Differential Geometry on Foliations and Submanifolds

Niels Slotboom
slotboom.n@gmail.com

Contents

1 Introduction

To formulate general relativity as a Hamiltonian field theory within the ADM formalism, one must choose a direction of evolution—essentially a preferred timelike parameter within spacetime. This is achieved through the use of a hypersurface foliation. These notes aim to present the concepts I have learned in the process of achieving my goal of deriving the ADM decomposition of the Einstein-Hilbert action, a result often referred to as the ADM action. This decomposition serves as the starting point for the Hamiltonian perspective on general relativity.

In working towards this goal, the notes introduce a variety of essential geometric structures. We begin with more general theory by defining submanifolds, pushforwards, pullbacks, and projectors onto tangent and normal spaces. After a brief exploration of vector bundles, the concept of foliations is introduced, followed by a detailed look at the structures they carry: the induced metric, projected covariant derivatives, as well as the intrinsic and extrinsic curvature tensors. Additionally, we decompose the tangential part of the ambient curvature into intrinsic and extrinsic components in the form of the Gauss-Codazzi equation. This result is what will enable us to derive the ADM action.

While I've aimed to make these notes as approachable and didactic as possible, they should remain to be seen as part of my personal learning journey. In some cases the path to a given result might not be the most direct, but it will always be my own. Rather than relying on existing derivations, I've opted to find them myself.

2 Basics

2.1 Definition of a Submanifold

In the following, $\mathcal{M}$ will denote a ((pseudo)-Riemannian) m-manifold, and x^μ local coordinates on it. We denote the components of the metric tensor on $\mathcal{M}$ by $g_{\mu\nu}(x)$, with $\mu, \nu, ...$ indicating coordinate indices ranging from $1, ..., m$. The metric tensor defines a line element on $\mathcal{M}$,

$$ds^2 = g = g_{\mu\nu}\, dx^\mu \otimes dx^\nu. \tag{2.1.1}$$

A *submanifold* of $\mathcal{M}$ is a subset $\mathcal{S} \subset \mathcal{M}$ that is itself a smooth manifold, say of dimension $s = \dim \mathcal{S} \leq m$. Here it is understood that the topological structure on $\mathcal{S}$ is the induced subtopology from $\mathcal{M}$. Additional structures—such as the induced metric, connections and curvature—will be introduced gradually in the following sections.

We call the quantity

$$\operatorname{codim}_{\mathcal{M}} \mathcal{S} = \dim \mathcal{M} - \dim \mathcal{S} \tag{2.1.2}$$

the *codimension* of $\mathcal{S}$. This is a simple but useful notion that conveys "how many dimensions less" the submanifold has in comparison to the ambient manifold $\mathcal{M} \supset \mathcal{S}$, which sometimes is more important than the dimension itself.

For example, a *hypersurface* is, by definition, a submanifold $\mathcal{S} \subset \mathcal{M}$ of codimension 1, and will be particularly important in the development of foliations and the ADM formalism. A simple way of generating (smooth) hypersurfaces is by considering a function $f : \mathcal{M} \to \mathbb{R}$ and constructing a level surface as

$$\mathcal{S} = f^{-1}(\{0\}) = \{p \in \mathcal{M} \mid f(p) = 0\} \subset \mathcal{M}. \tag{2.1.3}$$

For $\mathcal{S}$ defined in this way to be a proper, smooth submanifold, 0 has to be a regular value of f. This means that

$$df|_{\mathcal{S}} \neq 0 \tag{2.1.4}$$

or in words, the differential (or equivalently, the gradient) of f must not vanish on $\mathcal{S}$.

More generally, given a collection of smooth functions $f_i : \mathcal{M} \to \mathbb{R}$ for $i = 1, ..., k$, the common level set

$$\mathcal{S} = \bigcap_i f_i^{-1}(\{0\}) \tag{2.1.5}$$

is a submanifold of codimension k, provided the f_i are functionally independent, $\mathcal{S}$ is non-empty and 0 is a regular value of the map $f = (f_1, ..., f_k)$.

2.2 Induced Metric

So far, we have established that a submanifold $\mathcal{S} \subset \mathcal{M}$ inherits its topological structure from the subspace topology induced by $\mathcal{M}$. This makes $\mathcal{S}$ a topological manifold. However, since we are studying submanifolds of a (pseudo-)Riemannian manifold, it should come to no surprise that we also wish for $\mathcal{S}$ to carry a (pseudo-)Riemannian structure.

As a smooth manifold in its own right, $\mathcal{S}$ admits local coordinates y^i, where $i = 1, ..., s = \dim \mathcal{S}$. In principle, endowing $\mathcal{S}$ with a (pseudo-)Riemannian metric is straightforward: one simply picks a symmetric tensor field γ taking values in $\gamma_p \in T_p^*\mathcal{S} \otimes T_p^*\mathcal{S}$ with the desired signature and defines the corresponding line element

$$d\bar{s}^2 = \gamma = \gamma_{ij}\, dy^i \otimes dy^j. \tag{2.2.1}$$

While this makes $\mathcal{S}$ a (pseudo-)Riemannian manifold, it does so independently of the geometry on the ambient manifold $\mathcal{M} \supset \mathcal{S}$—it disregards the structure already present on $\mathcal{M}$.

Yet $\mathcal{S}$ is fully contained in $\mathcal{M}$; any point $p \in \mathcal{S}$ is also a point of $\mathcal{M}$. That much is obvious—but it matters, because the metric on $\mathcal{M}$ carries geometric information we may wish to preserve. What, after all, does a metric provide? In particular, it can be used to assign lengths to curves. If we consider a path lying in $\mathcal{S}$, then this path is also a path in $\mathcal{M}$. Naturally, we would want its length to be the same whether we compute it using the metric $ds^2 = g$ of the ambient manifold or the metric $d\bar{s}^2 = \gamma$ of the submanifold.

This requirement reflects a choice, but a canonical one: the geometry of $\mathcal{S}$ should arise from that of $\mathcal{M}$ by restriction, not by arbitrary redefinition. That is, the structures on $\mathcal{M}$ should be *induced* on $\mathcal{S}$, not constructed from scratch.

Requiring that all lengths of curves in $\mathcal{S}$ agree under both metrics leads to the precise condition that

$$d\bar{s}^2 = ds^2|_{\mathcal{S}}. \tag{2.2.2}$$

This is no longer just a heuristic, but an equation—it specifies how the metric on the submanifold must be related to the metric of the ambient space.

Though precise, the definition in relation (2.2.2) is not especially practical for computations. Every time we wish to measure the length of a path on $\mathcal{S}$, we must return to the structure on $\mathcal{M}$ and consult what *it* "thinks" the length is. We would prefer to treat $\mathcal{S}$ as a manifold in its own right, without this constant indirection via $\mathcal{M}$. To achieve that, we need explicit knowledge of the components γ_{ij} of the *induced metric* on $\mathcal{S}$—the symmetric tensor satisfying

$$\gamma_{ij}\, dy^i \otimes dy^j = d\bar{s}^2 \overset{!}{=} ds^2|_{\mathcal{S}} = \left(g_{\mu\nu}\, dx^\mu \otimes dx^\nu\right)|_{\mathcal{S}}. \tag{2.2.3}$$

The remainder of this section is devoted to deriving a general formula for γ_{ij} in terms of the ambient metric $g_{\mu\nu}$.

To do so, observe that for any point $p \in \mathcal{S}$, we have two coordinate systems at our disposal: the intrinsic coordinates $y^i(p)$ on $\mathcal{S}$, and the ambient coordinates $x^\mu(p)$ from $\mathcal{M}$. More formally, since $\mathcal{S} \subset \mathcal{M}$, we obtain an *injective coordinate map*

$$y^i \mapsto x^\mu(y^i), \tag{2.2.4}$$

expressing the ambient coordinates as functions of the submanifold coordinates. This has a consequence for the differential of x^μ restricted to $\mathcal{S}$, namely

$$dx^\mu|_{\mathcal{S}} = d(x^\mu(y^i))|_{\mathcal{S}} = \frac{\partial x^\mu}{\partial y^i}\, dy^i =: E^\mu_i\, dy^i, \tag{2.2.5}$$

where we defined the so-called *pushforward matrix* or *vielbein*

$$E^\mu_i = \frac{\partial x^\mu}{\partial y^i}. \tag{2.2.6}$$

Using this relation, we can rewrite the restriction of the line element as

$$\begin{aligned} \gamma_{ij}\, dy^i \otimes dy^j \,=\, d\bar{s}^2 \,=\, ds^2|_{\mathcal{S}} \\ &= \big(g_{\mu\nu}\, dx^\mu \otimes dx^\nu\big)|_{\mathcal{S}} \\ &= g_{\mu\nu}(x^\mu(y^i))E^\mu_i\, dy^i \otimes E^\nu_j\, dy^j \\ &= \big(E^\mu_i E^\nu_j g_{\mu\nu}\big)\, dy^i \otimes dy^j \end{aligned} \tag{2.2.7}$$

From this, we identify the components of the induced metric as

$$\gamma_{ij} = E^\mu_i E^\nu_j g_{\mu\nu}|_{\mathcal{S}} = \frac{\partial x^\mu}{\partial y^i}\frac{\partial x^\nu}{\partial y^j} g_{\mu\nu}|_{\mathcal{S}}. \tag{2.2.8}$$

This gives a general formula for the induced metric on a submanifold, expressed in terms of the ambient metric and a choice of compatible[1] coordinate systems on both $\mathcal{M}$ and $\mathcal{S}$.

This expression for γ_{ij} guarantees that the length of any path measured intrinsically on $\mathcal{S}$, via the line element $d\bar{s}^2$, matches the length of the same path computed in the ambient manifold via ds^2. It is, of course, a coordinate-dependent representation—whereas the earlier definition in equation (2.2.2) was fully coordinate-independent.

2.3 Example Submanifold: $S^2 \subset \mathbb{R}^3$

Before diving into more abstract constructions, let us ground our intuition with a concrete example of a manifold and a natural submanifold. To this end, we consider the flat Euclidean space $\mathcal{M} = \mathbb{R}^3$, with Cartesian coordinates $x^\mu = (x, y, z) \in \mathbb{R}^3$ and endowed with the Euclidean metric

$$ds^2 = dx^2 + dy^2 + dz^2 \quad \Leftrightarrow \quad g_{\mu\nu} = \delta_{\mu\nu}. \tag{2.3.1}$$

As our submanifold $\mathcal{S}$, we consider the 2-sphere S^2, defined by

$$S^2 = \big\{(x, y, z) \in \mathbb{R}^3 \,\big|\, x^2 + y^2 + z^2 = 1\big\}. \tag{2.3.2}$$

It is clear that this is a submanifold, as it is the preimage of $\{0\}$ under the map

$$f(x, y, z) = x^2 + y^2 + z^2 - 1, \tag{2.3.3}$$

which has 0 as a regular value, i.e., the gradient vanishes nowhere where $f(x, y, z) = 0$.

[1]Obviously, the coordinate charts must overlap appropriately and be smooth enough for the construction to be valid.

The 2-sphere is most conveniently parameterised by two angular coordinates $y^i = (\theta, \varphi) \in (0, \pi) \times (0, 2\pi)$, with the embedding into the ambient space $\mathbb{R}^3$ given by

$$\begin{aligned} x(\theta, \varphi) &= \sin\theta \cos\varphi, \\ y(\theta, \varphi) &= \sin\theta \sin\varphi, \\ z(\theta, \varphi) &= \cos\theta. \end{aligned} \tag{2.3.4}$$

We now compute the induced metric on S^2 using two methods: first by means of the pushforward matrix, and then in a more direct way involving differentials of the above relations.

The components of $E_i^\mu = \frac{\partial x^\mu}{\partial y^i}$ are given by

$$\begin{aligned} E_\theta^x &= \frac{\partial x}{\partial \theta} = \cos\theta\cos\varphi, & E_\varphi^x &= \frac{\partial x}{\partial \varphi} = -\sin\theta\sin\varphi, \\ E_\theta^y &= \frac{\partial y}{\partial \theta} = \cos\theta\sin\varphi, & E_\varphi^y &= \frac{\partial y}{\partial \varphi} = \sin\theta\cos\varphi, \\ E_\theta^z &= \frac{\partial z}{\partial \theta} = -\sin\theta, & E_\varphi^z &= \frac{\partial z}{\partial \varphi} = 0. \end{aligned} \tag{2.3.5}$$

This allows us to compute the components of the induced metric as

$$\begin{aligned} \gamma_{\theta\theta} &= E_\theta^\mu E_\theta^\mu \delta_{\mu\nu} = E_\theta^x E_\theta^x + E_\theta^y E_\theta^y + E_\theta^z E_\theta^z \\ &= \cos^2\theta\cos^2\varphi + \cos^2\theta\sin^2\varphi + \sin^2\theta \\ &= 1, \end{aligned} \tag{2.3.6}$$

$$\begin{aligned} \gamma_{\varphi\varphi} &= E_\varphi^\mu E_\varphi^\mu \delta_{\mu\nu} = E_\varphi^x E_\varphi^x + E_\varphi^y E_\varphi^y + E_\varphi^z E_\varphi^z \\ &= \sin^2\theta\sin^2\varphi + \sin^2\theta\cos^2\varphi \\ &= \sin^2\theta, \end{aligned} \tag{2.3.7}$$

$$\begin{aligned} \gamma_{\theta\varphi} = \gamma_{\varphi\theta} &= E_\theta^x E_\varphi^x + E_\theta^y E_\varphi^y + E_\theta^z E_\varphi^z \\ &= -\sin\theta\cos\theta\sin\varphi\cos\varphi + \sin\theta\cos\theta\sin\varphi\cos\varphi \\ &= 0. \end{aligned} \tag{2.3.8}$$

Using this, the induced line element becomes

$$d\bar{s}^2 = \gamma_{ij}\, dy^i \otimes dy^j = d\theta \otimes d\theta + \sin^2\theta\, d\varphi \otimes d\varphi \tag{2.3.9}$$

—an anticipated result, the standard metric on S^2.

We now briefly go over how one can compute the induced metric slightly more directly by considering the differentials of the embeddings (2.3.4). These differentials read

$$\begin{aligned} dx|_{S^2} &= \cos\theta\cos\varphi\, d\theta - \sin\theta\sin\varphi\, d\varphi, \\ dy|_{S^2} &= \cos\theta\sin\varphi\, d\theta + \sin\theta\cos\varphi\, d\varphi, \\ dz|_{S^2} &= -\sin\theta d\theta. \end{aligned} \tag{2.3.10}$$

We can then simply insert these relations into the Euclidean line element (2.3.1) to find (writing $dx^\mu \otimes dx^\nu$ as $dx^\mu dx^\nu$ for brevity)

$$
\begin{aligned}
d\bar{s}^2 = ds^2|_{S^2} &= dx^2|_{S^2} + dy^2|_{S^2} + dz^2|_{S^2} \\
&= (\cos\theta\cos\varphi\, d\theta - \sin\theta\sin\varphi\, d\varphi)^2 \\
&\quad +(\cos\theta\sin\varphi\, d\theta + \sin\theta\cos\varphi\, d\varphi)^2 \\
&\quad +(-\sin\theta d\theta)^2 \\
&= \cos^2\theta\cos^2\varphi\, d\theta^2 - \cancel{2\cos\theta\sin\theta\cos\varphi\sin\varphi\, d\theta d\varphi} + \sin^2\theta\sin^2\varphi\, d\varphi^2 \\
&\quad + \cos^2\theta\sin^2\varphi\, d\theta^2 + \cancel{2\cos\theta\sin\theta\cos\varphi\sin\varphi\, d\theta d\varphi} + \sin^2\theta\cos^2\varphi\, d\varphi^2 \\
&\quad + \sin^2\theta\, d\theta^2 \\
&= (\cos^2\theta(\cos^2\varphi + \sin^2\varphi) + \sin^2\theta)\, d\theta^2 + \sin^2\theta(\sin^2\varphi + \cos^2\varphi)\, d\varphi^2 \\
&= d\theta^2 + \sin^2\theta\, d\varphi^2.
\end{aligned}
\tag{2.3.11}
$$

This alternative method avoids explicitly constructing the pushforward matrix, relying instead on computing the differentials of the embedding directly and inserting them into the ambient metric —though of course, it is fully equivalent to the first approach. *Which* of the two approaches is better comes down to being mostly situational, and naturally, personal preference.

3 Tangent Spaces, Pushforwards, Pullbacks, and Projections

3.1 Preliminaries: Maps Between Manifolds and Their Derivatives

Before diving into the formalism of pushforwards and pullbacks, which is the goal of this section, let us first understand the geometric motivation behind them.

Suppose we have two smooth manifolds $\mathcal{M}$ and $\mathcal{N}$, with dimensions m and n, and $\varphi : \mathcal{M} \to \mathcal{N}$ a smooth map between them. This map assigns to every point $p \in \mathcal{M}$ a point $\varphi(p) \in \mathcal{N}$. But more than just matching up points, φ also relates the local *geometry* near those points, i.e., the spatial relationships between nearby points.

To illustrate this, imagine a flexible sheet of rubber ($\mathcal{M}$) that we smoothly press onto another, possibly curved surface $\mathcal{N}$. Every point on the rubber sheet ends up on some point of the surface, but the way the rubber sheet bends or stretches as it moves gives us more than just a positional correspondence—it tells us how *directions* and *infinitesimal displacements* on $\mathcal{M}$ are transformed under the map. For example, a tiny arrow drawn on the rubber sheet (a tangent vector) might point in some new direction or have a different length when the rubber sheet is fitted onto the surface.

This is the essence of what the pushforward captures: it tells us how tangent vectors on $\mathcal{M}$ —representing directions of motion away from a point—are mapped to tangent vectors on $\mathcal{N}$. Formally, the *pushforward of* φ, denoted by φ_*, is a map between tangent spaces,

$$
\varphi_* : T_p\mathcal{M} \to T_{\varphi(p)}\mathcal{N}. \tag{3.1.1}
$$

We now develop this map explicitly and precisely.

Definition: (Pushforward) We begin by recalling what we want to construct: given a smooth map $\varphi : \mathcal{M} \to \mathcal{N}$, we aim to associate to any vector $X \in T_p\mathcal{M}$ a vector $\varphi_* X \in T_{\varphi(p)}\mathcal{N}$, which we call the *pushforward* of X under φ. This should represent the idea of how the direction X "looks" after the map φ has distorted the space onto $\mathcal{N}$.

But how do we compare directions at two different points of two different manifolds? We us the fact that vectors act on functions as differential operators—they are directional derivatives. A vector $X \in T_p\mathcal{M}$ acting on a function on $\mathcal{M}$ tells us how it changes as we move in the direction of X. Similarly, $\varphi_* X \in T_{\varphi(p)}\mathcal{N}$ should tell us how a function on $\mathcal{N}$ changes in the "corresponding direction" under φ.

There is a caveat, though; we cannot directly evaluate how X acts on a function on $\mathcal{N}$, because X lives on $\mathcal{M}$. But we *can* make any function on $\mathcal{N}$ into a function on $\mathcal{M}$ by composing it with φ. That is, given any smooth function $f : \mathcal{N} \to \mathbb{R}$, the composition $f \circ \varphi : \mathcal{M} \to \mathbb{R}$ is a new function defined on $\mathcal{M}$ to which we *can* apply X.

The function $f \circ \varphi$ is, in essence, just f, but with the codomain rearranged according to φ. The function values don't change—only *where* in the domain they are sampled from does. This captures the distortion φ induces: how orientation, stretching, or compression affect the function's positions of values (but not the values themselves) when viewed on $\mathcal{M}$. Hence, taking $X[f \circ \varphi]$ tells us how f changes along the image of the direction X under the map φ.

This leads us naturally to the definition

$$(\varphi_* X)[f] := X[f \circ \varphi], \quad \forall f \in C^\infty(\mathcal{N}). \tag{3.1.2}$$

That is, $\varphi_* X$ is the unique vector at $\varphi(p)$ whose action on any function f on $\mathcal{N}$ agrees with the action of X on the pulled-back version of f. This expression captures exactly what the pushforward is doing: it tells us how to "translate" a directional derivative on $\mathcal{M}$ to one on $\mathcal{N}$, while preserving how it affects functions.

This is all very abstract, so let us make it more concrete with an analogy. Let us think of $\mathcal{M}$ as the flat Mercator map, a rectangle showing latitude and longitude lines. Then, we think of $\mathcal{N}$ as the actual globe, a 2-sphere. The map $\varphi : \mathcal{M} \to \mathcal{N}$ we envision as the wrapping map, which tells us how to take the rectangle and bend it into a globe. Suppose f assigns temperature values across the globe—a real-valued function on $\mathcal{N}$. Then $f \circ \varphi$ is the pulled-back function: it contains the same information, the same temperature values, but now shown on the Mercator map instead of the globe. This function $f \circ \varphi$ lets us look at the globe's data f as if it lived on the Mercator map. We then ask: how does the temperature change when moving along some vector X on that map? This is what $X[f \circ \varphi]$ computes: how the pulled-back temperature map changes as we follow X on the rectangle.

Now, on the globe, there should be a corresponding direction, $\varphi_* X$, along which the same change occurs—after all, we have the same function values, just arranged differently in a smooth way. This is the direction we have to move towards on the sphere to see the same change in temperature. This should not just happen for this one temperature map—for any function f on the globe, the pushforward $\varphi_* X$ must give us the corresponding direction whose action on f replicates the directional derivative seen on the Mercator map. This precisely is what the definition (3.1.2) encodes.

For explicit calculations, it is typically more convenient to move away from the abstract, coordinate-independent definition (3.1.2), and instead adopt a coordinate-based approach. To this end, let $p \in \mathcal{M}$ a point, with local coordinates y^α around p, and let x^μ be local coordinates on $\mathcal{N}$ around the image point $\varphi(p)$. Then φ induces a local map between coordinate systems,

$$x^\mu(y^\alpha) := x^\mu(\varphi(p(y^\alpha))) : \mathbb{R}^m \to \mathbb{R}^n, \tag{3.1.3}$$

which is differentiable in the standard sense. That is, composing the chart inverse on $\mathcal{M}$, the map φ, and the chart on $\mathcal{N}$, we obtain a smooth map between subsets of Euclidean space.

