

Faculté des sciences

Symplectic structures preserved by geodesic symmetries

Auteur·es : Maxime Willaert

Promoteur·rices : Pierre Bieliavsky

Lecteur·rices : Heiner Olbermann, Pierre-Emmanuel Caprace

Année académique 2023-2024

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
FACULTÉ DES SCIENCES
ÉCOLE DE MATHÉMATIQUES

Symplectic structures preserved by geodesic symmetries

Maxime Willaert

Advisor: Pierre Bieliavsky

Readers: Heiner Olbermann, Pierre-Emmanuel Caprace

Master thesis
September 2024

Remerciements

Merci Pierre pour ta patience, ta bienveillance, ton indéfectible bonne humeur et ta passion si contagieuse.

Merci aussi aux Professeurs Olbermann et Caprace, qui ont accepté de lire ce mémoire (mes excuses, surtout, pour avoir perdu contrôle de sa longueur).

Merci Logan pour ces sessions bibli's mémorables. Ton humour restera plus mystérieux que toute théorie mathématique et me procurera au moins autant de plaisir.

Merci à toi aussi Téo, grâce à toi la solitude ne s'est pas ajoutée à la charge que représentaient ces dernières semaines de rédaction.

Merci à mon père et ma soeur, modèles de gentillesse et d'écoute, dont l'intérêt et l'enthousiasme n'ont pas été découragés par la difficulté d'accès de mon travail.

Merci à ma mère, pour sa présence toujours aussi structurante.

Merci à toi aussi Maud, pour tes précieux conseils, et à vous mes petits frères, Pierre et Mathias, pour votre esprit toujours aussi vif et inquisiteur.

Et à toutes les personnes qui m'ont accompagné durant ces deux années d'études ainsi que celles qui ont précédé, mais que les limites d'espace et de bienséance ne m'ont pas permis de citer, je vous remercie.

Introduction

If we had to summarize this thesis in one sentence, it would be “an exploration of the interplay between certain geometric structures”. Chief among these structures are affine connections, symplectic and complex structures, as well as Riemannian metrics. The interactions between these geometries are numerous and subtle, which makes their study highly rewarding, but complicates the problem of finding a good starting point as one can easily get overwhelmed by the terminology used to describe these many configurations: “Kähler manifolds, complex affine locally symmetric spaces, preferred symplectic connections, . . .”

Perhaps a good place to start would be Riemannian symmetric spaces. Elie Cartan achieved something akin to legendary status by providing a classification of all simple real Lie algebras in 1914, only to prove twelve years later that his work on said algebras could be used to give a complete classification of irreducible (globally) symmetric Riemannian spaces (of course this does not even begin to exhaust his many accomplishments). This naturally generated great interest for symmetric Riemannian spaces, and today they appear in fields as diverse as harmonic analysis and representation theory. A definition of these spaces could be sketched as follows: around any point p of a Riemannian manifold, one can define an involution of a neighborhood of p (for which p is an isolated fixed point) by ‘inverting’ the geodesics emanating from p . We call this involution the *local geodesic symmetry* centered at p , or the global symmetry if it is defined on the entire manifold. A Riemannian manifold is then said to be *locally symmetric* if all of its geodesic symmetries are isometric (despite this long-winded definition, these spaces admit a very simple characterization in terms of their curvature), or *globally symmetric* if their geodesic symmetries are globally defined in addition to being isometric.

As is known by any geometer or practitioner of general relativity, the geodesics (and curvature) of a (pseudo-)Riemannian manifold are defined by means of the unique affine connection (or covariant derivative) induced by the metric: the Levi-Civita connection. As it turns out, an affine connection is the only thing one really needs for all of this to ‘work’, and this entire discussion carries over to any affinely connected manifold (even the ones that do not arise from a metric), leading to the notion of a (locally) affine symmetric space. In their simplest forms, locally symmetric spaces are thus manifolds with affine connections whose geodesic symmetries are affine transformations (a classification of such spaces has yet to be achieved, in contrast to the Riemannian setting). Affine symmetric spaces can then be ‘flavored’ by adding new geometric structures, and requiring that these structures be compatible with the local geodesic symmetries. Riemann symmetric

spaces are obtained precisely through that process, but this scheme can be applied to a wide range of structures. Most importantly for our purposes, this scheme can be applied to symplectic structures, resulting in the (locally) symmetric symplectic spaces that Pierre Bieliavsky introduced in his PhD thesis [4].

The study of symplectic forms began with Hamiltonian mechanics, indeed the phase space of a mechanical system (as modeled in Hamiltonian mechanics) carries a canonical symplectic form, and this symplectic form is the ‘hidden structure’ behind the renowned Poisson brackets. One might glance at the definition of a symplectic form and see nothing but a ‘skew-symmetric metric’, and there do exist certain analogies between Riemannian and symplectic geometry, but in truth their behaviors vary widely (symplectic structures are in fact closer to complex structures, as will be made apparent later). Most notably, symplectic geometry is *non-local*, as encapsulated by Darboux’s theorem which states that any two symplectic manifolds of equal dimension are locally isomorphic. In particular symplectic geometry only admits *global* invariants (there cannot exist *local* symplectic invariants) which is the reason why many speak of symplectic *topology* and not of symplectic *geometry*. This is to be confronted to the Riemann curvature tensor of Riemannian geometry, the prototypical local invariant.

This non-locality of symplectic structures manifests in a striking manner when we introduce affine connections. In stark contrast with the unique Levi-Civita connection of Riemannian geometry, the set of symplectic connections on a given symplectic manifold forms an ∞ -dimensional affine space. Affine connections compatible with the symplectic structure, so called *symplectic connections*, constitute a fairly natural object of investigation to the differential geometer interested in symplectic geometry. Their study is further motivated by the role they play in quantization, the process through which one obtains the quantum mechanical description of a physical system from its classical description, a classical description that will take the form of a symplectic manifold in our case (see [1] for an introduction to the Fedosov construction in formal quantization). In any case, regardless of the path that might lead one to symplectic connections, when faced their overabundance a natural reflex is to try and isolate the most interesting ones through judicious criteria. We will encounter several examples of such criteria in the coming pages, namely:

- The already mentioned (locally) symmetric symplectic connections;
- Preferred connections, which Bourgeois and Cahen obtained by means of a variational principle in [6];
- Kähler connections, which arise when there exist a metric and a complex structure that are highly compatible with the symplectic structure;
- And lastly, the original notion of symplectic connections of type S, which we define to be connections whose geodesic symmetries preserve the symplectic form (but are not required to be affine transformations).

Our interest in type S connections was sparked by a theorem of Sekigawa and Vanhecke [2] stating that Kähler manifolds (manifolds equipped with mutually compatible Riemannian, complex and symplectic structures) whose local geodesic symmetries are symplectic (meaning that they preserve the symplectic form) are locally symmetric, i.e. all three of their structures are preserved by the local symmetries. This rather surprising result, given the loose relationship between a symplectic form and its symplectic connections, lead us to the following question: are type S connections necessarily locally symmetric?

We answered this question in the negative by providing a family of counter-examples (extracted from the article of Bourgeois and Cahen [6]), and in the process obtained this sequence of strict inclusions:

$$\begin{array}{c}
\text{Locally symmetric symplectic spaces} \\
\cap \\
\text{Symplectic connections of type S} \\
\cap \\
\text{Preferred symplectic connections}
\end{array}$$

The main challenge we faced while trying to construct these counter-example was the lack of manageable sufficient conditions for a symplectic connection to be of type S. A solution to this problem was found in the work of D'Atri and Nickerson [8] on Riemannian manifolds with divergence-preserving symmetries (so-called *D'Atri spaces*). In [8] D'Atri and Nickerson describe an infinite sequence of curvature condition, called the Ledger conditions (introduced by Ledger in his PhD thesis on *harmonic spaces* [16]), which are defined through a recurrence formula. In the case of a symplectic surface (in other words, a 2-dimensional symplectic manifold) with an *analytic* symplectic connection, the Ledger conditions of odd order turn out to be sufficient for our connection to be of type S.

Building on the work of Cahen and Bourgeois, we then discovered that a preferred connection over a symplectic surface must always satisfy the odd Ledger conditions. Putting these two results together, we obtained the counter-examples we sought in the form of the analytic, non-locally-symmetric, preferred symplectic connections over $\mathbb{R}^2$ described in [6].

Let us also mention that the odd Ledger conditions are necessarily satisfied by any torsionless connection whose geodesic symmetries preserve a volume form (something we will prove in the penultimate chapter). This applies in particular to connections of type S, given the existence of the symplectic volume form. The first of these odd Ledger conditions, denoted by (L3), is precisely the condition used by Bourgeois and Cahen to define their preferred symplectic connections. Therefore a symplectic connection of type

S is necessarily preferred (this realization is what lead us to the work of Bourgeois and Cahen in the first place). We show that this inclusion is also a strict one by confronting the Sekigawa-Vanhecke theorem to the existence of non-locally-symmetric Ricci-parallel Kähler manifolds.

Let us now describe the layout of this thesis.

The purpose of chapter 1 is to recall the basic properties of the tensor algebra of a finite-dimensional vector space, notions that will be used systematically throughout this entire thesis. We begin by introducing some of the equivalent point of views one can adopt regarding tensors, and define the tensor product and contraction operation (which can be combined to produce almost all the operations of linear algebra). We then describe the process through which isomorphisms and endomorphisms of vector spaces induce isomorphism and endomorphisms of their tensor algebras. We close this chapter with a reminder on polarization, a process that turns homogeneous polynomials into multilinear symmetric maps by adding variables, and which will be of use to us in the last chapters as the Ledger conditions take the form of homogeneous polynomial equations.

In chapter 2 we invest a sizeable amount of time into the development of the general theory of *connections on vector bundles*. Vector bundles are 'smooth' families of vector spaces indexed by the points of a manifold, which are locally isomorphic to direct products (which is why they are sometimes called *twisted products*). We define vector bundles, their sections, and maps between their sections. We then describe operations one can apply to vector bundles, most notably: the bundle equivalent of the tensor algebra, and the pullback of a vector bundle along a map of base manifolds.

We then introduce the notion of a connection on a vector bundle, a tool used to create links between the fibers of a bundle. Using the pullback operation, we define the covariant derivative along a curve and parallel transport. We introduce the curvature of a connection, and present its interpretation in terms of infinitesimal holonomy. We give a description of the connections induced on tensor bundles (as well as their curvature). We end the chapter by introducing an alternative point of view on connections, that of *Ehresmann connections* which are defined by introducing a horizontal distribution. We take the opportunity to explain why flat connections are locally trivial.

In chapter 3 we apply our general theory of connections to the case of the tangent bundle, leading us to the notion of an *affine connection*. We describe the local invariants unique to affine connections: torsion and the Ricci curvature, the latter of which we relate to the curvature on the determinant bundle. Most importantly we develop the theory of geodesics (by means of the *geodesic spray*): we introduce the exponential mapping, normal coordinates, Jacobi fields and, finally, *geodesic symmetries*.

This leads us to the notion of *symmetric tensors* ('tensors compatible with the geodesic symmetries'), and through 'radial power expansion in normal coordinates' we obtain our first necessary curvature condition for the existence of a symmetric tensor (this curvature

condition will later be used to prove the Sekigawa-Vanhecke theorem, and the inclusion "type S $\subset$ Preferred").

In chapter 4 we give a brief overview of Riemannian, symplectic and complex structures, their interactions with each other and with affine connections, as well as their combination in the form of Kähler manifolds. The most important result in this chapter are:

- The inclusion "type S $\subset$ Preferred", obtained by applying the condition derived at the end of chapter 3 to the symplectic volume form.
- The symmetries verified by the curvature of a Kähler manifold, which will play a key role in the proof of the Sekigawa-Vanhecke theorem.
- The characterization of locally symmetric spaces in terms of their curvature (and torsion, which must vanish).

In chapter 5 we state and prove the Sekigawa-Vanhecke theorem. At that point, most of the work has already been done, and the proof is rather brief. Relying on the Sekigawa-Vanhecke theorem, we also find examples of preferred symplectic connections which are not of type S in the form of non-locally-symmetric Kähler-Einstein manifolds.

In chapter 6, we introduce an original notion of divergence, which we show to be equivalent to a choice of connection on the determinant bundle. Using this notion of divergence, we extend the work of D'Atri and Nickerson [8] to the affine setting by showing that a torsionless affine connection with divergence-preserving symmetries must necessarily satisfy all odd Ledger conditions, and that these conditions are in fact sufficient when the connection is analytic and locally equiaffine (meaning that there exist local parallel volume forms).

Chapter 7 is our most original contribution. The great simplicity of the geometry of preferred connections on symplectic surfaces, uncovered by Bourgeois and Cahen in [6], allows us to show that such a connection must necessarily verify all odd Ledger conditions.

A family of non-locally-symmetric symplectic connections of type S is then easily found, thanks to the very explicit expressions provided by Bourgeois and Cahen.

We also include an appendix on distributions and the Frobenius theorem (appendix A), as they intervene when we introduce horizontal distributions in chapter 2 (the Ehresmann point of view), as well as an appendix on analytic connections (appendix B), since they play a key role in our construction of non-symmetric connections of type S.

We assume our readers to be acquainted with the notions that an introductory class on differential geometry should include, such as abstract manifolds, vector fields and their flows, Lie derivatives, differential forms, and so forth. All the prerequisites will be found in [19].

We finish this introduction with an overview of the general references used in this work.

- As we already mentioned, for any question concerning the foundations of differential geometry, we turn to [19], a very thorough introduction with chapters on vector bundles, the Frobenius theorem and even symplectic manifolds.
- For the elementary notions of Riemannian geometry, we rely on [18] from the same author, which contains some computations that carry over to affinely connected manifolds.
- For the theory of connections, our ultimate reference is [12, 13] however the reader should be warned: the point of view adopted in this text is that of principal connections (on principal bundles) which are not covered here. In [26] the theory of connections is developed with vector bundles as a starting point. The same goes for [17] which includes a section on Ehresmann connection (an alternative, very general approach to connections).
- For the theory of affine connections, their curvature, geodesics and local symmetries we consult [12, 13] and [24]. [24] in particular takes affine connections as a starting point. In it the reader will find an extensive and self-contained account of Riemannian symmetric spaces, including an outline of their classification (together with an introduction to the theory of semisimple Lie algebras).
- [21], [7] and [22] cover everything pertaining to symplectic geometry (excluding symplectic connections), and almost everything concerning complex manifolds.
- As for Kähler manifolds, their connection and curvature, we once again use [13] (vol II.).
- The survey article [5] covers the basic properties of symplectic connections.
- Finally, for the more original sections of this thesis, the relevant articles are [2] (the Sekigawa-Vanhecke theorem), [6] (preferred connections on symplectic surfaces) and [8] (D'Atri spaces and the Ledger conditions). We also look in [3] to find examples of preferred symplectic connections that are not of type S.

Contents

Remerciements	5
Introduction	6
1 Algebraic Preliminaries	14
1.1 Tensor algebra	14
1.1.1 Induced morphisms of algebras	22
1.2 Polarization of a homogeneous polynomial	26
2 Vector bundles and connections	30
2.1 Vector Bundles	30
2.1.1 Sections of a bundle	33
2.1.2 Local operators and bundle morphisms	39
2.1.3 Vector subbundles	44
2.1.4 Tensor bundles and bundle-valued forms	45
2.1.5 Pullback bundle	52
2.2 Connections on vector bundles	57
2.2.1 Covariant derivative along a path and parallel transport	67
2.2.2 Curvature	72
2.2.3 Connections on tensor bundles	76
2.2.4 Horizontal distribution and vertical projection	80
3 Affine Connections	90
3.1 Definition, torsion and Bianchi identities	90
3.2 Ricci curvature and locally equiaffine manifolds	93
3.3 Geodesics, exponential mapping and normal coordinates	95
3.3.1 Geodesic symmetries and symmetric tensors	102
4 Riemannian, Symplectic and Kähler structures	112
4.1 Riemannian Connections	112
4.1.1 Curvature of the Levi-Civita connection	115
4.2 Symplectic Connections	118
4.2.1 Symplectic vector spaces	118
4.2.2 Elements of symplectic geometry	121
4.2.3 The space of symplectic connections	123

4.3	(Almost) Complex, Hermitian and Kähler Structures	127
4.3.1	Linear structures	129
4.3.2	(Almost) complex structures	133
4.3.3	Hermitian and Kähler metrics	136
4.4	Locally symmetric spaces	140
5	The Sekigawa-Vanhecke theorem	144
6	Divergence-preserving geodesic symmetries	148
6.1	Divergence and connections on the determinant bundle	148
6.2	Divergence-preserving symmetries	155
7	Preferred connections on symplectic surfaces	163
7.1	Odd Ledger conditions	163
7.2	A family of type S connections which are not locally symmetric	166
	Conclusion	168
A	Distributions, foliations and the Frobenius theorem	169
B	Analytic connections	172

Chapter 1

Algebraic Preliminaries

Differential geometry and differential calculus in general draw much of their power from the following intriguing phenomenon: "things" tend to become linear at the infinitesimal scale, meaning that they can be described through the theory of (finite-dimensional) vector spaces. In differential geometry, this idea is formalized via the notion of a tangent space: given a point in manifold, the space tangent to that point (a finite-dimensional vector space) can be thought of as a model for the 'infinitesimal' piece of the manifold surrounding that point. In that way, a lot is often gained by first understanding the linear structure of the tangent spaces, then mediating between the global and infinitesimal scales through the use of tensor fields (and more generally, sections of vector bundles). This theme, which we will encounter more than once in the following pages, might serve as an explanation for the ubiquity of tensor fields throughout differential geometry (and certainly throughout this thesis).

With this in mind, we begin with a preliminary chapter investigating the theory of tensors on finite-dimensional vector space. The purpose of this chapter is first and foremost to remind the reader of the basic properties of the tensor (and exterior) algebra of a finite-dimensional vector space, as such most proofs will be omitted. Very careful treatment of these subjects (which include the construction of the tensor and exterior algebra) can be found in [19], [26] or [27].

1.1 Tensor algebra

We begin by presenting the different point of views that we will adopt regarding tensors throughout this text. Interplay between these equivalent point of views will prove fruitful on numerous occasions.

Theorem 1.1.1. *Let $V_1, \dots, V_k$ be finite dimensional vector spaces.*

- (i) *(Tensors via the universal property) The tensor product $V_1 \otimes \dots \otimes V_k$ comes equipped with a multilinear map*

$$\otimes : V_1 \times \dots \times V_k \rightarrow V_1 \otimes \dots \otimes V_k.$$

This map is universal in the following sense: given a vector space W and a multilinear map

$$f : V_1 \times \dots \times V_k \rightarrow W$$

there exists a unique linear map

$$\bar{f} : V_1 \otimes \dots \otimes V_k \rightarrow W$$

such that

$$\begin{array}{ccc} V_1 \otimes \dots \otimes V_k & \xrightarrow{\bar{f}} & W \\ \uparrow \otimes & \nearrow f & \\ V_1 \times \dots \times V_k & & \end{array}$$

commutes, meaning that $\bar{f}$ sends $v_1 \otimes \dots \otimes v_k$ to $f(v_1, \dots, v_k)$ for all $v_1, \dots, v_k \in V$.

More concisely, we have a canonical linear isomorphism

$$\text{Lin}(V_1 \otimes \dots \otimes V_k, W) \simeq \text{Lin}(V_1, \dots, V_k; W)$$

where $\text{Lin}(V_1, \dots, V_k; W)$ denotes the space of multilinear maps

$$V_1 \times \dots \times V_k \rightarrow W$$

(and for two vector spaces A, B , $\text{Lin}(A, B)$ denotes the space of linear maps $A \rightarrow B$).

(ii) (Tensors as multilinear maps) There exists a unique linear isomorphism

$$V_1^* \otimes \dots \otimes V_k^* \simeq \text{Lin}(V_1, \dots, V_k; \mathbb{R})$$

that sends $\alpha^1 \otimes \dots \otimes \alpha^k$ (for $\alpha^j \in V_j^*$) to the multilinear map

$$V_1 \times \dots \times V_k \rightarrow \mathbb{R}, (v_1, \dots, v_k) \rightarrow \alpha^1(v_1) \dots \alpha^k(v_k)$$

which we also denote by $\alpha^1 \otimes \dots \otimes \alpha^k$. Equivalently, since V_j and $(V_j^*)^*$ are canonically isomorphic, we also have a unique isomorphism

$$V_1 \otimes \dots \otimes V_k \simeq \text{Lin}(V_1^*, \dots, V_k^*; \mathbb{R})$$

that sends $v_1 \otimes \dots \otimes v_k$ (for $v_j \in V_j$) to the multilinear map

$$V_1^* \times \dots \times V_k^* \rightarrow \mathbb{R}, (\alpha^1, \dots, \alpha^k) \rightarrow \alpha^1(v_1) \dots \alpha^k(v_k).$$

(iii) (Tensors through coordinates) For each $1 \leq i \leq k$, let $(e_1^{(i)}, \dots, e_{n_i}^{(i)})$ be a basis of V_i .

Then

$$\{e_{i_1}^{(1)} \otimes \dots \otimes e_{i_k}^{(k)} \mid 1 \leq i_1 \leq n_1, \dots, 1 \leq i_k \leq n_k\}$$

is a basis of $V_1 \otimes \dots \otimes V_k$.

Next we present what one might call the "commutativity" and "associativity" of the tensor product.

Proposition 1.1.1. *Let U, V, W be vector spaces. There exists canonical isomorphisms*

$$\begin{aligned} U \otimes V &\simeq V \otimes U \\ U \otimes (V \otimes W) &\simeq U \otimes V \otimes W \\ (U \otimes V) \otimes W &\simeq U \otimes V \otimes W \end{aligned}$$

that respectively send $u \otimes v$ to $v \otimes u$, $(u \otimes v) \otimes w$ to $u \otimes v \otimes w$ and $u \otimes (v \otimes w)$ to $u \otimes v \otimes w$ for all $u \in U$, $v \in V$ and $w \in W$.

Remark. When we talk of the tensor product being commutative, we refer to the aforementioned isomorphism $U \otimes V \simeq V \otimes U$, *not* to a supposed commutativity of the operation $\otimes : U \times V \rightarrow U \otimes V$. Indeed, for $u \in U$ and $v \in V$, $u \otimes v$ and $v \otimes u$ should not be expected to be equal, as the second expression will not even be well-defined in most cases (that is, when $U \neq V$). Even in the event that $U = V$, the equality will not hold in general.

Making use of the universal property of the tensor product, we can combine a family of linear maps $L_j : V_j \rightarrow W_j$ into a single linear map $V_1 \otimes \dots \otimes V_k \rightarrow W_1 \otimes \dots \otimes W_k$.

Proposition 1.1.2. *Let $V_1, \dots, V_k$ and $W_1, \dots, W_k$ be two families of vector spaces, and for each $1 \leq i \leq k$, let $L_i : V_i \rightarrow W_i$ be a linear map. There exists a unique linear map*

$$(L_1 \otimes \dots \otimes L_k) : (V_1 \otimes \dots \otimes V_k) \rightarrow (W_1 \otimes \dots \otimes W_k)$$

such that $(L_1 \otimes \dots \otimes L_k)(v_1 \otimes \dots \otimes v_k) = L_1(v_1) \otimes \dots \otimes L_k(v_k)$, for $v_j \in V_j$.

Before defining the tensor algebra of a vector space, we mention the following result, which will allow us to see the theory of linear maps as a special case of our theory of tensor products (and dual spaces).

Proposition 1.1.3. *Let V, W be finite-dimensional vector spaces. There exists a canonical isomorphism*

$$W \otimes V^* \rightarrow \text{Lin}(V, W)$$

that sends $w \otimes \alpha$ to

$$w\alpha(-) : V \rightarrow W, v \mapsto w\alpha(v)$$

for all $w \in W$, $\alpha \in V^*$.

Definition 1.1.1 (Tensor algebra). Let V be a vector space and let r, s be two natural numbers, we define

- (i) $T^r V := V^{\otimes r} = V \otimes V \otimes \dots \otimes V$ (r times), the tensor product of r copies of V .
- (ii) $T_s V := (V^*)^{\otimes s} = V^* \otimes V^* \otimes \dots \otimes V^*$ (s times), the tensor product of s copies of V^* , the vector space dual to V .
- (iii) $T_s^r V := T^r V \otimes T_s V$.

And by convention we set

$$T^0 V = T_0 V = T_0^0 V := \mathbb{R}.$$

We define the **tensor algebra** $T(V)$ of V to be the direct sum

$$T(V) := \bigoplus_{(r,s) \in \mathbb{N}^2} T_s^r V$$

(Notice that we take the direct sum and not the direct product, so that an element of $T(V)$ is a finite sum $\sum_{i=1}^m A_i$ with $A_i \in T_{s_i}^{r_i} V$). An element of $T_s^r V$ is said to be homogeneous of type (or degree) (r, s) .

To endow $T(V)$ with an internal multiplication $\otimes$, we start by defining a map

$$\otimes : T_s^r V \times T_l^k V \rightarrow T_{s+l}^{r+k} V$$

for each possible pair $(r, s), (k, l) \in \mathbb{N}^2$, then extend $\otimes$ bilinearly. To the different interpretations of tensors described in theorem 1.1.1 correspond different avenues for defining

$$\otimes : T_s^r V \times T_l^k V \rightarrow T_{s+l}^{r+k} V$$

- (i) (Via the universal property) We can use the natural isomorphisms

$$\begin{aligned} T_s^r V \otimes T_l^k V &= (T^r V \otimes T_s V) \otimes (T^k V \otimes T_l V) \\ &\simeq (T^r V \otimes T^k V) \otimes (T_s V \otimes T_l V) \simeq T^{r+k} V \otimes T_{s+l} V = T_{s+l}^{r+k} V \end{aligned}$$

meaning that we define $\otimes$ to be the unique bilinear map such that

$$(T \otimes \alpha) \otimes (S \otimes \beta) = (T \otimes S) \otimes (\alpha \otimes \beta)$$

for all $T \in T^r V$, $S \in T^k V$, $\alpha \in T_s V$ and $\beta \in T_l V$.

- (ii) (As multilinear maps) A tensor $T \in T_s^r V$ can be seen as a multilinear map

$$T : (V^*)^r \times V^s \rightarrow \mathbb{R}.$$

For $T \in T_s^r V$, $S \in T_l^k V$, the action of $T \otimes S$ as a multilinear map on $(V^*)^{r+k} \times V^{s+l}$ is given by

$$\begin{aligned} (T \otimes S)(\alpha^1, \dots, \alpha^{r+k}, v_1, \dots, v_{s+l}) \\ = T(\alpha^1, \dots, \alpha^r, v_1, \dots, v_s) S(\alpha^{r+1}, \dots, \alpha^{r+k}, v_{s+1}, \dots, v_{s+l}) \end{aligned}$$

for all $\alpha^j \in V^*$ and $v_j \in V$.

(iii) (Via components) Given a basis $(e_i)_{i=1}^m$ of V with dual basis $(e_*^i)_{i=1}^m$

$$\{e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e_*^{j_1} \otimes \dots \otimes e_*^{j_s} \mid 1 \leq i_1 \leq m, \dots, 1 \leq i_r \leq m, 1 \leq j_1 \leq m, \dots, 1 \leq j_s \leq m\}$$

is a basis of $T_s^r V$. For an element T of $T_s^r V$, we denote the coordinates of T in that basis by $T_{j_1 \dots j_s}^{i_1 \dots i_r}$, so that

$$T = \sum_{i_1, \dots, i_r, j_1, \dots, j_s} T_{j_1 \dots j_s}^{i_1 \dots i_r} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e_*^{j_1} \otimes \dots \otimes e_*^{j_s}$$

and call these coordinates the components of T for the basis $(e_i)_{i=1}^m$. The map $\otimes : T_s^r V \times T_l^k V \rightarrow T_{s+l}^{r+k} V$ can then be described as follows: for $T \in T_s^r V$ and $S \in T_l^k V$ the components of $S \otimes T$ are given by

$$(S \otimes T)_{j_1 \dots j_{s+l}}^{i_1 \dots i_{r+k}} = S_{j_1 \dots j_s}^{i_1 \dots i_r} T_{j_{s+1} \dots j_{s+l}}^{i_{r+1} \dots i_{r+k}}.$$

$T(V)$ endowed with this product is an $\mathbb{N}^2$ -graded $\mathbb{R}$ -algebra, meaning that for all $(r, s), (k, l) \in \mathbb{N}^2$,

$$T_s^r V \otimes T_l^k V \subseteq T_{s+l}^{r+k} V$$

(and in this case the inclusion is in fact an equality).

Remark. Notice that we have a canonical isomorphism

$$T_1^1 V = V \otimes V^* \simeq \text{End}(V)$$

where $\text{End}(V)$ denotes the space of linear endomorphisms of V .

We already know how to switch between the 'universal property', and 'multilinear maps' point of views on tensors, as the action of $\alpha^1 \otimes \dots \otimes \alpha^k$ on $V_1 \times \dots \times V_k$ was described in theorem 1.1.1, and at this point the relationship between the action of a tensor $T \in T_s^r V$ as a multilinear map and its components in a given basis could be derived without difficulty. Nevertheless, since we will make heavy use of said relationship in what is to come, we think it best to provide an explicit description.

Proposition 1.1.4. *Let V be finite-dimensional vector space with basis $(e_i)_{i=1}^m$ and dual basis $(e_*^i)_{i=1}^m$. Then a tensor $T \in T_s^r V$ of type (r, s) is uniquely written*

$$T = \sum_{i_1, \dots, i_r, j_1, \dots, j_s} T_{j_1 \dots j_s}^{i_1 \dots i_r} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e_*^{j_1} \otimes \dots \otimes e_*^{j_s}$$

with

$$T_{j_1 \dots j_s}^{i_1 \dots i_r} = T(e_*^{i_1}, \dots, e_*^{i_r}, e_{j_1}, \dots, e_{j_s}).$$

Now that covariant tensors have been introduced and interpreted as multilinear forms, we are ready to define the subclasses of symmetric and skew-symmetric tensors. The latter, which will become the exterior algebra once we endow it with the appropriate product, is arguably the most important one (for our purposes at least) as it gives rise to the theory of differential forms, hence we treat it with a bit more care.

Definition 1.1.2.

- (i) We call S_k the permutation group of degree k . For $\sigma \in S_k$ and $\alpha \in T_k V$ we define $\alpha^\sigma \in T_k V$ by

$$(\alpha^\sigma)(v_1, \dots, v_k) := \alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)})$$

for all $v_1, \dots, v_k \in V$.

- (ii) A covariant k -tensor α on V is said to be **symmetric** if $\alpha^\sigma = \alpha$ for all $\sigma \in S^k$. We denote by $S^k V^*$ the space of such tensors. We define a linear projection $\text{Sym} : T_k V \rightarrow S^k V^*$ which we call **symmetrization**. The symmetrization of a covariant k -tensor $\alpha \in T_k V$ is given by

$$\text{Sym}(\alpha) := \frac{1}{k!} \sum_{\sigma \in S_k} \alpha^\sigma.$$

- (iii) A covariant k -tensor α on V is said to be **skew-symmetric** (or **alternating**, or **antisymmetric**) if $\alpha^\sigma = (-1)^\sigma \alpha$ for all $\sigma \in S^k$, where $(-1)^\sigma$ denotes the signature of σ . The space of such tensors is denoted by $\Lambda^k V^*$. In analogy with the symmetric case, we define a linear projection $\text{Alt} : T_k V \rightarrow \Lambda^k V^*$ called the **alternation** (or **antisymmetrization**, or **skew-symmetrization**). The alternation of a covariant k -tensor $\alpha \in T_k V$ is given by

$$\text{Alt}(\alpha) := \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^\sigma \alpha^\sigma.$$

We define the exterior algebra $\Lambda(V)$ of V to be the direct sum

$$\Lambda(V) := \bigoplus_{k \in \mathbb{N}} \Lambda^k V^*.$$

We endow $\Lambda(V)$ with an internal multiplication $\wedge$, called the exterior or wedge product by first setting

$$\alpha \wedge \beta := \frac{(k+l)!}{k!l!} \text{Alt}(\alpha \otimes \beta)$$

for $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^l V^*$, then extending $\wedge$ bilinearly. Equipped with the wedge product, $\Lambda(V)$ becomes an anticommutative $\mathbb{N}$ -graded algebra, meaning that

$$(\Lambda^k V^*) \wedge (\Lambda^l V^*) \subseteq \Lambda^{k+l} V^*$$

(once again, the inclusion is in fact an equality), and that for all $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^l V^*$, we have

$$\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha.$$

Let us finish our overview of the exterior algebra by providing a basis for the space $\Lambda^k V^*$. Let $(e_i)_{i=1}^m$ be a basis of V , with dual basis $(e_*^i)_{i=1}^m$. Then an element α of $\Lambda^k V^*$ is uniquely determined by the set of numbers

$$\{\alpha(e_{i_1}, \dots, e_{i_k}) \mid \text{for } 1 \leq i_1 < i_2 < \dots < i_k \leq m\}.$$

This suggests the following choice of basis: for $1 \leq i_1 < i_2 < \dots < i_k \leq m$, and $1 \leq j_1 < j_2 < \dots < j_k \leq m$, the skew-symmetric tensor $e_*^{i_1} \wedge \dots \wedge e_*^{i_k}$ satisfies

$$(e_*^{i_1} \wedge \dots \wedge e_*^{i_k})(e_{j_1}, \dots, e_{j_k}) = \delta_{j_1}^{i_1} \dots \delta_{j_k}^{i_k}.$$

Thus we see that the set

$$\{e_*^{i_1} \wedge \dots \wedge e_*^{i_k} \mid \text{for } 1 \leq i_1 < i_2 < \dots < i_k \leq m\}.$$

is a basis of $\Lambda^k V^*$ and that any skew-symmetric k -tensor $\alpha \in \Lambda^k V^*$ can be uniquely written as

$$\alpha = \sum_{i_1 < i_2 < \dots < i_k} \alpha_{i_1 \dots i_k} e_*^{i_1} \wedge \dots \wedge e_*^{i_k}$$

with $\alpha_{i_1 \dots i_k} = \alpha(e_{i_1}, \dots, e_{i_k})$.

Example 1.1.1. Given $\alpha, \beta \in V^*$ and $v, w \in V$ we have

$$(\alpha \wedge \beta)(v, w) = \alpha(v)\beta(w) - \alpha(w)\beta(v).$$

An additional property of skew-symmetric (covariant) tensors that will be of use to us later is their universality as alternating maps, which is encapsulated in the following proposition.

Proposition 1.1.5. *Let V and W be finite-dimensional vector spaces. Then there exists a natural isomorphism*

$$W \otimes \Lambda^k V^* \simeq \text{Alt}(V^k; W)$$

(where $\text{Alt}(V^k; W)$ denotes the space of alternating multilinear maps $V^k \rightarrow W$) which sends $w \otimes \alpha$ ($w \in W, \alpha \in \Lambda^k V^*$) to the map

$$V^k \rightarrow W, (v_1, \dots, v_k) \rightarrow w\alpha(v_1, \dots, v_k).$$

We finish this section by defining the contraction operation on tensors. What motivates this definition is the fact that we do not know of a 'common' operation involving tensors that cannot be reduced to a combination of contractions and tensor products. Thus for each operation acting on tensors that we will introduce in later chapters (such as covariant differentiation), we will gain essentially complete knowledge of the 'algebraic' behaviour of these operations just by investigating their interaction with contraction and the tensor product.

Definition 1.1.3. Let $1 \leq i \leq r$ and $1 \leq j \leq s$. The (i, j) **contraction operation** is a linear map $C_j^i : T_s^r V \rightarrow T_{s-1}^{r-1} V$ that can be defined in the following equivalent ways:

- (i) The unique linear map obtained by applying the canonical pairing

$$V \otimes V^* \rightarrow \mathbb{R}, (v, \alpha) \rightarrow \langle v, \alpha \rangle := \alpha(v)$$

to the i th V factor and j th V^* factor of $T_s^r V = V \otimes \dots \otimes V \otimes V^* \otimes \dots \otimes V^*$ and the identity on all other factors. In other words C_j^i is the unique linear map $T_s^r V \rightarrow T_{s-1}^{r-1} V$ such that for all $v_1, \dots, v_r \in V, \alpha^1, \dots, \alpha^s \in V^*$

$$C_j^i(v_1 \otimes \dots \otimes v_r \otimes \alpha^1 \otimes \dots \otimes \alpha^s) = \langle v_i, \alpha^j \rangle v_1 \otimes \dots \otimes \hat{v}_i \otimes \dots \otimes v_r \otimes \alpha^1 \otimes \dots \otimes \hat{\alpha}^j \otimes \dots \otimes \alpha^s$$

where the hats indicate that v_i and α^j are omitted.

- (ii) Let $(e_k)_{k=1}^m$ and $(e_*^k)_{k=1}^m$ be a basis of V with dual basis $(e_*^k)_{k=1}^m$. Interpreting tensors as multilinear maps, for $T \in T_s^r V$, $C_j^i T$ is defined by

$$\begin{aligned} (C_j^i T)(\alpha^1, \dots, \hat{\alpha}^i, \dots, \alpha^r, v_1, \dots, \hat{v}_j, \dots, v_s) \\ = \sum_{k=1}^m T(\alpha^1, \dots, \alpha^{i-1}, e_*^k, \alpha^{i+1}, \dots, \alpha^r, v_1, \dots, v_{j-1}, e_k, v_{j+1}, \dots, v_s) \end{aligned}$$

for all $\alpha^1, \dots, \hat{\alpha}^i, \dots, \alpha^r \in V^*$ and $v_1, \dots, \hat{v}_j, \dots, v_s \in V$.

- (iii) In terms of components: let $T \in T_s^r V$ and write

$$T = \sum_{i_1, \dots, i_r, j_1, \dots, j_s} T_{j_1 \dots j_s}^{i_1 \dots i_r} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e_*^{j_1} \otimes \dots \otimes e_*^{j_s}.$$

Then the components of $C_j^i T$ are given by

$$(C_j^i T)_{j_1 \dots j_{s-1}}^{i_1 \dots i_{r-1}} = \sum_{k=1}^m T_{j_1 \dots k \dots j_{s-1}}^{i_1 \dots k \dots i_{r-1}}$$

where the superscript k appears in the i th position and the subscript k appears in the j th position.

For $1 \leq i_1 < i_2 < \dots < i_k \leq r$ and $1 \leq j_1 < j_2 < \dots < j_k \leq s$ we define $C_{j_1 \dots j_k}^{i_1 \dots i_k}$ to be the composition

$$C_{j_1 \dots j_k}^{i_1 \dots i_k} := C_{j_1}^{i_1} \circ C_{j_2}^{i_2} \circ \dots \circ C_{j_k}^{i_k} : T_s^r V \rightarrow T_{s-k}^{r-k} V.$$

Example 1.1.2. As stated earlier, almost all the natural operations one can apply to tensors are obtained by combining contraction and the tensor product.

- (i) Through the canonical isomorphism $\text{End}(V) \simeq V \otimes V^*$ we have for all $A, B \in \text{End}(V)$, $v \in V$, $\alpha \in V^*$

$$\begin{aligned} C_1^1 A &= \text{tr} A \\ C_1^2(A \otimes B) &= A \circ B \\ C_1^2(A \otimes v) &= Av \\ C_2^1(A \otimes \alpha) &= A^* \alpha. \end{aligned}$$

- (ii) For all $T \in T_s^r V$, $\alpha^1, \dots, \alpha^r \in V^*$, $v_1, \dots, v_s \in V$ we have

$$C_{1,2,\dots,2s}^{1,2,\dots,2r}(\alpha^1 \otimes \dots \otimes \alpha^r \otimes T \otimes v_1 \otimes \dots \otimes v_s) = T(\alpha^1, \dots, \alpha^r, v_1, \dots, v_s).$$

1.1.1 Induced morphisms of algebras

The construction of the tensor algebra of a vector space V relies solely on the linear structure of V , so it should come as no surprise that an isomorphism of vector spaces $V \rightarrow W$ induces an isomorphism $T(V) \rightarrow T(W)$ that preserves all the relevant structures, namely contraction and the tensor product. Still, this process of turning a linear isomorphism into a morphism of tensor algebras will make enough appearances in the present work to warrant a dedicated subsection. One last note before we begin, when $V = W$ we will obtain an 'infinitesimal' version of this process, which we will also describe.

Definition 1.1.4. Let $L : V \rightarrow W$ be a linear isomorphism. Then for any pair of natural numbers r, s , L induces a linear isomorphism, $L \cdot (-) : T_s^r V \rightarrow T_s^r W$ that can be defined in the following equivalent ways:

- (i) The unique linear map obtained by applying L to all the V factors of $T_s^r V = V \otimes \dots \otimes V \otimes V^* \otimes \dots \otimes V^*$ and $(L^{-1})^*$ to all V^* factors. In other words $L \cdot (-)$ is the unique linear map $T_s^r V \rightarrow T_s^r W$ such that for all $v_1, \dots, v_r \in V$, $\alpha^1, \dots, \alpha^s \in V^*$

$$L \cdot (v_1 \otimes \dots \otimes v_r \otimes \alpha^1 \otimes \dots \otimes \alpha^s) = (Lv_1) \otimes \dots \otimes (Lv_r) \otimes ((L^{-1})^* \alpha^1) \otimes \dots \otimes ((L^{-1})^* \alpha^s).$$

- (ii) If we view an element $T \in T_s^r V$ as a multilinear map $(V^*)^r \times V^s \rightarrow \mathbb{R}$, then $L \cdot T$ is defined by

$$(L \cdot T)(\beta_1, \dots, \beta_r, w^1, \dots, w^s) := T(L^* \beta_1, \dots, L^* \beta_r, L^{-1} w^1, \dots, L^{-1} w^s)$$

for all $\beta_1, \dots, \beta_r \in W^*$ and $w^1, \dots, w^s \in W$ (here L^* denotes the adjoint of L).

- (iii) In terms of components: let $(e_k)_{k=1}^m, (f_k)_{k=1}^m$ be bases of V and W respectively. Write $Le_j = \sum_{i=1}^k L_j^i f_i$, $L^{-1} f_j = \sum_{i=1}^k (L^{-1})_j^i e_i$ and

$$T = \sum_{i_1, \dots, i_r, j_1, \dots, j_s} T_{j_1 \dots j_s}^{i_1 \dots i_r} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e_*^{j_1} \otimes \dots \otimes e_*^{j_s}.$$

Then the components of $L \cdot T$ (in the basis $(f_k)_{k=1}^m$) are given by

$$(L \cdot T)_{j_1 \dots j_s}^{i_1 \dots i_r} = \sum_{t_1, \dots, t_r, q_1, \dots, q_s} L_{t_1}^{i_1} \dots L_{t_r}^{i_r} (L^{-1})_{j_1}^{q_1} \dots (L^{-1})_{j_s}^{q_s} T_{q_1 \dots q_s}^{t_1 \dots t_r}.$$

Example 1.1.3. Let us provide a few special cases of the process described above. Fix a linear isomorphism $L : V \rightarrow W$.

- (i) By convention, the action of L on $T_0^0 V = \mathbb{R} = T_0^0 W$ is the identity map.
- (ii) The induced map $T^1 V = V \rightarrow W = T^1 W$ is of course L itself.
- (iii) Given a covariant k -tensor $\alpha \in T_k V$, $(L \cdot \alpha) = (L^{-1})^* \alpha$, where $(L^{-1})^* \alpha$ denotes the pullback of α along L^{-1} . In particular, the isomorphism $L \cdot (-) : T_k V \rightarrow T_k W$ maps $\Lambda^k V^*$ and $S^k V^*$ isomorphically onto $\Lambda^k W^*$ and $S^k W^*$ respectively.
- (iv) A special case of (iii) for $V = W$ which will be very important in the course of this thesis is that of top skew-symmetric tensors on V , i.e. elements of $\Lambda^m V^*$ with $m := \dim V$. Since $\Lambda^m V^*$ has dimension 1, there must exist a unique real number $d(L)$ such that $L \cdot \omega = d(L) \omega$ for all $\omega \in \Lambda^m V^*$. This number is non-other than $\det(L)^{-1}$, thus in summary: given a linear automorphism $L \in GL(V)$, a top skew-symmetric tensor $\omega \in \Lambda^m V^*$ and $v^1, \dots, v^m \in V$ we have

$$(L \cdot \omega)(v^1, \dots, v^m) = \omega(L^{-1} v^1, \dots, L^{-1} v^m) = \det(L)^{-1} \omega(v^1, \dots, v^m).$$

- (v) Through the canonical isomorphisms $\text{End}(V) \simeq V \otimes V^*$ and $\text{End}(W) \simeq W \otimes W^*$, the map $L \cdot (-) : \text{End}(V) \rightarrow \text{End}(W)$ is the one we would expect, i.e. for $A \in \text{End}(V)$ we have $L \cdot A = L \circ A \circ L^{-1}$.

Remark. One advantage of starting with the case $V \neq W$ is that the presence of a (-1) exponent in (iii) and (iv) is not surprising. When $V = W$, the presence of this exponent remains necessary in order for the corresponding $GL(V)$ -action to be a left-action (and not a right-action).

As announced, all the relevant structures are preserved: the induced map $T(V) \rightarrow T(W)$ is an isomorphism of graded algebras, and preserves contraction.

Proposition 1.1.6. *Let $L : V \rightarrow W$ be a linear isomorphism.*

(i) $L \cdot (T \otimes S) = (L \cdot T) \otimes (L \cdot S)$ for all $T \in T_s^r V$, $S \in T_l^k V$.

(ii) For all $\alpha \in T_k V$ and $\sigma \in S_k$ we have

$$\begin{aligned} L \cdot (\alpha^\sigma) &= (L \cdot \alpha)^\sigma \\ \text{Sym}(L \cdot \alpha) &= L \cdot \text{Sym}(\alpha) \\ \text{Alt}(L \cdot \alpha) &= L \cdot \text{Alt}(\alpha). \end{aligned}$$

(iii) $L \cdot (\alpha \wedge \beta) = (L \cdot \alpha) \wedge (L \cdot \beta)$ for all $\alpha \in \Lambda^k V^*$, $\beta \in \Lambda^l V^*$.

(iv) Let $1 \leq i_1 < i_2 < \dots < i_k \leq r$ and $1 \leq j_1 < j_2 < \dots < j_k \leq s$. Then for all $T \in T_s^r V$ we have

$$C_{j_1 \dots j_k}^{i_1 \dots i_k}(L \cdot T) = L \cdot C_{j_1 \dots j_k}^{i_1 \dots i_k} T.$$

When $V = W$, definition 1.1.4 gives us a linear action of $GL(V)$ on $T_s^r V$, or equivalently a Lie group morphism $GL(V) \rightarrow GL(T_s^r V)$. This Lie group morphism induces a morphism of Lie algebras $\text{End}(V) \rightarrow \text{End}(T_s^r V)$, and in that way a linear endomorphism A of V induces a linear endomorphism of $T_s^r V$.

Definition 1.1.5. A linear map $A : V \rightarrow V$ induces a linear map $A \cdot (-) : T_s^r V \rightarrow T_s^r V$ which we also call A . Given $T \in T_s^r V$, $A \cdot T$ is defined by

$$A \cdot T := \left. \frac{d}{dt} \right|_{t=0} e^{tA} \cdot T.$$

In analogy with the three cases of definition 1.1.4 we have the following equivalent definitions of $A \cdot T$:

(i) The unique linear map $T_s^r V \rightarrow T_s^r V$ such that for all $v_1, \dots, v_r \in V$, $\alpha^1, \dots, \alpha^s \in V^*$

$$\begin{aligned} A \cdot (v_1 \otimes \dots \otimes v_r \otimes \alpha^1 \otimes \dots \otimes \alpha^s) &= \sum_{i=1}^r v_1 \otimes \dots \otimes (A v^i) \otimes \dots \otimes v_r \otimes \alpha^1 \otimes \dots \otimes \alpha^s \\ &\quad - \sum_{i=1}^s v_1 \otimes \dots \otimes v_r \otimes \alpha^1 \otimes \dots \otimes (A^* \alpha^i) \otimes \dots \otimes \alpha^s. \end{aligned}$$

- (ii) If we view tensors as multilinear mappings, then for $T \in T_s^r V$, $\alpha^1, \dots, \alpha^r \in V^*$ and $v_1, \dots, v_s \in V$ we have

$$(A \cdot T)(\alpha^1, \dots, \alpha^r, v_1, \dots, v_s) = \sum_{i=1}^r T(\alpha^1, \dots, A^* \alpha^i, \dots, \alpha^r, v_1, \dots, v_s) \\ - \sum_{i=1}^s T(\alpha_1, \dots, \alpha_r, v^1, \dots, A v^i, \dots, v^s)$$

for all $\alpha^1, \dots, \alpha^r \in V^*$ and $v_1, \dots, v_s \in V$.

- (iii) In terms of components, if $T \in T_s^r V$ has components $T_{j_1 \dots j_s}^{i_1 \dots i_r}$ in a basis $(e_k)_{k=1}^m$, then in that same basis the components of $A \cdot T$ are given by

$$(A \cdot T)_{j_1 \dots j_s}^{i_1 \dots i_r} = \sum_{k=1}^r \sum_t A_t^{i_k} T_{j_1 \dots j_s}^{i_1 \dots i_{k-1} t i_{k+1} \dots i_r} \\ - \sum_{k=1}^s \sum_q A_{j_k}^q T_{j_1 \dots j_{k-1} q j_{k+1} \dots j_s}^{i_1 \dots i_r}$$

where the matrix (A_j^i) is defined by $A e_j = A_j^i e_i$.

Remark. Note that for any differentiable map $L : J \rightarrow GL(V), t \rightarrow L_t$ (J an open interval containing 0) with $L_0 = I$ and $\left. \frac{d}{dt} \right|_{t=0} L_t = A \in \text{End}(V)$ we have

$$\left. \frac{d}{dt} \right|_{t=0} L_t \cdot T = A \cdot T$$

for all $T \in T_s^r V$.

Each property and example given for the $GL(V)$ -action gives rise to a property or example for the action of $\text{End}(V)$ through differentiation. In particular, for $A \in \text{End}(V)$, the induced action $A \cdot (-)$ is an 'infinitesimal graded-algebra automorphism', i.e. a graded-derivation of degree 0.

Proposition 1.1.7. *Let $A : V \rightarrow V$ be a linear map. Then the induced map $A \cdot (-) : T_k V \rightarrow T_k V$ preserves $\Lambda^k V^*$ and $S^k V^*$. Furthermore*

$$(i) \quad A \cdot (T \otimes S) = (A \cdot T) \otimes S + T \otimes (A \cdot S) \text{ for all } T \in T_s^r V, S \in T_l^k V.$$

(ii) For all $\alpha \in T_k V$ and $\sigma \in S_k$ we have

$$A \cdot (\alpha^\sigma) = (A \cdot \alpha)^\sigma \\ \text{Sym}(A \cdot \alpha) = A \cdot \text{Sym}(\alpha) \\ \text{Alt}(A \cdot \alpha) = A \cdot \text{Alt}(\alpha).$$

- (iii) $A \cdot (\alpha \wedge \beta) = (A \cdot \alpha) \wedge \beta + \alpha \wedge (A \cdot \beta)$ for all $\alpha \in \Lambda^k V^*$ and $\beta \in \Lambda^l V^*$.
- (iv) Let $1 \leq i_1 < i_2 < \dots < i_k \leq r$ and $1 \leq j_1 < j_2 < \dots < j_k \leq s$. Then for all $T \in T_s^r V$ we have

$$C_{j_1 \dots j_k}^{i_1 \dots i_k}(A \cdot T) = A \cdot C_{j_1 \dots j_k}^{i_1 \dots i_k} T.$$

Example 1.1.4. Fix $A \in \text{End}(V)$, this time we list some special cases of the action of A on tensor spaces of V .

- (i) The action of A on $T_0^0 V = \mathbb{R}$ is the zero map. Once again the induced map $T^1 V = V \rightarrow V = T^1 V$ is simply A itself. Notice however that the induced map $V^* \rightarrow V^*$ is not A^* but $-A^*$.
- (ii) For a top skew-symmetric tensor $\omega \in \Lambda^m V^*$ we have $A\omega = -(tr A)\omega$, or in more details: for $v^1, \dots, v^m \in V$:

$$(A \cdot \omega)(v^1, \dots, v^m) = - \sum_{i=1}^m \omega(v^1, \dots, Av^i, \dots, v^m) = -(tr A)\omega(v^1, \dots, v^m).$$

As stated in example 1.1.3 (iv), this will be of particular importance for us.

- (iii) Through the canonical isomorphism $\text{End}(V) \simeq V \otimes V^*$, for $B \in \text{End}(V)$ we have

$$A \cdot B = A \circ B - B \circ A.$$

Remark. Once (iv) in example 1.1.3 is known, (ii) in the list above becomes equivalent to

$$\left. \frac{d}{dt} \right|_{t=0} \det(e^{tA}) = tr(A)$$

for all $A \in \text{End}(V)$. Note also that the awkward ‘-’ sign that appears in a few places corresponds to the (-1) exponent for the $GL(V)$ -action, whose presence we justified in a previous remark.

1.2 Polarization of a homogeneous polynomial

The Ledger curvature conditions we will encounter in the last chapters will take the form of homogeneous polynomial equations, and we will sometimes need to turn them into linear equations, a process known as **polarization**.

We begin by making precise what we mean by a (homogeneous) polynomial function on an arbitrary finite-dimensional vector space.

Definition 1.2.1. Let V be a real vector space of dimension $m < \infty$. A basis $e = (e_k)_{k=1}^m$ of V induces a linear isomorphism

$$e : \mathbb{R}^m \rightarrow V, (x^k)_{k=1}^m \rightarrow \sum_{k=1}^m x^k e_k$$

(in fact the datum of a basis of V is equivalent to the datum of a linear isomorphism $\mathbb{R}^m \rightarrow V$). A (real) polynomial on V is a map $f : V \rightarrow \mathbb{R}$ such that for any basis e of V , the composition

$$f \circ e : \mathbb{R}^m \rightarrow \mathbb{R}$$

is a (real) polynomial in m variables. One can show that given a polynomial function $g : \mathbb{R}^m \rightarrow \mathbb{R}$ and a linear isomorphism $L : \mathbb{R}^m \rightarrow \mathbb{R}^m$, the composition $g \circ L : \mathbb{R}^m \rightarrow \mathbb{R}$ is a polynomial function. From this we deduce that a map $f : V \rightarrow \mathbb{R}$ is a polynomial as soon as there exists a single basis e of V for which $f \circ e$ is polynomial.

