$$\mathcal{N}(\mathfrak{g}_{\mathcal{S}}) \cong E(\mathcal{S}, \mathcal{N}(\mathfrak{g})). \tag{4.16}$$

Example 4.17. This bundle is related to Examples 4.4, 4.5, 4.7 and 4.8 as follows:

(i) For a vector space V endowed with an inner product, one has
$$\mathcal{N}(\mathrm{gl}(V,g)) \cong S^2V^*,$$
induced by the map $\phi : \mathrm{gl}(V) \to S^2V^*$
$$\phi(T)(u,v) = g(T(u),v) + g(u,T(v)).$$
Hence, in this case
$$\mathcal{N}(\mathfrak{g}_{\mathcal{S}}) \cong S^2T^*M.$$

(ii) For a vector space V endowed with an almost symplectic structure Ω the story is completely similar, with
$$\mathcal{N}(\mathrm{gl}(V,\Omega)) \cong \Lambda^2V^*,$$
and this completely agrees with the spaces arising in Example 4.5.

(iii) For a vector space V endowed with an almost complex structure J one has
$$\mathcal{N}(\mathrm{gl}(V,J)) \cong \mathrm{Hom_J(V,V)}$$
induced by the map $\phi : \mathrm{gl}(V) \to \mathrm{Hom_J(V,V)}$,
$$\phi(T)(u) = T(J(u)) - J(T(u)).$$
Again, this completely agrees with the spaces arising in Example 4.7.

(iv) For a vector space V together with a subspace W one has
$$\mathcal{N}(\mathrm{gl}(V,W)) \cong \mathrm{Hom(W,V/W)}$$
induced by the map
$$\phi(T)(w) = T(w) \text{ mod W } \in \mathrm{V/W}.$$
Again, this completely agrees with the spaces arising in Example 4.8.

In general, we have:

Proposition 4.18. *For G-structures $\mathcal{S}$ on M and connections ∇ on TM, one has an associated element*
$$\nabla(\mathcal{S}) \in \Omega^1(M, \mathcal{N}(\mathfrak{g}_{\mathcal{S}}))$$
with the property that ∇ is compatible with $\mathcal{S}$ if and only if $\nabla(\mathcal{S}) = 0$.

Of course, it is interesting to also see the actual construction of $\nabla(\mathcal{S})$. There are several ways in which one can introduce $\nabla(\mathcal{S})$- ranging from a more geometric but also more involved one, up to the easiest one given in Remark 4.21 but which lacks any intuition.

The most intuitive one arises by analyzing the variation of $\mathcal{S}$ with respect to the parallel transport induced by ∇. Let us look at a curve γ in M starting at $x \in M$, and a curve u in $\mathcal{S}$ sitting above γ. Then $u(t) \in \mathcal{S}_{\gamma(t)}$ is a frame at $\gamma(t)$ and we can use the inverse of the parallel transport along γ,
$$T_\gamma^{t,0} = (T_\gamma^{0,t})^{-1} : T_{\gamma(t)}M \to T_xM,$$
to move back to a frame at x; thinking of a frame at a point y as a linear isomorphism $\mathbb{R}^n \to T_yM$, the new frame at x is
$$T_\gamma^{t,0} \circ u(t) \in \mathrm{Fr(T_xM)}.$$

The question is whether this still belongs to $\mathcal{S}_x$. This can be measured by comparing this frame with $u(0)$, i.e. by looking at

$$A_u(t) = T_\gamma^{t,0} \circ u(t) \circ u(0) \in \mathrm{GL}(T_x M);$$

the question is whether this path takes values in $\mathrm{GL}(T_x M, \mathcal{S}_x)$. Nota that this path starts, at $t = 0$, with the identity. Hence the variation (at 0) that we are interested in is measured by the derivative of A_u, modulo $\mathrm{gl}(T_x M, \mathcal{S}_x) = \mathfrak{g}_{\mathcal{S},x}$:

$$\left[\frac{dA_u}{dt}(0)\right] \in \mathrm{gl}(T_x M)/\mathfrak{g}_{\mathcal{S},x} = \mathcal{N}(\mathfrak{g}_{\mathcal{S},x}). \tag{4.19}$$

The following is a nice exercise:

Lemma 4.20. *The expression (4.19) does only depends on $X_x := \dot{\gamma}(0)$ and not on the choice of u sitting above γ.*

This allows us to define $\nabla(\mathcal{S}) \in \Omega^1(M, \mathcal{N}(\mathfrak{g}_{\mathcal{S}}))$ which sends $X_x \in T_x M$ to

$$\nabla_{X_x}(\mathcal{S}) := \left[\frac{dA_u}{dt}(0)\right] \in \mathcal{N}(\mathfrak{g}_{\mathcal{S},x}).$$

Unraveling this, one obtains a more algebraic description of $\nabla(\mathcal{S})$, which could itself be taken as the definition. This is based on the characterization from (iii) of Proposition 4.3: interpreting an arbitrary ∇ as a connection 1-form on Fr(TM), the failure of the compatibility is the failure of $\omega|_{\mathcal{S}} \in \Omega^1(\mathcal{S}, gl_n)$ to be $\mathfrak{g}$-valued, i.e. the non-vanishing of the class induced in the quotient

$$[\omega|_{\mathcal{S}}] \in \Omega^1(S, gl_n)/\Omega^1(S, \mathfrak{g}) = \Omega^1(\mathcal{S}, gl_n/\mathfrak{g}).$$

Moreover, this form on $\mathcal{S}$ is clearly basic, hence it can be interpreted as a 1-form on M with coefficients in the vector bundle $E(\mathcal{S}, \mathcal{N}(\mathfrak{g}))$. Via the identification (4.16), $[\omega|_{\mathcal{S}}]$ can be interpreted as an element $\nabla(\mathcal{S})$ as in the statement.

4.1.4. The space of compatible connections

Here we discuss the freedom one has in choosing a connection compatible with a G-structure $\mathcal{S}$. In the case $\mathcal{S} = \mathrm{Fr(TM)}$, i.e. usual connections on TM with no extra-condition, the outcome is that the space of connections ∇ on TM is an affine space modeled by the vector space $\Omega^1(M, \mathrm{End(TM)})$. This is due to (and translates into) two remarks:

(i) while the space of connections is not a vector space, any two connections on TM, ∇^1 and ∇^2, give rise to a vector

$$\xi(\nabla^1, \nabla^2) \in \Omega^1(M, \mathrm{End(TM)})$$

(that we could denote by $\overrightarrow{\nabla^1 \nabla^2}$). This is simply the difference between the two operators,

$$\xi(\nabla^1, \nabla^2) := \nabla^2 - \nabla^1 : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM);$$

the $C^\infty(M)$-non-linearity of connections in the second argument cancels out and allows us to interpret their difference $\xi(\nabla^1, \nabla^2)$ as desired.

(ii) Conversely, for any connection ∇ and any vector $\xi \in \Omega^1(M, \mathrm{End(TM)})$ we have a new connection

$$\nabla^\xi : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM), \quad \nabla^\xi_X(Y) = \nabla_X(Y) + \xi(X)(Y).$$

The choice of a connection ∇ can be seen as the choice of an origin in the affine space, and gives rise to a bijection between $\Omega^1(M, \mathrm{End}(\mathrm{TM}))$ and the space of connections (given by $\xi \mapsto \nabla^{\xi}$).

Of course, in the case of a G-structure $\mathcal{S}$ on M, the space of connections compatible with $\mathcal{S}$ is smaller. However, it continues to have the same type of structure. And here is where the infinitesimal automorphism bundle $\mathfrak{g}_{\mathcal{S}} \subset \mathrm{End}(\mathrm{TM})$ enters:

Lemma 4.21. *The space of connections compatible with a G-structure $\mathcal{S}$ is an affine space with underlying vector space $\Omega^1(M, \mathfrak{g}_{\mathcal{S}})$. More explicitly: given a connection ∇ compatible with $\mathcal{S}$, ∇^{ξ} is still compatible with $\mathcal{S}$ for any $\xi \in \Omega^1(M, \mathfrak{g}_{\mathcal{S}})$, and any other connection compatible with $\mathcal{S}$ arises in this way.*

Proof. Just use the previous discussion and the characterization (ii) from the Proposition 4.3. □

Remark 4.22 (the compatibility tensor again)**.** Note that one can now an even simpler algebraic description (but also an even less insightful one!) for the compatibility tensor from Proposition 4.18. Namely, for a connection ∇ on TM and a G-structure $\mathcal{S}$ on M,

(i) choose a connection ∇^0 compatible with $\mathcal{S}$,
(ii) consider $\xi := \nabla - \nabla^0 \in \Omega^1(M, gl(TM)$ and its class modulo $\Omega^1(M, \mathfrak{g}_{\mathcal{S}})$,
$$[\xi] \in \Omega^1(M, \mathcal{N}(\mathfrak{g}_{\mathcal{S}})).$$

Then $[\xi]$ does not depend on the choice of ∇^0 and coincides with the compatibility tensor $\nabla(\mathcal{S})$ from Proposition 4.18.