To derive the coordinate expression for the pushforward, consider the coordinate bases

$$\left\{ \partial_\alpha = \frac{\partial}{\partial y^\alpha} \right\} \text{ of } \mathcal{M} \quad \text{and} \quad \left\{ \partial_\mu = \frac{\partial}{\partial x^\mu} \right\} \text{ of } \mathcal{N}. \tag{3.1.4}$$

A vector $X \in T_p\mathcal{M}$ can be written as $X = X^\alpha \partial_\alpha$, and its pushforward as $\varphi_* X = (\varphi_* X)^\mu \partial_\mu$. The definition (3.1.2) then turns into

$$
\begin{aligned}
(\varphi_* X)^\mu \frac{\partial}{\partial x^\mu} f(x) &= X^\alpha \frac{\partial}{\partial y^\alpha} f(x(y)) \\
&= X^\alpha \frac{\partial x^\mu}{\partial y^\alpha} \frac{\partial}{\partial x^\mu} f(x(y)),
\end{aligned}
\tag{3.1.5}
$$

where we applied the standard multidimensional chain rule on the right-hand side. We may hence identify the components of the pushforward $\varphi_* X$ as

$$
(\varphi_* X)^\mu = \frac{\partial x^\mu}{\partial y^\alpha} X^\alpha \tag{3.1.6}
$$

This shows that the pushforward acts linearly on the components of the vector, via the Jacobian matrix

$$
(\varphi_*)^\mu{}_\alpha = \frac{\partial x^\mu}{\partial y^\alpha} = \frac{\partial}{\partial y^\alpha} x^\mu(\varphi(p(y^\alpha))), \tag{3.1.7}
$$

i.e.

$$
(\varphi_* X)^\mu = (\varphi_*)^\mu{}_\alpha X^\alpha. \tag{3.1.8}
$$

This expression makes the pushforward entirely concrete: it is simply a linear transformation of the vector components under the coordinate ap induced by φ. In terms of implementation and computation, it behaves exactly like a Jacobian—and it reproduces the familiar coordinate transformation rule for vector components when $\varphi : \mathcal{M} \to \mathcal{M}$ is a diffeomorphism.

The ability to relate tangent vectors between the two manifolds $\mathcal{M}$ and $\mathcal{N}$ suggests that a similar association may be possible between covectors or 1-forms. We develop this now by establishing a dual relation.

Definition (Canonical Pairing) For a vector $X \in T_p\mathcal{M}$ and a 1-form $\omega \in T_p^*\mathcal{M}$, we define their *canonical pairing* (sometimes called an inner product)

$$
\langle \cdot, \cdot \rangle_{\mathcal{M}} : T_p^*\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R} \tag{3.1.9}
$$

by

$$
\langle \omega, X \rangle_{\mathcal{M}} := \omega(X). \tag{3.1.10}
$$

Here, $\omega(X)$ denotes the natural pairing between a covector and a vector—i.e., the evaluation of the linear map $\omega : T_p\mathcal{M} \to \mathbb{R}$ on the argument $X \in T_p\mathcal{M}$. In coordinate components, this reads $\omega(X) = \omega_\mu X^\mu$. Notice that this pairing is defined independently of any metric structure. We will use it below to define the pullback φ^*, the dual operation to the pushforward, acting on cotangent spaces.

Definition (Pullback) We define the *pullback* associated to φ as the unique linear map

$$
\varphi^* : T_{\varphi(p)}^*\mathcal{N} \to T_p^*\mathcal{M} \tag{3.1.11}
$$

satisfying

$$
\langle \varphi^*\omega, X \rangle_{\mathcal{M}} = \langle \omega, \varphi_* X \rangle_{\mathcal{N}}, \quad \forall X \in T_p\mathcal{M}, \quad \omega \in T_{\varphi(p)}^*\mathcal{N}. \tag{3.1.12}
$$

That is, the canonical pairing between a covector and a vector is preserved under the pushforward-pullback action of φ. This is a natural requirement: having defined how directions deform under φ, we want to ensure that objects measuring those directions (i.e. 1-forms) adapt in a compatible way. Equivalently, this can be written as

$$(\varphi^*\omega)(X) = \omega(\varphi_* X), \quad \forall X \in T_p\mathcal{M}, \quad \omega \in T^*_{\varphi(p)}\mathcal{N}. \tag{3.1.13}$$

Another perspective is the following: a 1-form projects a component from a vector. But if the vector is altered—say, by a pushforward—then extracting the "same" component now requires a different projection. The pullback gives us precisely this adjusted 1-form: one that reproduces the same scalar when applied to the original vector as the original 1-form applied to the pushforward.

As with the pushforward, let us now extract a coordinate-based expression for the pullback. This case turns out to be even simpler. Using the same coordinate setup—local coordinates y^α around $p \in \mathcal{M}$, and x^μ around $\varphi(p) \in \mathcal{N}$—we express

$$X = X^\alpha \partial_\alpha, \quad \omega = \omega_\mu \, dx^\mu. \tag{3.1.14}$$

We begin by expanding the left-hand side of the definition (3.1.12),

$$\langle \varphi^*\omega, X\rangle_{\mathcal{M}} = (\varphi^*\omega)_\alpha X^\alpha. \tag{3.1.15}$$

On the right-hand side, we compute the action of ω on the pushforward of X,

$$\langle \omega, \varphi_* X\rangle_{\mathcal{N}} = \omega_\mu (\varphi_* X)^\mu = \omega_\mu (\varphi_*)^\mu{}_\alpha X^\alpha. \tag{3.1.16}$$

Equating both sides and using the arbitrariness of X^α, we identify the components of the pullback:

$$(\varphi^*\omega)_\alpha = (\varphi_*)^\mu{}_\alpha \omega_\mu = \frac{\partial x^\mu}{\partial y^\alpha}\omega_\mu. \tag{3.1.17}$$

In words: the pullback transforms a 1-form using the same Jacobian matrix as the pushforward acts with on vectors, but the roles of the indices reversed—covariant versus contravariant transformation. Equivalently, the pullback acts via the transpose of the Jacobian of φ, in keeping with the duality between tangent and cotangent spaces.

Remark *Pullbacks of Vectors?* We have seen the pushforward as a map

$$\varphi_* : T_p\mathcal{M} \to T_{\varphi(p)}\mathcal{N}. \tag{3.1.18}$$

A natural question is whether this relationship can be reversed: can we pull back vectors from $T_{\varphi(p)}\mathcal{N}$ to $T_p\mathcal{M}$?

In general, the answer is no. Algebraically, this follows from the fact that φ_* is not necessarily invertible. As a linear map, invertibility requires its component representation

$$(\varphi_*)^\mu{}_\alpha = \frac{\partial x^\mu}{\partial y^\alpha} \tag{3.1.19}$$

to be an invertible matrix. This is only possible if $\dim \mathcal{M} = \dim \mathcal{N}$, since otherwise the Jacobian is rectangular and hence cannot be inverted. And even when the dimensions match, the Jacobian must be non-singular (i.e. its determinant must be nonzero). Thus, φ_* is only invertible when φ is a diffeomorphism—a smooth bijective map with smooth inverse.

This is the algebraic reasoning. Fortunately, there's also geometric intuition to support it. The pushforward tracks how a direction away from $p \in \mathcal{M}$ translates to a direction away from $\varphi(p) \in \mathcal{N}$, consistent with the local deformation φ induces. In that process, we may preserve the number of independent directions, or lose some—but we can never gain new ones. The rank of a linear map $\varphi_* : \mathbb{R}^m \to \mathbb{R}^n$ is at most m, so if $n > m$, its image lies in a proper subspace of $\mathbb{R}^n$. In other words, a pushforward cannot invent new degrees of freedom.

In the best-case scenario $\dim \operatorname{im} \varphi_* = m$, we retain full information and the map φ_* is invertible. But if information is lost—i.e. if φ_* collapses multiple directions into the same image—then we can no longer recover those original directions when attempting to go back. We do not know how to "blow it back up again"—that is, how to reconstruct the full-dimensional direction space of $T_p\mathcal{M}$. A helpful instance to picture is a pushforward collapsing $T_p\mathcal{M}$ onto a lower-dimensional plane inside $T_{\varphi(p)}\mathcal{N}$; there is no way to lift this back up uniquely.

Therefore, in general, it is not possible to pull back vectors from $T_{\varphi(p)}\mathcal{N}$ to $T_p\mathcal{M}$.

Generalisation to Tensors Now that we know how we should associate vectors and 1-forms/ covectors on two manifolds $\mathcal{M}$ and $\mathcal{N}$ between which we have an association of points given by $\varphi : \mathcal{M} \to \mathcal{N}$, we can generalise to tensors. Due to the difference in "functorial direction" between the pushforward and pullback, it is not possible to define a combined map for arbitrary, mixed tensors. For fully covariant or fully contravariant tensors, however, it is possible—or, if φ is a diffeomorphism. This latter case we will not consider, though—this is because it essentially reduces to a change of coordinates.

Let us begin with the pullback of covariant tensors $T \in T^{(0,s)}_{\varphi(p)}\mathcal{N}$. Such a tensor can be written as

$$T = T_{\mu_1 \dots \mu_s} \, dx^{\mu_1} \otimes \dots \otimes dx^{\mu_s}, \tag{3.1.20}$$

which allows us to naturally pair it with s vectors $X_{(i)} \in T_{\varphi(p)}\mathcal{N}$ to form a scalar

$$T\Big(X_{(1)}, \dots, X_{(s)}\Big) = T_{\mu_1 \dots \mu_s} X^{\mu_1}_{(1)} \dots X^{\mu_s}_{(s)}. \tag{3.1.21}$$

To ensure compatibility between T and its pullback $\varphi^* T \in T^{0,s}_p\mathcal{M}$, we impose that for all vectors $X_{(1)}, \dots X_{(s)} \in T_p\mathcal{M}$, we have

$$(\varphi^* T)\Big(X_{(1)}, \dots, X_{(s)}\Big) = T\Big(\varphi_* X_{(1)}, \dots, \varphi_* X_{(s)}\Big) \tag{3.1.22}$$

In words, we require the pullback of T to "have the same opinion" about a set of vectors as the original tensor has about their pushforwards—which it should, if they are to encode the "same" tensorial structure on the two spaces related by φ. It shouldn't matter whether we ask $\varphi^* T$ on the original vectors for a scalar value, or first squish the space around and then ask the same question to T, with the correspondents $\varphi_* X_{(i)}$ as arguments.

Of course, this definition can also be given in terms of coordinate components. Expanding both sides of the equation, we get

$$(\varphi^* T)_{\alpha_1 \dots \alpha_s} X^{\alpha_1}_{(1)} \dots X^{\alpha_s}_{(s)} = T_{\mu_1 \dots \mu_s} (\varphi_*)^{\mu_1}{}_{\alpha_1} X^{\alpha_1}_{(1)} \dots (\varphi_*)^{\mu_s}{}_{\alpha_s} X^{\alpha_s}_{(s)}. \tag{3.1.23}$$

This identifies the components of the pullback as

$$(\varphi^* T)_{\alpha_1 \dots \alpha_s} = (\varphi_*)^{\mu_1}{}_{\alpha_1} \dots (\varphi_*)^{\mu_s}{}_{\alpha_s} T_{\mu_1 \dots \mu_s}. \tag{3.1.24}$$

Hence, the pullback simply acts multilinearly on all indices by contraction with the pushforward matrix $(\varphi_*)^{\mu}{}_{\alpha}$.

We now proceed to the pushforward for contravariant objects, i.e. of tensors $T \in T^{(r,0)}_p\mathcal{M}$. These can be represented as

$$T = T^{\alpha_1 \dots \alpha_r} \, \partial_{\alpha_1} \otimes \dots \otimes \partial_{\alpha_r} \tag{3.1.25}$$

in terms of the coordinate basis $\partial_{\alpha_1} = \frac{\partial}{\partial y^{\alpha_1}}$. Such a tensor is a multilinear map, taking r 1-forms $\omega^{(i)} \in T^*_p\mathcal{M}$ and mapping them to the scalar

$$
\begin{aligned}
T\left(\omega^{(1)}, ..., \omega^{(r)}\right) &= T^{\alpha_1 ... \alpha_r}\, \omega^{(1)}\left(\partial_{\alpha_1}\right) ... \omega^{(r)}\left(\partial_{\alpha_r}\right) \\
&= T^{\alpha_1 ... \alpha_r}\, \omega^{(1)}_{\alpha_1} ... \omega^{(r)}_{\alpha_r}.
\end{aligned}
\tag{3.1.26}
$$

We define the pushforward $\varphi_* T \in T^{r,0}_{\varphi(p)}\mathcal{N}$ as the unique tensor satisfying

$$
T\left(\varphi^* \omega^{(1)}, ..., \varphi^* \omega^{(r)}\right) = (\varphi_* T)\left(\omega^{(1)}, ..., \omega^{(r)}\right) \quad \forall \omega^{(i)} \in T^*_{\varphi(p)}\mathcal{N}. \tag{3.1.27}
$$

In words: The pushforward $\varphi_* T$ "says" the same about the $\omega^{(i)}$ as the original tensor T would have to say about their pullbacks. In terms of coordinates, we verify that this really is just the multilinear extension of what happens to vectors: Inserting coordinate expressions for both sides we get

$$
T^{\alpha_1 ... \alpha_r} (\varphi_*)^{\mu_1}{}_{\alpha_1} \omega^{(1)}_{\mu_1} ... (\varphi_*)^{\mu_r}{}_{\alpha_r} \omega^{(r)}_{\mu_r} = (\varphi_* T)^{\mu_1 ... \mu_r} \omega^{(1)}_{\mu_1} ... \omega^{(r)}_{\mu_r}. \tag{3.1.28}
$$

We may thus identify the components of $\varphi_* T$ as

$$
(\varphi_* T)^{\mu_1 ... \mu_r} = (\varphi_*)^{\mu_1}{}_{\alpha_1} ... (\varphi_*)^{\mu_r}{}_{\alpha_r} T^{\alpha_1 ... \alpha_r}, \tag{3.1.29}
$$

which is simply just the index-wise transformation with the pushforward matrix $(\varphi_*)^{\mu}{}_{\alpha}$.

3.2 Special Case: Submanifolds and the Pushforward as Inclusion

Now that we have established some general theory of smooth maps between manifolds—along with their associated pushforwards and pullbacks—we are ready to specialise to a particularly important class: the inclusion maps of submanifolds. Given a submanifold $\mathcal{S} \subset \mathcal{M}$, the inclusion map $\iota : \mathcal{S} \to \mathcal{M}$ allows us to view $\mathcal{S}$ as embedded in the ambient space $\mathcal{M}$

This setup will naturally recover familiar constructions from Riemannian geometry. Most notably, we will see that the induced metric on the submanifold arises as the pullback of the ambient metric under ι. Framed this way, the induced geometry of submanifolds is revealed to be not an isolated trick, but part of the broader formalism of pullbacks—showing how submanifold geometry fits seamlessly into the general machinery of differential geometry.

Definition (Inclusion Map for Submanifolds) Let $\mathcal{M}$ be an m-dimensional manifold and $\mathcal{S} \subset \mathcal{M}$ an s-dimensional submanifold. We call the map

$$
\iota : \mathcal{S} \to \mathcal{M}, \quad p \mapsto p \tag{3.2.1}
$$

the *inclusion map* or simply *inclusion* of $\mathcal{S}$ into $\mathcal{M}$.

At first glance, this may not seem particularly interesting—it merely sends each point to itself. But despite its simplicity, ι is a smooth map between manifolds, and as such, all the machinery developed previously (pushforward, pullback, tensor behaviour) can and does apply. This makes the inclusion map a powerful conceptual tool, especially when studying how geometric or tensorial structures on the ambient manifold $\mathcal{M}$ restrict or induce structure on the submanifold $\mathcal{S}$. So, let us now examine ι through this lens.

The Pushforward ι_* Firstly, we consider the pushforward ι_*. Before diving into concrete definitions, it helps to build some intuition. In general, the pushforward φ_* of a smooth map $\varphi : \mathcal{M} \to \mathcal{N}$ maps tangent vectors from $T_p\mathcal{M}$ to $T_{\varphi(p)}\mathcal{N}$, capturing how directions around p deform under φ.

In our case, $\varphi = \iota$ is the inclusion of a submanifold $\mathcal{S} \subset \mathcal{M}$. Since $\iota(p) = p$, the point itself remains unchanged, and so do the local relationships between points. A tangent vector $X \in T_p\mathcal{S}$ should therefore remain to be the same geometric object under the pushforward, just now interpreted as living inside the larger ambient space $T_p\mathcal{M}$. We are thus led to expect that ι_* simply embeds $T_p\mathcal{S}$ into $T_p\mathcal{M}$.

This is what we will now make precise. For that, we consider the pushforward $\iota_* X$ of some vector $X \in T_p\mathcal{S}$. By definition, for any function $f \in C^\infty(\mathcal{M})$, its action is given by

$$(\iota_* X)[f] = X[f \circ \iota]. \tag{3.2.2}$$

Since $\iota(p) = p$ for all $p \in \mathcal{S}$, the composition $f \circ \iota$ is simply the restriction of f to the submanifold, i.e. $f|_{\mathcal{S}}$. That is,

$$(\iota_* X)[f] = X[f|_{\mathcal{S}}]. \tag{3.2.3}$$

This shows that $\iota_* X \in T_p\mathcal{M}$ acts on ambient functions $f \in C^\infty(\mathcal{M})$ exactly as $X \in T_p\mathcal{S}$ acts on their restrictions to $\mathcal{S}$. In other words, the pushforward simply embeds the vector X into the ambient tangent space in a way that preserves its actions on functions.

We therefore conclude that the pushforward acts as the canonical inclusion of tangent spaces,

$$\iota_* : T_p\mathcal{S} \to T_p\mathcal{M}, \quad X \mapsto X, \quad \text{for all } p \in \mathcal{S}. \tag{3.2.4}$$

That is, we naturally identify each tangent space $T_p\mathcal{S}$ with a subspace of the ambient tangent space $T_p\mathcal{M}$, and the map ι_* serves as a pointwise injective linear embedding.

Let us now consider how the pushforward looks in coordinate components. To this end, let $p \in \mathcal{S} \subset \mathcal{M}$, and suppose y^i, $i = 1, ..., s$ and x^μ, $1, ..., m$, are local coordinates around p on $\mathcal{S}$ and $\mathcal{M}$, respectively. We consider the pushforward of a vector $X \in T_p\mathcal{S}$, written in the coordinate basis $\partial_i = \frac{\partial}{\partial y^i}$ as

$$X = X^i \partial_i \tag{3.2.5}$$

From the general theory of pushforwards established in the previous section, its image under the inclusion ι is given by

$$\iota_* X = (\iota_* X)^\mu \partial_\mu = (\iota_*)^\mu{}_i X^i \partial_\mu \tag{3.2.6}$$

where $\partial_\mu = \frac{\partial}{\partial x^\mu}$ is the coordinate basis of $T_p\mathcal{M}$, and the pushforward matrix $(\iota_*)^\mu{}_i$ is given by the Jacobian matrix

$$(\iota_*)^\mu{}_i = \frac{\partial x^\mu}{\partial y^i}. \tag{3.2.7}$$

This is the same object we previously denoted by $E^\mu{}_i$ in equation (2.2.6), and we now return to that notation. Thus, the pushforward becomes

$$\iota_* X = \left(E^\mu{}_i X^i\right)\partial_\mu \quad \Leftrightarrow \quad (\iota_* X)^\mu = E^\mu{}_i X^i. \tag{3.2.8}$$

The Pullback $\boldsymbol{\iota^*}$ We now turn to the pullback $\iota^* : T_p^*\mathcal{M} \to T_p^*\mathcal{S}$ associated with the immersion $\iota : \mathcal{S} \to \mathcal{M}$. By definition, for any $X \in T_p\mathcal{S}$ and $\omega \in T_p^*\mathcal{M}$, it satisfies

$$\langle \iota^* \omega, X \rangle_{\mathcal{S}} = \langle \omega, \iota_* X \rangle_{\mathcal{M}}. \tag{3.2.9}$$

In terms of components, this reads

$$(\iota^* \omega)_i X^i = \omega_\mu (\iota_* X)^\mu = \omega_\mu E^\mu{}_i X^i, \tag{3.2.10}$$

from which we identify the components of the pullback as

$$(\iota^* \omega)_i = E^\mu{}_i \omega_\mu. \tag{3.2.11}$$

The Induced Metric As we have seen in the previous section, we can also define the pushforward and pullback for purely contra- or covariant tensors, by applying the transformation with $E^\mu{}_i$ to each index separately. This arose from the requirement of compatibility under the map between manifolds of the multilinear map a tensor defines. A tensor of particular interest to differential geometry is the metric. Since it is a $(0,2)$-tensor, it is purely covariant, and we can apply the pullback to the metric on $\mathcal{M}$ to get a $(0,2)$-tensor on $\mathcal{S}$. Intuition would tell us that this is the induced metric, but let us go through this calmly. The metric on $\mathcal{M}$ is the symmetric tensor

$$ds^2 = g = g_{\mu\nu}\, dx^\mu \otimes dx^\nu. \tag{3.2.12}$$

It is a symmetric bilinear map of two vectors $A, B \in T_p\mathcal{M}$ to $\mathbb{R}$, with

$$g(A, B) = g_{\mu\nu}\, dx^\mu(A)\, dx^\nu(B) = g_{\mu\nu} A^\mu B^\nu. \tag{3.2.13}$$

Its pullback $\iota^* g$ is defined by the relationship

$$(\iota^* g)(X, Y) = g(\iota_* X, \iota_* Y), \quad X, Y \in T_p\mathcal{S}. \tag{3.2.14}$$

Let us briefly interpret this, then we move to the coordinate expression which will match the induced metric we derived in Section 2.2. Equation (3.2.14) tells us that the the pullback $\iota^* g$ is a symmetric bilinear map that simply uses the metric on $\mathcal{M}$ to measure the pushforwards $\iota_* X$ and $\iota_* Y$. Recall, however, that we had previously found $\iota_* X$ and $\iota_* Y$ to be nothing more than the natural embeddings of X and Y in $T_p\mathcal{M}$. So, in essence, the equation states that $(\iota^* g)$ simply returns the same value as g would, if X and Y were to be seen as vectors in $T_p\mathcal{M}$, which they can as $T_p\mathcal{S}$ is embedded in it as a subspace by ι_*. This narrative of "asking the ambient metric what it thinks and reproducing that" is precisely the reasoning we used in Section 2.2 to motivate its definition—but now, we have rediscovered it in a much more general context, in that of pushforwards and pullbacks associated to smooth maps between manifolds.