Definition 1.2.2. A homogeneous polynomial of degree k on V is a polynomial map $f : V \rightarrow \mathbb{R}$ such that $f(\lambda v) = \lambda^k f(v)$ for all $v \in V$ and $\lambda \in \mathbb{R}$.

Remark. Composition with a linear isomorphism $\mathbb{R}^m \rightarrow \mathbb{R}^m$ (i.e. a change of basis) preserves the degree of a polynomial, so that we could still speak of the degree of a non-homogeneous polynomial function $f : V \rightarrow \mathbb{R}$.

We are now in a position to define polarization which, to be completely rigorous, is not a process one applies to *equations* but to homogeneous polynomials *themselves*. It is the process through which we identify a homogeneous polynomial of degree k on V with a symmetric covariant k -tensor on V .

Theorem 1.2.1. Let $f : V \rightarrow \mathbb{R}$ be a homogeneous polynomial of degree k . There exists a unique symmetric k -tensor $F \in S^k V^*$ such that for all $v \in V$,

$$f(v) = F(v, v, \dots, v).$$

We call F the **polar** (or **linear**) form of f . The process of going from f to F is called **polarization** or **linearization**.

Proof. Let $e = (e_k)_{k=1}^m$ be a basis of V . There exists a unique family of coefficients $(a_\alpha)_{\alpha \in \mathbb{N}^r, |\alpha|=k}$ (where we use the multi-index notation $|\alpha| = \sum_{i=1}^r \alpha_i$ for $\alpha \in \mathbb{N}^r$) such that for all $v = \sum_{i=1}^r x^i e_i \in V$

$$f(v) = \sum_{|\alpha|=k} a_\alpha x^\alpha \tag{1.1}$$

with $x^\alpha := (x^1)^{\alpha_1} (x^2)^{\alpha_2} \dots (x^r)^{\alpha_r}$.

Now let, $F \in S^k V^*$ be a symmetric k -tensor on V , with components $F_{i_1 \dots i_k}$ in the basis e . For a multi-index $\alpha \in \mathbb{N}^r$ with $|\alpha| = k$ we define F_α by

$$F_\alpha := F \underbrace{11\dots 1}_{(\alpha_1 \text{ times})} \underbrace{22\dots 2}_{(\alpha_2 \text{ times})} \dots \underbrace{rr\dots r}_{(\alpha_r \text{ times})}.$$

So for example, for $r = 4$, $\alpha = (4, 0, 0, 0)$ and $\beta = (1, 2, 1, 0)$ we have

$$\begin{aligned} F_\alpha &= F_{1111} \\ F_\beta &= F_{1223}. \end{aligned}$$

Then for all $v = \sum_{i=1}^r x^i e_i$ we have

$$F(v, v, \dots, v) = \sum_{|\alpha|=k} \binom{r}{\alpha} F_\alpha x^\alpha \quad (1.2)$$

with the multinomial coefficient $\binom{r}{\alpha} = \frac{r!}{\alpha_1! \dots \alpha_k!}$ (the number of ways to place r items into k bins with α_i items in the i th bin).

Comparing equations 1.1 and 1.2, we see that the polar form F of f is the unique symmetric k -tensor such that

$$F_\alpha = \left[1 / \binom{r}{\alpha} \right] a_\alpha$$

for all $\alpha \in \mathbb{N}^r$ with $|\alpha| = k$. □

In the applications relevant to us, the homogeneous polynomial in question will always be defined using an initial (not necessarily symmetric) tensor. In that particular instance, polarization coincides with an operation we already introduced: symmetrization.

Theorem 1.2.2. *Let α be a covariant k -tensor on V . Then the map*

$$\bar{\alpha} : V \rightarrow \mathbb{R}, v \rightarrow \alpha(v, v, \dots, v)$$

is a homogeneous polynomial of degree k , and its polar form is $\text{Sym}(\alpha)$.

Proof. To see that $\bar{\alpha}$ is a polynomial map, simply write $\alpha(v, v, \dots, v)$ (for $v \in V$) in component form after a choice of basis. The fact that $\bar{\alpha}(\lambda v) = \lambda^k \bar{\alpha}(v)$ for $\lambda \in \mathbb{R}$, $v \in V$ is an immediate consequence of the multilinearity of α .

Lastly, $\text{Sym}(\alpha)$ is symmetric by construction, and since $\alpha^\sigma(v, v, \dots, v) = \alpha(v, v, \dots, v)$ for all $\sigma \in S_k$, $v \in V$, one easily checks that

$$\bar{\alpha}(v) = \alpha(v, v, \dots, v) = \text{Sym}(\alpha)(v, v, \dots, v)$$

for all $v \in V$ and the conclusion follows from the uniqueness in theorem [1.2.1](#). $\square$

Corollary 1.2.1. *Let α be a covariant k -tensor on V . If $\alpha(v, v, \dots, v) = 0$ for all $v \in V$, then $\text{Sym}(\alpha) = 0$. In particular if α is already symmetric and if $\alpha(v, v, \dots, v) = 0$ for all $v \in V$ then $\alpha = 0$.*

Chapter 2

Vector bundles and connections

We now leave the realm of pure algebra and enter that of differential geometry proper. In this chapter, much time will be spent developing the general theory of connections on vector bundles. This theory will then find its incarnation in subsequent chapters in the form of affine connections, which are connections on the tangent bundle.

Of course we could have chosen affine connections as our starting point, thereby avoiding a lot of work, and perhaps this would have been wiser. But we believe that this more general approach presents a number of advantages, the first being that an introduction to the theory of connections on vector bundles, a very powerful and far-reaching theory, could prove useful to the reader (this however would depend heavily on the author's abilities). Another benefit comes from the fact that, although our main object of study will remain affine connections (that is, connections on the tangent bundle), we will still come into contact with vector bundles other than the tangent bundle through tensor fields, and we will need to describe how these affine connections act on said tensor fields. In that regard, a general theory of connections on vector bundles will allow for a more unified and streamlined treatment. Lastly, time invested in a more general presentation is often rewarded by greater clarity, as the most essential and salient properties are laid bare.

Going back to more concrete matters, the main references used in this section are [26], [19] and [17] (this last one applies mainly to the last section, where we introduce the Ehresmann point of view).

2.1 Vector Bundles

As mentioned in the beginning of chapter 1, the geometry of a manifold becomes linear at the infinitesimal scale (when we 'zoom in' enough on a given point) thus each point of a manifold come with its own little linear structure. But to make this idea precise, and to treat all of these pointwise structures at once, and in such a way that we take the smooth structure of our manifold into account, we would need a notion of

'parametrized linear algebra' (an expression we will borrow from the dedicated nLab page).

We could attempt to achieve this by simply taking the direct product of our manifold with a given vector space. Certainly this would give us a 'smooth' (in this case essentially constant) family of vector spaces parametrized by our manifold. And indeed, it is a step in the right direction, and such direct products will form an important class of vector bundles, called **trivial** vector bundles to emphasize that they are the simplest of their kind. But this approach will turn out to be too restrictive. The ideal compromise is reached instead by requiring a vector bundle to be only *locally* isomorphic to a direct product (or equivalently, locally trivial). This is the essence behind the definition of a vector bundle, which we will now state.

Definition 2.1.1. A (real) **vector bundle of rank r** over M is a smooth surjection $\pi : E \rightarrow M$ such that

- (i) For every $p \in M$, the set $E_p := \pi^{-1}(p)$ is a real vector space of dimension r .
- (ii) Every point $p \in M$ has an open neighbourhood U such that there exists a fiber-preserving diffeomorphism

$$\phi_U : E_U := \pi^{-1}(U) \rightarrow U \times \mathbb{R}^r$$

that restricts to a linear isomorphism $E_q \rightarrow q \times \mathbb{R}^r$ for each $q \in U$.

E is called the **total space**, M is called the **base space**, and the space E_p the **fiber above p** . We call the open subset U in (ii) a **trivializing open subset for E** , and ϕ_U a **trivialization of $\pi^{-1}(U)$** . A trivializing open cover for E is a collection $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ where for each $\alpha \in A$, ϕ_α is a trivialization of E over the open subset U_α , and such that the U_α cover M .

We immediately give the correct notions of morphism and isomorphism for vector bundles.

Definition 2.1.2. Let $\pi_E : E \rightarrow M$ and $\pi_F : F \rightarrow N$ be two vector bundles. A **morphism of vector bundles** from E to F is a pair of smooth maps $\Phi : E \rightarrow F$, $\phi : M \rightarrow N$ such that

- (i) The diagram

$$\begin{array}{ccc} E & \xrightarrow{\Phi} & F \\ \pi_E \downarrow & & \downarrow \pi_F \\ M & \xrightarrow{\phi} & N \end{array}$$

commutes.

(ii) The map Φ restricts to a linear map $\Phi_p : E_p \rightarrow F_{\phi(p)}$ of fibers for each $p \in M$.

We often refer to Φ as the vector bundle morphism, which is not a terrible mistake since ϕ can be recovered from Φ . We call Φ a **vector bundle morphism over ϕ** . An important special case is when $M = N$ and $\phi = id_M$, we then speak of a **vector bundle morphism over M** . Since vector bundles are the only kind of fiber bundles we will treat in this thesis, we will sometimes refer to them simply as 'bundles' (e.g. by speaking of 'bundle morphisms' instead of 'vector bundle morphisms').

Bundle isomorphisms are defined in the usual way: as invertible bundle morphisms. More explicitly, given two vector bundles $\pi_E : E \rightarrow M$, $\pi_F : F \rightarrow N$ and a diffeomorphism $\phi : M \rightarrow N$, a bundle isomorphism $E \rightarrow F$ over ϕ is a bundle morphism $\Phi : E \rightarrow F$ over ϕ for which there exists a bundle morphism $\Phi^{-1} : F \rightarrow E$ over ϕ^{-1} such that

$$\begin{aligned}\Phi \circ \Phi^{-1} &= id_F \\ \Phi^{-1} \circ \Phi &= id_E.\end{aligned}$$

When $M = N$ we often speak of bundle isomorphisms $E \rightarrow F$ with no further specifications. In that case it is assumed that the isomorphism lies over id_M .

Example 2.1.1.

- (i) In most introductions to differential geometry, the first vector bundles one encounters are the tangent bundle TM of a smooth manifold, whose local trivializations are provided by coordinate charts, and its tensor bundles $T_s^r M$ (we will come back to tensor bundles shortly). Very important as well are the exterior powers $\Lambda^k T^*M$ of the cotangent bundle T^*M . The **determinant bundle** of M in particular, that is the top exterior power $\Lambda^{top} T^*M = \Lambda^m T^*M$ ($m = \dim(M)$), will play a special role in this thesis.
- (ii) The simplest examples of vector bundles however, are trivial bundles. The trivial bundle of rank r over M is simply the direct product $M \times \mathbb{R}^r$ with the obvious smooth structure, fiberwise linear structure and projection. We use the notation $\mathbb{R}_M^r$ (or $\mathbb{R}^r$ if there is no confusion about the base manifold) to refer to the trivial bundle of rank r over M . Though they are simple, they are of central importance, as they provide the local model for vector bundles. Indeed, given a vector bundle E of rank r and an open subset U of M , a trivialization of E over U is precisely a bundle isomorphism between E_U and the trivial bundle of rank r over U .

But as we already mentioned, they can only serve as *local* models: most vector bundles are not *globally* trivial (meaning they are not globally isomorphic to a trivial bundle), which is one of the main motivations behind the notion of a vector bundle. The example *par excellence* is the tangent bundle of the 2-sphere S^2 , which is non-trivial as a consequence of the hairy ball theorem.

We finish this introductory section with a little criterion telling us that a bundle morphism is a bundle isomorphism as soon as it acts isomorphically on fibers.

Proposition 2.1.1. *Let E and F be two vector bundles over M . Let $\Phi : E \rightarrow F$ be a bundle morphism (over M). Then Φ is a bundle isomorphism if and only if it is fiberwise isomorphic, meaning for each $p \in M$, $\Phi_p : E_p \rightarrow F_p$ is an isomorphism of vector spaces.*

Proof. Clearly if Φ is a bundle isomorphism, then it must act isomorphically on each fiber.

We need only prove the converse locally, so we can assume that E and F are trivial. They must be of equal rank, so we can write $E = \mathbb{R}_M^r = F$ (recall that $\mathbb{R}_M^r := M \times \mathbb{R}^r$). Since Φ is morphism of vector bundles, and since it restricts to an isomorphism on each fiber, there exists a unique smooth map $g : M \rightarrow GL_r(\mathbb{R})$ such that for any $p \in M$ and $v_p \in \mathbb{R}^r = E_p$

$$\Phi_p(v_p) = g_p v_p.$$

For all $p \in M$, $w_p \in \mathbb{R}^r = W_p$ we must then have

$$\Phi_p^{-1}(w_p) = g_p^{-1} w_p.$$

The inversion map $GL_r(\mathbb{R}) \rightarrow GL_r(\mathbb{R})$ is smooth, so g^{-1} is a smooth map $M \rightarrow GL_r(\mathbb{R})$, and we see that $\Phi^{-1} : F = \mathbb{R}_M^r \rightarrow \mathbb{R}_M^r = E$ is smooth. $\square$

Remark. The reader who finds this proof too concise might wish to consult the proof of proposition 2.1.2, where we proceed in a very similar manner, but with more details.

2.1.1 Sections of a bundle

Sections of a vector bundle are of such importance that vector bundles could be seen simply as means to define sections¹. Except for a few rare exceptions, all of our statements regarding vector bundles from here onwards will make mention of sections in some way.

Definition 2.1.3 ((Local) sections). Let U be an open subset of M . A **section of E over U** is a smooth map $s : U \rightarrow E$ such that $\pi \circ s = id_U$. We denote by $\Gamma(U, E)$ the space of such sections. An important special case is the space of sections over M , which we denote by $\Gamma(E)$. Pointwise addition and scalar multiplication naturally make $\Gamma(U, E)$ into a $C^\infty(U)$ -module.

Remark. Given a local section $s \in \Gamma(U, E)$ and a point $p \in U$, we will use indifferently s_p and $s(p)$ to denote the value of s at p .

¹This idea can be made precise by exhibiting an equivalence of categories between the category of vector bundles over a manifold M , and the category of locally free sheaves of $C^\infty(M)$ -modules over M , see [20] p. 268 for a brief overview.

Example 2.1.2. The following examples of vector bundle sections should be found in any introductory lectures on differential geometry:

- (i) **Smooth functions** on M , which are sections of the trivial bundle of rank 1, $M \times \mathbb{R} = \underline{\mathbb{R}}^1$. We use the special notation $C^\infty(M)$ to denote the space of smooth functions on M . More generally, a section of the trivial bundle $\underline{\mathbb{R}}^r$ of rank r is a family of r smooth functions $M \rightarrow \mathbb{R}$, in other words $\Gamma(\underline{\mathbb{R}}^r) = C^\infty(M)^r$.
- (ii) **Vector fields**, which are sections of the tangent bundle TM . We use the special notation $\mathfrak{X}(M)$ to denote the space of vector fields on M .
- (iii) **Differential k -forms**, which are sections of $\Lambda^k T^*M$. The space of such sections is denoted by $\Omega^k(M)$. In particular, we have $\Omega^1(M)$, the space of 1-forms which are sections of the **cotangent bundle** T^*M . Of special importance to us as well is the space of **top forms** $\Omega^{\text{top}}(M) = \Omega^m(M)$ ($m = \dim(M)$).
- (iv) Sections of $T_s^r M$, which are called **tensor fields of type (r, s)** on M . For these we will not use any special notation.
- (v) Symmetric covariant k -tensors, which are sections of $S^k T^*M$. No special notation for them either.

As a first application, we give a little criterion that will facilitate the task of checking whether a given map of vector bundles is smooth.

Proposition 2.1.2. *Let E, F be two vector bundles over M , and let $\Phi : E \rightarrow F$ be a fiber-preserving map such that for all $p \in M$, the restriction*

$$\Phi_p : E_p \rightarrow F_p$$

is linear. Then Φ is smooth, and thus a bundle morphism, if and only for any open subset $U \subseteq M$ and any smooth local section $s \in \Gamma(U, E)$, the section $\Phi \circ s : U \rightarrow F$ is smooth.

Proof. If Φ is smooth, then for any local section $s \in \Gamma(U, E)$, $\Phi \circ s$ is smooth as a composition of smooth maps.

For the converse, we proceed as follows: given a point $p \in M$, we can find an open neighborhood U of p such that there exist local trivializations

$$\begin{aligned}\phi : E_U &\rightarrow U \times \mathbb{R}^r = \underline{\mathbb{R}}_U^r \\ \psi : F_U &\rightarrow U \times \mathbb{R}^l = \underline{\mathbb{R}}_U^l\end{aligned}$$

of E and F respectively. Since ϕ and ψ are diffeomorphisms, we need only check that $\bar{\Phi} := \psi \circ \Phi \circ \phi^{-1} : \underline{\mathbb{R}}_U^r \rightarrow \underline{\mathbb{R}}_U^l$ is smooth to obtain that $\Phi : E_U \rightarrow F_U$ is smooth.

We denote the standard basis of $\mathbb{R}^m$ by $(b_i^0)_{i=1}^m$, and define $b_i^0 \in \Gamma(\mathbb{R}_U^m)$ to be the constant (and clearly smooth) section $U \rightarrow \mathbb{R}^m$ with value b_i^0 .

$\bar{\Phi}$ shares the properties of Φ , meaning that $\bar{\Phi}$ is a fiber-preserving map that acts linearly on fibers and sends smooth sections to smooth sections. In particular, for all $1 \leq j \leq r$, $\bar{\Phi} \circ b_j^0$ is smooth, so there must exist a family $(A_j^i)_{i=1}^l \in C^\infty(U)$ of smooth functions such that

$$\bar{\Phi} \circ b_j^0 = \sum_{i=1}^l A_j^i b_i^0.$$

We collect the A_j^i into a smooth matrix-valued map

$$A : U \rightarrow M_{l \times r}(\mathbb{R}), q \rightarrow (A_j^i(q))_{i,j}$$

and since $\bar{\Phi}$ acts linearly on fibers we have for any $q \in U$ and $v \in \mathbb{R}^r$

$$\bar{\Phi}_q v = A(q)v$$

which shows that $\bar{\Phi}$ is smooth.

From this we deduce that Φ is smooth over U , and since M can be covered by such opens, we conclude that Φ is smooth. $\square$

Local frames are families of local sections of a particular kind. Their importance lies in the fact that they are equivalent to trivializations. We will state their definition and main property, and expand on their role afterwards.

Definition 2.1.4 (Local frames). Let U be an open subset of M . A **local frame** of E over U is a family of local sections $(e_k)_{k=1}^r \subseteq \Gamma(U, E)$ such that $(e_k(p))_{k=1}^r$ is a linear basis of E_p for each $p \in U$.

Proposition 2.1.3. *Let E be a vector bundle over M , and let U be an open subset of M . There is a one-to-one correspondence between trivializations of E over U and local frames of E over U . To a local frame $(e_k)_{k=1}^r$ over U we associate the bundle isomorphism*

$$U \times \mathbb{R}^r \rightarrow \pi^{-1}(U), (p, (x^k)_{k=1}^r) \rightarrow \sum_{k=1}^r x^k e_k(p).$$

Conversely, to a bundle isomorphism $\phi : U \times \mathbb{R}^r \rightarrow \pi^{-1}(U)$ (or equivalently, a trivialization of E over U), we associate the local frame $(e_k)_{k=1}^r$ defined by

$$e_k(p) := \phi(p, b_k^0), \text{ for all } p \in U$$

where $(b_k^0)_{k=1}^r$ denotes the standard basis of $\mathbb{R}^r$.

Proof. We need only check that, given a local frame $(e_k)_{k=1}^r$ of E over U , the map

$$\Phi : U \times \mathbb{R}^r \rightarrow \pi^{-1}(U), (p, x) \rightarrow \sum_{k=1}^r x^k e_k(p)$$

is a bundle isomorphism, and this we obtain by combining propositions 2.1.1 and 2.1.2. $\square$

Local frames (or trivializations) are to vector bundles what bases are to vector spaces. And just like for vector spaces, fixing a local frame allows for the usage of coordinates. More precisely, once a local frame $e = (e_k)_{k=1}^r$ of E has been chosen over an open subset U , any local section $s \in \Gamma(U, E)$ can be written

$$s = \sum_{k=1}^r s^k e_k$$

for a unique family of smooth maps $s^k \in C^\infty(U)$, $k = 1, \dots, r$ which we call the components of s in the trivialization e . Thus, through the local frame e , a section s of E over U correspond to a smooth map $\underline{s} : U \rightarrow \mathbb{R}^r$. This is coherent with our interpretation of trivializations as isomorphisms with trivial bundles: sections of E_U become sections of the trivial bundle of rank r over U .

This interplay between bundle-valued object and their real-valued counterparts (i.e. components) will be a recurrent theme, thus we find it useful to fix some notation. We will often use underlining $\underline{(-)}$ to designate the components of a bundle-valued object and $\equiv$ to indicate the correspondence between such an object and its components. So in this instance, we would write $s \equiv \underline{s}$.

Just as two bases of a finite-dimensional vector space are related by a transition matrix, so are two overlapping trivializations of a given vector bundle, except that in the case of bundles, they are linked by a transition function: a matrix-valued smooth function. More formally, given local frames $e = (e_k)_{k=1}^r$ and $e' = (e'_k)_{k=1}^r$ of E over U and V respectively, and assuming $U \cap V \neq \emptyset$, there exists a unique smooth map

$$g_{UV} : U \cap V \rightarrow GL_r(\mathbb{R})$$

called the transition map (between e and e') such that

$$e'_l = \sum_{k=1}^r (g_{UV})_l^k e_k$$

for all $1 \leq l \leq r$. Given a section $s \in \Gamma(U \cap V, E)$ with components $\underline{s} = (s^k)_{k=1}^r$ and $\underline{s}' = (s'^k)_{k=1}^r$ in e and e' respectively we have

$$s = \sum_{k=1}^r s'^k e'_k = \sum_{k,l=1}^r (g_{UV})_k^l s'^k e_l$$

which shows that

$$\underline{s} = (g_{UV})\underline{s}'$$

if we see $\underline{s}$ and $\underline{s}'$ as $\mathbb{R}^r$ -valued functions, g_{UV} as a $GL_r(\mathbb{R})$ -valued function and use matrix multiplication.

We now introduce the perhaps not-so-widespread notion of germ. Though their definition might seem cumbersome at first glance, germs are an excellent way to describe local phenomena in a rigorous and formal fashion.

Definition 2.1.5 (Germs). Let p be a point of M , let U, V be open neighborhoods of p and let $s \in \Gamma(U, E)$, $t \in \Gamma(V, E)$ be local sections of E . We say that s and t have the same germ at p , written $s \sim_p t$, if there exists an open W with $p \in W \subseteq U \cap V$ such that

$$s|_W = t|_W.$$

One easily sees that $\sim_p$ defines an equivalence relation on the set

$$\{s \in \Gamma(U, E) | U \text{ is an open nbd. of } p\}$$

of local sections defined around p . For such a local section $s \in \Gamma(U, E)$ the equivalence class of s for $\sim_p$, denoted by $[s]_p$, is called the **germ** of s at p . In other words an element of $[s]_p$ is a local section t defined on an open neighborhood V of p such that $s|_W = t|_W$ for some open W with $p \in W \subseteq U \cap V$. The set of germs of E at p (i.e. the set of equivalence classes for $\sim_p$) is denoted by $\Gamma_p(E)$, so

$$\Gamma_p(E) := \{[s]_p | s \in \Gamma(U, E) \text{ with } U \text{ an open nbd. of } p\}.$$

$\Gamma_p(E)$ is sometimes called the **stalk** of E at p .

Since scalar multiplication and pointwise addition are compatible with $\sim_p$, $\Gamma_p(E)$ naturally possesses the structure of an $\mathbb{R}$ -module, in fact of a $C_p^\infty(M)$ module, where $C_p^\infty(M)$ denotes the space of germs of *smooth functions* at p .

The following proposition tells us that a vector bundle has 'many' global sections, enough that most properties of a vector bundle can be understood simply by looking at its global section. This follows essentially from the existence of partitions of unity on a smooth manifold.

Proposition 2.1.4. *Let p be a point of M and let s be a local section of E defined on an open neighborhood of p . Then there exists a global section $\bar{s} \in \Gamma(E)$ such that s and $\bar{s}$ have the same germ at p , i.e. there exists an open V with $p \in V \subseteq U$ such that $s|_V = \bar{s}|_V$.*

Equivalently, one might state that all germs at p can be obtained as germs of global sections of E . More formally

$$\Gamma_p(E) = \{[s]_p | s \in \Gamma(E)\}.$$

Proof. Take a point $p \in M$ and a local section $s \in \Gamma(U, E)$ with U an open neighbourhood of p . We can always find an open neighbourhood V of p such that the closure $\bar{V}$ of V is included in U (for example choose a small disk around p in a coordinate chart). We can then choose a partition of unity (ρ, ρ') subordinate to the open cover $(U, M \setminus \bar{V})$ of M . The support of ρ is contained in U , so we can define the smooth global section

$$\rho s : M \rightarrow E, q \mapsto \begin{cases} \rho(q)s_q, & \text{if } q \in U \\ 0, & \text{otherwise.} \end{cases}$$

Since ρ' vanishes on $\bar{V}$, for every $q \in \bar{V} \subseteq U$ we have $\rho(q) = 1$ which implies

$$(\rho s)(p) = \rho(p)s_p = s_p$$

and we see that $\rho s \sim_p s$. □

Corollary 2.1.1. *Let E be a vector bundle over M . Then for any $p \in M$ and $e \in E_p$ there exists a global section $s \in \Gamma(E)$ such that $s(p) = e$.*

Proof. Take a point $p \in M$ and an element $e \in E_p$. Choose a local trivialization (U, ϕ) of E over an open neighborhood U of p . Then clearly, there exists a section s of E over U such that $s(p) = e$. Proposition 2.1.4 then guarantees the existence of a global section $\tilde{s}$ such that $\tilde{s} \sim_p s$, in particular we must have

$$\tilde{s}(p) = s(p) = e.$$

□

As an immediate application of proposition 2.1.4, we can improve the criterion we gave in proposition 2.1.2.

Proposition 2.1.5. *Let E, F be two vector bundles over M , and let $\Phi : E \rightarrow F$ be a fiber-preserving map such that for all $p \in M$, the restriction*

$$\Phi_p : E_p \rightarrow F_p$$

is linear. Then Φ is smooth, and thus a bundle morphism, if and only for any smooth global section $s \in \Gamma(E)$, the section $\Phi \circ s : M \rightarrow F$ is smooth.

Proof. Once again, if Φ is smooth then so is the composition $\Phi \circ s$ for any $s \in \Gamma(E)$.

We will obtain the converse by proving that Φ verifies the criterion given in proposition 2.1.2. So let $s \in \Gamma(U, E)$ be a smooth local section of E , and let p be a point of U . By

proposition 2.1.4, there exists an open neighborhood V of p contained in U and a global section $\tilde{s} \in \Gamma(E)$ such that $s|_V = \tilde{s}|_V$. By hypothesis, $\Phi \circ \tilde{s} : M \rightarrow F$ is smooth, so

$$(\Phi \circ s)|_V = (\Phi \circ \tilde{s})|_V$$

is smooth as well. Since this holds for any $p \in U$, this shows that $\Phi \circ s$ is smooth. $\square$

2.1.2 Local operators and bundle morphisms

As stated at the end of the previous section, global sections are numerous and flexible enough to register local phenomena, which makes them great tools of investigation when studying vector bundles. This will often be done by looking at maps that send global sections to other global sections (possibly of distinct bundles).

With this in mind, our purpose in this section is to present two families of such maps:

1. Local operators, which as their name indicate only witness what happens at the local level (that is, at the level of germs). 'Differential operators', that is operators that generalize the notion of derivative in a certain sense, often belong to this class², and most examples of local operators we will encounter will be of this kind.
2. Pointwise operators, which only witness what happens at the level of fibers. Most importantly, they will be shown to be equivalent both to morphisms of $C^\infty(M)$ -modules and bundle morphisms.

We begin with the notion of a local operator *at a point*.

Definition 2.1.6 (Local operator at a point). Let E, F be two vector bundles over p . A **local operator $E \rightarrow F$ at a point p of M** is an $\mathbb{R}$ -linear map

$$\alpha_p : \Gamma_p(E) \rightarrow F_p.$$

that goes from the stalk of E at p into the fiber of F above p .

As we'll see, proposition 2.1.4 allows us to see a local operator at a point as an operator acting on global sections.

Proposition 2.1.6. *Let $\alpha : \Gamma(E) \rightarrow F_p$ be an $\mathbb{R}$ -linear map compatible with $\sim_p$, meaning that $\alpha(s) = \alpha(t)$ for any pair $s, t \in \Gamma(E)$ such that $s \sim_p t$. Then there exists a unique local operator*

$$\alpha_p : \Gamma_p(E) \rightarrow F_p$$

²A clear definition of 'differential' operators exists and makes this idea precise. This can be done e.g. through the theory of jet spaces, see [25].

at p such that

$$\alpha(s) = \alpha_p([s]_p)$$

for all $s \in \Gamma(E)$.

Proof. Let $\mathfrak{s}_p \in \Gamma_p(E)$ be a germ of E at p , by proposition 2.1.4 there exists $s \in \Gamma(E)$ such that $[s]_p = \mathfrak{s}_p$, so we can set

$$\alpha_p(\mathfrak{s}_p) := \alpha(s).$$

Since α is compatible with $\sim_p$, the choice of s is inconsequential and we obtain a well-defined map

$$\alpha_p : \Gamma_p(E) \rightarrow F_p.$$

The $\mathbb{R}$ -linearity of α_p , the fact that $\alpha_p([s]_p) = \alpha(s)$ for any $s \in \Gamma(E)$, as well as the uniqueness of α_p are quite clear, and are left as exercises to the reader. $\square$

Remark. Conversely, if $\alpha_p : \Gamma_p(E) \rightarrow F_p$ is a local operator at p , then the map

$$\alpha : \Gamma(E) \rightarrow F_p, s \mapsto \alpha_p([s]_p)$$

is compatible with $\sim_p$. Clearly this process and the one described in proposition 2.1.6 are inverse to each other, so that we have a one-to-one correspondence between local operators $\Gamma_p(E) \rightarrow F_p$ and $\mathbb{R}$ -linear maps $\Gamma(E) \rightarrow F_p$ compatible with $\sim_p$.

Having defined local operators at a point, we move on to local operators acting on global sections, as well as pointwise operators.

Definition 2.1.7 (Local and pointwise operators). Let E, F be two vector bundles over a smooth manifold M . An $\mathbb{R}$ -linear map $\alpha : \Gamma(E) \rightarrow \Gamma(F)$ is:

- (i) A **local operator** if whenever a section $s \in \Gamma(E)$ vanishes on an open subset $U \subseteq M$, then $\alpha(s)$ also vanishes on U .
- (ii) **Pointwise** if whenever a section $s \in \Gamma(E)$ vanishes at a point $p \in M$, then $\alpha(s)$ also vanishes at p .

More generally, if $E_1, \dots, E_k$ and F are vector bundles over M , an $\mathbb{R}$ -multilinear map

$$\alpha : \Gamma(E_1) \times \dots \times \Gamma(E_k) \rightarrow \Gamma(F)$$

is said to be pointwise if for any point $p \in M$ and any family of section $s_j \in \Gamma(E_j)$, $\alpha(s_1, \dots, s_k)$ vanishes at p as soon as one of the sections $s_1, \dots, s_k$ vanishes at p .

The distinctive feature of local and pointwise operators (which is essentially a consequence of proposition 2.1.4) is that even though they are defined as operators acting on *global* sections, they can be restricted to open subsets (for local operators) or even fibers (for pointwise operators).

Proposition 2.1.7 (Restriction of local and pointwise operators). *Let E, F be two bundles over M and let $\alpha : \Gamma(E) \rightarrow \Gamma(F)$ be an $\mathbb{R}$ -linear map.*

(i) *If α is a local operator, then for any $p \in M$ there exists a unique local operator*

$$\alpha_p : \Gamma_p(E) \rightarrow F_p$$

at p such that for all $s \in \Gamma(E)$

$$\alpha(s)_p = \alpha_p([s]_p).$$

(ii) *If α is a local operator, then for any open subset U of M , there exists a unique $\mathbb{R}$ -linear map $\alpha|_U : \Gamma(U, E) \rightarrow \Gamma(U, F)$ such that for all $s \in \Gamma(E)$*

$$(\alpha|_U)(s|_U) = \alpha(s)|_U.$$

$\alpha|_U$ is then also a local operator, and for any point $p \in U$, $\alpha|_U$ and α induce the same local operator $\Gamma_p(E) \rightarrow F_p$ at p .

(iii) *If α is a pointwise operator, then for any point $p \in M$, there exists a unique $\mathbb{R}$ -linear map $\alpha_p : E_p \rightarrow F_p$ such that for all $s \in \Gamma(E)$*

$$\alpha(s)_p = \alpha_p(s_p).$$

(iv) *More generally, if*

$$\alpha : \Gamma(E_1) \times \dots \times \Gamma(E_k) \rightarrow \Gamma(F)$$

is pointwise, then for any point $p \in M$, there exists a unique $\mathbb{R}$ -multilinear map $\alpha_p : E_{1,p} \times \dots \times E_{k,p} \rightarrow F_p$ such that for all $s_j \in \Gamma(E_j)$

$$\alpha(s_1, \dots, s_k)_p = \alpha_p(s_1(p), \dots, s_k(p)).$$

Proof. (i) For any $p \in M$, the map

$$\Gamma(E) \rightarrow F_p, s \rightarrow \alpha(s)_p$$

is compatible with $\sim_p$, so the statement is an immediate consequence of proposition 2.1.6.

(ii) The map $\alpha|_U$ must be given by

$$(\alpha|_U)(s)_p = \alpha_p([s]_p)$$

for all $s \in \Gamma(U, E)$. Conversely, we can use this as a definition for $\alpha|_U$ and check that the resulting map has the desired properties.

(iii) For $p \in M$ and $e \in E_p$, we know by corollary 2.1.1 that there exists a global section $s \in \Gamma(E)$ such that $s(p) = e$, so we can set

$$\alpha_p(e) := \alpha(s)_p.$$

If $s' \in \Gamma(E)$ is another global section verifying $s'(p) = e$ then $(s - s')$ vanishes at p , and we get

$$\alpha(s)_p - \alpha(s')_p = \alpha(s - s')_p = 0$$

since α is pointwise. Thus the map

$$\alpha_p : E_p \rightarrow F_p$$

is well-defined. Its linearity follows from the linearity of α .

(iv) is treated analogously. □

We are now in a position to fulfill our stated goals: we give characterizations of pointwise operators, first as $C^\infty(M)$ -linear maps, then as bundle morphisms, and give a sufficient condition for an operator to be local by introducing the Leibniz rule (which should be interpreted as a way to generalize first order differential operators).

Proposition 2.1.8. *Let E, F be two bundles over M and let $\alpha : \Gamma(E) \rightarrow \Gamma(F)$.*

- (i) *Given a vector field $X \in \mathfrak{X}(M)$, an $\mathbb{R}$ -linear map $\alpha_X : \Gamma(E) \rightarrow \Gamma(F)$ verifies **the Leibniz rule (for X)** if for any $f \in C^\infty(M)$ and $s \in \Gamma(E)$ we have*

$$\alpha_X(fs) = (Xf)s + f\alpha_X(s).$$

Such an operator is always local.

- (ii) *An $\mathbb{R}$ -linear map $\alpha : \Gamma(E) \rightarrow \Gamma(F)$ is pointwise if and only if it is $C^\infty(M)$ -linear.*
 (iii) *Once again, we can extend to the multilinear setting: assuming that $E_1, \dots, E_k$ and F are vector bundles over M , an $\mathbb{R}$ -multilinear map*

$$\alpha : \Gamma(E_1) \times \dots \times \Gamma(E_k) \rightarrow \Gamma(F)$$

is pointwise if and only if it is $C^\infty(M)$ -multilinear.

Proof. (i) Assume that s vanishes over an open $U \subseteq M$, and let $p \in U$. Using partitions of unity, we can find an open neighborhood V of p contained in U and a smooth map $f \in C^\infty(M)$ that vanishes over V and takes value 1 outside of U . Then $s = fs$, so by the Leibniz rule we have

$$\alpha_X(s) = \alpha_X(fs) = (Xf)s + f\alpha_X(s)$$

and since Xf and f both vanish over V , so must $\alpha_X(s)$. As this holds for every $p \in M$, we obtain $\alpha_X(s)|_U = 0$.

(ii) Assume that α is $C^\infty(M)$ -linear and that $s \in \Gamma(E)$ vanishes at p . Choose a local frame $(e_k)_{k=1}^r$ of E defined on an open neighborhood U of p , and choose a smooth function $f \in C^\infty(M)$ with $\text{supp } f \subseteq U$ and $f(p) = 1$ (e.g. using partitions of unity). Let $(s^k)_{k=1}^r \subseteq C^\infty(U)$ be the unique smooth functions such that

$$s|_U = \sum_{k=1}^r s^k e_k.$$

Since $s(p) = 0$ we have $s^k(p) = 0$ for every $1 \leq k \leq r$.

Using the $C^\infty(M)$ -linearity of α , we then obtain

$$\begin{aligned} \alpha(s) &= \alpha[(1 - f^2)s + f^2s] = (1 - f^2)\alpha(s) + \alpha[f^2s] \\ &= (1 - f^2)\alpha(s) + \alpha\left(\sum_{k=1}^r (fs^k)(fe_k)\right) \\ &= (1 - f^2)\alpha(s) + \sum_{k=1}^r (fs^k)\alpha(fe_k) \end{aligned}$$

which yields $\alpha(s)_p = 0$ when evaluated at p , since $1 - f^2(p) = 1 - 1 = 0$ and $f(p)s^k(p) = 1 \cdot 0 = 0$ for all $1 \leq k \leq r$.

Conversely, if α is pointwise then, by proposition 2.1.7, for every $p \in M$ there exists a unique $\mathbb{R}$ -linear map

$$\alpha_p : E_p \rightarrow F_p$$

such that $\alpha(s)_p = \alpha_p(s_p)$ for all $s \in \Gamma(E)$. Consequently, for $f \in C^\infty(M)$ and $s \in \Gamma(E)$ we have that for every $p \in M$

$$\alpha(fs)_p = \alpha_p(f(p)s_p) = f(p)\alpha_p(s_p)$$

which shows that $\alpha(fs) = f\alpha(s)$.

(iii) is treated similarly. □

Proposition 2.1.9. *Let E, F be two vector bundles over M . There is an one-to-one correspondence between bundle morphisms from $E \rightarrow F$ over M and $C^\infty(M)$ -linear maps $\Gamma(E) \rightarrow \Gamma(F)$. Through this correspondence, the bundle morphism $\Phi : E \rightarrow F$ is identified with the map of sections sending $s \in \Gamma(E)$ to $\Phi \circ s : M \rightarrow F$.*

Proof. Given a bundle morphism $\Phi : E \rightarrow F$, the map

$$\Gamma(E) \rightarrow \Gamma(F), s \mapsto \Phi \circ s$$

is clearly $C^\infty(M)$ -linear.

Conversely, let $\alpha : \Gamma(E) \rightarrow \Gamma(F)$ be a $C^\infty(M)$ -linear map. By proposition 2.1.8, α is pointwise, so for every $p \in M$ there must exist a unique $\mathbb{R}$ -linear map

$$\Phi_p : E_p \rightarrow F_p$$

such that $\alpha(s)_p = \Phi_p(s_p)$ for all $s \in \Gamma(E)$ (proposition 2.1.7). Clearly $\alpha(s) = \Phi \circ s$ for every $s \in \Gamma(E)$, in particular Φ sends smooth sections to smooth sections, so Φ is a bundle morphism by proposition 2.1.5. The fact that Φ is the only bundle morphism such that $\alpha(s) = \Phi \circ s$ for every $s \in \Gamma(E)$ is an obvious consequence of the uniqueness in proposition 2.1.7. $\square$

This last, very beautiful result is perhaps the clearest sign that a vector bundle is 'determined' by its sections. A similar interpretation for $C^\infty(M)$ -multilinear maps of sections will be made available once we introduce tensor bundles.

2.1.3 Vector subbundles

We dedicate this very brief section to vector subbundles, the correct notion of subobject for vector bundles, and the bundle equivalent of linear subspaces. The key takeaway here is the **subbundle criterion** which provides a very straightforward characterization of vector subbundles in terms of local frames.

Definition 2.1.8 (Vector subbundles). Let E be a vector bundle on M . A **(vector) subbundle** of E is a vector bundle $\rho : F \rightarrow M$ such that

- (i) F is an embedded submanifold of E .
- (ii) The inclusion map $i : F \rightarrow E$ is a bundle morphism.

Remark. A smooth section of F is then simply a smooth section of E that takes value in F .

Theorem 2.1.1 (Vector subbundle criterion). *For each $p \in M$, let F_p be a k -dimensional linear subspace of E_p . Then the set $F := \bigsqcup_{p \in M} F_p$ equipped with the obvious projection is a rank k subbundle of E if and only if for each $p \in M$ there exists a family $e_1, \dots, e_k$ of (linearly independent) smooth sections of E defined on an open neighborhood U of p such that for each $q \in U$, F_q is spanned by $e_1(q), \dots, e_k(q)$.*

Proof. That the criterion is necessary is clear enough. To see that it is sufficient, the idea is to complete a local frame of F into a local frame of E (similar to how a linearly independent family in a vector space can be completed into a basis), possibly restricting to a smaller open in the process. Details can be found in [19] (lemma 10.32, p.265). $\square$

2.1.4 Tensor bundles and bundle-valued forms

All the common operations one can use to obtain new vector spaces from old ones have their analogues for vector bundles. Of main concern to us, this holds for the operations we introduced in chapter 1, namely the tensor and wedge product (as well as Lin). Starting from a vector space operation (such as the tensor product), one obtains the corresponding bundle operation by following this two-step procedure:

1. Constructing the fibers of our new vector bundle is straightforward: we simply apply the "vector space operation" fiberwise.
2. After step 1, we have the underlying set of our new bundle. Endowing this set with the appropriate smooth structure requires a bit more work. We obtain an atlas through a judicious choice of local trivializations. That is, we once again use our "vector space operation" to combine smooth local frames of our "old bundles", which gives us "local frames" for our new bundle (between quotes since our new bundle does not yet have a smooth structure, so that these local frames are only set-theoretic maps for now). These local frames induce bijections between (restrictions of) our new bundle and (local) trivial bundles. We endow our new bundle with a smooth structure by requiring that these bijections be diffeomorphisms.

Since our local frames were constructed using smooth local frames and a "natural vector space operation"³, the transition maps between these local trivializations are smooth, and our smooth structure is well-defined.

The theorem below is what will allow us to make our informal discussion concrete.

Theorem 2.1.2 (Constructing vector bundles). *Let M be a smooth manifold. For each $p \in M$, let E_p be a real vector space of dimension r . We define $E := \bigsqcup_{p \in M} E_p$ and the obvious projection $\pi : E \rightarrow M$. Let $(U_\alpha)_{\alpha \in A}$ be an open cover of M . For each $\alpha \in A$, let $e^\alpha := (e_k^\alpha)_{k=1}^r$ be a family of local sections $U_\alpha \rightarrow E$ of π (for now π and the e_k^α are only set maps), such that for each $p \in U_\alpha$, $(e_k^\alpha(p))_{k=1}^r$ is a linear basis of E_p .*

For each $\alpha, \beta \in A$ with $U_{\alpha\beta} := U_\alpha \cap U_\beta \neq \emptyset$, there exists a unique map $g_{\alpha\beta} : U_{\alpha\beta} \rightarrow GL_r(\mathbb{R})$ such that

$$e_k^\alpha = \sum_{l=1}^r (g_{\alpha\beta})_l^k e_l^\beta$$

over $U_{\alpha\beta}$ for all $1 \leq k \leq r$. If the map $g_{\alpha\beta} : U_{\alpha\beta} \rightarrow GL_r(\mathbb{R})$ is smooth for each pair $\alpha, \beta \in A$ with $U_{\alpha\beta} \neq \emptyset$, then there exists a unique smooth structure on E for which $\pi : E \rightarrow M$ is a vector bundle of rank r , and such that for each $\alpha \in A$, e^α is a smooth local frame of E . Note that we also require the vector space structure on the fibers of E to be the one provided in the beginning of our (admittedly long) statement.

³To be a little less vague: the precise notion of "natural vector space operation" we require to ensure that our recipe works well is that of a **smooth functor**.

Proof. See [19] (lemma 10.6 'Vector Bundle Chart Lemma', p.253). $\square$

Definition 2.1.9 (Operations on vector bundles). Let E, F be vector bundles over M . In what follows we describe several natural operations one can apply to E and F in order to obtain new vector bundles. For each of these operations, we describe the underlying set of the new bundle, provide characterizations for the smooth structure of the new bundle as well as its smooth sections, and methods to obtain local frames of our new bundle from local frames of E and F . Additionally, when pertinent, we describe analogues to the tensor algebra operations we introduced in chapter 1.

- (i) The **direct sum** $E \oplus F$. As a set $E \oplus F := \bigsqcup_{p \in M} (E_p \oplus F_p)$. To make $E \oplus F$ into a smooth vector bundle, we endow it with the unique smooth structure such that given $e \in \Gamma(E)$ and $f \in \Gamma(F)$,

$$e \oplus f : M \rightarrow E \oplus F, p \rightarrow e(p) \oplus f(p)$$

is a smooth section of $E \oplus F$. The existence and uniqueness of such a smooth structure is guaranteed by theorem 2.1.2.

A smooth section of $E \oplus F$ is always the direct sum of a smooth section of E with a smooth section of F . In other words, we have a canonical isomorphism

$$\Gamma(E \oplus F) \simeq \Gamma(E) \oplus \Gamma(F).$$

For an open subset U of M , a local frame $e = (e_k)_{k=1}^r$ of E over U and a local frame $f = (f_k)_{k=1}^s$ of F over U , the family

$$e \oplus f := ((e_k \oplus 0)_{k=1}^r, (0 \oplus f_k)_{k=1}^s)$$

constitutes a local frame of $E \oplus F$.

- (ii) The **dual** E^* . As a set $E^* := \bigsqcup_{p \in M} E_p^*$. The smooth structure with which we equip E^* is uniquely characterized by following property: a global section $\alpha : M \rightarrow E^*$ is smooth if and only if for all $e \in \Gamma(E)$,

$$\alpha(e) : M \rightarrow \mathbb{R}, p \rightarrow \alpha_p(e_p)$$

is smooth.

There is a one-to-one correspondence between smooth sections of E^* and $C^\infty(M)$ -linear maps

$$\Gamma(E) \rightarrow C^\infty(M).$$

In other words, $\Gamma(E^*)$ is the dual of $\Gamma(E)$ as a $C^\infty(M)$ -module. To a section α of E^* , we associate the map

$$\alpha(-) : \Gamma(E) \rightarrow C^\infty(M), e \rightarrow \alpha(e).$$

Given a local frame $e = (e_k)_{k=1}^r$ of E over an open subset U , if we define $(e_*^k(p))_{k=1}^r$ to be the basis dual to $(e_k(p))_{k=1}^r$ for all $p \in U$, then $(e_*^k)_{k=1}^m$ is a local frame of E^* , which we call the coframe of $(e_k)_{k=1}^m$.

- (iii) The **tensor product** $E \otimes F$. As a set $E \otimes F := \bigsqcup_{p \in M} (E_p \otimes F_p)$, and we endow $E \otimes F$ with the unique smooth structure such that for any $e \in \Gamma$ and $f \in \Gamma(F)$,

$$e \otimes f : M \rightarrow E \otimes F, p \rightarrow e(p) \otimes f(p)$$

is a smooth section of $E \otimes F$.

Given local frames $(e_k)_{k=1}^r, (f_l)_{l=1}^s$ of E and F over U , the family

$$e \otimes f := (e_k \otimes f_l)_{1 \leq k \leq r, 1 \leq l \leq s}$$

is a local frame of $E \otimes F$ over U .

In general, a global section of $E \otimes F$ will not be the tensor product of a global section of E with a global section of F , nor even a finite sum of such tensor products. However, as evidenced by the local frame we exhibited in the previous paragraph, a global section of $E \otimes F$ will always coincide *locally* with a finite sum of tensor products of sections of E and F . More formally, for any $t \in \Gamma(E \otimes F)$ and any $p \in M$, there exists $e_1, \dots, e_k \in \Gamma(E)$ and $f_1, \dots, f_k \in \Gamma(F)$ such that

$$t \sim_p \sum_{i=1}^k e_i \otimes f_i.$$

Alternatively, the "multilinear map" approach to tensors outlined in theorem 1.1.1 combined with the characterization of pointwise maps via $C^\infty(M)$ -linearity (proposition 2.1.8) gives us a very fruitful point of view on sections of $E \otimes F$: there is a one-to-one correspondence between smooth sections of $E \otimes F$ and $C^\infty(M)$ -multilinear maps

$$\Gamma(E^*) \times \Gamma(F^*) \rightarrow C^\infty(M).$$

To $e \otimes f$ ($e \in \Gamma(E), f \in \Gamma(F)$) we associate the map

$$\Gamma(E^*) \times \Gamma(F^*) \rightarrow C^\infty(M), (\alpha, \beta) \rightarrow \alpha(e)\beta(f).$$

(Note that we could have chosen this correspondence to characterize the smooth structure of $E \otimes F$).

- (iv) Iterating we construct $T_s^r E := \underbrace{E \otimes \dots \otimes E}_{r \text{ times}} \otimes \underbrace{E^* \otimes \dots \otimes E^*}_{s \text{ times}}$. Reproducing the previous point, given a local frame $e = (e_k)_{k=1}^m$ of E with local coframe $(e_*^k)_{k=1}^m$, the set $(e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e_*^{j_1} \otimes \dots \otimes e_*^{j_s})_{i_1 \dots i_r, j_1, \dots, j_s}$ forms a local frame of $T_s^r E$. As in the previous point, this implies that a section of $T_s^r E$ will always coincide locally with

a finite sum of tensor products of sections of E and E^* . The components of T in that local frame are denoted by $T_{j_1 \dots j_s}^{i_1 \dots i_r}$.

We can once again invoke the "multilinear maps" point of view for tensors, and this time it tells us that there is a one-to-one correspondence between global sections of $T_s^r E$ and $C^\infty(M)$ -multilinear maps

$$\Gamma(E^*)^r \times \Gamma(E)^s \rightarrow C^\infty(M).$$

To $e_1 \otimes \dots \otimes e_r \otimes \alpha^1 \otimes \dots \otimes \alpha^s$ ($e_j \in \Gamma(E)$, $\alpha^j \in \Gamma(E^*)$) we associate the map

$$\begin{aligned} \Gamma(E^*)^r \times \Gamma(E)^s &\rightarrow C^\infty(M), \\ (\beta^1, \dots, \beta^r, f_1, \dots, f_s) &\rightarrow \beta^1(e_1) \dots \beta^r(e_r) \alpha^1(f_1) \dots \alpha^s(f_s). \end{aligned}$$

Very important as well is the fact that the operations we defined in chapter 1 for the tensor algebra of a vector space can be applied fiberwise, which gives us maps of smooth sections. In that way, given a section T of $T_s^r E$ and a section S of $T_l^k E$ we can construct the section $T \otimes S$ which is defined by

$$(T \otimes S)(p) = T(p) \otimes S(p)$$

for all $p \in M$. Similarly, given $1 \leq i_1 < i_2 < \dots < i_k \leq r$, $1 \leq j_1 < j_2 < \dots < j_k \leq s$ and a section T of $T_s^r E$, we can define the section $C_{j_1 \dots j_k}^{i_1 \dots i_k} T$ of $T_{s-k}^{r-k} E$ by contraction:

$$(C_{j_1 \dots j_k}^{i_1 \dots i_k} T)(p) = C_{j_1 \dots j_k}^{i_1 \dots i_k}(T_p)$$

for all points $p \in M$. The symmetrization Sym , the alternation Alt as well as the $\cdot$ action of linear isomorphism and endomorphisms (definitions 1.1.4 and 1.1.5) can be treated similarly.

- (v) The space of **linear maps** $\text{Lin}(E, F)$. As a set $\text{Lin}(E, F) := \bigsqcup_{p \in M} \text{Lin}(E_p, F_p)$, where $\text{Lin}(E_p, F_p)$ denotes the space of linear maps $E_p \rightarrow F_p$. For each $p \in M$ there is a canonical isomorphism $\text{Lin}(E_p, F_p) \simeq F_p \otimes E_p^*$, and together they give us a bijection $\text{Lin}(E, F) \simeq F \otimes E^*$. We endow $\text{Lin}(E, F)$ with a smooth structure by requiring that this bijection be a diffeomorphism. In that way the fiberwise linear isomorphisms $\text{Lin}(E_p, F_p) \simeq F_p \otimes E_p^*$ assemble into a bundle isomorphism $\text{Lin}(E, F) \simeq F \otimes E^*$.

Equivalently the smooth structure on $\text{Lin}(E, F)$ is uniquely characterized by the following property: a section $A : M \rightarrow \text{Lin}(E, F)$ is smooth if and only if for all $e \in \Gamma(E)$

$$A(e) : M \rightarrow F, p \rightarrow A_p(e_p)$$

is a smooth section of F .