Note also the the previous discussion can be carried out at the level of $\mathcal{S}$, where we look at principal G-bundle connections $\omega \in \Omega^1(\mathcal{S}, \mathfrak{g})$. Due to the condition $\omega(a(v)) = v$ for connections, we see again that connections on $\mathcal{S}$ do not form a vector space but an affine space. The difference of two connections ω^1 and ω^2 gives now an element
$$\xi(\omega^1, \omega^2) = \omega^2 - \omega^1 \in \Omega^1(\mathcal{S}, \mathfrak{g})_{\mathrm{bas}}$$
and a discussion completely similar to the one above shows that the space of connections on $\mathcal{S}$ is an affine space with underlying vector space $\Omega^1(\mathcal{S}, \mathfrak{g})_{\mathrm{bas}}$. Of course, the two discussions are just two faces of the same phenomenon, one carried out at the level of TM and one at the level of $\mathcal{S}$. Indeed, while the previous proposition gives a bijection $\nabla \longleftrightarrow \omega$ between connections compatible with $\mathcal{S}$ and principal bundle connections on $\mathcal{S}$, the identification between the underlying vector spaces,
$$\Omega^1(M, \mathfrak{g}_{\mathcal{S}}) \cong \Omega^1(\mathcal{S}, \mathfrak{g})_{\mathrm{bas}}$$
is just the canonical identification described in Proposition 2.49, applied to $P = \mathcal{S}$ and $V = \mathfrak{g}$, which give rise to the vector bundle $E = \mathfrak{g}_{\mathcal{S}}$ (cf. Proposition 4.12).

4.2. Torsion and curvature (and their relevance to integrability)

4.2.1. Some linear algebra that comes with G-structures

In this subsection we describe several new vector bundles associated to a G-structure. The reason to have this subsection is to make the exposition in the next subsections more fluent. However, some of these constructions may be a bit too hard to digest at once, without any motivation, so it may be a good idea to look at them in more detail only whenever they are actually needed (when they do arise naturally!).

The vector bundles that we need all fit the general construction from Remark 4.13, hence what we have to do here is to describe the linear algebra behind them. Therefore, we place ourselves in the linear context and start with a pair $(\mathfrak{g}, V)$ as in (4.14).

The first two constructions will become relevant when, starting with a connection ∇ compatible with a G-structure $\mathcal{S}$ and its torsion, we want to extract the intrinsic information that the torsion encodes (intrinsic in the sense that it does not depend on the choice of the connection). We consider the linear map

$$\partial : \mathrm{Hom}(V, \mathfrak{g}) \longrightarrow \mathrm{Hom}(\Lambda^2 V, V),\ \partial(\phi)(u, v) = \phi(u)(v) - \phi(v)(u). \tag{4.23}$$

Definition 4.24. *For a linear subspace $\mathfrak{g} \subset gl(V)$, we define*

1. the first prolongation of $\mathfrak{g}$, denoted

$$\mathfrak{g}^{(1)} := \mathrm{Ker}(\partial) \subset \mathrm{Hom}(\mathrm{V}, \mathfrak{g}). \tag{4.25}$$

2. the torsion space of $\mathfrak{g}$, defined as the cokernel of ∂:

$$\mathcal{T}(\mathfrak{g}) := \frac{Hom(\Lambda^2 V, V)}{\partial(Hom(V, \mathfrak{g}))}. \tag{4.26}$$

We will apply these constructions to the fibers of the infinitesimal automorphism bundle $\mathfrak{g}_{\mathcal{S}}$ of a G-structure $\mathcal{S}$,

$$\mathfrak{g}_{\mathcal{S},x} \subset \mathrm{End}(\mathrm{T_xM}) = \mathrm{gl}(\mathrm{T_xM}).$$

They give rise to the first prolongation of the infinitesimal automorphism bundle

$$\mathfrak{g}_{\mathcal{S}}^{(1)} \subset \mathrm{Hom}(\mathrm{TM}, \mathfrak{g}_{\mathcal{S}})$$

and to what we will call the torsion bundle of $\mathcal{S}$, denoted $\mathcal{T}_{\mathcal{S}}$, with

$$\mathcal{T}_{\mathcal{S},x} := \mathcal{T}(\mathfrak{g}_{\mathcal{S},x}). \tag{4.27}$$

Note that, at the level of sections, ∂ becomes a map $\partial : \Omega^1(M, \mathfrak{g}_{\mathcal{S}}) \to \Omega^2(M, TM)$, and its kernel and cokernel are precisely the spaces of sections of $\mathfrak{g}_{\mathcal{S}}^{(1)}$ and $\mathcal{T}_{\mathcal{S}}$, respectively.

Remark 4.13 tells us also that these two vector bundles over M are obtained from the principal G-bundle $\mathcal{S}$ by attaching $\mathfrak{g}^{(1)}$ and $\mathcal{T}(\mathfrak{g})$, viewed as representations of G. For the action of G, one starts with the canonical action of G on $V_0 = \mathbb{R}^n$ and the adjoint action of G on $\mathfrak{g}$, to obtain actions of G on $\mathrm{Hom}(\Lambda^2\mathrm{V_0}, \mathrm{V_0})$ and on $\mathrm{Hom}(\mathrm{V_0}, \mathfrak{g})$. The map ∂ is clearly G-equivariant hence its kernel $\mathfrak{g}^{(1)}$ and its cokernel $\mathcal{T}(\mathfrak{g})$ become representations of G. We deduce:

Proposition 4.28. *For any G-structure $\mathcal{S}$ on M one has*

$$\mathfrak{g}_{\mathcal{S}}^{(1)} \cong E(\mathcal{S}, \mathfrak{g}^{(1)}), \quad \mathcal{T}_{\mathcal{S}} \cong E(\mathcal{S}, \mathcal{T}(\mathfrak{g})).$$

Next, we discuss higher versions of $\mathfrak{g}^{(1)}$ and $\mathcal{T}(\mathfrak{g})$; they will be relevant when discussing the intrinsic curvature and the higher versions of intrinsic torsion. First of all, using $\mathfrak{g}^{(1)} \subset \mathrm{Hom}(\mathrm{V}, \mathfrak{g})$, we have an analogue of the map (4.23),

$$\partial' : \mathrm{Hom}(\mathrm{V}, \mathfrak{g}^{(1)}) \to \mathrm{Hom}(\Lambda^2, \mathfrak{g}), \quad \partial'(\eta)(\mathrm{u}, \mathrm{v}) = \eta(\mathrm{u})\mathrm{v} - \eta(\mathrm{v})\mathrm{u}. \tag{4.29}$$

Define then the second prolongation of $\mathfrak{g}$ as the kernel of this map

$$\mathfrak{g}^{(2)} := \mathrm{Ker}(\partial') \subset \mathrm{Hom}(\mathrm{V}, \mathfrak{g}^{(1)}).$$

The second torsion space is a bit more subtle, as it is not just the cokernel of ∂'. Geometrically, the main reason for this is the Bianchi identity that curvatures satisfy (Lemma 4.46 below). In general, we define

$$\partial' : \mathrm{Hom}(\Lambda^2\mathrm{V}, \mathfrak{g}) \to \mathrm{Hom}(\Lambda^3\mathrm{V}, \mathrm{V}), \tag{4.30}$$

$$\partial(\xi)(u, v, w) = \xi(u, v)w + \xi(v, w)u + \xi(w, u)v.$$

Note that the composition of this map with (4.29),

$$\mathrm{Hom}(\mathrm{V}, \mathfrak{g}^{(1)}) \xrightarrow{\partial'} \mathrm{Hom}(\Lambda^2, \mathfrak{g}) \xrightarrow{\partial'} \mathrm{Hom}(\Lambda^3\mathrm{V}, \mathrm{V}),$$

vanishes.

Definition 4.31. *For a finite dimensional vector space V and a linear subspace $\mathfrak{g} \subset gl(V)$ we define the second torsion space of $\mathfrak{g}$ as*

$$\mathcal{T}^{(1)}(\mathfrak{g}) := \frac{\mathrm{Ker}(\partial' : \mathrm{Hom}(\Lambda^2, \mathfrak{g}) \to \mathrm{Hom}(\Lambda^3\mathrm{V}, \mathrm{V}))}{\mathrm{Im}(\partial' : \mathrm{Hom}(\mathrm{V}, \mathfrak{g}^{(1)}) \to \mathrm{Hom}(\Lambda^2, \mathfrak{g}))}.$$