Going through the component expressions, we find

$$\begin{aligned} (\iota^* g)(X, Y) &= g(\iota_* X, \iota_* Y) = g_{\mu\nu} E^\mu{}_i E^\nu{}_j X^i Y^j \\ &= \left(g_{\mu\nu} E^\mu{}_i E^\nu{}_j\right) X^i X^j = \gamma_{ij} X^i Y^j = \gamma(X, Y). \end{aligned} \tag{3.2.15}$$

This confirms explicitly that $\iota^* g = \gamma$; in other words, the induced metric on $\mathcal{S}$ is simply the pullback of the metric from the ambient manifold $\mathcal{M}$. We have come full circle: the geometric idea that guided our definition of the induced metric has now emerged naturally from algebraic considerations grounded in a broader theoretical framework. That coherence gives us confidence to admit the idea into our formal foundations—when algebra and intuition converge, we are likely on the right path.

3.3 The Left-Inverse: Projections onto Tangent Spaces

We have previously remarked that for the pushforward

$$\varphi_* : T_p\mathcal{M} \to T_{\varphi(p)}\mathcal{N} \tag{3.3.1}$$

associated with a smooth map $\varphi : \mathcal{M} \to \mathcal{N}$, it is in general not possible to define an inverse

$$(\varphi_*)^{-1} : T_{\varphi(p)}\mathcal{N} \to T_p\mathcal{M}. \tag{3.3.2}$$

However, if φ_* is injective, then it *is* possible to define a *left-inverse.* A particularly relevant case where this holds is that of immersion maps, such as

$$\varphi = \iota : \mathcal{S} \to \mathcal{M}, \tag{3.3.3}$$

where we have previously seen that the associated pushforward

$$\iota_* : T_p\mathcal{S} \to T_p\mathcal{M} \tag{3.3.4}$$

is an injective linear embedding.

Let us now examine this case in detail. We may define a (non-unique) left-inverse[2]

$$(\iota_*)^{-1} : T_p\mathcal{M} \to T_p\mathcal{S} \tag{3.3.5}$$

as a linear map satisfying

$$(\iota_*)^{-1} \circ \iota_* = \mathrm{id}_{T_p\mathcal{S}} . \tag{3.3.6}$$

Though this constrains $(\iota_*)^{-1}$ fully on $\mathrm{im}(\iota_*)$, its action on the complement of $\mathrm{im}(\iota_*)$ remains arbitrary—hence we have no uniqueness.

In components, using the same coordiante systems as before, let us expand both sides of this identity for a vector $X \in T_p\mathcal{S}$. On the right-hand side, we get

$$\mathrm{id}_{T_p\mathcal{S}}\, X = X = X^i\partial_i = \delta^i_j X^j \partial_i \tag{3.3.7}$$

For the left-hand side, we use the component expression for the pushforward to find

$$((\iota_*)^{-1} \circ \iota_*)(X) = (\iota_*)^{-1}(\iota_* X) = (\iota_*)^{-1}\left(E^\mu{}_i X^i \partial_\mu\right) = ((\iota_*)^{-1})^j{}_\mu E^\mu{}_i X^i \partial_j. \tag{3.3.8}$$

Let us define

$$E^i{}_\mu = ((\iota_*)^{-1})^i{}_\mu \tag{3.3.9}$$

as the component representation of the left-inverse. Equating both of the above equations (3.3.7) and (3.3.8) allows us to rewrite equation (3.3.6) in components as

$$E^i{}_\mu E^\mu{}_j X^j \partial_i = \delta^i_j X^j \partial_i. \tag{3.3.10}$$

By arbitrariness of X, we conclude the matrix identity

$$E^i{}_\mu E^\mu{}_j = \delta^i_j. \tag{3.3.11}$$

This is exactly the condition for $E^i{}_\mu$ to be a left-inverse of $E^\mu{}_i$ in the standard sense of matrix algebra. Notice that this implies that both the ranks of $E^\mu{}_i$ and $E^i{}_\mu$ must be $s = \dim\mathcal{S} = \operatorname{rank}\delta^i_j$.

The left-inverse $(\iota_*)^{-1}$ might not appear very useful, since all it does is allow us to form an identity on $T_p\mathcal{S}$. We can, however, attempt to use it as a right-inverse, to define the map

$$P := \iota_* \circ (\iota_*)^{-1} : T_p\mathcal{M} \to \mathrm{im}(\iota_*) \subset T_p\mathcal{M}. \tag{3.3.12}$$

Clearly, the component representation of this map acting on $X \in T_p\mathcal{M}$ is

$$P(X) = \left(\iota_* \circ (\iota_*)^{-1}\right)\left(X^\mu \partial_\mu\right) = \left(E^\mu{}_i E^i{}_\nu X^\nu\right)\partial_\mu = \left(P^\mu{}_\nu X^\nu\right)\partial_\mu, \tag{3.3.13}$$

with the components of P denoted by $P^\mu{}_\nu := E^\mu{}_i E^i{}_\nu$. Since P is a linear map, and

$$P^2 = \iota_* \circ \underbrace{(\iota_*)^{-1} \circ \iota_*}_{=\mathrm{id}_{T_p\mathcal{S}}} \circ (\iota_*)^{-1} = \iota_* \circ (\iota_*)^{-1} = P, \tag{3.3.14}$$

we infer that P is a projection. Moreover, since we have $\operatorname{rank} P = \operatorname{rank}\iota_* = \dim\mathcal{S} = s$, we infer that

[2] We use the notation $(\cdot)^{-1}$ here to indicate just a left- and not a proper inverse.

$$P : T_p\mathcal{M} \to \mathrm{im}(\iota_*) \tag{3.3.15}$$

is an injective projector of $T_p\mathcal{M}$ onto the embedding $\mathrm{im}(\iota_*)$ of $T_p\mathcal{S}$ in $T_p\mathcal{M}$.

Recall that the left-inverse $(\iota_*)^{-1}$ is non-unique—it depends on the choice of complement to $\mathrm{im}(\iota_*)$ during its construction. It is always possible to choose it such that P satisfies the condition

$$g(P(X), Y) = g(X, P(Y)), \quad X, Y \in T_p\mathcal{M}, \tag{3.3.16}$$

turning P into an orthogonal projection.[3]

In words, P tells us the following: For any vector $X \in T_p\mathcal{M}$, the associated $P(X) \in \mathrm{im}(\iota_*)$ represents the part of X that aligns with the tangent space $T_p\mathcal{S}$ of the submanifold. This construction of P allows us to split any vector $X \in T_p\mathcal{M}$ into two parts: a component tangent to the submanifold $\mathcal{S}$, and a component orthogonal to it. That is,

$$X = P(X) + (X - P(X)) \tag{3.3.17}$$

where $P(X) \in \mathrm{im}(\iota_*)$ and $X - P(X) \in \mathrm{im}(\iota_*)^{\perp}$, where the complement is taken with respect to g.

What we should take away from this section is the following. The left-inverse $(\iota_*)^{-1}$ is not particularly interesting on its own, as it is inherently non-unique—its definition depends on an arbitrary choice of complement to $\mathrm{im}(\iota_*)$. However, it does enable the construction a projection $P : T_p\mathcal{M} \to \mathrm{im}(\iota_*)$, which can be made into an orthogonal projection by condition (3.3.16), equivalent to the symmetry condition[4]

$$P_{\mu\nu} = P_{\nu\mu} \tag{3.3.18}$$

where $P_{\mu\nu} = g_{\mu\lambda}P^{\lambda}{}_{\nu}$. This condition ensures that P is self-adjoint with respect to the metric g, and hence defines an orthogonal projection.

When P is symmetric in this sense, we can interpret it as specifying an orthogonal splitting of the tangent space,

$$T_p\mathcal{M} = \underbrace{\mathrm{im}(\iota_*)}_{\text{tangent to } \mathcal{S}} \oplus \underbrace{\ker(P)}_{\text{normal to } \mathcal{S}}, \tag{3.3.19}$$

where the symbol $\oplus$ indicates a direct sum of orthogonal subspaces. The image of ι_* is the embedding of $T_p\mathcal{S}$ in $T_p\mathcal{M}$), while the kernel of P corresponds to vectors in $T_p\mathcal{M}$ orthogonal to $\mathcal{S}$ —i.e., the *normal directions*.

3.4 Ambient Metric Decomposition and the Pullback of $P_{\mu\nu}$

In the previous section, we saw that for P to define an *orthogonal* projection, its components must be symmetric with respect to the metric g. Since the metric tensor itself is symmetric, this hints at the possibility of a decomposition of g in terms of P and a complementary projection Q. The goal of this section is to examine the properties and consequences of such a decomposition.

The symmetry condition for P was previously written in components as

$$P_{\mu\nu} = P_{\nu\mu}, \quad \text{with} \quad P_{\mu\nu} = g_{\mu\lambda}P^{\lambda}{}_{\nu}, \tag{3.4.1}$$

but this can be expressed more naturally in coordinate-free language, which we now introduce.

To do so, we reinterpret P, originally a $(1,1)$-tensor (a linear map on $T_p\mathcal{M}$) as a $(0,2)$-tensor by lowering one index using the metric. That is, we define the bilinear form $\tilde{P}$

[3] We give an explicit construction of this in terms of components in equation (6.5.8)

[4] $g(P(X), Y) = g_{\mu\nu}P^{\mu}{}_{\lambda}X^{\lambda}Y^{\nu} = P_{\nu\mu}X^{\mu}Y^{\nu}$, and $g(X, P(Y)) = ... = P_{\mu\nu}X^{\mu}Y^{\nu}$.

$$\tilde{P}(X,Y) = g(P(X),Y), \tag{3.4.2}$$

for all $X, Y \in T_p\mathcal{M}$. In components, this reads

$$\tilde{P}(X,Y) = \big(g_{\mu\lambda}P^{\lambda}{}_{\nu}\big)X^{\mu}Y^{\nu} \tag{3.4.3}$$

so the components of $\tilde{P}$ are precisely what we earlier denoted by $P_{\mu\nu}$—the object that must be symmetric. Thus, the requirement of symmetry of P turns into

$$\tilde{P}(X,Y) = \tilde{P}(Y,X), \quad X,Y \in T_p\mathcal{M}. \tag{3.4.4}$$

We will retain the tilde notation for $\tilde{P}$ throughout this section for clarity. However, it should be understood that identifications between $(1,1)$- and $(0,2)$-tensors via the metric are always possible, and we may implicitly make such conversions in later sections, slightly abusing notation for brevity.

Let us now approach the decomposition of the metric in terms of $\tilde{P}$. We introduce it as

$$g(X,Y) = \tilde{P}(X,Y) + \tilde{Q}(X,Y), \quad X,Y \in T_p\mathcal{M} \tag{3.4.5}$$

where $\tilde{Q}$, trivially given by

$$\tilde{Q}(X,Y) = g(X,Y) - \tilde{P}(X,Y), \quad X,Y \in T_p\mathcal{M}, \tag{3.4.6}$$

is necessarily symmetric as well—due to symmetry of both $\tilde{P}$ and g. The components of $\tilde{Q}$ (and its associated $(1,1)$-tensor Q) are given by

$$\tilde{Q}_{\mu\nu} = g_{\mu\nu} - P_{\mu\nu} \quad \Leftrightarrow \quad Q^{\mu}{}_{\nu} = g^{\mu\lambda}\tilde{Q}_{\lambda\nu} = \delta^{\mu}_{\nu} - P^{\mu}{}_{\nu}. \tag{3.4.7}$$

Notice that hence, $Q(X) = X - P(X)$ is nothing but the projection of X onto the normal space $\ker(P)$ in the decomposition (3.3.19), and $\tilde{Q}$ its associated bilinear form. The projections P and Q are orthogonally complete, as we have (this is easily verifiable by plugging in definitions)

$$P + Q = \mathrm{id}_{T_p\mathcal{M}}, \qquad P \circ Q = Q \circ P = 0. \tag{3.4.8}$$

What we learned thus far is the following: Writing the metric in terms of $\tilde{P}$ and collecting the remaining parts into $\tilde{Q}$ naturally decomposes it into a projection onto the embedding of $T_p\mathcal{S}$ into $T_p\mathcal{M}$ and its orthogonal complement—the *normal space*

$$N_p\mathcal{S} := \ker(P). \tag{3.4.9}$$

Now that we have decomposed the metric, we can reconsider the pullback of this decomposition onto $\mathcal{S}$. Since we have established $\iota^*g = \gamma$, i.e., that the ambient metric pulls back to the induced metric, it is natural to ask: which part of γ arises from $\tilde{P}$, and which from $\tilde{Q}$? We approach this question from two perspectives: firstly, via a coordinate-independent formulation; then, secondly, by a component-based calculation.

By linearity of ι^*, and since $\iota^*g = \gamma$, it is sufficient to compute $\iota^*\tilde{P}$—this immediately gives us the pullback of $\tilde{Q}$ via $\iota^*\tilde{Q} = \gamma - \iota^*\tilde{P}$. We compute this now; let $X, Y \in T_p\mathcal{S}$—then, by definition,

$$\begin{aligned}
\big(\iota^*\tilde{P}\big)(X,Y) &= \tilde{P}(\iota_*X, \iota_*Y) = g(P(\iota_*X), \iota_*Y) = g\Big(\big(\iota_* \circ \underbrace{(\iota_*)^{-1} \circ \iota_*}_{=\mathrm{id}_{T_p\mathcal{S}}}\big)X, \iota_*Y\Big) \\
&= g(\iota_*X, \iota_*Y) = (\iota^*g)(X,Y) = \gamma(X,Y); \\
\Longrightarrow \qquad \iota^*\tilde{Q} &= \gamma - \gamma = 0.
\end{aligned} \tag{3.4.10}$$

In summary, we have found

$$\gamma = \iota^* g = \iota^* \tilde{P}. \tag{3.4.11}$$

Thus, $\iota^* g$ only depends on the tangential part $\tilde{P}$—the normal contribution $\tilde{Q}$ vanishes under pullback. Notice that $\operatorname{rank} P = \operatorname{rank} \gamma = \dim \mathcal{S}$; in light of the above, this implies that γ and $\tilde{P}$ represent the same bilinear map on $T_p\mathcal{S}$ and its embedding $\operatorname{im}(\iota_*) \subset T_p\mathcal{M}$.

Let us now repeat the same calculation in components. We find

$$\begin{aligned}\left(\iota^* \tilde{P}\right)_{ij} &= E^\mu{}_i E^\nu{}_j P_{\mu\nu} = E^\mu{}_i E^\nu{}_j g_{\mu\lambda} P^\lambda{}_\nu = E^\mu{}_i E^\nu{}_j g_{\mu\lambda} E^\lambda{}_k E^k{}_\nu \\ &= E^\mu{}_i \delta^k_j E^\lambda{}_k g_{\mu\lambda} = E^\mu{}_i E^\nu{}_j g_{\mu\nu} = \gamma_{ij}.\end{aligned} \tag{3.4.12}$$

This reproduces the identity (3.4.11) in terms of components as

$$\gamma_{ij} = E^\mu{}_i E^\nu{}_j g_{\mu\nu} = E^\mu{}_i E^\nu{}_j P_{\mu\nu}. \tag{3.4.13}$$

This result offers a clear structural interpretation of the induced metric γ: it is precisely the tangential part of the ambient metric g, isolated by pulling back the tangential projection P. The normal component Q plays no role in the geometry intrinsic to the submanifold, as expected—γ contains only the information relevant to distances and angles *within* $\mathcal{S}$.

4 Bundles

This section aims to introduce the notion of *vector bundles* one can define on a smooth manifold. Though this is not strictly necessary to study submanifolds, it seems like it would be a useful digression to prepare for the differential geometry lecture in Part III, so I will go over it briefly here.

4.1 Vector Bundles: Intuition and Definitions

So far, when we discussed vectorial (or tensorial) objects, our expressions have been entirely pointwise. We have considered, for example, maps from $T_p\mathcal{M}$ to $T_{\varphi(p)}\mathcal{N}$, where $\mathcal{N}$ is another manifold and $\varphi : \mathcal{M} \to \mathcal{N}$ is a smooth map between them. These are relations between tangent spaces at individual points.

However, since there exists a tangent space at every point $p \in \mathcal{M}$, it is natural to seek a way to *assemble* or *bundle together* all these tangent spaces into a single structure. Intuitively, we take the manifold $\mathcal{M}$ and, at each point, attach the to it its corresponding tangent space. This yields a new, higher-dimensional manifold-like object that encodes all the tangent spaces and their relation to points in $\mathcal{M}$. This construction is known as the *tangent bundle*.

Of course, this is just a loose and purely intuitive description. In this section, we aim to make rigorous the idea of "attaching a vector space to each point of a manifold" by introducing the notion of a *vector bundle*. In the next section, we will see that the tangent (and cotangent) bundles —central objects in differential geometry—are but special instances of this general concept.

Definition (Vector Bundle) Let $\mathcal{M}$ be a smooth manifold of dimension m. A *smooth real vector bundle* of rank n over $\mathcal{M}$ is a triple $(\mathcal{E}, \pi, \mathcal{M})$, where:

1. $\mathcal{E}$ is a smooth manifold of dimension $n + m$, called the *total space*.
2. $\pi : \mathcal{E} \to \mathcal{M}$ is a smooth surjective map, called the *bundle projection*.
3. For each $p \in \mathcal{M}$, the *fibre* $\mathcal{E}_p := \pi^{-1}(\{p\})$ is equipped with the structure of a real vector space of dimension n.
4. We have *local triviality*, i.e. for each $p \in \mathcal{M}$ there exists an open neighbourhood $U \subset \mathcal{M}$ of p and a diffeomorphism

$$\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^n \tag{4.1.1}$$

such that

- Φ is a fibre-preserving vector space isomorphism on fibres, meaning that for each $q \in U$,

$$\Phi|_{\mathcal{E}_q} : \mathcal{E}_q \to \{q\} \times \mathbb{R}^n \simeq \mathbb{R}^n \tag{4.1.2}$$

is a vector space isomorphism.

- the diagram

$$\begin{tikzcd} \mathcal{E} \supset \pi^{-1}(U) \arrow[r, "\Phi"] \arrow[d, "\pi"'] & U \times \mathbb{R}^n \arrow[dl, "P_1"] \\ U & \end{tikzcd} \tag{4.1.3}$$

commutes, where P_1 is the projection onto the first component of the Cartesian product $((a, b) \mapsto a)$.

Let us now go through this definition calmly, and explain the meaning and intuition behind each of the constructions separately.

Firstly, we should give a summary of what a vector bundle is supposed to be. Intuitively, a vector bundle over $\mathcal{M}$ is a smooth family of vector spaces $\{\mathcal{E}_p\}_{p\in\mathcal{M}}$, smoothly parameterised by $\mathcal{M}$, such that near each point, the collection of fibres looks like a product $U \times \mathbb{R}^n$. This means we can locally identify each fibre with $\mathbb{R}^n$ in a way that varies smoothly with the base point, and respects the vector space operations.

Now, let us go over each part of the definition in detail.

1. *Total Space*: The total space $\mathcal{E}$ can be viewed as the union of all the vector spaces at each point. Formally, it is the disjoint union of all fibres,

$$\mathcal{E} = \bigsqcup_{p\in\mathcal{M}} \mathcal{E}_p. \tag{4.1.4}$$

This is an $(m+n)$-dimensional manifold, since the base manifold has dimension $\mathcal{M}$, and attaching an n-dimensional vector space at each point increases the dimension by n.

Very loosely, one might imagine this as dragging a window (the vector space) across a screen (the manifold) in Windows XP, with the bug that leaves behind a smeared trail of it. As you drag it, the trail being formed represents this union of all these smeared copies (recounting overlapping regions appropriately), which is what is known as the total space.

2. *Bundle Projection*: Consider the total space $\mathcal{E}$ of a vector bundle. Since it can be viewed as the union of all tangent spaces, any point in it contains two pieces of information: A vector in one of the tangent spaces, as well as the point to which that tangent space is attached to. In other words, an element of $\mathcal{E}$ is naturally written as a pair (p, X_p), where p denotes a point in the manifold $\mathcal{M}$, and X_p is a vector in the fibre attached to $\mathcal{M}$ at p, i.e., $X_p \in \mathcal{E}_p$.

To visualise this, imagine attaching copies of $\mathbb{R}^2$ perpendicular to a straight line (which can be represented by $\mathbb{R}$). The result is a space that is isomorphic to $\mathbb{R}^3$, where each point in this total space consists of three real numbers. The first number indicates the position along the line (i.e., in which "paper sheet" in the stack we are on), and the last two numbers describe a vector in $\mathbb{R}^2$ (i.e., the direction of the vector in the plane attached at that position).

The role of the bundle projection map $\pi : \mathcal{E} \to \mathcal{M}$ is straightforward: it takes such a pair $\left(p, X_p\right)$ and projects it onto the point $p \in \mathcal{M}$ to which the vector space is attached. In our analogy, this means that the bundle projection "reads" the position within the stack of paper and returns the corresponding piece of paper, or more formally, the point on the line to which the sheet is attached.

This map must be surjective: every point $p \in \mathcal{M}$ must have a corresponding vector space attached to it. If it were not surjec tive, there would exist a point $p \in \mathcal{M}$ without any associated vector, which would contradict the idea of "attaching a vector space to each point of the manifold".

3. *Fibres*: The fibre at $p \in \mathcal{M}$ is the collection of all vectors attached to p, i.e. its preimage under the bundle projection π. More precisely,

$$\mathcal{E}_p = \pi^{-1}(\{p\}). \tag{4.1.5}$$

Since we want to attach not just *any* kind of fibre, but specifically an n-dimensional vector space, we impose the additional requirement that each fibre $\mathcal{E}_p$ carries the structure of a real vector space of dimension n.

