Department of Mathematics

Torsion and Non-Metricity: An Introduction to Metric-Affine Gravity

Yusuf Emre Baran

Supervisor: Yaghoub Heydarzade

Course: Math 291

Semester: Summer 2025

Contents

1 Introduction **2**

2 Scope and Limitations **2**

3 Connections **3**

3.1 The Directional Derivative in $\mathbb{R}^n$ 4

3.2 Further Properties of $\overline{\nabla}$ 6

3.2.1 Torsion 7

3.2.2 Metric-Compatibility 8

3.2.3 Curvature 9

3.3 Connections on a Vector Bundle 10

3.4 Connections on a Smooth Manifold 13

3.5 Covariant Derivative Along Curves 15

3.6 Curvature 17

3.7 Torsion 20

3.7.1 Geometrical Interpretation of Torsion 22

3.8 Metric-Compatibility 23

3.9 Levi-Civita Connection 26

4 Connection, Torsion, Curvature and Non-Metricity Forms **27**

4.1 Connection 1-Forms 28

4.2 Torsion 2-Forms 28

4.3 Curvature 2-Forms 31

4.4 Non-Metricity 1-Forms 32

5 General Relativity and Metric-Affine Gravity **32**

References **35**

1 Introduction

While the theory of general relativity (GR) has proven itself by correctly describing gravitational phenomena, it faces both theoretical and observational challenges, such as the quantization of gravity and the cosmological constant problem [15, 11], that motivate extensions of the theory. Metric-Affine Gravity (MAG) is one such generalization, relaxing the usual assumption that the connection depends on the metric. In MAG, the connection is promoted to an independent dynamical field, leading to two important geometric quantities: torsion and non-metricity.

This project aims to provide an introduction to the mathematical framework of MAG, with a focus on connections, torsion, and non-metricity tensors with their associated differential forms. Familiarity with smooth manifolds is assumed.[1][2] A basic understanding of vector bundles, Lie brackets, and metrics is also needed. To maintain the flow, these background topics have been omitted. See [8, 12, 13] for a brief treatment of these topics.

The central object of study in this project is connections on smooth manifolds. Accordingly, the discussion begins with the notion of directional derivative on $\mathbb{R}^n$ and its special properties. This is followed by an introduction to connections on vector bundles, with the primary focus quickly restricted to connections on tangent bundles. After developing the necessary tools, geometric interpretations of torsion and non-metricity are developed. We then show that information about torsion and non-metricity can be encoded in differential forms. The project concludes with a concise comparison of the field equations of GR and MAG.

Throughout this report, we mostly follow [8, 9].

2 Scope and Limitations

The initial aim of this project was to derive the general form of the field equations of MAG. The usual recipe for deriving the field equations of a field theory is to extremize a quantity known as the action, which requires tools from the calculus of variations. In the "classical" calculus of variations, one seeks a function that extremizes an integral subject

[1] Throughout this work, any object possessing a notion of smoothness will be assumed to be smooth, regardless of whether smoothness is mentioned explicitly.

[2] A smooth manifold will always refer to a smooth manifold without boundary.

to certain constraints. The main idea is to define a notion of (functional) derivative measuring the integral's rate of change with respect to functions. Then, similar to how one solves optimization problems in calculus, setting the derivative of the functional to zero gives the critical points, which are now functions.

The action in field theories is given by an integral of a function called the Lagrangian, which can be of ever-increasing complexity. In the standard formulation of general relativity, one seeks a metric that extremizes the integral of the curvature scalar. While the problem is easy to state, giving a precise and mathematically acceptable solution is not at all straightforward. The main idea is again to define a (functional) derivative measuring the integral's rate of change with respect to the metric. However, making these ideas mathematically precise is considerably more involved and requires further prerequisites. Additionally, the techniques needed to develop the field equations of MAG require even more advanced background and thus exceeded the time available for this project.

For the reasons outlined above, and to avoid treating them superficially or imprecisely, functional derivatives have been excluded from this project. As a result, the field equations of MAG, which are direct applications of functional derivatives, are not derived but taken as given to compare them with the field equations of GR. It should also be mentioned that although the title mentions MAG, the focus of the project is more on the mathematical structures of GR and MAG themselves, rather than on their physical interpretations. Consequently, no attempt is made to provide the reader with a physical intuition of these theories, and such interpretations are reserved for a potential follow-up project.

3 Connections

After studying the theory of smooth manifolds, a natural next step is to develop the notion of differentiation. However, differentiation on manifolds presents itself with some crucial problems that were not apparent in $\mathbb{R}^n$. In particular, one faces problems in adding a vector to a point and in comparing two vectors at different points.

To illustrate these challenges, let X be a vector field on $\mathbb{R}^n$, v a vector at $p \in \mathbb{R}^n$. Then, the directional derivative of X at p in the direction of v, denoted by $\nabla_v X$ is defined as [6]

$$\overline{\nabla}_v X := \lim_{t\to 0} \frac{X(p+tv) - X(p)}{t}. \tag{3.1}$$

Looking at the above expression, to take a directional derivative, one first adds tv to the point p. Then, one subtracts the vector $X(p)$ at the point p from the vector $X(p+tv)$ at the point $p+tv$. Finally, one divides by t and takes the limit. We refer to the subtraction operation here when we mean the comparison of two vectors at different points.

The first problem with this definition is that vectors cannot be added to points. One gets away with it by naturally identifying points of $\mathbb{R}^n$ with arrows whose tail is at the origin and the tip is at the given point. In other words, we identify $\mathbb{R}^n$ with its tangent spaces via translation. The second problem is that the vectors $X(p+tv)$ and $X(p)$ are at different points, and tangent spaces at distinct points are disjoint. Hence, the subtraction does not make sense, and one cannot compare two vectors at different points. This issue is also avoided in $\mathbb{R}^n$ because tangent spaces at each point are canonically isomorphic to $\mathbb{R}^n$, so one generally does not make a distinction between $\mathbb{R}^n$ and $T_p\mathbb{R}^n$. Thus, both problems are circumvented in $\mathbb{R}^n$ by the natural identification $\mathbb{R}^n \cong T_p\mathbb{R}^n$.

Despite these problems, there is a way to make sense of differentiation on manifolds, called a connection. Before introducing connections, we inspect the properties of the directional derivative defined by (3.1) further.

3.1 The Directional Derivative in $\mathbb{R}^n$

Let us unpack (3.1) using multivariable calculus on $\mathbb{R}^n$.

$$\begin{aligned}
\overline{\nabla}_v X = \lim_{t\to 0} \frac{X(p+tv) - X(p)}{t} &= \left(\lim_{t\to 0} \frac{X^i(p+tv) - X^i(p)}{t} \right) \partial_i\big|_p \\
&= \left(\frac{d}{dt}\bigg|_{t=0} X^i(p+tv) \right) \partial_i\big|_p \\
&= \left(\frac{\partial X^i}{\partial x^j}\bigg|_p \cdot \frac{dx^j}{dt}\bigg|_{t=0} \right) \partial_i\big|_p \quad \text{(chain rule)} \\
&= \left(\frac{\partial X^i}{\partial x^j}\bigg|_p \cdot v^j \right) \partial_i\big|_p \\
&= \left(\left(v^j \frac{\partial}{\partial x^j}\bigg|_p \right) X^i \right) \partial_i\big|_p.
\end{aligned}$$

More succinctly, we have

$$\overline{\nabla}_v X = \big(v(X^i)\big)\, \partial_i\big|_p. \tag{3.2}$$

Motivated by this calculation, we generalize the directional derivative in $\mathbb{R}^n$ to the derivative of a vector field with respect to another vector field, which we also call a directional derivative, as follows. For $X, Y \in \mathfrak{X}(\mathbb{R}^n)$, define a new vector field $\overline{\nabla}_X Y$ on $\mathbb{R}^n$ by

$$(\overline{\nabla}_X Y)_p = \overline{\nabla}_{X_p} Y \quad \text{for all } p \in \mathbb{R}^n. \tag{3.3}$$

By (3.2), one has

$$(\overline{\nabla}_X Y)_p = \overline{\nabla}_{X_p} Y = (X_p(Y^i))\, \partial_i\big|_p \tag{3.4}$$

Hence, as p runs over $\mathbb{R}^n$, we obtain

$$\overline{\nabla}_X Y = X(Y^i)\, \partial_i \tag{3.5}$$

Remark 3.1. The resemblance of the above equation to the action of a vector field on a function suggests that we make the definition $\overline{\nabla}_X f \coloneqq X(f)$. In fact, $\overline{\nabla}_X Y$ can be seen as the directional derivative of components of Y.

By (3.5), if $X, Y \in \mathfrak{X}(\mathbb{R}^n)$, then $\overline{\nabla}_X Y \in \mathfrak{X}(\mathbb{R}^n)$. Thus, the directional derivative on $\mathbb{R}^n$ induces a map

$$\overline{\nabla} : \mathfrak{X}(\mathbb{R}^n) \times \mathfrak{X}(\mathbb{R}^n) \to \mathfrak{X}(\mathbb{R}^n),$$

written as $\overline{\nabla}_X Y$ instead of $\overline{\nabla}(X, Y)$. We now list some properties of this new directional derivative that will guide us in formulating the axioms of a connection.

Proposition 3.2. Let $X, Y \in \mathfrak{X}(\mathbb{R}^n)$. Then, the following properties hold:
(i) $\overline{\nabla}_X Y$ is $C^\infty(\mathbb{R}^n)$-linear in X.
(ii) $\overline{\nabla}_X Y$ is $\mathbb{R}$-linear in Y.
(iii) For $f \in C^\infty(\mathbb{R}^n)$,

$$\overline{\nabla}_X (fY) = (Xf)Y + f\overline{\nabla}_X Y.$$

Proof. (i) Let $f \in C^\infty(\mathbb{R}^n)$ and $p \in \mathbb{R}^n$. Then,

$$(\overline{\nabla}_{fX} Y)_p = \overline{\nabla}_{(fX)(p)} Y = \overline{\nabla}_{f(p)X_p} Y = f(p)\overline{\nabla}_{X_p} Y = (f\overline{\nabla}_X Y)_p,$$

where we used the $\mathbb{R}$-linearity of $\overline{\nabla}_{X_p}Y$ in X_p, established by (3.4), in the second-to-last equality. Next, let $Z \in \mathfrak{X}(\mathbb{R}^n)$. Then,

$$(\overline{\nabla}_{X+Z}Y)_p = \overline{\nabla}_{(X+Z)(p)}Y = \overline{\nabla}_{X_p+Z_p}Y = \overline{\nabla}_{X_p}Y + \overline{\nabla}_{Z_p}Y = (\overline{\nabla}_X Y + \overline{\nabla}_Z Y)_p,$$

where we again used (3.4) in the second-to-last equality. Because $p \in \mathbb{R}^n$ was arbitrary, this completes the proof.