Again, these constructions will be applied to the fibers of the infinitesimal automorphism bundle $\mathfrak{g}_{\mathcal{S}} \subset \mathrm{End}(\mathrm{TM})$ of a G-structure $\mathcal{S}$ giving rise to the second prolongation bundle

$$\mathfrak{g}_{\mathcal{S}}^{(2)} \subset \mathrm{Hom}(\mathrm{TM}, \mathfrak{g}_{\mathcal{S}}^{(1)})$$

and the second torsion bundle of $\mathcal{S}$, denoted

$$\mathcal{T}_{\mathcal{S}}^{(1)}. \tag{4.32}$$

Remark 4.33 (higher prolongations and higher torsion spaces)**.** Of course, these constructions can be continued by defining inductively

$$\mathrm{Hom}(\mathrm{V}, \mathfrak{g}^{(\mathrm{k})}) \xrightarrow{\partial^{(\mathrm{k})}} \mathrm{Hom}(\Lambda^2\mathrm{V}, \mathfrak{g}^{(\mathrm{k}-1)}) \xrightarrow{\partial^{(\mathrm{k})}} \mathrm{Hom}(\Lambda^3\mathrm{V}, \mathfrak{g}^{(\mathrm{k}-2)}), \tag{4.34}$$

by the same formulas as above so that the $(k+1)$-th prolongation is defined as the kernel of the first map,

$$\mathfrak{g}^{(k+1)} := \mathrm{Ker}(\partial^{(\mathrm{k})}) \subset \mathrm{Hom}(\mathrm{V}, \mathfrak{g}^{(\mathrm{k})}),$$

and the $(k+1)$-th torsion bundle as the cohomology at the middle term $\mathrm{Hom}(\Lambda^2\mathrm{V}, \mathfrak{g}^{(\mathrm{k}-1)})$ of the previous sequence,

$$\mathcal{T}^{(k)}(\mathfrak{g}) := \frac{\mathrm{Ker}(\partial^{(\mathrm{k})})}{\mathrm{Im}(\partial^{(\mathrm{k})})}.$$

Actually, (4.34) can easily be extended to the right to a cochain complex, so that one can talk about cohomology groups in all degrees. The cohomology based at place $\mathrm{Hom}(\Lambda^{\mathrm{l}}\mathrm{V}, \mathfrak{g}^{(\mathrm{k}-1)})$ is denoted $H^{k,l}(\mathfrak{g})$ so that, with this notation,

$$\mathcal{T}^{(k)}(\mathfrak{g}) = H^{k,2}(\mathfrak{g}).$$

4.2.2. The torsion of compatible connections; the intrinsic torsion

This subsection is based on a simple remark: given a G-structure $\mathcal{S}$ on a manifold M, choosing a connection ∇ compatible with $\mathcal{S}$, although its torsion

$$T_{\nabla} \in \Omega^2(M, TM)$$

depends on ∇, it contains "torsion information" about $\mathcal{S}$ itself, independent of the choice of ∇, which gives a measure to the failure of $\mathcal{S}$ from being integrable. With this in mind, the resulting notion of intrinsic notion of $\mathcal{S}$ arises naturally when investigating the change that occurs in the torsion T_∇ when changing ∇. Recall that, for $\mathcal{S}$-compatible connections, ∇ can only be changed by an element $\xi \in \Omega^1(M, \mathfrak{g}_{\mathcal{S}})$ (see Lemma 4.21). The resulting change in the torsion is rather simple and reveals the map ∂ given by (4.23) and its version on sections

$$\partial : \Omega^1(M, \mathfrak{g}_{\mathcal{S}}) \to \Omega^2(M, TM). \tag{4.35}$$

Lemma 4.36. *Let ∇ be a connection compatible with a G-structure $\mathcal{S}$. Then, for $\xi \in \Omega^1(M, \mathfrak{g}_{\mathcal{S}})$, the torsion of ∇^ξ is related to the torsion of ∇ by*

$$T_{\nabla^\xi} = T_\nabla + \partial(\xi).$$

Proof. This is a straightforward computation:

$$T_{\nabla^\xi}(X,Y) = \nabla^\xi_X(Y) - \nabla^\xi_Y(X) - [X,Y] = T_\nabla(X,Y) + \xi(X)(Y) - \xi(Y)(X)$$

where $\xi(X)(Y) - \xi(Y)(X)$ is precisely $\partial(\xi)(X,Y)$ (see (4.23)). □

We see that, in order to force intrinsic information, we have to pass to the torsion bundle $\mathcal{T}_{\mathcal{S}}$ (see (4.27) and (4.26)). Recall that its space of sections is precisely the cokernel of (refpartial-on-sections).

Definition 4.37. *The intrinsic torsion of a G-structure $\mathcal{S}$ on M,*

$$T^{\mathrm{intr}}_{\mathcal{S}} \in \Gamma(\mathcal{T}_{\mathcal{S}}) = \frac{\Omega^2(M, TM)}{\partial(\Omega^1(M, \mathfrak{g}_{\mathcal{S}}))}$$

is the element induced by $T_\nabla \in \Omega^2(M, TM)$, where ∇ is a/any connection compatible with $\mathcal{S}$.

Lemma 4.36 implies that, indeed, $T^{\mathrm{intr}}_{\mathcal{S}}$ does not depend on the choice of the connection. Also the following, although important, is immediate:

Theorem 4.38. *If a G-structure is integrable then $T^{\mathrm{intr}}_{\mathcal{S}} = 0$.*

To appreciate how much information this intrinsic torsion actually contains, we will have to look at some examples; we then very often find that the intrinsic torsion is the only obstruction to integrability (i.e. also the converse of this theorem holds). E.g., as we shall see, this happens for foliations (when the converse is the Frobenius theorem), symplectic structures (when the converse is the Darboux theorem), complex structures (when the converse is the Newlander-Nirenberg theorem).

Finally, we discuss the existence/uniqueness of torsion free connections compatible with a G-structure. This is where also the first prolongation construction (4.25) enters in.

Proposition 4.39. *Given $G \subset GL_n$:*

(i) If $\mathfrak{g}^{(1)} = 0$, then for any G-structure $\mathcal{S}$ on M, M admits a at most one torsion-free connection compatible with $\mathcal{S}$.

(ii) If $\mathcal{T}(\mathfrak{g}) = 0$ then for any G-structure $\mathcal{S}$ on M, M admits a torsion-free connection compatible with $\mathcal{S}$.

Proof. For (i), note actually that the computation from Lemma 4.36 (see also Lemma 4.21) implies that the space of torsion free compatible connections is an affine space with underlying vector space $\Gamma(\mathfrak{g}^{(1)}_{\mathcal{S}})$. For (ii), one starts with any compatible ∇ and the question is whether one finds $\xi \in \Omega^1(M, \mathfrak{g}_{\mathcal{S}})$ so that $\partial(\xi) = T_\nabla$, i.e. whether $[T_\nabla] \in \Gamma(\mathcal{T}_{\mathcal{S}})$ vanishes. □

Note that the previous proof shows a bit more- the equivalence of 1 and 2 below; the local nature of 1 allows us to formulate also condition 3:

Corollary 4.40. *For a G-structure $\mathcal{S}$, the following are equivalent:*

1. *its intrinsic torsion vanishes.*
2. *M admits a torsion free connection compatible with $\mathcal{S}$.*
3. *locally, M admits torsion free connections compatible with $\mathcal{S}$.*

4.2.3. The curvature of compatible connections; the intrinsic curvature

Given an arbitrary connection ∇ on TM we have its curvature

$$K_{\nabla} \in \Omega^2(M, \mathrm{End}(TM)).$$

Of course, if ∇ is compatible with a G-structure $\mathcal{S}$, we expect K_{∇} to live in a smaller space. And that will be indeed the case: it will be a form with values in the infinitesimal automorphism bundle $\mathfrak{g}_{\mathcal{S}} \subset \mathrm{End(TM)}$. This is to be expected because of the reinterpretation of $\mathcal{S}$-compatible connections ∇ as connections on the principal G-bundle $\mathcal{S}$,

$$\omega_{\mathcal{S}} \in \Omega^1(\mathcal{S}, \mathfrak{g}).$$

Indeed, while the curvature of ∇ is of course related to the curvature of the corresponding $\omega_{\mathcal{S}}$, ones has

$$K_{\omega_{\mathcal{S}}} \in \Omega^2(\mathcal{S}, \mathfrak{g}_{\mathcal{S}})_{\mathrm{bas}}.$$

Corollary 4.41. *For any connection ∇ compatible with the G-structure $\mathcal{S}$ on M,*

$$K_{\nabla} \in \Omega^2(M, \mathfrak{g}_{\mathcal{S}})$$

and, modulo the isomorphism

$$\pi^{\bullet} : \Omega(M, aut(\mathcal{S})) \xrightarrow{\sim} \Omega(\mathcal{S}, \mathfrak{g})_{bas}$$

(see Proposition 2.49 and Proposition 4.12), K_{∇} is identified with

$$K_{\omega_{\mathcal{S}}} \in \Omega^2(\mathcal{S}, \mathfrak{g})_{bas},$$

the curvature of the principal bundle connection $\omega_{\mathcal{S}} \in \Omega^1\mathcal{S}, \mathfrak{g})$ corresponding to ∇.