4. *Local Triviality*: This is likely the most convoluted part of the definition, but can also be broken down intuitively. What local triviality demands is that *locally*, in some neighbourhood $U \subset \mathcal{M}$ of $p \in \mathcal{M}$, the total space $\mathcal{E}$ "looks like" the space $U \times \mathbb{R}^n$. This is the simplest way of "attaching vector spaces to each point"—the Cartesian product does exactly that. In more formal terms, "looks like" is replaced by the notion of the diffeomorphism Φ. Since $\mathbb{R}^n$ is the concrete representation of the attached vector space, we would also like the fibres $\mathcal{E}_q$, $q \in U$ to map to $\mathbb{R}^n$ under Φ in a way that respects the algebraic structure—hence the condition on $\Phi|_{\mathcal{E}_q}$. The requirement that the diagram (4.1.3) commutes then further ensures that the fibres get attached to the correct points on $\mathcal{M}$—we don't want to, for example, end up attaching the tangent space at a point p to another point q.

Now that we have defined vector bundles, let us introduce a notion that makes use of it. Specifically, we consider so-called *smooth sections* of vector bundles. A smooth section is, intuitively speaking, the selection of one vector in the fibre $\mathcal{E}_p$ at each $p \in \mathcal{M}$, in a way that creates a smooth surface in the total space $\mathcal{E}$. Such a surface can be viewed as a vector field, since it maps each point on the manifold to one vector in its fibre.

Definition (Smooth Section) Let $\mathcal{M}$ be a smooth manifold and $(\mathcal{E}, \pi, \mathcal{M})$ a smooth real vector bundle over $\mathcal{M}$. A *smooth section* of $\mathcal{E}$ is a smooth map

$$\sigma : \mathcal{M} \to \mathcal{E} \tag{4.1.6}$$

such that

$$\pi \circ \sigma = \mathrm{id}_{\mathcal{M}} \tag{4.1.7}$$

In other words, for each $p \in \mathcal{M}$, the map σ selects a vector $\sigma(p) \in \mathcal{E}_p$ lying in the fibre over p, and this assignment varies smoothly with p. In particular, for $\pi \circ \sigma$ to be the identity on $\mathcal{M}$, σ must be injective.

We denote the set of all smooth sections of $\mathcal{E}$ by

$$\begin{aligned} \Gamma(\mathcal{E}) &= \{\sigma : \mathcal{M} \to \mathcal{E} \mid \pi \circ \sigma = \mathrm{id}_{\mathcal{M}}, \sigma \text{ smooth}\}, \\ &= \{\sigma : \mathcal{M} \to \mathcal{E} \mid \sigma \text{ smooth section on } \mathcal{E}\}. \end{aligned} \tag{4.1.8}$$

This is a real vector space under pointwise addition and scalar multiplication. In the case where $\mathcal{E} = T\mathcal{M}$ is the tangent bundle (which we introduce in the next section), $\Gamma(T\mathcal{M})$ is the space of smooth vector fields on $\mathcal{M}$.

We are now prepared for the next section, in which we will define the tangent and cotangent bundles $T\mathcal{M}$ and $T^*\mathcal{M}$, and use smooth sections to give an alternative perspective on vector fields and differential 1-forms.

4.2 The Tangent and Cotangent Bundles

The tangent bundle $T\mathcal{M}$ and the cotangent bundle $T^*\mathcal{M}$ are, arguably, the two most important vector bundles in differential geometry. In essence, these are the special cases of vector bundles where one chooses the fibres $\mathcal{E}_p$ to be the (co)-tangent spaces $T_p\mathcal{M}$ and $T_p^*\mathcal{M}$, respectively—let us now introduce this rigorously.

Definition (Tangent Bundle) Let $\mathcal{M}$ be a smooth manifold of dimension m and denote by $T_p\mathcal{M}$ its tangent space at any point $p \in \mathcal{M}$. Define the $2m$-dimensional total space $T\mathcal{M}$ by the disjoint union

$$T\mathcal{M} = \bigsqcup_{p \in \mathcal{M}} T_p\mathcal{M}. \tag{4.2.1}$$

We define the bundle projection

$$\pi : T\mathcal{M} \to \mathcal{M} \tag{4.2.2}$$

as the map

$$\pi(p, X) = p, \tag{4.2.3}$$

choosing to write elements of $T\mathcal{M}$ as pairs (p, X) with $X \in T_p\mathcal{M}$.

Remarks: A few direct consequences follow from this construction:

- *Fibres*: The fibre over a point $p \in \mathcal{M}$ is given by

  $$(T\mathcal{M})_p = \pi^{-1}(\{p\}) = T_p\mathcal{M}, \tag{4.2.4}$$

  meaning that we are indeed attaching the tangent space $T_p\mathcal{M}$ to each point p.
- *Local triviality and charts on $T\mathcal{M}$*: Given a coordinate chart (U, x^μ) on $\mathcal{M}$, we can define a chart on the preimage $\pi^{-1}(U) \subset T\mathcal{M}$ by identifying

  $$\pi^{-1}(U) \simeq U \times \mathbb{R}^m, \quad (p, V) \mapsto (x^\mu(p), V^\mu), \tag{4.2.5}$$

  where V^μ are the components of the tangent vector $V = V^\mu \partial_\mu$ in the coordinate basis induced by x^μ. This provides $T\mathcal{M}$ with a smooth manifold structure of dimension $2m$.
- *Zero section*: We may define the zero section

  $$\sigma_0 : \mathcal{M} \to T\mathcal{M}, \quad p \mapsto (p, 0), \tag{4.2.6}$$

  or, in coordinates,

  $$x^\mu \mapsto (x^\mu, 0), \tag{4.2.7}$$

  where 0 denotes the zero vector in $T_p\mathcal{M}$. This map is a smooth embedding, and its image forms a submanifold of $T\mathcal{M}$ diffeomorphic to $\mathcal{M}$ itself.

Intuitively, this makes perfect sense: if we attach a vector space to each point of the manifold in such a way that their origins coincide with the base points, then the collection of those origins sweeps out a copy of $\mathcal{M}$ embedded inside $T\mathcal{M}$.

- *Vector fields as section*: Any smooth vector field on $\mathcal{M}$, i.e. a smooth assignment $X : p \mapsto X_p \in T_p\mathcal{M}$, is precisely a smooth section of the tangent bundle,

$$X : \mathcal{M} \to T\mathcal{M}, \quad p \mapsto \big(p, X_p\big) \quad \text{with} \quad \pi \circ X = \mathrm{id}_{\mathcal{M}} \,. \tag{4.2.8}$$

  That is, $X \in \Gamma(T\mathcal{M})$. This is the prototypical example of a smooth section, and shows how the familiar notion of a vector field fits directly into the general formalism of vector bundles.

Definition (Cotangent Bundle) The definition of the cotangent bundle is analogous to that of the tangent bundle, with the distinction being that we use

$$T^*\mathcal{M} = \bigsqcup_{p \in \mathcal{M}} T_p^*\mathcal{M} \tag{4.2.9}$$

as our total space, and the bundle projection

$$\pi : T^*\mathcal{M} \to \mathcal{M} \tag{4.2.10}$$

now given by the map

$$\pi(p, \omega) = p, \tag{4.2.11}$$

writing elements of $T^*\mathcal{M}$ as pairs (p, ω) with $\omega \in T_p^*\mathcal{M}$. The same remarks as for the tangent bundle hold, with the occasional change of terminology from vectors to 1-forms.

4.3 Bundle Maps and Vector Bundle Morphisms

Now that we have introduced the notion of vector bundles as well as the concrete (and important) tangent bundle $T\mathcal{M}$ and the cotangent bundle $T^*\mathcal{M}$, we can begin considering maps between them. A map between two tangent bundles becomes particularly interesting if it respects both the geometric structure of the underlying manifold(s) as well as the algebraic structure of the fibres —i.e., if it maps vectors in one fibre linearly to vectors in another. Such maps we will refer to as *bundle morphisms*. We first introduce the more general concept of a bundle map, and then impose additional algebraic structure to define the notion of bundle morphisms.

Definition (Bundle Map) Let $\mathcal{M}$, $\mathcal{N}$ be smooth manifolds and $(\mathcal{E}, \pi_{\mathcal{M}}, \mathcal{M})$ as well as $(\mathcal{F}, \pi_{\mathcal{N}}, \mathcal{N})$ vector bundles on them. Further, let

$$\varphi : \mathcal{M} \to \mathcal{N} \tag{4.3.1}$$

be a smooth map from one manifold to the other. We call a map

$$\Phi : \mathcal{E} \to \mathcal{F} \tag{4.3.2}$$

a *bundle map* if the following diagram commutes:

$$\begin{tikzcd} \mathcal{E} \arrow[r, "\Phi"] \arrow[d, "\pi_{\mathcal{M}}"'] & \mathcal{F} \arrow[d, "\pi_{\mathcal{N}}"] \\ \mathcal{M} \arrow[r, "\varphi"'] & \mathcal{N} \end{tikzcd} \tag{4.3.3}$$

This is equivalent to the condition that

$$\pi_{\mathcal{N}} \circ \Phi = \varphi \circ \pi_{\mathcal{M}}. \tag{4.3.4}$$

Without further interpretation, the definition above may seem opaque—so let us walk through it and develop some intuition. Consider an element $(p, X) \in \mathcal{E}$, where $p \in \mathcal{M}$ and $X \in \mathcal{E}_p$, the fibre over p. We examine how the two sides of the condition (4.3.4) act on such a point.

Starting with the right-hand side, we have

$$(\varphi \circ \pi_{\mathcal{M}})(p, X) = \varphi(\pi_{\mathcal{M}}(p, X)) = \varphi(p). \tag{4.3.5}$$

this means we first project (p, X) onto its base point $p \in \mathcal{M}$ via $\pi_{\mathcal{M}}$, then apply φ to obtain a point in $\mathcal{N}$.

Now for the left-hand side:

$$(\pi_{\mathcal{N}} \circ \Phi)(p, X) = \pi_{\mathcal{N}}(\Phi(p, X)) = \pi_{\mathcal{N}}(q, Y) = q. \tag{4.3.6}$$

Here, we write $\Phi(p, X)$ as some pair $(q, Y) \in \mathcal{F}$, and $\pi_{\mathcal{N}}$ returns the base point $q \in \mathcal{N}$.

Thus, the commutativity condition simplifies to

$$q = \varphi(p). \tag{4.3.7}$$

In other words, a bundle map Φ must map $(p, X) \in \mathcal{E}$ to $(\varphi(p), Y) \in \mathcal{F}$; the base point of the image is determined entirely by the underlying map φ. The vector part Y can be chosen freely (within the fibre over $\varphi(p)$), but the association between fibres is rigidly tied to that between base points.

An equivalent way to phrase this is to demand that Φ maps each fibre $\mathcal{E}_p$ into the fibre $\mathcal{F}_{\varphi(p)}$. That is,

$$\Phi|_{\mathcal{E}_p} : \mathcal{E}_p \to \mathcal{F}_{\varphi(p)}. \tag{4.3.8}$$

The content of the definition is precisely this: a bundle map over φ is one that maps vectors attached to a point $p \in \mathcal{M}$ into vectors attached to $\varphi(p) \in \mathcal{N}$, without violating the structure of the fibration.

Besides the structure of the fibration, a vector bundle has additional algebraic structure that one could demand to be preserved; each fibre is a vector space, and we could require a bundle map to respect it by imposing linearity. This notion, called *bundle morphisms*, is what we now define rigorously.

Definition (Vector Bundle Morphism) Let $\mathcal{M}$, $\mathcal{N}$ be smooth manifolds, and let $(\mathcal{E}, \pi_{\mathcal{M}}, \mathcal{M})$ and $(\mathcal{F}, \pi_{\mathcal{N}}, \mathcal{N})$ be smooth real vector bundles over them. Suppose we are given a smooth map $\varphi : \mathcal{M} \to \mathcal{N}$ and a bundle map $\Phi : \mathcal{E} \to \mathcal{F}$ covering φ, i.e. a map Φ which satisfies

$$\pi_{\mathcal{N}} \circ \Phi = \varphi \circ \pi_{\mathcal{M}}. \tag{4.3.9}$$

We say that Φ is a *vector bundle morphism* (over φ) if, for each $p \in \mathcal{M}$, the induced map on the fibres

$$\Phi|_{\mathcal{E}_p} : \mathcal{E}_p \to \mathcal{F}_{\varphi(p)} \tag{4.3.10}$$

is a linear map of vector spaces.

Intuitively, a vector bundle morphism over φ can be seen as a fibrewise linear transformation that "respects the base": it linearly transforms each vector in a fibre over $p \in \mathcal{M}$ to a vector in the fibre over $\varphi(p) \in \mathcal{N}$.

While there are many abstract vector bundle morphisms on could define and study, there is a particularly natural one associated with a smooth map $\varphi : \mathcal{M} \to \mathcal{N}$ that we have already encountered: the pushforward. Though we initially introduced the pushforward $\varphi_* : T_p\mathcal{M} \to T_{\varphi(p)}\mathcal{N}$ as a pointwise linear map between tangent spaces, it readily extends to a global vector bundle morphism.

To see this, consider the map

$$\Phi : T\mathcal{M} \to T\mathcal{N}, \quad (p, X) \mapsto (\varphi(p), \varphi_* X), \quad p \in \mathcal{M}, X \in T_p\mathcal{M}. \tag{4.3.11}$$

This construction maps each element of the tangent bundle $T\mathcal{M}$ to the tangent bundle $T\mathcal{N}$ by pushing forward the vector X and sending its base point p to $\varphi(p)$. It is easy to check that this satisfies the bundle map condition

$$\pi_{\mathcal{N}} \circ \Phi = \varphi \circ \pi_{\mathcal{M}}, \tag{4.3.12}$$

and that the fibrewise maps $\Phi|_{T_p\mathcal{M}} = \varphi_* : T_p\mathcal{M} \to T_{\varphi(p)}\mathcal{N}$ are linear. Hence, Φ defines a (vector) bundle morphism from $T\mathcal{M}$ to $T\mathcal{N}$ over φ. Because of this, it is common to write the pushforward as a map

$$\varphi_* : T\mathcal{M} \to T\mathcal{N}, \tag{4.3.13}$$

considering it as a vector bundle morphism.

4.4 The Normal Bundle and Orthogonal Decomposition

In this section, we explore how the tangent bundle $T\mathcal{S}$ of a submanifold $\mathcal{S} \subset \mathcal{M}$ can be understood as a subbundle of the restriction $T\mathcal{M}|_{\mathcal{S}}$ of the tangent bundle $T\mathcal{M}$ to $\mathcal{S}$. Further, we define the normal bundle $N\mathcal{S}$ and explain how the pointwise projections P and Q we introduced earlier extend naturally to smooth vector bundle morphisms between $T\mathcal{M}|_{\mathcal{S}}$, $T\mathcal{S}$ and $N\mathcal{S}$. While this may seem like an unnecessary abstraction at first glance, it will turn out to offer geometric clarity and prepare us for later constructions involving intrinsic curvature.

Definition (Restriction of vector bundles to submanifolds) Let $\mathcal{M}$ be a smooth manifold of dimension m, and let $\mathcal{S} \subset \mathcal{M}$ be a submanifold of dimension $s < m$. Given a vector bundle $(\mathcal{E}, \pi, \mathcal{M})$ over $\mathcal{M}$, its *restriction to* $\mathcal{S}$ is defined as the triple

$$(\mathcal{E}|_{\mathcal{S}}, \pi|_{\mathcal{S}}, \mathcal{S}), \quad \text{where } \mathcal{E}|_{\mathcal{S}} := \pi^{-1}(\mathcal{S}) \subset \mathcal{E}. \tag{4.4.1}$$

This construction simply discards all fibres of $\mathcal{E}$ lying over points $p \in \mathcal{M} \setminus \mathcal{S}$, retaining only the portion of the bundle sitting above $\mathcal{S}$.

As an example, consider the tangent bundle $T\mathcal{M}$. Its restriction to $\mathcal{S}$, denoted by $T\mathcal{M}|_{\mathcal{S}}$, consists of the collection of tangent spaces $T_p\mathcal{M}$ for $p \in \mathcal{S}$. This restricted bundle cannot be the same as the tangent bundle $T\mathcal{S}$ of $\mathcal{S}$—the fibres of $T\mathcal{M}|_{\mathcal{S}}$ are m-dimensional, while the fibres of $T\mathcal{S}$ are only s-dimensional.

However, the inclusion map $\iota : \mathcal{S} \to \mathcal{M}$ induces a smooth injective bundle morphism

$$\iota_* : T\mathcal{S} \to T\mathcal{M}|_{\mathcal{S}} \tag{4.4.2}$$

which embeds each fibre $T_p\mathcal{S} \subset T_p\mathcal{M}$ for $p \in \mathcal{S}$. As a result, $T\mathcal{S}$ is realised as a smooth vector subbundle of $T\mathcal{M}|_{\mathcal{S}}$. This perspective will be useful not only for defining and understanding the normal bundle but also for lifting previously pointwise constructions—such as the projections P and Q—into global morphisms between bundles.

Recall that for any submanifold $\mathcal{S} \subset \mathcal{M}$ of a (pseudo-)Riemannian manifold $\mathcal{M}$, we obtain a canonical orthogonal decomposition of the tangent space at each point $p \in \mathcal{S}$,

$$T_p\mathcal{M} = T_p\mathcal{S} \oplus N_p\mathcal{S}, \tag{4.4.3}$$

where $T_p\mathcal{S} \subset T_p\mathcal{M}$ is identified via the pushforward ι_* of the inclusion $\iota : \mathcal{S} \to \mathcal{M}$, and the normal space $N_p\mathcal{S}$ is defined as the orthogonal complement

$$N_p\mathcal{S} = \left(T_p\mathcal{S}\right)^\perp \subset T_p\mathcal{M} \tag{4.4.4}$$

with respect to the metric. Note that $\dim N_p\mathcal{S} = \operatorname{codim}_{\mathcal{M}} \mathcal{S}$, so that $\dim T_p\mathcal{S} + \dim N_p\mathcal{S} = \dim T_p\mathcal{M}$.

Just as bundling the tangent spaces $T_p\mathcal{S}$ yields the tangent bundle $T\mathcal{S}$, we may bundle the normal spaces to obtain another smooth vector bundle over $\mathcal{S}$. We now introduce this object more precisely.

Definition (Normal bundle) Let $\mathcal{S} \subset \mathcal{M}$ be an embedded submanifold of a (pseudo-)Riemannian manifold $\mathcal{M}$. The *normal bundle* $N\mathcal{S}$ is the smooth vector bundle over $\mathcal{S}$ defined by

$$N\mathcal{S} := \bigsqcup_{p\in\mathcal{S}} N_p\mathcal{S}, \qquad \pi : N\mathcal{S} \to \mathcal{S}, \quad (p, X) \mapsto p, \tag{4.4.5}$$

choosing the normal spaces at each $p \in \mathcal{S}$ as the fibres.

Since the decomposition

$$T_p\mathcal{M} = T_p\mathcal{S} \oplus N_p\mathcal{S} \tag{4.4.6}$$

is orthogonal and varies smoothly with p, the total bundle $T\mathcal{M}|_{\mathcal{S}}$ likewise decomposes as a direct sum of vector bundles,

$$T\mathcal{M}|_{\mathcal{S}} = T\mathcal{S} \oplus N\mathcal{S}. \tag{4.4.7}$$

Here, the direct sum $\oplus$ is understood as the fibrewise orthogonal sum within each $T_p\mathcal{M}$.

Let us make a final addition to the bundle perspective we have been building in this section. Specifically, let us consider how the orthogonal projections

$$\begin{aligned} P &: T_p\mathcal{M} \to \operatorname{im}(P) = T_p\mathcal{S} \quad \text{and} \\ Q &: T_p\mathcal{M} \to \operatorname{im}(Q) = \ker(P) = N_p\mathcal{S} \end{aligned} \tag{4.4.8}$$

are lifted from their pointwise definitions to smooth vector bundle morphisms. Since these projections were defined for each $p \in \mathcal{S}$ using the orthogonal decomposition (4.4.6) and since this splitting varies smoothly over $\mathcal{S}$, we obtain globally defined, smooth bundle morphisms

$$P : T\mathcal{M}|_{\mathcal{S}} \to T\mathcal{S}, \quad Q : T\mathcal{M}|_{\mathcal{S}} \to N\mathcal{S}, \tag{4.4.9}$$

which act fibrewise as the orthogonal projections onto $T_p\mathcal{S}$ and $N_p\mathcal{S}$, respectively. Both P and Q are idempotent ($P^2 = P$, $Q^2 = Q$) and satisfy

$$\ker P = \operatorname{im} Q \quad \text{and} \quad \ker Q = \operatorname{im} P, \tag{4.4.10}$$

pointwise. This allows us to view the splitting (4.4.7) not merely as a statement about individual tangent spaces, but as a decomposition of vector bundles, mediated by smooth projections.

5 Foliations

5.1 Motivation and Definition of Foliations

In the previous sections, we introduced and examined submanifolds of (pseudo-)Riemannian manifolds in detail, including the tensorial structures they inherit, such as the induced metric. We now broaden our perspective by turning to decompositions of *entire* manifolds (or open subsets thereof) into smooth families of non-intersecting submanifolds—as opposed to just a single isolated submanifold. These decompositions, referred to as *foliations*, naturally arise in a variety of contexts. For example, in the ADM formalism, isolating a direction of dynamical evolution —typically a timelike one—foliates spacetime into spacelike hypersurfaces facilitates . In other settings, such as the study of flows or congruences, one may be interested in decompositions into curves or integral lines, which are families of one-dimensional submanifolds.

Regardless of where they appear, the concept of splitting manifolds into lower-dimensional submanifolds is fundamental, and we now develop this notion of foliations. In this section, we first provide the definition of foliations of arbitrary codimension, as well as coordinates adapted to them. We then specialise to the codimension-1 case of *hypersurface foliations*, which are particularly relevant in the ADM formalism and for which certain equations and identities take on a simpler form.

Definition (Foliation) Let $\mathcal{M}$ be a smooth manifold. A *foliation of codimension k* is a k-parameter family $\{\Sigma_{t_0}\}_{t_0\in\mathbb{R}^k}$ of smooth, embedded submanifolds $\Sigma_{t_0}\subset\mathcal{M}$ of codimension k such that

$$\Sigma_{t_0}\cap\Sigma_{t_0'}=\emptyset \quad \text{for} \quad t_0\neq t_0', \quad \text{and} \quad \mathcal{M}=\bigcup_{t_0\in\mathbb{R}^k}\Sigma_{t_0}. \tag{5.1.1}$$

The submanifolds Σ_{t_0} are referred to as the *leaves* of the foliation.