(ii) Let $a_1, a_2 \in \mathbb{R}$ and let $Z \in \mathfrak{X}(\mathbb{R}^n)$. For $p \in \mathbb{R}^n$, (3.4) yields

$$\begin{aligned}(\overline{\nabla}_X(a_1Y + a_2Z))_p = X_p((a_1Y + a_2Z)^i)\,\partial_i\big|_p &= (X_p((a_1Y)^i) + X_p((a_2Z)^i))\,\partial_i\big|_p \\ &= (X_p(a_1Y^i) + X_p(a_2Z^i))\,\partial_i\big|_p \\ &= (a_1X_p(Y^i) + a_2X_p(Z^i))\,\partial_i\big|_p \\ &= (a_1X_p(Y^i))\,\partial_i\big|_p + (a_2X_p(Z^i))\,\partial_i\big|_p \\ &= (a_1\overline{\nabla}_X Y)_p + (a_2\overline{\nabla}_X Z)_p,\end{aligned}$$

where we used the fact that X_p is a derivation at p. As $p \in \mathbb{R}^n$ was arbitrary, $\overline{\nabla}_X Y$ is $\mathbb{R}$-linear in Y.

(iii) Let $Y = Y^i\partial_i$, where $Y^i \in C^\infty(\mathbb{R}^n)$. Then, (3.4) gives

$$\begin{aligned}(\overline{\nabla}_X(fY))_p = X_p((fY)^i)\,\partial_i\big|_p = X_p(fY^i)\,\partial_i\big|_p &= (f(p)X_p(Y^i) + X_p(f)Y^i(p))\,\partial_i\big|_p \\ &= f(p)X_p(Y^i)\,\partial_i\big|_p + X_p(f)\,Y^i(p)\,\partial_i\big|_p \\ &= f(p)(\overline{\nabla}_X Y)_p + (Xf)(p)Y_p,\end{aligned}$$

where we again used the fact that X_p is a derivation at p. Since $p \in \mathbb{R}^n$ was arbitrary, this finishes the proof. □

3.2 Further Properties of $\overline{\nabla}$

This section is based on [13].

Although Proposition 3.2 is sufficient to define a connection on a smooth manifold, examining the directional derivative on $\mathbb{R}^n$ more closely leads to special classes of connections that will be our main subject in the subsequent discussion. For this reason, we postpone the definition of a connection to the next subsection and consider further properties of $\overline{\nabla}$. The calculations to be presented in this subsection are mostly computational and therefore lack geometric motivation. The reason for this is that we prefer to defer geometrical interpretations to a more general setting that extends to smooth manifolds.

3.2.1 Torsion

We start by noting that $\overline{\nabla}$ is $\mathbb{R}$-bilinear by (3.4). We then ask whether it is symmetric, i.e., whether $\overline{\nabla}_X Y = \overline{\nabla}_Y X$ for all $X, Y \in \mathfrak{X}(\mathbb{R}^n)$. Let $X = X^i \partial_i$ and $Y = Y^j \partial_j$. By (3.5), we have

$$\overline{\nabla}_X Y = X(Y^i)\,\partial_i = (X^j \partial_j)(Y^i)\,\partial_i = X^j(\partial_j Y^i)\,\partial_i.$$

Observe that

$$X^j(\partial_j Y^i)\,\partial_i = X^j \partial_j (Y^i \partial_i) - X^j Y^i \partial_j \partial_i.$$

Hence,

$$\overline{\nabla}_X Y = X^j \partial_j (Y^i \partial_i) - X^j Y^i \partial_j \partial_i.$$

Interchanging X and Y,

$$\overline{\nabla}_Y X = Y^j \partial_j (X^i \partial_i) - Y^j X^i \partial_j \partial_i.$$

Thus,

$$\begin{aligned}\overline{\nabla}_X Y - \overline{\nabla}_Y X &= X^j \partial_j (Y^i \partial_i) - X^j Y^i \partial_j \partial_i - Y^j \partial_j (X^i \partial_i) + Y^j X^i \partial_j \partial_i \\ &= X^j \partial_j (Y^i \partial_i) - X^j Y^i \partial_j \partial_i - Y^j \partial_j (X^i \partial_i) + Y^j X^i \partial_i \partial_j,\end{aligned}$$

where we used the fact that partial derivatives on $\mathbb{R}^n$ commute. Relabelling the indices in the last term shows that it cancels with the second term. Hence,

$$\begin{aligned}\overline{\nabla}_X Y - \overline{\nabla}_Y X &= X^j \partial_j (Y^i \partial_i) - Y^j \partial_j (X^i \partial_i) \\ &= XY - YX \\ &= [X, Y],\end{aligned}$$

where $[X, Y] \coloneqq XY - YX$ is knwon as the Lie bracket of X and Y. The expression

$$T(X, Y) \coloneqq \overline{\nabla}_X Y - \overline{\nabla}_Y X - [X, Y]$$

is said to be the torsion of $\overline{\nabla}$.

3.2.2 Metric-Compatibility

From the theory of curves and surfaces in $\mathbb{R}^3$, we are accustomed to the following product rule for the standard inner product:

$$\frac{d}{dt}\langle X, Y\rangle = \left\langle \frac{dX}{dt}, Y \right\rangle + \left\langle X, \frac{dY}{dt} \right\rangle, \tag{3.6}$$

where $X = X(t)$ and $Y = Y(t)$ are vector fields along a curve $\gamma : I \to \mathbb{R}^3$. We ask whether (3.6) is satisfied by $\overline{\nabla}$, i.e., whether

$$\overline{\nabla}_X\langle Y, Z\rangle = \langle \overline{\nabla}_X Y, Z\rangle + \langle Y, \overline{\nabla}_X Z\rangle$$

holds for all $X, Y, Z \in \mathfrak{X}(\mathbb{R}^n)$.

Let $X = X^i\partial_i, Y = Y^i\partial_i$ and $Z = Z^k\partial_k$. Then,

$$\langle Y, Z\rangle = \langle Y^i\partial_i, Z^j\partial_j\rangle = Y^iZ^j\langle\partial_i, \partial_j\rangle = Y^iZ^j\delta_{ij} \in C^\infty(\mathbb{R}^n),$$

where we used the fact that the coordinate basis (∂_i) is orthonormal with respect to the standard inner product on $\mathbb{R}^n$. Then, by Remark 3.1,

$$\overline{\nabla}_X\langle Y, Z\rangle = X\langle Y, Z\rangle = X(Y^iZ^j\delta_{ij}).$$

We now compute the right-hand side using (3.5).

$$\begin{aligned}
\langle \overline{\nabla}_X Y, Z\rangle + \langle Y, \overline{\nabla}_X Z\rangle &= \langle X(Y^i)\partial_i, Z^j\partial_j\rangle + \langle Y^i\partial_i, X(Z^j)\partial_j\rangle \\
&= X(Y^i)Z^j\langle\partial_i, \partial_j\rangle + Y^iX(Z^j)\langle\partial_i, \partial_j\rangle \\
&= \left(X(Y^i)Z^j + Y^iX(Z^j)\right)\delta_{ij} \\
&= X(Y^iZ^j\delta_{ij}),
\end{aligned}$$

where we used the facts that vector fields on $\mathbb{R}^n$ are derivations of $C^\infty(\mathbb{R}^n)$ and $X(\delta_{ij}) = 0$. Therefore,

$$\overline{\nabla}_X\langle Y, Z\rangle = \langle \overline{\nabla}_X Y, Z\rangle + \langle Y, \overline{\nabla}_X Z\rangle$$

for all $X, Y, Z \in \mathfrak{X}(\mathbb{R}^n)$, and we say that $\overline{\nabla}$ is metric-compatible.

3.2.3 Curvature

We know that the mixed partial derivatives of smooth functions commute. This begs the question whether the mixed directional derivatives of a smooth vector field commute, that is, whether

$$\overline{\nabla}_X \overline{\nabla}_Y Z = \overline{\nabla}_Y \overline{\nabla}_X Z$$

holds for all $X, Y, Z \in \mathfrak{X}(\mathbb{R}^n)$. Equivalently,

$$[\overline{\nabla}_X, \overline{\nabla}_Y] Z = 0.$$

Let $X = X^i \partial_i, Y = Y^i \partial_i$ and $Z = Z^k \partial_k$. Then, (3.5) gives

$$\begin{aligned} \overline{\nabla}_X \overline{\nabla}_Y Z = \overline{\nabla}_X (Y(Z^i)\,\partial_i) &= X(Y(Z^i))\,\partial_i \\ &= X(Y^j \partial_j Z^i)\,\partial_i \\ &= X^k \partial_k (Y^j \partial_j Z^i)\,\partial_i \\ &= (X^k \partial_k Y^j \partial_j Z^i + X^k Y^j \partial_k \partial_j Z^i)\,\partial_i. \end{aligned}$$

$\overline{\nabla}_Y \overline{\nabla}_X Z$ can be calculated similarly. Taking their difference,

$$\begin{aligned} [\overline{\nabla}_X, \overline{\nabla}_Y] Z &= (X^k \partial_k Y^j \partial_j Z^i + X^k Y^j \partial_k \partial_j Z^i)\,\partial_i - (Y^k \partial_k X^j \partial_j Z^i + Y^k X^j \partial_k \partial_j Z^i)\,\partial_i \\ &= (X^k \partial_k Y^j \partial_j Z^i - Y^k \partial_k X^j \partial_j Z^i)\,\partial_i + (X^k Y^j \partial_k \partial_j Z^i - Y^k X^j \partial_k \partial_j Z^i)\,\partial_i. \end{aligned}$$

We relabel the indices of the last term and use the symmetry of mixed partials to obtain

$$X^k Y^j \partial_k \partial_j Z^i - Y^k X^j \partial_k \partial_j Z^i = 0.$$

Thus,

$$\begin{aligned} [\overline{\nabla}_X, \overline{\nabla}_Y] Z &= (X^k \partial_k Y^j \partial_j Z^i - Y^k \partial_k X^j \partial_j Z^i)\,\partial_i \\ &= (X(Y(Z^i)) - Y(X(Z^i)))\,\partial_i \\ &= (XY - YX)(Z^i)\,\partial_i \\ &= \overline{\nabla}_{XY-YX} Z \\ &= \overline{\nabla}_{[X,Y]} Z. \end{aligned}$$

The expression

$$R(X,Y) \coloneqq [\overline{\nabla}_X, \overline{\nabla}_Y] - \overline{\nabla}_{[X,Y]}$$

is called the curvature of $\overline{\nabla}$.