As pointed out in the previous subsection (see Theorem 4.38), the vanishing of the intrinsic torsion can be seen as the first obstruction to integrablity. We now explain that, when this condition is fulfilled, the curvature of compatible connections gives rise to a new intrinsic obstruction. So, we will now restrict to the case when

$$T^{\mathrm{intr}}_{\mathcal{S}} = 0$$

so that we may, an we will, work only with torsion-free connections compatible with $\mathcal{S}$. Similar to the discussion on the intrinsic torsion, when investigating the dependence of K_{∇} on ∇ we discover the map (4.29) and its version on sections:

$$\partial' : \Omega^1(M, \mathfrak{g}^{(1)}_{\mathcal{S}}) \to \Omega^2(M, \mathfrak{g}), \quad \partial'(\eta)(u, v) = \eta(u)v - \eta(v)u. \tag{4.42}$$

The first step in understanding the intrinsic curvature is:

Lemma 4.43. *Given a G-structure $\mathcal{S}$ and two compatible torsion free connections, the difference between their curvatures belongs to the image of ∂.*

Proof. Let us fix first some notations. For a connection ∇ on TM, it induces a natural connection on End(TM) by

$$\nabla_X^{\mathrm{End}}(A)(Y) = \nabla_X(A(Y)) - A(\nabla_X(Y))$$

for A any section of the endomorphism bundle. If the original ∇ is compatible with $\mathcal{S}$ it follows that the ∇^{End} restricts to a connection $\nabla^{\mathfrak{g}}$ on $\mathfrak{g}_{\mathcal{S}}$,

$$\nabla^{\mathfrak{g}} : \mathcal{X}(M) \times \Gamma(\mathfrak{g}_{\mathcal{S}}) \to \Gamma(\mathfrak{g}_{\mathcal{S}}).$$

In turn, this together with ∇ induces a connection ∇' on $\mathrm{Hom}(\mathrm{TM}, \mathfrak{g}_{\mathcal{S}})$

$$\nabla'_X(T')(Y) = \nabla_X^{\mathfrak{g}_{\mathcal{S}}}(T'(Y)) - T'(\nabla_X(Y).$$

Again, this restricts to a connection on $\mathfrak{g}_{\mathcal{S}}^{(1)}$,

$$\nabla^{\mathfrak{g}^{(1)}} : \mathcal{X}(M) \times \Gamma(\mathfrak{g}_{\mathcal{S}}^{(1)}) \to \Gamma(\mathfrak{g}_{\mathcal{S}}^{(1)});$$

To see this, we look at the resulting formula for

$$\nabla_X^{\mathfrak{g}^{(1)}}(\xi) \in \Gamma(\mathrm{Hom}(\mathrm{TM}, \mathfrak{g}_{\mathcal{S}}) \tag{4.44}$$

where $\xi \in \Gamma(\mathfrak{g}_{\mathcal{S}}^{(1)})$. We find

$$\nabla_X^{\mathfrak{g}^{(1)}}(\xi)(Y, Z) = \nabla_X(\xi(Y, Z)) - \xi(Y, \nabla_X(Z)) - \xi(\nabla_X(Y), Z); \tag{4.45}$$

since this is symmetric in Y and Z, it follows that (4.44) belongs indeed to $\Gamma(\mathfrak{g}_{\mathcal{S}}^{(1)})$.

We now return to our result. By Lemma 4.21 and Lemma 4.36, the freedom in changing a torsion free compatible connection ∇ to another one comes from the elements $\xi \in \Omega^1(M, \mathfrak{g}_{\mathcal{S}})$ with $\partial(\xi) = 0$, i.e.

$$\xi \in \Gamma(\mathfrak{g}_{\mathcal{S}}^{(1)}) \subset \Omega^1(M, \mathfrak{g}_{\mathcal{S}}).$$

We fix ∇ and ξ and we analyze the curvature of ∇^{ξ}. For ξ we will use the notation

$$\xi_X := \xi(X) \in \mathfrak{g}_{\mathcal{S}}$$

for $X \in TM$ and, using $\mathfrak{g}_{\mathcal{S}} \subset \mathrm{End}(\mathrm{TM})$,

$$\xi(X, Y) := \xi_X(Y) \in TM$$

for $X, Y \in TM$. The fact that ξ is in the prolongation ensures that the last expressions are symmetric in X and Y.

Compute now the curvature $K_{\nabla^{\xi}}$ in terms of K_{∇}. We find that

$$\begin{gathered}(K_{\nabla^{\xi}} - K_{\nabla})(X, Y)Z = \\ \nabla_X(\xi(Y, Z)) - \nabla_Y(\xi(X, Z)) + \\ +\xi(X, \nabla_Y(Z)) - \xi(Y, \nabla_X(Z)) + \\ +\xi(X, \xi(Y, Z)) - \xi(Y, \xi(X, Z)) - \xi([X, Y], Z).\end{gathered}$$

Combining the first and fourth term of the right hand side with (4.45) and similarly for the second and the third terms, we find

$$\begin{gathered}\nabla_X^{\mathfrak{g}^{(1)}}(\xi)(Y, Z) + \xi(\nabla_X(Y), Z) - \\ -\nabla_Y^{\mathfrak{g}^{(1)}}(\xi)(X, Z) - \xi(\nabla_Y(X), Z) + \\ +\xi(X, \xi(Y, Z)) - \xi(Y, \xi(X, Z)) - \xi([X, Y], Z).\end{gathered}$$

Note that the last terms of the lines cancel because ∇ is torsion free. Therefore

$$(K_{\nabla^{\xi}} - K_{\nabla})(X, Y) = \nabla_X^{\mathfrak{g}^{(1)}}(\xi)(Y) - \nabla_Y^{\mathfrak{g}^{(1)}}(\xi)(X) + [\xi_X, \xi_Y] \in \mathfrak{g}_{\mathcal{S}}.$$

To see that $K_{\nabla^\xi} - K_\nabla$ is in the image of ∂' we look at the two types of terms. The first type (containing the connection) indicate to look at

$$d_{\nabla^{\mathfrak{g}^{(1)}}}(\xi) \in \Omega^1(M, \mathfrak{g}^{(1)}) \quad (\text{sending} \quad \mathrm{X} \mapsto \nabla^{\mathfrak{g}^{(1)}}_{\mathrm{X}}(\xi));$$

computing $\partial' d_{\nabla^{\mathfrak{g}^{(1)}}}(\xi)(X,Y)$ we find indeed the desired terms. To handle the term $[\xi_X, \xi_Y]$ we now consider

$$A : TM \to \mathrm{Hom}(\mathrm{TM}, \mathfrak{g}_{\mathcal{S}}), \quad \mathrm{X} \mapsto \left(\mathrm{Y} \mapsto \mathrm{A(X)(Y)} := \frac{1}{2}\left([\xi_{\mathrm{X}}, \xi_{\mathrm{Y}}] - \xi_{\xi(\mathrm{X},\mathrm{Y})}\right)\right).$$

We first claim that $A(X) \in \mathfrak{g}^{(1)}_{\mathcal{S}}$, i.e. that for X, $A(X)(Y)Z$ is symmetric in Y and Z. This is immediate computing

$$\begin{aligned} 2A(X)(Y)Z &= [\xi_X, \xi_Y](Z) - \xi_{\xi(X,Y)}(Z) = \\ &= \xi(\xi(Y,Z),X) - \xi(\xi(X,Z),Y) - \xi(\xi(X,Y),Z). \end{aligned}$$

Hence $A \in \Omega^1(M, \mathfrak{g}^{(1)}_{\mathcal{S}})$. Computing the corresponding $\partial'(A) \in \Omega^2(M, \mathfrak{g}_{\mathcal{S}})$ we find precisely what we wanted:

$$\partial'(A)(X,Y) = [\xi_X, \xi_Y].$$

In conlcusion,

$$K_{\nabla^\xi} - K_\nabla = \partial'\left(d_{\nabla^{\mathfrak{g}^{(1)}}}(\xi) + A\right).$$

□

For a correct understanding of the space that accommodates the intrinsic curvature we need the so-called Bianchi identity:

Lemma 4.46. *For any connection ∇ on TM, considering the induced differential*

$$d_\nabla : \Omega^\bullet(M, TM) \to \Omega^{\bullet+1}(M, TM),$$

the differential of the torsion, $d_\nabla(T_\nabla) \in \Omega^3(M, TM)$, can be computed in terms of the curvature by:

$$d_\nabla(T_\nabla)(X,Y,Z) = K_\nabla(X,Y)Z + K_\nabla(Y,Z)X + K_\nabla(Z,X)Y.$$

Proof. Straightforward computation. □

This implies that, for torsion free connections, the curvature satisfies further (linear!) conditions: it is killed by

$$\partial' : \Omega^2(M, \mathfrak{g}_{\mathcal{S}}) \to \Omega^3(M, TM),$$

which is (4.30) applied to the (section of the) infintesimal automorphism bundle (therefore explaining why we considered that map). Therefore, the relevant space is the space of sections of the second torsion bundle (4.32),

$$\Gamma(\mathcal{T}^{(1)}_{\mathcal{S}}) = \frac{\mathrm{Ker}(\partial' : \Omega^2(\mathrm{M}, \mathfrak{g}_{\mathcal{S}}) \to \Omega^3(\mathrm{M}, \mathrm{TM}))}{\mathrm{Im}(\partial' : \Omega^1(\mathrm{M}, \mathfrak{g}^{(1)}_{\mathcal{S}}) \to \Omega^2(\mathrm{M}, \mathfrak{g}_{\mathcal{S}}))}.$$

Definition 4.47. *Given a G-structure $\mathcal{S}$ on a manifold M whose intrinsic torsion vanishes, define its intrinsic curvature, also called the intrinsic second torsion,*

$$K^{\mathrm{intr}}_{\mathcal{S}} \in \Gamma(\mathcal{T}^{(1)}_{\mathcal{S}})$$

as the element represented by $K_\nabla \in \Omega^2(M, TM)$, where ∇ is a/any connection compatible with $\mathcal{S}$.