This definition, while clean, hides a more powerful and flexible characterisation. Since the leaves are disjoint and cover all of $\mathcal{M}$, each point $p\in\mathcal{M}$ lies in a unique leaf Σ_{t_0}. We can therefore associate to each point its corresponding label $t_0=\left(t_0^A\right)$, $A=1,...,k$, giving rise to a map

$$t=\left(t^A\right):\mathcal{M}\to\mathbb{R}^k, \quad t(p)=t_0 \text{ such that } p\in\Sigma_{t_0}. \tag{5.1.2}$$

This defines k smooth scalar fields $t^A(p)$ on $\mathcal{M}$, allowing the leaves of the foliation to be expressed as the family of level sets

$$\Sigma_{t_0}=t^{-1}(t_0)=\bigcap_{A=1}^{k}\left(t^A\right)^{-1}(t_0). \tag{5.1.3}$$

In order for each Σ_{t_0} to be a smooth submanifold, we require t_0 to be a regular value of all component maps t^A of $t=\left(t^A\right)$. For this to hold for all t_0, we demand each t^A be a *submersion*, i.e.

$$dt^A\neq 0 \quad \text{everywhere} \quad \forall A=1,...,k. \tag{5.1.4}$$

We thus arrive at an equivalent perspective: a codimension-k foliation of $\mathcal{M}$ may be defined by a set of k (functionally independent, i.e. $\left\{dt^A\right\}$ is linearly independent) scalar fields $t=\left(t^A\right):\mathcal{M}\to\mathbb{R}^k$, $A=1,...,k$, whose differentials are nowhere vanishing. The intersection of their level sets then define the leaves of the foliation.

Note that because $dt^A\neq 0$, the map $t=\left(t^A\right)$ can be extended to a coordinate chart on $\mathcal{M}$, where the coordinates $\left(t^A,y^i\right)$ consist of both the foliation parameters t^A and the coordinates y^i on the leaves. The number of transverse coordiantes y^i is given by $\dim\Sigma_t=\dim\mathcal{M}-k$. Such coordinates are called *weakly adapted* to the foliation. In particular, fixing $t^A=t_0^A$ to some constant value t_0^A yields a coordinate chart (y^i) on the leaf Σ_{t_0}. For this reason, the y^i are referred to as *transverse coordinates.*

Definition (Hypersurface Foliation) We call a foliation $\Sigma=\{\Sigma_t\}$ of codimension 1 a *hypersurface foliation.*

A hypersurface foliation is defined by a single scalar field

$$t : \mathcal{M} \to \mathbb{R}, \quad dt \neq 0 \text{ everywhere}, \tag{5.1.5}$$

which can be extended to a coordinate chart as (t, y^i), $i = 1, ..., \dim \mathcal{M} - 1$.

5.2 The Normal 1-Form and Normal Vector Field

The choice of a foliation of a manifold $\mathcal{M}$ naturally gives rise to a host of associated mathematical objects, each playing its own role in its geometry. Among the most central of these is the normal vector field, together with its corresponding one-form. In what follows, we shall first build some intuition for why such an object may be constructed, before proceeding to define it rigorously.

As we have seen previously, a submanifold $\mathcal{S} \subset \mathcal{M}$ induces, at each point $p \in \mathcal{S}$, a decomposition

$$T_p\mathcal{M} = T_p\mathcal{S} \oplus N_p\mathcal{S}, \tag{5.2.1}$$

splitting the tangent space of $\mathcal{M}$ into components tangent and normal to $\mathcal{S}$. However, this decomposition is inherently local to $\mathcal{S}$; at points $p \in \mathcal{M}$ outside of $\mathcal{S}$, the notions $T_p\mathcal{S}$ and $N_p\mathcal{S}$ have no meaning. Hence one cannot define a vector field on $\mathcal{M}$ with the property that it is normal to $\mathcal{S}$ everywhere—this statement simply does not make sense away from $\mathcal{S}$.

A foliation $\Sigma = \{\Sigma_t\}_{t\in\mathbb{R}}$ changes this picture dramatically. Given such a foliation, every point $p \in \mathcal{M}$ belongs to a unique leaf $\Sigma_{t(p)}$, for a unique $t(p) \in \mathbb{R}$. Consequently, we may write

$$T_p\mathcal{M} = T_p\Sigma_{t(p)} \oplus N_p\Sigma_{t(p)}, \tag{5.2.2}$$

thereby defining a decomposition into tangent and normal directions at *every* point of $\mathcal{M}$. In other words, a foliation allows us to speak globally about directions that are tangent or normal to the slices Σ_t, since every point of $\mathcal{M}$ lies in precisely one such submanifold.

At this point it is natural to briefly revisit the notion of vector bundles. Using the decomposition (5.2.2), we may define two vector bundles over $\mathcal{M}$ that together decompose the tangent bundle $T\mathcal{M}$. Namely,

$$T\Sigma = \bigsqcup_{p\in\mathcal{M}} T_p\Sigma_{t(p)}, \qquad N\Sigma = \bigsqcup_{p\in\mathcal{M}} N_p\Sigma_{t(p)}, \tag{5.2.3}$$

equipped with their canonical projections onto $\mathcal{M}$, yield the tangent bundle $T\Sigma$ and normal bundle $N\Sigma$ of the foliation, respectively. It is then immediate that the tangent bundle of the manifold decomposes as

$$T\mathcal{M} = T\Sigma \oplus N\Sigma. \tag{5.2.4}$$

We may now also formalise the notion of a *normal vector field.* Recall that a vector field on $\mathcal{M}$ is a smooth section of the tangent bundle $T\mathcal{M}$. Given the decomposition (5.2.4) we call a vector field *tangent* to the foliation Σ if it is a smooth section of $T\Sigma$, and normal to Σ if it is a smooth section of $N\Sigma$. In more elementray terms, a vector field X is tangent (respectively normal), if at each point $p \in \mathcal{M}$, the vector $X(p)$ lies in the subspace $T_p\Sigma_t$ (respectively, $N_p\Sigma_t$) associated to the unique slice Σ_t containing p.

In the specific case of a codimension-one foliation, the normal spaces $N_p\Sigma_{t(p)}$ are one-dimensional at every point. This observation is powerful. It implies that any vector field normal to the foliation can be written as a scalar multiple of a single, globally defined, nowhere-vanishing basis vector field. Concretely, we may express any such field X as

$$X = \lambda\, n^\sharp, \tag{5.2.5}$$

where $n^\sharp$ denotes a chosen normal vector field, and λ is a smooth scalar function on $\mathcal{M}$. The use of the musical isomorphism in this notation is deliberate: rather than constructing $n^\sharp$ directly, it is often more natural to begin with a normal one-form n, and then obtain the vector field $n^\sharp$ via the defining identity

$$g\big(n^\sharp, X\big) = n(X) \quad \forall X \quad \Leftrightarrow \quad (n^\sharp)^\mu = n^\mu = g^{\mu\nu} n_\nu. \tag{5.2.6}$$

If the normal subspaces $N_p\Sigma_{t(p)}$ are timelike (or spacelike) everywhere, then we may impose a canonical normalisation to fix $n^\sharp$ uniquely (up to sign),

$$g\big(n^\sharp, n^\sharp\big) = \begin{cases} -1, & \text{timelike}, \\ +1, & \text{spacelike}. \end{cases} \tag{5.2.7}$$

Since the normal bundle has one-dimensional leaves, this condition pins down a distinguished unit normal vector field (up to sign).

We have not yet addressed how to construct such a form or field in practice—let us now turn to this task. For a vector field $n^\sharp$ to be normal to the foliation Σ, by the definition of $N_p\Sigma_{t(p)}$ as $(T_p\Sigma_{t(p)})^\perp$ it must satisfy

$$g\big(n^\sharp, X\big) = 0, \quad \forall X \in T_p\Sigma_{t(p)} \tag{5.2.8}$$

at every point $p \in \mathcal{M}$; that is, $n^\sharp$ must be orthogonal to all vectors tangent to the slice Σ_t through p. By the defining property of the musical isomorphism, this is equivalent to requiring

$$n(X) = 0 \quad \forall X \in T_p\Sigma_{t(p)} \tag{5.2.9}$$

i.e. the 1-form n must annihilate all tangent vectors to the foliation.

This insight significantly simplifies the problem in adapted coordinates. Let $x^\mu = (t, y^i)$ be a coordinate chart adapted to the foliation, so that each slice Σ_t is locally given by $t = \text{const.}$, and the y^i serve as coordinates on each leaf. Then the tangent space to Σ_t at p is spanned by the coordinate basis vectors ∂_i, i.e.

$$T_p\Sigma_{t(p)} = \operatorname{span}\left\{ \partial_i = \frac{\partial}{\partial y^i} \,\middle|\, i = 1, \ldots, \dim\mathcal{M} - 1 \right\}. \tag{5.2.10}$$

We thus seek a 1-form n satisfying

$$n(\partial_i) = 0 \quad i = 1, \ldots, \dim\mathcal{M} - 1. \tag{5.2.11}$$

The standard coordinate duality relation

$$dx^\mu(\partial_\nu) = \delta^\mu_\nu \tag{5.2.12}$$

immediately suggests a solution: the 1-form

$$n = \alpha\, dt, \quad \alpha \in C^\infty(\mathcal{M}) \tag{5.2.13}$$

clearly satisfies condition (5.2.11), as

$$n(\partial_i) = \alpha\, dt(\partial_i) = \delta^t_i = 0. \tag{5.2.14}$$

Here, α is a smooth scalar field on $\mathcal{M}$, necessary to adjust the magnitude in order to satisfy the normalisation condition (5.2.7). If we require $n^\sharp$ to be unit-normalised, this yields the relation

$$\pm 1 =: \varepsilon = g\big(n^\sharp, n^\sharp\big) = g^{\mu\nu} n_\mu n_\nu = \alpha^2 g^{tt} \quad \Leftrightarrow \quad g^{tt} = \frac{\varepsilon}{\alpha^2}. \tag{5.2.15}$$

Moreover, the components of $n^\sharp$ can be written in terms of the metric as well,

$$(n^\sharp)^\mu = n^\mu = g^{\mu\nu} n_\nu = \alpha g^{\mu\nu} \delta^t_\nu = \alpha g^{\mu t}. \tag{5.2.16}$$

In hindsight, it is only natural that the function $t : \mathcal{M} \to \mathbb{R}$ which defines the foliation Σ plays a central role in the construction of the normal form. After all, moving along the gradient of t corresponds to moving between leaves, so it makes sense that for an arbitrary set of coordinates $\bar{x}^\alpha$,

$$dt = \partial_\alpha t \, d\bar{x}^\alpha \tag{5.2.17}$$

somehow encodes the normal direction.

This construction of n as being proportional to dt was rather simple in coordinates. Keeping with the spirit of these notes, though, we should also show the equivalence

$$dt_p(X) = 0 \quad \Leftrightarrow \quad X \in T_p \Sigma_{t(p)} \tag{5.2.18}$$

in a coordinate-independent way. This is not that hard, and gives some additional geometric intuition as it makes use of the definition of tangent spaces in terms of derivatives along curves.

Proof: (of eq. (5.2.18)) We first prove the implication "$\Leftarrow$". Let $X \in T_p \Sigma_{t(p)}$. Then there exists a smooth curve $\gamma : (-\varepsilon, \varepsilon) \to \Sigma_{t(p)} \subset \mathcal{M}$ with $\gamma(0) = p$ and $\dot{\gamma}(0) = X$, such that for any $f \in C^\infty(\mathcal{M})$,

$$X[f] = \frac{d}{ds} f(\gamma(s))|_{s=0}. \tag{5.2.19}$$

In particular,

$$dt(X) = X[t] = \frac{d}{ds} t(\gamma(s)) = \frac{d}{ds} t(p) = 0, \tag{5.2.20}$$

since $\gamma(s) \subset \Sigma_t = t^{-1}(t(p))$ implies $t(\gamma(s)) = t(p) = \text{const.}$ identically.

For the converse, note that $dt \neq 0$ everywhere by construction, and $\operatorname{codim} T_p \Sigma_t(p) = 1$. Hence, for any vector $X \in T_p \mathcal{M} \setminus T_p \Sigma_{t(p)}$, we must have $dt(X) \neq 0$, otherwise dt_p would vanish on all of $T_p \mathcal{M}$, contradicting $dt \neq 0$. This implies

$$T_p \Sigma_{t(p)} = \ker\big(dt_p\big), \tag{5.2.21}$$

which establishes the equivalence. □

We can even take the pointwise identification (5.2.21) one step further. To this end, let us define the notion of a kernel of a differential 1-form $\omega \in \Gamma(T^* \mathcal{M})$. At any point p, we have

$$\ker\big(\omega_p\big) = \big\{ X \in T_p \mathcal{M} \mid \omega_p(X) = 0 \big\}, \tag{5.2.22}$$

the standard definition of the kernel from linear algebra. We can extend this to the entire differential form ω by attaching to each point in $\mathcal{M}$ the corresponding subspace $\ker\big(\omega_p\big) \subset T_p \mathcal{M}$, yielding the subbundle

$$\ker(\omega) := \bigsqcup_{p \in \mathcal{M}} \ker\big(\omega_p\big) \subset T\mathcal{M}, \tag{5.2.23}$$

together with the canonical projection map onto $\mathcal{M}$. It is now straightforward to see that when choosing $\omega = dt$, due to the identification (5.2.21) we find

$$\ker(dt) = T\Sigma \subset T\mathcal{M}. \tag{5.2.24}$$

The tangent bundle of a foliation Σ generated by the level sets of a scalar function $t : \mathcal{M} \to \mathbb{R}$ is hence simply the kernel of the 1-form dt associated to it by the exterior derivative.

5.3 Hypersurface-Orthogonal Distributions

In the previous section, we defined what it means for a vector field to be normal to a hypersurface foliation Σ generated by a scalar function $t \in C^\infty(\mathcal{M})$ with $dt \neq 0$ everywhere. In particular, we constructed an explicit example of such a vector field $n^\sharp$ by setting

$$n = \alpha\, dt \quad \text{such that} \quad g\big(n^\sharp, n^\sharp\big) = \varepsilon. \tag{5.3.1}$$

This addressed the question *"Given a hypersurface foliation, can we find a vector field that is normal to it everywhere?"*.

In this section, we turn that question around: *"Given a vector field, does there exist a hypersurface foliation to which it is everywhere normal?"*. This leads us to the notion of *hypersurface-orthogonal* vector fields—those that are locally normal to a family of hypersurfaces. Without proof, we will also give the so-called *Frobenius condition* which can be used to check this property directly.

Definition (Hypersurface-Orthogonal Vector Fields) Let $X \in \Gamma(T\mathcal{M})$ be a smooth vector field on a (pseudo-)Riemannian manifold $\mathcal{M}$. We say that X is *hypersurface-orthogonal* if, for every point $p \in \mathcal{M}$, there exists a neighbourhood $U \subset \mathcal{M}$ of p and a local foliation $\Sigma = \{\Sigma_t\}_{t\in\mathbb{R}}$ of U with the property that

$$g\big(X_p, Y\big) = 0 \quad \text{for all } Y \in T_p\Sigma_{t(p)}. \tag{5.3.2}$$

In words: X is hypersurface-orthogonal if, at every point, it is orthogonal to the leaves of some local foliation.

Though geometrically intuitive, this definition of hypersurface-orthogonality is rather difficult to verify in practice. We now derive an equivalent *analytic* criterion that allows us to check whether a given vector field $X \in \Gamma(T\mathcal{M})$ is hypersurface-orthogonal.

The derivation proceeds in a few steps:

- First, recall that any local foliation $\Sigma = \{\Sigma_t\}$ defined on an open set $U \subset \mathcal{M}$ can be represented by a smooth function $\varphi : U \to \mathbb{R}$ with $d\varphi \neq 0$, whose level sets define the leaves as $\Sigma_{\varphi(p)} = \varphi^{-1}(\varphi(p))$. Conversely, any such function defines a local foliation by level sets. In the previous section, we saw that the tangent spaces to the leaves are given by the pointwise kernels of the differential,

$$\ker\big(d\varphi_p\big) = T_p\Sigma_{\varphi(p)} \tag{5.3.3}$$

- A vector field X is orthogonal to the foliation if and only if for all $Y \in T_p\mathcal{M}$

$$0 = g\big(X_p, Y\big) = X_p^\flat(Y) \quad \Leftrightarrow \quad Y \in T_p\Sigma_{\varphi(p)}, \tag{5.3.4}$$

  i.e. $X_p^\flat$ annihilates all vectors tangent to the leaf through p. But by equation (5.3.3), those tangent vectors are precisely the kernel of $d\varphi_p$. Hence, the two kernels must agree,

$$\ker\big(X_p^\flat\big) = \ker\big(d\varphi_p\big). \tag{5.3.5}$$

- Since both $X_p^\flat$ and $d\varphi_p$ are non-vanishing 1-forms, they must be pointwise proportional,

$$X_p^\flat \propto d\varphi_p. \tag{5.3.6}$$

  That is, there exists a smooth function λ such that

$$X^\flat = \lambda\, d\varphi \tag{5.3.7}$$

  on U.

We have thus shown that X is hypersurface-orthogonal if and only if, for every point $p \in \mathcal{M}$, there exist smooth functions $\varphi, \lambda \in C^\infty(U)$ on a neighbourhood $U \subset \mathcal{M}$ of p such that

$$X^\flat = \lambda d\varphi. \tag{5.3.8}$$

This provides a concrete and (more) checkable analytic characterisation of hypersurface-orthogonality.

Equation (5.3.8) is a condition on the 1-form $X^\flat$ associated to the vector field X. Conversely, one can also say that if a 1-form ω can locally (say, on $U \subset \mathcal{M}$) be expressed as

$$\omega = \lambda d\varphi, \quad \lambda, \varphi \in C^\infty(U), \tag{5.3.9}$$

then the associated vector field $X = \omega^\sharp$ is hypersurface-orthogonal. An immediate example is when $\omega = dx^{\mu_0}$ is the differential of a fixed coordinate x^{μ_0}, which trivially has the above form and hence generates the hypersurface-orthogonal vector field

$$X = \omega^\sharp = (dx^{\mu_0})^\sharp = g^{\mu_0\nu}\partial_\nu, \tag{5.3.10}$$

which is orthogonal to the foliation generated by fixing the coordinate x^{μ_0}, i.e.

$$\Sigma_t = \{p \in U \subset \mathcal{M} \mid x^{\mu_0}(p) = t\}. \tag{5.3.11}$$

However, there are many vector fields or 1-forms for which it is nontrivial to determine whether a representation as in condition (5.3.8) exists. In such cases, the so-called *Frobenius condition* comes into play, which we formulate (but not prove) below.

Theorem (Frobenius Condition) Let $\mathcal{M}$ be a smooth manifold, $U \subset \mathcal{M}$ an open neighbourhood, and ω a differential 1-form on U. Then, ω admits a local expression of the form

$$\omega = \lambda d\varphi \tag{5.3.12}$$

for some $\lambda, \varphi \in C^\infty(U)$ if and only if the condition

$$\omega \wedge d\omega = 0 \tag{5.3.13}$$

holds.

This is now a purely computational way to verify whether a vector field X is hypersurface-orthogonal—in components, this condition reads

$$0 \stackrel{!}{=} X^\flat \wedge dX^\flat = X_\mu dx^\mu \wedge (\partial_\nu X_\lambda) dx^\nu \wedge dx^\lambda = X_\mu \partial_\nu X_\lambda dx^\mu \wedge dx^\nu \wedge dx^\lambda, \tag{5.3.14}$$

which yields the standard identity

$$X_{[\mu}\nabla_\nu X_{\lambda]} \stackrel{!}{=} 0. \tag{5.3.15}$$

Notice that we were able to replace partial by covariant derivatives due to the antisymmetrisation, which removes the (here torsion-free) connection components from the expression.

As announced, we will not prove the Frobenius condition here, at least not the difficult implication that $\omega \wedge d\omega = 0$ guarantees a decomposition into $\omega = \lambda d\varphi$. To make the claim somewhat more plausible, though, we can opt to briefly examine the converse direction, which is immediate from a direct computation—let us assume $\omega = \lambda d\varphi$, and plug in:

$$\omega \wedge d\omega = \lambda d\varphi \wedge d(\lambda d\varphi) = \lambda d\varphi \wedge d\lambda \wedge d\varphi = 0, \tag{5.3.16}$$

where the last equality holds by anti-symmetry of the wedge product.

5.4 Decomposition of the Metric under Foliation

In many applications such as ADM, it is essential to decompose a metric in terms of its contributions tangential and normal to the leaves of a hypersurface foliation. In this section, we approach this decomposition from two angles, both from a coordinate-independent and -dependent perspective.

5.4.1 Coordinate-Independent Perspective

We begin by recalling from Section 5.2 that given a hypersurface foliation $\Sigma = \{\Sigma_t\}$ of a (pseudo-)Riemannian manifold $\mathcal{M}$, the tangent bundle can be decomposed as

$$T\mathcal{M} = T\Sigma \oplus N\Sigma. \tag{5.4.1}$$

This decomposition is implemented by the corresponding bundle maps

$$P : T\mathcal{M} \to T\Sigma, \quad Q : T\mathcal{M} \to N\Sigma \tag{5.4.2}$$

—pointwise, these are simply the orthogonal projections

$$P_p : T_p\mathcal{M} \to T_p\Sigma_{t(p)}, \quad Q_p : T_p\mathcal{M} \to N_p\Sigma_{t(p)}, \quad p \in \mathcal{M}. \tag{5.4.3}$$

For any vector field $X \in \Gamma(T\mathcal{M})$, we have

$$X = P(X) + Q(X) \tag{5.4.4}$$

where $P(X) \in \Gamma(T\Sigma)$ and $Q(X) \in \Gamma(N\Sigma)$.

Alternatively, in the pointwise view, for any $X_p \in T_p\mathcal{M}$, it holds that

$$X_p = P_p(X_p) + Q_p(X_p), \tag{5.4.5}$$

such that $P_p(X_p) \in T_p\Sigma_{t(p)}$ and $Q_p(X_p) \in N_p\Sigma_{t(p)}$, and hence

$$g(P_p(X_p), Q_p(X_p)) = 0. \tag{5.4.6}$$

Further recall that the associated $(0,2)$-tensors of P and Q decompose the metric as

$$g = P + Q. \tag{5.4.7}$$

Having set the stage for orthogonal decompositions of the metric, we now turn to the special case where Σ is a hypersurface foliation generated by $t \in C^\infty(\mathcal{M})$ with $dt \neq 0$ everywhere. Our goal is to derive the orthogonal projection Q onto $N\Sigma$, as this will enable us to decompose the metric as $P + Q$ using $P = g - Q$.