3.3 Connections on a Vector Bundle

Using the properties of the directional derivative in $\mathbb{R}^n$ listed in Proposition 3.2 as a prototype, we are now ready to define and delve into connections.

Definition 3.3 (Connections)**.** Let (E, M, π) be a real vector bundle of rank k. A *connection in* E is a map

$$\nabla : \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E),$$

written $(X, Y) \mapsto \nabla_X Y$, such that
(i) $\nabla_X Y$ is $C^\infty(M)$-linear in X.
(ii) $\nabla_X Y$ is $\mathbb{R}$-linear in Y.
(iii) ∇ satisfies the following product rule: for $f \in C^\infty(M)$,

$$\nabla_X(fY) = (Xf)Y + f\nabla_X Y.$$

We call $\nabla_X Y$ the covariant derivative of Y in the direction X.

Example 3.4 (Euclidean Connection on $\mathbb{R}^n$)**.** As shown in Proposition 3.2, the directional derivative on $\mathbb{R}^n$ defined by (3.3) is a connection on $\mathbb{R}^n$, also known as the Euclidean connection on $\mathbb{R}^n$. To distinguish, we denote the Euclidean connection on $\mathbb{R}^n$ by $\overline{\nabla}$.

A surprising fact about connections is that, while they are global objects in the sense that they act on global sections, they are local operators. We first define what we mean by local operators.

Definition 3.5 (Local Operators [13])**.** Let (E, M, π) be a real vector bundle. An $\mathbb{R}$-linear map $F : \Gamma(E) \to \Gamma(E)$ is called a local operator if whenever a section $s \in \Gamma(E)$ vanishes on an open subset $U \subseteq M$, then $F(s) \in \Gamma(E)$ also vanishes on U.

Since a connection in E is $\mathbb{R}$-bilinear, we must separately show that $X \mapsto \nabla_X Y$ and $Y \mapsto \nabla_X Y$ are local operators for all $X \in \mathfrak{X}(M)$ and $Y \in \Gamma(E)$, which can be combined into the following lemma:

Lemma 3.6 (Locality of Connections)**.** Let M be a smooth manifold. For every $X \in \mathfrak{X}(M), Y \in \Gamma(E)$ and $p \in M$, if $X = \tilde{X}$ and $Y = \tilde{Y}$ on a neighborhood of p for some $\tilde{X}, \tilde{Y} \in \mathfrak{X}(M)$, then $\nabla_X Y\big|_p = \nabla_{\tilde{X}} \tilde{Y}\big|_p$.

Proof. We first show that $\nabla_X Y\big|_p = \nabla_X \tilde{Y}\big|_p$, which is equivalent to $\nabla_X (Y - \tilde{Y})\big|_p = 0$ by linearity. Let ψ be a smooth bump function with support in U such that $\psi(p) = 1$. Note that such a choice for ψ is possible since M is a Hausdorff space and singletons are closed in Hausdorff spaces. Then, $\psi(Y - \tilde{Y}) \equiv 0$ on M because $Y - \tilde{Y} \equiv 0$ on U and $\operatorname{supp}(\psi) \subseteq U$. Then, for every $X \in \mathfrak{X}(M)$,

$$\nabla_X(\psi(Y - \tilde{Y})) = \nabla_X(0 \cdot \psi(Y - \tilde{Y})) = 0\nabla_X(\psi(Y - \tilde{Y})) = 0.$$

In addition, the product rule gives

$$0 = \nabla_X(\psi(Y - \tilde{Y})) = (X\psi)(Y - \tilde{Y}) + \psi\nabla_X(Y - \tilde{Y}).$$

One has $(X\psi)(Y - \tilde{Y}) = 0$ on M since $Y - \tilde{Y} = 0$ on U and ψ vanishes outside U. Hence, $\psi\nabla_X(Y - \tilde{Y}) = 0$. Evaluating at p, we obtain $\psi(p)\nabla_X(Y - \tilde{Y})\big|_p = \nabla_X(Y - \tilde{Y})\big|_p = 0$.

We now show that $\nabla_X Y\big|_p = \nabla_{\tilde{X}} Y\big|_p$, which is equivalent to $\nabla_{X - \tilde{X}} Y\big|_p = 0$ by linearity. Similarly, we choose a bump function ψ with support in U such that $\psi(p) = 1$. Then, $\psi(X - \tilde{X}) \equiv 0$ on M and for every $Y \in \mathfrak{X}(M)$, the product rule gives

$$\nabla_{\psi(X - \tilde{X})} Y = \nabla_{0(X - \tilde{X})} Y = 0\nabla_X Y - 0\nabla_{\tilde{X}} Y = 0.$$

On the other hand,

$$\nabla_{\psi(X - \tilde{X})} = \psi\nabla_X Y - \psi\nabla_{\tilde{X}} Y.$$

Evaluating the above at p, we get

$$\psi(p)\nabla_X Y\big|_p - \psi(p)\nabla_{\tilde{X}} Y\big|_p = \nabla_X Y\big|_p - \nabla_{\tilde{X}} Y\big|_p = 0.$$

Combining our results,

$$\nabla_X Y = \nabla_X \tilde{Y} = \nabla_{\tilde{X}} \tilde{Y}.$$

□

We now prove an important result that will enable us to represent connections by functions called connection coefficients.

Proposition 3.7 (Restriction of a Connection)**.** Let ∇ be a connection in (E, M, π). For every open subset $U \subseteq M$, there exists a unique connection, denoted by ∇^U, on the restricted bundle $E\big|_U$ such that

$$\nabla^U_{(X|_U)}(Y\big|_U) = (\nabla_X Y)\big|_U$$

for all $X \in \mathfrak{X}(M)$ and $Y \in \Gamma(E)$.

Proof. Before starting the proof, observe that for $p \in U$, $\nabla^U_X Y\big|_p$ is meaningful by Lemma 3.6. despite ∇^U not being defined on $\Gamma(E)$.

We first establish uniqueness. Given $p \in U$, take a bump function ψ equal to 1 in some neighborhood of p with support in U. Then, define $\tilde{X} \in \mathfrak{X}(M), \tilde{Y} \in \Gamma(E)$ as $\tilde{X} = \psi X$ and $\tilde{Y} = \psi Y$. Then, $X = \tilde{X}$ and $\tilde{Y} = Y$ on some neighborhood of p. Thus, Lemma 3.6 applies and we have

$$\nabla^U_X Y\big|_p = \nabla^U_{(\tilde{X}|_U)}(\tilde{Y}\big|_U)\big|_p = \nabla_{\tilde{X}} \tilde{Y}\big|_p, \tag{3.7}$$

which shows uniqueness since the right-hand side is independent of ∇^U, i.e., determined only by ∇.

For existence, construct $\tilde{X}$ and $\tilde{Y}$ as above. Then, define $\nabla^U_X Y\big|_p$ by (3.7). By Lemma 3.6, $\nabla^U_X Y$ does not depend on the choices of $\tilde{X}$ and $\tilde{Y}$. It remains to show that ∇^U is a connection, which simply follows from the inherited properties of ∇. $\square$

Corollary 3.8. Assume the conditions of Lemma 3.6. Then, $\nabla_X Y\big|_p$ depends only on the value of X at p and the values of Y in a neighborhood of p.

Proof. By Lemma 3.6, it suffices to show that the value of $\nabla_X Y\big|_p$ only depends on the value of X at p. Suppose now that $\tilde{X}_p = X_p$ for some $\tilde{X} \in \mathfrak{X}(M)$. By linearity, we must prove that $\nabla_{(X-\tilde{X})} Y\big|_p = 0$. We have $X - \tilde{X} = (X^i - \tilde{X}^i)E_i$ in some local frame for U. By Proposition 3.7, we can work with a restricted connection ∇^U with U being some neighborhood of p. Then, for all $Y \in \Gamma(E\big|_U)$, one has

$$\nabla^U_{(X-\tilde{X})} Y\big|_p = (X^i(p) - \tilde{X}^i(p))\nabla_{E_i} Y\big|_p = 0 \nabla_{E_i} Y\big|_p = 0.$$

$\square$

Let $v \in T_pM$ and let Y be a local section of E. By this corollary, $\nabla_v Y$ is now a meaningful expression in the following sense. Choose a vector field X defined on a neighborhood of p with $X_p = v$, and define $\nabla_v Y \coloneqq \nabla_X Y\big|_p$.

3.4 Connections on a Smooth Manifold

In this project, we are mostly interested in connections in the tangent bundle, simply called connections on M. Since the tangent bundle of a manifold is also a vector bundle, all of our results from subsection 3.3 carry over directly.

By Proposition 3.7, we can represent a connection locally. To see this, consider a local frame (E_i) for TM over an open subset $U \subseteq M$. Then, we can write $\nabla_{E_i} E_j$ as a linear combination of the same local frame:

$$\nabla_{E_i} E_j = \Gamma_{ij}^k E_k,$$

where $\Gamma_{ij}^k : U \to \mathbb{R}$ are called the connection coefficients of ∇ with respect to the frame (E_i). Here is the main result about connection coefficients.

Proposition 3.9. Let ∇ be a connection on M. and let (E_i) be a local frame for an open subset $U \subseteq M$. For $X, Y \in \mathfrak{X}(U)$, written in terms of the local frame as $X = X^i E_i$ and $Y = Y^j E_j$, one has

$$\nabla_X Y = (X(Y^k) + X^i Y^j \Gamma_{ij}^k) E_k. \tag{3.8}$$

Proof. We use the definition of a connection.

$$\begin{aligned}
\nabla_X Y = \nabla_{X^i E_i}(Y^j E_j) &= X^i \nabla_{E_i}(Y^j E_j) \\
&= X^i (E_i(Y^j) E_j + Y^j \nabla_{E_i} E_j) \\
&= X^i E_i(Y^j) E_j + X^i Y^j \Gamma_{ij}^k E_k \\
&= X(Y^j) E_j + X^i Y^j \Gamma_{ij}^k E_k.
\end{aligned}$$

Relabelling the indices in the first term, we obtain the desired result. □

As a special case, let $X = \partial_i$ and $Y = Y^j \partial_j$. Then, (3.8) gives

$$\nabla_X Y = \nabla_{\partial_i}(Y^j \partial_j) = (\partial_i Y^k + Y^j \Gamma_{ij}^k)\, \partial_k.$$

We introduce the notation $\nabla_i Y^k := E_i(Y^k) + \Gamma^k_{ij} Y^j$ (we usually work with coordinate frames so that $E_i = \partial_i$), which can be thought of as the k-th component of the covariant derivative Y along ∂_i. For this reason, we call it the component notation. Moreover, this component form can be seen as the generalization of the partial derivatives of components of vector fields, where the first term is the familiar partial differentiation and the second term captures the change of basis vectors.