Remark 4.48 (higher torsion)**.** The intrinsic torsion and curvature are just part of a sequence of "higher intrinsic torisons,

$$T^{\text{intr}}_{\mathcal{S}}, K^{\text{intr}}_{\mathcal{S}} = T^{\text{intr},(1)}_{\mathcal{S}}, T^{\text{intr},(2)}_{\mathcal{S}}, \ldots$$

in which each

$$T^{\text{intr},(\text{k})}_{\mathcal{S}} \in \Gamma(\mathcal{T}^{(k)}_{\mathcal{S}})$$

is defined whenever the previous ones vanish. Here $\mathcal{T}^{(k)}_{\mathcal{S}}$ are the higher torsion bundles, whose fibers are the higher torsion spaces of $\mathfrak{g}_{\mathcal{S},x}$ (see Remark 4.33).

With these, a G-structure is called formally integrable if all its higher torsions vanish. Of course, this condition is necessary for integrability, and one of most interesting questions in the theory of G-structures is whether formal integrability implies integrability. Such a statement is not true in full generality, but it is true for large classes of groups G and for many of the geometries that one encounters.

4.2.4. Example: Riemannian metrics

Let us first look at Riemannian metrics interpreted as $G = O(n)$-structures (Example 3.2.8).

Lemma 4.49. *$o(n)^{(1)} = 0$ and $\mathcal{T}(o(n)) = 0$.*

Proof. For the first part let $\xi \in o(n)^{(1)}$, i.e. $\xi : V \longrightarrow o(n)$ ($V = \mathbb{R}^n$) so that $\xi(u)(v) = \xi(v)(u)$ for all $u, v \in V$. The fact that ξ takes values in $o(n)$ can be expressed using the standard metric $\langle \cdot, \cdot \rangle$ on V by saying:

$$\langle \xi(u)(v), w \rangle + \langle v, \xi(u)(w) \rangle = 0 \tag{4.50}$$

for all u, v, w. Apply now repeatedly these formulas to compute $\langle \xi(u)(v), w \rangle$ and we obtain

$$-\langle v, \xi(u)(w) \rangle = -\langle v, \xi(w)(u) \rangle = \langle \xi(w)(v), u \rangle = \langle \xi(v)(w), u \rangle = -\langle w, \xi(v)(u) \rangle$$

hence $-\langle w, \xi(u)(v) \rangle$, i.e., by the symmetry of $\langle \cdot, \cdot \rangle$, we obtain

$$\langle \xi(u)(v), w \rangle = -\langle \xi(u)(v), w \rangle$$

for all u, v, w hence $\xi = 0$. For the second part, let $\phi \in \text{Hom}(\Lambda^2 V, V)$ arbitrary. We look for $\xi : V \longrightarrow o(n)$ so that

$$\phi(u, v) = \xi(u)(v) - \xi(v)(u)$$

for all u and v. The fact that ξ takes values in $o(n)$ means that ξ should satisfy (4.50). For $u \in V$, consider $\phi_u(\cdot) = \phi(u, \cdot) : V \longrightarrow V$, denote by ϕ^*_u its adjoint, and then one immediately checks that

$$\xi(u)(v) := \frac{1}{2}(\phi(u, v) - \phi^*_u(v) - \phi^*_v(u)$$

satisfies both equations (how did we “guess” this? do you see any resemblance with the proof of the existence/uniqueness of the Levi-Civita connection?). □

Note that the previous lemma, together with Proposition 4.39, give a more conceptual explanation for the existence and uniqueness of the Levi-Civita connection associated to a Riemannian metric.

The previous lemma also shows that the intrinsic torsion does not bring anything new in this case: it is automatically zero- therefore it provides us with no obstruction to integrability. The intrinsic curvature however is basically just the curvature of the Levi-Civita connection (and higher torsions, in the sense of Remark 4.48 will automatically be zero, simply because the spaces that they belong to vanish- and this follows immediately from the previous lemma). Note that this fits well with the fact that we have already mentioned, that the intergarbility of g is equivalent to the fact that the curvature of the Levi-Civita connection vanishes.

4.2.5. Example of the intrinsic torsion: symplectic forms

Let us now pass to (almost) symplectic structures; hence we assume that $n = 2k$, $G = Sp_k$. To keep the notations coordinate free, we work in the general context of a vector space V endowed with a linear symplectic form Ω, and the associated Lie algebra

$$\mathrm{gl}(V,\Omega) = \{A \in \mathrm{gl}(V) : \Omega(A(u), v) + \Omega(u, A(v)) = 0, \ \forall\, u, v \in V\}$$

Lemma 4.51. *For the symplectic Lie algebra* $\mathrm{gl}(V,\Omega)$, $\mathrm{gl}(V,\Omega)^{(1)}$ *is isomorphic to* S^3V^* *and one has an isomorphism*

$$\mathcal{T}(\mathrm{gl}(V,\Omega)) \cong \Lambda^3V^*$$

by the isomorphism which associates to the class of $\phi \in Hom(\Lambda^2V, V)$ *modulo the image of* ∂ *the bilinear map* $\partial_\Omega(\phi)$ *given by*

$$\partial_\Omega(\phi)(u,v,w) = \Omega(\phi(u,v),w) + \Omega(\phi(v,w),u) + \Omega(\phi(w,u),v).$$

Proof. We have an exact sequence

$$0 \longrightarrow S^3V^* \stackrel{i}{\longrightarrow} \mathrm{Hom}(V, \mathrm{gl}(V,\Omega)) \stackrel{\partial}{\longrightarrow} \mathrm{Hom}(\Lambda^2V, V) \stackrel{\partial_\Omega}{\longrightarrow} \Lambda^3V^* \longrightarrow 0$$

where ∂_Ω is as in the statement and where, for $p \in S^3V^*$, $i(p)$ is given as follows: $i(p)(u)(v) \in V$ is the unique element in V with the property that

$$\Omega(i(p)(u)(v), w) = p(u,v,w)$$

for all $w \in V$ (which exists since Ω is non-degenerate). □

Corollary 4.52. *Given the presymplectic form* Ω *on* M *interpreted as a* Sp_k*-structure (see subsection 3.2.8), its intrinsic torsion (see Definition 4.37) can be identified with* $d\Omega \in \Omega^3(M)$.

Proof. The isomorphism from the previous proposition gives an identification:

$$\mathcal{T}_{\mathcal{S},x} = \mathcal{T}(\mathrm{gl}(T_xM, \Omega_x)) \stackrel{\sim}{\longrightarrow} \Lambda^3T^*_xM$$

hence the intrinsic torsion can be seen as

$$T^{\mathrm{intr}}_{\mathcal{S}} \in \Gamma(\mathcal{T}_{\mathcal{S}}) \cong \Omega^3(M)$$

where the last identification is given by applying fiberwise ∂_{Ω_x}. To compute it, we need to use a compatible connection ∇, consider its torsion $T(X,Y) = \nabla_X(Y) - \nabla_Y(X) - [X,Y]$ and then apply ∂_Ω to it. We find the 3-form given by

$$\partial_\Omega(T)(X,Y,Z) = \Omega(\nabla_X(Y) - \nabla_Y(X) - [X,Y], Z) +$$
$$+\Omega(\nabla_Y(Z) - \nabla_Z(Y) - [Y,Z], X) + \Omega(\nabla_Z(X) - \nabla_X(Z) - [Z,X], Y)$$

and then, after using the compatibility of ∇ with Ω (see Example 4.5, equation (4.6)) to eliminate ∇ we find precisely the Koszul formula for the differential of Ω applied to (X,Y,Z). □

Corollary 4.53. *The converse of Theorem 4.38 for almost symplectic structures holds true and is the Darboux theorem (Theorem 3.54): Ω is integrable if and only if it is closed.*

4.2.6. Example of the intrinsic torsion: complex structures

We now look at the case of complex structures, when $G = GL_k(\mathbb{C})$, $n = 2k$. Again, to keep the notations coordinate free, we work with a vector space V endowed with a linear complex structure $J : V \to V$, and we consider the associated Lie algebra

$$\mathrm{gl}(V, J) = \{A \in \mathrm{gl}(V) : AJ = JA\}.$$

To compute its torsion space we introduce the space

$$\mathrm{Hom}_J(\Lambda^2 V, V) := \{\phi : \Lambda^2 V \longrightarrow V : \phi(Ju, v) = \phi(u, Jv) = -J\phi(u, v)\}.$$

Lemma 4.54. *One has*

$$\mathcal{T}(\mathrm{gl}(V, J)) \cong Hom_J(\Lambda^2 V, V)$$

by the isomorphism which associates to the class of $\phi \in Hom(\Lambda^2 V, V)$ modulo the image of ∂ the bilinear map $t_J(\phi)$ given by

$$t_J(\phi)(u, v) = \phi(u, v) + J(\phi(Ju, v) + \phi(u, Jv)) - \phi(Ju, Jv).$$

Proof. We claim there is an exact sequence

$$\mathrm{Hom}(V, gl(V, J)) \xrightarrow{\partial} \mathrm{Hom}(\Lambda^2 V, V) \xrightarrow{t_J} \mathrm{Hom}(\Lambda^2 V, V) \xrightarrow{\partial_J} \mathrm{Hom}(\Lambda^2 V, V)$$

where ∂_J is given by

$$\partial_J(\phi)(u, v) = \phi(Ju, v) + \phi(u, Jv) + 2J\phi(u, v)$$

and has as kernel precisely the space $\mathrm{Hom}_J(\Lambda^2 V, V)$ from the statement.