To make this decomposition explicit, we now compute the form of Q in terms of the normal 1-form n we introduced in Section 5.2. More precisely, it is given by

$$n = \alpha\, dt, \tag{5.4.8}$$

with $\alpha \in C^\infty(\mathcal{M})$ fixed by the condition

$$g(n^\sharp, n^\sharp) = \varepsilon = \pm 1, \tag{5.4.9}$$

where the sign ε reflects whether $n^\sharp$ is spacelike or timelike (we do not consider the null case here). Further recall that n annihilates vectors tangent to Σ, i.e.

$$g(n^\sharp, X) = n(X) = 0 \quad \forall X \in \Gamma(T\Sigma). \tag{5.4.10}$$

Let us consider how n acts on an arbitrary vector $X \in \Gamma(N\Sigma)$. For our case of a hypersurface foliation, $N\Sigma$ is a vector bundle with one-dimensional fibres. Since $n^\sharp \in \Gamma(N\Sigma)$ is nowhere vanishing, we can express X as a scalar multiple of it, i.e.

$$X = \lambda n^\sharp, \quad \text{for some } \lambda \in C^\infty(\mathcal{M}). \tag{5.4.11}$$

This allows us to derive the action of the normal 1-form n on any $X \in \Gamma(N\Sigma)$ as

$$n(X) = \lambda n(n^\sharp) = \lambda g(n^\sharp, n^\sharp) = \varepsilon\lambda. \tag{5.4.12}$$

Multiplying this equation by $\varepsilon n^\sharp$ yields

$$\varepsilon n^\sharp \cdot n(X) = \varepsilon^2 \lambda n^\sharp = X. \tag{5.4.13}$$

Let us briefly digest this last result. It tells us that the map

$$\varepsilon n^\sharp \otimes n : T\mathcal{M} \to N\Sigma \tag{5.4.14}$$

acts as the identity when restricted to $N\Sigma$. Moreover, since n annihilates any vector in $\Gamma(T\Sigma)$ and due to the orthogonal decomposition (5.4.1), we conclude that $\varepsilon n^\sharp \otimes n$ is the orthogonal projection of $T\mathcal{M}$ onto $N\Sigma$. Consequently, we have

$$Q = \varepsilon n^\sharp \otimes n. \tag{5.4.15}$$

The metric hence decomposes as

$$\begin{aligned} g &= P + Q \\ &= (g - \varepsilon n \otimes n) + \varepsilon n \otimes n \\ &= (g - \varepsilon \alpha^2 \, dt \otimes dt) + \varepsilon \alpha^2 \, dt \otimes dt \end{aligned} \tag{5.4.16}$$

where we identify

$$P = g - \varepsilon \alpha^2 \, dt \otimes dt \tag{5.4.17}$$

as the orthogonal projector onto $T\Sigma$. Recall that P is the tangential component of the metric, pulling back to the induced metric on Σ_t as

$$\gamma = \iota^* g = \iota^* P \tag{5.4.18}$$

due to $\iota^* Q = 0$.

5.4.2 ADM-Type Metric Decomposition in Coordinates

We again assume the (pseudo-)Riemannian manifold $\mathcal{M}$ to be equipped with a hypersurface foliation $\Sigma = \{\Sigma_t\}$ generated by a scalar function $t \in C^\infty(\mathcal{M})$ with $dt \neq 0$ everywhere. We use t as a coordinate, extending it to a full, local coordinate system $x^\mu = (t, y^i)$, $i = 1, ..., \dim \mathcal{M} - 1$ by transverse coordinates y^i. Let us again consider the normal 1-form

$$n = \alpha \, dt, \tag{5.4.19}$$

subject to the condition

$$g(n^\sharp, n^\sharp) = \varepsilon. \tag{5.4.20}$$

Our goal will be to write the metric g in terms of the components of $n^\sharp$—i.e. of the function α and a vector β^i we will introduce momentarily—as well as the induced metric on the leaves, which has the components

$$\gamma_{ij} := \gamma(\partial_i, \partial_j) = g(\partial_i, \partial_j). \tag{5.4.21}$$

The vector $n^\sharp$ associated to the normal 1-form can be written in terms of its coordinate components, yielding

$$n^\sharp = \alpha\delta^t_\mu g^{\mu\nu}\partial_\nu = \alpha g^{t\mu}\partial_\mu = \alpha g^{tt}\partial_t + \alpha g^{ti}\partial_i. \tag{5.4.22}$$

Making use of the normalisation condition (5.4.20), we find

$$\varepsilon = g\big(n^\sharp, n^\sharp\big) = n\big(n^\sharp\big) = \alpha^2 g^{tt}\underbrace{dt(\partial_t)}_{=1} + \alpha^2 g^{ti}\underbrace{dt(\partial_i)}_{=0} = \alpha^2 g^{tt}, \tag{5.4.23}$$

or equivalently,

$$g^{tt} = \frac{\varepsilon}{\alpha^2}. \tag{5.4.24}$$

This is our first direct relationship between α and a component of the (inverse) metric. Inserting this back into the expansion for $n^\sharp$ above, we get

$$n^\sharp = \frac{\varepsilon}{\alpha}\partial_t + \alpha g^{ti}\partial_i. \tag{5.4.25}$$

We now introduce the *shift vector* β^i as

$$\beta^i = -\varepsilon\alpha^2 g^{ti}, \tag{5.4.26}$$

where the factor is chosen such one can extract ε/α as a common factor in $n^\sharp$, i.e.

$$n^\sharp = \frac{\varepsilon}{\alpha}(\partial_t - \beta^i\partial_i). \tag{5.4.27}$$

Equivalently, we may express ∂_t in terms of $n^\sharp$ and $\beta = \beta^i\partial_i$,

$$\partial_t = \varepsilon\alpha n^\sharp + \beta. \tag{5.4.28}$$

This allows us to derive both g_{tt} and g_{ti} via

$$\begin{aligned} g_{tt} &= g(\partial_t, \partial_t) = \varepsilon^2\alpha^2\underbrace{g\big(n^\sharp, n^\sharp\big)}_{=\varepsilon} + 2\varepsilon\alpha\underbrace{g\big(n^\sharp, \beta\big)}_{=0} + \underbrace{g(\beta, \beta)}_{=\gamma_{ij}\beta^i\beta^j} \\ &= \varepsilon\alpha^2 + \gamma_{ij}\beta^i\beta^j, \\ g_{ti} &= g(\partial_t, \partial_i) = \varepsilon\alpha\underbrace{g\big(n^\sharp, \partial_i\big)}_{=0} + g(\beta, \partial_i) = \beta^i g\big(\partial_i, \partial_j\big) = \gamma_{ij}\beta^j \end{aligned} \tag{5.4.29}$$

Here, we made use of the fact that $n^\sharp$ is normal to the foliation, whereas the ∂_i and hence also β are tangent to it—or, algebraically,

$$g\big(n^\sharp, \partial_i\big) = n(\partial_i) = \alpha\underbrace{dt(\partial_i)}_{=0} = 0. \tag{5.4.30}$$

In summary, the components of the metric are given by

$$g_{tt} = \varepsilon\alpha^2 + \gamma_{ij}\beta^i\beta^j, \qquad g_{ti} = \gamma_{ij}\beta^j, \qquad g_{ij} = \gamma_{ij}, \tag{5.4.31}$$

or in block matrix form,

$$g_{\mu\nu} = \begin{pmatrix} \varepsilon\alpha^2 + \gamma_{ij}\beta^i\beta^j & \gamma_{ij}\beta^j \\ \gamma_{ij}\beta^j & \gamma_{ij} \end{pmatrix}. \tag{5.4.32}$$

The metric tensor is hence given by

$$g = g_{\mu\nu}\, dx^\mu \otimes dx^\nu = \varepsilon\alpha^2\, dt \otimes dt + \gamma_{ij}(dy^i + \beta^i\, dt)(dy^j + \beta^j\, dt). \tag{5.4.33}$$

Remarks

- **The Lapse** In the literature, the function α is known as the *lapse function.* To understand its meaning (assuming t is a timelike coordinate), consider an observer moving orthogonally to the spatial slices Σ_t, i.e. following the integral curves of the normal vector field $n^\sharp$. Since $n^\sharp$ is normalised, the tangent vector to such a path $\gamma(\tau)$, parameterised by proper time τ, satisfies

$$\dot{\gamma} = n^\sharp. \tag{5.4.34}$$

Thus, the tangent vector is normal to the foliation, and when computing the elapsed proper time $d\tau$ between nearby leaves, only the normal part of the metric contributes. Hence, for such a path, we have

$$d\tau^2 = -g = -Q = n \otimes n = \alpha^2\, dt^2. \tag{5.4.35}$$

This implies

$$d\tau = \pm\alpha\, dt. \tag{5.4.36}$$

In other words, when moving from one hypersurface Σ_t to the next along the normal flow, the lapse function α gives the rate at which proper time τ advances with respect to coordinate time t. The larger the lapse, the "more" proper time elapses between adjacent slices for such an observer. Naturally, if one moves between slices while also displacing in tangential directions, additional contributions appear—the lapse governs the conversion only for motion orthogonal to the foliation.

- **The Shift** Given the definition of the normal 1-form,

$$n = \alpha\, dt, \tag{5.4.37}$$

one might naïvely expect the associated vector to be proportional to ∂_t, perhaps something like

$$n^\sharp = \frac{\varepsilon}{\alpha}\partial_t. \tag{5.4.38}$$

However, as we have seen before, the correct expression is

$$n^\sharp = \frac{\varepsilon}{\alpha}(\partial_t - \beta), \tag{5.4.39}$$

which includes an additional component from β along the tangential directions ∂_i. Since $n^\sharp$ is normal to the foliation, this tells us that ∂_t itself is generally *not* normal. The shift vector $\beta = \beta^i \partial_i$ encodes precisely this discrepancy—it provides the spatial correction needed to turn ∂_t into a vector normal to the leaves.

Put differently, the shift quantifies the failure of ∂_t to be orthogonal to the foliation. It tells us how much spatial displacement occurs—within the leaves themselves—when evolving forward in coordinate time. Alternatively, we can re-express this as

$$\partial_t = \varepsilon\alpha n^\sharp + \beta, \tag{5.4.40}$$

making explicit that ∂_t consists of both a normal and a tangential component. From this viewpoint, the shift is simply the tangential part of the time flow vector ∂_t.

Motion along the flow of ∂_t involves keeping the coordinates y^i fixed. Therefore, the choice of the coordinates y^i is intimately related to the shift vector β—in particular, as we will see in an upcoming example, the coordinates y^i can (typically) be chosen such that the shift is zero.

- **Induced Metric vs. Tangential Projector** We have seen before that under the orthogonal decomposition

$$g = P + Q, \tag{5.4.41}$$

with P the tangential and Q the normal projector (or rather, their associated bilinear forms), the induced metric γ is related via

$$\gamma = \iota^* g = \iota^* P + \underbrace{\iota^* Q}_{=0} = \iota^* P. \tag{5.4.42}$$

In most contexts, it is fine to say that $\gamma = P$, since they act the same on $T\Sigma$. However, it should be kept in mind that P is a bilinear form on $T\mathcal{M}$ whereas γ is only defined on $T\Sigma$. In other words, P "knows" how to handle vectors with components in $N\Sigma$—namely by mapping those components to zero—while γ, by definition, does not.

This manifests itself when considering the expression (5.4.17) for P after inserting the ADM decomposition for g:

$$P = \gamma_{ij}\big(dy^i + \beta^i\,dt\big) \otimes \big(dy^j + \beta^j\,dt\big). \tag{5.4.43}$$

This is in comparison to the induced metric on the leaves, which simply reads

$$\gamma = \gamma_{ij}\,dy^i \otimes dy^j. \tag{5.4.44}$$

The dt terms equip P with the (correct) means to handle ∂_t, which generically has both tangential ($\propto \beta$) and normal components. The tangential components are what P needs to isolate, which is why the dt bits are necessary.

However, note here that ∂_t lies outside of $T\Sigma$. Hence, it is a foreign object to the induced metric γ, and it "does not have to know" how to deal with it—it is safe to discard the dt pieces. It need not concern itself with vectors pointing "off the leaf"—that is, directions leading from one slice to another—even if they do have tangential components. The induced metric is content with only measuring vectors fully tangential to the leaf.

If the shift is zero, though, then $P = \gamma$ not just in effect but in substance. One can see this either algebraically, or intuitively as follows: when the shift vanishes, we have $\partial_t \propto n^\sharp$, meaning ∂_t is fully normal. In that case, P never has to process ∂_t at all—there is no tangential contribution to extract. Hence, no dt terms need to appear, and P reduces directly to γ.

- **Inverse Metric** Besides the ADM split of the metric, (5.4.33), one frequently needs to use the inverse metric,

$$g^{-1} = g^{\mu\nu}\partial_\mu \otimes \partial_\nu, \tag{5.4.45}$$

as well. Its components are given by

$$g^{tt} = \frac{\varepsilon}{\alpha^2}, \quad g^{ti} = -\frac{\varepsilon}{\alpha^2}\beta^i, \quad g^{ij} = \gamma^{ij} + \frac{\varepsilon}{\alpha^2}\beta^i\beta^j, \tag{5.4.46}$$

where γ^{ij} is inverse to γ_{ij} in the sense that

$$\gamma^{ik}\gamma_{kj} = \delta^i_j. \tag{5.4.47}$$

As a block matrix, this reads

$$g^{\mu\nu} = \begin{pmatrix} \varepsilon/\alpha^2 & -\varepsilon\beta^i/\alpha^2 \\ -\varepsilon\beta^i/\alpha^2 & \gamma^{ij} + \varepsilon\beta^i\beta^j/\alpha^2 \end{pmatrix}. \tag{5.4.48}$$

Moreover, the inverse metric tensor can be written as

$$\begin{aligned} g^{-1} &= \frac{\varepsilon}{\alpha^2}(\partial_t - \beta^i\partial_i) \otimes (\partial_t - \beta^j\partial_j) + \gamma^{ij}\partial_i \otimes \partial_j \\ &= \varepsilon n^\sharp \otimes n^\sharp + \gamma^{ij}\partial_i \otimes \partial_j. \end{aligned} \tag{5.4.49}$$

Notice that also here, we have a clean separation into a normal and a tangential part.

- **Metric Determinant** Another commonly encountered object related to the metric is its determinant (or the square root thereof)—let us compute it here. Recall from linear algebra the formula for the determinant of a block matrix,

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A - CD^{-1}B)\det(D), \tag{5.4.50}$$

where D is assumed to be invertible. We may apply this to the ADM form (5.4.32) of the metric to obtain

$$\begin{aligned} \det g &= \det\Big(\varepsilon\alpha^2 + \gamma_{ij}\beta^i\beta^j - \underbrace{(\gamma_{ij}\beta^j)\gamma^{ik}(\gamma_{k\ell}\beta^\ell)}_{=\gamma_{ij}\beta^i\beta^j}\Big)\det\gamma \\ &= \varepsilon\alpha^2\det\gamma. \end{aligned} \tag{5.4.51}$$

Consequently, the square root of its absolute value, $\sqrt{g} = \sqrt{|\det g|}$, is given by

$$\sqrt{g} = \alpha\sqrt{\gamma}, \tag{5.4.52}$$

where $\sqrt{\gamma} = \sqrt{|\det\gamma|}$.

5.5 Example: Foliation of $\mathbb{R}^3 \setminus \{0\}$ into Spheres

At this point, we should consider an example to solidify our grasp of hypersurface foliations and the ADM decomposition of the metric. To this end, we examine a simple (but still nontrivial) foliation of the manifold $\mathcal{M} = \mathbb{R}^3 \setminus \{0\}$, equipped with the flat Euclidean metric

$$g = dx \otimes dx + dy \otimes dy + dz \otimes dz \tag{5.5.1}$$

where x, y, z are the Cartesian coordinates on $\mathcal{M}$. We foliate it into origin-centered spheres of varying radii. Here, we will derive the objects we defined generally in the previous sections—the normal 1-form, its associated normal vector, the lapse, the shift, and the ADM decomposition of the metric, as well as the projectors P and Q onto the tangent and normal bundles of the foliation.

First, let us define the foliation by introducing the function

$$r : \mathcal{M} \to \mathbb{R}, \quad r(p) = \sqrt{x^2 + y^2 + z^2}, \quad p = (x, y, z) \tag{5.5.2}$$

and specify the leaves of the foliation $\Sigma = \{\Sigma_r\}$ to be its level sets,

$$\Sigma_{r_0} := \{p \in \mathcal{M} \mid r(p) = r_0\} = r_0 S^2. \tag{5.5.3}$$

Notice that this does indeed define a foliation, since

$$dr = \frac{x}{r}dx + \frac{y}{r}dy + \frac{z}{r}dz \neq 0 \tag{5.5.4}$$

everywhere on $\mathcal{M}$—each of the $\Sigma_r \subset \mathcal{M}$ is a submanifold in its own right.

We begin by analysing the normal 1-form and the associated vector, which also yields an explicit expression for the lapse funtion α. In general, the normal 1-form is given by

$$n = \alpha\, dr. \tag{5.5.5}$$

Since the (inverse) metric in Cartesian coordinates is trivial, we have

$$n^\sharp = \alpha\, dr^\sharp = \alpha\Big(\frac{x}{r}\partial_x + \frac{y}{r}\partial_y + \frac{z}{r}\partial_z\Big). \tag{5.5.6}$$

The normalisation condition on n hence yields

$$\begin{aligned} 1 \overset{!}{=} g\big(n^\sharp, n^\sharp\big) &= n\big(n^\sharp\big) = \alpha^2\, dr\Big(\frac{x}{r}\partial_x + \frac{y}{r}\partial_y + \frac{z}{r}\partial_z\Big) \\ &= \alpha^2\left(\frac{x^2}{r^2}\,dx(\partial_x) + \frac{y^2}{r^2}\,dy(\partial_y) + \frac{z^2}{r^2}\,dz(\partial_z)\right) \\ &= \frac{\alpha^2}{r^2}\underbrace{(x^2+y^2+z^2)}_{=r^2} = \alpha^2, \end{aligned} \tag{5.5.7}$$

which fixes the lapse as $\alpha = \pm 1$—we choose the positive sign, i.e. $\alpha = +1$. Moreover, we now have the concrete expressions

$$\begin{aligned} n =& dr = \frac{x}{r}\,dx + \frac{y}{r}\,dy + \frac{z}{r}\,dz, \\ n^\sharp =& \frac{x}{r}\partial_x + \frac{y}{r}\partial_y + \frac{z}{r}\partial_z. \end{aligned} \tag{5.5.8}$$

Let us now consider the shift vector β. For this, we need to introduce additional coordinates q^i, $i = 1, 2$ to amend r into a full coordinate system. Naturally, a great candidate for the q^i are the angles θ, φ used with spherical coordinates—for now though, let us remain general. What we have now is a relationship between two sets of coordinates,

$$\left.\begin{aligned} r &= r(x,y,z) \\ q^i &= q^i(x,y,z) \end{aligned}\right\} \quad \longleftrightarrow \quad \begin{cases} x = x(r, q^i) \\ y = y(r, q^i) \\ z = z(r, q^i) \end{cases} \tag{5.5.9}$$

Recall from the previous section that the shift β is given by

$$\beta = \partial_r - \underbrace{\varepsilon\alpha}_{=1}\, n^\sharp = \partial_r - n^\sharp, \tag{5.5.10}$$

quantifying how much tangential motion (since $\beta = \beta^i\partial_i \in \Gamma(T\Sigma)$) displacement along the flow of ∂_r entails. In terms of the Cartesian coordinate basis $\{\partial_x, \partial_y, \partial_z\}$, we may express ∂_r as

$$\partial_r = \frac{\partial x}{\partial r}\partial_x + \frac{\partial y}{\partial r}\partial_y + \frac{\partial z}{\partial r}\partial_z, \tag{5.5.11}$$

employing the multi-dimensional chain rule. We can now explicitly write down the shift as

$$\beta = \partial_r - n^\sharp = \left(\frac{\partial x}{\partial r} - \frac{x}{r}\right)\partial_x + \left(\frac{\partial y}{\partial r} - \frac{y}{r}\right)\partial_y + \left(\frac{\partial z}{\partial r} - \frac{z}{r}\right)\partial_z. \tag{5.5.12}$$

At this point, we cannot really proceed without some sort of assumption on the functions $x(r, q^i)$, $y(r, q^i)$ and $z(r, q^i)$, and by that, without assumption on the q^i. So, let us do this: we require that $\beta = 0$. In other words, we want the q^i to be such that we have zero shift, that ∂_r is normal to $T\Sigma$.

This requires that each of the coefficients of β above must be zero. Notice that

$$\frac{\partial x}{\partial r} - \frac{x}{r} = 0 \quad \Leftrightarrow \quad x(r, q^i) = r f_x(q^i) \tag{5.5.13}$$

for some function $f_x(q_i)$, and analogously for y and z. Thus, we must parameterise each of our leaves $\Sigma_r = rS^2$ with the same coordinates q^i, simply stretched by the sphere's radius. Put differently, requiring $\beta = 0$ implies that the coordinates x, y, z are simply parametrisations of the 2-sphere S^2 scaled linearly by the radius r.