In this case, there is a one-to-one correspondence between global sections of $\text{Lin}(E, F)$

and $C^\infty(M)$ -linear maps

$$\Gamma(E) \rightarrow \Gamma(F)$$

which are equivalent to bundle morphisms $E \rightarrow F$ (proposition 2.1.9).

An important special case is $\text{End}(E) := \text{Lin}(E, E) \simeq E \otimes E^*$.

- (vi) The **exterior or wedge power** $\Lambda^k E^*$. As a set $\Lambda^k E^* := \bigsqcup_{p \in M} \Lambda^k E_p^*$. We endow $\Lambda^k E^*$ with the unique smooth structure such that for all $\alpha^1, \dots, \alpha^k \in \Gamma(E^*)$

$$\alpha^1 \wedge \dots \wedge \alpha^k : M \rightarrow \Lambda^k E^*, p \mapsto \alpha^1(p) \wedge \dots \wedge \alpha^k(p)$$

is a smooth section of $\Lambda^k E^*$.

There is a one-to-one correspondence between global sections of $\Lambda^k E^*$ and skew-symmetric $C^\infty(M)$ -multilinear maps

$$\Gamma(E)^k \rightarrow C^\infty(M).$$

Given a section α of $\Lambda^k E^*$, and $e_1, \dots, e_k \in \Gamma(E)$, the smooth function $\alpha(e_1, \dots, e_k)$ is defined by

$$\alpha(e_1, \dots, e_k)_p := \alpha_p(e_{1,p}, \dots, e_{k,p})$$

for all $p \in M$.

Given a local frame $(e_k)_{k=1}^r$ of E over U with coframe $(e_*^k)_{k=1}^r$, a section α of $\Lambda^k E^*$ over U can be uniquely written

$$\alpha = \sum_{1 \leq i_1 < \dots < i_k \leq r} \alpha_{i_1 \dots i_k} e_*^{i_1} \wedge \dots \wedge e_*^{i_k}$$

with $\alpha_{i_1 \dots i_k} = \alpha(e_{i_1}, \dots, e_{i_k}) \in C^\infty(U)$. Applying the subbundle criterion (theorem 2.1.1) to $(e_*^{i_1} \wedge \dots \wedge e_*^{i_k})_{i_1 < \dots < i_k}$, we see that $\Lambda^k E^*$ is a **vector subbundle** of $T_k E$.

Operating fiberwise, for $\alpha \in \Lambda^k E^*$ and $\beta \in \Lambda^l E^*$ we can define $\alpha \wedge \beta \in \Lambda^{k+l} E^*$.

Proof that the constructions are well-defined. We will treat the case of $T_s^r E$ since it is the most important for our purposes.

We define the set

$$T_s^r E := \bigsqcup_{p \in M} T_s^r E_p$$

And endow $T_s^r E$ with a vector bundle structure using theorem 2.1.2, and local frames of E : given a local frame $e = (e_k)_{k=1}^r$ of E over an open subset $U \subseteq M$ with coframe $e_* = (e_*^k)_{k=1}^r$, for every $p \in U$ the family $(e_{i_1}(p) \otimes \dots \otimes e_{i_r}(p) \otimes e_*^{j_1}(p) \otimes \dots \otimes e_*^{j_s}(p))_{i_1 \dots i_r, j_1, \dots, j_s}$ forms a basis of $T_s^r E_p$. Furthermore, given another local frame $\bar{e} = (\bar{e}_k)_{k=1}^r$ of E over an

open subset $V \subseteq M$, if we denote the transition map between e and $\bar{e}$ by $g : U \cap V \rightarrow GL_r(\mathbb{R})$ so that

$$\bar{e}_l = \sum_{k=1}^r g_l^k e_k$$

we obtain

$$\bar{e}_{i_1} \otimes \dots \otimes \bar{e}_{i_r} \otimes \bar{e}_*^{j_1} \otimes \dots \otimes \bar{e}_*^{j_s} = \sum_{t_1, \dots, t_r, q_1, \dots, q_s} g_{i_1}^{t_1} \dots g_{i_r}^{t_r} (g^{-1})_{q_1}^{j_1} \dots (g^{-1})_{q_s}^{j_s} e_{t_1} \otimes \dots \otimes e_{t_r} \otimes e_*^{q_1} \otimes \dots \otimes e_*^{q_s}$$

and since the components $g_{i_1}^{t_1} \dots g_{i_r}^{t_r} (g^{-1})_{q_1}^{j_1} \dots (g^{-1})_{q_s}^{j_s}$ are smooth, by theorem 2.1.2 we see that $T_s^r E$ admits a unique vector bundle structure such that for any local frame $(e_k)_{k=1}^r$ of E , $(e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e_*^{j_1} \otimes \dots \otimes e_*^{j_s})_{i_1, \dots, i_r, j_1, \dots, j_s}$ is a local frame of $T_s^r E$.

(Notice that we now have a bundle structure for E^* as the special case $E^* = T_1 E$).

Clearly, if we equip $T_s^r E$ with this structure, then for any $e_i \in \Gamma(E)$ and $\alpha^j \in \Gamma(E^*)$, $e_1 \otimes \dots \otimes e_r \otimes \alpha^1 \otimes \dots \otimes \alpha^s$ is a smooth section of $T_s^r E$ (simply by observing that the components of such a section in the local frames of $T_s^r E$ given above are smooth). Conversely, if we endow $T_s^r E$ with a smooth structure possessing this property, then given *local* sections $e_i \in \Gamma(U, E)$ and $\alpha^j \in \Gamma(U, E^*)$ over an open U , we can show that $e_1 \otimes \dots \otimes e_r \otimes \alpha^1 \otimes \dots \otimes \alpha^s$ is a smooth section of $T_s^r E$ over U using proposition 2.1.4. This implies that the local frames characterizing the first smooth structure on $T_s^r E$ (i.e. those induced by local frames of E) are also local frames for this new smooth structure on $T_s^r E$, so that these two smooth structures are in fact equal.

Lastly, let us prove the characterization of smooth sections of $T_s^r E$. Consider a $C^\infty(M)$ -multilinear map $T : \Gamma(E^*)^r \times \Gamma(E)^s \rightarrow C^\infty(M)$. By proposition 2.1.8, T is pointwise so by proposition 2.1.7 there must exist for every $p \in M$ a unique multilinear map

$$T_p : (E_p^*)^r \times (E_p)^s \rightarrow \mathbb{R}$$

such that $T(\alpha^1, \dots, \alpha^r, e_1, \dots, e_s)_p = T_p(\alpha^1(p), \dots, \alpha^r(p), e_1(p), \dots, e_s(p))$ for $\alpha^j \in \Gamma(E^*)$ and $e_j \in \Gamma(E)$. Such a multilinear map T_p is precisely an element of $T_s^r E_p$ (theorem 1.1.1) thus we have a set-theoretic section $T : M \rightarrow T_s^r E, p \rightarrow T_p$. We still need to ascertain that this section is smooth. To that end, for a point $p \in M$ we choose a local frame $e = (e_k)_{k=1}^r$ of E over an open neighborhood U of p . The components of T in the frame e are given by

$$T_{j_1 \dots j_s}^{i_1 \dots i_r} = T(e_*^{i_1}, \dots, e_*^{i_r}, e_{j_1}, \dots, e_{j_s}).$$

We choose $\tilde{e}_j \in \Gamma(E)$ and $\tilde{e}_*^j \in \Gamma(E^*)$ such that $\tilde{e}_j \sim_p e_j$ and $\tilde{e}_*^j \sim_p e_*^j$. Then on a neighborhood of p , $T_{j_1 \dots j_s}^{i_1 \dots i_r}$ coincides with the smooth function $T(\tilde{e}_*^{i_1}, \dots, \tilde{e}_*^{i_r}, \tilde{e}_{j_1}, \dots, \tilde{e}_{j_s})$, thus showing that the section T is smooth on that same neighborhood of p . Since this holds for all $p \in M$, T is smooth.

We leave it to the reader to check that, conversely, a smooth section of $T_s^r E$ induces a $C^\infty(M)$ -multilinear maps $\Gamma(E^*)^r \times \Gamma(E)^s \rightarrow C^\infty(M)$. We also let the reader convince

themselves that the contraction and tensor product of smooth sections of $T_s^r E$ are well-defined. $\square$

Example 2.1.3. The only application of these operations we will genuinely be concerned with in subsequent chapters are the tensor bundles $T_s^r M$ and exterior powers $\Lambda^k T^* M$ of the (co)tangent bundle.

Local frames (i.e. trivializations) of the tangent bundle are usually obtained through coordinate charts. Given a coordinate chart $(U, x^1, \dots, x^m)$ on M , the standard vector fields $(\partial_k)_{k=1}^m$ induced by the coordinate chart form a local frame of TM over U . The associated coframe of $T^* M$ over U is given by $(dx^k)_{k=1}^m$.

Given a tensor field $T \in \Gamma(T_s^r M)$ and a k -form $\alpha \in \Omega^k(M)$, the components of T and α in that coordinate chart are the unique smooth functions $(T_{j_1 \dots j_s}^{i_1 \dots i_r})$, $(\alpha_{i_1 \dots i_k})$ over U such that

$$\begin{aligned} T &= \sum_{i_1, \dots, i_r, j_1, \dots, j_s} T_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s} \\ \alpha &= \sum_{i_1 < \dots < i_k} \alpha_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} \end{aligned}$$

and are given by

$$\begin{aligned} T_{j_1 \dots j_s}^{i_1 \dots i_r} &= T(dx^{i_1}, \dots, dx^{i_r}, \partial_{j_1}, \dots, \partial_{j_s}) \\ \alpha_{i_1 \dots i_k} &= \alpha(\partial_{i_1}, \dots, \partial_{i_k}). \end{aligned}$$

Lastly, if $(V, x'^1, \dots, x'^m)$ is another coordinate chart with $U \cap V$, we have

$$\partial'_l = \sum_{k=1}^m \frac{\partial x^k}{\partial x'^l} \partial_k \quad (2.1)$$

so the transition map between the local frames associated to the charts (U, x) and (V, x') is given by

$$U \cap V \rightarrow GL_m(\mathbb{R}), p \rightarrow \left(\frac{\partial x^k}{\partial x'^l}(p) \right)_{k,l=1}^m.$$

Using 2.1 and $dx'^k = \sum_{l=1}^m \frac{\partial x'^k}{\partial x^l} dx^l$ we see that the components T in the charts (U, x) and (V, x') are related by

$$T_{j_1 \dots j_s}^{i_1 \dots i_r} = \sum_{q_1, \dots, q_r, t_1, \dots, t_s} T'_{t_1 \dots t_s}{}^{q_1 \dots q_r} \frac{\partial x^{i_1}}{\partial x'^{q_1}} \dots \frac{\partial x^{i_r}}{\partial x'^{q_r}} \frac{\partial x'^{t_1}}{\partial x^{j_1}} \dots \frac{\partial x'^{t_s}}{\partial x^{j_s}}$$

(a similar expression can be obtained for the components of α).

We've seen how, in perfect analogy with the linear case, tensor bundles (or more precisely, their sections) can be used to encode $C^\infty(M)$ -multilinear maps of sections. Sim-

ilarly, taking advantage of the universality of the exterior power (proposition 1.1.5), we can encode *skew-symmetric* $C^\infty(M)$ -multilinear maps of sections using **bundle-valued forms**.

Definition 2.1.10 (Bundle-valued forms). Let E be a vector bundle over M . An E -valued k -form is a section of $E \otimes \Lambda^k T^*M$. Given an open subset $U \subseteq M$, we denote the space of E -valued k -forms over U by $\Omega^k(U, E)$, or simply $\Omega^k(E)$ when $U = M$.

Remark. Using the vector bundle criterion 2.1.1, we can see that $E \otimes \Lambda^k E^*$ is a vector subbundle of $E \otimes T_k E$. In particular, an E -valued k -form is a smooth section ω of $E \otimes T_k E$ such that for every $p \in M$, $\omega_p \in E_p \otimes \Lambda^k E_p^*$.

Proposition 2.1.10. Let E, F be vector bundles over M .

- (i) There is a one-to-one correspondence between elements of $\Omega^k(E)$ and alternating $C^\infty(M)$ -multilinear maps $\mathfrak{X}(M) \times \mathfrak{X}(M) \times \dots \times \mathfrak{X}(M) \rightarrow \Gamma(E)$.
- (ii) There is a one-to-one correspondence between $C^\infty(M)$ -multilinear maps

$$\mathfrak{X}(M) \times \mathfrak{X}(M) \times \dots \times \mathfrak{X}(M) \times \Gamma(E) \rightarrow \Gamma(F)$$

which are alternating in their first k arguments and elements of $\Omega^k(\text{Lin}(E, F))$.

Proof. We will prove the more difficult implication in (ii). Let ω be a $C^\infty(M)$ -multilinear map

$$\mathfrak{X}(M) \times \mathfrak{X}(M) \times \dots \times \mathfrak{X}(M) \times \Gamma(E) \rightarrow \Gamma(F)$$

alternating in its first k arguments. A reasoning similar to the one we used to characterize sections of $T_s^r E$ shows that there exists a unique smooth section ω of $F \otimes E^* \otimes T_k M$ such that

$$\omega(X_1, \dots, X_k, e)_p = \omega_p(X_1(p), \dots, X_k(p), e(p))$$

for $X_j \in \mathfrak{X}(M)$, $e \in \Gamma(E)$. The fact that ω is skew-symmetric in its first k arguments implies that ω_p is skew-symmetric in its first k -arguments, so $\omega_p \in F_p \otimes E_p^* \otimes \Lambda^k T_p^* M$ (proposition 1.1.5) and the conclusion follows. $\square$

2.1.5 Pullback bundle

The bundle operations listed in the previous section all have their analogues for vector spaces, and as such they apply to vector bundles lying over a single, fixed, manifold. The **pullback** on the other hand, the operation we will introduce here, has no analogue for vector spaces as it involves a smooth map between two base manifolds. The idea is the following: given a smooth map $\phi : M \rightarrow N$ and a vector bundle E lying over N , one can 'pull E back along ϕ ' to obtain a vector bundle $\phi^* E$ over M together with a bundle morphism $\Phi : \phi^* E \rightarrow E$ universal among bundle morphisms over ϕ in a sense that will

be made precise shortly. This pullback process happens fiberwise: each point p of M is sent to single point $\phi(p)$ of N , and we obtain the fiber of ϕ^*E over p by 'transporting' the fiber of E over $\phi(p)$ so that $(\phi^*E)_p \simeq E_{\phi(p)}$.

This idea is simple but quite deep: it tells us that vector bundles are, in their essence, *contravariant* objects, objects that are to be 'pulled back' along maps⁴. This contravariant nature becomes obvious if we see vector bundles as primarily determined by their sections, since sections are very clearly prone to being 'pulled back along maps' via precomposition.

Definition 2.1.11 (Pullback bundle). Let E be a vector bundle over N and let $\phi : M \rightarrow N$ be a smooth map. There exists a pair (ϕ^*E, Φ) , consisting of a vector bundle ϕ^*E over M and a bundle morphism $\Phi : \phi^*E \rightarrow E$ over ϕ , that possesses the **universal property of the pullback**, meaning that for any vector bundle F over M and any bundle morphism $\Psi : F \rightarrow E$ over ϕ , there exists a unique bundle morphism (over M) $\bar{\Psi} : F \rightarrow \phi^*E$ such that

$$\begin{array}{ccc} \phi^*E & \xrightarrow{\Phi} & E \\ \bar{\Psi} \uparrow & \nearrow \Psi & \\ F & & \end{array}$$

commutes. We call ϕ^*E the **pullback (bundle) of E along ϕ** . We sometimes say that $\bar{\Psi}$ **factors** uniquely through Φ and call $\bar{\Psi}$ the **factoring map**.

As with all universal properties, (ϕ^*E, Φ) is unique up to canonical isomorphism, meaning that if (ϕ^*E', Φ') is another pair that possesses the universal property of the pullback, then there exists a unique bundle isomorphism $\Psi : \phi^*E \rightarrow \phi^*E'$ such that

$$\begin{array}{ccc} \phi^*E' & \xrightarrow{\Phi'} & E \\ \Psi \uparrow & \nearrow \Phi & \\ \phi^*E & & \end{array} \quad \begin{array}{ccc} \phi^*E' & \xrightarrow{\Phi'} & E \\ \Psi^{-1} \downarrow & \nearrow \Phi' & \\ \phi^*E & & \end{array}$$

commute (note that as soon as one of the above diagrams commutes, the other must as well).

Proof of the existence of the pullback bundle. We can construct a 'standard version' of the pullback bundle using the same approach as in definition 2.1.9. As a set $\phi^*E := \bigsqcup_{p \in M} E_{\phi(p)}$ (in other words, the fiber of ϕ^*E at p is chosen to be the fiber of E at

⁴For the readers who are aware, this property is shared by cohomology classes, and this observation announces the theory of characteristic classes. An introduction to the subject will be found in [26].

$\phi(p)$), and we define $\Phi : \phi^*E \rightarrow E$ to be the unique map that restricts to the identity $id : \phi^*E_p = E_{\phi(p)} \rightarrow E_{\phi(p)}$ on each fiber.

Notice that a set-theoretic section of ϕ^*E is simply a map $s : M \rightarrow E$ such that $\pi \circ s = \phi$ (with π the projection of E), in particular for any section $s \in \Gamma(E)$, we can construct the 'pullback section' $\phi^*s := s \circ \phi$ of ϕ^*E .

We endow ϕ^*E with a smooth structure by pulling back the local frames of E . Given a local frame $e = (e_k)_{k=1}^r$ of E over U , the collection $\phi^*e := (\phi^*e_k)_{k=1}^r$ forms a local frame of ϕ^*E over $\phi^{-1}(U)$. Transition maps between such "pulled back" local frames are very easily described. If $e' = (e'_k)_{k=1}^r$ is another local frame of E over an open V with $U \cap V \neq \emptyset$, and if we denote the transition map between e and e' by

$$g_{UV} : U \cap V \rightarrow GL_r(\mathbb{R})$$

then the transition map between ϕ^*e and ϕ^*e' is given by $g_{UV} \circ \phi$ (defined over $\phi^{-1}(U \cap V)$). Clearly $g_{UV} \circ \phi$ is smooth, so by theorem 2.1.2, ϕ^*E admits a unique bundle structure such that for any local frame e of E , ϕ^*e is a local frame of ϕ^*E .

Given a bundle F over M and a bundle morphism $\Psi : F \rightarrow E$ lying over ϕ , it is then clear what the factoring map $\bar{\Psi} : F \rightarrow \phi^*E$ ought to be: it is the only fiber-preserving map such that for any point $p \in M$, $\bar{\Psi}_p : F_p \rightarrow \phi^*E_p = E_{\phi(p)}$ is equal to $\Psi_p : F_p \rightarrow E_{\phi(p)}$. We leave it to the reader to check that this map, as well as the previously defined map $\Phi : \phi^*E \rightarrow E$ are bundle morphisms. $\square$

Remark. The smooth structure on ϕ^*E is uniquely characterized by the following property: a section s of ϕ^*E is smooth if and only if the composition

$$\Phi \circ s : M \rightarrow E$$

is smooth.

Remark. It shouldn't be too difficult for the reader to convince themselves (possibly making use of proposition 2.1.12) of the functoriality of the pullback operation, which means that given a chain of maps $M \xrightarrow{\psi} N \xrightarrow{\phi} K$ and a vector bundle E over K , if we let (ϕ^*E, Φ) be the pullback of E along ϕ and $(\psi^*\phi^*E, \Psi)$ be the pullback of ϕ^*E along ψ , we have a canonical isomorphism

$$\psi^*\phi^*E \simeq (\phi \circ \psi)^*E$$

through which the universal morphism $(\phi \circ \psi)^*E \rightarrow E$ is given by $\Phi \circ \Psi$.

We use the notation $\Gamma_\phi(E)$ to denote the space of sections of ϕ^*E . Notice that if we use the standard pullback bundle, a section of ϕ^*E is a **section of E over ϕ** , that is a smooth map $s : M \rightarrow E$ such that $\pi \circ s = \phi$. As we already alluded to, precomposition by ϕ allows us to turn sections on E into sections of ϕ^*E .

Proposition 2.1.11. *For any section s of E there is a unique section ϕ^*s of ϕ^*E verifying $\Phi \circ \phi^*s = s \circ \phi$. We call ϕ^*s the **pullback of s along ϕ** .*

Remark. If we use the standard version of the pullback bundle, then the pullback section ϕ^*s is in fact equal to $s \circ \phi$.

Drawing on the 'fiberwise' criterion for vector bundle isomorphisms (proposition 2.1.1), we obtain a similar fiberwise characterization of the pullback bundle.

Proposition 2.1.12. *Let E be a vector bundle over N and let $\phi : M \rightarrow N$ be a smooth map. A pair (F, Ψ) consisting of a vector bundle F over M and a bundle morphism $\Psi : F \rightarrow E$ over ϕ is a pullback of E along ϕ if and only if for any point $p \in M$, $\Psi_p : F_p \rightarrow E_{\phi(p)}$ is a linear isomorphism.*

Proof. Let $\Phi : F \rightarrow E$ be a bundle morphism over ϕ such that for any $p \in M$, $\Psi_p : F_p \rightarrow E_{\phi(p)}$ is a linear isomorphism. Then the factoring map $\bar{\Psi} : F \rightarrow \phi^*E$ acts isomorphically on fibers - this can be seen by taking the standard pullback and using the explicit description of the factoring map in that case. By proposition 2.1.1, $\bar{\Psi}$ must then be a bundle isomorphism. From there, proving that (F, Ψ) possesses the universal property of the pullback bundle is a simple exercise. $\square$

The pullback operation is compatible with all the vector bundle operations we introduced in the previous section, essentially because both these operations and the pullback process happen fiberwise.

Proposition 2.1.13. *Let $\phi : M \rightarrow N$ be a smooth map and let E be a vector bundle over N . Then for any $0 \leq r, s$ there exists a canonical isomorphism*

$$\phi^*T_s^r E \simeq T_s^r(\phi^*E).$$

Through this isomorphism we have for any $T \in \Gamma(T_s^r E)$, $S \in \Gamma(T_l^k E)$, $1 \leq i_1 < i_2 < \dots < i_k \leq r$ and $1 \leq j_1 < j_2 < \dots < j_k \leq s$

$$(i) \quad \phi^*(T \otimes S) = (\phi^*T) \otimes (\phi^*S);$$

$$(ii) \quad \phi^*(C_{j_1 \dots j_k}^{i_1 \dots i_k} T) = C_{j_1 \dots j_k}^{i_1 \dots i_k} (\phi^*T).$$

Similarly we have

$$\phi^*\Lambda^k E^* \simeq \Lambda^k(\phi^*E)^*$$

and

$$\phi^*(\alpha \wedge \beta) = (\phi^*\alpha) \wedge (\phi^*\beta)$$

for any $\alpha \in \Gamma\Lambda^k E^*$, $\beta \in \Gamma\Lambda^l E^*$.

Proof. Exercise. This is essentially a consequence of proposition 2.1.12, and is most easily seen when the pullback is the 'standard one', in which case sections of $\phi^*T_s^r E$ are simply smooth maps $s : M \rightarrow T_s^r E$ such that $\pi \circ s = \phi$. $\square$

Example 2.1.4. In what follows, the examples of pullback bundles we will encounter will mostly be of the form $\gamma^*T_s^r M$ for a curve γ , that is a smooth map $\gamma : I \rightarrow M$ with I an interval of $\mathbb{R}$.

Sections of γ^*TM in particular are commonly called **vector fields over γ** , so a vector field X over γ corresponds to a choice of $X(t) \in T_{\gamma(t)}M$ for all $t \in I$ such that the resulting map $\gamma : I \rightarrow TM$ is smooth. A natural example is given by the **velocity field** $\dot{\gamma}$ of γ , defined by

$$\dot{\gamma}(t) := d\gamma_t(d/dt) = \left. \frac{d}{dt} \right|_{t=0} \gamma(t) \in T_{\gamma(t)}M$$

where d/dt denotes the standard coordinate vector field on $\mathbb{R}$.

Finally, extending the pullback operation for (real-valued) differential forms, bundle-valued forms also be pulled back, and this requires pulling back the bundle they take value in.

Definition 2.1.12 (Pullback of bundle-valued forms). Let E be a vector bundle over N , $\omega \in \Omega^k(N, E)$ and $\phi : M \rightarrow N$ be a smooth map. The **pullback** of ω along ϕ is the unique form $\phi^*\omega \in \Omega^k(M, \phi^*E)$ such that for any point $p \in M$ and any family of tangent vectors $X_1, \dots, X_k \in T_pM$, $\Phi_p(\phi^*\omega)_p(X_1, \dots, X_k) = \omega_{\phi(p)}(d\phi_p X_1, \dots, d\phi_p X_k)$.

Alternatively, if we view bundle-valued forms as skew-symmetric $C^\infty(M)$ -multilinear maps acting on vector fields (proposition 2.1.10), the pullback $\phi^*\omega$ of $\omega \in \Omega^k(N, E)$ is defined by

$$(\phi^*\omega)(X_1, \dots, X_k) = \phi^*(\omega(Y_1, \dots, Y_k))$$

for vector fields $X_j \in \mathfrak{X}(M)$, $Y_j \in \mathfrak{X}(N)$ such that Y_j is ϕ -related to X_j , meaning that $Y_{j, \phi(p)} = d\phi_p X_{j, p}$ for all $p \in M$.

And in the spirit of our remark on the functoriality of the pullback operation, we have the following.

Proposition 2.1.14. Consider the sequence of maps $M \xrightarrow{\psi} N \xrightarrow{\phi} K$, let E be a vector bundle over K and let $\omega \in \Omega^k(E)$. Then $\psi^*\phi^*\omega = (\phi \circ \psi)^*\omega$.

2.2 Connections on vector bundles

One of the main limitations we face when working 'plain' vector bundle E (i.e. with no additional structure) is the fact that there exists no clear relationship between distinct fibers of E , no canonical way of transporting vectors from one fiber to another. Connections are tailor-made to fill that gap. As their name indicates, they provide a way to 'connect' fibers to each other.

This will be made explicit in the next subsection, by introducing the notion of parallel transport along a path, and will not be so clear until then. The starting definition we choose for connections (which is the common one, to our knowledge), introduces them instead as a kind of derivative (which in this case means 'operator satisfying the Leibniz rule') acting on sections, a point of view which should be understood as a description of parallel transport at the infinitesimal scale.

Definition 2.2.1 (Connection on a vector bundle). Let E be a vector bundle over M a **connection**, or **covariant derivative** on E is an $\mathbb{R}$ -linear map

$$\nabla : \Gamma(E) \rightarrow \Omega^1(M, E)$$

such that for all $f \in C^\infty(M)$ and $s \in \Gamma(E)$

$$\nabla(fs) = df \otimes s + f\nabla s.$$

Given $s \in \Gamma(E)$ and $X \in \mathfrak{X}(M)$ we write $(\nabla s)(X) \stackrel{n}{=} \nabla_X s$.

Equivalently, a connection on E can be defined as an $\mathbb{R}$ -bilinear map

$$\nabla : \mathfrak{X}(M) \times \Gamma(E) \rightarrow \Gamma(E), (X, s) \rightarrow \nabla_X s$$

such that for all $s \in \Gamma(E)$, $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$:

- (i) $\nabla_{fX}s = f\nabla_X s$.
- (ii) $\nabla_X(fs) = (Xf)s + f\nabla_X s$.

We finish this definition by introducing isomorphisms for vector bundles with connections. So let $\pi_E : E \rightarrow M$ and $\pi_F : F \rightarrow N$ be vector bundles over the manifolds M and N , equipped with connections ∇_E and ∇_F respectively. Let $\Phi : E \rightarrow F$ be a vector bundle isomorphism over the diffeomorphism $\phi : M \rightarrow N$. Then Φ is said to preserve the connections (that is, Φ is an isomorphism of bundles with connections) if for all $s \in \Gamma(E)$ and $X \in \mathfrak{X}(M)$ we have

$$\Phi \circ (\nabla_X s) \circ \phi^{-1} = \nabla_{\phi_* X} (\Phi \circ s \circ \phi^{-1}). \quad (2.2)$$

(If this formula does not seem natural, recall that through the isomorphism $\Phi : E \rightarrow F$, a section s of E is mapped to the section $\Phi \circ s \circ \phi^{-1}$ of F).

Remark. It should be clear that, given an isomorphism of vector bundles $\Phi : E \rightarrow F$, if either E or F is equipped with a connection, then there exists a unique connection on the other member of the pair (E, F) for which Φ becomes an isomorphism of bundles with connection. This connection is essentially defined by formula 2.2 (as we'll soon see, this construction is a special case of the pullback operation one can apply to connections on vector bundles).

If we view the connection, or covariant derivative, as a kind of differentiation for sections of a vector bundle, the corresponding notion of a 'constant section' is that of a parallel section (a section whose covariant derivative vanishes).

Definition 2.2.2 (Parallel sections). A section $s \in \Gamma(E)$ is said to be **parallel** if $\nabla s = 0$, or equivalently if $\nabla_X s = 0$ for all $X \in \mathfrak{X}(M)$.

As we'll see, these parallel sections tend to be rare, even locally, except in certain special cases. Most notably among the latter, parallel sections are found in abundance on a vector bundle over a curve.

Example 2.2.1.

- (i) The trivial vector bundle of rank r over M admits a standard connection, which we denote by d . A section of $\underline{\mathbb{R}}_M^r$ is a smooth map

$$\underline{s} : M \rightarrow \mathbb{R}^r$$

or equivalently a family of smooth functions $s^k : M \rightarrow \mathbb{R}, k = 1, \dots, r$, while an $\underline{\mathbb{R}}_M^r$ -valued 1-form $\alpha \in \Omega^1(M, \underline{\mathbb{R}}_M^r)$ is simply a collection of (real-valued) 1-forms $\alpha^1, \dots, \alpha^k \in \Omega^1(M)$. The standard connection d on $\underline{\mathbb{R}}_M^r$ sends

$$\underline{s} : M \rightarrow \mathbb{R}^r$$

to

$$d\underline{s} := (ds^k)_{k=1, \dots, r}$$

where $ds^k \in \Omega^1(M)$ denotes the usual differential of s^k as a smooth function. We leave it as an exercise to the reader to check that this does indeed define a connection on $\underline{\mathbb{R}}_M^r$. When M is connected, the parallel sections of $\underline{\mathbb{R}}_M^r$ are the constant maps $\underline{s} : M \rightarrow \mathbb{R}^r$ (in particular, the connection d admits "many" parallel sections, something we will come back to).

- (ii) The main objects of study in this work will be connections on the tangent bundle TM , which are called **affine connections**.

If we fix a vector field X on M , then the map $\nabla_X : \Gamma(E) \rightarrow \Gamma(E)$ is a local operator, as it satisfies the Leibniz rule for X . In particular, ∇_X induces a local operator at p (see definition 2.1.7). On the other hand, if we fix a section s , the map

$$\mathfrak{X}(M) \rightarrow \Gamma(E), X \rightarrow \nabla_X s$$

is pointwise, as a $C^\infty(M)$ -linear map of sections. In particular $(\nabla_X s)_p$ only depends on the value of X at p for any $X \in \mathfrak{X}(M)$ and $s \in \Gamma(E)$. Combining these facts we get the following:

Proposition 2.2.1. *There exists a unique $\mathbb{R}$ -bilinear map*

$$\nabla : T_p M \times \Gamma_p(E) \rightarrow E_p, (X_p, \mathfrak{s}_p) \rightarrow \nabla_{X_p} \mathfrak{s}_p$$

such that for all $X \in \mathfrak{X}(M)$ and $s \in \Gamma(E)$

$$(\nabla_X s)_p = \nabla_{X_p} [s]_p$$

In what follows we will simply denote $\nabla_{X_p} [s]_p$ by $\nabla_{X_p} s$.

Remark. The moral of this somewhat convoluted development is that even though ∇ was introduced as an operator acting on global sections, it turns out that $\nabla_{X_p} s \in E_p$ is perfectly well-defined for any tangent vector $X_p \in T_p M$ and any local section s defined on a neighborhood of p . Furthermore, $\nabla_{X_p} s$ and $\nabla_{X_p} t$ are equal whenever the local sections s and t coincide on a neighborhood of p .

This local nature of the connection allows us to define its restriction to an open subset.

Proposition 2.2.2. *Let $\pi : E \rightarrow M$ be a vector bundle over M equipped with a connection ∇ . Then for any open subset U of M , there exists a unique connection ∇^U on $E_U := \pi^{-1}(U)$, called the **restriction of ∇ to U** such that for any $s \in \Gamma(E)$, $\nabla^U(s|_U) = (\nabla s)|_U$. This restriction is completely determined by the fact that for all $X \in \mathfrak{X}(U)$, $s \in \Gamma(U, E)$ and $p \in U$*

$$(\nabla_X^U s)_p = \nabla_{X_p} s$$

where ∇_{X_p} is the operator defined in proposition 2.2.1.

Proof. As indicated, for $p \in U$, $X \in \mathfrak{X}(U)$ and $s \in \Gamma(U, E)$, the only possible candidate for ∇^U will be defined by the formula

$$(\nabla_X^U s)(p) = (\nabla_{\tilde{X}} \tilde{s})(p)$$

with $\tilde{X} \in \mathfrak{X}(M)$ a global vector field and $\tilde{s} \in \Gamma(E)$ a global section of E such that $\tilde{X} \sim_p X$ and $\tilde{s} \sim_p s$. Such $\tilde{X}$ and $\tilde{s}$ always exist by proposition 2.1.4, and the choice of

$\tilde{X}$ and $\tilde{s}$ is inconsequential as we already explained. Showing that ∇^U possesses the properties of a connection is then a routine exercise. $\square$

Conversely, if we cover our manifold with mutually compatible local connections, we obtain a globally defined connection.

Proposition 2.2.3. *Let $(U_\alpha)_{\alpha \in A}$ be an open cover of M , and assume that for each $\alpha \in A$ we are given a connection ∇^α on $\pi^{-1}(U_\alpha) \subseteq E$ in such a way that for any $\alpha, \beta \in A$ with $U_\alpha \cap U_\beta \neq \emptyset$, the restrictions of ∇^α and ∇^β to $U_\alpha \cap U_\beta$ coincide. Then there exists a unique connection ∇ on M such that for all $\alpha \in A$ the restriction of ∇ to U_α is ∇^α .*

Proof. Clearly the only possible candidate for ∇ is determined by

$$(\nabla s)|_{U_\alpha} = \nabla^\alpha(s|_{U_\alpha})$$

for all $s \in \Gamma(E)$ and $\alpha \in A$. Alternatively, taking this as a definition of ∇ - which is unambiguous, since the ∇^α are mutually compatible - showing that ∇ possesses the property of a connection is a routine exercise. $\square$

Let us now investigate the structure of the set of connections on a given vector bundle. We begin by observing that what distinguishes a connection on E from an element $\Omega^1(M, \text{End}(E))$ is the presence of the $(Xf)s$ term in the Leibniz rule (which keeps ∇ from being $C^\infty(M)$ -linear in its $\Gamma(E)$ argument). This term, however, does not depend on the connection, meaning we should be able to 'cancel it out' by taking the difference of two connections, thereby obtaining an element of $\Omega^1(M, \text{End}(E))$. This suggests that the set of connections on E has the structure of an affine space modeled on $\Omega^1(M, \text{End}(E))$, under the condition that E admits at least one connection, something we will show to be true by patching together trivial connections using a partition of unity.

Theorem 2.2.1. *Let E be a vector bundle over M . The set of connections on E is an affine space modeled on $\Omega^1(M, \text{End}(E))$. In more words, a vector bundle always admits a connection, and given a connection ∇ on E , a map*

$$\nabla' : \Gamma(E) \rightarrow \Omega^1(M, E)$$

is a connection on E if and only if

$$\mathfrak{X}(M) \times \Gamma(E) \rightarrow \Gamma(E), (X, s) \rightarrow (\nabla'_X s - \nabla_X s)$$

determines a section of $\Omega^1(M, \text{End}(E))$ (i.e. is $C^\infty(M)$ -linear).

Proof. We begin by constructing a connection ∇ on E . To that end, we fix a trivializing open cover $(U_\alpha, \phi_\alpha)_{\alpha \in S}$ of E . For each $\alpha \in S$, we denote by ∇^α the connection on $\pi^{-1}(U_\alpha) \subseteq E$ obtained by pulling back the standard connection on the trivial bundle $U_\alpha \times \mathbb{R}^r$ over r (we assume E to be of rank r) along the bundle isomorphism

$$\phi_\alpha : \pi^{-1}(U_\alpha) \rightarrow U_\alpha \times \mathbb{R}^r.$$

Next we choose a partition of unity $(\rho_\alpha)_{\alpha \in S}$ subordinate to $(U_\alpha)_{\alpha \in S}$. For $X \in \mathfrak{X}$, $s \in \Gamma(E)$ we denote the restriction of s to U_α by s^α , and write $\nabla_X^\alpha \stackrel{n}{=} \nabla_{X|_{U_\alpha}}^\alpha$. The formula

$$\nabla_X s := \sum_{\alpha \in S} \rho_\alpha \nabla_X^\alpha s^\alpha$$

then defines a connection on M . We let the reader verify that ∇ exhibits the properties of a connection (the reasoning is in fact similar to the one we use to prove the next proposition - prop. 2.2.4).

We now move on to the second part of the statement (which states that the direction of the set of connections - as an affine space - is given by $\Omega^1(\text{End}(E))$). Consider two maps

$$\begin{aligned} \nabla : \mathfrak{X}(M) \times \Gamma(E) &\rightarrow \Gamma(E), (X, s) \rightarrow \nabla_X s \\ \nabla' : \mathfrak{X}(M) \times \Gamma(E) &\rightarrow \Gamma(E), (X, s) \rightarrow \nabla'_X s \end{aligned}$$

and define

$$A : \mathfrak{X}(M) \times \Gamma(E), (X, s) \rightarrow A(X, s) := (\nabla'_X s - \nabla_X s).$$

Assume that ∇ is a connection on E . Then if ∇' is a connection, we have for any $X \in \mathfrak{X}(M)$, $s \in \Gamma(E)$ and $f \in C^\infty(M)$

$$\begin{aligned} A(fX, s) &= \nabla'_{fX} s + \nabla_{fX} s = f(\nabla'_X s - \nabla_X s) \\ &= fA(X, s) \end{aligned} \tag{2.3}$$

as well as

$$\begin{aligned} A(X, fs) &= \nabla'_X(fs) - \nabla_X(fs) \\ &= (Xf)s + f\nabla'_X s - (Xf)s - f\nabla_X s \\ &= f(\nabla'_X s - \nabla_X s) \\ &= fA(X, s). \end{aligned} \tag{2.4}$$

Conversely, one easily sees that ∇' is a connection whenever 4.1 and 4.2 hold. Thus ∇' is a connection if and only if A is $C^\infty(M)$ -linear in each of its arguments, or equivalently, if and only if A is a section of $\Omega^1(\text{End}(E))$ (proposition 2.1.10). $\square$

In particular, as elements of an affine space, connections are stable under convex combinations.

Proposition 2.2.4. *Let $(\nabla^i)_{i=1,\dots,n}$ be a finite family of connections on E and let $(\lambda_i)_{i=1,\dots,n}$ be a family of positive smooth functions on M such that $\sum_{i=1}^n \lambda_i = 1$. Then $\nabla := \sum_{i=1}^n \lambda_i \nabla^i$ is a connection on E .*

Proof. For $s \in \Gamma(E)$, $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$ we have

$$\begin{aligned} \nabla_{fX}s &= \sum_{i=1}^n \lambda_i \nabla_{fX}^i s = \sum_{i=1}^n f \lambda_i \nabla_X^i s \\ &= f \left(\sum_{i=1}^n \lambda_i \nabla_X^i s \right) \\ &= f \nabla_X s. \end{aligned}$$

as well as

$$\begin{aligned} \nabla_X(fs) &= \sum_{i=1}^n \lambda_i \nabla_X^i(fs) \\ &= \sum_{i=1}^n \lambda_i \left((Xf)s + f \nabla_X^i s \right) \\ &= \left(\sum_{i=1}^n \lambda_i \right) (Xf)s + f \left(\sum_{i=1}^n \lambda_i \nabla_X^i s \right) \\ &= (Xf)s + f \nabla_X s. \end{aligned}$$

Note that, as we expected, the convexity hypothesis ($\sum_{i=1}^n \lambda_i = 1$) was only needed to ensure that ∇ satisfy the Leibniz rule. Once again, we see that the Leibniz rule is the reason why the set of connections possesses the structure of an affine space (as opposed to a vector space). $\square$

Example 2.2.2. Theorem 2.2.1 applied to the case $E = TM$ tells us that the set of affine connections on a manifold is an affine space modeled on the space $\Gamma(T_2^1 M)$ of tensor fields of type $(1, 2)$.

We will now investigate the local expressions of connections, i.e. their representations in local trivializations. So let $e = (e_k)_{k=1}^r$ be a local frame for E over U . Since ∇ satisfies the Leibniz rule, the restriction of ∇ to U (which we will also call ∇) is completely known once ∇e_k is fixed for every $k = 1, \dots, r$. We could choose this as a starting point to find the object representing ∇ over U , but we will explore another path, closer to the spirit of theorem 2.2.1.

We begin by defining d to be the connection on $\pi^{-1}(U) \subseteq E$ obtained through the induced trivialization

$$\phi : \pi^{-1}(U) \rightarrow U \times \mathbb{R}^r$$

and the standard connection on the trivial bundle $U \times \mathbb{R}^r$ (see example 2.2.1). Notice that d is completely characterized by

$$de_k = 0, k = 1, \dots, r.$$

Theorem 2.2.1 then provides us with the local representation of ∇ : it is the unique $\text{End}(E)$ -valued 1-form $A \in \Omega^1(U, \text{End}(E))$, called the **local connection form**, such that $\nabla = d + A$. The relationship between A and ∇ is given by

$$\nabla e_k = (d + A)e_k = Ae_k, k = 1, \dots, r$$

thus we see that they determine each other. If we express A in components

$$A = \sum_{k,l=1}^r e_k \otimes e_*^l \otimes A_l^k.$$

the 1-forms $A_l^k \in \Omega^1(U)$ are defined by

$$\nabla e_l = \sum_{k=1}^r e_k \otimes A_l^k.$$

As mentioned earlier, knowing the ∇e_k allows us to compute ∇s for any local section $s \in \Gamma(U, E)$ by making use of the Leibniz rule. So for a general section $s \in \Gamma(U, E)$ with components $\underline{s} = (s^k)_{k=1}^r$, the components of ∇s are given by

$$\nabla s = \sum_{k=1}^r (d + A)(s^k e_k) = \sum_{k=1}^r e_k \otimes \left(ds^k + \sum_{l=1}^r A_l^k s^l \right)$$

Thus if collect the components of A in a matrix-valued 1-form (or matrix of 1-forms) $\underline{A} = (A_l^k)_{k,l=1}^r \in \Omega^1(U, M_r(\mathbb{R}))$, then for any vector field $X \in \mathfrak{X}(U)$ we have

$$\nabla_X s \equiv X \underline{s} + \underline{A}(X) \underline{s} \tag{2.5}$$

where we apply X to each component of $\underline{s}$ in the first term, and use matrix multiplication in the second term. If we allow for matrix multiplication between 1-forms and functions we can reexpress this as

$$\nabla s \equiv d \underline{s} + \underline{A} \underline{s}.$$

Now that we know how a connection is represented over a local trivialization (by a matrix A of 1-forms), let us look at how this representation varies under changes of trivialization. So let $e' = (e'_k)_{k=1}^r$ be another local frame of E over an open V such that $U \cap V \neq \emptyset$. Let $A' = \sum_{k,l=1}^r e'_k \otimes e_*'^l \otimes A_l'^k$ be the matrix of 1-forms representing ∇ in e' , and let

$$g_{UV} : U \cap V \rightarrow GL_r(\mathbb{R})$$

be the transition map between e and e' so that

$$e'_l = \sum_{k=1}^r (g_{UV})_l^k e_k.$$

Then we have

$$\begin{aligned} \sum_{v=1}^r e'_v \otimes A_l'^v &= \nabla e'_l = \sum_{k=1}^r \nabla \left((g_{UV})_l^k e_k \right) \\ &= \sum_{k=1}^r e_k \otimes \left(d(g_{UV})_l^k + \sum_{u=1}^r (g_{UV})_l^u A_u^k \right) \\ &= \sum_{k,v=1}^r (g_{UV}^{-1})_k^v e'_v \otimes \left(d(g_{UV})_l^k + \sum_{u=1}^r (g_{UV})_l^u A_u^k \right) \end{aligned}$$

Thus we see that A' and A are related by the equation

$$A_l'^v = \sum_{k=1}^r (g_{UV}^{-1})_k^v \left(d(g_{UV})_l^k + \sum_{u=1}^r (g_{UV})_l^u A_u^k \right) \quad (2.6)$$

which can be rewritten more concisely in the following way

$$\underline{A}' = (g_{UV})^{-1} d(g_{UV}) + (g_{UV})^{-1} \underline{A} (g_{UV}).$$

Conversely, in view of proposition 2.2.3, we can construct a connection on E by choosing a trivializing cover for E and local connection forms for each trivialization in such a way that these local forms are compatible with each other (i.e. transform according to equation 2.6). This is a rather common way of constructing connections.

Theorem 2.2.2. *Let $(U_\alpha, \phi_\alpha)_{\alpha \in S}$ be a trivializing open cover of E . Assume that for each $\alpha \in A$ we are given an $\text{End}(E)$ -valued form A_α over U_α , and assume that for each $\alpha, \beta \in S$ with $U_\alpha \cap U_\beta \neq \emptyset$, A_α and A_β are related by equation 2.6. Then there exists a unique connection ∇ on M such that for each $\alpha \in S$, A_α is the local connection form of ∇ in the trivialization (U_α, ϕ_α) .*

Example 2.2.3. Once again, let us apply our discussion to the case of an affine connection ∇ on TM . Given a coordinate chart $(U, x^1, \dots, x^m)$ on M , with associated local frame $(\partial_i)_{i=1}^m$, the most standard way of representing ∇ over this chart is through the **Christoffel symbols**, a family of smooth functions

$$\Gamma_{ij}^k : U \rightarrow \mathbb{R}, 1 \leq i, j, k \leq m$$

which are related to the local connection form by

$$A_j^k = \sum_{i=1}^m \Gamma_{ij}^k dx^i. \quad (2.7)$$

Equivalently, the Christoffel symbols can be defined by

$$\nabla_{\partial_i} \partial_j = \sum_{k=1}^m \Gamma_{ij}^k \partial_k$$

so that for arbitrary vector fields $X = \sum_{i=1}^m X^i \partial_i$ and $Y = \sum_{i=1}^m Y^i \partial_i$ over U we have

$$\nabla_X Y = \sum_{k=1}^m \left(\sum_{i=1}^m X^i \partial_i Y^k + \sum_{i,j=1}^m \Gamma_{ij}^k X^i Y^j \right) \partial_k. \quad (2.8)$$

Given another chart $(V, x'^1, \dots, x'^m)$ we've already seen that the transition map is given by

$$g_l^k = \frac{\partial x^k}{\partial x'^l}.$$

Injecting the following equations

$$\begin{aligned} (g^{-1})_l^k &= \frac{\partial x'^k}{\partial x^l} \\ dg_l^k &= \sum_{j=1}^m \frac{\partial^2 x^k}{\partial x'^j \partial x'^l} dx'^j \\ dx^i &= \sum_{j=1}^m \frac{\partial x^i}{\partial x'^j} dx'^j \end{aligned}$$

into 2.7 and 2.6 we obtain the usual transformation law for the Christoffel symbols of an affine connection

$$\Gamma_{ij}^k = \sum_{l=1}^m \frac{\partial x'^k}{\partial x^l} \frac{\partial^2 x^l}{\partial x'^i \partial x'^j} + \sum_{p,q,r=1}^m \frac{\partial x'^k}{\partial x^p} \frac{\partial x^q}{\partial x'^i} \frac{\partial x^r}{\partial x'^j} \Gamma_{qr}^p. \quad (2.9)$$

In keeping with the idea that vector bundles are objects 'to be pulled back', connections can be pulled back as well. We prove this through the procedure outlined in theorem 2.2.2, applied to a pulled back trivializing cover.

Corollary 2.2.1. *Let E be a vector bundle over M equipped with a connection ∇ , and let $\phi : M \rightarrow N$ be a smooth map. Then there exists a unique connection ∇^ϕ on ϕ^*E , called the **pullback connection** such that for any $s \in \Gamma(E)$, $X_p \in T_p M$, $\nabla_{X_p}^\phi(\phi^*s) = \nabla_{d\phi_p X_p} s$.*

Equivalently, ∇^ϕ is the unique connection on ϕ^*E such that for any pair of ϕ -related vector fields $X \in \mathfrak{X}(M)$, $Y \in \mathfrak{X}(N)$ and any section $s \in \Gamma(E)$ we have

$$\nabla_X(\phi^*s) = \phi^*(\nabla_Y s).$$

Proof. We choose a trivializing open cover $(U_\alpha, e_\alpha)_{\alpha \in S}$ of E , where for each $\alpha \in S$, $e_\alpha = (e_{\alpha,k})_{k=1}^r$ is a local frame for E over U_α . For $\alpha \in S$ we denote the local connection

form of ∇ in (U_α, e_α) by A_α , in components

$$A_\alpha = \sum_{k,l=1}^r e_{\alpha,k} \otimes e_{*\alpha}^l \otimes A_{\alpha,l}^k.$$

Then $(\phi^{-1}(U_\alpha), \phi^*e_\alpha)_{\alpha \in S}$ is a trivializing open cover for ϕ^*E , and if a connection ∇^ϕ on ϕ^*E is to possess the properties of the pullback connection, then for each $\alpha \in S$, its local connection form in $(\phi^{-1}(U_\alpha), \phi^*e_\alpha)$ must be given by

$$\phi^*A_\alpha = \sum_{k,l=1}^r (\phi^*e)_{\alpha,k} \otimes (\phi^*e)_{*\alpha}^l \otimes \phi^*A_{\alpha,l}^k \quad (2.10)$$

which already shows uniqueness. To prove existence, we proceed as follows: we denote by $g_{\alpha\beta}$ the transition map between (U_α, e_α) and (U_β, e_β) . Using the fact that the forms A_α transform according to 2.6, and the fact that the transition map between $(\phi^{-1}(U_\alpha), \phi^*e_\alpha)$ and $(\phi^{-1}(U_\beta), \phi^*e_\beta)$ is given by $g_{\alpha\beta} \circ \phi$, we see that the local forms ϕ^*A_α also transform according to 2.6, so that they define a unique connection ∇^ϕ on ϕ^*E (theorem 2.2.2). Using the expression 2.10 for the local connection forms, it is then straightforward to show that ∇^ϕ is the sought after pullback connection. $\square$

Remark. In a more intrinsic fashion, if e is a local frame of E over U , and $A \in \Omega^1(U, \text{End}(E))$ is the local connection form of ∇ in this trivialization, then the local connection form of ∇^ϕ in the local frame ϕ^*e is given by ϕ^*A (this is essentially the way we constructed ∇^ϕ in the proof above).

Remark. Extending an earlier remark, we can show that the pullback of connections is functorial as well, meaning that given a sequence of maps $M \xrightarrow{\psi} N \xrightarrow{\phi} K$ and a vector bundle E over K with connection ∇ , pulling ∇ back along $\phi \circ \psi$ is equivalent to first pulling ∇ back along ϕ , then pulling ∇^ϕ back along ψ (this could be proven using proposition 2.1.14 applied to the local connection form).

Remark. Going back to definition 2.2.1, observe that, given two vector bundles $\pi_E : E \rightarrow M$, $\pi_F : F \rightarrow N$ equipped with connections ∇_E and ∇_F respectively, then a bundle isomorphism $\Phi : E \rightarrow F$ lying over $\phi : M \rightarrow N$ preserves the connection if and only if the pullback of ∇_F along ϕ coincides with ∇_E (by proposition 2.1.12, (E, Φ) is a pullback of F along ϕ).

We finish this section with an easy but useful result.

Proposition 2.2.5. *Let s be a parallel section of E . Then ϕ^*s is a parallel section of ϕ^*E .*

2.2.1 Covariant derivative along a path and parallel transport

We are now ready to define parallel transport. We begin by applying the pullback machinery (for bundles and connections) to curves, resulting in the notion of the **covariant derivative along a curve**.

Definition 2.2.3 (Covariant derivative along a curve). Let E be a vector bundle over M equipped with a connection ∇ , and let $\gamma : I \rightarrow M$ be a smooth curve on M (with I an interval of any kind). For any section $s \in \Gamma_\gamma(E)$ over γ , we define

$$\nabla_{\dot{\gamma}} s := \nabla_{\frac{d}{dt}}^\gamma s$$

with d/dt the standard vector field on $\mathbb{R}$. We call $\nabla_{\dot{\gamma}}$ the **covariant derivative along γ** . Notice that a section $s \in \Gamma_\gamma(E)$ over γ is parallel for ∇^γ if and only if $\nabla_{\dot{\gamma}} s = 0$.

Remark. The notation we use here is motivated by the fact that, for a section $s \in \Gamma(E)$ the covariant derivative of $\gamma^* s$ is given by

$$\nabla_{\dot{\gamma}}(\gamma^* s)(t) = \nabla_{\dot{\gamma}(t)} s, \text{ for all } t \in I.$$

Let us also mention the peculiar form that the Leibniz rule will take in this instance: for a smooth function $f : I \rightarrow \mathbb{R}$ and a section $s \in \Gamma_\gamma(E)$ along γ we have

$$\nabla_{\dot{\gamma}}(fs) = \dot{f}s + f\nabla_{\dot{\gamma}} s$$

with $\dot{f} = \frac{df}{dt}$ (the application of the vector field d/dt to f , which coincides with the usual derivative of f).