First, it is straightforward to check that $t_J \circ \partial = 0$; in particular, $\mathrm{Im}(\partial) \subset \mathrm{Ker}(t_J)$. Let us prove the converse inclusion. Hence assume that ϕ is in the kernel of t_J, with $\phi : \Lambda^2 V \longrightarrow V$. We are looking for $\xi : V \longrightarrow gl(V, J)$ such that $\partial(\xi) = \phi$. We view ξ as a function of two variable $\xi(u, v) = \xi(u)(v)$, so that the fact that ξ takes values in $gl(V, J)$ is equivalent to

$$\xi(u, Jv) = J\xi(u, v).$$

The condition $\partial(\xi) = \phi$ is then written as

$$\phi(u, v) = \xi(u, v) - \xi(v, u).$$

We look for solutions ξ (of these two equations) of type

$$\xi(u, v) = a\phi(u, v) + bJ\phi(Ju, v) + cJ\phi(u, Jv),$$

with $a, b, c \in \mathbb{R}$ to be determined. The second equation give the condition $a = \imath 2$, $b + c = 0$. For the first equation, after we write it down and we use $t_J(\phi)$ to eliminate the expressions of type $\phi(Ju, Jv)$, we find the condition $b = a + c$. Alltogether, we find $a = \frac{1}{2}$, $b = \frac{1}{4}$, $c = -\frac{1}{4}$.

The next level is similar: first, $\partial_J \circ t_J = 0$ is straightforward; this implies $\mathrm{Im}(t_J) \subset \mathrm{Ker}(\partial_J)$. To prove the other inclusion, we first show that $\mathrm{Ker}(\partial_J) = \mathrm{Hom}_J(\Lambda^2 V, V)$. On one hand, if ϕ is in the right hand side, then it is immediate to check that it is also in the kernel of ∂_J. On the other hand, if ϕ is in the kernel, we have

$$\phi(Ju, v) + \phi(u, Jv) + 2J\phi(u, v) = 0$$

for all u and v. Replacing u by Ju (and keeping v), and then v by Jv (and keeping u), and subtracting the resulting two equations, we find that $\phi(u, Jv) = \phi(Ju, v)$. Combined with the original equation, we find that $\phi \in \mathrm{Hom}_J(\Lambda^2 V, V)$. Finally, note that this also shows that $\phi = t_J(\frac{1}{2}\phi) \in \mathrm{Im}(t_J)$. □

Note that, as in the symplectic case, this lemma implies that, for almost complex structures, the intrinsic torsion can be seen as an element

$$T_{\mathcal{S}}^{\mathrm{intr}} \in \Omega^2_J(M, TM) = \{\omega \in \Omega^2(M, TM) : \omega(JX, Y) = \omega(X, JY) = -J\omega(X, Y)\}$$

which can be computed by choosing a compatible connection ∇ and applying the map t_J. A straightforward computations gives:

Corollary 4.55. *Given the almost complex structure J on M interpreted as a GL_k-structure (see subsection 3.2.7), its intrinsic torsion (see Definition 4.37) can be identified, up to a minus sign, with the Nijenhuis torsion of J, $\mathcal{N}_J \in \Omega^2(M, TM)$ (see Definition 3.48).*

Hence, again,

Corollary 4.56. *The converse of Theorem 4.38 for almost complex structures holds true and is the Nirenberg-Newlander theorem (Theorem 3.49): J is integrable if and only if its Nijenhuis torsion vanishes.*

4.2.7. Example of the intrinsic torsion: foliations

The story for foliations is completely similar to the one for symplectic structures and complex structures, with the conclusion that also in this case the converse of Theorem 4.38 does hold- and it is the Frobenius theorem. Hence in this case one starts with a pair (V, W) with W a p-dimensional subspace of V and the associated Lie algebra

$$\mathrm{gl}(V, W) = \{T \in \mathrm{gl}(V) : T(W) \subset W\}.$$

Lemma 4.57. *One has*

$$\mathcal{T}(\mathrm{gl}(V, W)) \cong \mathit{Hom}(\Lambda^2 W, V/W)$$

by the isomorphism which associates to the class of $\phi \in \mathit{Hom}(\Lambda^2 V, V)$ modulo the image of ∂ the element $(\phi|_W \bmod \mathrm{W}) \in \mathit{Hom}(\Lambda^2\mathrm{W}, \mathrm{V}/\mathrm{W})$.

Proof. Simple exercise. □

Passing to manifolds M, hence to p-dimensional distributions $\mathcal{F} \subset TM$, to relate the resulting intrinsic torsion with the involutivity of $\mathcal{F}$ it is useful to remark that the involutivity itself is controlled by a tensor. More precisely, in complete analogy with the compatibility tensor, one consider the involutivity tensor that arises by looking at expressions of type

$$[V, W] \bmod \Gamma(\mathcal{F}) \in \Gamma(\mathrm{TM}/\mathcal{F})$$

with $V, W \in \Gamma(\mathcal{F})$. This defines a tensor

$$\iota_{\mathcal{F}} \in \Gamma(\mathrm{Hom}(\Lambda^2\mathcal{F}, \Gamma(\mathrm{TM}/\mathcal{F})))$$

and, by the very construction, the involutivity of $\mathcal{F}$ is equivalent to $\iota_{\mathcal{F}} = 0$. Again, a simple computation implies:

Corollary 4.58. *Given a p-dimensional distribution $\mathcal{F}$ on M interpreted as a $GL(p, n-p)$-structure (see subsection 3.2.5), its intrinsic torsion (see Definition 4.37) can be identified with the involutivity tensor of $\mathcal{F}$.*

Corollary 4.59. *The converse of Theorem 4.38 for p-dimensional distributions holds true and is the Frobenius theorem (Theorem thm:Frobenius): $\mathcal{F}$ is integrable if and only if it is involutive.*

4.3. Torsion and curvature via structure equations

4.3.1. The tautological form

The notion of torsion does not make sense on arbitrary vector bundles or on arbitrary principal bundles. In particular, for connections compatible with a G-structure, unlike the case of the curvature which was discussed e.g. in the first part of subsection 4.2.3, we had no similar discussion for the torsion. We would like to clarify what is the structure present on Fr(TM) (or on a general G-structure $\mathcal{S}$, interpreted as a principal bundle) that allows us to talk about the torsion of principal bundle connections. The resulting structure (the tautological form described below) turns out to be extremely useful also for other purposes. As before, $G \subset GL_n$ will be a fixed Lie subgroup.

Definition 4.60. *The tautological form of a G-structure $\mathcal{S}$ on M is the 1-form*

$$\theta_{\mathcal{S}} \in \Omega^1(\mathcal{S}, \mathbb{R}^n)$$

which takes a tangent vector $X_u \in T_u\mathcal{S}$ ($u \in \mathcal{S}$), projects it to $(d\pi)_u(X_u) \in T_xM$, and then uses the interpretation of the frame u as an isomorphism $u : \mathbb{R}^n \longrightarrow T_xM$ to move back to $\mathbb{R}^n$:

$$\theta_{\mathcal{S}}(X_u) := u^{-1}((d\pi)_u(X_u)) \in \mathbb{R}^n.$$

To describe the main (abstract) properties of $\theta_{\mathcal{S}}$ we will use the terminology from Definition 2.48 on equivariant, horizontal and basic forms. Furthermore, we will say that a form $\theta \in \Omega^1(P, V)$ on a principal G-bundle P with coefficients in a vector space V is strictly horizontal if the kernel of each θ_p ($p \in P$) is precisely the space $T_p^{\vee}P$ of vertical vectors at p. It is straightforward to check that the tautological form is G-invariant and strictly horizontal. The main point is that this property, as well as the one from Lemma 4.2, characterize the principal G-bundles that come from G-structures.

Proposition 4.61. *Let $G \subset GL_n$ be a closed subgroup. Then any principal G-bundle $\pi : P \longrightarrow M$ together with a G-invariant, strictly horizontal 1-form $\theta \in \Omega^1(P, \mathbb{R}^n)$ must be isomorphic to $(\mathcal{S}, \theta_{\mathcal{S}})$ for some G-structure $\mathcal{S}$ on M. More precisely:*

(i) There exists a unique smooth map $\Phi : P \to \mathrm{Fr(TM)}$ covering the identity and such that $\theta = \Phi^*\theta_{\mathrm{Fr(TM)}}$.
(ii) Φ is actually an embedding whose image $\mathcal{S}$ is a G-structure on M.