Remark 3.10. In the physics literature, the component notation is often referred to concisely as the covariant derivative.

Remark 3.11. The concept of covariant derivative can be generalized to tensor fields. We will not cover the construction of covariant derivatives of tensor fields (see Prop. 4.15 and Prop. 4.16 of [9] for details.) Instead, we will use the results. Specifically, if A is a mixed tensor field of rank (k, l) written locally as

$$A = A^{i_1 \ldots i_k}_{j_1 \ldots j_l} \, E_i \otimes \cdots \otimes E_{i_k} \otimes d\varepsilon^{j_1} \otimes \cdots \otimes d\varepsilon^{j_l},$$

the covariant derivative of A is defined as

$$\begin{aligned} \nabla_X A = \left(X(A^{i_1 \ldots i_k}_{j_1 \ldots j_l}) + \sum_{s=1}^{k} X^m A^{i_1 \ldots p \ldots i_k}_{j_1 \ldots j_l} \Gamma^{i_s}_{mp} - \sum_{s=1}^{l} X^m A^{i_1 \ldots i_k}_{j_1 \ldots p \ldots j_l} \Gamma^{p}_{m j_s} \right) \times \\ E_{i_1} \otimes \cdots \otimes E_{i_k} \otimes \varepsilon^{j_1} \otimes \cdots \otimes \varepsilon^{j_l}. \end{aligned} \tag{3.9}$$

Equivalently, in component notation, one has [3]

$$\nabla_a A^{i_1 \ldots i_k}_{j_1 \ldots j_l} = E_a \left(A^{i_1 \ldots i_k}_{j_1 \ldots j_l} \right) + \sum_{s=1}^{k} A^{i_1 \ldots p \ldots i_k}_{j_1 \ldots j_l} \mathit{\Gamma}^{i_s}_{ap} - \sum_{s=1}^{l} A^{i_1 \ldots i_k}_{j_1 \ldots p \ldots j_l} \mathit{\Gamma}^{p}_{a j_s}.$$

Specifically, if $f \in C^\infty(M)$ (a tensor of rank (0,0)), then

$$\nabla_X f = X f.$$

In order to introduce torsion, metric-compatibility, and curvature in a more general setting, we need to do some more work.

3.5 Covariant Derivative Along Curves

One is usually interested in differentiating a vector field along a curve rather than along a vector field, which helps to extend the concept of parallel vectors to manifolds. Let us first make precise what is meant by a vector field along a curve.

Definition 3.12 (Vector Fields along Curves)**.** Let $\gamma : I \to M$ be a smooth curve on a smooth manifold M. A smooth map $V : I \to TM$ with $\pi \circ V = \gamma$, where $\pi : TM \to M$ is the natural projection map, is called a vector field along γ. In short, $V(t) \in T_{\gamma(t)}M$ for all $t \in I$.

The set of smooth vector fields along γ, denoted by $\mathfrak{X}(\gamma)$, is a vector space over $\mathbb{R}$ under pointwise addition and scalar multiplication.

Example 3.13. Let $(U, (x^i))$ be a coordinate chart M, and let $X = \partial_i$ be the corresponding coordinate vector field on U. Let $\gamma : I \to M$ be a smooth curve with $\gamma(I) \subseteq U$. Then, the map $V : I \to TM$ given by

$$V(t) = X_{\gamma(t)} = \partial_i|_{\gamma(t)}$$

is clearly smooth. Thus, coordinate vector fields restrict to smooth vector fields along any smooth curve $\gamma : I \to M$ with $\gamma(I) \subseteq U$.

Definition 3.14. Let $V \in \mathfrak{X}(\gamma)$. We say that V is extendible if there exists a vector field $\tilde{V}$ on a neighborhood of the image of γ such that $V = \tilde{V} \circ \gamma$.

Example 3.15. Consider Example 3.13. Working backwards, we see that a coordinate field along a curve can be naturally extended.

We can now make precise what we mean by a covariant derivative along a curve.

Theorem 3.16. Let ∇ be a connection on M. For each smooth curve $\gamma : I \to M$, there exists a unique operator $D_t : \mathfrak{X}(\gamma) \to \mathfrak{X}(\gamma)$, called the covariant derivative along γ, such that

(i) D_t is linear over $\mathbb{R}$.

(ii) $D_t(fV) = \dot{f}V + fD_tV$ for all $f \in C^\infty(I)$, where $\dot{}$ denotes differentiation with respect to t.

(iii) If $V \in \mathfrak{X}(\gamma)$ is extendible, then $D_tV(t) = \nabla_{\dot{\gamma}(t)}\tilde{V}$ for every extension $\tilde{V}$ of V.

Proof. We first show uniqueness. Since $D_tV(t) = \nabla_{\dot\gamma(t)}\tilde{V}$, the value of D_tV depends only on the values of V in a neighborhood of t by Corollary 3.8. Now choose a local coordinate frame (∂_i) for a neighborhood of $\gamma(t_0)$ and write $V(t) = V^j(t)\,\partial_j\big|_{\gamma(t)}$, where $V^j(t)$ are smooth functions defined on the neighborhood of t_0. Since ∂_j is extendible by Example 3.15, properties (i), (ii) and (iii) gives

$$\begin{aligned} D_tV(t) = D_t\left(V^j(t)\partial_j\big|_{\gamma(t)}\right) &= \dot V^j(t)\partial_j\big|_{\gamma(t)} + V^j(t)D_t\left(\partial_j\big|_{\gamma(t)}\right) \\ &= \dot V^j(t)\partial_j\big|_{\gamma(t)} + V^j(t)\nabla_{\dot\gamma(t)}\partial_j\big|_{\gamma(t)} \\ &= \dot V^j(t)\partial_j\big|_{\gamma(t)} + V^j(t)\nabla_{\dot\gamma(t)}\partial_j, \end{aligned}$$

where we abused notation and denoted the extension of ∂_j along γ by ∂_j. Now suppose that Y is a vector field defined on some neighborhood of $\gamma(t)$ with $Y_{\gamma(t)} = \dot\gamma(t)$. Then,

$$\begin{aligned} D_tV(t) = \dot V^j(t)\partial_j\big|_{\gamma(t)} + V\nabla_{\dot\gamma(t)}\partial_j &= \dot V^j(t)\partial_j\big|_{\gamma(t)} + V^j(t)\nabla_Y\partial_j\big|_{\gamma(t)} \\ &= \dot V^j(t)\partial_j\big|_{\gamma(t)} + V^j(t)(Y(\delta_j^l) + \Gamma_{il}^k\delta_j^lY^i)\partial_k\big|_{\gamma(t)} \\ &= \dot V^j(t)\partial_j\big|_{\gamma(t)} + \Gamma_{ij}^k(\gamma(t))V^j(t)\dot\gamma^i(t)\partial_k\big|_{\gamma(t)}. \end{aligned}$$

Renaming the indices, we obtain

$$D_tV(t) = (\dot V^k(t) + \Gamma_{ij}^k(\gamma(t))\dot\gamma^i(t)V^j(t))\,\partial_k\big|_{\gamma(t)}, \tag{3.10}$$

which shows that the right-hand side determines D_t uniquely.

For existence, define D_t by (3.10) if $\gamma(I)$ is contained in a single chart. It can be checked that D_t satisfies the required properties. In general, we can define D_t by (3.10) in each chart. Then, since we have shown uniqueness, different D_t's must be equal on the intersection of charts. Therefore, we can define D_t for any smooth curve in M. □

Remark 3.17. Covariant derivative of tensor fields along a curve can also be defined. In component notation, the analogue of (9) for a tensor field T of rank (k, l) along a smooth curve γ is given by [14]

$$D_tT(t) = \dot T^{j_1\ldots j_k}_{i_1\ldots i_l}(t) + \sum_{r=1}^{k}\Gamma^{j_r}_{mn}\dot\gamma^m(t)T^{j_1\ldots n\ldots j_k}_{i_1\ldots i_l}(t) - \sum_{s=1}^{l}\Gamma^n_{mi_s}\dot\gamma^m(t)T^{j_1\ldots j_k}_{i_1\ldots n\ldots i_l} \tag{3.11}$$

We can now define parallel vectors.

Definition 3.18. Let ∇ be a connection on M and suppose $\gamma : I \to M$ is a smooth curve. $V \in \mathfrak{X}(\gamma)$ is said to be parallel along γ if $D_t V = 0$ for all $t \in I$.

A parallel vector field along a curve can be visualized as follows:

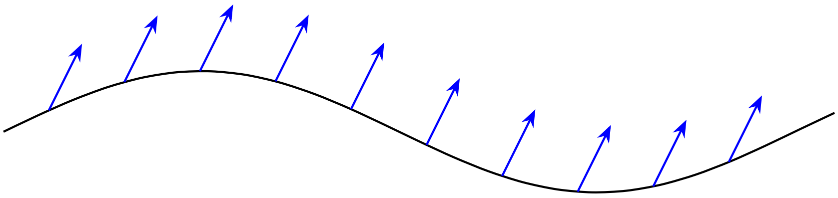

Figure 1: A parallel vector field along a curve (adapted from [9])

The next result shows that we can extend vectors on a curve in parallel.

Theorem 3.19 (Parallel Transport)**.** Let ∇ be a connection on M and let $\gamma : I \to M$ be a smooth curve with $t_0 \in I$. For $v \in T_{\gamma(t_0)}M$, there is a unique parallel vector field V along γ with $V(t_0) = v$.

Proof. Let $(v^1, \ldots, v^n)$ be the components of v. By (3.10), a vector field V is parallel along γ if and only if the components satisfy

$$\dot{V}^k(t) = -V^j(t)\dot{\gamma}^i(t)\Gamma_{ij}^k(\gamma(t)), \quad V^k(t_0) = v^k. \tag{3.12}$$

The theorem on the existence, uniqueness, and smoothness of linear ODEs (for instance, see Thm 4.31 of [9]) finishes the proof. □

3.6 Curvature

Definition 3.20. Let ∇ be a connection on M. The map $R : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ given by

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

is called the curvature of the connection.

Before showing that R is a tensor field, we need a preliminary result.