The condition $\nabla_{\dot{\gamma}} s = 0$ is shown to be a linear ODE of first order by choosing a local trivialization. Indeed, let $e = (e_k)_{k=1}^r$ be a local frame of E over an open subset U , let A be the local connection form of ∇ for this local frame. Referring back to the proof of corollary 2.2.1, we can pullback e and A along γ to obtain a local frame $\gamma^* e$ for $\gamma^* E$ over $J := \gamma^{-1}(U)$ together with the local connection form

$$\gamma^* A = \sum_{k,l=1}^r (\gamma^* e)_k \otimes (\gamma^* e)_*^l \otimes \gamma^* A_l^k$$

for ∇^γ in that local frame. Then given a local section

$$s = \sum_{k=1}^r s^k (\gamma^* e_k)$$

of $\gamma^* E$ over J with components $s^k : J \rightarrow \mathbb{R}, k = 1, \dots, r$ we have

$$\nabla_{\dot{\gamma}} s = \sum_{k=1}^r (\gamma^* e_k) \left(\dot{s}^k + \sum_{l=1}^r A_l^k(\dot{\gamma}) s^l \right)$$

where we use notations

$$\begin{aligned}\dot{s}^k &:= \frac{ds^k}{dt} \\ \dot{\gamma}(s) &:= d\gamma_s(d/dt) = \left. \frac{d}{dt} \right|_{t=s} \gamma(t), s \in J \\ A_l^k(\dot{\gamma})(t) &:= A_l^k|_{\gamma(t)}(\dot{\gamma}(t)), t \in J\end{aligned}$$

(recall that A_l^k is a 1-form). So locally the condition $\nabla_{\dot{\gamma}}s = 0$ becomes the system of linear ODE's

$$0 = \dot{s}^k + \sum_{l=1}^r A_l^k(\dot{\gamma})s^l, k = 1, \dots, r. \quad (2.11)$$

The usual existence and uniqueness results for linear ODE's then yield:

Theorem 2.2.3. *Let $\gamma : I \rightarrow M$ be a smooth curve (with I an interval of any kind). Then for any $t \in I$ and $e \in E_{\gamma(t)}$ there exists a unique parallel section s over γ such that $s(t) = e$.*

And from this we finally obtain the following theorem:

Theorem 2.2.4 (Parallel transport along smooth curves). *Let $\gamma : [a, b] \rightarrow M$ be a smooth curve going from $p \in M$ to $q \in M$ (meaning $\gamma(a) = p$, $\gamma(b) = q$). Then for any $s_p \in E_p$ there exists a unique parallel section s over γ such that $s(a) = s_p$.*

*We write $s(b) := P(\gamma)_a^b s_p$ (which only depends on s_p and γ as stated). The map $P(\gamma)_a^b : E_p \rightarrow E_q$ is a linear isomorphism which we call the **parallel transport**, or **translation** along γ .*

To prove the injectivity of $P(\gamma)_a^b$, we'll use the following lemma (which is an immediate consequence of the $\mathbb{R}$ -linearity of a connection):

Lemma 2.2.1. *Let E be a vector bundle over M equipped with a connection ∇ . Then the 0 section of E is parallel (the map $M \rightarrow E$ sending $p \in M$ to the origin of the fiber E_p over p).*

Proof of theorem 2.2.4. We take theorem 2.2.3 for granted, so we need only prove that $P(\gamma)_a^b$ is linear isomorphism $E_a \rightarrow E_b$.

The linearity of $P(\gamma)_a^b$ follows from the fact that for any two parallel sections s, t over γ and any $\lambda \in \mathbb{R}$ we have

$$\nabla_{\dot{\gamma}}(\lambda s + t) = \lambda \nabla_{\dot{\gamma}}s + \nabla_{\dot{\gamma}}t = 0$$

by the linearity of ∇ . To see that $P(\gamma)_a^b$ is injective (and thus an isomorphism), take $s_a \in E_a$ and assume that $P(\gamma)_a^b s_a = 0$. Then there exists a parallel section s over γ such

that $s(a) = s_a$ and $s(b) = 0$. But by combining theorem 2.2.3 and lemma 2.2.1 we see that s must be the 0 section, so that $s_a = s(a) = 0$.

□

We can show that parallel transport is not sensitive to reparametrization (this follows from the functoriality of the pullback operation for connections), and behaves very well with respect to partitions (the act of subdividing a curve into a sequence of curves). We can also show that reversal of paths inverts parallel transport.

Proposition 2.2.6. *Let $\gamma : [a, b] \rightarrow M$ be a smooth curve. Then*

(i) *For any smooth map $\phi : [a', b'] \rightarrow [a, b]$ with $\phi(a') = a$ and $\phi(b') = b$, we have $P(\gamma \circ \phi)_{a'}^{b'} = P(\gamma)_a^b$.*

(ii) *For any finite partition $a = t_0 < t_1 < \dots < t_n = b$ of $[a, b]$ we have*

$$P(\gamma)_a^b = P(\gamma_{n-1})_{t_{n-1}}^{t_n} \circ \dots \circ P(\gamma_0)_{t_0}^{t_1}$$

where $P(\gamma_i)_{t_i}^{t_{i+1}}$ denotes parallel transport along $\gamma_i := \gamma|_{[t_i, t_{i+1}]}$.

(iii) *If we define the reversal of γ , $\bar{\gamma} : [0, b-a] \rightarrow M$ by $\bar{\gamma}(t) := \gamma(b-t)$, then*

$$P(\bar{\gamma})_0^{b-a} = [P(\gamma)_a^b]^{-1}.$$

Proof. (i) By the functoriality of the pullback for connections (see the remark below corollary 2.2.1), the connection over $\gamma \circ \phi$ is obtained by pulling back ∇^γ along ϕ . (i) is then obtained by observing that for any parallel section s over γ , $\phi^*s = s \circ \phi$ is a parallel section over $\gamma \circ \phi$ (proposition 2.2.5).

(ii) is a consequence of the fact that given a parallel section s over γ , $s|_{[t_i, t_{i+1}]}$ is a parallel section over $\gamma|_{[t_i, t_{i+1}]}$ for all $0 \leq i < n$.

As for (iii), observe that $\bar{\gamma} = \gamma \circ \tau$ with

$$\tau : [0, b-a] \rightarrow [a, b], t \rightarrow b-t$$

so that for any parallel section s over γ , $s \circ \tau$ is parallel along $\bar{\gamma}$. The conclusion then follows. □

This means that parallel transport along piecewise smooth curves is perfectly well-defined, and plays well with the concatenation of paths.

Definition 2.2.4 (Parallel transport along piecewise smooth curves). Let $\gamma : [a, b] \rightarrow M$ be a piecewise smooth curve, meaning that there exists a finite partition $a = t_0 < t_1 < \dots < t_n = b$ such that $\gamma_i := \gamma|_{[t_i, t_{i+1}]}$ is smooth for all $0 \leq i < n$. Then we define

$$P(\gamma)_a^b := P(\gamma_{n-1})_{t_{n-1}}^{t_n} \circ \dots \circ P(\gamma_0)_{t_0}^{t_1}.$$

which does not depend on our choice of partition.

Proposition 2.2.7. Let $\mu, \tau : [0, 1] \rightarrow M$ be two piecewise smooth curves with $\tau(1) = \mu(0)$, then

$$P(\mu \cdot \tau)_0^1 = P(\mu)_0^1 \circ P(\tau)_0^1$$

where $\mu \cdot \tau$ denotes the concatenation of μ and τ , defined by

$$\mu \cdot \tau : [0, 1] \rightarrow M, t \rightarrow \begin{cases} \tau(2t), & \text{for } 0 \leq t \leq \frac{1}{2} \\ \mu(2t - 1), & \text{for } \frac{1}{2} \leq t \leq 1 \end{cases}$$

Confirming the idea that parallel transport is at the core of the theory of connections, we show that *from* the induced parallel transport, one can recover the connection.

Proposition 2.2.8. Let $s \in \Gamma(U, E)$ and $X_p \in T_p M$ for $p \in U$. Let $\gamma : I \rightarrow M$ be a smooth curve such that $\dot{\gamma}(0) = X_p$. Then $\nabla_{X_p} s = \left. \frac{d}{dt} \right|_{t=0} P(\gamma)_t^0 (s \circ \gamma)(t)$.

Proof. We choose a basis $(e_{k,p})_{k=1}^r$ of E_p . For each $1 \leq k \leq r$, let e_k be the unique parallel section over γ such that $e_k(0) = e_{k,p}$, whose existence and uniqueness is guaranteed by theorem 2.2.3. Then by theorem 2.2.4, we see that $e = (e_k)_{k=1}^r$ is a global frame of $\gamma^* E$ and that for $t \in I$ and $v = \sum_{k=1}^r x^k e_k(t)$ we have

$$P(\gamma)_t^0(v) = \sum_{k=1}^r x^k e_k(0).$$

In particular, given a section $s = \sum_{k=1}^r s^k e_k$ over γ (with $s^k \in C^\infty(I)$) we have

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} P(\gamma)_t^0 s(t) &= \left. \frac{d}{dt} \right|_{t=0} \left(\sum_{k=1}^r s^k(t) e_k(0) \right) \\ &= \sum_{k=1}^r \dot{s}^k(0) e_k(0) \end{aligned} \tag{2.12}$$

On the other hand, from $\nabla_{\dot{\gamma}} e_k = 0$ and the Leibniz rule, we obtain

$$\nabla_{\dot{\gamma}} s = \sum_{k=1}^r \dot{s}^k e_k. \tag{2.13}$$

Comparing 2.12 and 2.13 we deduce

$$\nabla_{\dot{\gamma}} s|_{t=0} = \frac{d}{dt} \Big|_{t=0} P(\gamma)_t^0 s(t)$$

We can then conclude using the fact that for $s \in \Gamma(E)$

$$\nabla_{X_p} s = \nabla_{\dot{\gamma}}(s \circ \gamma)|_{t=0}$$

with $X_p = \dot{\gamma}(0)$. □

We finish this section by refining our earlier statement on the relative scarcity of parallel sections. As seen below, the connections which give rise to the most parallel sections are the trivial ones. In the next section we will introduce a new invariant, curvature, which will allow us to distinguish between trivial and non-trivial connections at the local scale (though this will only be made precise in the last section of the present chapter).

Theorem 2.2.5. *Let ∇ be a connection on a vector bundle E . The following are equivalent:*

- (i) *(E, ∇) is isomorphic to a trivial bundle equipped with its standard connection.*
- (ii) *E admits a global frame consisting of parallel sections.*
- (iii) *For any $p \in E$ and any $s_p \in E_p$ there exists a parallel section $s \in \Gamma(E)$ whose value at p is s_p .*

When conditions (i)-(iii) are fulfilled, we say that ∇ is a **trivial connection** on E .

Proof. (i) $\Rightarrow$ (iii) We need only show that a trivial bundle equipped with its standard connection verifies (iii). This is easily seen once one recalls that the parallel sections of the trivial bundle $\underline{\mathbb{R}}^r$ with its standard connection are the constant maps $M \rightarrow \mathbb{R}^r$ (when M is connected).

(iii) $\Rightarrow$ (ii). We assume that M is connected, if not we repeat the process for each connected component of M . Choose a point $p \in M$ and a basis $(e_{k,p})_{k=1}^r$ of E_p . For each $1 \leq k \leq r$, let $e_k \in \Gamma(E)$ be a parallel section such that $e_k(p) = e_{k,p}$. For any $q \in M$ there exists a path $\gamma : [0, 1] \rightarrow M$ going from p to q , and for all $1 \leq k \leq r$, $e_k \circ \gamma$ is parallel over γ , so $e_k(q) = P(\gamma)_0^1 e_k(p)$. Since $P(\gamma)_0^1$ is a linear isomorphism $E_p \rightarrow E_q$, $e_k(q)$ must form a basis of E_q , and we see that $(e_k)_{k=1}^r$ is a global frame of E consisting of parallel sections.

(ii) $\Rightarrow$ (i). Let $e = (e_k)_{k=1}^r$ be a global frame of E consisting of parallel sections. This global frame induces a bundle isomorphism $\Phi : E \rightarrow \underline{\mathbb{R}}^r$ that sends e_k to b_k^0 (the constant section of $\underline{\mathbb{R}}^r$ with value b_k^0 , with $b^0 = (b_k^0)_{k=1}^r$ the standard basis of $\mathbb{R}^r$). Thus for

all $1 \leq k \leq r$, b_k^0 must be parallel for the connection on $\underline{\mathbb{R}}^r$ induced by ∇ and the isomorphism Φ . The only connection on $\underline{\mathbb{R}}^r$ with this property is the trivial one. $\square$

Remark. If E is a bundle over an interval $I \subseteq \mathbb{R}$, any connection E is trivial as shown earlier. As we'll see, this is coherent with the fact that the curvature of any such section must vanish, as a 2-form over a 1-dimensional manifold.

2.2.2 Curvature

We define curvature by means of its action on sections. This action is proven to be $C^\infty(M)$ -linear with no difficulty, thus showing that the curvature is a bundle-valued 2-form (proposition 2.1.10).

Definition 2.2.5 (Curvature). Let E be a vector bundle over M equipped with a connection ∇ . The **curvature** R of ∇ is the map

$$\begin{aligned} \mathfrak{X}(M) \times \mathfrak{X}(M) \times \Gamma(E) &\rightarrow \Gamma(E) \\ (X, Y, s) &\rightarrow \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X, Y]} s. \end{aligned}$$

The notation

$$R(X, Y)s := ([\nabla_X, \nabla_Y] - \nabla_{[X, Y]})s$$

is also common.

Proposition 2.2.9. *The curvature R is $C^\infty(M)$ -linear in each of its arguments and skew-symmetric in its first two arguments and thus corresponds to an element of $\Omega^2(M, \text{End}(E))$, which we also call R .*

Let us derive the expression of the curvature in a local trivialization. Let $e = (e_k)_{k=1}^r$ be a local frame for E over U . We are looking for the matrix-valued 2-form $\underline{R} = (R_l^k)_{k, l=1}^r \in \Omega^2(U, M_r(\mathbb{R}))$ such that

$$R = \sum_{k, l=1}^r e_k \otimes e_*^l \otimes R_l^k.$$

We denote the components of the local connection form $A \in \Omega(U, \text{End}(E))$ by $\underline{A} = (A_l^k)_{k, l=1}^r \in \Omega^1(U, M_r(\mathbb{R}))$ so that

$$A = \sum_{k, l=1}^r e_k \otimes e_*^l \otimes A_l^k.$$

Fix two vector fields X, Y over U , and a local section $s \in \Gamma(U, E)$ whose components we denote by $\underline{s} = (s^k)_{k=1}^r \in C^\infty(U)$. The components of $\nabla_X s$ are given by

$$\nabla_Y s \equiv Y \underline{s} + \underline{A}(Y) \underline{s} \tag{2.14}$$

where we apply Y to each component of $\underline{s}$ and use matrix multiplication in the second term. Applying ∇_X to 2.14 we get

$$\begin{aligned}\nabla_X \nabla_Y s &\equiv X(Y\underline{s} + \underline{A}(Y)\underline{s}) + \underline{A}(X)(Y\underline{s} + \underline{A}(Y)\underline{s}) \\ &\equiv XY\underline{s} + X[\underline{A}(Y)]\underline{s} + \underline{A}(Y)X\underline{s} \\ &\quad \underline{A}(X)Y\underline{s} + \underline{A}(X)\underline{A}(Y)\underline{s}.\end{aligned}\tag{2.15}$$

Exchanging X and Y in 2.15 and replacing Y with $[X, Y]$ in 2.14 we obtain

$$\begin{aligned}R(X, Y)s &= \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X, Y]}s \\ &\equiv \cancel{XY\underline{s}} + X[\underline{A}(Y)]\underline{s} + \cancel{\underline{A}(Y)X\underline{s}} + \cancel{\underline{A}(X)Y\underline{s}} + \underline{A}(X)\underline{A}(Y)\underline{s} \\ &\quad - \cancel{YX\underline{s}} - Y[\underline{A}(X)]\underline{s} - \cancel{\underline{A}(X)Y\underline{s}} - \cancel{\underline{A}(Y)X\underline{s}} - \underline{A}(Y)\underline{A}(X)\underline{s} \\ &\quad - \cancel{[X, Y]\underline{s}} - \underline{A}([X, Y])\underline{s} \\ &= (X[\underline{A}(Y)] - Y[\underline{A}(X)] - \underline{A}([X, Y]))\underline{s} + [\underline{A}(X), \underline{A}(Y)]\underline{s} \\ &= (d\underline{A}(X, Y) + [\underline{A}(X), \underline{A}(Y)])\underline{s}.\end{aligned}\tag{2.16}$$

where we use Cartan's magic formula in the last step, and where $[-, -]$ denotes the anticommutator of matrices. From 2.16 we obtain

$$\begin{aligned}R_l^k(X, Y) &= dA_l^k(X, Y) + \sum_{h=1}^r (A_h^k(X)A_l^h(Y) - A_h^k(Y)A_l^h(X)) \\ &= dA_l^k(X, Y) + \sum_{h=1}^r (A_h^k \wedge A_l^h)(X, Y)\end{aligned}$$

So we see that the 2-forms representing R are given by

$$R_l^k = dA_l^k + \sum_{h=1}^r A_h^k \wedge A_l^h.\tag{2.17}$$

Example 2.2.4. For an affine connection whose connection form A is encoded by the Christoffel symbols

$$A_l^k = \sum_{j=1}^m \Gamma_{jl}^k dx^j$$

equation 2.17 becomes

$$\begin{aligned}R_l^k &= d\left(\sum_{j=1}^m \Gamma_{jl}^k dx^j\right) + \sum_{h=1}^m \left(\sum_{i=1}^m \Gamma_{ih}^k dx^i\right) \wedge \left(\sum_{j=1}^m \Gamma_{jl}^h dx^j\right) \\ &= \sum_{j=1}^m d\Gamma_{jl}^k \wedge dx^j + \sum_{i,j,h=1}^m \Gamma_{ih}^k \Gamma_{jl}^h dx^i \wedge dx^j \\ &= \sum_{i,j=1}^m \left(\partial_i \Gamma_{jl}^k + \sum_{h=1}^m \Gamma_{ih}^k \Gamma_{jl}^h\right) dx^i \wedge dx^j\end{aligned}$$

and if we use $dx^i \wedge dx^j = dx^i \otimes dx^j - dx^j \otimes dx^i$ we obtain the usual expression of the curvature in terms of the Christoffel symbols:

$$R_{ijl}^k = \partial_i \Gamma_{jl}^k - \partial_j \Gamma_{il}^k + \sum_{h=1}^m (\Gamma_{ih}^k \Gamma_{jl}^h - \Gamma_{jh}^k \Gamma_{il}^h).$$

As a bundle-valued form, the curvature interacts with the pullback operation as we would expect.

Proposition 2.2.10. *Let $\phi : M \rightarrow N$ be a smooth map and let E be a vector bundle over N equipped with a connection ∇ , with curvature R . Then the curvature R^ϕ of the pullback connection ∇^ϕ is the pullback of R along ϕ , i.e. $R^\phi = \phi^* R$ (here we use the canonical isomorphism $\phi^* \text{End}(E) \simeq \text{End}(\phi^* E)$).*

Proof. This is obtained by putting together the expression for the local connection form of the pullback connection (see the proof of corollary 2.2.1, and the remarks below) and the expression of the curvature in terms of the local connection form (equation 2.17). $\square$

We now provide an illuminating interpretation of curvature as **infinitesimal holonomy**. Holonomy is the name we give to the following phenomenon: in general (that is, for an arbitrary connection), transporting a vector parallelly along a loop will not leave this vector unchanged. This is formalized by introducing the holonomy group.

Definition 2.2.6 (Holonomy group). Let ∇ be a connection on E , and let p be a point of M . The **holonomy group** of ∇ at p is the subgroup of $GL(E_p)$ defined by

$$\text{Hol}_p(E, \nabla) := \{P(\gamma)_0^1 | \gamma : [0, 1] \rightarrow M \text{ is a piecewise smooth curve based at } p\}.$$

(A loop based at p is a curve $\gamma : [a, b] \rightarrow M$ such that $\gamma(a) = p = \gamma(b)$).

The fact that $\text{Hol}_p(E, \nabla)$ is an algebraic subgroup of $GL(E_p)$ should be clear from propositions 2.2.6 and 2.2.7. The fact that it is a Lie subgroup of $GL(E_p)$ is a very deep result, which we will state at the end of this section.

But first, we present a theorem that makes clear the interpretation of curvature as ‘holonomy along an infinitesimal loop’.

Theorem 2.2.6. *Let $X, Y \in T_p M$ and let $e \in E_p$. Assume that U is an open neighborhood of the origin in $\mathbb{R}^2$ and that $f : U \rightarrow M$ is a smooth map such that $f(0) = p$, $df_0(\partial_1(0)) = X_p$, and $df_0(\partial_2(0)) = Y_p$, where ∂_1 and ∂_2 designate the standard vector fields on $\mathbb{R}^2$. For s, t sufficiently small, let e_{st} be the element of E_p obtained by parallel transporting e along the piecewise smooth loop consisting of the following four curves (see figure 2.1):*

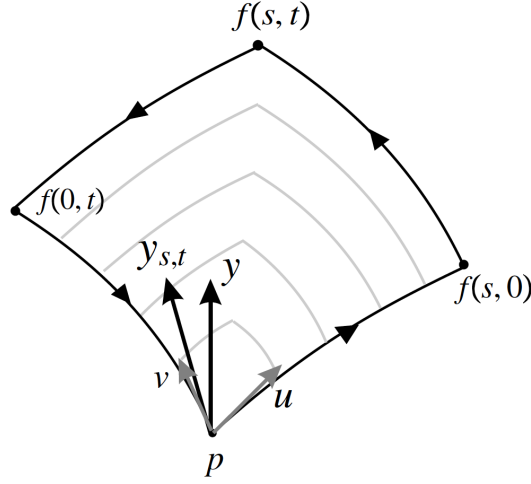


Figure 2.1: Parallel translation along a loop, from [17], sect. 12.5.

- (1) $c_1^{s,t} : \sigma \rightarrow f(\sigma, 0)$ for $0 \leq \sigma \leq s$;
- (2) $c_2^{s,t} : \tau \rightarrow f(s, \tau)$ for $0 \leq \tau \leq t$;
- (3) $c_3^{s,t} : \sigma \rightarrow f(s - \sigma, t)$ for $0 \leq \sigma \leq s$;
- (4) $c_4^{s,t} : \tau \rightarrow f(0, t - \tau)$ for $0 \leq \tau \leq t$.

Then

$$R(X_p, Y_p)e = - \lim_{s,t \rightarrow 0} \frac{e_{st} - e}{st}.$$

Proof. See [17] (theorem 12.47, section 12.5, p.527) or [18] (theorem 7.11, p.201). $\square$

To finish, we state the very important theorem of Ambrose and Singer which, among other things, strengthens our interpretation of curvature as a measure of infinitesimal holonomy.

Theorem 2.2.7 (Ambrose-Singer). *$\text{Hol}_p(E, \nabla)$ is a Lie subgroup of $GL(E_p)$, and its Lie algebra (a Lie subalgebra of $\text{End}(E_p)$) is spanned by elements of the form $R(X_p, Y_p)$ for $X_p, Y_p \in T_p M$.*

Proof. See [12] (vol. I, chapter II, theorem 4.2 p.73 and theorem 8.1 p.89, same volume and chapter). $\square$

2.2.3 Connections on tensor bundles

We will now describe how connections interact with the operations listed in definition 2.1.9.

Definition 2.2.7. Let E, F be vector bundles over M equipped with connections (we will use the same symbol ∇ to designate most of the connections appearing in the span of this definition). Then the connections on E and F induce canonical connections on all the bundles one can construct from E and F through the operations listed in 2.1.9.

- (i) On $E \oplus F$ the induced connection ∇ is defined by

$$\nabla_X(e \oplus f) = (\nabla_X e) \oplus (\nabla_X f)$$

for all $X \in \mathfrak{X}(M)$, $e \in \Gamma(E)$ and $f \in \Gamma(F)$.

- (ii) The induced connection ∇^* on E^* is defined by

$$(\nabla_X^* \alpha)(e) = X[\alpha(e)] - \alpha(\nabla_X e)$$

for all $X \in \mathfrak{X}(M)$, $e \in \Gamma(E)$ and $\alpha \in \Gamma(E^*)$ (in the equation above, $\alpha(e)$ is a smooth function, which is why we can apply the vector field X to it).

- (iii) The induced connection ∇ on $E \otimes F$ is defined by we

$$(\nabla_X W)(\alpha, \beta) = X[W(\alpha, \beta)] - W(\nabla_X^* \alpha, \beta) - W(\alpha, \nabla_X^* \beta)$$

for all $X \in \mathfrak{X}(M)$, $W \in \Gamma(E \otimes F)$, $\alpha \in \Gamma(E^*)$ and $\beta \in \Gamma(F^*)$, where we use the induced connections on E^* and F^* (see previous point).

Through repeated applications of (i) and (ii), the connection on E also induces a connection on $T_s^r E$ for all $r, s \geq 0$.

- (iv) Using the canonical isomorphism $\text{Lin}(E, F) \simeq F \otimes E^*$, we also obtain a canonical connection on $\text{Lin}(E, F)$ from what came before. This connection is defined by

$$(\nabla_X A)(e) = \nabla_X[A(e)] - A(\nabla_X e)$$

for all $A \in \Gamma \text{Lin}(E, F)$, $e \in \Gamma(E)$ (in the equation above, $A(e)$ is a section of F , which is why we can take its covariant derivative $\nabla_X[A(e)]$).

- (v) The connection on $T_k E$ descends to a connection on $\Lambda^k E^*$, defined by

$$(\nabla_X \alpha)(e_1, \dots, e_k) = X[\alpha(e_1, \dots, e_k)] - \sum_{i=1}^k \alpha(e_1, \dots, \nabla_X e_i, \dots, e_k)$$

for $\alpha \in \Gamma \Lambda^k E^*$, $e_j \in \Gamma(E)$.

Remark. Notice that the connection on $T_s^r E$ is defined so as to enforce a kind of Leibniz rule.

The task of verifying in each case that the formula provided does indeed define a connection on the corresponding bundle is left as an exercise.

The reader will now discover the reason why we chose to dedicate an entire section to the actions of $GL(V)$ and $\text{End}(V)$ on $T(V)$. The language developped in said section will enable us to describe the induced parallel transport on $T_s^r E$ in a very natural manner.

Theorem 2.2.8. *Let $\gamma : [a, b] \rightarrow M$ be a path going from $p \in M$ to $q \in M$, and let*

$$P(\gamma)_a^b : E_p \rightarrow E_q$$

denote parallel translation along γ . Then for any $T \in T_s^r E_p$, the parallel transport of T along γ is given by $P(\gamma)_a^b \cdot T \in T_s^r E_q$, where $\cdot$ is the operation defined in 1.1.4.

Proof. We leave it to the reader to argue that the pullback and tensor operations for connections are compatible, just as they are for vector bundles.

We begin with the dual E^* . Choose $\alpha \in \Gamma_\gamma(E^*)$ and $e \in \Gamma_\gamma(E)$ be parallel sections over γ . By definition of the induced connection on $\gamma^* E^* \simeq (\gamma^* E)^*$ we have

$$0 = (\nabla_{\dot{\gamma}} \alpha)(e) = \frac{d}{dt}[\alpha(e)]$$

so $\alpha(e)$ is constant over $[a, b]$. By definition of the parallel transport, we then have

$$(P^*(\gamma)_a^b \alpha_a)(e_b) = \alpha_b(e_b) = \alpha_a(e_a) = \alpha_a([P(\gamma)_a^b]^{-1} e_b)$$

where P^* denotes parallel transport on E^* . From this we deduce

$$P^*(\gamma)_a^b \alpha_a = [(P(\gamma)_a^b)^{-1}]^* \alpha_a = P(\gamma)_a^b \cdot \alpha_a.$$

We are now ready to treat the case $T_s^r E$. let $T \in \Gamma_\gamma(T_s^r E)$, $\alpha^j \in \Gamma_\gamma(E^*)$ and $e_j \in \Gamma_\gamma(E)$ be parallel sections over γ . Once again we can compute

$$0 = (\nabla_{\dot{\gamma}} T)(\alpha^1, \dots, \alpha^r, e_1, \dots, e_s) = \frac{d}{dt}[T(\alpha^1, \dots, \alpha^r, e_1, \dots, e_s)]$$

and see that $T(\alpha^1, \dots, \alpha^r, e_1, \dots, e_s)$ is constant over $[a, b]$. Denoting parallel transport on

$T_s^r E$ by $\bar{P}$ we then have

$$\begin{aligned}
(\bar{P}(\gamma)_a^b T_a)(\alpha^1(b), \dots, \alpha^r(b), e_1(b), \dots, e_s(b)) &= T_b(\alpha^1(b), \dots, \alpha^r(b), e_1(b), \dots, e_s(b)) \\
&= T_a(\alpha^1(a), \dots, \alpha^r(a), e_1(a), \dots, e_s(a)) \\
&= T_a([P(\gamma)_a^b]^* \alpha^1(b), \dots, [P(\gamma)_a^b]^* \alpha^r(b), [P(\gamma)_a^b]^{-1} e_1(b), \dots, [P(\gamma)_a^b]^{-1} e_s(b)) \\
&= (P(\gamma)_a^b \cdot T_a)(\alpha^1(b), \dots, \alpha^r(b), e_1(b), \dots, e_s(b))
\end{aligned}$$

and from this we conclude $\bar{P}(\gamma)_a^b T_a = P(\gamma)_a^b \cdot T_a$. $\square$

Combining this with proposition 1.1.6 on the compatibility of " $\cdot$ " and the structure of the tensor algebra, and extracting the connection from the parallel transport as in 2.2.8 we can determine how the connection interacts with $\otimes$, $\wedge$ and contraction.

Corollary 2.2.2. *Consider $T \in \Gamma T_s^r E$, $S \in \Gamma T_l^k E$ $1 \leq i_1 < \dots < i_k \leq r$, and $1 \leq j_1 < \dots < j_k \leq s$. Then for all $X \in \mathfrak{X}(M)$ we have*

$$\begin{aligned}
\nabla_X(S \otimes T) &= (\nabla_X S) \otimes T + S \otimes (\nabla_X T) \\
\nabla_X(C_{j_1 \dots j_k}^{i_1 \dots i_k} T) &= C_{j_1 \dots j_k}^{i_1 \dots i_k} (\nabla_X T).
\end{aligned}$$

Similarly for all $\alpha \in \Gamma \Lambda^k E^*$, $\beta \in \Gamma \Lambda^l E^*$ we have

$$\nabla_X(\alpha \wedge \beta) = (\nabla_X \alpha) \wedge \beta + \alpha \wedge (\nabla_X \beta).$$

Example 2.2.5. Expanding on the example of an affine connection: given a tensor field of type $(r, s-1)$ on a manifold M equipped with an affine connection ∇ , let us derive relationships between the components of T and those of the (r, s) tensor ∇T in a coordinate chart $(U, x^1, \dots, x^m)$.

We can begin by describing the action of $\nabla_{\partial_{j_1}}$ on the local frame of $T_{s-1}^r M$. For this corollary 2.2.2 is very useful as it gives us

$$\begin{aligned}
&\nabla_{\partial_{j_1}} (\partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes dx^{j_s}) \\
&= \sum_{k=1}^r \partial_{i_1} \otimes \dots \otimes \nabla_{\partial_{j_1}} \partial_{i_k} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes dx^{j_s} \\
&+ \sum_{k=2}^s \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes \nabla_{\partial_{j_1}} dx^{j_k} \otimes \dots \otimes dx^{j_s}.
\end{aligned} \tag{2.18}$$

By definition of the Christoffel symbols we have

$$\nabla_{\partial_i} \partial_j = \sum_{k=1}^m \Gamma_{ij}^k \partial_k \tag{2.19}$$

and by definition of the connection on T^*M (definition 2.2.7) we have

$$(\nabla_{\partial_i} dx^k)(\partial_j) = \underbrace{\partial_i[dx^k(\partial_j)]}_{=0} - dx^k[\nabla_{\partial_i} \partial_j] = -\Gamma_{ij}^k$$

from which we deduce

$$\nabla_{\partial_i} dx^k = - \sum_{j=1}^m \Gamma_{ij}^k dx^j \quad (2.20)$$

and so equation 2.18 becomes

$$\begin{aligned} & \nabla_{\partial_{j_1}} (\partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes dx^{j_s}) \\ &= \sum_{l=1}^r \sum_{k=1}^m \Gamma_{j_1 i_k}^l \partial_{i_1} \otimes \dots \otimes \partial_l \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes dx^{j_s} \\ & - \sum_{k=2}^s \sum_{l=1}^m \Gamma_{j_1 l}^{j_k} \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes dx^l \otimes \dots \otimes dx^{j_s}. \end{aligned} \quad (2.21)$$

On the other hand, over U we can write

$$T = \sum_{i_1, \dots, i_r, j_2, \dots, j_s} T_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes dx^{j_s}$$

and by the Leibniz rule we obtain

$$\begin{aligned} \nabla_{\partial_{j_1}} T &= \sum_{i_1, \dots, i_r, j_2, \dots, j_s} \left(\partial_{j_1} T_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes dx^{j_s} \right. \\ & \left. + T_{j_1 \dots j_s}^{i_1 \dots i_r} \nabla_{\partial_{j_1}} (\partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_2} \otimes \dots \otimes dx^{j_s}) \right). \end{aligned} \quad (2.22)$$

Finally, putting together equations 2.21 and 2.22 we get

$$\begin{aligned} (\nabla T)_{j_1 \dots j_s}^{i_1 \dots i_r} &= \partial_{j_1} T_{j_2 \dots j_s}^{i_1 \dots i_r} + \sum_{k=1}^r \sum_{l=1}^m \Gamma_{j_1 l}^{i_k} T_{j_2 \dots j_s}^{i_1 \dots l \dots i_r} \\ & - \sum_{k=2}^s \sum_{l=1}^m \Gamma_{j_1 j_k}^l T_{j_2 \dots l \dots j_s}^{i_1 \dots i_r}. \end{aligned} \quad (2.23)$$

From this last equation we can easily compute the components of $\nabla_X T$ for a vector field $X = \sum_{i=1}^m X^i \partial_i$, we need only use $\nabla_X = \sum_{i=1}^m X^i \nabla_{\partial_i}$ which gives us

$$\begin{aligned} (\nabla_X T)_{j_2 \dots j_s}^{i_1 \dots i_r} &= X T_{j_2 \dots j_s}^{i_1 \dots i_r} + \sum_{k=1}^r \sum_{l, h=1}^m \Gamma_{hl}^{i_k} X^h T_{j_2 \dots j_s}^{i_1 \dots l \dots i_r} \\ & - \sum_{k=2}^s \sum_{l, h=1}^m \Gamma_{h j_k}^l X^h T_{j_2 \dots l \dots j_s}^{i_1 \dots i_r}. \end{aligned} \quad (2.24)$$

where we use $X T_{j_2 \dots j_s}^{i_1 \dots i_r} = \sum_{h=1}^m X^h \partial_h T_{j_2 \dots j_s}^{i_1 \dots i_r}$.

Finally, since parallel transport on $T_s^r E$ is given by the $\cdot$ operation, and since curvature is infinitesimal holonomy, it should come as no surprise that the curvature $\bar{R}$ on $T_s^r E$ is induced by the curvature R on E via $\cdot$ (referring to the $\text{End}(E)$ action this time). This can indeed be proven using theorems 2.2.6 and 2.2.8.

Corollary 2.2.3. *Let E be a vector bundle over M with connection ∇ and curvature R . Let $\bar{\nabla}$ denote the induced connection and $T_s^r E$ and $\bar{R}$ denote the curvature of $\bar{\nabla}$. Then for $X, Y \in \mathfrak{X}(M)$ and $W \in \Gamma(T_s^r E)$*

$$\bar{R}(X, Y)W = R(X, Y) \cdot W$$

where $\cdot$ denotes the action of $\text{End}(E)$ on $T_s^r E$.

Proof. Fix $p \in M$ and $X_p, Y_p \in T_p M$. Let $P_{st} : E_p \rightarrow E_p$ and $\bar{P}_{st} : T_s^r E_p \rightarrow T_s^r E_p$ denote the parallel translation along the four curves $c_1^{s,t}$, $c_2^{s,t}$, $c_3^{s,t}$ and $c_4^{s,t}$ described in theorem 2.2.6. We define $L_t := P_{\sqrt{t}\sqrt{t}}$ (note that $L_0 = I$).

Then by theorem 2.2.8 and 2.2.6 we have for any $W_p \in T_s^r E_p$

$$\begin{aligned} \bar{R}(X_p, Y_p)W_p &= -\lim_{t \rightarrow 0} \frac{1}{t} (\bar{P}_{\sqrt{t}\sqrt{t}} - I)W_p = -\lim_{t \rightarrow 0} \frac{1}{t} (P_{\sqrt{t}\sqrt{t}} - I) \cdot W_p \\ &= -\left. \frac{d}{dt} \right|_{t=0} (L_t \cdot W_p) = R(X_p, Y_p) \cdot W_p \end{aligned}$$

where in the last equality we use the fact that for any $e_p \in E_p$

$$-\left. \frac{d}{dt} \right|_{t=0} L_t e_p = -\lim_{t \rightarrow 0} \frac{1}{t} (P_{\sqrt{t}\sqrt{t}} - I)e_p = R(X_p, Y_p)e_p.$$

□

Corollary 2.2.4. *Let E be a vector bundle over M with connection ∇ and curvature R . Let W be a parallel section of $T_s^r E$ (for the induced connection on $T_s^r E$). Then*

$$R(X, Y) \cdot W = 0$$

for any $X, Y \in \mathfrak{X}(M)$.

2.2.4 Horizontal distribution and vertical projection

In the last section of this chapter, we explore a completely different point of view on connections: that of **Ehresmann connections**. Contrary to covariant derivatives, which rely the linear structure of vector bundles in an essential way, Ehresmann connections can be defined on any fiber bundle, with no additional structure.

In the Ehresmann approach, a connection on a fiber bundle $\pi : F \rightarrow M$ is encoded in a **horizontal distribution** on F (see appendix A for the definition of a distribution), the idea being the following: given a point $p \in M$ and an element f of the fiber F_p above p , the differential of the projection $d\pi_f : T_f F \rightarrow T_p M$ is a surjection whose kernel is $T_f(F_p)$, the space tangent to the fiber F_p at f , which we denote by $V_f F \subseteq T_f F$ and call the **vertical subspace at f** . So we have $T_f F / V_f F \simeq T_p M$, meaning that any space complementary to $V_f F$ would be mapped isomorphically onto $T_p M$ by $d\pi_f$. The 'problem' is that there is no canonical choice of such a complementary subspace, and this is precisely the gap that the horizontal distribution HF will fill.

We collect the vertical subspaces $V_f F$ into the vertical distribution $VF := \bigsqcup_{f \in F} V_f F$ (an involutive distribution whose leaves are given by the fibers of F), and a horizontal distribution HF is a distribution on F that is *complementary* to VF . That is, HF is obtained choosing, in a 'smooth fashion', a subspace $H_f F \subseteq T_f F$ complementary to $V_f F$ (which we call the **horizontal subspace at f**) for each $f \in F$.

Parallel transport is then obtained as follows: as we already mentioned, for an element $f \in F$, the horizontal subspace $H_f F$ is mapped isomorphically onto $T_p M$ by the projection (in fact $(HF, d\pi)$ is a pullback of TM along π as per proposition 2.1.12), meaning that given a tangent vector $X_p \in T_p M$, there exists a unique horizontal tangent vector $h_f X_p \in H_f F$, called the **horizontal lift of X_p at f** , that is mapped onto X_p by $d\pi_f$. If we think of tangent vectors as infinitesimal paths, then $h_f X_p$ is to be thought of as the unique parallel (or horizontal) section over X_p starting at f .

And more generally, if we're careful about the domains and images of our maps, then given a curve $\gamma : [0, 1] \rightarrow M$, and an element f of the fiber above $\gamma(0)$, γ will admit a unique horizontal (think parallel) lift $\tilde{\gamma} : [0, 1] \rightarrow F$ starting at f (where we say that a curve $\tilde{\gamma}$ on F is horizontal if its velocity field is horizontal at all points). This unique lift (which is to be compared to the unique parallel section of theorem 2.2.4) will provide the notion of parallel transport we need to connect distinct fibers of F .

If our fiber bundle is equipped with additional structures, as is the case for principal or vector bundles, we can require the Ehresmann connection to be compatible with these additional structures. This is an avenue one could choose to introduce principal connections, or indeed connections on vector bundles (this point of view is explored in [17] section 12). We will not give a full proof of the equivalence between the 'Ehresmann' and 'covariant derivative' point of views, however we will explain how one can obtain the horizontal distribution from the covariant derivative, and relate the curvature of our covariant derivative to the integrability of this horizontal distribution (once again, see appendix A for the notion of integrability for distributions).

Before diving in, let us mention one last thing: just as a choice of a subspace $H_f F \subseteq T_f F$ complementary to $V_f F$ is equivalent to a choice of linear projection $T_f F \rightarrow V_f F$, fixing a horizontal distribution HF complementary to VF is equivalent to fixing a projection $TF \rightarrow VF$ of vector bundles over F (that is a bundle morphism $P_V : TF \rightarrow VF$ such that $P_V|_{VF} = id_{VF}$) which we call the **vertical projection**.

Proposition 2.2.11. *Let $\pi : E \rightarrow M$ be a vector bundle over M , and let p be a point of M . As a finite-dimensional vector space, E_p has a canonical smooth structure. Equipped with this smooth structure, E_p is a properly embedded submanifold of E , and for all $e \in E$ we have*

$$T_e E_p = \ker d\pi_e.$$

Proof. This property is local and is easily verified for a trivial vector bundle. $\square$

Definition 2.2.8. Let E be a vector bundle over M . The **vertical bundle** on E is the smooth distribution VE on E defined by

$$V_e E := \ker d\pi_e$$

From 2.2.11 we see that VE is integrable, and that the fibers of E are its leaves (i.e. maximal connected integral submanifolds).

Theorem 2.2.9. *Let E be a vector bundle over M .*

(i) *Let p be a point of M and let e, v be two elements of E_p . The curve*

$$\mathbb{R} \rightarrow E, t \rightarrow e + tv$$

takes value in E_p and so $\frac{d}{dt}\Big|_{t=0} (e + tv)$ is an element of $V_e E$. This allows us to define the map

$$\zeta : \pi^* E \rightarrow VE, (e, v) \rightarrow \zeta_e v := \frac{d}{dt}\Big|_{t=0} (e + tv)$$

which is an isomorphism of vector bundles over E .

(ii) *Let p be a point of M and let e be an element of the fiber E_p over p . There is a unique linear map*

$$h_e : T_p M \rightarrow T_e E$$

such that for any section $\Gamma(E)$ that takes value e at p we have

$$h_e(X_p) = ds_p(X_p) - \zeta_e \nabla_{X_p} s.$$

*We call h_e the **horizontal lifting at e** and we define the **horizontal subspace at e** by $H_e E := h_e(T_p M)$.*

(iii) *Then the union $HE := \bigsqcup_{e \in E} H_e E$ is a smooth distribution of rank $m = \dim(M)$ over E called the **horizontal distribution** (for ∇). For each $e \in E$, $T_e E = V_e E \oplus H_e E$, $d\pi_e$ maps $H_e E$ isomorphically onto $T_p M$, and h_e is its inverse.*

(iv) There exists a unique bundle projection $p_v : TE \rightarrow VE$, called the **vertical projection** (for ∇) whose kernel is the horizontal distribution HE . Through the canonical isomorphism $VE \simeq \pi^*E$, the vertical projection $p_v : TE \rightarrow VE$ corresponds to a bundle morphism $\kappa : TE \rightarrow E$ over $\pi : E \rightarrow M$, which we call the **connector** of ∇ . Finally, given a tangent vector $X_p \in T_pM$ and a section $s \in \Gamma(E)$ we have

$$\nabla_{X_p} s = \kappa(ds_p X_p).$$

Proof. It suffices to prove the theorem locally, meaning we can assume that $M = \mathbb{R}^m$ and that E is the direct product $\mathbb{R}^m \times \mathbb{R}^r$, that is the trivial bundle of rank r over $\mathbb{R}^m$.

We denote the standard coordinates on $M = \mathbb{R}^m$ by $(x^i)_{i=1}^m$, and the corresponding standard vector fields $(\partial_{x^i})_{i=1}^m$. Similarly we denote the standard coordinates on $E = \mathbb{R}^m \times \mathbb{R}^r$ by $((x^i)_{i=1}^m, (v^\alpha)_{\alpha=1}^r)$ and the induced standard vector fields by $((\partial_{x^i})_{i=1}^m, (\partial_{v^\alpha})_{\alpha=1}^r)$. We see that $TM = T\mathbb{R}^m$ and $TE = T(\mathbb{R}^m \times \mathbb{R}^r)$ are isomorphic to $\mathbb{R}^m \times \mathbb{R}^m$ and $(\mathbb{R}^m \times \mathbb{R}^r) \times (\mathbb{R}^m \times \mathbb{R}^r)$ respectively, through the isomorphisms

$$\mathbb{R}^m \times \mathbb{R}^m \rightarrow TM, (x, X) \rightarrow \sum_{i=1}^m X^i \partial_{x^i}(x)$$

and

$$(\mathbb{R}^m \times \mathbb{R}^r) \times (\mathbb{R}^m \times \mathbb{R}^r) \rightarrow TE, (x, v, X, V) \rightarrow \sum_{i=1}^m X^i \partial_{x^i}(x, v) + \sum_{\alpha=1}^r V^\alpha \partial_{v^\alpha}(x, v).$$

We can already verify (i): here the vertical bundle VE is isomorphic to $\mathbb{R}^m \times \mathbb{R}^r \times \{0\} \times \mathbb{R}^r$, that is

$$VE = \left\{ \sum_{\alpha=1}^r V^\alpha \partial_{v^\alpha}(x, v) \mid (x, v, V) \in \mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^r \right\}$$

while the pullback bundle π^*E is isomorphic to $\mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^r$. The map $\zeta : \pi^*E \rightarrow VE$ is then simply the obvious bundle morphism $\mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^r \simeq \mathbb{R}^m \times \mathbb{R}^r \times \{0\} \times \mathbb{R}^r$, indeed for $(x, v, V) \in \pi^*E = \mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^r$ we have

$$\zeta_{(x,v)} V = \left. \frac{d}{dt} \right|_{t=0} (x, v + tV) = \sum_{\alpha=1}^r V^\alpha \partial_{v^\alpha}(x, v).$$

A vector field X on $M = \mathbb{R}^m$ is uniquely represented by a family of smooth maps $X^i : \mathbb{R}^m \rightarrow \mathbb{R}$ such that $X(x) = \sum_{i=1}^m X^i(x) \partial_{x^i}(x)$ and a section s of $E = \mathbb{R}^m \times \mathbb{R}^r$ is uniquely represented by a family smooth maps $s^\alpha : \mathbb{R}^m \rightarrow \mathbb{R}$ $1 \leq \alpha \leq r$ such that $s(x) = (x, (s^\alpha(x))_{\alpha=1}^r)$. The connection ∇ is then encoded by a family of smooth maps $A_{\beta j}^\alpha : M = \mathbb{R}^m \rightarrow \mathbb{R}$ indexed by $1 \leq \alpha, \beta \leq r$ and $1 \leq j \leq m$, such that for a vector field

X and a section s , we have

$$(\nabla_X s)^\alpha = \sum_{i=1}^m X^i \frac{\partial s^\alpha}{\partial x^i} + \sum_{\beta=1}^r \sum_{j=1}^m A_{\beta j}^\alpha X^j s^\beta$$

for all $1 \leq \alpha \leq r$ (see equation 2.5).

For $e = (x, v) \in E = \mathbb{R}^m \times \mathbb{R}^r$, a section s with $s(x) = e$ (such a section always exists) and a tangent vector $X_x \in T_x M \simeq \mathbb{R}^m$ we have

$$ds_x(X_x) = \sum_{i=1}^m X_x^i \partial_{x^i}(x, v) + \sum_{\alpha=1}^r X_x^i \frac{\partial s^\alpha}{\partial x^i}(x) \partial_{v^\alpha}(x, v) \quad (2.25)$$

while

$$\zeta_e \nabla_{X_x} s = \sum_{\alpha=1}^r \left(\sum_{i=1}^m X_x^i \frac{\partial s^\alpha}{\partial x^i}(x) + \sum_{\beta=1}^r \sum_{j=1}^m A_{\beta j}^\alpha(x) X_x^j s^\beta(x) \right) \partial_{v^\alpha}(x, v)$$

so that

$$h_e(X_x) := ds_x(X_x) - \zeta_e \nabla_{X_x} s = \sum_{i=1}^m X_x^i \partial_{x^i}(x, v) - \sum_{\alpha, \beta=1}^r \sum_{j=1}^m A_{\beta j}^\alpha(x) X_x^j v^\beta \partial_{v^\alpha}(x, v). \quad (2.26)$$

From this we can easily see that $ds_x(X_x) - \zeta_e \nabla_{X_x} s$ does not depend on the choice of section s , so that h_e is well defined. From the expression of h_e we also check immediately that h_e is linear and that $d\pi_e(h_e(X_x)) = X_x$, in particular h_e is injective. Furthermore if $h_e(X_x)$ belongs to $V_e E = \ker d\pi_e$ then

$$X_x = d\pi_e(h_e(X_x)) = 0$$

and $h_e(X_x) = h_e(0) = 0$, thus if we define

$$H_e E := h_e(T_p M) = \left\{ \sum_{i=1}^m X^i \partial_{x^i}(e) - \sum_{\alpha, \beta=1}^r \sum_{j=1}^m A_{\beta j}^\alpha(x) X^j v^\beta \partial_{v^\alpha}(e) \mid X \in \mathbb{R}^m \right\}$$

we have $H_e E \cap V_e E = 0$. Combining this with the injectivity of h_e we get

$$m = \dim(T_x M) \leq \dim(H_e E) \leq \dim(T_e E) - \dim(V_e E) = (m + r) - r = m$$

from which we deduce that $\dim(H_e E) = m$, $T_e E = V_e E \oplus H_e E$, and that $h_e : T_x M \rightarrow H_e E$ is an isomorphism with inverse $d\pi_e$.

Let us now verify that HE is a smooth distribution. For $1 \leq i \leq m$ we define the vector field $\tilde{\partial}_{x^i}$ on E by

$$\tilde{\partial}_{x^i}(x, v) := h_{(x, v)}(\partial_{x^i}(x)) = \partial_{x^i}(x, v) - \sum_{\alpha, \beta=1}^r A_{\beta i}^\alpha(x) v^\beta \partial_{v^\alpha}(x, v)$$

for all $(x, v) \in E$. Clearly the $(\tilde{\partial}_{x^i})_{i=1}^m$ are smooth and for each $(x, v) \in E$, $(\tilde{\partial}_{x^i}(x, v))_{i=1}^m$ is a basis of $H_{(x,v)}E$ (since the $\partial_{x^i}(x)$ form a basis of T_xM). By the subbundle criterion (theorem 2.1.1), this shows that HE is a smooth distribution of rank m on E .

Lastly, let us examine the vertical projection. For $e = (x, v) \in E$, since $T_eE = V_eE \oplus H_eE$, there exists a unique linear projection $p_{v,e} : T_eE \rightarrow V_eE$ whose kernel is H_eE . Since $d\pi_e \circ h_e = id_{T_xM}$ we have for all $W \in T_eE$

$$d\pi_e(W) = (d\pi_e \circ h_e \circ d\pi_e)(W)$$

implying that $W - (h_e \circ d\pi_e)(W)$ is vertical, so

$$W = (W - (h_e \circ d\pi_e)(W)) + (h_e \circ d\pi_e)(W)$$

where the left term is vertical and the right term horizontal. This yields

$$p_{v,e}(W) = W - (h_e \circ d\pi_e)(W)$$

and using the explicit expression for h_e obtained in 2.26, we get

$$p_{v,e} \left(\sum_{i=1}^m X^i \partial_{x^i}(e) + \sum_{\alpha=1}^r V^\alpha \partial_\alpha(e) \right) = \sum_{\alpha=1}^r \left(V^\alpha + \sum_{\beta=1}^r \sum_{i=1}^m A_{\beta i}^\alpha(x) X^i v^\beta \right) \partial_\alpha(e).$$

Using this last formula, it is not difficult to convince oneself that p_v is indeed a smooth morphism of vector bundles. The connector κ is given by

$$\begin{aligned} \kappa : TE &\rightarrow E, \\ \sum_{i=1}^m X^i \partial_{x^i}(x, v) + \sum_{\alpha=1}^r V^\alpha \partial_\alpha(x, v) &\rightarrow \left(V^\alpha + \sum_{\beta=1}^r \sum_{i=1}^m A_{\beta i}^\alpha(x) X^i v^\beta \right)_{\alpha=1, \dots, r} \in E_x \end{aligned}$$

for all $(x, v) \in E$, $X \in \mathbb{R}^m$ and $V \in \mathbb{R}^r$. From this and the expression for $ds_x(X_x)$ computed in 2.25 we obtain that

$$\nabla_{X_x} s = \kappa ds_x(X_x)$$

thus concluding the proof. $\square$

Let us make precise the notion of **horizontal lift** of a tangent vector or vector field. A tangent vector $W \in T_eE$ on E is said to be horizontal if $W \in H_eE$, and a vector field W on E is said to be horizontal if $W(e)$ is horizontal for all $e \in E$.

Given a tangent vector $X_p \in T_pM$, and an element e of the fiber E_p over p , $\tilde{X}_e := h_e(X_p)$ is the unique horizontal vector $\tilde{X}_e \in H_eE$ such that $d\pi_e(\tilde{X}_e) = X_p$. We call $\tilde{X}_e$ the horizontal lift of X_p at e . Given a smooth vector field X on M ,

$$\tilde{X} : E \rightarrow TE, e \rightarrow h_e(X_p)$$

defines the unique horizontal vector field $\tilde{X}$ on E that projects onto X . We call $\tilde{X}$ the horizontal lift of X . Starting from the computations in the proof of theorem 2.2.9, little work is needed to obtain an local expression for $\tilde{X}$, and one can check that $\tilde{X}$ is smooth.