All that is left to do now is to choose one's favourite parametrisation of S^2. Here, we opt to use the standard coordinates θ, φ which parameterise the 2-sphere of radius r as

$$\begin{aligned} x &= r \sin\theta \cos\varphi, \\ y &= r \sin\theta \sin\varphi, \\ z &= r \cos\theta. \end{aligned} \tag{5.5.14}$$

Let us now turn to the decomposition of the metric. We already have the lapse and shift, $\alpha = 1$ and $\beta = 0$—all that remains to be derived are the transverse metric components

$$\gamma_{ij} = g\big(\partial_i, \partial_j\big), \quad i, j \in \{\theta, \varphi\}. \tag{5.5.15}$$

The relevant vectors are

$$\begin{aligned} \partial_\theta &= r\cos\theta\big(\cos\varphi\,\partial_x + \sin\varphi\,\partial_y\big) - r\sin\theta\,\partial_z, \\ \partial_\varphi &= r\sin\theta\big(-\sin\varphi\,\partial_x + \cos\varphi\,\partial_y\big). \end{aligned} \tag{5.5.16}$$

A further mechanical computation reveals that

$$\gamma_{\theta\theta} = r^2, \quad \gamma_{\theta\varphi} = 0, \quad \gamma_{\varphi\varphi} = r^2 \sin^2\theta. \tag{5.5.17}$$

We are now ready to write down the full ADM decomposition of the metric. Inserting into the general result (5.4.33), we find

$$\begin{aligned} g &= \varepsilon\alpha^2\, dr \otimes dr + \gamma_{ij}\big(dq^i + \beta^i\, dr\big)\big(dq^j + \beta^j\, dr\big) \\ &= dr \otimes dr + r^2\, d\theta \otimes d\theta + r^2 \sin^2\theta\, d\varphi \otimes d\varphi \end{aligned} \tag{5.5.18}$$

At this point, one might reasonably point out that we've essentially just recovered the familiar expression for the flat metric in spherical coordinates through a rather elaborate detour—suspiciously close to plainly reinventing them. That, however, would be a disheartening conclusion after having put in all this effort—so let us not argue that. Instead, let us appreciate having seen the ADM decomposition machinery at work in a setting where the outcome is familiar, but derived by taking the scenic route. After all, we've just learned that spherical coordinates arise naturally when seeking a zero-shift ADM decomposition of the flat Euclidean metric on concentric spheres. This, as every self-respecting postgraduate knows, is the entire point of advanced studies: to rediscover well-known results in a more baroque fashion—ideally in a way that makes the undergrads look impressed.

Now that we have derived a particular ADM decomposition of the flat Euclidean metric on $\mathcal{M} = \mathbb{R}^3 \setminus \{0\}$, we can also briefly consider the bilinear forms P, Q associated to the orthogonal decomposition

$$T\mathcal{M} = T\Sigma \oplus N\Sigma. \tag{5.5.19}$$

The normal part of the metric, Q, is simply given by

$$Q = \varepsilon n \otimes n = dr \otimes dr. \tag{5.5.20}$$

Correspondingly, the tangential projection P—or equivalently (in this case, since $\beta = 0$), the induced metric γ on the leaves—is given by

$$\gamma = P = g - Q = r^2\big(d\theta \otimes d\theta + \sin^2\theta\, d\varphi \otimes d\varphi\big). \tag{5.5.21}$$

This is what one would expect: an appropriately scaled metric on S^2.

6 Covariant Derivatives on Foliations and Submanifolds

In addition to the metric, a manifold can be endowed with various other geometric structures. This section focuses on *affine connections*, with an emphasis on how a connection on a manifold $\mathcal{M}$ induces a connection on the leaves of a foliation Σ of $\mathcal{M}$. We begin by reviewing the definition of a connection ∇ (briefly, familiarity with the concept is assumed), and then explore how the conditions of vanishing torsion and metric compatibility uniquely determine the Levi-Civita connection. Next, we define the induced connection $\bar{\nabla}$ on Σ as the tangential projection of ∇ on $\mathcal{M}$, and demonstrate that if ∇ is Levi-Civita, then so is ∇.

6.1 Review: Affine Connections

To review the notion of affine connections, we must first introduce some additional notation. Let

$$T^{(r,s)}\mathcal{M} = (T\mathcal{M})^{\otimes r} \otimes (T^*\mathcal{M})^{\otimes s} \tag{6.1.1}$$

denote the *bundle of* (r,s)*-tensors on* $\mathcal{M}$, where the tensor product between vector bundles is to be understood pointwise—i.e., as a tensor product of the fibres at each point. Tensor fields are smooth sections of this bundle; in other words, an (r,s)-tensor field T is a section of $T^{(r,s)}\mathcal{M}$, that is, $T \in \Gamma\big(T^{(r,s)}\mathcal{M}\big)$. With this notational preface out of the way, we are now ready to give the definition of affine connections.

Definition (Affine Connection) Let $\mathcal{M}$ be a smooth manifold, and consider a linear map

$$\nabla : \Gamma\big(T^{(r,s)}\mathcal{M}\big) \to \Gamma\big(T^{(r,s+1)}\mathcal{M}\big) = \Gamma\big(T^*\mathcal{M} \otimes T^{(r,s)}\mathcal{M}\big), \tag{6.1.2}$$

which maps (r,s)-tensors to $(r,s+1)$-tensors.

With the interpretation of a tensor as a multilinear map on a set of vectors $Y_1, ..., Y_s \in \Gamma(T\mathcal{M})$ and covectors $\omega_1, ..., \omega_r \in \Gamma(T^*\mathcal{M})$, we introduce the shorthand notation

$$\nabla_X T(\omega_1, ..., \omega_r, Y_1, ..., Y_r) = (\nabla T)(X, \omega_1, ..., \omega_r, Y_1, ..., Y_r), \tag{6.1.3}$$

which simply indicates that ∇_X populates the additional vector argument—introduced by passing from T to ∇T—with the vector X.

We call ∇ an *affine connection* or *covariant derivative* if it satisfies:

1. Reduction to the exterior derivative on functions $\varphi \in C^\infty(\mathcal{M})$,

$$\nabla\varphi = d\varphi, \tag{6.1.4}$$

2. Tensorial Leibniz rule,

$$\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S), \tag{6.1.5}$$

3. and compatibility with contraction[5] of a contra- and a covariant index,

$$\nabla\Big(\mathrm{tr}^i_j\, T\Big) = \mathrm{tr}^i_j(\nabla T). \tag{6.1.6}$$

This is a rather abstract (though likely familiar, if you got this far in these notes) definition, which benefits from building some intuition. Let us make a few remarks:

[5]The contraction operator tr^i_j acts on a tensor by contracting its i-th contravariant with the j-th covariant index when written in components.

- **Connection as Derivative** The first axiom is the most concrete—requiring that the connection reduce to the exterior derivative when acting on functions immediately introduces a differential structure to ∇. In this context, it makes sense to consider the associated map ∇_X for a vector field $X \in \Gamma(T\mathcal{M})$, and observe what it reduces to on a smooth function $\varphi \in C^\infty(\mathcal{M})$. By definition,

$$\nabla_X \varphi = d\varphi(X) = X[\varphi], \tag{6.1.7}$$

so the covariant derivative along X coincides with the familiar directional derivative. This anchors the definition in standard calculus and motivates the interpretation of ∇ as a generalisation of differentiation to tensor fields. Since both ∇ and ∇_X map tensors to tensors of the same type (modulo the extra slot), they provide a coordinate-independent extension of the partial derivative which respects the underlying bundle structure.

- **Role of the Connection on Vectors** In components associated to some coordinate system x^μ, we may locally expand a vector field $X \in T\mathcal{M}$ as

$$X = X^\mu \partial_\mu = X^\mu \otimes \partial_\mu. \tag{6.1.8}$$

While the tensor product notation is formally unnecessary here—since we are simply multiplying smooth functions with the coordinate basis vector fields—it clarifies the use of the Leibniz rule in what follows.

It is best to think of ∇ as acting on X by acting separately on the component functions $X^\mu \in C^\infty(\mathcal{M})$ and the basis vector fields ∂_μ,

$$\nabla X = \nabla\big(X^\mu \otimes \partial_\mu\big) = dX^\mu \otimes \partial_\mu + X^\mu \otimes \nabla\partial_\mu. \tag{6.1.9}$$

The components X^μ are smooth functions, so by the first axiom of the connection, $\nabla X^\mu = dX^\mu$. However, the axioms impose no direct condition on $\nabla\partial_\mu$. Its role is to restore tensoriality: the first term alone is not a tensor, which we now demonstrate.

Consider the term $dX^\mu \otimes \partial_\mu$ in components, i.e.

$$dX^\mu \otimes \partial_\mu = (\partial_\nu X^\mu)\, dx^\nu \otimes \partial_\mu. \tag{6.1.10}$$

Under a coordinate transformation $x^\mu \to y^\alpha$ with Jacobians

$$J^\mu_\alpha = \frac{\partial x^\mu}{\partial y^\alpha}, \quad J^\alpha_\mu = \frac{\partial y^\alpha}{\partial x^\mu}, \tag{6.1.11}$$

the transformed components of a vector field read $X^\alpha = J^\alpha_\mu X^\mu$. Differentiating yields

$$\partial_\alpha X^\beta = J^\nu_\alpha \partial_\nu\big(J^\beta_\mu X^\mu\big) = J^\nu_\alpha J^\beta_\mu \partial_\nu X^\mu + X^\mu\big(J^\nu_\alpha \partial_\nu J^\beta_\mu\big). \tag{6.1.12}$$

The first term transforms as expected, but the second term spoils tensoriality—it introduces an inhomogeneous piece. Therefore,

$$dX^\mu \otimes \partial_\mu \text{ is not a tensor.} \tag{6.1.13}$$

This is precisely the failure remedied by the $X^\mu \otimes \nabla\partial_\mu$ term. Writing out the coordinate-transformed derivative in the new basis leads to

$$\big(\partial_\alpha X^\beta\big)\, dy^\alpha \otimes \partial_\beta = \big[J^\nu_\alpha J^\beta_\mu \partial_\nu X^\mu + X^\mu\big(J^\nu_\alpha \partial_\nu J^\beta_\mu\big)\big]\, dx^\nu \otimes \partial_\mu. \tag{6.1.14}$$

We see that the connection term must absorb the inhomogeneous contribution. That is,

$$X^\alpha \otimes \nabla\partial_\alpha = X^\mu \otimes \nabla\partial_\mu - X^\mu\big(J^\nu_\alpha \partial_\nu J^\beta_\mu\big)\, dx^\nu \otimes \partial_\mu. \tag{6.1.15}$$

The final term determines how $\nabla\partial_\mu$ must transform in order to make ∇X a genuine tensor. We will return to this structure momentarily when introducing the connection coefficients explicitly.

- **Connection Coefficients** As we have seen in the previous remark, the object $\nabla\partial_\mu$ is of particular interest—we should write it in terms of components. To this end, let us introduce additional shorthand notation for covariant derivatives along the coordinate directions ∂_μ, as

$$\nabla_\mu := \nabla_{\partial_\mu}. \tag{6.1.16}$$

From the definition of affine connections and the derivations in the previous remark, it is clear that the object $\nabla\partial_\nu$ must have the component form

$$\nabla\partial_\nu = \left(\nabla_\mu\partial_\nu\right) \otimes dx^\mu = \Gamma^\lambda{}_{\mu\nu}\, dx^\mu \otimes \partial_\lambda, \quad \text{for coeffecients} \quad \Gamma^\lambda{}_{\mu\nu} \in C^\infty(\mathcal{M}). \tag{6.1.17}$$

This implicitly defines the *connection coefficients* $\Gamma^\lambda{}_{\mu\nu}$ through

$$\Gamma^\lambda{}_{\mu\nu}\partial_\lambda = \nabla_\mu\partial_\nu. \tag{6.1.18}$$

Fully explicitly, they are given by

$$\Gamma^\lambda{}_{\mu\nu} = dx^\lambda\left(\nabla_\mu\partial_\nu\right). \tag{6.1.19}$$

Using these coefficients, we can write the action of the connection on a vector $X = X^\mu\partial_\mu \in \Gamma(T\mathcal{M})$ concretely as

$$\begin{aligned} \nabla X &= \nabla(X^\nu\partial_\nu) = dX^\nu \otimes \partial_\nu + X^\nu\nabla\partial_\nu \\ &= \left(\partial_\mu X^\lambda\right) dx^\mu \otimes \partial_\lambda + \left(X^\nu\Gamma^\lambda{}_{\mu\nu}\right) dx^\mu \otimes \partial_\lambda \\ &= \left(\partial_\mu X^\lambda + \Gamma^\lambda{}_{\mu\nu}X^\nu\right) dx^\mu \otimes \partial_\lambda. \end{aligned} \tag{6.1.20}$$

Along the coordinate directions, this turns into

$$\nabla_\mu X = \left(\partial_\mu X^\lambda + \Gamma^\lambda{}_{\mu\nu}X^\nu\right)\partial_\lambda. \tag{6.1.21}$$

This is nothing but the partial derivative of X, amended by a term linear in X that ensures tensoriality of ∇X.

- **Interpretation of the Connection** In analogy to our first remark that the connection along some vector $X \in \Gamma(T\mathcal{M})$ is a covariant generalisation of the directional derivative along X, we can give an interpretation to the equality

$$\nabla_\mu\partial_\nu = \Gamma^\lambda{}_{\mu\nu}\partial_\lambda. \tag{6.1.22}$$

On the left-hand side, we compute a generalised directional derivative along the coordinate direction ∂_μ. In other words, we ask ourselves: how does the basis vector ∂_ν change as we move along the direction ∂_μ? This question is answered by the right-hand side, which tells us, for the given combination of μ and ν, what the rate of change is, in terms of a linear combination of the coordinate basis vectors. So, the component $\Gamma^\lambda{}_{\mu\nu}$ encodes information about "by how much of ∂_λ does ∂_ν change when moving along the flow of ∂_μ?". A connection hence imposes a relationship between vectors in infinitesimally neighbouring tangent spaces.

A priori, there exists no preferred way to specify this relationship between vectors at different points of the manifold—that is, to compare vectors in distinct tangent spaces and transport them around the manifold. While the transformation law of the connection coefficients is constrained by the requirement that ∇T be tensorial for all tensors T, the connection itself remains arbitrary otherwise. However, by imposing additional (natural) conditions on ∇, one can uniquely determine a distinguished connection: the *Levi-Civita connection.* We will pursue this in the next section.

- **Coordinate Transformation Behaviour** From equation (6.1.15) it can be derived that under a coordinate transformation $x^\mu \to y^\alpha$, the connection coefficients

$$dx^\lambda\big(\nabla_\mu \partial_\nu\big) = \Gamma^\lambda{}_{\mu\nu} \tag{6.1.23}$$

must transform as

$$\Gamma^\lambda{}_{\mu\nu} \to \Gamma^\alpha{}_{\beta\gamma} = J^\alpha_\lambda J^\mu_\beta J^\nu_\gamma \Gamma^\lambda{}_{\mu\nu} + J^\alpha_\lambda \partial_\beta J^\lambda_\gamma. \tag{6.1.24}$$

The second term is what makes the connection coefficients transform non-tensorially. Unlike the first term, which is the usual triple contraction with (inverse) Jacobians for a $(1,2)$-tensor, the inhomogeneous second term violates tensorial transformation behaviour.

- **Action on Arbitrary Tensors** Since we know how ∇ acts on the coordinate basis vector fields ∂_μ, we can readily extend it to any contravariant tensor $T \in \Gamma\big(T^{(r,0)}\mathcal{M}\big)$ by using the Leibniz rule,

$$\begin{aligned}
\nabla T &= \nabla\Big(T^{\mu_1\cdots\mu_r}\partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_r}\Big) \\
&= dT^{\mu_1\cdots\mu_r} \otimes \partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_r} + T^{\mu_1\cdots\mu_r}\Big(\nabla\partial_{\mu_1}\Big) \otimes \partial_{\mu_2} \otimes \ldots \otimes \partial_{\mu_r} + \ldots \\
&= \Big(\partial_\nu T^{\mu_1\cdots\mu_r} + \Gamma^{\mu_1}{}_{\nu\lambda} T^{\lambda\mu_2\cdots\mu_r} + \ldots + \Gamma^{\mu_r}{}_{\nu\lambda} T^{\mu_1\cdots\mu_{r-1}\nu}\Big)\, dx^\nu \otimes \partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_r}.
\end{aligned} \tag{6.1.25}$$

In essence, every contravariant index produces a term where the connection coefficients are contracted with the tensor components, as by the Leibniz rule, ∇ acts on each of the basis vectors once.

For tensors that are not purely contravariant, i.e., for general tensors $T \in \Gamma\big(T^{(r,s)}\mathcal{M}\big)$, the general basis expansion reads

$$T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s}\partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_r} \otimes dx^{\nu_1} \otimes \ldots \otimes dx^{\nu_s}. \tag{6.1.26}$$

To be able to extend the action of ∇ for general tensors—as above, by means of the Leibniz rule—we must know what

$$\nabla\, dx^\mu \tag{6.1.27}$$

evaluates to. This is where the (so far unused!) third connection axiom comes into play, the compatibility with contractions: together with the Leibniz rule, it is possible to derive an explicit expression for the above from what we already know. Consider an arbitrary vector $X = X^\mu\partial_\mu \in \Gamma(T\mathcal{M})$ as well as an arbitrary 1-form $\omega = \omega_\mu\, dx^\mu \in \Gamma(T^*\mathcal{M})$. Clearly, since $\omega(X) \in C^\infty(\mathcal{M})$, we have

$$\mathrm{tr}^1_2\, \nabla(\omega \otimes X) = \nabla(\omega(X)) = d(\omega(X)) = d\big(\omega_\mu X^\mu\big). \tag{6.1.28}$$

Let us take a closer look at the left-hand side. We can derive

$$\begin{aligned}
&\mathrm{tr}^1_2\, \nabla(\omega \otimes X) = \mathrm{tr}^1_2\, \nabla\big(\omega_\mu X^\nu\, dx^\mu \otimes \partial_\nu\big) \\
&= \mathrm{tr}^1_2\big[d\big(\omega_\mu X^\nu\big) \otimes dx^\mu \otimes \partial_\nu + \omega_\mu X^\nu((\nabla\, dx^\mu) \otimes \partial_\nu + dx^\mu \otimes \nabla\partial_\nu)\big] \\
&= d\big(\omega_\mu X^\mu\big) + \omega_\mu X^\nu\, \mathrm{tr}^1_2((\nabla\, dx^\mu) \otimes \partial_\nu + dx^\mu \otimes \nabla\partial_\nu).
\end{aligned} \tag{6.1.29}$$

Inserting this back into equation (6.1.28) above and canceling the exterior derivative terms, by arbitrariness of ω_μ and X^μ we get

$$\mathrm{tr}^1_2[(\nabla dx^\mu)\otimes \partial_\nu] = -\,\mathrm{tr}^1_2[dx^\mu \otimes \nabla\partial_\nu]. \tag{6.1.30}$$

Evaluating this along a coordinate direction ∂_λ, we find

$$\begin{aligned}\mathrm{tr}^1_1[\nabla_\lambda dx^\mu \otimes \partial_\nu] &= -\,\mathrm{tr}^1_1[dx^\mu \otimes \nabla_\lambda\partial_\nu] = -\,\mathrm{tr}^1_1\left[dx^\mu \otimes \Gamma^\rho{}_{\lambda\nu}\partial_\rho\right] \\ &= -\Gamma^\rho{}_{\lambda\nu}\, dx^\mu\big(\partial_\rho\big) = -\Gamma^\rho{}_{\lambda\nu}\delta^\mu_\rho = -\Gamma^\mu{}_{\lambda\nu}.\end{aligned} \tag{6.1.31}$$

The left hand side simplifies as

$$\mathrm{tr}^1_1[\nabla_\lambda dx^\mu \otimes \partial_\nu] = (\nabla_\lambda dx^\mu)(\partial_\nu), \tag{6.1.32}$$

which implies that overall,

$$\nabla_\lambda dx^\mu = -\Gamma^\mu{}_{\lambda\nu}\, dx^\nu. \tag{6.1.33}$$

Thus, in summary, ∇ acts on the coordinate basis vectors and 1-forms as

$$\begin{aligned}\nabla\partial_\mu &= \Gamma^\lambda{}_{\nu\mu}\, dx^\nu \otimes \partial_\lambda, \\ \nabla dx^\mu &= -\Gamma^\mu{}_{\nu\lambda}\, dx^\nu \otimes dx^\lambda,\end{aligned} \tag{6.1.34}$$

or, along a coordinate direction ∂_μ,

$$\begin{aligned}\nabla_\mu\partial_\nu &= \Gamma^\lambda{}_{\mu\nu}\partial_\lambda, \\ \nabla_\mu dx^\nu &= -\Gamma^\nu{}_{\mu\lambda}\, dx^\lambda.\end{aligned} \tag{6.1.35}$$

This, together with the axiom that $\nabla\varphi = d\varphi$ (or $\nabla_\mu\varphi = \partial_\mu\varphi$), allows one to evaluate the action of ∇ (or ∇_μ) on arbitrary tensors.

To reiterate: the expressions for $\nabla_\mu\partial_\nu$ and $\nabla_\mu dx^\nu$ above inform us what the rate of change of the basis vectors ∂_ν and 1-forms dx^ν are as one moves along the coordinate direction ∂_μ.

6.2 Review: Levi-Civita Connection

Now that we have introduced the general notion of a connection—a way of encoding the change of basis vectors and 1-forms as one moves through a manifold—we turn to the most prominent example: the *Levi-Civita connection*, which is uniquely determined by two conditions. These are:

1. Vanishing torsion,
2. Metric compatibility.

We will now introduce the meaning of both conditions rigorously and examine their consequences for the connection coefficients $\Gamma^\lambda{}_{\mu\nu}$, which ultimately leads to the *Christoffel Symbols*, which are the coefficients of the Levi-Civita connection.

Vanishing Torsion A connection provides us with a means of comparing vectors at nearby points and describing how they change as we move infinitesimally along a given direction. Given two vector fields $X, Y \in \Gamma(T\mathcal{M})$, the covariant derivatives $\nabla_X Y$ and $\nabla_Y X$ describe how Y changes along the flow of X, and how X changes along the flow of Y, respectively.

On a flat space, one can think of the vectors $X, Y, X + \varepsilon\nabla_Y X$, and $Y + \varepsilon\nabla_X Y$ as forming a parallelogram for infinitesimal α, assuming $[X, Y] = 0$. However, on a general manifold with arbitrary connection, this parallelogram may fail to close. There are two distinct reasons for this failure:

1. The flow paths of X and Y do not commute, i.e., following X then Y leads to a different point than following Y then X. This is encoded by the Lie bracket $[X, Y]$.

2. The change in the transported vector fields differs, i.e., $\nabla_X Y \neq \nabla_Y X$. This is an intrinsic feature of the connection.

The *torsion* is designed to isolate the second phenomenon: it captures the failure of symmetry in the connection itself, independent of the non-commutativity of the vector fields.