Lemma 3.21. Let M be a smooth manifold. For all $X, Y, Z \in \mathfrak{X}(M)$ and $f, g \in C^\infty(M)$,

$$[fX, gY] = fg[X,Y] + (fXg)Y - (gYf)X.$$

Proof. Let $h \in C^\infty(M)$. We first show that $g(Yh) = (gY)h$. Let $p \in M$. Then,

$$(g(Yh))(p) = g(p)(Yh)(p) = g(p)Y_p h = (gY)_p h = ((gY)h)(p).$$

Since $p \in M$ was arbitrary, we have $g(Yh) = (gY)h$. Then,

$$\begin{aligned}
[fX, gY]h &= (fX)(gY)h - (gY)(fX)h \\
&= (fX)(g(Yh)) - (gY)(f(Xh)) \\
&= g(fX(Yh)) + Yh(fX(g)) - f(gY(Xh)) - Xh(gY(f)) \\
&= (gfX)(Yh) - (fgY)(Xh) + (fXg)(Yh) - (gYf)(Xh) \\
&= fg(XYh - YXh) + (fXg)Yh - (gYf)Xh \\
&= fg[X,Y]h + (fXg)Yh - (gYf)Xh.
\end{aligned}$$

Since this holds for any $h \in C^\infty(M)$, we obtain our result. □

Remark 3.22. Since $Xg = 0$ for a constant function $g \in C^\infty(M)$, we get

$$[fX, Y] = f[X,Y] - (Yf)X \tag{3.13}$$

for the special case $g = 1$.

Proposition 3.23. R is a tensor field on M.

Proof. We check linearity over $C^\infty(M)$ in each argument. Let $f, g \in C^\infty(M)$ and suppose $W \in \mathfrak{X}(M)$. Then,

$$\begin{aligned}
R(fX + gW, Y)Z &= \nabla_{fX+gW}\nabla_Y Z - \nabla_Y \nabla_{fX+gW} Z - \nabla_{[fX+gW,Y]}Z \\
&= f\nabla_X\nabla_Y Z + g\nabla_W\nabla_Y Z - \nabla_Y(f\nabla_X Z + g\nabla_W Z) \\
&\quad - \nabla_{[fX,Y]}Z - \nabla_{[gW,Y]}Z \\
&= f\nabla_X\nabla_Y Z + g\nabla_W\nabla_Y Z - \nabla_Y(f\nabla_X Z) - \nabla_Y(g\nabla_W Z) \\
&\quad - \nabla_{f[X,Y]-(Yf)X}Z - \nabla_{g[W,Y]-(Yg)W}Z \\
&= f\nabla_X\nabla_Y Z + g\nabla_W\nabla_Y Z - (Yf)\nabla_X Z - f\nabla_Y\nabla_X Z - (Yg)\nabla_W Z \\
&\quad - g\nabla_Y\nabla_W Z - f\nabla_{[X,Y]}Z + (Yf)\nabla_X Z - g\nabla_{[W,Y]}Z + (Yg)\nabla_W Z \\
&= fR(X,Y)Z + gR(W,Y)Z,
\end{aligned}$$

where we have used (3.13) in the third line. Linearity in Y is a simple result of $R(X,Y)Z = -R(Y,X)Z$. For linearity in Z, we have

$$\begin{aligned}
R(X,Y)(fZ+gW) &= \nabla_X\nabla_Y(fZ+gW) - \nabla_Y\nabla_X(fZ+gW) - \nabla_{[X,Y]}(fZ+gW)\\
&= \nabla_X(\nabla_Y fZ + \nabla_Y gW) - \nabla_Y(\nabla_X fZ + \nabla_X gW) - \nabla_{[X,Y]}fZ\\
&\quad - \nabla_{[X,Y]}gW\\
&= \nabla_X((Yf)Z + f\nabla_Y Z + (Yg)W + g\nabla_Y W) - \nabla_Y((Xf)Z)\\
&\quad - \nabla_Y(f\nabla_X Z + (Xg)W + g\nabla_X W) - ([X,Y]f)Z - f\nabla_{[X,Y]}Z\\
&\quad - ([X,Y]g)W - g\nabla_{[X,Y]}W\\
&= X(Yf)Z + (Yf)\nabla_X Z + (Xf)\nabla_Y Z + f\nabla_X\nabla_Y Z + X(Yg)W\\
&\quad + (Yg)\nabla_X W + (Xg)\nabla_Y W + g\nabla_X\nabla_Y W - Y(Xf)Z - (Xf)\nabla_Y Z\\
&\quad - (Yf)\nabla_X Z - f\nabla_Y\nabla_X Z - Y(Xg)W - (Xg)\nabla_Y W - (Yg)\nabla_X W\\
&\quad - g\nabla_Y\nabla_X W - ([X,Y]f)Z - f\nabla_{[X,Y]}Z - ([X,Y]g)W - g\nabla_{[X,Y]}W\\
&= f(\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z)\\
&\qquad + g(\nabla_X\nabla_Y W - \nabla_Y\nabla_X W - \nabla_{[X,Y]}W)\\
&= fR(X,Y)Z + gR(X,Y)W.
\end{aligned}$$

□

One can show that the curvature tensor is a (1,3)-tensor (for instance, see Lemma B.6 of [9]). Thus, in a coordinate chart $(U,(x^i))$, one can write

$$R = R^l{}_{ijk}\,\partial_l \otimes dx^i \otimes dx^j \otimes dx^k,$$

where $R^l{}_{ijk} \in C^\infty(U)$ can be found by the action of R on the basis vectors. Namely,

$$\begin{aligned}
R(\partial_a,\partial_b)\,\partial_c &= R^l{}_{ijk}\,\partial_l\,dx^i(\partial_a)\,dx^j(\partial_b)\,dx^k(\partial_c)\\
&= R^l{}_{ijk}\,\partial_l\delta^i_a\delta^j_b\delta^k_c\\
&= R^l{}_{abc}\,\partial_l.
\end{aligned}$$

Proposition 3.24. Let ∇ be a connection on M. If $(U,(x^i))$ is a coordinate chart, the components of the curvature tensor of the connection read as

$$R^l{}_{ijk} = \partial_i \Gamma^l_{jk} - \partial_j \Gamma^l_{ik} + \Gamma^m_{jk}\Gamma^l_{im} - \Gamma^m_{ik}\Gamma^l_{jm}. \tag{3.14}$$

Proof. We determine the action of R on (∂_i). Using Definition 3.20, one has

$$R(\partial_i, \partial_j)\,\partial_k = \nabla_{\partial_i}\nabla_{\partial_j}\partial_k - \nabla_{\partial_j}\nabla_{\partial_i}\partial_k - \nabla_{[\partial_i,\partial_j]}\partial_k.$$

Since the Lie bracket of the coordinate vector fields is zero, the last term vanishes, and we have

$$\begin{aligned} R^l{}_{ijk}\,\partial_l &= \nabla_{\partial_i}(\Gamma^m_{jk}\partial_m) - \nabla_{\partial_j}(\Gamma^m_{ik}\partial_m) \\ &= \partial_i\Gamma^m_{jk}\,\partial_m + \Gamma^m_{jk}\nabla_{\partial_i}\partial_m - \partial_j\Gamma^m_{ik}\partial_m - \Gamma^m_{ik}\nabla_{\partial_j}\partial_m \end{aligned}$$

Then,

$$R^l{}_{ijk}\,\partial_l = \partial_i\Gamma^m_{jk}\,\partial_m + \Gamma^m_{jk}\Gamma^l_{im}\partial_l - \partial_j\Gamma^m_{ik}\partial_m - \Gamma^m_{ik}\Gamma^l_{jm}\partial_l.$$

Renaming the indices, we obtain

$$R^l{}_{ijk}\,\partial_l = (\partial_i \Gamma^l_{jk} - \partial_j \Gamma^l_{ik} + \Gamma^m_{jk}\Gamma^l_{im} - \Gamma^m_{ik}\Gamma^l_{jm})\,\partial_l.$$

□

Since this project focuses on connections, torsion, and non-metricity, we will not provide an interpretation of curvature tensor and refer to [4, 14] for a geometrical interpretation.

3.7 Torsion

Definition 3.25. Let ∇ be a connection on M. The map $T : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ defined by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \tag{3.15}$$

is called the torsion tensor of ∇.

Proposition 3.26. The map T defined by (3.15) is a tensor field.

Proof. We simply check linearity over $C^\infty(M)$ in each argument. By the observation $T(X, Y) = -T(Y, X)$, it suffices to prove linearity in the first argument. By properties of a connection, we have

$$\begin{aligned}
T(fX+gY,Z) &= \nabla_{fX+gY}Z - \nabla_Z(fX+gY) - [fX+gY,Z] \\
&= f\nabla_X Z + g\nabla_Y Z - \nabla_Z(fX) - \nabla_Z(gY) - [fX+gY,Z] \\
&= f\nabla_X Z + g\nabla_Y Z - (Zf)X - f\nabla_Z X - (Zg)Y - g\nabla_Z Y \\
&\quad - [fX,Z] - [gY,Z] \\
&= f\nabla_X Z + g\nabla_Y Z - (Zf)X - f\nabla_Z X - (Zg)Y - g\nabla_Z Y \\
&\quad - (f[X,Z] - (Zf)X) - (g[Y,Z] - (Zg)Y)
\end{aligned}$$

$$\begin{aligned}
T(fX+gY,Z) &= f\nabla_X Z + g\nabla_Y Z - f\nabla_Z X - g\nabla_Z Y - f[X,Z] - g[Y,Z] \\
&= f(\nabla_X Z - \nabla_Z X - [X,Z]) + g(\nabla_Y Z - \nabla_Z Y - [Y,Z]) \\
&= fT(X,Z) + gT(Y,Z),
\end{aligned}$$

where we used (3.13) in the fourth line. □

Definition 3.27 (Symmetric Connections)**.** A connection on M is called as symmetric if its torsion tensor vanishes identically for all $X, Y \in \mathfrak{X}(M)$.

Proposition 3.28 (Characterization of Symmetric Connections)**.** A connection on M is symmetric if and only if its connection coefficients with respect to every local coordinate frame are symmetric, that is, $\Gamma^k_{ij} = \Gamma^k_{ji}$.