Alternatively, by theorem 2.2.9, $d\pi : HE \rightarrow TM$ is a bundle morphism over $\pi : E \rightarrow M$ which is fiberwise isomorphic, so we have a natural isomorphism

$$HE \simeq \pi^*TM.$$

In other words $(HE, d\pi)$ is a pullback of TM along $\pi : E \rightarrow M$, and $\tilde{X}$ can be viewed as the pullback section induced by X .

Now we move on to the relationship between this horizontal distribution and parallel transport. We start an easy but important consequence of theorem 2.2.9, telling us that integral submanifolds of the horizontal distribution HE (appendix A) should be thought of as parallel sections of E (at least locally).

Corollary 2.2.5. *Let $s \in \Gamma(U, E)$ be a local section of E . The following are equivalent:*

- (i) s is parallel for ∇ .
- (ii) For all $p \in U$, the image of ds_p lies in the horizontal subspace $H_{s(p)}E$.
- (iii) The image of s is an integral submanifold of the horizontal distribution HE .

Proof. (i) $\Leftrightarrow$ (ii) is a consequence of $\nabla_{X_p}s = \kappa(ds_p X_p)$ in point (iv) of theorem 2.2.9.

(ii) $\Leftrightarrow$ (iii) is clear once one recalls that s is an injective immersion of rank m (the rank of the horizontal distribution HE). $\square$

As we defined parallel transport through the pullback connection along a curve, we need to understand how the horizontal distribution behaves under pullbacks. Thankfully, this behaviour is very natural.

Theorem 2.2.10. *Let E be a vector bundle over N with connection ∇ . Let $\phi : M \rightarrow N$ be a smooth map, let (ϕ^*E, Φ) be the pullback of E along ϕ and denote the pullback connection on ϕ^*E by ∇^ϕ .*

Then for any tangent vector $W \in T_e\phi^*E$ on ϕ^*E , W is horizontal if and only if $d\Phi_e W \in T_{\Phi(e)}E$ is horizontal.

Proof. Fix a point p of M and an element e of the fiber $(\phi^*E)_p$ over p . We can assume that (ϕ^*E, Φ) is the standard pullback of E , so Φ restricts to the identity on fibers and e is an element of $E_{\phi(p)}$. Choose a section s of E such that s takes the value e at $\phi(p)$. Then the pullback section $\phi^*s = \Phi \circ \phi^*s = s \circ \phi$ is a section of ϕ^*E that takes value e at p .

Let $h_e : T_{\phi(p)}N \rightarrow T_eE$ and $h_e^\phi : T_pM \rightarrow T_e\phi^*E$ be the horizontal liftings at e for ∇ and ∇^ϕ respectively, and let ζ, ζ^ϕ be the bundle isomorphisms

$$\begin{aligned}\zeta : \pi^*E &\xrightarrow{\sim} VE \\ \zeta^\phi : \pi^*\phi^*E &\xrightarrow{\sim} V\phi^*E\end{aligned}$$

For $v \in \phi^*E_p = E_{\phi(p)}$ we can check that

$$d\Phi_e \zeta_e^\phi(v) = \left. \frac{d}{dt} \right|_{t=0} \Phi(e + tv) = \left. \frac{d}{dt} \right|_{t=0} (e + tv) = \zeta_e(v)$$

and by definition of the pullback connection we have for all $X_p \in T_pM$

$$\nabla_{X_p}^\phi(\phi^*s) = \nabla_{d\phi_p X_p} s$$

so we can compute

$$\begin{aligned}d\Phi_e h_e^\phi(X_p) &= d(\Phi \circ \phi^*s)_p(X_p) - d\Phi_e \zeta_e^\phi \nabla_{X_p}^\phi(\phi^*s) \\ &= ds_{\phi(p)} d\phi_p(X_p) - \zeta_e \nabla_{d\phi_p X_p} s \\ &= h_e(d\phi_p X_p).\end{aligned}$$

where we use $\Phi \circ \phi^*s = s \circ \phi$. This shows that

$$d\Phi_e \circ h_e^\phi = h_e \circ d\phi_p \tag{2.27}$$

which immediately implies that $d\Phi_e$ maps the horizontal subspace $H_e\phi^*E = h_e^\phi(T_pM)$ into the horizontal subspace $H_eE = h_e(T_{\phi(p)}N)$.

Conversely, let $W \in T_e\phi^*E$ and assume that $d\Phi_e W \in T_eE$ is horizontal. Define $X_p := d\pi_e W$. Since $d\pi_e h_e^\phi = id_{T_pM}$ we get

$$0 = d\pi_e(W - h_e^\phi X_p) \tag{2.28}$$

so $W - h_e^\phi X_p$ lies in the vertical subspace $V_e\phi^*E$. Furthermore, since $d\Phi_e W$ is horizontal we have

$$\begin{aligned}d\Phi_e W &= (h_e \circ d\pi_e)(d\Phi_e W) \\ &= h_e(d\pi_e \circ d\Phi_e)W \\ &= h_e(d\phi_p \circ d\pi_e)W \\ &= (h_e \circ d\phi_p)X_p \\ &= d\Phi_e(h_e^\phi X_p)\end{aligned}$$

where we use $\pi \circ \Phi = \phi \circ \pi$ and 2.27. Combining this last equation with 2.26 we see that $W - h_e^\phi X_p$ lies in the intersection of the vertical subspace $V_e\phi^*E$ with the kernel of $d\Phi_e$. But $d\Phi_e$ maps $V_e\phi^*E$ isomorphically onto V_eE (as Φ maps ϕ^*E_p isomorphically

onto $E_{\phi(p)}$), so $\ker(d\Phi_e) \cap V_e\phi^*E = 0$ and $W = h_e^\phi X_p$ is horizontal. $\square$

Corollary 2.2.6. *A section s of π^*E is parallel for ∇^ϕ if and only if for all $p \in M$ the image of $d(\Phi \circ s)_p$ lies in the horizontal subspace $H_{(\Phi \circ s)(p)}E$.*

For a curve $\gamma : [a, b] \rightarrow M$, a horizontal lift of γ is defined to be a curve $s : [a, b] \rightarrow E$ such that

- (i) $\pi \circ s = \gamma$.
- (ii) $\dot{s}(t)$ is horizontal for all $t \in [a, b]$.

From corollary 2.2.6, we see that a map $s : [a, b] \rightarrow E$ is a horizontal lift of γ if and only if it is a parallel section over γ . Theorem 2.2.4 on the properties of parallel transport then give us:

Theorem 2.2.11. *Consider a curve $\gamma : [a, b] \rightarrow M$ and an element e of the fiber E_a . Then there exists a unique horizontal lift $s : [a, b] \rightarrow E$ of γ such that $s(a) = e$. Furthermore, the parallel translation of e along γ is given by*

$$P(\gamma)_a^t e = s(t)$$

for all $a \leq t \leq b$.

We conclude this section and this chapter by studying the *integrability* of the horizontal distribution which, in view of corollary 2.2.5, should give us some information on the abundance or scarcity of parallel sections. We will see that the curvature of our connection can be seen as a measure of said integrability.

We begin by showing that the problem of determining whether HE is involutive is a 'pointwise problem' (see appendix A for the notion of involutivity for distributions).

Proposition 2.2.12. *Let $\Gamma(H)$ and $\Gamma(V)$ denote the $C^\infty(E)$ -modules of horizontal and vector fields (that is, sections of HE and VE). The map*

$$C : \Gamma(H) \times \Gamma(H) \rightarrow \Gamma(V), (Z_1, Z_2) \rightarrow p_v([Z_1, Z_2])$$

is $C^\infty(E)$ -linear.

Proof. Since C is manifestly skew-symmetric, it suffices to show $C^\infty(E)$ -linearity for the first argument. So let $f \in C^\infty(E)$. We compute

$$\begin{aligned} C(fZ_1, Z_2) &= p_v([fZ_1, Z_2]) = p_v(f[Z_1, Z_2] - (Z_2f)Z_1) \\ &= fp_v([Z_1, Z_2]) - (Z_2f)p_v(Z_1) = fp_v([Z_1, Z_2]) \\ &= fC(Z_1, Z_2). \end{aligned}$$

$\square$

Corollary 2.2.7. C is a section of $VE \otimes \Lambda^2 HE^*$, in particular for $e \in E$ the value $C(Z_1, Z_2)(e) \in V_e E$ depends only on $Z_1(e)$ and $Z_2(e)$.

The core observation is then the following:

Theorem 2.2.12. Let $X_p, Y_p \in T_p M$ and $e \in E_p$, and let $X_e, Y_e \in H_e E$ denote the horizontal lifts of X_p and Y_p at e . Then

$$R(X_p, Y_p)e = -\zeta_e^{-1}C(X_e, Y_e)$$

Proof. See for example theorem 12.50 in [17] (section 12.5 p.529). $\square$

So we see that the horizontal distribution is integrable if and only if the connection is **flat**, meaning that the curvature vanishes ($R = 0$). By the Frobenius theorem, some theory of covering spaces and theorem 2.2.5, we then obtain.

Corollary 2.2.8. A connection ∇ on a vector bundle is **flat** if and only if it is **locally trivial**. More precisely, the following are equivalent:

- (i) The curvature R vanishes.
- (ii) The horizontal distribution HE is integrable.
- (iii) For any simply connected open subset U of M , the restriction of ∇ to $E_U = \pi^{-1}(U)$ is a trivial connection.
- (iv) For any simply connected open subset U of M , any point $p \in U$ and any element $e \in E_p$ there exists a unique local parallel section $s \in \Gamma(U, E)$ such that $s(p) = e$.

When conditions (i)-(iv) we say that the connection is **flat**.

Proof. (i) $\Leftrightarrow$ (ii) follows from theorem 2.2.12 and the Frobenius theorem (see appendix A).

(iii) $\Leftrightarrow$ (iv) is a consequence of theorem 2.2.5.

(iv) $\Rightarrow$ (ii) is clear enough given corollary 2.2.5.

(ii) $\Rightarrow$ (iv) requires a bit more care, the idea is the following: we take the maximal connected integral manifold L_e of HE_U containing e , deduce that $\pi : L_e \rightarrow U$ is a covering space from the fact that any smooth in U can be uniquely lifted to L_e (theorem 2.2.11) then use the simple connectedness of U to conclude that $\pi : L_e \rightarrow U$ is invertible, the section we seek is then $(\pi|_{L_e})^{-1}$. These two steps require some knowledge of the theory of covering spaces, we refer the reader to [17] (theorem 12.25, p.517) and the references therein. $\square$

Chapter 3

Affine Connections

In this chapter we bring to bear our theory of connections on vector bundles by applying it to the tangent bundle, leading us to the notion of an affine connection. The tautological nature of the tangent bundle grants affine connections with properties shared by no other kind of connections, chief among them: the existence of geodesics, i.e. curves with parallel (think 'constant') velocity field (as evidence of the importance of geodesics, we will learn that a torsionless affine connection is fully determined by its geodesics).

Through the theory of geodesics we will define geodesic symmetries (by which a torsionless connection is also determined) and symmetric tensors. In anticipation of subsequent chapters, we will obtain our first curvature condition linked to the existence of symmetric tensors.

The main references for this chapter are [12, 13] and [24].

3.1 Definition, torsion and Bianchi identities

Definition 3.1.1 (Affine connections). Let M be a smooth manifold. An **affine connection** on M is a connection on the tangent bundle TM . We will sometimes refer two manifolds equipped with affine connections as **affinely connected manifolds**.

Given two affinely connected manifolds (M, ∇) , (M', ∇') , an **affine transformation** $(M, \nabla) \rightarrow (M', \nabla')$ is a diffeomorphism $f : M \rightarrow M'$ such that the induced bundle isomorphism $df : TM \rightarrow TM'$ (which lies over f) preserves the connection in the sense of definition 2.2.1, meaning that for any pair of vector fields $X, Y \in \mathfrak{X}(M)$ we have

$$f_*(\nabla_X Y) = \nabla_{f_* X}(f_* Y).$$

For affine connections, in addition to the curvature, we can define a second local invariant, called the **torsion**.

Definition 3.1.2 (Torsion). Let ∇ be an affine connection on M . The map

$$T : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M), (X, Y) \rightarrow T(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y]$$

defines a $(1, 2)$ -tensor called the **torsion** of ∇ . We say that ∇ is **torsionless** when $T = 0$.

In this thesis we will be mainly concerned with torsionless connections.

Using $\nabla_{\partial_i} \partial_j = \sum_{k=1}^m \Gamma_{ij}^k \partial_k$ and $[\partial_i, \partial_j] = 0$ we get a very simple formula relating the components of T to the Christoffel symbols:

$$T_{ij}^k = dx^k T(\partial_i, \partial_j) = \Gamma_{ij}^k - \Gamma_{ji}^k.$$

Thus we see that requiring an affine connection to be torsionless is equivalent to requiring that its Christoffel symbols Γ_{ij}^k be symmetric their ij indices, i.e. that $\Gamma_{ij}^k = \Gamma_{ji}^k$. This suggests an obvious process for extracting a torsionless connection from an arbitrary one: could we simply take the symmetrization of the Christoffel symbols (for their 'lower' ij indices)? The answer to this question is affirmative. Indeed, if we define the family of symbols

$$\bar{\Gamma}_{ij}^k := \Gamma_{ji}^k$$

one can check that they transform according to equation 2.9 (example 2.2.3), so that they define an affine connection $\bar{\nabla}$ on M . Then, as we saw in proposition 2.2.4, the convex combination

$$\hat{\nabla} = \frac{1}{2}(\nabla + \bar{\nabla})$$

also defines an affine connection on M . It is now straightforward to check that the Christoffel symbols $\hat{\Gamma}_{ij}^k$ of $\hat{\nabla}$ are given by

$$\hat{\Gamma}_{ij}^k = \frac{1}{2}(\Gamma_{ij}^k + \bar{\Gamma}_{ij}^k) = \frac{1}{2}(\Gamma_{ij}^k + \Gamma_{ji}^k) \quad (3.1)$$

in particular, they are symmetric in their lower indices, which shows that $\hat{\nabla}$ is torsionless.

Before we start reviewing the distinctive features exhibited by the curvature of a (torisonless) affine connection, we give a very neat result relating a torisonless connection to the **exterior derivative** of differential forms.

Theorem 3.1.1. *Let ∇ be a torsionless connection on M then for a differential form ω of degree k , we have*

$$d\omega = (k+1) \text{Alt}(\nabla\omega).$$

In particular, if a differential form ω is parallel for a torsionless connection, then ω is closed.

Proof. Repeated applications of Cartan's magic formula yield the following 'intrinsic formula' for the exterior derivative

$$\begin{aligned} d\omega(X_1, \dots, X_{k+1}) &= \sum_{i=1}^{k+1} (-1)^{i-1} X_i[\omega(X_1, \dots, \hat{X}_i, \dots, X_{k+1})] \\ &\quad + \sum_{i < j} (-1)^{i+j} \omega([X_i, X_j], X_1, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_{k+1}) \end{aligned} \quad (3.2)$$

where once again $\hat{X}_i$ and $\hat{X}_j$ indicate that X_i and X_j are omitted. From 3.2, the vanishing of the torsion

$$[X_i, X_j] = \nabla_{X_i} X_j - \nabla_{X_j} X_i$$

and the following formula for the covariant derivative of differential forms (and more generally of covariant tensors - definition 2.2.7 (v))

$$\begin{aligned} (\nabla_{X_i} \omega)(X_1, \dots, \hat{X}_i, \dots, X_{k+1}) &= X_i[\omega(X_1, \dots, \hat{X}_i, \dots, X_{k+1})] \\ &\quad - \sum_{j < i} \omega(X_1, \dots, \nabla_{X_i} X_j, \dots, \hat{X}_i, \dots, X_{k+1}) \\ &\quad - \sum_{i < j} \omega(X_1, \dots, \hat{X}_i, \dots, \nabla_{X_i} X_j, \dots, X_{k+1}) \end{aligned}$$

it is a routine exercise to derive

$$d\omega(X_1, \dots, X_{k+1}) = \sum_{i=1}^{k+1} (-1)^{i-1} (\nabla_{X_i} \omega)(X_1, \dots, \hat{X}_i, \dots, X_{k+1}).$$

Finally, using the fact that ω is skew-symmetric (which implies that $\nabla\omega$ is skew-symmetric in its last k -arguments) one finds

$$(k+1) \text{Alt}(\nabla\omega)(X_1, \dots, X_{k+1}) = \sum_{i=1}^{k+1} (-1)^{i-1} (\nabla_{X_i} \omega)(X_1, \dots, \hat{X}_i, \dots, X_{k+1}).$$

□

Of central importance to the study of affine connections are the Bianchi identities (they will intervene in many places in the following pages), a series of 'symmetries' verified by the curvature tensor and its covariant derivative. When the connection is torsionless they take a particularly simple form.

Theorem 3.1.2 (Bianchi Identities). *Let R be the curvature of a torsionless affine connection on M . Then for $X, Y, Z \in \mathfrak{X}(M)$ we have*

$$(i) \text{ (Bianchi's first identity) } \mathfrak{S}_{X,Y,Z} R(X, Y)Z = 0.$$

$$(ii) \text{ (Bianchi's second identity) } \mathfrak{S}_{X,Y,Z} (\nabla_X R)(Y, Z) = 0.$$

where $\mathfrak{S}$ denotes the cyclic sum.

Proof. From the definitions of T and R , a straightforward computation shows that (i) is equivalent to the Jacobi identity for vector fields

$$\mathfrak{S}_{X,Y,Z}[[X, Y], Z] = 0.$$

(this computation is carried out in [18], see proof of proposition 7.12, p.203).

(ii) can be proven in a similar fashion, but a genuinely illuminating proof is obtained by introducing the curvature and torsion forms (as well as the exterior covariant derivative) and then taking advantage of the structure equations they satisfy. For lack of time, however, these will not be treated here. Instead we refer the reader to [12] (vol I, theorem 5.3, p.135). $\square$

3.2 Ricci curvature and locally equiaffine manifolds

When working with an affine connection, the curvature is tensor field of type $(1, 3)$, and through contraction we can construct the **Ricci curvature**.

Definition 3.2.1. Let R be the curvature of an affine connection ∇ on M . Then the map

$$r : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow C^\infty(M), (X, Y) \mapsto r(X, Y) := \text{tr}(Z \mapsto R(X, Z)Y)$$

defines a 2-covariant tensor called the Ricci curvature of ∇ . Using our contraction notation, we have

$$r := C_2^1 R.$$

The purpose of this section is to make apparent the relationship between the Ricci curvature and the curvature of the connection induced on the determinant bundle, at least for torsionless connections (the hint that this relationship exists is to be found in the presence of the trace operation both in the definition of the Ricci curvature and in the induced action " $\cdot$ " on top forms - see example 1.1.4 point (ii)).

We begin with the definition of (locally) equiaffine manifolds, which are manifolds with affine connections for which there exist parallel (local) volume forms.

Definition 3.2.2 (Equiaffine manifolds). Let ∇ be an affine connection on a smooth manifold M . (M, ∇) is **equiaffine** if there exists a parallel volume form on M .

Definition 3.2.3 (Locally equiaffine manifolds). A **locally equiaffine manifold** is a smooth manifold M together with an affine connection ∇ on M such that for any point $p \in M$ there exists a parallel volume form defined on a neighborhood of p .

Of central importance is the fact that the determinant bundle is a **line bundle**.

Definition 3.2.4 (Line bundles). A **(real) line bundle** over a manifold M is a rank 1 (real) vector bundle over M .

This means that a non-vanishing section is a trivialization, and that a parallel volume form exists precisely when the determinant bundle connection is trivial (theorem 2.2.5). Consequently, in view corollary 2.2.8, the existence of local parallel volume forms is equivalent to the flatness of the connection on the determinant bundle. In other words, for line bundles corollary 2.2.8 takes the following form:

Theorem 3.2.1. *Let L be a real line bundle over a smooth manifold M , and let ∇ be a connection on L . The following are equivalent:*

- (i) *For any point $p \in M$ there exists a parallel non-vanishing local section of L defined on a neighborhood of p .*
- (ii) *The connection on L is flat.*
- (iii) *For any point $p \in M$, any element $s_p \in L_p$ and any simply connected neighborhood U of p , there exists a parallel section s of L defined on U whose value at p is s_p . In particular, if $s_p \neq 0$, s is non-vanishing.*

Everything comes together neatly in the case of torsionless connections, as the first Bianchi identity yields a very simple expression of the curvature on the determinant bundle in terms of the Ricci curvature, and we obtain:

Corollary 3.2.1. *Let ∇ be a torsionless affine connection on a smooth manifold M . The following are equivalent:*

- (i) *(M, ∇) is locally equiaffine.*
- (ii) *The Ricci curvature tensor of (M, ∇) is symmetric.*
- (iii) *For any point $p \in M$, any element $\omega_p \in \Lambda^{\text{top}} T_p^* M$ and any simply connected neighborhood U of p , there exists a parallel top form ω defined on U whose value at p is ω_p . In particular, if $\omega_p \neq 0$, ω is a volume form on U .*

Proof. This will follow from theorem 3.2.1 once we show that the induced connection on the determinant bundle is flat if and only if the Ricci tensor r is symmetric.

Let R be the curvature tensor of ∇ and let $\bar{R}$ be the curvature of the induced connection on the determinant bundle. Given two vector fields $X, Y \in \mathfrak{X}(M)$ and a top form $\omega \in \Omega^{\text{top}}(M)$ we use corollary 2.2.3 (and example 1.1.4 (ii)) to compute

$$\bar{R}(X, Y)\omega = R(X, Y) \cdot \omega = -(tr R(X, Y))\omega$$

so the connection on the determinant bundle is flat if and only if $\text{tr}R(X, Y) = 0$ for all $X, Y \in \mathfrak{X}(M)$. Using Bianchi's first identity for a torsionless connection and the skew-symmetry of the curvature tensor we obtain

$$R(X, Y)Z = -R(Z, X)Y - R(Y, Z)X = R(X, Z)Y - R(Y, Z)X$$

for all $X, Y, Z \in \mathfrak{X}(M)$, which implies

$$\text{tr}R(X, Y) = r(X, Y) - r(Y, X)$$

for all $X, Y \in \mathfrak{X}(M)$, finishing the proof. $\square$

3.3 Geodesics, exponential mapping and normal coordinates

The most important feature of affine connections, as the title of this thesis suggests, is the fact that they allow for the definition of **geodesics**, that is curves with parallel (think 'constant') velocity field. These geodesics are to our affinely connected manifolds what straight lines are to euclidean spaces.

Definition 3.3.1 (Geodesics). Let ∇ be an affine connection on a manifold M . A **geodesic** on (M, ∇) is a curve $\gamma : I \rightarrow M$ (with I an interval of any kind) such that $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$. In other words a curve γ is a geodesic if the velocity field $\dot{\gamma}$ is parallel along γ .

One should keep in mind that geodesics are *parametrized* curves, and should not be strictly identified with their images. However, using the fact that parallel transport does not depend on parametrization, we can impose strong restrictions on reparametrizations of geodesics.

Proposition 3.3.1 (Reparametrization of geodesics). *Let $\gamma : I \rightarrow M$ be a (non-constant) geodesic and let $\phi : J \rightarrow I$ be a smooth map (with J another interval). Then $\gamma \circ \phi$ is a geodesic if and only if ϕ is an affine function, meaning that there exists constants $\alpha, \beta \in \mathbb{R}$ such that*

$$\phi(t) = \alpha t + \beta$$

for all $t \in J$.

Proof. Let $\gamma : I \rightarrow M$ be a geodesic, and assume that $\gamma' := \gamma \circ \phi$ is a geodesic for some smooth map $\phi : J \rightarrow I$. Without loss of generality, assume that $0 \in J$. Then for all $t \in J$,

$$\dot{\gamma}'(t) = \dot{\phi}(t)(\dot{\gamma} \circ \phi)(t).$$

Since parallel transport is not sensitive to reparametrization (proposition 2.2.6), we must have for all $t \in J$

$$\dot{\phi}(0)(\dot{\gamma} \circ \phi)(0) = P(\gamma \circ \phi)_t^0 \dot{\gamma}'(t) = \dot{\phi}(t)(\dot{\gamma} \circ \phi)(0).$$

Assuming that γ is non-constant (so that $(\dot{\gamma} \circ \phi)(0) \neq 0$), $\dot{\phi}$ must then be constant, which implies that ϕ is an affine map. $\square$

Choosing a coordinate chart $(U, x^1, \dots, x^m)$ allows us to characterize geodesics locally in terms of a second order linear ODE. To achieve this, we simply adjust equation 2.11 to the case $A_l^k = \sum_{i=1}^m \Gamma_{il}^k dx^i$ and $s^k(t) = dx^k(\dot{\gamma}(t)) =: \dot{\gamma}^k(t)$, and we see that a smooth curve $\gamma : J \rightarrow U$ is a geodesic if and only if

$$0 = \ddot{\gamma}^k + \sum_{i,j=1}^m (\Gamma_{ij}^k \circ \gamma) \dot{\gamma}^i \dot{\gamma}^j, k = 1, \dots, m. \quad (3.3)$$

This and standard theory of linear ODE's is sufficient to develop the theory of geodesics on an affinely connected manifold. Nonetheless, we would like to present another, arguably more geometric approach, based on the notion of a *spray*.

Definition 3.3.2 (Geodesic spray). Let ∇ be an affine connection on M . The **geodesic spray** of ∇ is the unique smooth horizontal vector field W on TM such that for any $\xi \in T_p M$, $\pi(W_\xi) = \xi$. Equivalently W is defined by $W_\xi = h_\xi(\xi)$ for any tangent vector $\xi \in TM$, where h_ξ denotes the horizontal lifting at ξ .

Proof. The existence and uniqueness of the field W are not in question, and its smoothness should come as no surprise given the 'naturalness' of its definition. Still, we provide a local expression for W , which will allow us ascertain smoothness, but will also prove useful in later endeavours.

This is easily done by taking equation 2.26 in the proof of theorem 2.2.9, and injecting $A_l^k = \sum_{i=1}^m \Gamma_{il}^k dx^i$ which gives us

$$W_\xi = h_\xi(\xi) = \sum_{i=1}^m \xi^i \partial_{x^i}(\xi) - \sum_{i,j,k=1}^m (\Gamma_{ij}^k \circ \pi)(\xi) \xi^i \xi^j \partial_{v^k}(\xi) \quad (3.4)$$

for $\xi = \sum_{i=1}^m \xi^i \partial_i(p)$ (for any $p \in M$), where we use coordinates

$$(x, v) \rightarrow \sum_{i=1}^m v^i \partial_i(x)$$

on TM . $\square$

What makes the geodesic spray extremely useful for our purposes is that geodesics correspond exactly to integral curves of the geodesic spray.

Theorem 3.3.1. *There is a one-to-one correspondence between integral curves of the geodesic spray W and geodesics on M :*

- (i) *To a geodesic $\gamma : I \rightarrow M$ we associate the curve $\dot{\gamma} : I \rightarrow TM$ which is an integral curve of W .*
- (ii) *To an integral curve $\tilde{\gamma} : I \rightarrow TM$ of W we associate the projection $\pi \circ \tilde{\gamma} : I \rightarrow M$, which is a geodesic.*

And these two maps are inverse to each other.

Remark. The use of sprays extends beyond the theory of geodesics. They are a general tool for turning second-order ODE's on a manifold into first order ODE's on its tangent bundle (via the machinery of flows). This is analogous to the process whereby one can reduce the order of an ODE by adding variables (see [15], chapter IV, Sect. 3).

Proof. Let $\gamma : I \rightarrow M$ be a geodesic. $\dot{\gamma}$ is a section of TM over γ (that is a section of γ^*TM), so immediately we have $\pi(\dot{\gamma}) = \gamma$, and since $\dot{\gamma}$ is parallel along γ , by corollary 2.2.6, $\dot{\gamma}$ is horizontal, meaning that $\frac{d}{dt}\Big|_{t=s} \dot{\gamma}(t) \in T_{\dot{\gamma}(s)}TM$ lies in the horizontal subspace $H_{\dot{\gamma}(s)}TM$ for all $s \in I$. Finally for all $s \in I$ we have

$$\begin{aligned} d\pi_{\dot{\gamma}(s)} \left(\frac{d}{dt}\Big|_{t=s} \dot{\gamma}(t) \right) &= \frac{d}{dt}\Big|_{t=s} (\pi \circ \dot{\gamma})(t) \\ &= \frac{d}{dt}\Big|_{t=s} \gamma(t) \\ &= \dot{\gamma}(s) \end{aligned}$$

which, when combined with the fact that $\frac{d}{dt}\Big|_{t=s} \dot{\gamma}(t)$ is horizontal, shows that $\frac{d}{dt}\Big|_{t=s} \dot{\gamma}(t) = W_{\dot{\gamma}(s)}$. Thus $\dot{\gamma}$ is an integral curve of W .

Conversely, let $\tilde{\gamma} : I \rightarrow TM$ be an integral curve of W , and define $\gamma := \pi \circ \tilde{\gamma}$. Then for all $s \in I$, we have

$$\begin{aligned} \dot{\gamma}(s) &= \frac{d}{dt}\Big|_{t=s} \pi \circ \tilde{\gamma}(t) \\ &= d\pi_{\tilde{\gamma}(s)} \dot{\tilde{\gamma}}(s) \\ &= d\pi_{\tilde{\gamma}(s)} W_{\tilde{\gamma}(s)} \\ &= \dot{\gamma}(s) \end{aligned}$$

so $\dot{\gamma} = \tilde{\gamma}$ (which, among other things, shows that the two maps described in our statement are indeed inverses to each other). Lastly, as an integral curve of a horizontal vector field, $\dot{\gamma} = \tilde{\gamma}$ is parallel by corollary 2.2.6 and from this we conclude that γ is a geodesic. $\square$

Examining the equations 3.3 and 3.4 which determine geodesics and the geodesic spray, we see that Christoffel symbols appear in these equations as coefficients of a homogeneous polynomial of degree 2. And since

$$\sum_{i,j=1}^m \Gamma_{ij}^k y^i y^j = \sum_{i \leq j} \frac{1}{2} (\Gamma_{ij}^k + \Gamma_{ji}^k) y^i y^j$$

we see that what geodesics witness is precisely the symmetrization of the Christoffel symbols we used to extract a torsionless connection from an arbitrary connection in equation (3.1). This observation is enshrined in the following result.

Theorem 3.3.2. *Two torsionless affine connections on a manifold with the same geodesic spray (or equivalently, with the same geodesics) are equal. More generally, for any affine connection ∇ on a manifold M , there exists a unique torsionless affine connection on M with the same geodesic spray (or geodesics) as ∇ , whose Christoffel symbols are given by equation (3.1).*

In particular we see that a torsionless connection is completely determined by its geodesics.

Let us now bring forth the full power of the theory of flows through the geodesic spray.

Corollary 3.3.1. *Let $X_p \in T_p M$, then there exists a unique maximal geodesic $\gamma : I \rightarrow M$ such that $\dot{\gamma}(0) = X_p$. Furthermore, any other geodesic $\gamma' : J \rightarrow M$ such that $\dot{\gamma}'(0) = X_p$ is a restriction of γ , meaning that $J \subseteq I$ and $\gamma|_J = \gamma'$.*

Proof. This follows from theorem 3.3.1 together with the fact that there exists a unique maximal integral curve of W starting from X_p which extends all other integral curves of W that start from X_p . $\square$

We can see M as a properly embedded submanifold of TM via the zero section $0 : M \rightarrow TM$, this means identifying M with the submanifold

$$\{0_p \in T_p M | p \in M\} \simeq M \times \{0\}$$

one obtains by collecting the zero's of all the spaces tangent to M (here we use the subscript p to indicate that 0_p is the origin of the tangent space $T_p M$).

Let Θ_t denote the flow of the geodesic spray W . In virtue of the fundamental theorem for flows, there exists an open subset N of TM over which Θ_1 is well-defined. It turns out that M (viewed as a submanifold of TM in the natural way described above) is included in the open N . Indeed, we can see that W vanishes on M from its definition, which implies that for $p \in M$ the maximal integral curve of W starting from $p \simeq 0_p \in T_p M$ is simply the constant map

$$\mathbb{R} \rightarrow TM, t \rightarrow p \simeq 0_p.$$

In particular the integral curve starting from p is defined on the whole of $\mathbb{R}$ so the flow Θ_t is well-defined on M for all $t \in \mathbb{R}$. This holds in particular for Θ_1 , hence N must be an open neighborhood of M . By projecting Θ_1 onto M we obtain the exponential map:

Definition 3.3.3 (Exponential mapping). The (geodesic) **exponential mapping** of (M, ∇) is defined to be the composition

$$\exp := \pi \circ \Theta_1 : N \subseteq TM \rightarrow M.$$

We can interpret the exponential mapping in the following way: for a point $p \in M$ and a tangent vector $X_p \in T_p M \cap N$, if $\gamma : I \rightarrow M$ is the unique (maximal) geodesic such that $\dot{\gamma}(0) = X_p$, then I contains 1 and

$$\exp(X_p) = \gamma(1).$$

Remark. Notice that for $X_p \in T_p M$, if we define I to be the interval

$$I := \{t \in \mathbb{R} | tX_p \in N\}$$

then $I \rightarrow M, t \rightarrow \exp(tX_p)$ is the maximal geodesic with initial velocity X_p (here we use the fact that we can reparametrize geodesics as in proposition 3.3.1).

Combining the exponential mapping with the *local inversion theorem*, we can obtain coordinate charts which are highly attuned to the affine connection.

Theorem 3.3.3 (Normal Coordinates). *Let p be a point of M . There exists an open neighbourhood V of 0 in $T_p M$, and an open neighbourhood U of p in M such that $\exp : V \rightarrow U$ is a diffeomorphism. We call such an open neighborhood of U a **normal neighborhood** of p .*

If in addition we choose a basis $b = (b_i)_{i=1}^m$ of $T_p M$, we obtain a linear isomorphism $b : \mathbb{R}^m \rightarrow T_p M$ and if we define $\tilde{V} := b^{-1}(V)$ then the diffeomorphism

$$\tilde{V} \subseteq \mathbb{R}^m \rightarrow U, x \rightarrow \exp \left(\sum_{i=1}^m x^i b_i \right)$$

*provides us with a coordinate chart on U . We call such a coordinate chart **normal coordinate system** centered at p and denote it by $(U, x^1, \dots, x^m)$.*

Proof. The tangent space of $T_p M$ at 0 is identified in a canonical way with $T_p M$ and for any $X_p \in T_p M$ we have $d\exp_p(X_p) = \frac{d}{dt} \Big|_{t=0} \exp(tX_p) = X_p$, so the differential of $\exp|_{T_p M}$ (the restriction of $\exp$ on $T_p M$) at p is nonsingular and the result follows from the local inversion theorem. $\square$

One striking property of these normal coordinates is that they are 'flat at the origin', where 'flat' refers to the vanishing of the Christoffel symbols.

Proposition 3.3.2. *Let ∇ be a torsionless connection on a manifold M . Let $(U, x^1, \dots, x^m)$ be a normal coordinate system centered at $p \in M$, then the Christoffel symbols Γ_{ij}^k of ∇ in these coordinates vanish at p .*

Proof. By definition of the exponential mapping, we see that for any $a = (a^i)_{i=1}^m \in \mathbb{R}^m$, the curve whose description in normal coordinates is given by

$$\gamma_a(t) = (ta^1, \dots, ta^m)$$

(with $t \in]-\epsilon, \epsilon[$ for some $\epsilon > 0$) is a geodesic. Using 3.3, the equation characterizing geodesics in a coordinate chart, we then have

$$0 = \sum_{i,j=1}^m \Gamma_{ij}^k(p) a^i a^j, k = 1, \dots, m.$$

Since this holds for any $a \in \mathbb{R}^m$, and since $\Gamma_{ij}^k(p) = \Gamma_{ji}^k(p)$ (we assume that ∇ is torsionless), the proposition follows. $\square$

This preliminary section on geodesics will end with the introduction of **Jacobi fields**. Jacobi fields are vector fields along geodesics that are defined by a second order linear ODE (called the **Jacobi equation**). We will make immediate use of Jacobi fields in the next section, when we derive the first terms of the "**radial power expansion**" of a tensor in normal coordinates.

In general, Jacobi fields are important because they represent **infinitesimal geodesic variations**, meaning that they are the 'derivatives' of 'paths in the space of geodesics', in a sense that will be made precise shortly.

Definition 3.3.4 (Jacobi fields). Let ∇ be a torsionless connection on M , and let $\gamma : I \rightarrow M$ be a geodesic. A vector field Z along γ is called a **Jacobi field** if it satisfies the following second order linear ODE, called the **Jacobi equation**:

$$\nabla_{\dot{\gamma}}^2 Z + R(Z, \dot{\gamma})\dot{\gamma} = 0$$

where $\nabla_{\dot{\gamma}}^2$ indicates that we apply $\nabla_{\dot{\gamma}}$ twice: first to the field Z to obtain $\nabla_{\dot{\gamma}} Z$, then to the field $\nabla_{\dot{\gamma}} Z$ to obtain $\nabla_{\dot{\gamma}}^2 Z := \nabla_{\dot{\gamma}}(\nabla_{\dot{\gamma}} Z)$.

Routinely, the theory of linear ODE's elucidates the structure of the space of Jacobi fields.

Proposition 3.3.3. *For $t_0 \in I$ and $Z_0, Z_1 \in T_{\gamma(t_0)}M$ there exists a unique Jacobi field Z along γ such that*

$$\begin{aligned} Z(t_0) &= Z_0 \\ (\nabla_{\dot{\gamma}} Z)(t_0) &= Z_1. \end{aligned}$$

In particular the set of Jacobi vector fields along γ is a real vector space of dimension $2m$.

We end on the interpretation of Jacobi fields as infinitesimal geodesic variations.

Definition 3.3.5 ((Infinitesimal) geodesic variations). Let $\gamma : I \rightarrow M$ be a geodesic.

(i) A **geodesic variation** of γ is a smooth map

$$f : I \times]-\epsilon, \epsilon[\rightarrow M, (t, s) \rightarrow f(t, s)$$

(for some $\epsilon > 0$) such that

$$f_s : I \rightarrow M, t \rightarrow f(t, s)$$

is a geodesic for any $s \in]-\epsilon, \epsilon[$, and such that $f_0 = \gamma$.

(ii) A vector field Z along γ is called an **infinitesimal geodesic variation** if there exists a geodesic variation

$$f : I \times]-\epsilon, \epsilon[\rightarrow M$$

such that for all $t \in I$,

$$Z(t) = \left. \frac{d}{ds} \right|_{s=0} f(t, s).$$

Theorem 3.3.4. *A vector field Z along γ is a Jacobi field if and only if it is an infinitesimal geodesic variation.*

Proof. We only show that an infinitesimal geodesic variation is a Jacobi field. The proof of the converse relies on the existence of geodesically convex neighborhoods, see [13] (vol. II, chapter VIII, theorem 1.2, p.64).

So let $f : I \times]-\epsilon, \epsilon[\rightarrow M, (t, s) \rightarrow f(t, s)$ be a geodesic variation of γ . We work on the pullback bundle f^*TM over $I \times]-\epsilon, \epsilon[$ equipped with the pullback connection ∇^f . We denote by ∂_t and ∂_s the coordinate vector fields of $I \times]-\epsilon, \epsilon[$ and define f_t, f_s to be the following two vector fields over f (i.e. sections of f^*TM):

$$\begin{aligned} f_t(t, s) &:= df_{(t,s)}(\partial_t) = \left. \frac{d}{dh} \right|_{h=0} f(t+h, s) \\ f_s(t, s) &:= df_{(t,s)}(\partial_s) = \left. \frac{d}{dh} \right|_{h=0} f(t, s+h). \end{aligned}$$

Finally we use the notations $\nabla_t := \nabla_{\partial_t}^f$ and $\nabla_s := \nabla_{\partial_s}^f$.

Now, since f is a geodesic variation, we have $\nabla_t f_t = 0$. Using proposition 2.2.10 (the curvature of the pullback connection) and recalling that $[\partial_t, \partial_s] = 0$ we obtain

$$R(f_t, f_s)f_t = R^f(\partial_t, \partial_s)f_t = \nabla_t \nabla_s f_t \quad (3.5)$$

where R^f denotes the curvature of ∇^f .

Lastly, using the fact that ∇ is torsionless and that $[\partial_t, \partial_s] = 0$ we get

$$\nabla_s f_t = \nabla_t f_s. \quad (3.6)$$

Combining 3.5 and 3.6 yields

$$0 = \nabla_t^2 f_s + R(f_s, f_t)f_t$$

and we obtain the Jacobi equation by evaluating at $s = 0$. □

3.3.1 Geodesic symmetries and symmetric tensors

Finally, more than halfway into the thesis, we are in a position to define the first expression that appears in title we chose.

Definition 3.3.6 (Local geodesic symmetries). Let U be a connected normal neighborhood of p , and write $V := \exp^{-1}(U)$. We say that U is **symmetric** if $-V = V$, meaning that for all $X_p \in V$, $-X_p \in V$ (we might also say that V is symmetric).

On such a symmetric normal neighborhood, the local geodesic symmetry centered at p is the involutive diffeomorphism $s_p : U \rightarrow U$ defined by $s_p(\exp(X)) = \exp(-X)$ for all $X \in V$.

One last thing of note: one easily checks that $s_p(p) = p$ and $d(s_p)_p = -I$, i.e. the differential of s_p at p is the opposite of the identity on $T_p M$.

Remark. If U and V are two symmetric normal neighborhoods of p , then the involutions $U \rightarrow U$ and $V \rightarrow V$ defined above coincide on $U \cap V$, hence why it makes sense to speak of *the* (local) geodesic symmetry centered at p .

We will now invest a bit of work to arrive at the correct notion of 'compatibility' between a tensor field and local geodesic symmetries. We begin by defining the pushforward of a tensor field along a diffeomorphism.

Definition 3.3.7 (Pushforward of a tensor along a diffeomorphism). Given a diffeomorphism $f : M \rightarrow N$ and a tensor field W of type (r, s) on M , the **pushforward of W along f** , which we denote by fW , is the tensor field of type (r, s) on N defined by

$$(fW)_y := df_{f^{-1}(y)} \cdot W_{f^{-1}(y)}$$

(where " $\cdot$ " indicates the operation defined in 1.1.4, see chapter 1).

Example 3.3.1. Some special cases of this pushforward operation:

(i) For a smooth function $\phi \in C^\infty(M)$, we have

$$f\phi = \phi \circ f^{-1}.$$

(ii) For a vector field $X \in \mathfrak{X}(M)$, fX coincides with the usual pushforward of X along f i.e. we have

$$fX = f_*X.$$

(iii) For a covariant k -tensor field α , the pushforward along f corresponds to the pull-back along f^{-1} i.e. we have

$$f\alpha = (f^{-1})^*\alpha.$$

This notion of pushforward suggests a first, and most obvious notion of compatibility: requiring that the tensors be **preserved** by the local geodesic symmetries.

Definition 3.3.8 (Tensor preserved by geodesic symmetries). A tensor field W on M is said to be **preserved by local geodesic symmetries** if for any point $p \in M$ there exists a symmetric normal neighbourhood U of p such that $s_p(W|_U) = W|_U$.

But since $d(s_p)_p = -I_p$, a tensor of odd degree preserved by geodesic symmetries must necessarily vanish, making this notion of compatibility excessively stringent. Nonetheless, this fact proves useful on several occasions.

Proposition 3.3.4. *Let W be a tensor field of odd degree, that is of type (r, s) with $r + s$ odd. If W is preserved by local geodesic symmetries, then $W = 0$.*

Proof. For any $p \in M$ we have

$$\begin{aligned} W_p &= d(s_p)_p \cdot W_p \\ &= (-I) \cdot W_p = (-1)^{r+s} W_p \\ &= -W_p \end{aligned}$$

so $W_p = 0$. □

The weaker notion of a **symmetric tensor** is the correct one for requiring compatibility.

Definition 3.3.9 (Symmetric Tensor). A tensor field W of type (r, s) is said to be **symmetric** if for any point $p \in M$ there exists a symmetric normal neighborhood U such that $s_p(W|_U) = (-1)^{r+s}W|_U$.

Remark. Notice that for a tensor W of even degree (that is of type (r, s) with $r + s$ even), W is symmetric if and only if W is preserved by local geodesic symmetries.

We now present the first theorem that could be described as something other than 'exposition of basic theory' (although we do not claim it to be unknown by any means), and which could furthermore serve as motivation for the study of geodesic symmetries.

Theorem 3.3.5. *Let ∇ be a torsionless connection on M , and let W be a tensor field of type (r, s) on M . Then for any vector field $X \in \mathfrak{X}(M)$ and any point $p \in M$ we have*

$$\nabla_{X_p} W = \frac{1}{2} \mathcal{L}_X (W - (-1)^{r+s} s_p W)_p.$$

Remark. When applied to vector fields, this theorem implies

$$\nabla_{X_p} Y = \frac{1}{2} [X, Y + s_{p*} Y]_p$$

showing that a torsionless connection is completely determined by its geodesic symmetries (more specifically the *germs* of its geodesic symmetries)¹.

The proof of this theorem relies on the flatness of normal coordinates at the origin (proposition 3.3.2) which holds only for torsionless connections, as well as the following two lemmas:

Lemma 3.3.1. *Let $(U, x^1, \dots, x^m)$ be a coordinate chart on M and write*

$$\begin{aligned} W &= \sum_{i_1, \dots, i_r, j_1, \dots, j_s} W_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s} \\ X &= \sum_i X^i \partial_i. \end{aligned}$$

¹In fact it can be shown that one can start from a family of local symmetries to obtain a torsionless connection. This relationship is made exact, and extended to connections *with* torsion in [23] using groupoids and jet spaces.

Then the components of $\mathcal{L}_X W$ in that coordinate chart are given by

$$\begin{aligned} (\mathcal{L}_X W)_{j_1 \dots j_s}^{i_1 \dots i_r} &= X(W_{j_1 \dots j_s}^{i_1 \dots i_r}) + \sum_{t=1}^s \sum_{k=1}^m W_{j_1 \dots j_{t-1} k j_{t+1} \dots j_s}^{i_1 \dots i_r} \partial_{j_t} X^k \\ &\quad - \sum_{t=1}^r \sum_{k=1}^m W_{j_1 \dots j_s}^{i_1 \dots i_{t-1} k i_{t+1} \dots i_r} \partial_k X^{i_t} \end{aligned}$$

where $X(W_{j_1 \dots j_s}^{i_1 \dots i_r}) = \sum_{k=1}^m X^k \partial_k W_{j_1 \dots j_s}^{i_1 \dots i_r}$.

Proof. The Lie derivative is a 0-derivation of the tensor algebra, so we have

$$\begin{aligned} \mathcal{L}_X W &= \sum_{i_1, \dots, i_r, j_1, \dots, j_s} X(W_{j_1 \dots j_s}^{i_1 \dots i_r}) \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s} \\ &\quad + \sum_{t=1}^r \sum_{i_1, \dots, i_r, j_1, \dots, j_s} W_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1} \otimes \dots \otimes \mathcal{L}_X \partial_{i_t} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s} \\ &\quad + \sum_{t=1}^s \sum_{i_1, \dots, i_r, j_1, \dots, j_s} W_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \dots \otimes \mathcal{L}_X dx^{j_t} \otimes \dots \otimes dx^{j_s}. \end{aligned}$$

Proving the statement is then just a matter of checking that for all $1 \leq k \leq m$

$$\begin{aligned} \mathcal{L}_X \partial_k &= [X, \partial_k] = - \sum_{l=1}^m \partial_k X^l \partial_l \\ \mathcal{L}_X dx^k &= d(\mathcal{L}_X x^k) = d(X x^k) = dX^k = \sum_{l=1}^m \partial_l X^k dx^l. \end{aligned}$$

□

Lemma 3.3.2. Let $(U, x^1, \dots, x^m)$ be a symmetric normal coordinate system centered at p and assume that U is symmetric. Let $W_{j_1 \dots j_s}^{i_1 \dots i_r}$ be the components of W in these coordinates. Then the components of $s_p W$ in these same coordinates are given by

$$(s_p W)_{j_1 \dots j_s}^{i_1 \dots i_r} = (-1)^{r+s} W_{j_1 \dots j_s}^{i_1 \dots i_r} \circ s_p$$

Proof. By the linearity of the $\cdot$ operation, point (i) in definition 1.1.4 and the fact that $(s_p)^{-1} = s_p$ we have

$$s_p W = \sum_{i_1, \dots, i_r, j_1, \dots, j_s} (W_{j_1 \dots j_s}^{i_1 \dots i_r} \circ s_p) (s_{p*} \partial_{i_1}) \otimes \dots \otimes (s_{p*} \partial_{i_r}) \otimes (s_p^* dx^{j_1}) \otimes \dots \otimes (s_p^* dx^{j_s})$$

Now since we're working in a normal coordinate system centered at p , s_p sends the point with coordinates $(x^1, \dots, x^m)$ to the point with coordinates $(-x^1, \dots, -x^m)$ so we get for

all $1 \leq k \leq r$

$$\begin{aligned} s_{p*}\partial_k &= -\partial_k \\ s_p^*dx^k &= d(s_p^*x^k) = -dx^k \end{aligned}$$

allowing us to conclude. $\square$

Proof of theorem 3.3.5. Combining lemma 3.3.1 and 3.3.2 (and recalling that $s_p(p) = p$), we get the following formula for the components of $(\mathcal{L}_X s_p W)_p$:

$$\begin{aligned} (-1)^{r+s}(\mathcal{L}_X s_p W)_{j_1 \dots j_s}^{i_1 \dots i_r}(p) &= X_p(W_{j_1 \dots j_s}^{i_1 \dots i_r} \circ s_p) + \sum_{t=1}^s \sum_{k=1}^m W_{j_1 \dots j_{t-1} k j_{t+1} \dots j_s}^{i_1 \dots i_r}(p) \partial_{j_t} X^k(p) \\ &\quad - \sum_{t=1}^r \sum_{k=1}^m W_{j_1 \dots j_s}^{i_1 \dots i_{t-1} k i_{t+1} \dots i_r}(p) \partial_k X^{i_t}(p) \end{aligned}$$

and since $d(s_p)_p = -I$ we have

$$X_p(W_{j_1 \dots j_s}^{i_1 \dots i_r} \circ s_p) = d(s_p)_p(X_p)W_{j_1 \dots j_s}^{i_1 \dots i_r} = -X_p W_{j_1 \dots j_s}^{i_1 \dots i_r}.$$

Thus, using lemma 3.3.1 a second time we find

$$\frac{1}{2}\mathcal{L}_X(W - (-1)^{r+s}s_p W)_{j_1 \dots j_s}^{i_1 \dots i_r}(p) = X_p W_{j_1 \dots j_s}^{i_1 \dots i_r}.$$

Lastly since the Christoffel symbols vanish at p (proposition 3.3.2) we have

$$(\nabla_{X_p} W)_{j_1 \dots j_s}^{i_1 \dots i_r} = X_p W_{j_1 \dots j_s}^{i_1 \dots i_r}$$

(see equation 2.23 for the component expression of the covariant derivative) which finishes the proof. $\square$

As an immediate and very useful consequence of theorem 3.3.5 we have:

Corollary 3.3.2. *A tensor that is symmetric for a torsionless connection is parallel for that connection.*

We finish this section by presenting a necessary curvature condition implied by the existence of a symmetric covariant tensor. This condition is the first in an infinite sequence of curvature conditions one obtains by computing the Taylor expansion of said tensor at the origin of a normal coordinate system (making use of the Jacobi equation - definition 3.3.4) then noticing that, in virtue of the tensor being symmetric, either all odd or all even (depending on the parity of the tensor) terms of the aforementioned expansion must vanish.

Even with this very informal description, it should be clear to the reader that this infinite sequence of curvature conditions becomes sufficient (for the tensor to be symmetric) when the structures involved are *analytic* (see appendix B) and thus coincide locally with their Taylor expansion. We will not go further than this first term however, in part because this will already be enough to fulfill our purposes, but also because computations become tedious very quickly as we move to higher orders.

This theme will make a second appearance in the last two chapters of this thesis, through the odd Ledger conditions. There, no mention of Jacobi fields will be made, only volume forms will be treated and higher order computations will be rendered tractable by means of a recurrence formula.

Without further ado, let us state this curvature condition.

Theorem 3.3.6. *Let ∇ be a torsionless connection on M . Given a vector field $X \in \mathfrak{X}(M)$, we define the endomorphism field $P_X := (\nabla_X R)(X, -)X$ whose action on vector fields is given by*

$$P_X : \mathfrak{X}(M) \rightarrow \mathfrak{X}(M), Y \rightarrow (\nabla_X R)(X, Y)X.$$

Then for any symmetric covariant tensor field α we have

$$P_X \cdot \alpha = 0$$

for all $X \in \mathfrak{X}(M)$, where $P_X \cdot \alpha$ is defined by $(P_X \cdot \alpha)(p) := (P_X)(p) \cdot \alpha(p)$ for all points $p \in M$ (see definition 1.1.5 for the action of a linear map $A : V \rightarrow V$ on the tensor algebra of V).

Proof. We fix a point $p \in M$ and a symmetric normal neighborhood U of p such that $s_p(\alpha|_U) = (-1)^k \alpha|_U$ (assuming that α is a covariant k -tensor). We also fix a non-zero tangent vector $X_p \in T_p M$ and a basis $b = (b_k)_{k=1}^m$ of $T_p M$ with $b_1 = X_p$. This basis induces a normal coordinate system $(x^1, \dots, x^m)$ on U . We denote the components of α in that coordinate system by $\alpha_{i_1 \dots i_k}$. Given a smooth function ϕ on U , we say that ϕ is even if $\phi \circ s_p = \phi$. Then $s_p(\alpha|_U) = (-1)^k \alpha|_U$ is equivalent to the components of α being even, by lemma 3.3.2.