The main message of the previous proposition is that, if we want to say something about G-structure that depends on more than just the principal bundle structure, then one has to use the tautological form $\theta_{\mathcal{S}}$ which, together with its properties, fully encodes the situation. Hence it is not surprising that the tautological forms can also be used to tackle the equivalence problem for G-structures: deciding when two G-structures are (locally) isomorphic.

Proposition 4.62. *Let $G \subset GL_n$ be a closed, connected, subgroup and assume that $\mathcal{S}_i$ is a G-structure on M_i, for $i \in \{1, 2\}$. Then the two G-structures are isomorphic if and only if there exists a diffeomorphism*

$$\Psi : \mathcal{S}_1 \longrightarrow \mathcal{S}_2$$

with the property that

$$\Psi^*(\theta_{\mathcal{S}_2}) = \theta_{\mathcal{S}_1}.$$

Proof. To come (this is not completely trivial). □

4.3.2. The torsion revisited

Let us now return to our original motivation for the previous subsection and explain that the tautological form is precisely what allows us to talk about the torsion of a connection ω on the principal G-bundle $\mathcal{S}$ that comes from a G-structure: while the curvature of ω was arising as the horizontal part of $d\omega$ (see (i) of Proposition 2.58), the torsion of ω can now be defined similarly as the horizontal part of $d\theta$. I.e. define

$$T_\omega \in \Omega^2(\mathcal{S}, \mathbb{R}^n)_{\text{bas}}$$

as the unique horizontal form on $\mathcal{S}$ satisfying

$$T_\omega(h(X), h(Y)) = (d\theta)(h(X), h(Y)),$$

where h is the horizontal lift operation associated to ω.

Of course, starting with a connection ∇ on TM which is compatible with $\mathcal{S}$, one expects that, analogous to the case of curvatures, the torsion of ∇

$$T_\nabla \in \Omega^2(M, TM)$$

is basically the same thing as the torsion of the corresponding principal bundle connection $\omega_{\mathcal{S}}$. Of course, this identification will go again via an isomorphism arising from Proposition 2.49 , namely

$$\pi^\bullet : \Omega^2(M, TM) \xrightarrow{\sim} \Omega(\mathcal{S}, \mathbb{R}^n)_{\text{bas}}.$$

The following is an easy exercise which follows by a computation completely similar to that done for the curvature (Proposition 2.58)

Lemma 4.63. *For any connection ∇ compatible with a G-structure $\mathcal{S}$, via the previous isomorphism, its torsion T_∇ coincides with the torsion $T_{\omega_{\mathcal{S}}}$ of the principal bundle connection $\omega_{\mathcal{S}}$ associated to ∇.*

4.3.3. Structure equations; the point of view of coframes

We now further discuss the curvature and torsion of compatible connections by working directly at the level of $\mathcal{S}$; we will show that they arise very naturally, when looking for "structure equations". Therefore, we fix a G-structure $\mathcal{S}$ on M and we also fix a connection compatible with $\mathcal{S}$. The connection will be encoded in the corresponding connection form, denoted simply

$$\omega \in \Omega^1(\mathcal{S}, \mathfrak{g}).$$

We will denote by K_ω and T_ω the curvature and the torsion of ω. Taking advantage of the fact that the points of $\mathcal{S}$ are frames, they can be handled more directly, as functions instead of sections of vector bundles or differential forms. For K_ω, at any $\phi \in \mathcal{S}_x$, due to the horizontality of K_ω can now be seen as a map from $\Lambda^2 T_x^* M$ to

$\mathfrak{g}$; moreover, since ϕ defines an identification of T_xM with V, we find that K_ω can be seen as a smooth map, denoted

$$K \in C^\infty(\mathcal{S}, \mathrm{Hom}(\Lambda^2 V, \mathfrak{g})). \tag{4.64}$$

Similarly, the torsion can be reinterpreted as a smooth map

$$T \in C^\infty(\mathcal{S}, \mathrm{Hom}(\Lambda^2 V, V)). \tag{4.65}$$

Here are the structure equations we mentioned, in the most compact formulation; more explanations/interpretations will come after the statement.

Proposition 4.66. *One has*

$$d\theta = T(\theta \wedge \theta) + \omega \wedge \theta \quad (\textit{in } \Omega^2(\mathcal{S}, V)),$$

$$d\omega = K(\theta \wedge \theta) - \frac{1}{2}[\omega, \omega] \quad (\textit{in } \Omega^2(\mathcal{S}, \mathfrak{g})).$$

Remark 4.67 (the notations)**.** Let us first explain the notations used. They are all particular cases of a general construction: whenever one has three vector spaces W_1, W_2 and W and a bilinear map

$$f : W_1 \times W_2 \longrightarrow W$$

one has an wedge-product operations

$$\Omega^p(\mathcal{S}, W_1) \otimes \Omega^q(\mathcal{S}, W_2) \longrightarrow \Omega^{p+q}(\mathcal{S}, W),\ (\omega, \eta) \mapsto \omega \wedge_f \eta$$

given by the standard formula (1.13), but using f to pair the values of ω with those of η. Most of the times f is clear from the context and one just omits it from the notation. This applies in particular to

- the evaluation operation $\mathfrak{g} \times V \to V$, $(A, v) \mapsto A(v)$ which induces
$$\Omega^1(\mathcal{S}, \mathfrak{g}) \times \Omega^1(\mathcal{S}, V) \longrightarrow \Omega^2(\mathcal{S}, V)$$
and which explains the meaning of $\omega \wedge \theta$ (sometimes also denoted $\omega \circ \theta$).
- the Lie algebra operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$ which induces
$$\Omega^1(\mathcal{S}, \mathfrak{g}) \times \Omega^1(\mathcal{S}, \mathfrak{g}) \longrightarrow \Omega^2(\mathcal{S}, \mathfrak{g}),$$
which is usually denoted still by $[\cdot, \cdot]$ and which explains the term $[\omega, \omega]$.
- the triple operation $\mathrm{Hom}(\Lambda^2 V, V) \times V \times V \to V$, $(\xi, v_1, v_2) \mapsto \xi(v_1, v_2)$, which induces
$$C^\infty(\mathcal{S}, \mathrm{Hom}(\Lambda^2 \mathrm{V}, \mathrm{V})) \times \Omega^1(\mathcal{S}, \mathrm{V}) \times \Omega^1(\mathcal{S}, \mathrm{V}) \longrightarrow \Omega^2(\mathcal{S}, \mathrm{V}).$$
explaining the meaning of $T(\theta \wedge \theta)$.
- similarly for
$$C^\infty(\mathcal{S}, \mathrm{Hom}(\Lambda^2 \mathrm{V}, \mathfrak{g})) \times \Omega^1(\mathcal{S}, \mathrm{V}) \times \Omega^1(\mathcal{S}, \mathrm{V}) \longrightarrow \Omega^2(\mathcal{S}, \mathfrak{g}),$$
explaining the notation $K(\theta \wedge \theta)$.

Remark 4.68 (more insight)**.** Here is some more insight into the previous proposition. The starting point is the remark that ω together with the tautological form θ,

$$\Theta = (\omega, \theta) \in \Omega^1(\mathcal{S}, \mathfrak{g} \oplus V)$$

$(V = \mathbb{R}^n)$, induces at each point $u \in \mathcal{S}$, it gives an isomorphism

$$\Theta_u : T_u\mathcal{S} \xrightarrow{\sim} \mathfrak{g} \oplus V.$$

Hence, Θ defines a parallelism (or an e-structure) on $\mathcal{S}$; to obtain an actual coframe as discussed in subsection 3.2.2 we would have to fix a basis of $\mathfrak{g} \oplus V$, but let us proceed more intrinsically first. The main message of subsection 3.2.2 was that e-structures should be studied via their structure functions, obtained by expressing $d\Theta$ in terms of $\Theta \wedge \Theta$. The outcome, described by the previous proposition, is quite beautiful: the resulting structure functions for our Θ contains many of the objects that we have around: the Lie algebra structure of $\mathfrak{g}$, the representation of $\mathfrak{g}$ on $\mathbb{R}^n$, K and T.

Let us now reformulate the previous discussion in a more explicit way, by fixing bases. We consider

$$\{e_1, \ldots, e_n\} - \text{ the standard basis of } V = \mathbb{R}^n \tag{4.69}$$

and, denoting by N the dimension of G, we also consider and we also consider

$$\{A(1), \ldots, A(N)\} - \text{ a fixed basis of } \mathfrak{g}. \tag{4.70}$$

Hence each $A(\alpha)$ has a matrix representation

$$A(\alpha) = (A^i_j(\alpha))_{1 \leq i,j \leq n} \in \mathfrak{g} \subset gl_n.$$

With this, $\omega \in \Omega^1(\mathcal{S}; \mathfrak{g})$ is described by its component with respect to the basis (4.70), i.e. by 1-forms

$$\omega^\alpha \in \Omega^1(\mathcal{S}), \quad \alpha \in \{1, \ldots, N\},$$

so that

$$\omega = \omega^1 A(1) + \ldots + \omega^N A(N).$$

Similarly, the tautological form $\theta \in \Omega^1(\mathcal{S}, V)$ has component

$$\theta^1, \ldots, \theta^n \in \Omega^1(\mathcal{S})$$

with respect to the basis (4.69)- so that

$$\theta = \theta^1 e_1 + \ldots + \theta^n e_n.$$

The previous remark on Θ translates into:

Proposition 4.71. *The collection of 1-forms*

$$\omega^1, \ldots, \omega^N, \theta^1, \ldots, \theta^n \in \Omega^1(\mathcal{S}) \tag{4.72}$$

is a coframe on $\mathcal{S}$.