Proof. Let (∂_i) be a local coordinate frame. Writing $X = X^i\partial_i$ and $Y = Y^j\partial_j$, we have

$$\begin{aligned}
T(X,Y) &= \nabla_X Y - \nabla_Y X - [X,Y] \\
&= (X(Y^k) + X^iY^j\Gamma^k_{ij} - Y(X^k) - Y^iX^j\Gamma^k_{ij})\,\partial_k - (X(Y^k) - Y(X^k))\,\partial_k \\
&= (X^iY^j - Y^iX^j)\,\Gamma^k_{ij}\,\partial_k \\
&= (\Gamma^k_{ij} - \Gamma^k_{ji})X^iY^j\,\partial_k,
\end{aligned}$$

where we have written the Lie bracket in coordinates in the second line and renamed the indices in the third line to obtain the fourth line. By the above expression, it is clear that the torsion tensor vanishes if the connection coefficients are symmetric. Conversely, if the torsion tensor vanishes, we can choose X and Y such that $X^i = Y^j = 1$. Hence, the vanishing of the torsion tensor implies that $\Gamma^k_{ij} = \Gamma^k_{ji}$. □

3.7.1 Geometrical Interpretation of Torsion

In this subsection, we follow [10, 3, 4].

With the tools we have developed, we can now interpret torsion geometrically. While the calculations in the following are not rigorous, they give us a heuristic way to think about torsion.

Let ∇ be a connection on M. Assume $\gamma : I \to M$ and $\tilde{\gamma} : \tilde{I} \to M$ be two smooth curves on M such that $p \coloneqq \gamma(t_0) = \tilde{\gamma}(t_0)$ for some $t_0 \in I \cap \tilde{I}$. Suppose that $(U, (x^i))$ is a coordinate chart such that $p \in U$. Let $\epsilon > 0$ be small such that $(t_0, t_0 + \epsilon_1) \subseteq I \cap \tilde{I}$. Set $s \coloneqq \gamma(t_0 + \epsilon)$ and $q \coloneqq \tilde{\gamma}(t_0 + \epsilon)$.

By Theorem 3.19, we can find a unique parallel vector field $\tilde{W}$ along $\tilde{\gamma}$ such that $\tilde{W}(t_0) = \dot{\gamma}(t_0)$. In the coordinate chart $(U, (x^i))$, (3.12) reads as

$$\dot{\tilde{W}}^k(t) = -\Gamma^k_{ij}(\tilde{\gamma}(t))\dot{\tilde{\gamma}}^j(t)\tilde{W}^i(t).$$

Since ϵ is small, we can Taylor expand $\tilde{W}^k(t)$ at $t_0 + \epsilon$ as [3]

$$\tilde{W}^k(t_0 + \epsilon) = \tilde{W}^k(t_0) + \epsilon \left. \frac{d\tilde{W}^k}{dt} \right|_{t_0} + O(\epsilon^2),$$

where $O(\epsilon^2)$ denotes higher-order terms in ϵ. Combining, we obtain

$$\begin{aligned} \tilde{W}^k(t_0 + \epsilon) &= \dot{\gamma}^k(t_0) - \epsilon\, \Gamma^k_{ij}(p)\, \tilde{W}^i(t_0)\, \dot{\tilde{\gamma}}^j(t_0) + O(\epsilon^2) \\ &= \dot{\gamma}^k(t_0) - \epsilon\, \Gamma^k_{ij}(p)\, \dot{\gamma}^i(t_0)\, \dot{\tilde{\gamma}}^j(t_0) + O(\epsilon^2). \end{aligned}$$

Similarly, one can find a unique parallel vector field W along γ satisfying $W(t_0) = \dot{\tilde{\gamma}}(t_0)$. Parallel transporting $\tilde{\gamma}(t)$ along W from t_0 to $t_0 + \epsilon$ by carrying out an analogous calculation in local coordinates yields.

$$\begin{aligned} W^k(t_0 + \epsilon) &= \dot{\tilde{\gamma}}^k(t_0) - \epsilon\, \Gamma^k_{ij}(p)\, W^i(t_0)\, \dot{\gamma}^j(t_0) + O(\epsilon^2) \\ &= \dot{\tilde{\gamma}}^k(t_0) - \epsilon\, \Gamma^k_{ij}(p)\, \dot{\tilde{\gamma}}^i(t_0)\, \dot{\gamma}^j(t_0) + O(\epsilon^2). \end{aligned}$$

To give a geometric interpretation, we start by thinking of vectors as arrows. In the figure above ps and pq represent initial vectors $\dot{\gamma}(t_0)$ and $\dot{\tilde{\gamma}}(t)$, respectively. After

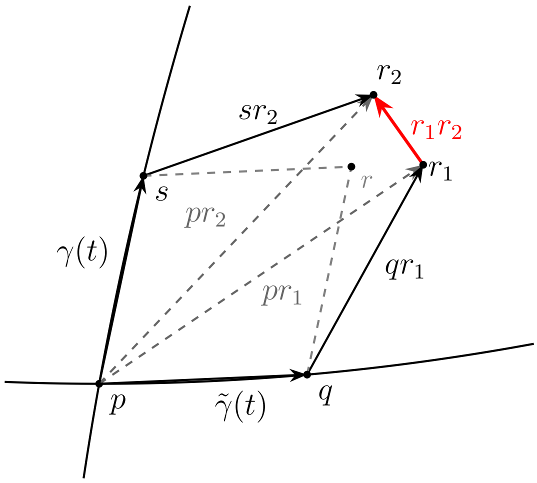

Figure 2: Torsion as the closure of infinitesimal parallelograms (adapted from [10])

parallel transport, we obtain the vectors W and $\tilde{W}$, represented by the vectors sr_2 and sr_1, respectively. Since ϵ is thought to be infinitesimal, we suppose $p = q = s$. Then, the difference r_1r_2 between pr_2 and pr_1 can be found by subtracting the sum of the components of $\dot{\gamma}(t)$ and $\tilde{W}$ from the sum of the components of $\gamma(t)$ and W. This difference is given by

$$\begin{aligned}(W^k(t_0+\epsilon)+\dot{\gamma}^k(t_0)) - (\tilde{W}^k(t_0+\epsilon)+\dot{\tilde{\gamma}}^k(t_0)) &= \epsilon\Gamma^k_{ij}(p)\dot{\gamma}^i(t_0)\dot{\tilde{\gamma}}^j(t_0) - \epsilon\Gamma^k_{ij}(p)\dot{\tilde{\gamma}}^i(t_0)\dot{\gamma}^j(t_0)\\ &= \epsilon(\Gamma^k_{ij}(p)-\Gamma^k_{ji}(p))\dot{\gamma}^i(t_0)\dot{\tilde{\gamma}}^j(t_0).\end{aligned}$$

From this expression, we see that $r_1r_2 = 0$ if the torsion tensor vanishes. Hence, geometrically, torsion prevents the closure of infinitesimal parallelograms.

3.8 Metric-Compatibility

Definition 3.29. Let ∇ be a connection on a Rimeannian manifold (M, g). We say that ∇ is compatible with g, or simply metric-compatible if g is understood, if

$$\nabla_X\langle Y, Z\rangle = \langle\nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle \tag{3.16}$$

for all $X, Y, Z \in \mathfrak{X}(M)$.

Remark 3.30. Since $\langle X, Y\rangle \in C^\infty(M)$, $\nabla_X\langle Y, Z\rangle = X\langle Y, Z\rangle$.

Proposition 3.31 (Characterization of Metric-Compatibility)**.** Let ∇ be a connection on (M, g). The connection is metric-compatible if and only if

$$\nabla_k g_{ij} = 0$$

for every local frame (E_i).

Proof. Let (E_i) be a local frame. Then,

$$\nabla_X \langle Y, Z \rangle = \nabla_X (g_{ij} Y^i Z^j) = X(g_{ij} Y^i Z^j) = X(g_{ij}) Y^i Z^j + g_{ij} X(Y^i) Z^j + g_{ij} Y^i X(Z^j).$$

On the other hand,

$$\begin{aligned}
\langle \nabla_X Y, Z \rangle &= \langle \nabla_{X^i E_i} (Y^j E_j), Z^k E_k \rangle \\
&= \langle X^i (Y^j \nabla_{E_i} E_j + E_i(Y^j) E_j), Z^k E_k \rangle \\
&= X^i Z^k (Y^j \langle \Gamma^l_{ij} E_l, E_k \rangle + E_i(Y^j) \langle E_j, E_k \rangle) \\
&= X^i Z^k (Y^j \Gamma^l_{ij} g_{lk} + E_i(Y^j) g_{jk}).
\end{aligned}$$

Renaming the indices,

$$\langle \nabla_X Y, Z \rangle = X^i Z^k (E_i(Y^j) + \Gamma^j_{il} Y^l) g_{jk}.$$

By symmetry, $\langle Y, \nabla_X Z \rangle = X^i Y^k (E_i(Z^j) + \Gamma^j_{il} Z^l) g_{jk}$. Then, the metric-compatibility condition reads as

$$\begin{aligned}
X(g_{ij}) Y^i Z^j + g_{ij} X(Y^i) Z^j + g_{ij} Y^i X(Z^j) = X^i Z^k (E_i(Y^j) + \Gamma^j_{il} Y^l) g_{jk} \\
+ X^i Y^k (E_i(Z^j) \Gamma^j_{il} Z^l) g_{jk}.
\end{aligned}$$

Renaming the indices and rearranging the terms, one obtains

$$\begin{gathered}
X^k Y^i Z^j (E_k(g_{ij}) - \Gamma^l_{ki} g_{lj} - \Gamma^l_{kj} g_{il}) = 0 \\
X^k Y^i Z^j \nabla_k g_{ij} = 0.
\end{gathered}$$

The above shows that if $\nabla_k g_{ij} = 0$, then (3.16) is satisfied. Conversely, if (3.16) holds, it holds for all $X, Y, Z \in \mathfrak{X}(U)$. So, we must have $\nabla_k g_{ij} = 0$ by the above expression. $\square$

The following result gives us a way to interpret metric-compatibility geometrically.

Proposition 3.32. Let ∇ be a connection on (M, g) and suppose $\gamma : I \to M$ is a smooth curve. If V and W are parallel vector fields along γ, then ∇ is compatible with g if and only if $\langle V, W\rangle$ is constant along γ.

Proof. See Prop. 5.5 of [9]. $\square$

The metric g enables us to define angles and lengths of vectors in the same way as the inner product on $\mathbb{R}^n$. Hence, Proposition 3.32 asserts that the angle between parallel-transported vectors does not change. If $V = W$, the length of the parallel transported vector does not change, as one would intuitively expect.

Motivated by Proposition 3.32, we define a new tensor measuring the failure of a connection being compatible with the metric.

Definition 3.33. Let ∇ be a connection on (M, g). The map $Q : \mathfrak{X}(M)\times\mathfrak{X}(M)\times\mathfrak{X}(M) \to C^\infty(M)$ defined by

$$Q(X, Y, Z) = -(\nabla_X g)(Y, Z)$$

for all $X, Y, Z \in \mathfrak{X}(M)$ is called the non-metricity tensor of ∇ with respect to g [4]. Similar calculations as before show that Q is a tensor.