Let $\gamma : I \rightarrow U, t \rightarrow \exp(tX_p) = \exp(tb_1)$ be the geodesic with initial velocity X_p (restricted so that it takes value in U). Along γ we define the vector fields $X = \dot{\gamma}$ and

$$Z_i(t) = t\partial_i(\gamma(t)), t \in I$$

for $2 \leq i \leq m$. Since γ is a geodesic we have $\nabla_{\dot{\gamma}} X = 0$, and by definition of the normal coordinates, we have for all $t \in I$ and $2 \leq i \leq m$

$$Z_i(t) = \left. \frac{d}{ds} \right|_{s=0} \exp(t(b_1 + sb_i))$$

hence Z_i is an infinitesimal variation of γ and so satisfies the Jacobi equation (theorem 3.3.4):

$$\nabla_{\dot{\gamma}}^2 Z_i = R(X, Z_i)X. \quad (3.7)$$

Reasoning inductively from 3.7, we can express $(\nabla_{\dot{\gamma}}^r Z_i)(0)$ in terms of the curvature tensor:

1. For $r = 0$ we simply have $Z_i(0) = 0$ from the definition.
2. For $r = 1$, Leibniz' rule gives us

$$\nabla_{\dot{\gamma}} Z_i = \nabla_{\dot{\gamma}}(t(\partial_i \circ \gamma)) = (\partial_i \circ \gamma) + t\nabla_{\dot{\gamma}}(\partial_i \circ \gamma)$$

and we get

$$\nabla_{\dot{\gamma}} Z_i(0) = \partial_i(p)$$

3. After these two initial steps, we can proceed inductively using 3.7. Indeed applying $\nabla_{\dot{\gamma}}^r$ to the aforementioned equation (and recalling $\nabla_{\dot{\gamma}} X = 0$) gives us

$$\nabla_{\dot{\gamma}}^{r+2} Z_i = \sum_{q=0}^r \binom{r}{q} (\nabla_{\dot{\gamma}}^{r-q} R)(X, \nabla_{\dot{\gamma}}^q Z_i)X$$

which expresses $\nabla_{\dot{\gamma}}^{r+2} Z_i$ in terms of covariant derivatives of lower order (and covariant derivatives of R). In our case we only need $r = 2$ and $r = 4$.

For $r = 2$ we have

$$\nabla_{\dot{\gamma}}^2 Z_i(0) = R(X_p, \underbrace{Z_i(0)}_{=0})X_p = 0$$

and for $r = 4$ we have

$$\begin{aligned} \nabla_{\dot{\gamma}}^4 Z_i(0) &= (\nabla_{\dot{\gamma}}^2 R)(X_p, \underbrace{Z_i(0)}_{=0})X_p + 2(\nabla_{X_p} R)(X_p, \nabla_{\dot{\gamma}} Z_i(0))X_p + R(X_p, \underbrace{\nabla_{\dot{\gamma}}^2 Z_i(0)}_{=0})X_p \\ &= 2(\nabla_{X_p} R)(X_p, \partial_i(p))X_p = 2P_{X_p}(\partial_i(p)). \end{aligned}$$

Now let us choose $0 \leq l \leq k$, $2 \leq i_1, \dots, i_l \leq m$ and define the function

$$h : I \rightarrow \mathbb{R}, t \rightarrow \alpha(Z_{i_1}(t), \dots, Z_{i_l}(t), X(t), \dots, X(t)).$$

Then

$$h(t) = t^l \alpha_{i_1 \dots i_l 1 \dots 1}(\gamma(t))$$

and since $\alpha_{i_1 \dots i_l 1 \dots 1}$ is even, $h(-t) = (-1)^l h(t)$. In particular we must have

$$\left. \frac{d^{3+l}}{dt^{3+l}} \right|_{t=0} h(t) = 0. \quad (3.8)$$

To compute the derivatives of h we can once again use Leibniz' rule for the covariant derivative together with $\nabla\alpha = 0$ (since α is symmetric, corollary 3.3.2) and $\nabla_{\dot{\gamma}}X = 0$, which gives us

$$\left. \frac{d^{3+l}}{dt^{3+l}} \right|_{t=0} h(t) = \sum_{\beta \in \mathbb{N}^l, |\beta|=3+l} \binom{3+l}{\beta} \alpha(\nabla_{\dot{\gamma}}^{\beta_1} Z_{i_1}(0), \dots, \nabla_{\dot{\gamma}}^{\beta_l} Z_{i_l}(0), X_p, \dots, X_p). \quad (3.9)$$

We will see that many of the terms in the sum on the right-hand side vanish. Take a multi-index $\beta \in \mathbb{N}^l$ such that $|\beta| = 3 + l$, and define

$$\alpha_\beta := \alpha(\nabla_{\dot{\gamma}}^{\beta_1} Z_{i_1}(0), \dots, \nabla_{\dot{\gamma}}^{\beta_l} Z_{i_l}(0), X_p, \dots, X_p).$$

We consider the following cases:

1. There exists $1 \leq q \leq l$ such that $\beta_q = 0$ or $\beta_q = 2$. Then $\nabla_{\dot{\gamma}}^{\beta_q} Z_{i_q}(0) = 0$ so $\alpha_\beta = 0$. Thus we need only consider the cases for which $\beta_q \geq 1$, and $\beta_q \neq 2$ for all $1 \leq q \leq l$. Since $|\beta| = 3 + l$, this leaves only one kind of configuration, which we examine in the next point.
2. There exists a unique q such that $\beta_q = 4$, and for $r \neq q$ we have $\beta_r = 1$. In that case we have

$$\alpha_\beta = 2\alpha(\partial_{i_1}(p), \dots, \underbrace{P_{X_p} \partial_{i_q}(p)}_{q\text{th position}}, \dots, \partial_{i_l}(p), X_p, \dots, X_p).$$

Putting together 1., 2., and equations 3.9, 3.8 we obtain

$$\begin{aligned} 0 &= \left. \frac{d^{3+l}}{dt^{3+l}} \right|_{t=0} h(t) \\ &= 2 \frac{(3+l)!}{4!} \sum_{q=1}^l \alpha(\partial_{i_1}(p), \dots, \underbrace{P_{X_p} \partial_{i_q}(p)}_{q\text{th position}}, \dots, \partial_{i_l}(p), X_p, \dots, X_p) \end{aligned}$$

and since $0 = P_{X_p} X_p = (\nabla_{X_p} R)(X_p, X_p) X_p$ (as R is skew-symmetric in its first two

arguments) we have (see definition 1.1.5)

$$\begin{aligned}
& - (P_{X_p} \cdot \alpha_p)(\partial_{i_1}(p), \dots, \partial_{i_l}(p), \partial_1(p), \dots, \partial_1(p)) \\
& = \sum_{q=1}^l \alpha(\partial_{i_1}(p), \dots, \underbrace{P_{X_p} \partial_{i_q}(p)}_{q\text{th position}}, \dots, \partial_{i_l}(p), X_p, \dots, X_p) \\
& + \sum_{q=k+1}^l \alpha(\partial_{i_1}(p), \dots, \partial_{i_l}(p), X_p, \dots, \underbrace{P_{X_p} X_p}_{q\text{th position}}, \dots, X_p) \\
& = \sum_{q=1}^l \alpha(\partial_{i_1}(p), \dots, \underbrace{P_{X_p} \partial_{i_q}(p)}_{q\text{th position}}, \dots, \partial_{i_l}(p), X_p, \dots, X_p) \\
& = 0
\end{aligned}$$

so

$$0 = (P_{X_p} \cdot \alpha_p)_{i_1 \dots i_l 1 \dots 1}. \quad (3.10)$$

We are now essentially ready to conclude. The $(k-l)$ "1" indices in 3.10 appear last because we chose to place the $X(t)$ arguments at the end when defining $h(t)$ (this was a matter of simplifying notation), but clearly this choice is inconsequential and our reasoning applies equally well with $(k-l)$ $X(t)$ arguments placed in arbitrary positions. From this we deduce that $P_{X_p} \cdot \alpha_p = 0$, as all of its components vanish. This holds for any $p \in M$ and any $X_p \in T_p M$ (the case $X_p = 0$ is trivial), so the theorem follows. $\square$

When α is a volume form, the curvature condition admits a very simple expression.

Definition 3.3.10 (Cyclic parallel Ricci curvature). Let ∇ be an affine connection on a manifold M . The Ricci curvature is said to be **cyclic parallel** if

$$\mathfrak{S}_{X,Y,Z}(\nabla_X r)(Y, Z) = 0$$

for all $X, Y, Z \in \mathfrak{X}(M)$.

Proposition 3.3.5. *If the Ricci curvature is symmetric (i.e. if (M, ∇) is locally equiaffine), then the Ricci curvature is cyclic parallel if and only if*

$$(\nabla_X r)(X, X) = 0$$

for all $X \in \mathfrak{X}(M)$.

Proof. Assume that the Ricci curvature tensor is symmetric, so that ∇r is symmetric in its last two arguments. From theorem 1.2.2 (in section 1.2, dedicated to polarization) we obtain that

$$(\nabla_X r)(X, X) = 0, \text{ for all } X \in \mathfrak{X}(M)$$

if and only if

$$\text{Sym}(\nabla r) = 0.$$

Then, using the fact that ∇r is symmetric in its last two arguments we see that for all $X, Y, Z \in \mathfrak{X}(M)$

$$3\text{Sym}(\nabla r)(X, Y, Z) = \mathfrak{S}_{X,Y,Z}(\nabla_X r)(Y, Z).$$

□

Corollary 3.3.3. *Let ∇ be a torsionless connection on M , and let ω be a volume form on M . If ω is symmetric then ∇ is equiaffine and has cyclic parallel Ricci curvature.*

Proof. Applying theorem 3.3.6 to ω and recalling example 1.1.4 (ii) we find that for all $X \in \mathfrak{X}(M)$,

$$0 = P_X \cdot \omega = -\text{tr}(P_X)\omega$$

which implies

$$0 = \text{tr}(P_X) = \text{tr}(\nabla_X R)(X, -)X = (\nabla_X r)(X, X).$$

Lastly, from corollary 3.3.2 ω is parallel so (M, ∇) is equiaffine and the statement follows from proposition 3.3.5. □

Chapter 4

Riemannian, Symplectic and Kähler structures

In this section we introduce the three geometries that will intervene in this thesis (besides the geometry of affine connections): Riemannian metrics, complex structures and symplectic forms. We highlight some similarities and points of divergence in their behaviour, present some of the relationships that exist between these geometries, in particular how they come together to form Kähler geometry. We also examine the interactions between these structures and affine connections. A point of key importance will be the description of the symmetries exhibited by the curvature of a Kähler manifold.

The last section is dedicated to locally symmetric spaces. For each geometry encountered previously we provide the corresponding notion of locally symmetric space, and we give a common characterization of these locally symmetric spaces in terms of their curvature.

4.1 Riemannian Connections

The main reference for this section is [\[18\]](#).

Among the geometric structures we will introduce in this chapter, the study of Riemannian metrics has been the most extensive by a wide margin, and spans the longest period of time. This is hardly surprising, as Riemannian metrics are also the most intuitive in many respects.

Simply put, Riemannian metrics are the structures that give rise to well-behaved notions of distance and angle on a manifold. As such, any manifold immersed in a euclidean space $\mathbb{R}^n$ (which encompasses most easily visualized manifolds) inherits a Riemannian metric from this ambient space.

As a first instance of a recurring phenomenon, linear structures serve as our infinitesimal models. It is widely known that the correct notions of distance and angle on a real vector

space are obtained by introducing a **real inner product**, which yields the notion of a (finite-dimensional) **euclidean space**.

Definition 4.1.1 (Euclidean vector spaces). Let V be a finite-dimensional real vector space. A **(real) inner product** on V is a non-degenerate, symmetric and positive-definite bilinear map $g : V \times V \rightarrow \mathbb{R}$. A **euclidean vector space** is a finite-dimensional real vector space equipped with an inner product.

Moving from vector spaces to smooth manifolds, a Riemannian metric will be a 'smooth choice' of inner product on each tangent space. A process which is made precise through the use of tensor fields.

Definition 4.1.2 (Riemannian metrics). A **Riemannian metric**, or just **metric**, on a smooth manifold M is a covariant 2-tensor g such that for all $p \in M$, g_p is an inner product on $T_p M$.

Given two Riemannian manifolds (M, g) , (M, g') , an **isometry** $(M, g) \rightarrow (M', g')$ is a diffeomorphism $f : M \rightarrow M'$ such that $f^* g' = g$.

We start with an easy result, which hints at the abundance of Riemannian structures.

Proposition 4.1.1. *Let M be a smooth manifold. Then M admits a Riemannian metric.*

Remark. This will be contrasted with the comparatively scarce symplectic and complex structures.

The proposition is essentially a consequence of the fact that the set of real inner products on a vector space V forms a convex subset of $S^2 V^*$. This is the content of the following lemma, the proof of which we leave as an exercise.

Lemma 4.1.1. *Let V be a finite-dimensional real vector space, let $(g_i)_{i=1}^k$ ($k \geq 1$) be a family of inner products on V and let $(\lambda_i)_{i=1}^k$ be a family of positive real numbers such that*

$$\sum_{i=1}^k \lambda_i = 1.$$

Then $\sum_{i=1}^k \lambda_i g_i$ is a real inner product on V .

The convexity of inner products then allows us to patch together local metrics using a partition of unity.

Proof of proposition 4.1.1. Let $(U_\alpha, \phi_\alpha)_{\alpha \in S}$ be an atlas on M so that for each $\alpha \in S$,

$$\phi_\alpha : U_\alpha \subseteq M \rightarrow \mathbb{R}^m$$

is a coordinate chart. Denote the standard real inner product on $\mathbb{R}^m$ by g_0 . Then for any $\alpha \in S$ the pullback

$$g_\alpha := \phi_\alpha^* g_0$$

is a riemannian metric on U_α . Let $(\rho_\alpha)_{\alpha \in S}$ be a partition of unity subordinate to $(U_\alpha)_{\alpha \in S}$. Then since $\sum_{\alpha \in S} \rho_\alpha = 1$, lemma 4.1.1 implies that

$$g := \sum_{\alpha \in S} \rho_\alpha g_\alpha$$

is a riemannian metric on M . □

The musical isomorphism between the a euclidean vector space and its dual space becomes an isomorphism between the tangent and cotangent bundle of a Riemannian manifold.

Proposition 4.1.2. *For a Riemannian manifold (M, g) the map*

$$\flat : \mathfrak{X}(M) \rightarrow \Omega^1(M), X \rightarrow g(X, -)$$

induces a bundle isomorphism $TM \simeq T^*M$.

Proof. $\flat$ is a $C^\infty(M)$ -linear map, so it corresponds to a vector bundle morphism $\flat : TM \rightarrow T^*M$ (proposition 2.1.9). Furthermore, since g is non-degenerate, the morphism $\flat : TM \rightarrow T^*M$ acts isomorphically on each fiber, so that it must in fact be a bundle isomorphism (proposition 2.1.1). □

Most importantly for us, a Riemannian manifold comes equipped with a canonical torsionless connection, called the **Levi-Civita** connection.

Theorem 4.1.1 (Levi-Civita connection). *Let (M, g) be a Riemannian manifold. There exists a unique torsionless connection ∇ on M such that $\nabla g = 0$. ∇ is called the **Levi-Civita connection** of (M, g) .*

Proof. See [18] (theorem 5.10, p.122). □

We put a little emphasis on the importance of this result. The moral here is that a Riemannian manifold possesses a natural connection, and that this connection is fully determined by the metric. In particular an isometry between two Riemannian manifolds will necessarily be an affine transformation in the sense of definition 3.1.1.

Corollary 4.1.1. *Let (M, g) and (M', g') be two Riemannian manifolds, and let $f : M \rightarrow M'$ be an isometry. Then f preserves the Levi-Civita connections of (M, g) and (M', g') .*

Proof. Let ∇^f be the pullback of the Levi-Civita connection ∇' on M' along the bundle isomorphism $df : TM \rightarrow TM'$. One need only notice that ∇^f is torsionless, and that

$$\nabla^f g = \nabla^f(f^*g') = f^*(\nabla'g') = 0.$$

We can then conclude by theorem 4.1.1. □

4.1.1 Curvature of the Levi-Civita connection

Here we review a number of algebraic identities that are satisfied by the curvature of a Riemannian manifold, and introduce the notion of **scalar curvature**. Some results presented here are included solely for the purpose of providing context, some will be useful in subsequent chapters.

We begin by proving that the Ricci curvature is symmetric by exhibiting a (local) parallel Riemannian volume form. As usual, we start with vector spaces.

Proposition 4.1.3. *Let V be euclidean vector space (meaning that V is equipped with a real inner product) of dimension $m < \infty$ which is also oriented. Then there exists a unique top form $\mu \in \Lambda^m V^*$ such that*

$$\mu(e_1, \dots, e_m) = 1$$

*for any positively-oriented orthonormal basis $(e_1, \dots, e_m)$ of V . We call μ the **euclidean volume form** of V .*

Applying this process pointwise, we see that an oriented Riemannian manifold admits a natural volume form. This Riemannian volume form is very naturally parallel for the Levi-Civita connection as it is completely determined by the metric, which is preserved by said connection (we let the readers fill in the details).

Definition 4.1.3 (Riemannian volume form). Let (M, g) be an *oriented* Riemannian manifold of dimension m (so that each tangent space has an orientation). For each $p \in M$, let μ_p be the euclidean volume form of $(T_p M, g_p)$, then

$$M \rightarrow \Lambda^m T^* M, p \rightarrow \mu_p$$

defines a smooth volume form μ on M called the **Riemannian volume form** of (M, g) . Furthermore, since g is parallel for the Levi-Civita connection, so is μ .

Remark. A proof that the Riemannian volume form is parallel is easily obtained once one observes that for any path $\gamma : [0, 1] \rightarrow M$ joining the points p and q , the parallel translation

$$P(\gamma)_0^1 : (E_p, g_p) \rightarrow (E_q, g_q)$$

is isometric, meaning that $(P(\gamma)_0^1)^* g_q = g_p$.

Since orientations can always be chosen locally, a Riemannian manifold, even when it does not admit a global orientation, is always *locally* equiaffine.

Corollary 4.1.2. *Any Riemannian manifold is locally equiaffine (an oriented Riemannian manifold is globally equiaffine). In particular the Ricci curvature of a Riemannian manifold is always symmetric.*

Next we examine the algebraic identities satisfied by the curvature of a Riemannian connection. Making use of the musical isomorphism $TM \simeq T^*M$ we turn the curvature R of the Levi-Civita into a fully covariant 4-tensor.

Definition 4.1.4 (Riemannian covariant curvature). Let (M, g) be a Riemannian manifold and let R be the curvature of its Levi-Civita connection. The **covariant curvature** of (M, g) , denoted by $\hat{R}$ is the $(0, 4)$ -tensor defined by

$$\hat{R}(X, Y, Z, W) = g(R(X, Y)Z, W)$$

for all $X, Y, Z, W \in \mathfrak{X}(M)$.

Equipped with the covariant curvature, we can express the announced identities in their most standard form. These identities, which transfer to the Kähler setting, will be of crucial importance in our proof of the Sekigawa-Vanhecke theorem.

Theorem 4.1.2. *Let $\hat{R}$ be the covariant curvature of a Riemannian manifold. Then for all $X, Y, Z, W, V \in \mathfrak{X}(M)$ we have*

- (i) $\hat{R}(X, Y, Z, W) = -\hat{R}(Y, X, Z, W)$.
- (ii) $\hat{R}(X, Y, Z, W) = -\hat{R}(X, Y, W, Z)$.
- (iii) $\hat{R}(X, Y, Z, W) = \hat{R}(Z, W, X, Y)$.
- (iv) $\mathfrak{S}_{X, Y, Z} \hat{R}(X, Y, Z, W) = 0$.
- (v) $\mathfrak{S}_{X, Y, Z} (\nabla_X \hat{R})(Y, Z, W, V) = 0$

Proof. (i), (iv) and (v) are restatements of the known properties of the curvature of a torsionless affine connection. (i) is simply the skew-symmetry of R in its first two arguments, (iv) is the first Bianchi identity, and (v) is obtained by combining the second Bianchi identity (theorem 3.1.2) with $\nabla g = 0$.

Applying corollary 2.2.4 to g yields (ii). Indeed, since $\nabla g = 0$ we get

$$\begin{aligned} 0 &= -(R(X, Y) \cdot g)(Z, W) = g(R(X, Y)Z, W) + g(Z, R(X, Y)W) \\ &= \hat{R}(X, Y, Z, W) + \hat{R}(X, Y, W, Z) \end{aligned}$$

where we use the symmetry of g in the last equality.

Lastly (iii) is obtained from (i), (ii), (iv) through pure algebraic manipulations (in fact any $(0, 4)$ -tensor verifying symmetries (i), (ii) and (iv) will satisfy (iii)). See for example [12] (vol. I, chapter V, proposition 1.1, p.198), or [18] (proof of proposition 7.12 p.204). $\square$

As a final point in this section: using the musical isomorphism once again, we can turn the Ricci curvature into an endomorphism field (called the **Ricci endomorphism**, unsurprisingly), then define the **scalar curvature** through the trace.

Definition 4.1.5 (Ricci endomorphism and scalar curvature). Let (M, g) be a Riemannian manifold. The (Riemannian) **Ricci endomorphism** $\hat{\rho}$ of (M, g) is the tensor field of type $(1, 1)$ defined by

$$r(X, Y) = g(X, \hat{\rho}Y)$$

for all $X, Y \in \mathfrak{X}(M)$. The **scalar curvature** S of (M, g) is then defined to be the trace of the Ricci endomorphism

$$S := \text{tr}(\hat{\rho}).$$

Remark. All applications of the musical isomorphism can be expressed using contraction. So here we have

$$\begin{aligned} \hat{R} &= C_2^1(g \otimes R) \\ r &= C_2^1(g \otimes \hat{\rho}). \end{aligned}$$

Remark. The curvature of the Levi-Civita provides a local invariant for Riemannian manifolds, i.e. a tool we can use to distinguish between Riemannian manifolds at the local scale. For example, the well-known fact that the earth cannot be mapped without distorting distances (even locally) can be proven by showing that the scalar curvature of the sphere is non-zero (implying that the sphere cannot be isometric, even locally, to the flat plane). This will be a major point of divergence between Riemannian and symplectic geometry.

4.2 Symplectic Connections

Symplectic forms are non-degenerate 2-forms that verify an integrability condition (namely, they are required to be closed). The contrast between symplectic geometry and riemannian geometry, which we try to illustrate as much as possible, might be one of most catastrophic examples of the effects that changing a '+' to a '-' can have.

The study of symplectic manifolds was originally motivated by Hamiltonian mechanics, but symplectic geometry (as well as its sisters, contact and Poisson geometry) has long since exceeded these bounds and is now one of the richest and most diverse fields of mathematics. Some subjects one might encounter while exploring symplectic geometry include: the theory of Hamiltonian actions, symplectic reduction and momentum maps; symplectic capacities; pseudoholomorphic curves; Kirillov's orbit method; Floer theory; Fukaya categories and mirror symmetry, to name only a few.

An introduction to these subject would be beyond the scope of this thesis however (and of the author's abilities), in actuality, as it pertains to symplectic geometry itself, we will need very little beyond the most elementary notions.

In the following pages we will first develop the theory of symplectic vector spaces, which serve as infinitesimal models for symplectic geometry, but also as *local* models (by which we mean, at the level of germs, or small neighborhoods) as shown by Darboux's theorem. We will then introduce symplectic manifolds give a few basic definitions and state Darboux's theorem to build intuition. The section most relevant to our task will certainly be the last one, where we introduce symplectic connections. We describe the structure of the set of symplectic connections on a symplectic manifold (which will turn out to be an ∞ -dimensional affine space), present a few properties exhibited by a symplectic curvature and define certain special classes of symplectic connections. We close the section with the proof that type S connections are preferred.

[22], [7] and [21] are perfectly adequate as introductions to the subject of symplectic geometry, and will be our main references in the first two subsections. In the last subsection we will mostly rely on [5].

4.2.1 Symplectic vector spaces

Definition 4.2.1 (Symplectic vector spaces). Let V be a finite-dimensional vector space. A **symplectic form** on V is a non-degenerate skew-symmetric bilinear form $\omega : V \times V \rightarrow \mathbb{R}$. A pair (V, ω) consisting of a finite-dimensional real vector space V together with a symplectic form ω on V is called a **symplectic vector space**.

Example 4.2.1. The quintessential examples of symplectic vector spaces are:

- (i) The vector space $\mathbb{R}^{2n}$, whose coordinates we denote by $(x^1, y^1, x^2, y^2, \dots, x^n, y^n)$, equipped with the symplectic form

$$\omega_0 := \sum_{i=1}^n dx^i \wedge dy^i.$$

- (ii) The direct sum $E \oplus E^*$ of a finite-dimensional real vector space E with its dual, equipped with the symplectic form ω defined by

$$\omega((v, v^*), (w, w^*)) = w^*(v) - v^*(w)$$

for all $v, w \in E, v^*, w^* \in E^*$.

Definition 4.2.2 (Isomorphism of symplectic vector spaces). Let $(V, \omega), (V', \omega')$ be two symplectic vector spaces. An isomorphism from (V, ω) to (V', ω') is a linear isomorphism

$$L : V \rightarrow V'$$

such that $L^*\omega' = \omega$.

The following theorem elucidates the structure of symplectic vector spaces. Major differences with the Riemannian setting already become apparent, with the necessity for symplectic vector spaces to be even-dimensional, and we see that symplectic vector spaces are classified by dimension (i.e. $(\mathbb{R}^{2n}, \omega_0)$ is essentially the only symplectic vector space of dimension $2n$), which prepares the terrain for Darboux's theorem (thm. 4.2.2).

Theorem 4.2.1. *Let (V, ω) be a symplectic vector space. There exists a basis $e_1, \dots, e_n, f_1, \dots, f_n$ of V such that*

- (i) $\omega(e_i, e_j) = 0$.
- (ii) $\omega(f_i, f_j) = 0$.
- (iii) $\omega(e_i, f_j) = \delta_{ij}$.

We call such a basis a symplectic basis. In particular V must be even-dimensional, and (V, ω) is symplectomorphic to $(\mathbb{R}^{2n}, \omega_0)$.

Proof. See [21] (theorem 2.1.3, p.39) or [7] (theorem 1.1, p.3) for a slightly more general statement. $\square$

The following corollary, though simple has a number of uses. It shows that a symplectic vector space comes equipped with a canonical volume form (which induces a canonical orientation) all the while providing a useful criterion for determining whether a skew-symmetric 2-tensor is symplectic.

Corollary 4.2.1 (Symplectic volume form). *Let V be a real vector space of dimension $2n$ and let ω be a skew-symmetric bilinear form on V . Then ω is non-degenerate if and only if*

$$\omega^n := \underbrace{\omega \wedge \dots \wedge \omega}_{n \text{ times}}$$

*is non-zero. In the event that ω is symplectic, we call ω^n the **symplectic volume form** of (V, ω) .*

Proof. If ω is degenerate, then there exists a non-zero vector $v \in V$ such that $\iota_v \omega = 0$ (where ι_v denotes interior multiplication). Using the fact that ι_v is a graded derivation of degree (-1) on $\Lambda(V)$ we get

$$\iota_v(\omega^n) = 0$$

which shows that $\omega^n = 0$ (for this last step one can complete $\{v\}$ into a basis of V).

Conversely, if ω is non-degenerate, then theorem 4.2.1 guarantees the existence of a symplectic basis $(e_1, f_1, e_2, f_2, \dots, e_n, f_n)$, and a straightforward computation (using the definition of the wedge product) yields

$$(\omega^n)(e_1, f_1, \dots, e_n, f_n) > 0.$$

□

We end the linear section by introducing the group of automorphisms of a symplectic vector space and its Lie algebra.

Definition 4.2.3 (Symplectic automorphisms). Let (V, ω) be a symplectic vector space. We define $Sp(V, \omega)$ to be the group of automorphisms of (V, ω) , that is

$$Sp(V, \omega) := \{L \in GL(V) | L^* \omega = \omega\}.$$

$Sp(V, \omega)$ is a closed Lie subgroup of $GL(V)$ and its Lie algebra $\mathfrak{sp}(V, \omega)$ is given by

$$\mathfrak{sp}(V, \omega) = \{A \in \text{End}(V) | 0 = \omega(A-, -) + \omega(-, A-)\}.$$

As special cases we define $Sp(2n) := Sp(\mathbb{R}^{2n}, \omega_0)$ and $\mathfrak{sp}(2n) := \mathfrak{sp}(\mathbb{R}^{2n}, \omega_0)$.

Any element of $Sp(V, \omega)$ preserves the volume form ω^n and so must have determinant 1 (in other words, $Sp(V, \omega)$ is a subgroup of the special linear group). Consequently, elements of $\mathfrak{sp}(V, \omega)$ have trace 0.

4.2.2 Elements of symplectic geometry

In analogy with the definition of Riemannian manifolds, a symplectic form on a manifold will be a 'smooth choice' of a symplectic structure on each tangent space, but in contrast to it, the definition of a symplectic manifold also includes an **integrability condition** (which is 'not pointwise') : the resulting 2-form must be *closed*.

Definition 4.2.4 (Symplectic manifolds). A **symplectic form** on a manifold M is a closed non-degenerate 2-form. A **symplectic manifold** is a smooth manifold M together with a symplectic form ω . Given two symplectic manifolds (M, ω) and (M', ω') , a **symplectomorphism** $(M, \omega) \rightarrow (M', \omega')$ is a diffeomorphism $\phi : M \rightarrow M'$ such that $\phi^*\omega' = \omega$.

Example 4.2.2. Expanding on our two examples of symplectic vector spaces.

- (i) While flat euclidean spaces serve as 'pointwise' or 'infinitesimal' models for Riemannian manifolds and not much more, the flat symplectic space $(\mathbb{R}^{2n}, \omega_0)$ is a genuine *local* model of $2n$ -dimensional symplectic manifolds as per Darboux's theorem (thm. 4.2.2), and thus plays a much bigger role in symplectic geometry.
- (ii) Extending the example of the direct sum $E \oplus E^*$ of a vector space with its dual, the cotangent bundle T^*M of any smooth manifold M comes equipped with a tautological 1-form λ , defined by

$$\lambda_\alpha(X_\alpha) := (\alpha \circ d\pi_\alpha)(X_\alpha)$$

for all $\alpha \in T^*M$, $X_\alpha \in T_\alpha T^*M$, whose exterior derivative $d\lambda$ can be shown to be a symplectic form on T^*M (only non-degeneracy remains to be shown, which can be done in the coordinates induced by a chart on M). This example is of particular historical importance, indeed Hamiltonian mechanics models a system by fixing a configuration space (the 'space of positions') in the form of a smooth manifold M , only to introduce the *phase space* (the 'space of positions and momenta') in the form of the cotangent space T^*M , which is where the study of the system's mechanical behaviour begins in earnest. The important role played by the canonical symplectic structure of T^*M in Hamiltonian mechanics (through the Poisson brackets most notably) was the original motivation behind the first forays into what we now call symplectic geometry.

We mention those properties of symplectic vector spaces that transfer seamlessly to manifolds.

Corollary 4.2.2. *Let (M, ω) be a symplectic manifold. Then M has dimension $2n$ for some natural n (M is even-dimensional) and ω^n is a volume form on M . In particular M is orientable.*

The reader can compare with proposition 4.1.1 and convince themselves that Riemannian and symplectic geometry differ at their very core. If that is not sufficient, more is to come.

Despite these differences, the non-degeneracy of the symplectic form means that we still enjoy the benefit of a musical isomorphism.

Proposition 4.2.1. *Let (M, ω) be a symplectic manifold. Then*

$$\flat : \mathfrak{X}(M) \rightarrow \Omega^1(M), X \mapsto \iota_X \omega = \omega(X, -)$$

induces a bundle isomorphism $TM \simeq T^*M$.

This musical isomorphism allows for the definition of the **symplectic gradient** of a function.

Definition 4.2.5 (Symplectic gradients). Let $f \in C^\infty(M)$, the **symplectic gradient** X_f of f is the unique smooth vector field on M such that $\iota_{X_f} \omega = df$.

Remark. Symplectic and Riemannian gradients are defined in perfectly analogous manners, however their behaviour have little in common. For one thing, f is a constant of motion for the flow induced by X_f , indeed

$$X_f f = df(X_f) = \omega(X_f, X_f) = 0.$$

We now state the much anticipated theorem of Darboux.

Theorem 4.2.2 (Darboux's Theorem). *Let (M, ω) be a symplectic manifold of dimension $2n$. Then for any point $p \in M$ there exists an open neighborhood U of p such that $(U, \omega|_U)$ is symplectomorphic to an open subset of $(\mathbb{R}^{2n}, \omega_0)$.*

Proof. See [21] (theorem 3.2.2, p.110). □

Remark. And so we see that two symplectic manifolds of equal dimension can never be distinguished at the local scale, as great a contrast with Riemannian geometry and its curvature invariants as one could ask for. Once again cartography is a good source of examples: though isometric map projections do not exist, *symplectic* projections, i.e. map projections that preserve *area* (also called *authalic* projections) are plenty.

4.2.3 The space of symplectic connections

We introduce symplectic connections and give an overview of their basic properties.

Definition 4.2.6 (Symplectic connections). Let (M, ω) be a symplectic manifold. A **symplectic connection** on (M, ω) is a torsionless affine connection ∇ on M such that $\nabla\omega = 0$.

Symplectic connections happen to come in great numbers, indeed they form an ∞ -dimensional affine space.

Theorem 4.2.3. *Let (M, ω) be a symplectic manifold.*

- (i) *(M, ω) admits a symplectic connection.*
- (ii) *Let ∇ be a symplectic connection on (M, ω) , and let ∇' be another affine connection on M . Then ∇' is symplectic if and only if the tensor W defined by*

$$W(X, Y, Z) = \omega(\nabla'_X Y - \nabla_X Y, Z) \text{ for all } X, Y, Z \in \mathfrak{X}(M)$$

is symmetric.

In particular, the set of symplectic connections on (M, ω) is an affine space modelled on S^3T^*M (the space of symmetric covariant 3-tensors).

Before giving the proof of theorem 4.2.3, we present a lemma describing the structure of the set of torsionless affine connections on a manifold. We've seen that the set of affine connections on a manifold is an affine space modeled on $\Gamma T_2^1 M$ (example 2.2.2, section 2.2), and the set of *torsionless* affine connections on M will form an affine subspace modeled on $\Gamma(TM \otimes S^2 T^*M)$ (the subspace of $\Gamma T_2^1 M$ consisting of (1,2)-tensors that are symmetric in their covariant indices).

Lemma 4.2.1. *Let ∇ be a torsionless affine connection on M . Let S be a map*

$$S : \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M).$$

Then $\nabla'_X Y := \nabla_X Y + S(X, Y)$ defines a torsionless affine connection ∇' if and only if S is a section of $TM \otimes S^2 T^*M$, or equivalently if and only if S is a (1,2)-tensor such that

$$S(X, Y) = S(Y, X), \text{ for all } X, Y \in \mathfrak{X}(M).$$

Remark. Since a manifold always admits a torsionless affine connection (we know that a manifold always admits an affine connection - specialize theorem 2.2.1 to $E = T^*M$ - and we can always extract a torsionless connection from an arbitrary affine connection using formula 3.1), we indeed see that the set of torsionless affine connections is an affine space modeled on $\Gamma(TM \otimes S^2 T^*M)$.

Proof of lemma 4.2.1. We know from theorem 2.2.1 that ∇' is an affine connection if and only if S is a (1,2)-tensor. The torsion T of ∇' is then given by

$$\begin{aligned} T(X, Y) &= \nabla'_X Y - \nabla'_Y X - [X, Y] \\ &= \underbrace{\nabla_X Y - \nabla_Y X - [X, Y]}_{=0} + S(X, Y) - S(Y, X) \\ &= S(X, Y) - S(Y, X) \end{aligned}$$

and the lemma follows. $\square$

Proof of theorem 4.2.3. (i) we choose a torsionless connection ∇^0 on M (for example the Levi-Civita connection of some metric on M). We define the (1,2)-tensor N by

$$(\nabla_X^0 \omega)(Y, Z) =: \omega(N(X, Y), Z).$$

Since ω is skew-symmetric we have

$$\omega(N(X, Y), Z) = (\nabla_X^0 \omega)(Y, Z) = -(\nabla_X^0 \omega)(Z, Y) = -\omega(N(X, Z), Y) \quad (4.1)$$

and since ω is closed, theorem 3.1.1 yields

$$\begin{aligned} 0 &= d\omega(X, Y, Z) = 3 \text{Alt}(\nabla^0 \omega)(X, Y, Z) \\ &= \mathfrak{S}_{X,Y,Z}(\nabla_X^0 \omega)(Y, Z) \\ &= \mathfrak{S}_{X,Y,Z} \omega(N(X, Y), Z). \end{aligned} \quad (4.2)$$

We define the affine connection ∇ by

$$\nabla_X Y := \nabla_X^0 Y + \frac{1}{3} (N(X, Y) + N(Y, X)).$$

Clearly $\frac{1}{3} (N(X, Y) + N(Y, X))$ is symmetric in X and Y , so by lemma 4.2.1, ∇ is torsionless. Furthermore, we have

$$\begin{aligned} (\nabla_X \omega)(Y, Z) &= X(\omega(Y, Z)) - \omega(\nabla_X Y, Z) - \omega(Y, \nabla_X Z) \\ &= (\nabla_X^0 \omega)(Y, Z) - \frac{1}{3} \omega(N(X, Y), Z) - \frac{1}{3} \omega(N(Y, X), Z) \\ &\quad - \frac{1}{3} \omega(Y, N(X, Z)) - \frac{1}{3} \omega(Y, N(Z, X)) \\ &= (\nabla_X^0 \omega)(Y, Z) - \frac{1}{3} \omega(N(X, Y), Z) - \frac{1}{3} \omega(N(Y, X), Z) \\ &\quad - \frac{1}{3} \omega(N(X, Y), Z) - \frac{1}{3} \omega(N(Z, Y), X) \quad (\text{eq. 4.1}) \\ &= (1 - \frac{1}{3} - \frac{1}{3}) \omega(N(X, Y), Z) + \frac{1}{3} \omega(N(X, Z), Y) \quad (\text{eq. 4.2}) \\ &= 0 \quad (\text{eq. 4.1}) \end{aligned}$$

so ∇ is symplectic. This takes care of (i).

We move on to (ii). By lemma 4.2.1, ∇' is a torsionless connection if and only if it can be written $\nabla'_X Y = \nabla_X Y + S(X, Y)$ with S such that $S(X, Y) = S(Y, X)$, and in that case we find

$$\begin{aligned} (\nabla'_X \omega)(Y, Z) &= \underbrace{(\nabla_X \omega)(Y, Z)}_{=0} - \omega(S(X, Y), Z) - \omega(Y, S(X, Z)) \\ &= -\omega(S(X, Y), Z) + \omega(S(X, Z), Y) \end{aligned}$$

so ∇' is symplectic if and only if

$$W(X, Y, Z) := \omega(S(X, Y), Z) = \omega(\nabla'_X Y - \nabla_X Y, Z)$$

is symmetric. □

This is another point of divergence with Riemannian geometry and its unique 'metric connection', but one that is not too surprising considering Darboux's theorem. Indeed, a unique symplectic connection, if it were to exist, would run the risk of introducing a forbidden local invariant in the form of its curvature.

The (symplectic) musical isomorphism can be used to define a (symplectic) covariant curvature tensor, which comes with symmetries of its own.

Theorem 4.2.4. *Let ∇ be a symplectic connection on (M, ω) with curvature R . The covariant curvature of (M, ω, ∇) is the $(0, 4)$ -tensor defined by*

$$\bar{R}(X, Y, Z, W) = \omega(R(X, Y)Z, W) \text{ for all } X, Y, Z, W \in \mathfrak{X}(M).$$

Then for all vector fields X, Y, Z, W, V on M we have

- (i) $\bar{R}(X, Y, Z, W) = -\bar{R}(Y, X, Z, W)$.
- (ii) $\bar{R}(X, Y, Z, W) = \bar{R}(X, Y, W, Z)$.
- (iii) $\mathfrak{S}_{X, Y, Z} \bar{R}(X, Y, Z, W) = 0$.
- (iv) $\mathfrak{S}_{X, Y, Z} (\nabla_X \bar{R})(Y, Z, W, V) = 0$.

Proof. Just as in theorem 4.1.2, (i), (iii) and (iv) follow from the skew-symmetry of R and Bianchi's identities (as well as $\nabla \omega = 0$ for (iv)).

(ii) is the application of corollary 2.2.4 to ω which, combined with the skew-symmetry of ω , yields

$$\begin{aligned} 0 &= -(R(X, Y) \cdot \omega)(Z, W) = \omega(R(X, Y)Z, W) + \omega(Z, R(X, Y)W) \\ &= \bar{R}(X, Y, Z, W) - \bar{R}(X, Y, W, Z). \end{aligned}$$

□

The Ricci curvature of a symplectic connection remains symmetric given the existence of the (clearly parallel) symplectic volume form ω^n , and can still be made into an endomorphism through the (symplectic) musical isomorphism.

Definition 4.2.7 (Symplectic Ricci endomorphism). The **(symplectic) Ricci endomorphism** is the $(1, 1)$ -tensor field ρ defined by

$$r(X, Y) = \omega(X, \rho Y)$$

for all $X, Y \in \mathfrak{X}(M)$.

But since r is symmetric while ω is antisymmetric we have

$$\omega(\rho X, Y) + \omega(X, \rho Y) = -r(Y, X) + r(X, Y) = 0$$

which shows that ρ_p is an element of $\mathfrak{sp}(T_p M, \omega_p)$ for all $p \in M$. In particular $\text{tr} \rho = 0$ (see definition 4.2.3), so there is *no* useful notion of scalar curvature for a symplectic connection.

Given the plethora symplectic connections that a given symplectic manifold admits, one would naturally want to isolate certain symplectic connections with special features. Much labor has already been invested in this enterprise, the fruits of which are neatly summarized in [5]. Here we only mention those special symplectic connections that are of import to our task.

Definition 4.2.8. Let ∇ be a symplectic connection on (M, ω) .

- (i) ∇ is said to be preferred if it has cyclic parallel Ricci curvature.
- (ii) We say that ∇ is of type S if ω is symmetric for ∇ .

And corollary 3.3.3 applied to the symplectic volume form ω^n immediately yields the following implication.

Corollary 4.2.3. *A symplectic connection of type S is necessarily preferred.*

4.3 (Almost) Complex, Hermitian and Kähler Structures

The governing example that the reader should keep in mind for the entirety of this section is that of the standard complex vector space $\mathbb{C}^n$. $\mathbb{C}^n$ comes with its standard sesquilinear Hermitian product

$$h_0(z, z') := \sum_{k=1}^n z_k \bar{z}'_k, \text{ for } z, z' \in \mathbb{C}^n.$$

If we identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ via $z_k = x_k + iy_k$, then h_0 can be written

$$\begin{aligned} h_0(z, z') &= \sum_{k=1}^n (x_k + iy_k)(x'_k - iy'_k) \\ &= \sum_{k=1}^n (x_k x'_k + y_k y'_k) + i \sum_{k=1}^n (y_k x'_k - x_k y'_k) \\ &= g_0((x, y), (x', y')) + i\omega_0((x, y), (x', y')) \end{aligned}$$

so we see that the standard inner product and symplectic structure of $\mathbb{R}^{2n}$

$$\begin{aligned} g_0((x, y), (x', y')) &:= \sum_{k=1}^n (x_k x'_k + y_k y'_k) \\ \omega_0((x, y), (x', y')) &:= \sum_{k=1}^n (y_k x'_k - x_k y'_k) \end{aligned}$$

are nothing but the real and imaginary parts of the Hermitian product h_0 :

$$\begin{aligned} g_0 &= \operatorname{Re} h_0 \\ \omega_0 &= \operatorname{Im} h_0. \end{aligned}$$

Furthermore, multiplication by i can be represented on $\mathbb{R}^{2n}$ by the automorphism

$$J_0(x, y) := (-y, x)$$

and the sesquilinearity of h_0 then yields

$$\begin{aligned} g_0(J_0-, J_0-) &= \operatorname{Re} h_0(i-, i-) = \operatorname{Re}[i(-i)h_0] = \operatorname{Re} h_0 = g_0 \\ g_0(-, J_0-) &= \operatorname{Re} h_0(-, i-) = \operatorname{Re}[(-i)h_0] = \operatorname{Im} h_0 = \omega_0 \end{aligned}$$

and similarly

$$\begin{aligned} \omega_0(J_0-, J_0-) &= \omega_0 \\ \omega_0(J_0-, -) &= g_0. \end{aligned}$$

Lastly, since g_0 and ω_0 are non-degenerate, the equation

$$\omega_0 = g_0(-, J_0-)$$

implies that J_0 is fully determined by g_0 and ω_0 .

So to summarize, the space $\mathbb{C}^n \simeq \mathbb{R}^{2n}$ supports a number of standard structures:

- An inner product g_0 .
- A symplectic structure ω_0 .
- A complex structure J_0 (meaning that J_0 encodes 'multiplication by i ').
- A Hermitian inner product h_0 .

The first three structures (g_0, ω_0, J_0) are mutually compatible in several ways. Firstly, both g_0 and ω_0 are preserved by J_0 :

$$\begin{aligned} g_0(J_0-, J_0-) &= g_0 \\ \omega_0(J_0-, J_0-) &= \omega_0. \end{aligned}$$

Secondly the pair of equations

$$\begin{aligned} \omega_0 &= g_0(-, J_0-) \\ g_0 &= \omega_0(J_0-, -) \end{aligned}$$

(which are in fact redundant) shows that any element of the triple (g_0, ω_0, J_0) is determined by the other two.

Finally, the information contained in the triple (g_0, ω_0, J_0) is exactly the information contained in the structure of $\mathbb{C}^n$ as a complex vector space equipped with the Hermitian inner product h_0 , since J_0 corresponds to multiplication by i and

$$h_0 = g_0 + i\omega_0 \Leftrightarrow (g_0 = \operatorname{Re} h_0, \omega_0 = \operatorname{Im} h_0).$$

This model example should serve as a good introduction to the major theme of this section: the rich interplay between Riemannian, symplectic, complex and Hermitian structures (only affine connections are missing).

This entire discussion transfers to manifolds, giving rise to Kähler manifolds when it succeeds 'perfectly'. However, at the level of manifolds, matters are somewhat complicated by the presence of **integrability conditions** for complex and symplectic structures, leading to the notions of **almost complex** and **almost symplectic** structures (with the latter defined as non-degenerate 2-forms).

As has become routine by now, we begin with a survey of the relevant linear structures. Most proofs in this first section are simple exercises in linear algebra and will thus be omitted. As for references, we rely on [7], [21] and [13] (vol. II).

4.3.1 Linear structures

A complex vector space has a natural real vector space structure. Going the other way, given a real vector space V , a complex vector space structure on V will be said to be **compatible** if the real vector space structure it induces is the one we started with.

Given a real vector space V , if we wish equip V with a compatible complex structure, it is necessary and sufficient for us to specify multiplication by i . This is the motivation behind the notion of a **complex structure** on V .

Definition 4.3.1 (Complex structures). A **complex structure** on a real vector space V is a (real) linear automorphism $J : V \rightarrow V$ such that $J^2 = -I$.

Theorem 4.3.1. *Let V be a real vector space. There is a one-to-one correspondence between compatible complex vector space structures on V and complex structures on V (in the sense of definition 4.3.1). The automorphism J corresponds to multiplication by i .*

In view of this, we will use the denomination "complex vector space" to refer both to complex vector spaces in the usual sense, as well as pairs (V, J) consisting of a real vector space V and a complex structure J on V .

A complex vector space of (finite) complex dimension n will have real dimension $2n$.

Of course this means that complex linear maps can be characterized via the complex structure J .

Proposition 4.3.1. *Given two complex vector spaces V, V' with complex structures J and J' , a real linear map $A : V \rightarrow V'$ is complex linear if and only if*

$$A \circ J = J' \circ A.$$

We now introduce the group of complex automorphisms.

Definition 4.3.2 (Group of complex automorphisms). Given a complex vector space V with a complex structure J , we denote the **group of complex automorphisms** of (V, J) by $GL(V, J)$, that is

$$GL(V, J) := \{L \in GL(V) \mid J \circ L = L \circ J\}.$$

And take advantage of its properties to endow a complex vector space with a canonical orientation.

Proposition 4.3.2. *Let (V, J) be a complex vector space of dimension n . Then*

- (i) *There exists a complex linear isomorphism $(V, J) \simeq (\mathbb{C}^n, J_0)$ where J_0 denotes the standard complex structure of $\mathbb{C}^n$.*
- (ii) *$GL(V, J)$ is connected, in particular all of its element have (strictly) positive determinant. In other words, they are orientation-preserving.*

Proof. The connectedness of $GL_n(\mathbb{C})$ can be proven in a very neat way: Given $A \in GL_n(\mathbb{C})$ we would like to find a continuous path $\gamma : [0, 1] \rightarrow \mathbb{C}$ going from 0 to 1 such that

$$[0, 1] \rightarrow M_n(\mathbb{C}), t \rightarrow A + \gamma(t)(I - A)$$

takes value in $GL_n(\mathbb{C})$ (this would provide a path in $GL_n(\mathbb{C})$ joining A and I). This is equivalent to requiring that $P \circ \gamma$ not vanish on $[0, 1]$, where $P : \mathbb{C} \rightarrow \mathbb{C}$ is the polynomial defined by

$$P(z) := \det[A + z(I - A)].$$

Since both A and I are invertible, $P(0) \neq 0$ and $P(1) \neq 0$, so P is a non-zero polynomial (of finite degree) and must have finitely many root. Consequently there exists a continuous path $\gamma : [0, 1] \rightarrow \mathbb{C}$ that goes from 0 to 1 and avoids all the roots of P . $\square$

Corollary 4.3.1. *Let V be a finite-dimensional real vector space equipped with a complex structure J . Then V must be even-dimensional, furthermore V admits a unique orientation for which any complex isomorphism*

$$(V, J) \rightarrow (\mathbb{C}^n, J_0)$$

is orientation-preserving (with n the complex dimension of (V, J)).

Remark. The proximity between symplectic and complex structures becomes apparent already at the linear level.

Next come Hermitian structures, which serve as the complex analogues of euclidean structures, and allow for the introduction of metric notions on complex vector spaces.

Definition 4.3.3 (Hermitian spaces). A **Hermitian inner product** on a complex vector space V is a map

$$h : V \times V \rightarrow \mathbb{C}$$

such that

- (i) (Conjugate symmetry) For all $x, y \in V$, $h(x, y) = \overline{h(y, x)}$, where $\overline{(-)}$ denotes complex conjugation.

- (ii) (Sesquilinear) h is complex linear in its first argument, and complex antilinear in its second argument. Meaning that for all $\lambda \in \mathbb{C}$, $x, x', y, y' \in V$ we have

$$\begin{aligned} h(\lambda x + x', y) &= \lambda h(x, y) + h(x', y) \\ h(x, \lambda y + y') &= \bar{\lambda} h(x, y) + h(x, y') \end{aligned}$$

- (iii) (Positive-definite) For all non-zero $x \in V$, $h(x, x) > 0$.

A **Hermitian vector space** is a finite-dimensional complex vector space equipped with a Hermitian inner product.

As we saw at the beginning of this section, a Hermitian inner product on a complex vector space V is fully determined by its real part. This real part will always be a real inner product compatible with the complex structure J of V . Conversely, such a compatible real inner product will induce a Hermitian inner product so that the correspondence is exact.

Definition 4.3.4 (Compatible real inner product). Let V be a real vector space equipped with a complex structure J . A real inner product g on V is said to be **compatible** with J if J preserves g , meaning that for all $x, y \in V$

$$g(Jx, Jy) = g(x, y).$$

Theorem 4.3.2. *Let V be a real vector space, and let J be a complex structure on V . There is a one-to-one correspondence between real inner products compatible with J , and Hermitian inner products on V equipped with the complex structure induced by J .*

- (i) *Given a Hermitian inner product h on V , if we define g to be the real part of h , meaning that*

$$g(x, y) := \operatorname{Re} h(x, y)$$

for all $x, y \in V$, then g is a compatible real inner product.

- (ii) *Given a compatible real inner product g , if we define the map h by*

$$h(x, y) := g(x, y) + ig(x, Jy)$$

for all $x, y \in V$, then h is a Hermitian inner product.

And these two maps are inverse to each other. Thus in what follows, we will use the name "Hermitian vector space" to refer both to Hermitian vector spaces in the sense of definition 4.3.3 as well as triples (V, J, g) consisting of a real vector space V , a complex structure J on V and a real inner product g on V that is compatible with J .

Remark. Many authors use the name 'Hermitian inner product' to designate the compatible inner product g .

We introduce the last element of the compatible triple: the symplectic structure induced by a Hermitian inner product, which we call the **Kähler form** in anticipation of what is to come.

Proposition 4.3.3. *Let (V, J, g) be a Hermitian vector space with Hermitian inner product h . The real-valued map ω defined by*

$$\omega(x, y) := g(x, Jy) = \operatorname{Im} h(x, y)$$

*for all $x, y \in V$ is a symplectic form, called the **Kähler form** of (V, J, g) . Furthermore, the orientations induced by ω^n and J coincide.*

Proof. Choose an orthonormal basis for the induced Hermitian inner product h , which will induce an isomorphism of complex inner product spaces

$$(V, h) \rightarrow (\mathbb{C}^n, h_0).$$

As a complex isomorphism and an isomorphism of symplectic spaces, this map preserves both the symplectic and complex orientations. We finish the proof by checking that the complex and symplectic orientations coincide on $(\mathbb{C}^n, h_0)$. $\square$

We conclude this section with a statement formalizing the observation that any element in a triple (g, ω, J) associated to a Hermitian vector space is determined by the other two members of said triple, and that this triple determines the Hermitian vector space structure.