Hence we are in the setting of coframes from subsection 3.2.2 and the interesting question is that of computing its structure functions. To make Proposition 4.66 more explicit, we expand the curvature (4.64) with respect to the given basis:

$$K(e_i, e_j) = \sum_\alpha K^\alpha_{i,j} A(\alpha)$$

and similarly for the torsion (4.65):

$$T(e_i, e_j) = \sum_k T^k_{i,j} e_k$$

so that the curvature and the torsion are encoded in their coefficients

$$\{K^\alpha_{i,j} \in C^\infty(\mathcal{S}) : 1 \leq i, j \leq n, 1 \leq \alpha \leq N\}, \ \{T^k_{i,j} \in C^\infty(\mathcal{S}) : 1 \leq i, j, k \leq n\}.$$

To write down explicitly the structure equations for the coframe from the previous proposition we also need the structure constants

$$\{c^{\alpha}_{\beta,\gamma} : 1 \leq \alpha, \beta, \gamma \leq N\}$$

of the Lie algebra $\mathfrak{g}$- defined by

$$[A(\beta), A(\gamma)] = \sum_{\beta,\gamma}^{\alpha} c^{\alpha}_{\beta,\gamma} A(\alpha).$$

Corollary 4.73. *The coframe (4.72) from the previous proposition satisfies the structure equations In terms of their components:*

$$d\theta^i = \sum_{j,k} T^i_{j,k} \theta^j \wedge \theta^k + \sum A^i_j(\alpha) \omega^{\alpha} \wedge \theta^j,$$

$$d\omega^{\alpha} = \sum_{j,k} K^{\alpha}_{j,k} \theta^j \wedge \theta^k - \sum_{\beta,\gamma} c^{\alpha}_{\beta,\gamma} \omega^{\beta} \wedge \omega^{\gamma}.$$

4.3.4. An integrability result for G-structures

Although the converse of Theorem 4.38 does not hold in general, let us at least prove a weaker version of it.

Theorem 4.74. *A G-structure $\mathcal{S}$ is integrable if and only if, locally, M admits flat torsion free connections compatible with $\mathcal{S}$.*

Proof. Since this is a local statement, we may assume that $M = \mathbb{R}^n$, we are looking around the point $0 \in \mathbb{R}^n$ and we have a connection as in the statement defined over the entire M. Note that the structure equations simplify to

$$d\theta = \omega \wedge \theta \quad (\text{in } \Omega^2(\mathcal{S}, V),$$

$$d\omega = -\frac{1}{2}[\omega, \omega] \quad (\text{in } \Omega^2(\mathcal{S}, \mathfrak{g}).$$

In components, we deal with structure equations with constant coefficients. These coefficients can be recognized as the structure constants of a larger Lie algebra; but let us proceed intrinsically. We consider the new Lie algebra (called the semi-direct product of the Lie algebra $\mathfrak{g}$ with its representation V),

$$\tilde{\mathfrak{g}} = \mathfrak{g} \oplus V$$

in which the Lie bracket is given by

$$[(A, v), (B, w)] = ([A, B], A(w) - B(v)).$$

Strictly speaking (because of the way we talked about Lie groups and their Lie algebras), we have to convince ourselves that $\tilde{g}$ comes from a Lie group $\tilde{G} \subset GL_{n+1}$. Although that is not really needed below, let us sketch the argument. For that we consider $\tilde{G} = G \times V$ with the groups structure

$$(g, v) \cdot (h, w) = (gh, h^{-1}(v) + w).$$

This is a subgroup of $GL(W)$ where $W = V^* \oplus \mathbb{R}$ (which is isomorphic to $\mathbb{R}^{n+1}$), where we see $(g, v) \in \tilde{G}$ acting on W by

$$(\xi, \lambda) \mapsto (\xi \circ g^{-1}, \lambda + \xi(v)).$$

You can check that the Lie algebra of $\tilde{G}$ is isomorphic to$\tilde{g}$.

Back to our connection, we put the forms θ, ω together into a 1-form

$$\Omega = (\omega, \theta) \in \Omega^1(\mathcal{S}, \tilde{\mathfrak{g}})$$

and then we immediately see that the structure equations can be re-written as

$$d\Omega + \frac{1}{2}[\Omega, \Omega] = 0.$$

We now need a general results about Lie groups and their Maurer-Cartan forms. Let $\tilde{G} \subset GL_N$ be a Lie group with Lie algebra denoted $\tilde{\mathfrak{g}}$. Then the (right-invariant) Maurer-Cartan form of $\tilde{G}$ is a 1-form

$$\omega_{\tilde{G}} \in \Omega^1(\tilde{G}, \tilde{\mathfrak{g}}).$$

One way to introduce it is by considering the map $\pi : \tilde{G} \longrightarrow pt$ into a one-point space and viewing it as a principal $\tilde{G}$-bundle (where $\tilde{G}$ acts on $P = \tilde{G}$ from the right by right translations). As such, $\tilde{G}$ admits a unique connection 1-form- and that is what $\omega_{\tilde{G}}$ is. Note that the associated infinitesimal action of $\tilde{\mathfrak{g}}$ on $\tilde{G}$ is given by

$$a_g : \tilde{\mathfrak{g}} \longrightarrow T_g\tilde{G}, \quad X \mapsto \frac{d}{dt}|_{t=0} g \cdot \exp(tX),$$

or, equivalently, using the left translations $L_g : \tilde{G} \longrightarrow \tilde{G}$, $h \mapsto hg$, as well as the identification of $\tilde{\mathfrak{g}}$ with $T_I\tilde{G}$, we see that

$$a_g = (dL_g)_I : \tilde{\mathfrak{g}} \longrightarrow T_g\tilde{G},$$

so that the formula for $\omega_{\tilde{G}}$ is

$$\omega_{\tilde{G}}(X_g) = (dL_{g^{-1}})_g(X_g) \ \ \forall \ X_g \in T_g\omega_{\tilde{G}}.$$

Moreover, since as a connection this is flat (all the 2-forms on the base are zero!), the structure equations do tell us that, indeed,

$$d\omega_{\tilde{G}} + \frac{1}{2}[\omega_{\tilde{G}}, \omega_{\tilde{G}}] = 0.$$

The result that we need is that $\omega_{\tilde{G}}$ is universal with this property. More precisely:

Lemma 4.75. *For any Lie group $\tilde{G}$ with Lie algebra $\tilde{\mathfrak{g}}$ and any 1-form $\Omega \in \Omega^1(M, \tilde{\mathfrak{g}})$ defined on a simply connected manifold N and satisfying*

$$d\Omega + \frac{1}{2}[\Omega, \Omega] = 0,$$

there exists a smooth map $F : N \longrightarrow \tilde{G}$ such that $\Omega = F^*(\omega_{\tilde{G}})$.

Proof. (of the lemma)(sketch) In the same way as with the bundle $\tilde{G} \longrightarrow pt$, one looks at how flat connections on the principal $\tilde{G}$-bundle $P := N \times \tilde{G} \longrightarrow N$ look like, and we see that, while they are uniquely determined by the connection condition on the vertical vector fields and by a form Ω as in the statement on the horizontal ones. Hence Ω induces a flat connection on P, and then an involutive distribution (the horizontal one) on P. By Frobenius theorem, the distribution is integrable and then we can take a leaf $L \subset N \times \tilde{G}$ of it. This leaf will play the role of the graph of the function we are looking for. More precisely, using that the differential of the projection $P \longrightarrow N$ maps the distribution isomorphically onto TN, one shows that its restriction $\mathrm{pr}_1 : L \longrightarrow N$ is a covering projection; since N is simply connected, it must be a diffeomorphism, and then the desired map will be the composition of the inverse of this map with the second projection into G. $\square$

Returning to the proof of our theorem one finds

$$F : \mathcal{S} \longrightarrow \tilde{G} = G \times V$$

so that Ω is the pull-back of the Maurer-Cartan form of $\tilde{G}$ (here we use that $M = \mathbb{R}^n$ is simply connected, but you should find the argument that shows that F exists on $\mathcal{S}$ in general, i.e. even when $\mathcal{S}$, or equivalently $\tilde{G}$, is not simply connected). It is not difficult to see that F is a local diffeomorphism which is G-equivariant and then, passing to the quotient modulo G, we obtain a local diffeomorphism

$$f : M \longrightarrow V$$

compatible with the G-structures. Locally, one can restrict to diffeomorphisms, proving integrability. $\square$