Proposition 3.34. In a local frame (E_i), the components of the non-metricity tensor are given by

$$Q_{ijk} = -\nabla_i g_{jk}.$$

Proof. Using (3.9), we have

$$-\nabla_{E_i} g = (-E_i(g_{ab}) + g_{cb}\Gamma^c_{ia} + g_{ac}\Gamma^c_{ib})(\varepsilon^a \otimes \varepsilon^b).$$

Then,

$$\begin{aligned}
Q_{ijk} = Q(E_i, E_j, E_k) &= -(\nabla_{E_i} g)(E_j, E_k) \\
&= (-E_i(g_{ab}) + g_{cb}\Gamma^c_{ia} + g_{ac}\Gamma^c_{ib})(\varepsilon^a \otimes \varepsilon^b)(E_j, E_k) \\
&= (-E_i(g_{ab}) + g_{cb}\Gamma^c_{ia} + g_{ac}\Gamma^c_{ib})\delta^a_j\delta^b_k \\
&= -E_i(g_{jk}) + \Gamma^l_{ij} g_{lk} + \Gamma^l_{ik} g_{jl} \\
&= -\nabla_i g_{jk}.
\end{aligned} \tag{3.17}$$

$\square$

3.9 Levi-Civita Connection

We have seen that there are mainly two types of connections. Then, a natural question is whether there is a connection that is metric-compatible and symmetric. The answer to this question summarizes our discussion about connections on Riemannian manifolds and leads to the theorem below.

Theorem 3.35 (Fundamental Theorem of Riemannian Geometry)**.** Let (M, g) be a Riemannian manifold. There exists a unique symmetric and metric-compatible connection on M, called the Levi-Civita connection of g.

Proof. We first show uniqueness. By metric-compatibility,

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.$$

Since the connection is also symmetric, the above becomes

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_Z X + [X, Z]\rangle, \quad (*)$$

where we used (3.15) with the left-hand side zero. Cyclic permutation of X, Y, Z yields two more equations:

$$\begin{aligned} Y\langle Z, X\rangle &= \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_X Y\rangle + \langle Z, [Y, X]\rangle \quad (**) \\ Z\langle X, Y\rangle &= \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Y Z\rangle + \langle X, [Z, Y]\rangle \quad (***) \end{aligned}$$

Adding $(*)$ and $(**)$ then subtracting $(***)$, one has

$$X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle = 2\langle \nabla_X Y, Z\rangle + \langle Y, [X, Z]\rangle + \langle Z, [Y, X]\rangle - \langle X, [Z, Y]\rangle.$$

Solving for $\langle \nabla_X Y, Z\rangle$, we obtain

$$\langle \nabla_X Y, Z\rangle \overset{(\cdot)}{=} \frac{1}{2}(X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle - \langle Y, [X, Z]\rangle - \langle Z, [Y, X]\rangle + \langle X, [Z, Y]\rangle).$$

This shows that if such a connection exists, it is unique since the right-hand side is independent of the connection.

To show existence, let $(U,(x^i))$ be a coordinate chart and let $X = \partial_i, Y = \partial_j$ and $Z = \partial_k$. Since the Lie bracket of coordinate vector fields vanish, $(\cdot)$ reduces to

$$\begin{aligned}
\langle \nabla_{\partial_i}\partial_j, \partial_k\rangle &= \frac{1}{2}(\partial_i\langle\partial_j,\partial_k\rangle + \partial_j\langle\partial_k,\partial_i\rangle - \partial_k\langle\partial_i,\partial_j\rangle) \\
\langle \Gamma^l_{ij}\partial_l, \partial_k\rangle &= \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}) \\
\Gamma^l_{ij} g_{lk} &= \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}).
\end{aligned}$$

Multiplying by g^{km}, one has

$$\Gamma^l_{ij}\,\delta^m_l = \Gamma^m_{ij} = \frac{1}{2}g^{km}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}).$$

One can check that $(\cdot)$ defines a connection by a straightforward calculation. The symmetry is apparent from the above form of the connection coefficients. To check metric-compatibility, we use Proposition 3.32:

$$\begin{aligned}
\partial_k g_{ij} - \mathit{\Gamma}^l_{ki} g_{lj} - \mathit{\Gamma}^l_{kj} g_{il} &= \partial_k g_{ij} - g_{lj}\frac{1}{2}g^{lm}(\partial_k g_{im} + \partial_i g_{mk} - \partial_m g_{ki}) \\
&\qquad - g_{il}\frac{1}{2}g^{lm}(\partial_k g_{jm} + \partial_j g_{mk} - \partial_m g_{kj}) \\
&= 0,
\end{aligned}$$

where we renamed the indices. □

To summarize, the Levi-Civita connection is the unique metric that preserves the inner product of parallel transported vectors and does not prevent the formation of infinitesimal parallelograms under parallel transport. In summary, the Levi-Civita connection behaves very nicely in the sense that it does not contradict one's geometrical intuition.

4 Connection, Torsion, Curvature and Non-Metricity Forms

In this section, we will see that information about curvature, torsion, and non-metricity can be encoded in differential forms. The main idea will be the same for all four forms that are to be introduced.

4.1 Connection 1-Forms

Let ∇ be a connection on M. Suppose (E_i) is a local frame on an open subset $U \subseteq M$, and denote the coframe dual to (E_i) by (ε^i). Then, we can write

$$\nabla_X E_i = \omega_i^{\ j}(X) E_j$$

for all $X \in \mathfrak{X}(U)$, where $\omega_i^{\ j}(X) \in C^\infty(U)$ depends on X.

Claim 4.1. $(\omega_i^{\ j})$ determine smooth 1-forms on U.

Proof. We first show $\mathbb{R}$-linearity. For $X_1, X_2 \in \mathfrak{X}(U)$ and $a_1, a_2 \in \mathbb{R}$,

$$\begin{aligned}\nabla_{a_1X_1+a_2X_2} E_i &= a_1\nabla_{X_1}E_i + a_2\nabla_{X_2}E_i \\ &= a_1\omega_i^{\ j}(X_1)E_j + a_2\omega_i^{\ j}(X_2)E_j \\ &= (a_1\omega_i^{\ j}(X_1) + a_2\omega_i^{\ j}(X_2))E_j.\end{aligned}$$

On the other hand, we have $\nabla_{a_1X_1+a_2X_2}E_i = \omega_i^{\ j}(a_1X_1+a_2X_2)E_j$. This proves $\mathbb{R}$-linearity.

Next, we show $C^\infty(U)$-linearity. Letting $f \in C^\infty(U)$, one has

$$\nabla_{fX}E_i = f\nabla_X E_i = f\omega_i^{\ j}(X)E_j.$$

On the other hand, $\nabla_{fX}E_i = \omega_i^{\ j}(fX)E_j$. This proves $C^\infty(U)$-linearity. $\square$

The smooth 1-forms $(\omega_i^{\ j})$ on U are called connection 1-forms, or simply connection forms.

4.2 Torsion 2-Forms

Similar to connection forms, we can write

$$T(X,Y) = T^i(X,Y)E_i$$

for all $X, Y \in \mathfrak{X}(U)$, where $T^i(X,Y) \in C^\infty(U)$ depends on X and Y.

Claim 4.2. (T^k) determine smooth 2-forms on U.

Proof. By a slight adjustment of Proposition 3.9 to non-coordinate local frames, we have

$$T^k(X,Y) = X^iY^j(\Gamma^k_{ij} - \Gamma^k_{ji} - c^k_{ij}),$$

where c^k_{ij} defined by $[E_i, E_j] = c^k_{ij}E_k$ are called structure coefficients [10]. It is clear from this expression that $T^k(X,Y)$ is bilinear over $\mathbb{R}$ and $C^\infty(U)$, and is smooth on U.

Antisymmetry follows from a direct calculation:

$$\begin{aligned} T^k(Y,X) &= Y^iX^j(\Gamma^k_{ij} - \Gamma^k_{ji}) \\ &= Y^jX^i(\Gamma^k_{ji} - \Gamma^k_{ij}) \\ &= -X^iY^j(\Gamma^k_{ij} - \Gamma^k_{ji}) \\ &= -T^k(X,Y), \end{aligned}$$

where we renamed the indices in the second line. Thus, T^k defines a smooth 2-form. □

The smooth 2-forms (T^k) on U are called torsion 2-forms, or simply torsion forms. Before stating the first main result of this subsection, we need a lemma.

Lemma 4.3. Let M be a smooth manifold and ω be a smooth 1-form. For $X, Y \in \mathfrak{X}(M)$,

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

Proof. Let $(U,(x^i))$ be a coordinate chart for M. Then, writing $\omega = \omega_i\, dx^i$, we have $d\omega = \partial_j\omega_i\, dx^j \wedge dx^i$. This gives

$$\begin{aligned} d\omega(X,Y) &= \partial_j\omega_i\,(dx^j \wedge dx^i)(X,Y) \\ &= \partial_j\omega_i(dx^j(X)\,dx^i(Y) - dx^j(Y)\,dx^i(X)) \\ &= \partial_j\omega_i(X^jY^i - Y^jX^i) \\ &= X^j\partial_j\omega_iY^i - Y^j\partial_j\omega_iX^i \\ &= (X\omega_i)Y^i - (Y\omega_i)X^i. \end{aligned}$$

On the other hand,

$$\begin{aligned}X\omega(Y) - Y\omega(X) &= X\omega_i\, dx^i(Y) - Y\omega_i\, dx^i(X)\\ &= X(\omega_i Y^i) - Y(\omega_i X^i)\\ &= (X\omega_i)Y^i + \omega_i XY^i - (Y\omega_i)X^i - \omega_i YX^i,\end{aligned}$$

and

$$\omega([X,Y]) = \omega_i\, dx^i(XY - YX) = \omega_i(XY^i - YX^i).$$

Hence, we obtain

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

□

Proposition 4.4 (Cartan's First Structure Equation)**.** The connection 1-forms and the torsion 2-forms satisfy the following equation:

$$d\varepsilon^j = \varepsilon^i \wedge \omega_i^{\ j} + T^j.$$

Proof. We start by noting that, similar to coordinate frames, $Y \in \mathfrak{X}(U)$ can be expressed as $Y = \varepsilon^i(Y)E_i$. Then,

$$\begin{aligned}\nabla_X Y = \nabla_X(\varepsilon^i(Y)E_i) &= (X\varepsilon^i(Y))E_i + \varepsilon^i(Y)\nabla_X E_i\\ &= (X\varepsilon^i(Y))E_i + \varepsilon^i(Y)\omega_i^{\ j}(X)E_j.\end{aligned}$$