Theorem 4.3.3 (2-out-of-3 property). *Let (V, J, g) be a finite-dimensional complex inner product space with Hermitian inner product space h and Kähler form ω . Then*

$$\begin{aligned} U(V, J, h) &= GL(V, J) \cap Sp(V, \omega) \cap O(V, g) \\ &= GL(V, J) \cap Sp(V, \omega) = GL(V, J) \cap O(V, g) = Sp(V, \omega) \cap O(V, g). \end{aligned}$$

where

- $O(V, g)$ is the group of real automorphisms of V that preserve g

$$O(V, g) := \{L \in GL(V) \mid L^*g = g\}.$$

- $U(V, J, h)$ is the group of complex automorphisms of (V, J) that preserve h

$$U(V, J, h) := \{L \in GL(V, J) \mid L^*h = h\}.$$

In particular, if we choose $V = \mathbb{C}^n$ with its canonical complex, euclidean, symplectic and Hermitian structures we obtain

$$U(n) = GL_n(\mathbb{C}) \cap O(2n) = GL_n(\mathbb{C}) \cap Sp(2n) = O(2n) \cap Sp(2n) = GL_n(\mathbb{C}) \cap O(2n) \cap Sp(2n).$$

4.3.2 (Almost) complex structures

We begin with the notion of an **almost complex structure**, which is nothing but a smooth choice of complex structure for each tangent space.

Definition 4.3.5 (Almost complex structures). An **almost complex structure** on a manifold M is a bundle automorphism $J : TM \rightarrow TM$ such that $J^2 = -1$. An **almost complex manifold** is a manifold equipped with an almost complex structure.

Corollary 4.3.2. *An almost complex manifold is even-dimensional and admits a standard orientation. The orientation is obtained by endowing each tangent space with its standard orientation as a complex vector space.*

Proof. We have all the tools needed to prove this statement, but we would prefer to wait until the next section, where we'll provide a proof relying on the Kähler form. $\square$

As their name indicates, almost complex structures are closely related to, yet distinct from complex structures proper. Defining the latter will necessitate a small detour through holomorphic maps (on opens of $\mathbb{C}^n$). An important fact regarding these holomorphic maps is that they can be defined via the standard complex structure of $\mathbb{C}^n$.

Definition 4.3.6 (Holomorphic maps). Consider two open subsets $U \subseteq \mathbb{C}^m$, $V \subseteq \mathbb{C}^n$. A smooth map $f : U \rightarrow V$ is said to be **holomorphic** if for all $p \in U$ the differential

$$df_p : T_p U \simeq \mathbb{C}^m \rightarrow \mathbb{C}^n \simeq T_p V$$

of f at p is complex linear. Equivalently, f is holomorphic if and only if

$$J_0 \circ df = df \circ J_0$$

(where we use J_0 to denote the standard complex structures of $\mathbb{C}^m$ and $\mathbb{C}^n$).

Equipped with this notion, we give a definition of complex manifolds. This definition replicates word for word the definition of smooth manifolds by means of atlases, except that transition maps are required to be holomorphic.

Definition 4.3.7 (Complex manifolds). Let M be a smooth manifold of dimension $2n$. A **complex structure** on M is a maximal holomorphic atlas, that is an atlas whose charts take value in $\mathbb{C}^n$ and have holomorphic transition maps, and which is maximal for these properties. Equipped with such a complex structure, M is called a **complex manifold** of complex dimension n . The charts belonging to the maximal holomorphic atlas are called holomorphic charts (on the complex manifold M).

The corresponding notion of morphism is the following:

Definition 4.3.8 (Holomorphic maps between manifolds). Let M, N be complex manifolds, of complex dimension m and n respectively. A smooth map $f : M \rightarrow N$ is said to be **holomorphic** if for any holomorphic charts $\phi : U \subseteq M \rightarrow \tilde{U} \subseteq \mathbb{C}^m$ and $\psi : V \subseteq N \rightarrow \tilde{V} \subseteq \mathbb{C}^n$ such that $f(U) \subseteq V$, the map

$$\psi \circ f \circ \phi^{-1} : \tilde{U} \subseteq \mathbb{C}^m \rightarrow \tilde{V} \subseteq \mathbb{C}^n$$

is holomorphic. A **biholomorphism** between two complex manifolds is a holomorphic diffeomorphism whose inverse is also holomorphic.

A complex manifold comes equipped with its almost complex structure.

Theorem 4.3.4. *Let M be a complex manifold of complex dimension n . Then there exists a unique almost complex structure J on M such that for any holomorphic chart $\phi : U \subseteq M \rightarrow \mathbb{C}^n$, and any point $p \in U$ the differential*

$$d\phi_p : (T_p M, J_p) \rightarrow (\mathbb{C}^n, J_0)$$

of ϕ at p is complex linear.

Alternatively, J can be characterized as follows: given a holomorphic chart $(U, z^1 = x^1 + iy^1, \dots, z^n = x^n + iy^n)$ on M , the restriction of J to U is defined by

$$\begin{aligned} J\partial_{x^k} &= \partial_{y^k} \\ J\partial_{y^k} &= -\partial_{x^k}. \end{aligned}$$

Proof. Given two holomorphic charts ϕ and ψ defined on a neighborhood of p ,

$$d\psi_p \circ (d\phi_p)^{-1} : \mathbb{C}^n \rightarrow \mathbb{C}^n$$

is an isomorphism of complex vector spaces, thus the $\mathbb{R}$ -linear isomorphisms

$$\begin{aligned} d\phi_p : T_p M &\rightarrow \mathbb{C}^n \\ d\psi_p : T_p M &\rightarrow \mathbb{C}^n \end{aligned}$$

induce the same complex structure on $T_p M$. The coordinate expression of J is a consequence of the formula

$$J_0(x, y) = (-y, x) \text{ for all } (x, y) \in \mathbb{R}^{2n}$$

for the standard complex structure J_0 on $\mathbb{C}^n$. $\square$

Holomorphic maps between complex manifolds can be fully characterized in terms of their associated almost complex structures. This indicates that a complex structure is in fact fully determined by its induced almost complex structure.

Proposition 4.3.4. *Let M and N be two complex manifolds, and denote both of their induced almost complex structure by the same letter J . Then a smooth map $f : M \rightarrow N$ is holomorphic if and only if the differential*

$$df_p : (T_p M, J_p) \rightarrow (T_{f(p)} N, J_{f(p)})$$

is complex linear. More concisely, f is holomorphic if and only if

$$J \circ df = df \circ J.$$

Proof. This is easily obtained by putting together the characterization of J as the almost complex structure on M obtained by pulling back J_0 along holomorphic charts, and the characterization of holomorphic maps $\mathbb{C}^n \rightarrow \mathbb{C}^m$ in terms of J_0 . $\square$

Corollary 4.3.3. *If two complex structures on a given smooth manifold induce the same almost complex structure, then they are identical.*

Proof. If two complex structures on M induce the same almost complex structures, then by proposition 4.3.4, the identity $id : M_1 \rightarrow M_2$ is a biholomorphism (M_1 denotes M equipped with one of the two complex structures, and M_2 denotes M equipped with the other complex structure). Thus if $\phi : U \subseteq M \rightarrow \mathbb{C}^n$ is a holomorphic chart for M_1 , then it is also a holomorphic chart for M_2 and vice-versa, which implies that the two complex structures are in fact equal. $\square$

We've established that a complex structure induces, and is fully determined by, an almost complex structure. This naturally poses the problem of determining which almost complex structures arise from complex structures. This question can be answered in a very satisfactory manner by means of the **Nijenhuis tensor** (also called the **torsion** of the almost complex structure).

Definition 4.3.9 (Integrability of almost complex structures). An almost complex structure on a manifold M is said to be **integrable** if it is induced by a complex structure. In that case, that complex structure is unique (corollary 4.3.3).

Definition 4.3.10 (Nijenhuis tensor). Let J be an almost complex structure on M . The **torsion**, or **Nijenhuis tensor** of J is the $(1, 2)$ -tensor defined by $N(X, Y) := [JX, JY] - J[JX, Y] - J[X, JY] - [X, Y]$.

Theorem 4.3.5. J is integrable if and only if $N = 0$.

Proof. See [13] (vol. II, chapter IX, theorem 2.5 p.124) and [10] (chapter 2). □

As a matter of fact the Nijenhuis is doubly miraculous as it behaves very well in the presence of affine connections.

Definition 4.3.11 (Almost complex connection). A connection ∇ on an almost complex manifold (M, J) is almost complex if $\nabla J = 0$.

Proposition 4.3.5. Let ∇ be an almost complex connection on an almost complex manifold (M, J) . Then $N(X, Y) = T(X, Y) + JT(JX, Y) + JT(X, JY) - T(JX, JY)$.

Proof. This is a straightforward application of the formula

$$[X, Y] = \nabla_X Y - \nabla_Y X - T(X, Y).$$

□

Corollary 4.3.4. If J is parallel for a torsionless connection on M , then J is integrable.

Remark. Putting this result and theorem 3.1.1 side by side, we see that both almost complex and almost symplectic structures are integrable as soon as they are parallel for a torsionless connection.

4.3.3 Hermitian and Kähler metrics

The tangent spaces of an almost complex manifold are complex vector spaces, thus a metric on such a manifold should take the form of a smooth choice of Hermitian inner product for each tangent space. As we established in our study of the corresponding linear structures (theorem 4.3.2) this is tantamount to selecting a compatible Riemannian structure (compatible with the almost complex structure), leading to the following definition of Hermitian metrics:

Definition 4.3.12 (Hermitian metrics). A **Hermitian metric** on an almost complex manifold (M, J) is a Riemannian metric g on M such that $g(JX, JY) = g(X, Y) \forall X, Y \in \mathfrak{X}(M)$. An **almost Hermitian manifold** is an almost complex manifold together with a Hermitian metric.

Proposition 4.3.6. *Let (M, J) be an almost complex manifold. Then (M, J) admits a hermitian metric.*

Proof. Choose a Riemannian metric g_0 on M . Then $g := g_0 + g_0(J-, J-)$ is a hermitian metric on (M, J) . $\square$

Remark. Notice that an almost Hermitian manifold comes equipped with a canonical torsionless connection: the Levi-Civita connection of its Hermitian metric.

As Hermitian vector spaces, the tangent spaces of an almost Hermitian manifold each admit a canonical linear symplectic structure (the Kähler form, proposition 4.3.3). These linear symplectic structures assemble into a non-degenerate 2-form, which will not be closed in general as this process occurs point-by-point. To emphasize the analogy with complex structures, non-degenerate 2-forms are sometimes called **almost symplectic structures** and are said to be **integrable** if they are closed (or equivalently, symplectic), a terminology we will adopt.

Proposition 4.3.7. *Let (M, J, g) be an almost Hermitian manifold. The formula*

$$\omega(X, Y) := g(X, JY)$$

*for $X, Y \in \mathfrak{X}(M)$ defines an **almost symplectic form** ω on M called the **Kähler form** of (M, J, g) .*

Furthermore, ω is compatible with J (or preserved by J) in the sense that

$$\omega(JX, JY) = \omega(X, Y)$$

for all $X, Y \in \mathfrak{X}(M)$.

Remark. Using contraction we can write $\omega = C_2^1(g \otimes J)$.

Corollary 4.3.5 (bis). *An almost complex manifold is even-dimensional and admits a canonical orientation, obtained by equipping each tangent space with its canonical orientation as a complex vector space.*

Proof. Let (M, J) be an almost complex manifold. The only statement left to prove is the one concerning the canonical orientation of (M, J) . Choose a hermitian metric g on (M, J) (such a metric always exists - proposition 4.3.6), and let ω be the associated Kähler form. Then ω^n is a volume form on M , and for each $p \in M$, the orientation induced by ω_p^n coincides with the orientation induced by J_p (proposition 4.3.3). $\square$

So we see that for almost Hermitian manifolds, the problem of integrability arises for both the Kähler and almost complex structures. All possible combinations occur and are labeled as follows:

- (i) If the Kähler form is integrable, but not necessarily the almost complex structure, one usually speaks of a symplectic manifold (M, ω) equipped with a compatible (or ω -compatible) almost complex structure J .
- (ii) If the almost complex structure is integrable, but not necessarily the Kähler form, one speaks of a Hermitian manifold (M, J, g) .
- (iii) If both the Kähler form *and* the almost complex structure are integrable, one speaks of a **Kähler manifold** (M, J, g) .

Only Kähler manifolds will intervene in this thesis, and no further mention will be made of the other two cases. Our remaining two objectives are:

- (i) Give a characterization of Kähler manifolds in terms of their Levi-Civita connection.
- (ii) Derive certain identities satisfied by the curvature of a Kähler manifold, which will be needed in the next chapter.

We are already in possession of most of what is needed to complete our first objective, the only missing ingredient is an equation relating the Levi-Civita connection of an almost Hermitian manifold with its other structures, which we state below.

Theorem 4.3.6. *Let (M, J, g) be an almost Hermitian manifold with Levi-Civita connection ∇ and Kähler form ω , and denote the torsion of J by N . Then*

$$d\omega(X, JY, JZ) - d\omega(X, Y, Z) = 2g(Y, (\nabla_X J)Z) + g(X, JN(Y, Z))$$

for all $X, Y, Z \in \mathfrak{X}(M)$.

Remark. We will apply this equation in the Kähler case (so when $d\omega = 0$ and $N = 0$) to obtain $\nabla J = 0$ (and by extension $\nabla\omega = 0$).

Proof. This is obtained through ingenious applications of formulae we already know, we refer the reader to [13] (vol. II, chapter IX, prop.4.2 p.148 - beware that the conventions used by Kobayashi and Nomizu to define the wedge product and exterior derivative differ from ours, which explains the difference in coefficients). $\square$

With this we are ready to give our characterization:

Theorem 4.3.7. *Let (M, J, g) be an almost Hermitian manifold with connection ∇ and Kähler form ω . The following are equivalent:*

- (i) (M, J, g) is Kähler.
- (ii) $\nabla J = 0$.

(iii) $\nabla\omega = 0$.

Thus both the symplectic and complex structure of a Kähler manifold are parallel for the Levi-Civita connection. In particular, the Levi-Civita connection of a Kähler manifold is symplectic for its Kähler form.

Proof. We begin by showing (ii) $\Leftrightarrow$ (iii). Note that the underlying theme is the 2-out-of-three property (theorem 4.3.3) which tells us in essence that as soon as 2 members of (g, J, ω) are parallel for a given connection ∇ , the third must be as well (and here we already know that $\nabla g = 0$).

(ii) $\Rightarrow$ (iii): Assume that $\nabla J = 0$. Then by combining $\nabla g = 0$, $\omega = C_2^1(g \otimes J)$ and corollary 2.2.2 we see that for all $X \in \mathfrak{X}(M)$

$$\begin{aligned}\nabla_X \omega &= \nabla_X C_2^1(g \otimes J) = C_2^1[\nabla_X(g \otimes J)] \\ &= C_2^1[\underbrace{(\nabla_X g)}_{=0} \otimes J + g \otimes \underbrace{(\nabla_X J)}_{=0}] \\ &= 0.\end{aligned}$$

(iii) $\Rightarrow$ (ii). Assume that $\nabla\omega = 0$, then a computation similar to the previous one yields

$$0 = \nabla_X \omega = C_2^1[g \otimes (\nabla_X J)]$$

so that for all $X, Y, Z \in \mathfrak{X}(M)$,

$$0 = g(Y, (\nabla_X J)Z)$$

and since g is non-degenerate, this implies $\nabla J = 0$.

Now (ii) $\Rightarrow$ (i) is clear: if (ii) holds, then (iii) holds as well so that we have both $\nabla J = 0$ and $\nabla\omega = 0$, and in virtue of corollary 4.3.4 and theorem 3.1.1, both J and ω must be integrable.

Lastly, let us show (i) $\Rightarrow$ (ii). If (M, J, g) is Kähler (and in particular almost hermitian), then by combining $d\omega = 0$, $N = 0$ (with N the torsion of J) and theorem 4.3.6 we have that for all $X, Y, Z \in \mathfrak{X}(M)$

$$0 = g(Y, (\nabla_X J)Z)$$

and we conclude that $\nabla J = 0$. □

The compatibility between the complex structure of a Kähler manifold and its Levi-Civita connection, together with the known symmetries of a metric curvature, then yield:

Theorem 4.3.8. *Let (M, J, g) be a Kähler manifold, and let $\hat{R}$ be the (Riemannian) covariant curvature tensor of its Levi-Civita connection. Then for all vector fields $X, Y, Z, W \in \mathfrak{X}(M)$ we have:*

$$\hat{R}(X, Y, Z, W) = \hat{R}(X, Y, JZ, JW) = \hat{R}(JX, JY, Z, W)$$

(in addition to the usual symmetries of the Riemannian covariant curvature tensor).

Proof. By theorem 4.3.7, $\nabla J = 0$, thus if we apply corollary 2.2.4 to J (and recall example 1.1.4 (iii)) we obtain

$$\begin{aligned} 0 &= (R(X, Y) \cdot J)Z = R(X, Y)(JZ) - JR(X, Y)Z \\ &\Leftrightarrow R(X, Y)(JZ) = JR(X, Y)Z \end{aligned}$$

which, when combined with the compatibility of g and J gives us

$$\begin{aligned} \hat{R}(X, Y, JZ, JW) &= g(R(X, Y)JZ, JW) = g(JR(X, Y)Z, JW) \\ &= g(R(X, Y)Z, W) = \hat{R}(X, Y, Z, W). \end{aligned}$$

Finally, using (iii) in theorem 4.1.2 we get

$$\hat{R}(X, Y, JZ, JW) = \hat{R}(JZ, JW, X, Y) = \hat{R}(Z, W, X, Y) = \hat{R}(X, Y, Z, W).$$

□

These identities, added to the usual identities of the metric curvature, will constitute what we call 'the symmetries of a Kähler curvature'.

4.4 Locally symmetric spaces

The reader looking for an introduction to symmetric spaces should see their wishes fulfilled by [24] or [12, 13].

In their barest forms, locally symmetric spaces are affinely connected manifolds whose local geodesic symmetries are affine transformation.

Definition 4.4.1 (Affine locally symmetric space). An manifold M equipped with an affine connection ∇ is said to be **affine locally symmetric** if its local geodesic symmetries are affine transformations, meaning that for any point $p \in M$ and any symmetric normal neighborhood U of p , the involution $s_p : U \rightarrow U$ preserves the affine connection.

Such spaces admit a strikingly simple characterization in terms of their torsion and curvature.

Theorem 4.4.1. *(M, ∇) is affine locally symmetric if and only if $T = 0$ and $\nabla R = 0$.*

Proof. We only prove that these conditions are necessary, for proof that they are sufficient, we refer the reader to [13] (vol. II, chapter IX, theorem 1.1, p.222) or [24] (chapter IV, theorem 1.1, p.163).

To see that they are necessary, simply notice that when (M, ∇) is affine locally symmetric, T and ∇R are tensors of odd degree which are preserved by the local geodesic symmetries, and must thus vanish by proposition 3.3.4. $\square$

Locally symmetric spaces can then be flavoured by adding geometric structures and requiring that these additional structure be compatible with the local geodesic symmetries. In our case these additional structures will take the form of tensor fields, which will be required to be symmetric.

In several instances these tensor fields will already be known to be parallel, which will allow us to invoke theorem 4.4.1 as a parallel tensor field defined over an affine locally symmetric space is automatically symmetric. This is the motivation behind the following sequence of results:

Proposition 4.4.1. *Let ∇ be an affine connection on a connected manifold M . Let W_1, W_2 be two parallel tensor fields of the same type on M . Assume that there exists a point $p \in M$ such that $W_1(p) = W_2(p)$. Then $W_1 = W_2$.*

Proof. Take any point $q \in M$. Since M is a connected manifold, there exists a smooth curve $\gamma : [0, 1] \rightarrow M$ such that $\gamma(0) = p$ and $\gamma(1) = q$. Since W_1 and W_2 are parallel, both γ^*W_1 and γ^*W_2 are parallel along γ (proposition 2.2.5), so by theorem 2.2.4 we have

$$W_2(q) = P(\gamma)_0^1 W_2(p) = P(\gamma)_0^1 W_1(p) = W_1(q).$$

$\square$

Corollary 4.4.1. *Let $(M, \nabla), (M', \nabla')$ be connected smooth manifolds equipped with affine connections. Let W be a parallel tensor field on M and W' be a parallel tensor field of the same type on M' . Let $f : M \rightarrow M'$ be a diffeomorphism preserving the connection. Assume there exists $p \in M$ such that $(fW)(p) = W'(f(p))$ then $W' = fW$.*

Proof. This follows immediately from proposition 4.4.1 once we notice that fW is parallel for ∇' . $\square$

Corollary 4.4.2. *Let (M, ∇) be an affine locally symmetric space and let W be a parallel tensor field on M . Then W is symmetric.*

Proof. We assume that W is of degree (r, s) . Take a point p of M . Since (M, ∇) is affine locally symmetric, there exists a symmetric normal neighborhood U of p such that $s_p : U \rightarrow U$ preserves the connection. Furthermore, since $d(s_p)_p = -I_p$ we have $(s_p W)(p) = (-1)^{r+s} W(p)$, and since any constant multiple of W is parallel, corollary 4.4.1 implies $s_p W = (-1)^{r+s} W$. $\square$

Let us now list the flavours of symmetric spaces relevant to our undertaking.

Definition 4.4.2 (Riemannian locally symmetric space). A Riemannian manifold (M, g) is locally symmetric if its local geodesic symmetries (for the Levi-Civita connection) are isometries, or equivalently if g is symmetric for the Levi-Civita connection (note that g is of even degree).

Remark. In that case, M equipped with the Levi-Civita connection of g is affine locally symmetric (this is an immediate consequence of corollary 4.1.1).

Definition 4.4.3 (Complex affine symmetric space). An almost complex manifold (M, J) together with an affine connection ∇ is **complex affine locally symmetric** if its local geodesic symmetries are affine and holomorphic transformations (that is, if they leave ∇ and J invariant). Equivalently, a complex affine locally symmetric space is an affine locally symmetric space with a symmetric almost complex structure (J has even degree).

Remark. In that case ∇ is torsionless as a locally symmetric connection, while J is parallel for ∇ as a symmetric tensor, thus J is integrable (corollary 4.3.4).

Definition 4.4.4 (Symplectic locally symmetric). A **symplectic locally symmetric space** is a symplectic manifold (M, ω) together with a connection ∇ whose local geodesic symmetries are affine transformations and symplectomorphisms. Equivalently a symplectic locally symmetric space is an affine locally symmetric space with a symmetric symplectic form (once again, ω is of even degree).

Remark. Such a connection is necessarily symplectic of type S (corollary 3.3.2).

Remark. Similar to complex affine locally symmetric spaces we could allow the 2-form to be *almost* symplectic and still obtain an equivalent definition, as this form would be parallel for a torsionless connection and thus closed (theorem 3.1.1).

At the intersection of all three of these classes lie **locally symmetric Kähler manifolds**, also called **Hermitian locally symmetric spaces**.

Theorem 4.4.2. *Let (M, J, g) be an almost Hermitian manifold with Levi-Civita connection ∇ . The following are equivalent:*

- (i) (M, J, ∇) is locally complex affine symmetric.
- (ii) (M, J, g) is a Kähler manifold which is locally symmetric as a Riemannian manifold.

Furthermore, when (i)-(ii) are satisfied, (M, ∇, ω) is symplectic locally symmetric, in particular ∇ is symplectic of type S.

A manifold verifying conditions (i) and (ii) is called a **Hermitian locally symmetric space**, or a **locally symmetric Kähler manifold**.

Proof. (i) $\Rightarrow$ (ii). Assume that (M, J, ∇) is locally complex affine symmetric. Then corollary 3.3.2 implies $\nabla J = 0$ from which we deduce that (M, J, g) is Kähler (theorem 4.3.7). Furthermore, since (M, ∇) is locally affine symmetric, and since $\nabla g = 0$, an application of corollary 4.4.2 shows that g is symmetric, so that (M, g) is a Riemannian locally symmetric space.

(ii) $\Rightarrow$ (i). As a Riemannian locally symmetric space (M, g) equipped with its Levi-Civita connection is necessarily locally affine symmetric. Furthermore, by theorem 4.3.7 we have $\nabla J = 0$, so J must be symmetric as a consequence of corollary 4.4.2.

Similarly, if (ii) holds we show that ω is symmetric for ∇ by combining theorem 4.3.7 and corollary 4.4.2. $\square$

We close this chapter by providing the promised characterizations of locally symmetric spaces one can obtain through a combination of corollary 4.4.2 and theorem 4.4.1.

Theorem 4.4.3. *Let ∇ be an affine connection on a smooth manifold M . Consider the following cases:*

- (i) ∇ is the Levi-Civita connection of a Riemannian metric g on M .
- (ii) ∇ is a torsionless complex connection for an almost complex structure J on M .
- (iii) ∇ is a symplectic connection for a symplectic form ω on M .
- (iv) ∇ is the Levi-Civita connection of a Kähler structure (J, g) on M .

To each of these cases, we will assign a statement that is equivalent to $\nabla R = 0$ (or equivalently, to (M, ∇) being affine locally symmetric, since the connection is torsionless in every case). In this manner we have in each respective case:

- (i) (M, g) is a locally symmetric Riemannian manifold if and only if $\nabla R = 0$.
- (ii) (M, J, ∇) is complex affine locally symmetric if and only if $\nabla R = 0$.
- (iii) (M, ω, ∇) is symplectic locally symmetric if and only if $\nabla R = 0$.
- (iv) (M, J, g) is a locally symmetric Kähler manifold if and only if $\nabla R = 0$.

Chapter 5

The Sekigawa-Vanhecke theorem

Earlier we saw that whenever a Kähler manifold is locally symmetric as a riemannian manifold, its local geodesic symmetries preserve all the structures present (complex and symplectic). This result could not be said to be genuinely surprising, given the strong link that exists between a Riemannian metric and its Levi-Civita connection. Less expected, if one is acquainted with the theory of symplectic connections, is the theorem of Sekigawa and Vanhecke [2] which offers a converse of sorts: if the local geodesic symmetries of an almost Hermitian manifold are symplectic, then the manifold is in fact Kähler and locally symmetric.

The Kähler part is obtained immediately from theorem 4.3.7 and the fact that symmetric tensors are parallel (corollary 3.3.2). Proving that the manifold is not only Kähler, but in fact locally symmetric is less straightforward, but not exceedingly so. To achieve this we start with the curvature equation we obtain by applying theorem 3.3.6 to the Kähler form, and deduce from it $\nabla R = 0$ (theorem 4.4.3) by repeated applications of polarization (corollary 1.2.1) and the symmetries of a Kähler curvature.

Theorem 5.0.1 (Sekigawa-Vanhecke [2]). *An almost Hermitian manifold whose local geodesic symmetries are symplectic (i.e. whose Kähler form is symmetric for the Levi-Civita connection) is a locally symmetric Kähler manifold.*

A note before we initiate the proof: earlier we saw that connections of type S are necessarily preferred. The Sekigawa-Vanhecke theorem enables us to show that this inclusion is a strict one.

Corollary 5.0.1. *There exists preferred symplectic connections that are not of type S.*

Proof by example. In virtue of the Sekigawa-Vanhecke theorem, it suffices to provide an example of a Kähler manifold whose Ricci curvature is parallel, but which is *not* locally symmetric. There are many such manifolds, a list of examples can be found in [3], theorem 14.98 (chapter 14, p.420). There the reader will find a class of non-locally-symmetric quaternion-Kähler manifolds. Quaternion-Kähler manifolds are automatically Einstein, which implies that their Ricci curvature is parallel (so cyclic parallel in particular). $\square$

After applying theorem 3.3.6 to the Kähler form, we will prove the Sekigawa-Vanhecke theorem by making use of the following two lemmas.

Lemma 5.0.1. *Let (V, J) be a finite-dimensional real vector space equipped with a complex structure J . A quadrilinear mapping*

$$R : V \times V \times V \times V \rightarrow \mathbb{R}$$

is said to possess the symmetries of a Kähler curvature if for all $X, Y, Z, W \in V$ we have

- (i) $R(X, Y, Z, W) = -R(Y, X, Z, W) = -R(X, Y, W, Z)$.
- (ii) $R(X, Y, Z, W) = R(Z, W, X, Y)$.
- (iii) $\mathfrak{S}_{X, Y, Z} R(X, Y, Z, W) = 0$.
- (iv) $R(X, Y, Z, W) = R(JX, JY, Z, W) = R(X, Y, JZ, JW)$.

So let R and T be two quadrilinear mappings on V possessing the symmetries of a Kähler curvature. If

$$R(X, JX, X, JX) = T(X, JX, X, JX)$$

for all $X \in V$, then $R = T$.

Proof. The proof is a sequence of algebraic manipulations that the reader will find in [13] (see proposition 7.1, Volume II, chapter IX, section 7, p. 165). $\square$

Remark. By putting together theorems 4.1.2 and 4.3.8, one can check that, given a Kähler manifold (M, J, g) with covariant Riemannian curvature tensor $\hat{R}$, for any point $p \in M$, $\hat{R}_p$ viewed as a quadrilinear mapping on $(T_p M, J_p)$ does indeed possess the symmetries of a Kähler curvature (bringing us much relief).

We bring attention to the fact that, since $\nabla J = 0$ on a Kähler manifold, for any vector field X on M the $(0, 4)$ -tensor $(\nabla_X \hat{R})$ also possesses the symmetries of a Kähler curvature.

Lemma 5.0.2. *Let (M, J, g) be a Kähler manifold with covariant Riemannian curvature tensor $\hat{R}$ and assume that for all vector fields $X \in \mathfrak{X}(M)$*

$$(\nabla_X \hat{R})(X, JX, X, JX) = 0. \tag{5.1}$$

Then (M, J, g) is a locally symmetric Kähler manifold.

Proof. Following [11] (see the proof of proposition 3.12, section 3, p. 351) we will first show that

$$(\nabla_X \hat{R})(Y, JY, Y, JY) = 0 \tag{5.2}$$

for all $X, Y \in \mathfrak{X}(M)$, then conclude using lemma 5.0.1.

So let us fix $X, Y \in \mathfrak{X}(M)$. Using Bianchi's second identity and the symmetries of the Kähler curvature we get

$$\begin{aligned} & (\nabla_X \hat{R})(Y, JY, Y, JY) \\ &= -(\nabla_Y \hat{R})(JY, X, Y, JY) - (\nabla_{JY} \hat{R})(X, Y, Y, JY) \\ &= (\nabla_Y \hat{R})(Y, JX, Y, JY) + (\nabla_{JY} \hat{R})(JY, JX, Y, JY). \end{aligned} \quad (5.3)$$

Next we polarize equation 5.1. By corollary 1.2.1, $\text{Sym}(\nabla \hat{R}) = 0$. But a simple computation yields

$$\begin{aligned} & \text{Sym}(\nabla \hat{R})(X, Y, Y, Y, Y) \\ &= \frac{1}{5} \left((\nabla_X \hat{R})(Y, JY, Y, JY) + (\nabla_Y \hat{R})(X, JY, Y, JY) + \dots + (\nabla_Y \hat{R})(Y, JY, Y, JX) \right) \end{aligned}$$

which, when combined with the symmetries of the Kähler curvature tensor, gives us

$$0 = (\nabla_X \hat{R})(Y, JY, Y, JY) + 2(\nabla_Y \hat{R})(X, JY, Y, JY) + 2(\nabla_Y \hat{R})(Y, JX, Y, JY). \quad (5.4)$$

We replace Y by JY in 5.4, add the result to 5.4 and use Bianchi's second identity (as in 5.3). This gives us

$$0 = 4(\nabla_X \hat{R})(Y, JY, Y, JY) + 2(\nabla_Y \hat{R})(Y, JX, Y, JY) + 2(\nabla_{JY} \hat{R})(JY, JX, Y, JY). \quad (5.5)$$

And combining 5.3 and 5.5 we obtain 5.2.

Finishing the proof is then just a matter of putting together the appropriate results. For all $p \in M$ and $X_p \in T_p M$ observe that $(\nabla_{X_p} \hat{R})$ is a quadrilinear mapping on $(T_p M, J_p)$ which possesses the symmetries of a Kähler curvature (apply ∇_{X_p} to said symmetries for $\hat{R}$ and use $\nabla J = 0$), so 5.2 and lemma 5.0.1 (setting $T = 0$) imply $\nabla \hat{R} = 0$. Since g is non-degenerate and parallel, this is equivalent to $\nabla R = 0$ and the conclusion follows from theorem 4.4.3 (the characterisation locally symmetric spaces).

□

Proof of the Sekigawa-Vanhecke theorem. Let (M, J, g) be an almost Hermitian manifold, and assume that the Kähler form ω is symmetric for the Levi-Civita connection ∇ . As ∇ is torsionless, ω being symmetric for ∇ implies $\nabla \omega = 0$ (corollary 3.3.2), so (M, J, g) is Kähler (theorem 4.3.7).

By theorem 3.3.6, ω being symmetric for the torsionless connection ∇ also yields

$$P_X \cdot \omega = 0$$

for all $X \in \mathfrak{X}(M)$, where P_X is the endomorphism field defined by $P_X := (\nabla_X R)(X, -)X$. In other words, for all vector fields X, Y, Z on M we have

$$0 = \omega((\nabla_X R)(X, Y)X, Z) + \omega(Y, (\nabla_X R)(X, Z)X).$$

Using the skew-symmetry of ω and $\omega = g(-, J-)$ this is equivalent to

$$0 = (\nabla_X \hat{R})(X, Y, X, JZ) - (\nabla_X \hat{R})(X, Z, X, JY).$$

Setting $Y = JX$ and $Z = X$ (and using $(\nabla \hat{R})(X, X, -, -) = 0$) we obtain

$$(\nabla_X \hat{R})(X, JX, X, JX) = 0$$

for all $X \in \mathfrak{X}(M)$. By lemma 5.0.2, this finishes the proof.

□

Chapter 6

Divergence-preserving geodesic symmetries

The main object of study in this chapter are manifolds with torsionless affine connections whose local geodesic symmetries are divergence-preserving, meaning that they preserve the induced connection on the determinant bundle. If our manifold admits a parallel volume form (or local parallel volume forms), in which case we say that the manifold is (locally) equiaffine, this is equivalent to requiring that the parallel volume form be symmetric.

The study of Riemannian manifolds possessing this property, called D'Atri spaces, was initiated by D'Atri and Nickerson in [8] and [9], which is where our 'divergence-preserving' denomination originates. Most importantly for our purposes, in their work D'Atri and Nickerson use radial power expansions in normal coordinates to show that Riemannian manifolds with divergence-preserving symmetries verify an infinite sequence of curvature conditions, called the odd Ledger conditions, and that these conditions are in fact sufficient in the analytic case. The curvature conditions of Ledger are defined by a recurrence formula, and were introduced in his PhD thesis on harmonic spaces [16].

As it turns out the derivation of D'Atri and Nickerson does not make use of the metric in any essential way and is readily extended to the affine setting. After carrying out this extension, which is the main purpose of this chapter, we will see that manifolds with torsionless affine connections whose local symmetries are divergence-preserving do still verify the odd Ledger conditions, and that, conversely, these conditions are sufficient when the connection is locally equiaffine and analytic.

6.1 Divergence and connections on the determinant bundle

As the name of their paper [8] suggests, central to the work of D'Atri and Nickerson is the notion of divergence on a Riemannian manifold, which is equivalently defined via

the Riemannian volume form or via the Levi-Civita connection.

In an effort to make our presentation as clear and general as possible, we introduce an original notion of divergence on a manifold, one that requires no additional structure (Riemannian, affine or otherwise). The defining characteristic of divergences on a manifold, as defined here, is the fact that they correspond exactly to connections on the determinant bundle of said manifold. Using this feature, we will show later in the section how the usual notions of divergence on a manifold (involving a volume form or an affine connection) are in fact special cases of the definition given here.

Definition 6.1.1 (Divergences). Let M be a smooth manifold. A **divergence** on M is an $\mathbb{R}$ -linear map

$$\operatorname{div} : \mathfrak{X}(M) \rightarrow C^\infty(M)$$

such that for all $X \in \mathfrak{X}(M)$, $f \in C^\infty(M)$,

$$\operatorname{div}(fX) = (Xf) + f\operatorname{div}(X).$$

Remark. The map $\operatorname{div} : \mathfrak{X}(M) \rightarrow C^\infty(M)$ is easily shown to be local in the sense of definition 2.1.7, by using a reasoning similar to the one we applied to operators verifying the Leibniz rule (proposition 2.1.7). In particular, a divergence admits a unique restriction to any open subset $U \subseteq M$ (proposition 2.1.7).

The correspondence between divergences as defined above and connections on the determinant bundle can be stated as follows:

Theorem 6.1.1. *Let M be a smooth manifold of dimension m and div be a divergence on M . Then the map*

$$\nabla : \mathfrak{X}(M) \times \Omega^m(M) \rightarrow \Omega^m(M), (X, \omega) \mapsto \nabla_X \omega := (\mathcal{L}_X - \operatorname{div}(X))\omega \quad (6.1)$$

is a connection on the determinant bundle of M . Conversely for any connection ∇ on the determinant bundle of M , there exists a unique divergence div that induces ∇ as in 6.1, so that we have a one-to-one correspondence between divergences on M and connections on the determinant bundle of M .

To prove this correspondence, we use the following three lemmas:

Lemma 6.1.1. *Let L be a line bundle over M . Then there is a one-to-one correspondence between bundle morphisms $L \rightarrow L$ and elements of $C^\infty(M)$. To $f \in C^\infty(M)$ we associate the bundle morphism that corresponds to the following $C^\infty(M)$ -linear map of sections:*

$$\Gamma(L) \rightarrow \Gamma(L), \sigma \mapsto f\sigma.$$

Lemma 6.1.2. *Let ∇ be a connection on $T_s^r M$. Then for any vector field $X \in \mathfrak{X}(M)$, $A_X := \mathcal{L}_X - \nabla_X : \Gamma(T_s^r M) \rightarrow \Gamma(T_s^r M)$ is $C^\infty(M)$ -linear and thus corresponds to a bundle morphism $T_s^r M \rightarrow T_s^r M$.*

Lemma 6.1.3. *For any vector field X and any top form ω on M*

$$\mathcal{L}_{fX}\omega = [(Xf) + f\mathcal{L}_X]\omega.$$

Proof. To prove lemma 6.1.1, notice first that $\text{Lin}(L, L)$ is a real line bundle, then notice that the identity section provides a global trivialization of $\text{Lin}(L, L)$. We then conclude using the correspondence between bundle morphisms $L \rightarrow L$ and global sections of $\text{Lin}(L, L)$.

Lemma 6.1.2 follows from

$$\mathcal{L}_X(fT) = (Xf)T + f\mathcal{L}_X T$$

for all $X \in \mathfrak{X}(M)$, $f \in C^\infty(M)$, $T \in \Gamma(T_s^r M)$.

To prove lemma 6.1.3: given a point $p \in M$, we choose a local frame $E_1, \dots, E_m$ of TM on a neighborhood of p , in which we write $X = \sum_{i=1}^m X^i E_i$. We then compute

$$\mathcal{L}_{fX}\omega(E_1, \dots, E_m) = fX(\omega(E_1, \dots, E_m)) - \sum_{i=1}^m \omega(E_1, \dots, [fX, E_i], \dots, E_m). \quad (6.2)$$

From $[fX, E_i] = f[X, E_i] - (E_i f)X$ and the skew-symmetry of ω we obtain

$$\begin{aligned} & \omega(E_1, \dots, [fX, E_i], \dots, E_m) \\ &= f\omega(E_1, \dots, [X, E_i], \dots, E_m) - (E_i f)\omega(E_1, \dots, X, \dots, E_m) \\ &= f\omega(E_1, \dots, [X, E_i], \dots, E_m) - \sum_{j=1}^m (X^j E_i f)\omega(E_1, \dots, E_j, \dots, E_m) \\ &= f\omega(E_1, \dots, [X, E_i], \dots, E_m) - (X^i E_i f)\omega(E_1, \dots, E_m) \end{aligned} \quad (6.3)$$

injecting 6.3 into 6.2 we get

$$\begin{aligned} \mathcal{L}_{fX}\omega(E_1, \dots, E_m) &= f \left(X(\omega(E_1, \dots, E_m)) - \sum_{i=1}^m \omega(E_1, \dots, [X, E_i], \dots, E_m) \right) \\ &\quad + \left(\sum_{i=1}^m X^i (E_i f) \right) \omega(E_1, \dots, E_m) \\ &= f(\mathcal{L}_X \omega)(E_1, \dots, E_m) + (Xf)\omega(E_1, \dots, E_m). \end{aligned}$$

(Note that every time we use the notation $\omega(E_1, \dots, A, \dots, E_m)$ in the preceding computations, A is the i th argument). $\square$

Proof of theorem 6.1.1. Let div be a divergence on M , and define

$$\nabla_X \omega := \mathcal{L}_X \omega - \text{div}(X)\omega \stackrel{n}{=} (\mathcal{L}_X - \text{div}(X))\omega$$

for $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^m(M)$.

The linearity of ∇ follows from the linearity of div and $\mathcal{L}$, and for any $f \in C^\infty(M)$, $X \in \mathfrak{X}(M)$, $\omega \in \Omega^m(M)$ we have by lemma 6.1.3 and by the definition of a divergence,

$$\begin{aligned} \nabla_X(f\omega) &= \mathcal{L}_X(f\omega) - f\text{div}(X)\omega \\ &= (Xf)\omega + f\mathcal{L}_X\omega - f\text{div}(X)\omega \\ &= (Xf)\omega + f\nabla_X\omega \end{aligned}$$

and

$$\begin{aligned} \nabla_{fX}\omega &= \mathcal{L}_{fX}\omega - \text{div}(fX)\omega \\ &= (Xf)\omega + f\mathcal{L}_X\omega - (Xf)\omega - f\text{div}(X)\omega \\ &= f\nabla_X\omega \end{aligned}$$

which shows that ∇ is a connection on the determinant bundle.

We finish the proof by showing that, conversely, given a connection ∇ on the determinant bundle, there exists a unique map $\text{div} : \mathfrak{X}(M) \rightarrow C^\infty(M)$ such that

$$\nabla_X\omega = (\mathcal{L}_X - \text{div}(X))\omega$$

and that this map is a divergence.

So let ∇ be such a connection. Then for any $X \in \mathfrak{X}(M)$,

$$A_X := \mathcal{L}_X - \nabla_X : \Omega^m(M) \rightarrow \Omega^m(M)$$

is a bundle morphism by lemma 6.1.2, and since the determinant bundle is a line bundle, by lemma 6.1.1 there exists a unique smooth function $\text{div}(X) \in C^\infty(M)$ such that $\nabla_X\omega = (\mathcal{L}_X - \text{div}(X))\omega$ for all $\omega \in \Omega^m(M)$.

The linearity of div follows from the linearity of ∇ and $\mathcal{L}$, and for $f \in C^\infty(M)$, $X \in \mathfrak{X}(M)$, $\omega \in \Omega^m(M)$, using lemma 6.1.3 we get

$$\begin{aligned} \text{div}(fX)\omega &= \mathcal{L}_{fX}\omega - \nabla_{fX}\omega \\ &= (Xf)\omega + f(\mathcal{L}_X\omega - \nabla_X\omega) \\ &= [(Xf) + f\text{div}(X)]\omega \end{aligned}$$

which shows that div is a divergence. □

Let us now examine two alternative definitions of divergence one can find in the literature. One of these alternatives relies on a choice of volume form, the other relies on a

choice of affine connection. Each of them corresponds to a choice of connection on the determinant bundle in a canonical way, and can thus be considered a special case of the notion of divergence introduced here.

Theorem 6.1.2. *Let L be a real line bundle on a smooth manifold M , and let s be a non-vanishing section of L . There exists a unique connection ∇ on L such that $\nabla s = 0$.*

Proof. The non-vanishing section s provides a global trivialization of L . Clearly the unique connection ∇ on L for which $\nabla s = 0$ is the pullback of the standard flat connection on the trivial real bundle of rank 1. $\square$

Corollary 6.1.1. *Let ω be a volume form on M (that is, a non-vanishing top form). There exists a unique connection ∇ on the determinant bundle for which ω is parallel. The corresponding divergence div is defined by $\text{div}(X)\omega = \mathcal{L}_X\omega$ for all $X \in \mathfrak{X}(M)$.*

Proof. By theorem 6.1.2 there exists a unique connection ∇ on the determinant bundle such that $\nabla\omega = 0$.

For any $X \in \mathfrak{X}(M)$, by definition of the divergence associated to ∇ and since $\nabla\omega = 0$ we have

$$\text{div}(X)\omega = (\mathcal{L}_X - \nabla_X)\omega = \mathcal{L}_X\omega$$

and since ω is non-vanishing, this defines $\text{div}(X)$. $\square$

Theorem 6.1.3. *Let ∇ be an affine connection on M . ∇ induces a connection on the determinant bundle of M and the corresponding divergence is given by*

$$\text{div}(X) = \text{tr}(T_X + \nabla X)$$

for any $X \in \mathfrak{X}(M)$, where the endomorphisms T_X and ∇X are defined by

$$\begin{aligned} T_X Y &:= T(X, Y) \\ (\nabla X)Y &:= \nabla_Y X \end{aligned}$$

in particular, if ∇ is torsionless then $\text{div}(X) = \text{tr}\nabla X = \text{tr}(Y \rightarrow \nabla_Y X)$.

Before proving this theorem, let us derive the expression of divergence in a coordinate chart $(U, x^1, \dots, x^m)$ for a torsionless connection, in terms of the Christoffel symbols Γ_{ij}^k . Using the coordinate expression for the covariant derivative of a vector field (equation 2.8), we see that for a vector field X with components

$$X = \sum_{i=1}^m X^i \partial_i$$

the components of ∇X are given by

$$(\nabla X)_j^i = \partial_j X^i + \sum_{k=1}^m \Gamma_{jk}^i X^k$$

and from this, we obtain the following expression for the divergence:

$$\operatorname{div}(X) = \sum_{i=1}^m \left(\partial_i X^i + \sum_{j=1}^m \Gamma_{ij}^i X^j \right). \quad (6.4)$$

Proof of theorem 6.1.3. We also denote the connection on the determinant bundle by ∇ . We fix $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^m(M)$. Given a point $p \in M$, we choose a local frame $(E_1, \dots, E_m)$ defined on an open neighborhood of p . Using $\nabla_X E_i = T_X E_i + \nabla_{E_i} X + [X, E_i]$ we compute

$$\begin{aligned} (\nabla_X \omega)(E_1, \dots, E_m) &= X(\omega(E_1, \dots, E_m)) - \sum_{i=1}^m \omega(E_1, \dots, \nabla_X E_i, \dots, E_m) \\ &= \left(X(\omega(E_1, \dots, E_m)) - \sum_{i=1}^m \omega(E_1, \dots, [X, E_i], \dots, E_m) \right) - \sum_{i=1}^m \omega(E_1, \dots, (T_X + \nabla X)E_i, \dots, E_m) \\ &= [\mathcal{L}_X - \operatorname{tr}(T_X + \nabla X)]\omega(E_1, \dots, E_m) \end{aligned}$$

in the last equality we use the fact that for $A \in \operatorname{End}(T_x M)$, $(A \cdot \omega_x) = -(\operatorname{tr} A)\omega_x$ (example 1.1.4, (ii)). This shows that $\operatorname{div}(X) = \operatorname{tr}(T_X + \nabla X)$. $\square$

Remark. Notice that at the intersection of these two special cases lie the equiaffine manifolds.

Lastly let us take a look at isomorphisms of manifolds with divergence, which we call divergence-preserving diffeomorphisms.

Definition 6.1.2 (Divergence-preserving diffeomorphisms). Let M, M' be smooth manifolds, let ∇, ∇' be connections on their respective determinant bundles. A diffeomorphism $f : M \rightarrow M'$ is said to be **divergence-preserving** if the isomorphism between the determinant bundles of M and M' induced by f preserves their connections, in other words if for all $\omega' \in \Omega^m(M')$ and $X \in \mathfrak{X}(M)$ we have

$$\nabla_X(f^* \omega') = f^*(\nabla_{f_* X} \omega').$$

Of course, since connections on the determinant bundle of a manifold correspond to divergences on said manifold, divergence-preserving diffeomorphism can equivalently be defined via divergences (which justifies our nomenclature). To obtain such a definition, we start by characterizing the pullback connection in terms of divergences.

Theorem 6.1.4. *Let M, M' be smooth manifolds, let ∇' be a connection on the determinant bundle of M' with divergence div' , and let $f : M \rightarrow M'$ be a diffeomorphism. f induces an isomorphism between the determinant bundles of M and M' and pulling ∇' back along this isomorphism we obtain a connection ∇ on the determinant bundle of M with divergence div . Then for all $X \in \mathfrak{X}(M)$*

$$\text{div}(X) = \text{div}'(f_*X) \circ f.$$

Proof. Let ω' be a top form on M' , by construction of the pullback connection we have

$$\nabla_X(f^*\omega') = f^*(\nabla_{f_*X}\omega')$$

combining this with the naturality of the Lie derivative, we compute

$$\begin{aligned} \text{div}(X)f^*\omega' &= (\mathcal{L}_X - \nabla_X)f^*\omega' = f^*[(\mathcal{L}_{f_*X} - \nabla_{f_*X})\omega'] \\ &= (\text{div}(f_*X) \circ f)f^*\omega'. \end{aligned}$$

□

As a corollary we obtain the following:

Corollary 6.1.2. *Let M, M' be smooth manifolds, let ∇, ∇' be connections on their respective determinant bundles and let div, div' be their corresponding divergences.*

Let $f : M \rightarrow M'$ be a diffeomorphism. The following two conditions are equivalent:

- (i) *f is divergence-preserving.*
- (ii) *The divergences are preserved by f in the sense that for all $X \in \mathfrak{X}(M)$ we have*

$$\text{div}(X) = \text{div}'(f_*X) \circ f.$$

Finally when parallel volume forms exist, they determine the connection on the determinant bundle as stated in corollary 6.1.1, and we can naturally use them to characterize pullback connections and divergence-preserving diffeomorphisms.

Theorem 6.1.5. *As in theorem 6.1.4, let $f : M \rightarrow M'$ be a diffeomorphism, let ∇' be a connection on the determinant bundle of M' and let ∇ be the connection on the determinant bundle of M obtained by pulling ∇' back along f .*

Let ω' be a parallel volume form on M' . Then ∇ is the unique connection on the determinant bundle of M for which $f^*\omega'$ is parallel.

Proof. By proposition 2.2.5, $\nabla'\omega' = 0$ implies $\nabla f^*\omega' = 0$. Since $f^*\omega'$ is a volume form on M , the theorem is then a consequence of corollary 6.1.1. □

Corollary 6.1.3. *Let $f : M \rightarrow M'$ be a diffeomorphism, let ∇ and ∇' be connections on the determinant bundles of M and M' respectively and let ω' be a parallel volume form on M' . Then f is divergence-preserving if and only if $f^*\omega'$ is parallel.*

6.2 Divergence-preserving symmetries

Equipped the proper notions of divergence and divergence-preserving diffeomorphisms, we begin the task which we assigned ourselves: showing that a torsionless connection whose local symmetries are divergence-preserving verifies the odd Ledger conditions, and that the converse holds when the connection is analytic and locally equiaffine. As stated before, this section is but a repetition of D'Atri and Nickerson's work [8], with only the slight modifications required to recast it in an affine setting. We start with some preliminary definitions.

Definition 6.2.1 (Normal balls). Let ∇ be an affine connection on a smooth manifold M , and let p be a point of M . A **normal ball** centered at p is an open neighborhood of U of p such that there exists normal coordinates $(x^1, \dots, x^m)$ centered at p whose domain include U and in which U corresponds to the set $B(p, \delta) := \{(x^1, \dots, x^m) \in \mathbb{R}^m \mid \sum_{i=1}^m (x^i)^2 < \delta^2\}$ for some $\delta > 0$. When working with normal balls, we will often implicitly assume that a choice of such a normal coordinate system has been made.

Definition 6.2.2 (Connection with divergence-preserving symmetries). Let ∇ be an affine connection on a smooth manifold M . We say that (M, ∇) has **divergence-preserving local geodesic symmetries** if for any point p of M there exists a normal ball U centered at p such that $s_p : U \rightarrow U$ is divergence-preserving, where s_p denotes the local geodesic symmetry with center p .

Before diving into the derivation of the Ledger curvature conditions, we briefly mention an easy consequence of corollary 6.1.3 which will become useful in the next chapter.

Theorem 6.2.1. *Let (M, ∇) be an equiaffine manifold of dimension m with parallel volume form ω . Then (M, ∇) has divergence-preserving symmetries if and only if ω is symmetric for ∇ .*

Proof. Assume that (M, ∇) has divergence-preserving symmetries, and let p be a point of M . Then there exists a normal ball U centered at p such that $s_p : U \rightarrow U$ is divergence-preserving. By corollary 6.1.3, $s_p^*\omega|_U$ is parallel, and since $d(s_p)_p = -I$ we have $(s_p^*\omega)_p = (-1)^m \omega_p$. Thus we see that $s_p^*\omega|_U$ and $(-1)^m \omega|_U$ are parallel volume forms, defined on the connected open U , which coincide at the point p . This implies $s_p^*\omega|_U = (-1)^m \omega|_U$ (proposition 4.4.1).

The converse is an immediate consequence of corollary 6.1.3.

□

Let us now fix the following data: a torsionless affine connection ∇ on a smooth manifold M , a point $p \in M$ and a normal coordinate system $(x^1, \dots, x^m)$ centered at p together with a normal ball U centered at p which corresponds to $B(p, \delta)$ for some $\delta > 0$ in the coordinates $(x^1, \dots, x^m)$, as in definition 6.2.1.

Definition 6.2.3 (Radial vector field). The **radial vector field** V on U is defined as follows: given a point $q \in U$, let $W := \exp_p^{-1}(q)$ so that $\gamma_q : [0, 1] \rightarrow U, t \rightarrow \exp_p(tW)$ is the geodesic joining p and q , then $V_q := \dot{\gamma}_q(1)$, or equivalently V_q is obtained by parallel transporting W along γ_q .