Similarly,

$$\nabla_Y X = (Y\varepsilon^i(X))E_i + \varepsilon^i(X)\omega_i^{\ j}(Y)E_j.$$

Combining,

$$\begin{aligned}T(X,Y) &= \nabla_X Y - \nabla_Y X - [X,Y]\\ &= (X\varepsilon^i(Y))E_i + \varepsilon^i(Y)\omega_i^{\ j}(X)E_j - (Y\varepsilon^i(X))E_i - \varepsilon^i(X)\omega_i^{\ j}(Y)E_j - [X,Y]\\ &= (X\varepsilon^i(Y))E_i + \varepsilon^i(Y)\omega_i^{\ j}(X)E_j - (Y\varepsilon^i(X))E_i - \varepsilon^i(X)\omega_i^{\ j}(Y)E_j - \varepsilon^i([X,Y])E_i\\ &= \left(X\varepsilon^i(Y) - Y\varepsilon^i(X) - \varepsilon^i([X,Y]) + \varepsilon^j(Y)\omega_j^{\ i}(X) - \varepsilon^j(X)\omega_j^{\ i}(Y)\right)E_i\\ &= (d\varepsilon^i + \omega_j^{\ i} \wedge \varepsilon^j)(X,Y)E_i,\end{aligned}$$

where we used Lemma 4.3 and the definition of the wedge product of 1-forms in the last equality. Writing $T(X,Y) = T^i(X,Y)E_i$, one has

$$T^i(X,Y)E_i = (d\varepsilon^i + \omega_j^{\ i} \wedge \varepsilon^j)(X,Y)E_i.$$

Since the above holds for all $X, Y \in \mathfrak{X}(U)$, we have $T^i = d\varepsilon^i + \omega_j^{\ i} \wedge \varepsilon^j$. The anticommutativity of the wedge product finishes the proof. $\square$

4.3 Curvature 2-Forms

Similar to torsion and connection forms, we can write $R(X,Y)E_j$ as a linear combination of the local frame as

$$R(X,Y)E_i = \Omega_i^{\ j}(X,Y)E_j$$

for all $X, Y \in \mathfrak{X}(U)$, where $\Omega_i^{\ j}(X,Y) \in C^\infty(U)$ depends on X and Y.

Claim 4.5. $(\Omega_i^{\ j})$ determine smooth 2-forms on U.

Proof. $\mathbb{R}$-bilinearity follows from $\mathbb{R}$-bilinearity of R in X and Y. For $C^\infty(U)$-bilinearity, let $f, g \in C^\infty(U)$. By $C^\infty(U)$-linearity of R in each argument,

$$\begin{aligned}
\Omega_i^{\ j}(fX + gY, Z)E_j &= R(fX + gY, Z)E_i \\
&= (fR(X,Z) + gR(Y,Z))E_i \\
&= (f\Omega_i^{\ j}(X,Z) + g\Omega_i^{\ j}(Y,Z))E_j.
\end{aligned}$$

Linearity in the second argument is similar. Finally, antisymmetry follows from the antisymmetry of R, i.e., $R(X,Y)Z = -R(Y,X)Z$. $\square$

The smooth 2-forms $(\Omega_i^{\ j})$ on U are called curvature 2-forms, or simply curvature forms. We can now state the second main result of this section.

Proposition 4.6 (Cartan's Second Structure Equation)**.** The curvature 2-forms satisfy the following equation:

$$\Omega_i^{\ j} = d\omega_i^{\ j} - \omega_i^{\ k} \wedge \omega_k^{\ j},$$

where $\omega_i^{\ j}$ are the connection 1-forms.

Proof. We have

$$\begin{aligned}
R(X,Y)E_i &= \nabla_X \nabla_Y E_i - \nabla_Y \nabla_X E_i - \nabla_{[X,Y]} E_i \\
&= \nabla_X (\omega_i{}^j(Y) E_j) - \nabla_Y (\omega_i{}^j(X) E_j) - \omega_i{}^j([X,Y]) E_j \\
&= X(\omega_i{}^j(Y)) E_j + \omega_i{}^j(Y) \nabla_X E_j - Y(\omega_i{}^j(X)) E_j - \omega_i{}^j(X) \nabla_Y E_j - \omega_i{}^j([X,Y]) E_j \\
&= (X(\omega_i{}^j(Y)) - Y(\omega_i{}^j(X)) - \omega_i{}^j([X,Y])) E_j - (\omega_i{}^j(X)\omega_j{}^k(Y) - \omega_i{}^j(Y)\omega_j{}^k(X)) E_k \\
&= (d\omega_i{}^k - \omega_i{}^j \wedge \omega_j{}^k)(X,Y) E_k \\
&= \Omega_i{}^k(X,Y) E_k,
\end{aligned}$$

where we have used Lemma 4.3 and renamed the indices in the fourth line. □

4.4 Non-Metricity 1-Forms

Similar as before, in a local frame (E_i), we can write

$$Q(X, E_i, E_j) = Q_{ij}(X).$$

Claim 4.7. (Q_{ij}) determine smooth 1-forms on U.

Proof. Similar to (3.17), we have

$$\begin{aligned}
Q_{ij}(X) = Q(X, E_i, E_j) &= -X g_{ij} + g(\nabla_X E_i, E_j) + g(E_i, \nabla_X E_j) \\
&= -X g_{ij} + g(\omega_i{}^k(X) E_k, E_j) + g(E_i, \omega_j{}^k(X) E_k) \\
&= -X g_{ij} + \omega_i{}^k(X) g(E_k, E_j) + \omega_j{}^k(X) g(E_i, E_k) \\
&= -X g_{ij} + \omega_i{}^k(X) g_{kj} + \omega_j{}^k(X) g_{ik}.
\end{aligned}$$

The above clearly shows that Q_{ij} is $\mathbb{R}$-linear and $C^\infty(U)$-linear. □

The smooth 1-forms Q_{ij} on U are called non-metricity 1-forms, or simply non-metricity forms [1].

5 General Relativity and Metric-Affine Gravity

In this final section, we will briefly compare the geometrical frameworks of general relativity (GR) and metric-affine gravity (MAG). From a purely mathematical point of view, the field equations of GR in vacuum without a cosmological constant are derived by extremizing the functional

$$S = \frac{1}{16\pi}\int_M R\,\mathrm{dvol}_g,$$

where $R = g^{ij}R^l{}_{ilj} =: g^{ij}R_{ij}$ is the curvature scalar, dvol_g is the volume form and M is the region of integration. By Proposition 3.24, $R^l{}_{ilj}$ only depends on the connection coefficients. Since we also multiply with g^{ij}, the curvature scalar R depends on the connection and the metric. Accordingly, the functional S depends on the connection and the metric. However, in GR, one assumes that the connection is the Levi-Civita connection. We have seen that the connection coefficients of the Levi-Civita connection are given in terms of the components of the metric. Consequently, R only depends on the metric. Thus, one extremizes the functional S only with respect to the metric. The metric extremizing S is given implicitly by [14]

$$R_{ij} - \frac{1}{2}Rg_{ij} = 0.$$

The above equations are called the Einstein Field Equations in vacuum without a cosmological constant.

In MAG, the field equations are again obtained by a functional given by [2]

$$S = \int_M L(g_{ij}, \varepsilon^i, Q_{ij}, T^i, \Omega_i{}^j),$$

where L is now a differential form. In MAG, one assumes that the connection is not necessarily the Levi-Civita connection of the metric. Accordingly, the metric and the connection are independent objects. An independent connection naturally introduces torsion and non-metricity. By Cartan's structure equations, the only independent variables are the metric, the coframe, and the connection 1-forms. Extremizing S with respect to these variables yields [2]

$$D\left(\frac{\partial L}{\partial Q_{ij}}\right) + \frac{\partial L}{\partial g_{ij}} = 0$$
$$D\left(\frac{\partial L}{\partial T^i}\right) + \frac{\partial L}{\partial \varepsilon^i} = 0$$
$$D\left(\frac{\partial L}{\partial \Omega_i{}^j}\right) + \frac{\partial L}{\partial \omega_i{}^j} = 0,$$

where D is called the exterior covariant derivative [5]. These equations are known as the field equations of MAG in vacuum.

As a short summary, metric-affine theories provide a richer physical and geometrical framework by relaxing the assumption that the connection is the Levi-Civita connection. This modification leads to new field equations that could potentially address issues that general relativity cannot answer.

References

[1] F. W. Hehl et al. Metric-affine gauge theory of gravity: Field equations, noether identities, world spinors, and breaking of dilation invariance. *Physics Reports*, 258:1–171, 1995.

[2] F. Gronwald. Metric-Affine Gauge Theory of Gravity: I. Fundamental Structures and Field Equations. *Int. J. Mod. Phys. D*, 6(3):263–303, 1997.

[3] D. Iosifidis. *Metric-Affine Gravity and Cosmology: Aspects of Torsion and Non-Metricity in Gravity Theories.* PhD thesis, Aristotle University of Thessaloniki, 2019.

[4] E. Jensko. *A Unified Approach to Geometric Modifications of Gravity.* PhD thesis, University College London, 2023.

[5] A. Jiménez-Cano. *Metric-Affine Gauge theories of gravity: Foundations and new insights.* PhD thesis, Universidad de Granada, 2021.

[6] J. Jost. *Riemannian Geometry and Geometric Analysis.* Springer, 7th edition, 2017.

[7] J. M. Lee. *Introduction to Topological Manifolds.* Springer, 2nd edition, 2011.

[8] J. M. Lee. *Introduction to Smooth Manifolds.* Springer, 2nd edition, 2012.

[9] J. M. Lee. *Introduction to Riemannian Manifolds.* Springer, 2nd edition, 2018.

[10] M. Nakahara. *Geometry, Topology and Physics.* IoP Publishing, 2nd edition, 2003.

[11] K. S. Stelle. Renormalization of higher-derivative of quantum gravity. *Physical Review D*, 16(4):953–969, 1977.

[12] L. W. Tu. *An Introduction to Manifolds.* Springer, 2nd edition, 2011.

[13] L. W. Tu. *Differential Geometry: Connections, Curvature, and Characteristic Classes.* Springer, 1st edition, 2017.

[14] R. M. Wald. *General Relativity.* The University of Chicago Press, 1st edition, 1984.

[15] S. Weinberg. The cosmological constant problem. *Reviews of Modern Physics*, 61(1):1–23, 1989.