Proposition 6.2.1. *Some immediate properties of the radial vector field:*

(i) *In the normal coordinate system, the radial vector field is expressed by*

$$V = \sum_{i=1}^m x^i \partial_i.$$

In particular, V is a smooth vector field.

(ii) *Let J be an interval and let $\gamma : J \rightarrow U$ be a geodesic starting from p , that is such that $\gamma(0) = p$. Then $V_{\gamma(t)} = t\dot{\gamma}(t)$ for all $t \in J$.*

The introduction of the radial vector field is motivated by the following result:

Theorem 6.2.2. *If the local geodesic symmetry $s_p : U \rightarrow U$ is divergence-preserving, then $\text{div}(V)$ is even in the sense that $\text{div}(V) = \text{div}(V) \circ s_p$. Conversely, if $\text{div}(V)$ is even and if (M, ∇) is locally equiaffine, then $s_p : U \rightarrow U$ is divergence-preserving.*

Proof. Assume that $s_p : U \rightarrow U$ is divergence-preserving. By corollary 6.1.2 we have $\text{div}(V) = \text{div}(s_{p*}V) \circ s_p$, thus it suffices to show that $s_{p*}V = V$. So let q be a point of U and let $\gamma_q : [0, 1] \rightarrow U$ be the unique geodesic such that $p = \gamma_q(0)$ and $q = \gamma_q(1)$. We have $V_q = \dot{\gamma}_q(1)$, which allows us to compute

$$(s_{p*}V)_{s_p(q)} = d(s_p)_q V_q = \left. \frac{d}{dt} \right|_{t=1} (s_p \circ \gamma_q)(t) = V_{s_p(q)}$$

using the fact that $(s_p \circ \gamma_q)$ is the geodesic joining p to $s_p(q)$ in the last equality. This shows that $s_{p*}V = V$.

Conversely, assume that $\text{div}(V) = \text{div}(V) \circ s_p$ and that (M, ∇) is locally equiaffine. Since U is simply connected, there exists a parallel volume form ω defined over U by (iii) in

corollary 3.2.1. In the normal coordinate system, we can write $\omega = \bar{\omega} dx^1 \wedge \dots \wedge dx^m$ for some non-vanishing smooth function $\bar{\omega} \in C^\infty(U)$. To show that $s_p : U \rightarrow U$ is divergence-preserving, we will start by showing that $\bar{\omega}$ is even (once again in the sense that $\bar{\omega} \circ s_p = \bar{\omega}$) from this we will deduce that $s_p^* \omega$ is parallel and conclude using corollary 6.1.3.

So let us show that $\bar{\omega}$ even. Using the coordinate expression of V given in point (i) of proposition 6.2.1 we get $Vx^i = x^i$ from which we deduce $\mathcal{L}_V dx^i = d(Vx^i) = dx^i$ for all $1 \leq i \leq m$. Combining this with $\nabla \omega = 0$ we obtain

$$\begin{aligned} \operatorname{div}(V)\omega &= \mathcal{L}_V \omega = (V\bar{\omega})dx^1 \wedge \dots \wedge dx^m + \bar{\omega} \left(\sum_{i=1}^m dx^1 \wedge \dots \wedge \mathcal{L}_V dx^i \wedge \dots \wedge dx^m \right) \\ &= [m\bar{\omega} + V\bar{\omega}]dx^1 \wedge \dots \wedge dx^m. \end{aligned}$$

which implies

$$V \ln(\bar{\omega}) = \frac{V\bar{\omega}}{\bar{\omega}} = \operatorname{div}(V) - m.$$

and since $\operatorname{div}(V)$ is even, $V \ln(\bar{\omega})$ is even as well. Let q be a point of U , and let $\gamma_q : [0, 1] \rightarrow U$ be the geodesic joining p to q as in the first paragraph. From Stokes' theorem and (ii) in proposition 6.2.1 we get

$$\begin{aligned} \ln(\bar{\omega})(q) - \ln(\bar{\omega})(p) &= \int_0^1 \gamma_q^* d \ln(\bar{\omega}) = \int_0^1 dt \dot{\gamma}_q(t) \ln(\bar{\omega}) \\ &= \int_0^1 \frac{dt}{t} (V \ln(\bar{\omega}))(\gamma_q(t)). \end{aligned}$$

Using the fact that $s_p \circ \gamma_q$ is the geodesic joining p to $s_p(q)$ and the fact that $V \ln(\bar{\omega})$ is even we obtain:

$$\begin{aligned} \ln(\bar{\omega})(s_p(q)) - \ln(\bar{\omega})(p) &= \int_0^1 \frac{dt}{t} (V \ln(\bar{\omega}))(s_p \circ \gamma_q(t)) \\ &= \int_0^1 \frac{dt}{t} (V \ln(\bar{\omega}))(\gamma_q(t)) \\ &= \ln(\bar{\omega})(q) - \ln(\bar{\omega})(p) \end{aligned}$$

so that $\ln(\bar{\omega})$ is even, which immediately implies that $\bar{\omega}$ is even.

We are now ready to conclude. For all $1 \leq i \leq m$ we have $s_p^* dx^i = d(x^i \circ s_p) = -dx^i$, combining this with $\bar{\omega} \circ s_p = \bar{\omega}$ we get

$$\begin{aligned} s_p^* \omega &= (\bar{\omega} \circ s_p)(s_p^* dx^1) \wedge \dots \wedge (s_p^* dx^m) \\ &= (-1)^m \bar{\omega} dx^1 \wedge \dots \wedge dx^m \\ &= (-1)^m \omega \end{aligned}$$

so $s_p^* \omega = (-1)^m \omega$ is parallel. $\square$

Somewhat informally, the basic observation that will allow us to restate theorem 6.2.2 in terms of a sequence of curvature conditions is that in order for $\text{div}(V)$ to be even, it is necessary for its derivatives of odd order to vanish, and that, conversely, the vanishing of these odd derivatives is sufficient when the objects involved are analytic. This is made precise in the proposition below:

Proposition 6.2.2.

(i) If $\text{div}(V)$ is even, then for any $r \in 2\mathbb{N} + 1$ and any $X_p \in T_p M$ we have

$$0 = \left. \frac{d^r}{dt^r} \text{div}(V)(\exp(tX_p)) \right|_{t=0}. \quad (6.5)$$

(ii) Conversely, if ∇ is analytic and if 6.5 is verified for all $X_p \in T_p M$ and $r \in 2\mathbb{N} + 1$, then there exist a normal ball $\tilde{U}$ centered at p , contained in U and such that $\text{div}(V)$ restricted to $\tilde{U}$ is even.

Proof. (i) Let $X_p \in T_p M$ and choose $\epsilon > 0$ such that $\exp(tX_p)$ is well-defined and belongs to U for all $-\epsilon < t < \epsilon$. Since $s_p \circ \exp(tX_p) = \exp(-tX_p)$ for all $-\epsilon < t < \epsilon$, and since $\text{div}(V)$ is even, the map $] -\epsilon, \epsilon[\rightarrow \mathbb{R}, t \rightarrow (\text{div}(V))(\exp(tX_p))$ is even and its derivatives of odd order must vanish at zero.

To prove (ii) we switch fully to the normal coordinate system $(x^1, \dots, x^m)$. In coordinates $\text{div}(V) \in C^\infty(U)$ is represented by a smooth map $B(p, \delta) \rightarrow \mathbb{R}$ which we also call $\text{div}(V)$ and s_p is represented by $-I : B(p, \delta) \rightarrow B(p, \delta), (x^1, \dots, x^m) \rightarrow (-x^1, \dots, -x^m)$ so $\text{div}(V)$ is even on a normal ball $B \subseteq B(p, \delta)$ if and only if $\text{div}(V)(x) = \text{div}(V)(-x)$ for all $x \in B$.

Observe first that if (M, ∇) is analytic, then $\text{div}(V)$ is analytic. Indeed, putting together the coordinate expression for the divergence (equation 6.4) with that of the radial vector field (proposition 6.2.1 (i)) we see that

$$\text{div}(V) = m + \sum_{i,j=1}^m \Gamma_{ij}^i x^j$$

and by corollary B.0.1 and theorem B.0.1 the Christoffel symbols of ∇ in normal charts are analytic.

Consequently there exists a power series $\sum_{\alpha \in \mathbb{N}^m} a_\alpha x^\alpha$ converging absolutely on $B(p, \eta)$ for some $0 < \eta < \delta$ such that $\text{div}(V)(x) = \sum_{\alpha \in \mathbb{N}^m} a_\alpha x^\alpha$ for all $x \in B(p, \eta)$ (here we use the multi-index notation $x^\alpha = x_1^{\alpha_1} \dots x_m^{\alpha_m}$ for $\alpha \in \mathbb{N}^m$ and $x \in \mathbb{R}^m$).

For $x \in \mathbb{R}^m$ with $1 = |x|^2 = \sum_{i=1}^m (x^i)^2$, corresponding to $X_p = \sum_{i=1}^m x^i \partial_i|_p$, the geodesic $t \rightarrow \exp(tX_p)$ is represented by $t \rightarrow (tx^1, \dots, tx^m)$ in coordinates, so for all $t \in]-\eta, \eta[$ we get

$$\operatorname{div}(V)(\exp(tX_p)) = \operatorname{div}(V)(tx) = \sum_{k=0}^{\infty} t^k \left(\sum_{|\alpha|=k} a_{\alpha} x^{\alpha} \right)$$

where $|\alpha| := \sum_{i=1}^m \alpha_i$, and 6.5 gives us $0 = \sum_{|\alpha|=r} a_{\alpha} x^{\alpha}$ for all $r \in 2\mathbb{N} + 1$ and all $x \in S(p, 1) := \{x \in \mathbb{R}^m | 1 = |x|^2\}$. Thus, if we fix $r \in 2\mathbb{N} + 1$, we see that the homogeneous polynomial of degree r in m variables $Q_r(x) := \sum_{|\alpha|=r} a_{\alpha} x^{\alpha}$ vanishes, which implies $a_{\alpha} = 0$ whenever $|\alpha| = r$.

Consequently all the terms of odd degree in the power series of $\operatorname{div}(V)$ vanish and on the normal ball $B(p, \eta)$, $\operatorname{div}(V)(x) = \sum_{k=0}^{\infty} \sum_{|\alpha|=2k} a_{\alpha} x^{\alpha}$ is even. $\square$

With proposition 6.2.2 in hand, our final task in the derivation of the odd Ledger conditions is reformulating 6.5 in terms of R , the curvature tensor of ∇ . Making use of the torsionless hypothesis on the connection ∇ , we start by reexpressing $\operatorname{div}(V)$ using the formula given in theorem 6.1.3 for the divergence associated to an affine connection. Following [8] and [9] we introduce the notation

Notation 6.2.1. $\Lambda := \nabla V \in \Gamma(U, \operatorname{End}(TM))$, so $\Lambda(Y) = \nabla_Y V$ for all $Y \in \mathfrak{X}(U)$. We have $\operatorname{div}(V) = \operatorname{tr} \Lambda$ (theorem 6.1.3).

We can immediately observe that

$$\Lambda_p = I. \quad (6.6)$$

Indeed for $X_p \in T_p M$, let γ be the geodesic through p such that $\dot{\gamma}(0) = X_p$ we have

$$\Lambda_p(X_p) = \nabla_{X_p} V = \nabla_{\dot{\gamma}}(t\dot{\gamma})|_{t=0} = \dot{\gamma}(0) = X_p.$$

We fix $X_p \in T_p M$ and $\epsilon > 0$ such that the geodesic $\gamma :]-\epsilon, \epsilon[\rightarrow U, t \rightarrow \exp(tX_p)$ is well-defined and takes value in U . From now on we work along γ , that is on the pullback bundle $\gamma^* TM$ equipped with the pullback connection, and its tensor bundles (Note: in what follows we often denote a tensor field over M and its pullback along γ by the same symbol. So for example we write R instead of $\gamma^* R$).

Along γ we define the field of endomorphisms $\Pi := R(\dot{\gamma}, -)\dot{\gamma}$, so for $-\epsilon < t < \epsilon$ and $Y \in T_{\gamma(t)} M$, $\Pi_t Y = R_{\gamma(t)}(\dot{\gamma}(t), Y)\dot{\gamma}(t)$. The core equation relating Π and Λ is:

$$t\nabla_{\dot{\gamma}} \Lambda = \Lambda - \Lambda \circ \Lambda + t^2 \Pi \quad (6.7)$$

Proof of 6.7. Let Y be a vector field on U . Using (ii) in proposition 6.2.1 we get

$$t^2 \Pi(Y) = R(V, Y)V = \nabla_V \nabla_Y V - \nabla_Y \nabla_V V - \nabla_{[V, Y]} V. \quad (6.8)$$

Along γ we have

$$\nabla_V V = t\nabla_{\dot{\gamma}}(tX) = tX = V$$

(in fact since this is valid along any geodesic through p , we have $\nabla_V V = V$ on U) and 6.8 can be rewritten

$$t^2\Pi(Y) = \nabla_V \nabla_Y V - \nabla_Y V - \nabla_{[V,Y]}V. \quad (6.9)$$

On the other hand we can also compute

$$(\nabla_V \Lambda)(Y) = \nabla_V(\Lambda(Y)) - \Lambda(\nabla_V Y) = \nabla_V \nabla_Y V - \nabla_{\nabla_V Y} V \quad (6.10)$$

and

$$(\Lambda \circ \Lambda)(Y) = \nabla_{\nabla_Y V} V. \quad (6.11)$$

Combining 6.10 and 6.12 we get

$$\begin{aligned} (\nabla_V \Lambda)(Y) + (\Lambda \circ \Lambda)(Y) - \Lambda(Y) &= \nabla_V \nabla_Y V - \nabla_Y V - \nabla_{\nabla_V Y - \nabla_Y V} V \\ &= \nabla_V \nabla_Y V - \nabla_Y V - \nabla_{[V,Y]} V \end{aligned} \quad (6.12)$$

where we once again we use the fact that ∇ is without torsion. Comparing 6.9 and 6.12 we get

$$t^2\Pi(Y) = (\nabla_V \Lambda)(Y) + (\Lambda \circ \Lambda)(Y) - \Lambda(Y)$$

and finally obtain 6.7 after noticing

$$\nabla_V \Lambda = t\nabla_{\dot{\gamma}} \Lambda.$$

□

Reasoning inductively from 6.6 and 6.7 we will obtain an expression for the derivatives of $tr\Lambda = div(V)$.

Applying $\nabla_{\dot{\gamma}}$ to 6.7 we get

$$t\nabla_{\dot{\gamma}}^2 \Lambda + \nabla_{\dot{\gamma}} \Lambda = \nabla_{\dot{\gamma}} \Lambda - \nabla_{\dot{\gamma}} \Lambda \circ \Lambda - \Lambda \circ \nabla_{\dot{\gamma}} \Lambda + t^2 \nabla_{\dot{\gamma}} \Pi + 2t\Pi.$$

which combined with 6.6 implies

$$\nabla_{\dot{\gamma}} \Lambda|_0 = 0. \quad (6.13)$$

If we keep applying $\nabla_{\dot{\gamma}}$ to 6.7 repeatedly and use 6.6 as well as 6.13 we obtain for all $r \geq 2$

$$\nabla_{\dot{\gamma}}^r \Lambda|_0 = P^r|_0 \quad (6.14)$$

where the endomorphism fields P^r , $r = 2, 3, \dots$ are defined by the recurrence formula

$$(r+1)P^r = r(r-1)\nabla_{\dot{\gamma}}^{r-2}\Pi - \sum_{q=2}^{r-2} \binom{r}{q} P^q \circ P^{r-q} \quad (6.15)$$

which defines the Ledger curvature conditions. An alternative expression one can obtain for P^r , $r = 2, 3, \dots$ from this recurrence formula is

$$P^r = \sum c_{i_1 \dots i_k}^r \nabla_{\dot{\gamma}}^{i_1} \Pi \circ \dots \circ \nabla_{\dot{\gamma}}^{i_k} \Pi \quad (6.16)$$

where we sum over $1 \leq k \leq \lfloor r/2 \rfloor$, $i_j \geq 0$ such that $r = i_1 + \dots + i_k + 2k$. The constants $c_{i_1 \dots i_k}^r$ are defined recursively over the same range of indices by

$$c_{i_1}^r = \frac{r(r-1)}{r+1}$$

for $k = 1$, and for $k > 1$ by

$$c_{i_1 \dots i_k}^r = -\frac{1}{r+1} \sum_{j=1}^{k-1} \binom{r}{q_j} c_{i_1 \dots i_j}^{q_j} c_{i_{j+1} \dots i_k}^{r-q_j}$$

with $q_j = i_1 + \dots + i_j + 2j$. This second expression will be of use to us later. Since

$$\frac{d^r}{dt^r} (tr \Lambda \circ \gamma)(t) = (tr \nabla_{\dot{\gamma}}^r \Lambda)(t)$$

by putting together 6.14, proposition 6.2.2 and theorem 6.2.2 we finally obtain the main result of this section:

Theorem 6.2.3. *If the local geodesic symmetry $s_p : U \rightarrow U$ is divergence-preserving then the odd Ledger conditions are verified at p , by which we mean that*

$$tr P^r|_0 = 0$$

for $r \in 2\mathbb{N} + 1$ and all choices of $X_p \in T_p M$. Conversely if (M, ∇) is analytic, locally equiaffine and satisfies the odd Ledger conditions at p , then there exists a normal ball $\tilde{U}$ centered at p , contained in U and such that $s_p : \tilde{U} \rightarrow \tilde{U}$ is divergence-preserving.

Corollary 6.2.1. *Let ∇ be a torsionless affine connection on ∇ . If (M, ∇) has divergence-preserving local geodesic symmetries, then it satisfies the odd Ledger conditions (at every point of M). Conversely, if (M, ∇) is analytic, locally equiaffine and satisfies the odd Ledger conditions, then (M, ∇) has divergence-preserving local geodesic symmetries.*

Corollary 6.2.2. *Let ∇ be a torsionless equiaffine analytic connection over an analytic manifold M with parallel volume form ω . Then ω is symmetric if and only if (M, ∇) satisfies all odd Ledger conditions.*

Remark. We can label the Ledger conditions by (Lr) for $r \geq 2$ and say for example that a manifold M with an affine connection ∇ satisfies $(L22)$ at a point $p \in M$ if

$$tr P^{22}|_0 = 0$$

for all choices of $X_p \in T_p M$.

Using $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ and reasoning inductively from the definition $\Pi = R(\dot{\gamma}, -)\dot{\gamma}$ we get the following expressions for $\nabla_{\dot{\gamma}}^r \Pi$:

Proposition 6.2.3.

- (i) For $r \geq 0$, $\nabla_{\dot{\gamma}}^r \Pi = (\nabla_{\dot{\gamma}}^r R)(\dot{\gamma}, -)\dot{\gamma} = (\nabla^r R)(\dot{\gamma}, -; \dot{\gamma}, \dots, \dot{\gamma})\dot{\gamma}$.
- (ii) In particular for $r \geq 0$, $\nabla_{\dot{\gamma}}^r \Pi|_0 = (\nabla^r R)_p(X_p, -; X_p, \dots, X_p)X_p$.

Remark. The difference between the last two expressions in (i) is not merely cosmetic, $\nabla_{\dot{\gamma}}^r R$, $r = 0, 1, \dots$ are tensor fields of type $(1, 3)$ along γ defined by $\nabla_{\dot{\gamma}}^{r+1} R := \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}}^r R$ (and $\nabla_{\dot{\gamma}}^0 R := R \circ \gamma$) while $\nabla^r R$, $r = 0, 1, \dots$ are tensor fields of type $(1, 3 + r)$ on M defined by $\nabla^{r+1} R := \nabla \nabla^r R$ (and $\nabla^0 R := R$), and the last equality in (i) does not hold for an arbitrary curve on M (one that is not a geodesic).

By injecting (ii) into 6.16 (evaluated at 0) we can obtain an invariant version of the Ledger conditions. Notably, the invariant form of (L3) is particularly simple.

Theorem 6.2.4. *Let M be a manifold with an affine connection ∇ . Then (M, ∇) satisfies (L3) if and only if for all vector fields X on M ,*

$$(\nabla_X r)(X, X) = 0.$$

In particular, if ∇ is locally equiaffine (i.e. has symmetric Ricci curvature), then (M, ∇) satisfies (L3) if and only if it has cyclic parallel Ricci curvature.

Proof. Using the recurrence formula 6.15 for the operators P^r , we get $P^3 = \frac{3}{2} \nabla_{\dot{\gamma}} \Pi$. Combining this with proposition 6.2.3, we see that (L3) is equivalent to

$$0 = \text{tr} P_X = (\nabla_X r)(X, X)$$

for all $X \in \mathfrak{X}(M)$, where $P_X := (\nabla_X R)(X, -)X$ is the operator we introduced in theorem 3.3.6 (the necessary curvature condition for a tensor to be symmetric). On a locally equiaffine manifold (i.e. a an affinely connected manifold with symmetric Ricci curvature), this is equivalent to requiring that the Ricci tensor be cyclic parallel (see proposition 3.3.5). $\square$

Chapter 7

Preferred connections on symplectic surfaces

Here we turn our attention to symplectic surfaces (2-dimensional symplectic manifolds) equipped with symplectic connections that satisfy the (L3) condition, so-called *preferred* symplectic connections (a nomenclature specific to the symplectic setting introduced by Cahen). Our objective is to show through a family of counter-examples, constructed by Bourgeois and Cahen in their general study of preferred symplectic connections [6], that there exist symplectic connections of type S which are not locally symmetric, thereby showing that the most obvious generalization of the Sekigawa-Vanhecke theorem to the symplectic (i.e. not necessarily Kähler) setting does not hold. We will be able to leverage our work on manifolds with divergence-preserving symmetries. Indeed, on a surface, a symplectic form is a volume form, so that a symplectic connection is of type S if and only if it has divergence-preserving local symmetries.

7.1 Odd Ledger conditions

As seen earlier, symplectic connections of type S are necessarily preferred, making preferred connections a natural subject of study in the context of this thesis. What makes symplectic *surfaces* with preferred connections particularly interesting in view of 6.2 is the fact that they satisfy *all* Ledger conditions of odd order, not just (L3). Showing this is the object of this section, and of fundamental importance in this task is the fact that the curvature tensor of a symplectic connection on a symplectic surface is determined by its Ricci curvature tensor. This property characterizes the broader class of symplectic connections of Ricci-type, which were also studied quite thoroughly by Cahen (see [5] section 4 for a detailed account).

Theorem 7.1.1. *Let (M, ω) be a symplectic surface, and let ∇ be a symplectic connection on (M, ω) . Then*

$$\underline{R} = -\omega \otimes r$$

or equivalently

$$R = \omega \otimes \rho$$

where $\rho \in \Gamma(\text{End}(TM))$ is the Ricci endomorphism defined by $\omega(X, \rho Y) := r(X, Y) \forall X, Y \in \mathfrak{X}(M)$.

Proof. Since M is 2-dimensional, the symplectic form ω is a volume form on M and we have

$$\begin{aligned} r(Z, T)\omega(X, Y) &= \text{tr}(R(Z, -)T)\omega(X, Y) = \omega(R(Z, X)T, Y) + \omega(X, R(Z, Y)T) \\ &= \underline{R}(Z, X, T, Y) - \underline{R}(Z, Y, T, X) = \underline{R}(Z, X, Y, T) + \underline{R}(Y, Z, X, T) \\ &= -\underline{R}(X, Y, Z, T) \end{aligned}$$

where we use the symmetries of the symplectic curvature (theorem 4.2.4). In particular, Bianchi's first identity is used in the last equality. $\square$

Theorem 7.1.2. *Let (M, ω) be a symplectic surface equipped with a preferred connection ∇ . Then (M, ∇) verifies the odd Ledger conditions.*

Proof. Let p be a point of M , let $X_p \in T_p M$ and take $\epsilon > 0$ such that $\gamma :] - \epsilon, \epsilon[\rightarrow M, t \rightarrow \exp(tX_p)$ is well-defined.

Reproducing section 6.2 we define the endomorphism field $\Pi := R(\dot{\gamma}, -)\dot{\gamma}$ along γ , as well as the sequence of endomorphism fields P^r , $r = 2, 3, \dots$ by the recurrence formula 6.16 (or equivalently by formula 6.15). The core observation is the following: for $q \geq 1$ and $j \geq 0$

$$\nabla_{\dot{\gamma}}^j \Pi \circ \nabla_{\dot{\gamma}}^q \Pi = 0 \tag{7.1}$$

which we prove in four steps:

1. Using $R = \omega \otimes \rho$ (theorem 7.1.1), $\nabla \omega = 0$ and $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$, we get

$$\nabla_{\dot{\gamma}}^j \Pi = \omega(\dot{\gamma}, -)(\nabla_{\dot{\gamma}}^j \rho)\dot{\gamma} \tag{7.2}$$

for all $j \geq 0$.

2. By definition of a preferred connection, $(\nabla_{\dot{\gamma}} r)(\dot{\gamma}, \dot{\gamma}) = 0$. Combining this with $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$, we deduce

$$(\nabla_{\dot{\gamma}}^q r)(\dot{\gamma}, \dot{\gamma}) = 0 \tag{7.3}$$

for all $q \geq 1$.

3. Using $\nabla \omega = 0$ once again, and $r = \omega(-, \rho-)$ we obtain

$$\nabla_{\dot{\gamma}}^j r = \omega(-, (\nabla_{\dot{\gamma}}^j \rho)-) \tag{7.4}$$

for all $j \geq 0$.

4. Let $q \geq 1$ and $j \geq 0$. Putting together 7.2, 7.3 and 7.4 we calculate

$$\begin{aligned}\nabla_{\dot{\gamma}}^j \Pi \circ \nabla_{\dot{\gamma}}^q \Pi &= \omega(\dot{\gamma}, -) \nabla_{\dot{\gamma}}^j \Pi ((\nabla_{\dot{\gamma}}^q \rho) \dot{\gamma}) \\ &= \omega(\dot{\gamma}, -) \omega(\dot{\gamma}, (\nabla_{\dot{\gamma}}^q \rho) \dot{\gamma}) (\nabla_{\dot{\gamma}}^j \rho) \dot{\gamma} \\ &= \omega(\dot{\gamma}, -) \underbrace{(\nabla_{\dot{\gamma}}^q r)(\dot{\gamma}, \dot{\gamma})}_{=0} (\nabla_{\dot{\gamma}}^j \rho) \dot{\gamma} = 0\end{aligned}$$

Furthermore, for $q \geq 1$ we have

$$tr \nabla_{\dot{\gamma}}^q \Pi = \frac{d^{q-1}}{dt^{q-1}} tr \nabla_{\dot{\gamma}} \Pi = \frac{d^{q-1}}{dt^{q-1}} (\nabla_{\dot{\gamma}} r)(\dot{\gamma}, \dot{\gamma}) = 0 \quad (7.5)$$

We are now in a position to show that (M, ∇) satisfies the odd Ledger conditions. By definition, ∇ satisfies (L3) so we need only show that

$$tr P^r|_0 = 0$$

for $r \geq 5$. We start from formula 6.16, which we partially restate here for convenience

$$P^r = \sum c_{i_1 \dots i_k}^r \nabla_{\dot{\gamma}}^{i_1} \Pi \circ \dots \circ \nabla_{\dot{\gamma}}^{i_k} \Pi$$

where the summation is done over $1 \leq k$, $0 \leq i_j$ such that $2k + i_1 + \dots + i_j = r$. Fix r odd with $r \geq 5$, $k \geq 1$ and $i_j \geq 0$ with $r = i_1 + \dots + i_k + 2k$. We will show that

$$tr(\nabla_{\dot{\gamma}}^{i_1} \Pi \circ \dots \circ \nabla_{\dot{\gamma}}^{i_k} \Pi) = 0 \quad (7.6)$$

thus finishing the proof.

Case 1: if $k = 1$ then $i_1 = r - 2 \geq 1$ and 7.6 becomes

$$tr \nabla_{\dot{\gamma}}^{r-2} \Pi = 0$$

which is a consequence of 7.5.

Case 2: if $k \geq 2$, since r is odd we cannot have $r = 2k$ so there must exist $1 \leq j \leq k$ such that $i_j \geq 1$. By the cyclic property of the trace, we can assume without loss of generality that $j = 2$. By 7.1 we have $\nabla_{\dot{\gamma}}^{i_1} \Pi \circ \nabla_{\dot{\gamma}}^{i_2} \Pi = 0$ so that

$$tr(\nabla_{\dot{\gamma}}^{i_1} \Pi \circ \nabla_{\dot{\gamma}}^{i_2} \Pi \circ \dots \circ \nabla_{\dot{\gamma}}^{i_k} \Pi) = 0.$$

□

Corollary 7.1.1. *Let M be an analytic surface, let ω be a smooth symplectic form on M and let ∇ be a preferred symplectic connection on M which is also analytic. Then ∇ is of type S .*

Proof. ω is a parallel volume form on (M, ∇) , in particular (M, ∇) is equiaffine, and a symplectic connection is torsionless by definition so the result is a consequence of corollary 6.2.2 and theorem 7.1.2. $\square$

Remark. Notice that we did not require ω to be analytic, nor does our proof require such a hypothesis. In any case, as a parallel section for an analytic connection, ω is analytic.

7.2 A family of type S connections which are not locally symmetric

The family of (geodesically complete) non-locally-symmetric preferred symplectic connections on $\mathbb{R}^2$ constructed by Bourgeois and Cahen in section 11 of [6] constitutes a class of non-locally-symmetric symplectic connections of type S. This is an immediate consequence of corollary 7.1.1 once one observes that these connections are analytic for the standard analytic structure on $\mathbb{R}^2$, and observing this poses no difficulty thanks to the very explicit formulae derived by Bourgeois and Cahen.

These explicit formulae are obtained by introducing a function β and a 1-form u which essentially control the geometry of a symplectic surface with preferred connection, as summarized in the theorem below. The proofs of the different statements listed here can be found in [6] (mainly section 5). See also [5] section 3.1 for a coordinate-free formulation of some of the results in [6].

Theorem 7.2.1. *Let (M, ω) be a symplectic surface equipped with a preferred connection ∇ . Then*

(i) *There exists a 1-form u such that for all $X, Y, Z \in \mathfrak{X}(M)$ one has*

$$(\nabla_X r)(Y, Z) = \omega(Y, X)u(Z) + \omega(Z, X)u(Y).$$

(ii) *There exists a function β such that*

$$\nabla_X u = \beta \omega.$$

(iii) *Defining the vector field $\bar{u}$ by*

$$i(\bar{u})\omega = u.$$

There exist real numbers A and B such that

$$\begin{aligned} r(\bar{u}, \bar{u}) &= \beta^2 + B \\ \frac{1}{4} \text{tr}[\rho^2] &= \beta + A. \end{aligned}$$

(iv) Assume $d\beta \neq 0$ and let $U = \{p \in M \mid r_p(\bar{u}, \bar{u}) \neq 0\}$. Then the preferred connection is given on U by

$$\begin{aligned}\nabla_{\bar{u}}\bar{u} &= -\beta\bar{u} \\ \nabla_{\bar{u}}X_\beta &= -\beta X_\beta \\ \nabla_{X_\beta}X_\beta &= (\beta^2 + 2A\beta - B)\bar{u}.\end{aligned}\tag{7.7}$$

Conversely, given a 1-form u and a non-constant function β such that $u(X_\beta) = \beta^2 + B$ for some real number B , equation 7.7 defines a preferred symplectic connection over $U := \{p \in M \mid \beta^2 \neq -B\}$.

(v) The symplectic connection ∇ is locally symmetric if and only if $\beta = 0$.

Let us now proceed with the construction of the examples. To avoid confusion between coordinate indices and exponents we denote the coordinates of $\mathbb{R}^2$ by (x, y) . We equip $\mathbb{R}^2$ with its standard constant symplectic form $\omega_0 = dx \wedge dy$. We choose real numbers A and B with $B > 0$ and endow $(\mathbb{R}^2, \omega_0)$ with the preferred symplectic connection induced by the function $\beta := y$ and the 1-form $u := (y^2 + B)dx = (\beta^2 + B)dx$, as described in (iv) of theorem 7.2.1.

We easily compute $X_\beta = \partial_x$ and $\bar{u} = -(y^2 + B)\partial_y$. From this we verify that

$$u(X_\beta) = y^2 + B = \beta^2 + B$$

ensuring that β and u do indeed induce a preferred connection ∇ on $\mathbb{R}^2$. Since $B > 0$, $\beta^2 + B$ never vanishes, so ∇ is defined on the whole of $\mathbb{R}^2$. In accordance with theorem 7.2.1 (iv) the Christoffel symbols of ∇ are given by:

$$\begin{aligned}\nabla_{\partial_x}\partial_x &= -(y^2 + B)(x^2 + 2Ay - B)\partial_y \\ \nabla_{\partial_x}\partial_y &= \nabla_{\partial_y}\partial_x = \frac{y}{y^2 + B}\partial_x \\ \nabla_{\partial_y}\partial_y &= -\frac{y}{y^2 + B}\partial_y.\end{aligned}$$

Clearly the Christoffel symbols are analytic, so ∇ is analytic (theorem B.0.1, appendix B). Finally, since $\beta = y$ is non-zero (with nowhere vanishing differential) by theorem 7.2.1 (v) (M, ∇) is not locally symmetric.

Conclusion

Our main objective was to determine whether non-locally-symmetric connections of type S exist. We answered the question in the positive by providing a family of examples, an achievement we owe in (very) great part to the work of D'Atri and Nickerson [8] on D'Atri spaces, and of Bourgeois and Cahen [6] on preferred symplectic connections.

On the journey that lead us to that point, we also showed that connections of type S are preferred, gave a proof of the Sekigawa-Vanhecke theorem and used said theorem to obtain examples of preferred symplectic connections that are not of type S. We introduced the affine equivalents to D'Atri spaces in the form of torsionless affine connections whose geodesic symmetries preserve the induced connection on the determinant bundle. The latter were shown to satisfy all odd Ledger conditions, and in turn the odd Ledger conditions were shown to be sufficient in the analytic, locally equiaffine case. To achieve this we introduced a new notion of divergence on a manifold, equivalent to a choice of connection on the determinant bundle.

Our work might be extended in a number of ways. The first point that comes to mind is that we might attempt to determine whether all preferred connections over symplectic surfaces are of type S. The fact that such connections satisfy all odd Ledger conditions (leaving only analyticity as an obstacle) gives credence to an affirmative answer. The work of Bourgeois and Cahen, who have more or less fully elucidated the geometry of such connections, also renders this problem very tractable.

Another problem we could tackle is that of determining which symplectic connections of Ricci-type (whose curvature is determined by the Ricci tensor - see [5]) are of type S. Indeed, the fact that the curvature of a symplectic connection over a symplectic surface is fully determined by its Ricci curvature played a key role in our computations.

One last obvious question is that of finding sufficient conditions for a symplectic connection to be of type S in dimension greater than 2. One of the main limitations of our derivation is the fact that the Ledger conditions arise for volume forms, which coincide with symplectic forms only in dimension 2. We could attempt to proceed in a manner similar to the one we used here: we could try to compute the power expansion of the symplectic form in normal coordinates, and hope that this expansion admits a nice description in the form of a reasonably simple recurrence formula. The less certain, but perhaps more exciting path would be to look for a sufficient condition that does not rely on this method of radial power expansion.

Appendix A

Distributions, foliations and the Frobenius theorem

In this appendix we state the theorem of Frobenius, which establishes an equivalence between foliations and involutive distributions, and recall the definitions necessary to understand its statement. Our reference is [19] chapter 19 'Distributions and Foliations'.

Foliations arise whenever a smooth manifold is subdivided into a family of submanifolds (which are called **leaves**). One can think of the level sets of a function, the orbits of a group action or the integral curves of a vector field. They are an essential tool in differential geometry, and extremely interesting structures in their own right. They can be defined through the notion of **flat charts**.

Definition A.0.1 (Flat charts). Let M be a smooth manifold of dimension m and let $\mathcal{F}$ be a family of k -dimensional submanifolds of M . A chart (U, ϕ) on M is said to be **flat** for $\mathcal{F}$ if $\phi(U)$ is a cube in $\mathbb{R}^m$ and if each submanifold in $\mathcal{F}$ intersects U in either the empty set or a countable union of k -dimensional slices of the form $x^{k+1} = c^{k+1}, \dots, x^m = c^m$ (for a family of constants $c^{k+1}, \dots, c^m$).

Definition A.0.2 (Foliations). A **foliation of dimension k of M** is a partition $\mathcal{F}$ of M into a collection of connected, immersed k -dimensional submanifolds of M such that any point of p is contained in the domain of a flat chart for $\mathcal{F}$. We call elements of $\mathcal{F}$ the **leaves** of the foliation.

One of the basic ideas behind the Frobenius theorem is that a foliation can be recovered from the collection of spaces tangent to its leaves, which form a 'smooth' family of subspaces in the tangent bundle, in other words they form a **distribution**:

Definition A.0.3 (Distributions). Let M be a smooth manifold. A **distribution of rank k on M** is a rank k subbundle D of TM . We denote by $\Gamma(D)$ the space of sections of D .

By the subbundle criterion, a distribution of rank k on M is the data for each $p \in M$ of a linear subspace $D_p \subseteq T_p M$ of dimension k , such that for any point $p \in M$, there exists a family of linearly independent smooth vector fields $E_1, \dots, E_k$ defined on an open neighborhood U of p such that for all $q \in U$, D_q is spanned by $E_1(q), \dots, E_k(q)$.

An element of $\Gamma(D)$ is then simply a smooth vector field X on M such that X_p belongs to D_p for all $p \in M$.

Example A.0.1. Let $\mathcal{F}$ be a foliation of dimension k of M . The collection of tangent spaces of the leaves of $\mathcal{F}$ is a rank k distribution called the **distribution tangent to $\mathcal{F}$** .

The Frobenius theorem answers the question of which distributions are induced by foliations. If a distribution is to arise as the distribution tangent to a foliation, there must at least exist submanifolds to which the distribution is tangent (so-called **integral manifolds**) at every point (locally). This leads to the notion of **integrable distributions**.

Definition A.0.4 (Integrable distributions). Let D be a rank k distribution on M . An **integral manifold of D** is an immersed k -dimensional submanifold $N \subseteq M$ such that for all $p \in N$, $T_p N = D_p$. A distribution is said to be **integrable** if any point $p \in M$ is contained in an integral manifold of D .

One (relatively) obvious necessary condition for a distribution to be integrable is **involutivity**. Given a submanifold $N \subseteq M$, a vector field $X \in \mathfrak{X}(M)$ is said to be tangent to N if $X_p \in T_p N$ for all $p \in N$. In that case $X|_N$ (the restriction of X to N) can be identified (locally) with a vector field on N . In particular, this implies that for any pair of vector fields $X, Y \in \mathfrak{X}(M)$ tangent to N , $[X, Y]$ is tangent to N , leading to the following:

Definition A.0.5 (Involutive distributions). A distribution D on M is said to be **involutive** if $\Gamma(D)$ is a Lie subalgebra of $\mathfrak{X}(M)$ for the usual Lie derivative of vector fields, meaning that for any $X, Y \in \Gamma(D)$, $[X, Y]$ is a section of D .

Theorem A.0.1. *An integrable distribution is necessarily involutive.*

So far, through a series of reasonably obvious steps we have arrived at the following scheme:

$$\text{Foliations} \implies \text{Integrable distributions} \subseteq \text{Involutive distributions}$$

Foliations induce integrable distributions, and in turn integrable distributions are necessarily involutive. The (very much *not* obvious) Frobenius theorem tells us that the correspondence is one-to-one at every step.

Theorem A.0.2 (Frobenius theorem). *Let D be an involutive distribution on M . Then there exists a unique foliation $\mathcal{F}$ of M whose tangent distribution is D . The leaves of $\mathcal{F}$ are the maximal connected integral manifolds of D .*

Remark. Implicit in that statement is the fact that, given an involutive distribution D on M , any point of M is contained in a unique maximal connected integral manifold of D .

Appendix B

Analytic connections

In this appendix we define analytic connections on analytic manifolds, and present the properties of analytic connections that are relevant to our work. Our references are [24] and [12].

In perfect analogy with complex and smooth manifolds, analytic manifolds are constructed by providing an atlas whose transition maps are, as their name suggests, analytic.

So let us start with the definition of analytic functions.

Definition B.0.1. Let U be an open subset of $\mathbb{R}^m$. A map

$$f : U \rightarrow \mathbb{R}$$

is said to be **analytic** if for any $x_0 \in U$, there exists a power series

$$\sum_{\alpha \in \mathbb{N}^m} a_\alpha (x - x_0)^\alpha$$

which converges absolutely on a small ball $B(x_0, \epsilon)$ (for some $\epsilon > 0$) centered at x_0 and coincides with f on that ball. A map

$$f : U \rightarrow \mathbb{R}^n$$

is said to be analytic if each of its components are analytic in the above sense.

Proposition B.0.1. *Sums, products, derivatives and compositions of analytic functions are analytic.*

From this the definition of analytic manifolds flows naturally.

Definition B.0.2. Let M be a smooth manifold. An analytic structure on M is a maximal analytic atlas, that is an atlas whose transition maps are analytic, and which is maximal for this properties. Equipped with such an analytic structure, M is called

an analytic manifold, and the charts belonging to the maximal analytic atlas are called analytic charts.

Example B.0.1.

- (i) $\mathbb{R}^m$ possesses a standard analytic structure, the maximal analytic atlas consists of all diffeomorphisms

$$\phi : U \subseteq \mathbb{R}^m \rightarrow V \subseteq \mathbb{R}^m$$

which are analytic in the sense of definition B.0.1. By extension, any finite-dimensional real vector space V admits a canonical analytic structure.

- (ii) Since any holomorphic map

$$f : U \subseteq \mathbb{C}^m \rightarrow V \subseteq \mathbb{C}^n$$

is analytic, any complex manifold admits a canonical analytic structure.

We define analytic maps between analytic manifolds in the obvious way.

Definition B.0.3. Let M and N be two analytic manifolds (of respective dimensions m and n). A map

$$f : M \rightarrow N$$

is said to be analytic if for any analytic charts $\phi : U \subseteq M \rightarrow \tilde{U} \subseteq \mathbb{R}^m$ and $\psi : V \subseteq N \rightarrow \tilde{V} \subseteq \mathbb{R}^n$ such that $f(U) \subseteq V$, the map

$$\psi \circ f \circ \phi^{-1} : \tilde{U} \subseteq \mathbb{R}^m \rightarrow \tilde{V} \subseteq \mathbb{R}^n$$

is analytic. In particular an analytic map is always smooth.

Remark. By proposition B.0.1, analytic maps between analytic manifolds are stable by composition.

Given an open subset U of an analytic manifold M , we denote the space of analytic maps

$$f : U \rightarrow \mathbb{R}$$

by $C^\omega(U)$ (where $\mathbb{R}$ is endowed with its standard analytic structure). By proposition B.0.1, $C^\omega(U)$ becomes an $\mathbb{R}$ -algebra (in particular a ring) when equipped with the obvious sum and product(s).

Remark. To give a more explicit characterization: a map $f : U \rightarrow \mathbb{R}$ is analytic if and only if for any analytic chart $\phi : V \rightarrow \tilde{V} \subseteq \mathbb{R}^m$ with $V \subseteq U$, $f \circ \phi^{-1} : \tilde{V} \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ is analytic in the sense of definition B.0.1.

Most, if not all, objects from *smooth* geometry have their analytic counterpart. To define these objects, two equivalent methods (at least) are available to us:

1. By way of their action on previously defined analytic objects (such as elements of $C^\omega(U)$), by requiring that they preserve the analyticity of the objects they act upon.
2. Through their components in analytic charts, by requiring that these components be analytic.

Here we treat the case of vector fields and affine connections, as they are the objects relevant to our work.

Definition B.0.4. A vector field X defined on an open subset U of an analytic manifold M is analytic if for any open subset $V \subseteq U$ and any $f \in C^\omega(V)$, Xf is analytic. We denote the set of such vector fields by $\mathfrak{X}^\omega(U)$.

Definition B.0.5. An affine connection ∇ on an analytic manifold M is analytic if for any open $U \subseteq M$, and any $X, Y \in \mathfrak{X}^\omega(U)$, $\nabla_X Y$ is analytic.

Theorem B.0.1. *A vector field on an analytic manifold M is analytic if and only if its components in any analytic chart are analytic. Similarly, an affine connection on M is analytic if and only if its Christoffel symbols in any analytic chart are analytic.*

Proof. Let X is an analytic vector field over U , and let $(V, x^1, \dots, x^m)$ be an analytic chart with $V \subseteq U$. Then X is uniquely written $X = \sum_{i=1}^m X^i \partial_i$ with $X^i = Xx^i$. For each $1 \leq i \leq m$, the map $x^i : U \rightarrow \mathbb{R}$ is analytic so $X^i = Xx^i$ is analytic.

Conversely, let X be a vector field over U whose components in any analytic chart are analytic, and let f be an analytic function over an open subset $V \subseteq U$. Then for any $p \in V$, we can find an analytic chart $(W, x^1, \dots, x^m)$ defined on an open neighborhood W of p contained in V , and X takes the form $X = \sum_{i=1}^m X^i \partial_i$ over W , with X^i analytic for every $1 \leq i \leq m$ by hypothesis. Consequently

$$(Xf)|_W = \sum_{i=1}^m X^i \partial_i f$$

is analytic by B.0.1. Since this holds for every $p \in V$, Xf is analytic.

The characterization of an analytic connection in terms of its Christoffel symbols is proven in a similar manner, using the formulae

$$\begin{aligned} \nabla_{\partial_i} \partial_j &= \sum_{k=1}^m \Gamma_{ij}^k \partial_k \\ \nabla_X Y &= \sum_{k=1}^m \left(\sum_{i=1}^m X^i \partial_i Y^k + \sum_{i,j=1}^m \Gamma_{ij}^k X^i Y^j \right) \partial_k \end{aligned}$$

together with the characterization of analytic vector fields we just established (notice that the coordinate vector fields ∂_i of an analytic chart $(U, x^1, \dots, x^m)$ are analytic). $\square$

The basic constructions of differential geometry (flows in particular) are proven to be well-defined and smooth by relying on existence, uniqueness and *regularity* theorems in the theory of ODE's, and as such, these theorems play a key role in smooth geometry. Thankfully, as the result below shows, these theorems extend to the analytic case which means that, in that respect at least, the transition from smooth to analytic geometry occurs without pain.

Theorem B.0.2 ([12] appendix 1 'Ordinary linear differential equations'). *Let $f(t, y, s)$ be a family of n functions defined in $|t| < \delta$ and $(y, s) \in D$, where D is an open subset of $\mathbb{R}^{n+m}$. If $f(t, y, s)$ is continuous in t and differentiable of class C^1 in y , then there exists a unique family $\phi(t, \eta, s)$ of n functions defined in $|t| < \delta'$ and $(\eta, s) \in D'$, where $0 < \delta' < \delta$ and D' is an open subset of D , such that*

- (i) $\phi(t, \eta, s)$ is differentiable of class C^1 in t and η .
- (ii) $\frac{\partial \phi}{\partial t}(t, \eta, s) = f(t, \phi(t, \eta, s), s)$.
- (iii) $\phi(0, \eta, s) = \eta$.

If $f(t, y, s)$ is differentiable of class C^p , $0 \leq p \leq \omega$, in t and of class C^q , $1 \leq q \leq \omega$ in y and s , then $\phi(t, \eta, s)$ is differentiable of class C^{p+1} in t and of class C^q in η and s .

Proof. See [14] p. 42. $\square$

Most importantly for us, we can apply the previous theorem to the ODE used to define geodesics

$$0 = \ddot{\gamma}^k + \sum_{i,j=1}^m (\Gamma_{ij}^k \circ \gamma) \dot{\gamma}^i \dot{\gamma}^j, k = 1, \dots, m$$

(see section 3.3) whose coefficients are analytic when the connection and chart are analytic, and combine it with the inverse mapping theorem for real analytic functions (see [14] p. 47) to obtain that the normal charts of an analytic connection are analytic, as we would expect.

Corollary B.0.1. *Let ∇ be an analytic connection on an analytic manifold M . Then for any $p \in M$, and any normal neighborhood U of p , the diffeomorphism*

$$\exp : V \subseteq T_p M \rightarrow U$$

is analytic (where $T_p M$ is equipped with its standard analytic structure as a finite-dimensional real vector space). In particular a normal coordinate system $(U, x^1, \dots, x^m)$ centered at p is an analytic chart on M , and the Christoffel symbols of ∇ in such a chart are analytic.

Bibliography

- [1] Mélanie Bertelson, Michel Cahen, and Simone Gutt. “Equivalence of star products”. In: vol. 14. 1A. Geometry and physics. 1997, A93–A107.
- [2] Mélanie Bertelson, Michel Cahen, and Simone Gutt. “Equivalence of star products”. In: vol. 14. 1A. Geometry and physics. 1997, A93–A107.
- [3] Arthur L. Besse. *Einstein manifolds*. Classics in Mathematics. Reprint of the 1987 edition. Springer-Verlag, Berlin, 2008, pp. xii+516. ISBN: 978-3-540-74120-6.
- [4] Pierre Bieliavsky. “Espaces Symétriques Symplectiques”. PhD thesis. Université Libre de Bruxelles, 1995.
- [5] Pierre Bieliavsky et al. “Symplectic connections”. In: *Int. J. Geom. Methods Mod. Phys.* 3.3 (2006), pp. 375–420. ISSN: 0219-8878,1793-6977.
- [6] F. Bourgeois and M. Cahen. “A variational principle for symplectic connections”. In: *J. Geom. Phys.* 30.3 (1999), pp. 233–265. ISSN: 0393-0440,1879-1662.
- [7] Ana Cannas da Silva. *Lectures on symplectic geometry*. Vol. 1764. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2001, pp. xii+217. ISBN: 3-540-42195-5.
- [8] J. E. D’Atri and H. K. Nickerson. “Divergence-preserving geodesic symmetries”. In: *J. Differential Geometry* 3 (1969), pp. 467–476. ISSN: 0022-040X,1945-743X.
- [9] J. E. D’Atri and H. K. Nickerson. “Geodesic symmetries in spaces with special curvature tensors”. In: *J. Differential Geometry* 9 (1974), pp. 251–262. ISSN: 0022-040X,1945-743X.
- [10] S. K. Donaldson and P. B. Kronheimer. *The geometry of four-manifolds*. Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1990, pp. x+440. ISBN: 0-19-853553-8.
- [11] Alfred Gray. “Riemannian manifolds with geodesic symmetries of order 3”. In: *J. Differential Geometry* 7 (1972), pp. 343–369. ISSN: 0022-040X,1945-743X.
- [12] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol. I*. Wiley Classics Library. Reprint of the 1963 original, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1996, pp. xii+329. ISBN: 0-471-15733-3.
- [13] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol. II*. Wiley Classics Library. Reprint of the 1969 original, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1996, pp. xvi+468. ISBN: 0-471-15732-5.

- [14] Steven G. Krantz and Harold R. Parks. *A primer of real analytic functions*. Vol. 4. Basler Lehrbücher [Basel Textbooks]. Birkhäuser Verlag, Basel, 1992, pp. x+184. ISBN: 3-7643-2768-5.
- [15] Serge Lang. *Fundamentals of differential geometry*. Vol. 191. Graduate Texts in Mathematics. Springer-Verlag, New York, 1999, pp. xviii+535. ISBN: 0-387-98593-X.
- [16] A. J. Ledger. “Harmonic Riemannian Spaces”. PhD thesis. University of Durham, 1954.
- [17] Jeffrey M. Lee. *Manifolds and differential geometry*. Vol. 107. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2009, pp. xiv+671. ISBN: 978-0-8218-4815-9.
- [18] John M. Lee. *Introduction to Riemannian manifolds*. Second. Vol. 176. Graduate Texts in Mathematics. Springer, Cham, 2018, pp. xiii+437. ISBN: 978-3-319-91754-2; 978-3-319-91755-9.
- [19] John M. Lee. *Introduction to smooth manifolds*. Vol. 218. Graduate Texts in Mathematics. Springer-Verlag, New York, 2003, pp. xviii+628. ISBN: 0-387-95495-3.
- [20] Krzysztof Maurin. *The Riemann legacy*. Vol. 417. Mathematics and its Applications. Riemannian ideas in mathematics and physics, Translated from the Polish by Jerzy Kowalski-Glikman. Kluwer Academic Publishers Group, Dordrecht, 1997, pp. xxii+717. ISBN: 0-7923-4636-X.
- [21] Dusa McDuff and Dietmar Salamon. *Introduction to symplectic topology*. Third. Oxford Graduate Texts in Mathematics. Oxford University Press, Oxford, 2017, pp. xi+623. ISBN: 978-0-19-879490-5; 978-0-19-879489-9.
- [22] Eckhard Meinrenken. *Symplectic Geometry*. 1998-2000.
- [23] Pierre Bieliavsky Mélanie Bertelson. “Affine connections and symmetry jets”. In: *Transactions of the American Mathematical Society* (2011).
- [24] D. J. Saunders. *The geometry of jet bundles*. Vol. 142. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1989, pp. viii+293. ISBN: 0-521-36948-7.
- [25] D. J. Saunders. *The geometry of jet bundles*. Vol. 142. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1989, pp. viii+293. ISBN: 0-521-36948-7.
- [26] Loring W. Tu. *Differential geometry*. Vol. 275. Graduate Texts in Mathematics. Connections, curvature, and characteristic classes. Springer, Cham, 2017, pp. xvi+346. ISBN: 978-3-319-55082-4; 978-3-319-55084-8.
- [27] Frank W. Warner. *Foundations of differentiable manifolds and Lie groups*. Vol. 94. Graduate Texts in Mathematics. Corrected reprint of the 1971 edition. Springer-Verlag, New York-Berlin, 1983, pp. ix+272. ISBN: 0-387-90894-3.