Definition (Torsion) Let $\mathcal{M}$ be a smooth manifold, and ∇ a connection on $\mathcal{M}$. The torsion is the vector-valued map defined by

$$\begin{aligned} T &: \Gamma(T\mathcal{M}) \otimes \Gamma(T\mathcal{M}) \to \Gamma(T\mathcal{M}), \\ T(X,Y) &:= \nabla_X Y - \nabla_Y X - [X,Y]. \end{aligned} \tag{6.2.1}$$

The first two terms, $\nabla_X Y - \nabla_Y X$, measure the asymmetry of the connection, but also include contributions from the possible non-closure of paths due to non-commuting vector fields. The Lie bracket $[X,Y]$ encodes this latter effect—by subtracting it, we isolate the connection's contribution to the failure of the parallelogram to close, i.e., the torsion.

If $T = 0$, the connection is said to be *torsion-free*. To make this condition explicit, and to check its consequences on the connection coefficients, we now express the torsion in terms of components. Take $X = X^\mu \partial_\mu$, $Y = Y^\nu \partial_\nu$; then

$$\begin{aligned} T(X,Y) &= X^\mu \nabla_\mu (Y^\nu \partial_\nu) - \big(X^\mu \partial_\mu Y^\nu\big)\partial_\nu - (X \leftrightarrow Y) \\ &= X^\mu \Big(\cancel{(\partial_\mu Y^\nu)}\partial_\nu + Y^\nu \underbrace{\nabla_\mu \partial_\nu}_{=\Gamma^\lambda{}_{\mu\nu}\partial_\lambda}\Big) - \cancel{\big(X^\mu \partial_\mu Y^\nu\big)}\partial_\nu - (X \leftrightarrow Y) \\ &= X^\mu Y^\nu \Gamma^\lambda{}_{\mu\nu} \partial_\lambda - (X \leftrightarrow Y) \\ &= X^\mu Y^\nu \big(\Gamma^\lambda{}_{\mu\nu} - \Gamma^\lambda{}_{\nu\mu}\big) \end{aligned} \tag{6.2.2}$$

Note that subtracting the Lie bracket was essential to remove contributions from non-commutativity of the vector fields X and Y.

The torsion thus has the following properties:

1. It is antisymmetric: $T(X,Y) = -T(Y,X)$;
2. It is $C^\infty(\mathcal{M})$-linear in both arguments;
3. It can be written in terms of components as

$$T(X,Y) = X^\mu Y^\nu T^\lambda{}_{\mu\nu} \partial_\lambda \quad \text{where} \quad T^\lambda{}_{\mu\nu} = T^\lambda{}_{[\mu\nu]} = \Gamma^\lambda{}_{\mu\nu} - \Gamma^\lambda{}_{\nu\mu}, \tag{6.2.3}$$

   and is therefore proportional to the anti-symmetric part of the connection coefficients in the lower indices;

4. It is tensorial, since the transformation behaviour under $x^\mu \to y^\alpha$ is given by

$$T^\lambda{}_{\mu\nu} \propto \Gamma^\lambda{}_{[\mu\nu]} \to \Gamma^\alpha{}_{[\beta\gamma]} = J^\alpha_\lambda J^\mu_\beta J^\nu_\gamma \Gamma^\lambda{}_{[\mu\nu]} + J^\alpha_\lambda \underbrace{\partial_{[\beta} J^\lambda_{\gamma]}}_{=0}. \tag{6.2.4}$$

   The inhomogeneous part cancels under anti-symmetrisation because

$$\partial_\beta J^\lambda_\gamma = \frac{\partial^2 x^\lambda}{\partial y^\beta \partial y^\gamma} \tag{6.2.5}$$

   is symmetric in β and γ.

This fourth property ensures that $T = 0$ is a meaningful, coordinate-independent condition. If torsion were not tensorial, the expression $T(X, Y) = 0$ could be true in one chart and false in another.

We are now ready to impose the condition of vanishing torsion. By the third property above, we have

$$T = 0 \quad \Leftrightarrow \quad T^{\lambda}{}_{\mu\nu} = \Gamma^{\lambda}{}_{\mu\nu} - \Gamma^{\lambda}{}_{\nu\mu} = 0 \quad \Leftrightarrow \quad \Gamma^{\lambda}{}_{\mu\nu} = \Gamma^{\lambda}{}_{(\mu\nu)}, \tag{6.2.6}$$

i.e., a connection is torsion-free if and only if the connection coefficients are symmetric in their lower indices.

Metric Compatibility Luckily, this second condition is simpler to motivate and introduce than that of vanishing torsion. Recall that any connection ∇ is, by definition, compatible with contractions between one co- and one contravariant index,

$$\nabla \operatorname{tr}^i_j T = \operatorname{tr}^i_j \nabla T. \tag{6.2.7}$$

But as we know, contractions are not restricted to mixed index types—the metric (or its inverse) also allows contractions of two covariant or two contravariant indices.

Coordinate-independently, these operations take the form

$$T \mapsto \operatorname{tr}^{i,j}_{1,2}(g \otimes T) \quad \text{or} \quad T \mapsto \operatorname{tr}^{1,2}_{i,j}(g^{-1} \otimes T), \tag{6.2.8}$$

depending on whether one is contracting co- or contravariant indices. Perhaps more concretely, in terms of components, this becomes

$$T^{\mu_1 \dots \mu_r}_{\nu_1 \dots \nu_s} \mapsto g_{\mu_i \mu_j} T^{\mu_1 \dots \mu_r}_{\nu_1 \dots \nu_s} \quad \text{or} \quad T^{\mu_1 \dots \mu_r}_{\nu_1 \dots \nu_s} \mapsto g^{\nu_i \nu_j} T^{\mu_1 \dots \mu_r}_{\nu_1 \dots \nu_s}. \tag{6.2.9}$$

It is not far-fetched to ask that the connection be compatible with such contractions as well—after all, contraction through raising and lowering indices is a natural and frequent operation.

To formalise this, we demand that contractions involving the metric commute with covariant differentiation. Explicitly, we require[6]

$$\nabla\Big(\operatorname{tr}^{i,j}_{1,2}\, g \otimes T\Big) = \operatorname{tr}^{i,j}_{1,2}\, g \otimes (\nabla T). \tag{6.2.10}$$

Using the axioms for ∇, we expand the left-hand side into

$$\nabla\Big(\operatorname{tr}^{i,j}_{1,2}\, g \otimes T\Big) = \operatorname{tr}^{i,j}_{1,2}\, \nabla(g \otimes T) = \operatorname{tr}^{i,j}_{1,2}\, g \otimes (\nabla T) + \operatorname{tr}^{i,j}_{1,2}(\nabla g) \otimes T \tag{6.2.11}$$

The first term matches the desired right-hand side, whereas the remaining second term vanishes if and only if $\nabla g = 0$.

Hence, *metric compatibility* is the requirement that the covariant derivative of the metric vanishes,

$$\nabla g = 0. \tag{6.2.12}$$

Notice that this also covers the case involving the inverse metric, since $\nabla g = 0 \Leftrightarrow \nabla g^{-1} = 0$. In components, the condition $\nabla g = 0$ reads

$$\begin{aligned} 0 \overset{!}{=} \nabla g = \nabla\big(g_{\mu\nu}\, dx^{\mu} \otimes dx^{\nu}\big) &= \Big(\partial_{\lambda} g_{\mu\nu} - g_{\rho\nu} \Gamma^{\rho}{}_{\lambda\mu} - g_{\mu\rho} \Gamma^{\rho}{}_{\lambda\nu}\Big)\, dx^{\lambda} \otimes dx^{\mu} \otimes dx^{\nu} \\ &= \big(\partial_{\lambda} g_{\mu\nu} - \Gamma_{\nu\lambda\mu} - \Gamma_{\mu\lambda\nu}\big)\, dx^{\lambda} \otimes x^{\mu} \otimes dx^{\nu}, \end{aligned} \tag{6.2.13}$$

[6]For clarity we suppress any shifts in index positions introduced by tensor products.

where we defined $\Gamma_{\lambda\mu\nu} = g_{\lambda\rho}\Gamma^{\rho}{}_{\mu\nu}$. Making use of the torsion-freeness—i.e., symmetry of the connection coefficients in the last two indices—this is equivalent to

$$g_{\mu\nu,\lambda} = \Gamma_{\mu\nu\lambda} + \Gamma_{\nu\mu\lambda}, \tag{6.2.14}$$

with $g_{\mu\nu,\lambda} = \partial_\lambda g_{\mu\nu}$. This is a direct relationship between the derivatives of the metric and the connection coefficients.

Since we are attempting to single out a specific connection by imposing constraints, the next step is to solve the equation above for $\Gamma_{\mu\nu\lambda}$, from which the explicit form of the connection coefficients can be extracted using the inverse metric. This derivation is largely an exercise in index manipulation and combining permutations of the above equation to isolate the connection. The result is the expression for the *Christoffel symbols*,

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}\left[g_{\rho\mu,\nu} + g_{\rho\nu,\mu} - g_{\mu\nu,\rho}\right]. \tag{6.2.15}$$

The fact that this is the correct expression can also be checked by inserting equation (6.2.14) to replace the expressions $g_{\mu\nu,\lambda}$ with the appropriate index permutations and verifying that the terms cancel correctly. This can be done in a few lines.

Remark The Christoffel symbols (6.2.15) above are the coefficients of the so-called *Levi-Civita connection*, which is the *unique* connection that is both torsion-free and metric-compatible.

6.3 Induced Connection on Foliations

The goal of this section is to define an affine connection on the submanifolds constituting a foliation Σ of a (pseudo-)Riemannian manifold $\mathcal{M}$. Just as with the ambient metric and the choice of induced metric on submanifolds, one is, in principle, free to assign any connection to these submanifolds. However, if $\mathcal{M}$ is already equipped with an affine connection ∇, it is natural to study a connection *induced* by this ambient structure, rather than constructing one from scratch. We begin by developing an intuitive, geometric picture of how such an induced connection should behave, before proceeding to formalise it in a precise definition.

As we have a seen, an affine connection is fully characterised by its action on vector fields, encoded via the connection coefficients $\Gamma^{\lambda}{}_{\mu\nu}$ appearing in

$$\nabla\partial_\nu = \Gamma^{\lambda}{}_{\mu\nu}\, dx^\mu \otimes \partial_\lambda. \tag{6.3.1}$$

Its action on functions $\varphi \in C^\infty(\mathcal{M})$, 1-forms, and general tensors is then determined entirely by the axioms of the connection. For example,

$$\nabla\varphi = d\varphi, \quad \nabla dx^\nu = -\Gamma^{\nu}{}_{\mu\lambda}\, dx^\mu \otimes dx^\lambda. \tag{6.3.2}$$

In other words, specifying how ∇ acts on vector fields suffices to determine its behaviour on all tensor fields.

Thus, the task of defining a geometrically meaningful connection on a foliation $\Sigma = \{\Sigma_t\}$ reduces to defining how it acts on vector fields tangent to the leaves, and extending that action to arbitrary tensors via the standard axioms of an affine connection.

Put differently, given a connection ∇ on a manifold $\mathcal{M}$ endowed with a foliation $\Sigma = \{\Sigma_t\}$, to define an induced connection $\bar{\nabla}$ on each leaf Σ_t we must provide a prescription for

$$\bar{\nabla}_X Y \quad \text{for any} \quad X, Y \in \Gamma(T\Sigma) \tag{6.3.3}$$

which relates back to ∇ in a geometrically meaningful way.

Before diving into a definition, let us first build some geometric intuition. Recall that the expression $\nabla_X Y$ describes how a vector field Y changes along the flow of X. Further, the pushforward ι_* of the inclusion map ι embeds $T\Sigma$ into the ambient tangent bundle $T\mathcal{M}$, allowing us to view $X, Y \in \Gamma(T\Sigma)$ as vector fields living in $T\mathcal{M}$ via their pushforwards $\iota_* X$ and $\iota_* Y$.

Since we want the connection $\bar{\nabla}$ on Σ to "mimic" the ambient connection ∇ on $\mathcal{M}$ as closely as possible, a natural first attempt to define $\bar{\nabla}$ would be

$$\bar{\nabla}_X Y = \nabla_{\iota_* X}(\iota_* Y). \tag{6.3.4}$$

In words, this says that $\bar{\nabla}$ measures change of Y along X in precisely the same way that the ambient connection ∇ measures change of the embedded version of Y along that of X within $T\mathcal{M}$.

This is a good starting point—essentially all of the geometric structure of ∇ is being transferred to $\bar{\nabla}$. However, this definition has a fundamental flaw. It is subtle but crucial: an induced connection on the foliation Σ must restrict to a connection on each individual leaf Σ_t, which is a map

$$\bar{\nabla} : \Gamma\big(T^{(r,s)}\Sigma_t\big) \to \Gamma\big(T^{(r,s+1)}\Sigma_t\big), \quad T \mapsto \bar{\nabla} T, \tag{6.3.5}$$

and, when acting on (and along) vector fields, is given more concretely by the map

$$\bar{\nabla} : \Gamma(T\Sigma_t) \times \Gamma(T\Sigma_t) \to \Gamma(T\Sigma_t), \quad (X, Y) \mapsto \bar{\nabla}_X Y. \tag{6.3.6}$$

In short, the connection must send vector fields tangent to the leaves to other vector fields tangent to the leaves. But $\nabla_{\iota_* X}(\iota_* Y)$, although well-defined in $\Gamma(T\mathcal{M})$, need not necessarily lie in the subbundle $T\Sigma$. The ambient connection ∇ is under no obligation to preserve tangency to the leaves —it can easily produce components orthogonal to them comparing vectors at different points of the submanifolds.

We are therefore forced to modify our first attempt so as to eliminate any normal components that may arise. For this purpose, recall the left-inverse $(\iota_*)^{-1} : T\mathcal{M} \to T\Sigma$ of the pushforward we introduced in Section 3.3. While there exist infintely many such left-inverses, we singled out a unique one by requiring that the projection $P = \iota_* \circ (\iota_*)^{-1}$ be orthogonal with respect to the ambient metric. This construction gives us precisely the tool we need: $(\iota_*)^{-1}$ acts as the identity on $T\Sigma$, while annihilating vectors in the normal bundle $N\Sigma$; that is,

$$\ker\big((\iota_*)^{-1}\big) = N\Sigma. \tag{6.3.7}$$

Thus, applying $(\iota_*)^{-1}$ to our first attempt yields a vector field in $\Gamma(T\Sigma)$, with all normal components stripped away. This motivates the corrected definition of the induced connection $\bar{\nabla}$ as

$$\bar{\nabla}_X Y = (\iota_*)^{-1} \nabla_{\iota_* X}(\iota_* Y). \tag{6.3.8}$$

This expression is admittedly cumbersome to read, but its geometric interpretation is clear. To compute $\bar{\nabla}_X Y$ for $X, Y \in \Gamma(T\Sigma)$, we push both vector fields forward (i.e., embed) into the ambient tangent bundle, measure the change of one along the other with the ambient connection, and then project the result back onto $T\Sigma$ using the orthogonal "projection" $(\iota_*)^{-1}$. This removes any component normal to the foliation that may have been introduced by ∇, while remaining true to the tangent contributions.

Let us now formalise this geometric construction as a rigorous definition.

Definition (Induced Connection on a Foliation) Let $\mathcal{M}$ be a smooth manifold, $\Sigma = \{\Sigma_t\}$ a foliation of $\mathcal{M}$, and

$$\nabla : \Gamma\big(T^{(r,s)}\mathcal{M}\big) \to \Gamma\big(T^{(r,s+1)}\mathcal{M}\big) \tag{6.3.9}$$

a connection on $\mathcal{M}$. We call the linear map

$$\bar{\nabla} : \Gamma\big(T^{(r,s)}\Sigma\big) \to \Gamma\big(T^{(r,s+1)}\Sigma\big) \tag{6.3.10}$$

the *induced connection on* Σ (associated to ∇) if

1. it reduces to the exterior derivative on functions $\varphi \in C^\infty(\mathcal{M})$,

$$\bar{\nabla}\varphi = d\varphi, \tag{6.3.11}$$

2. acts on vector fields $X, Y \in \Gamma(T\Sigma)$ as

$$\bar{\nabla}_X Y = (\iota_*)^{-1}\Big(\nabla_{\iota_* X}(\iota_* Y)\Big) \tag{6.3.12}$$

   where $(\iota_*)^{-1} : T\mathcal{M} \to T\Sigma$ is the left-inverse of ι_* such that

$$\ker\Big((\iota_*)^{-1}\Big) = N\Sigma, \tag{6.3.13}$$

   or equivalently, such that $P = \iota_* \circ (\iota_*)^{-1}$ is the orthogonal projection $P : T\mathcal{M} \to T\Sigma$;

3. satisfies a tensorial Leibniz rule,

$$\bar{\nabla}(T \otimes S) = \big(\bar{\nabla}T\big) \otimes S + T \otimes \big(\bar{\nabla}S\big), \tag{6.3.14}$$

   on arbitrary tensors T, S;

4. and is compatible with contractions,

$$\bar{\nabla}\Big(\mathrm{tr}^i_j\, T\Big) = \mathrm{tr}^i_j\big(\bar{\nabla}T\big), \tag{6.3.15}$$

   for any tensor T.

Remarks:

- The axioms 1., 3. and 4. guarantee that $\bar{\nabla}$ is itself an affine connection.
- Since the connection coefficients $\bar{\Gamma}^k{}_{ij}$ in coordinates $\big(t^A, y^i\big)$ adapted to Σ are defined by the action of the connection on basis vectors, axiom 2. allows for their explicit computation in terms of the coefficients $\Gamma^\lambda{}_{\mu\nu}$ of the ambient connection ∇, the pushforward matrix $E^\mu_i = \frac{\partial x^\mu}{\partial y^i}$ and its orthogonal left-inverse E^i_μ. Concretely, we may derive

$$\begin{aligned}
\bar{\Gamma}^k{}_{ij}\partial_k = \bar{\nabla}_i \partial_j &= (\iota_*)^{-1}\Big(\nabla_{\iota_*\partial_i}\big(\iota_*\partial_j\big)\Big) \\
&= (\iota_*)^{-1}\Big(\nabla_{E^\mu_i \partial_\mu}\big(E^\nu_j \partial_\nu\big)\Big) \\
&= E^\mu_i (\iota_*)^{-1}\big(\nabla_\mu\big(E^\nu_j \partial_\nu\big)\big) \\
&= E^\mu_i (\iota_*)^{-1}\Big(\big(\partial_\mu E^\nu_j\big)\partial_\nu + E^\nu_j \underbrace{\nabla_\mu \partial_\nu}_{=\Gamma^\lambda{}_{\mu\nu}\partial_\lambda}\Big) \\
&= E^\mu_i \Big(\big(\partial_\mu E^\lambda_j\big) + E^\nu_j \Gamma^\lambda{}_{\mu\nu}\Big)\underbrace{(\iota_*)^{-1}(\partial_\lambda)}_{=E^k_\lambda \partial_k} \\
&= \Big(E^k_\lambda E^\mu_i E^\nu_j \Gamma^\lambda{}_{\mu\nu} + E^k_\lambda \partial_i E^\lambda_j\Big)\partial_k.
\end{aligned} \tag{6.3.16}$$

  From this we conclude that the induced connection coefficients are given by

$$\bar{\Gamma}^k{}_{ij} = E^k_\lambda E^\mu_i E^\nu_j \Gamma^\lambda{}_{\mu\nu} + E^k_\lambda \partial_i E^\lambda_j. \tag{6.3.17}$$

In particular, this matches the structure of the coordinate transformation behaviour (6.1.24) of the connection.

- Since $\bar{\nabla}$ is an affine connection connection, the above relationship

$$\bar{\nabla}_i \partial_j = \bar{\Gamma}^k{}_{ij} \partial_k \tag{6.3.18}$$

on vectors extends to 1-forms by

$$\bar{\nabla}_i \, dy^j = -\bar{\Gamma}^j{}_{ik} \, dy^k \tag{6.3.19}$$

by the general result (6.1.34). Further, the linearity and Leibniz rule that hold for $\bar{\nabla}$ allow for the extension onto any tensor T.

6.4 Proof: Levi-Civita Property is Inherited

The previous section introduced the notion of a connection on the submanifolds of a foliation, induced from a connection on the ambient manifold. This construction defines a specific connection on each submanifold, but it is not immediately clear which properties of the ambient connection are preserved, or whether the induced connection might acquire new properties absent from the ambient. Such questions typically require case-by-case analysis.

There is, however, an important special case, which we explore in this section. Recall that the Levi-Civita is uniquely characterised by two conditions: vanishing torsion and compatibility with the metric. We will show that if the ambient connection satisfies these conditions, then the induced connection does as well—in this sense, they are inherited. It follows that the connection induced by a Levi-Civita connection is itself Levi-Civita. In particular, this yields an alternative to the projection of connection coefficients (6.3.17), as the coefficients of the induced connection can now be computed directly from linear combinations of partial derivatives of the induced metric.

In the following, let $\mathcal{M}$ denote a (pseudo-)Riemannian manifold with metric tensor $g \in \Gamma\big(T^{(0,2)}\mathcal{M}\big)$, $\Sigma = \{\Sigma_t\}$ a foliation of $\mathcal{M}$, and $\gamma = \iota^* g$ the induced metric on Σ_t. Further, let ∇ be a connection on $\mathcal{M}$ and $\bar{\nabla}$ the connection on Σ induced by ∇.

Vanishing Torsion Suppose ∇ has vanishing torsion. Concretely, this means that

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] = 0 \quad \forall X, Y \in \Gamma(T\mathcal{M}), \tag{6.4.1}$$

or equivalently,

$$\nabla_X Y - \nabla_Y X = [X,Y] \tag{6.4.2}$$

The torsion of the induced connection, evaluated on vector fields $X, Y \in \Gamma(T\Sigma)$ is given by

$$\begin{aligned} \bar{T}(X,Y) &= \bar{\nabla}_X Y - \bar{\nabla}_Y X - [X,Y] \\ &= (\iota_*)^{-1}\Big[\nabla_{\iota_* X}(\iota_* Y) - \nabla_{\iota_* Y}(\iota_* X)\Big] - [X,Y] \\ &= \underbrace{(\iota_*)^{-1}[\iota_* X, \iota_* Y]}_{=[X,Y] \quad (*)} - [X,Y] = 0. \end{aligned} \tag{6.4.3}$$

Thus, as a consequence of the vanishing torsion of ∇, the induced connection $\bar{\nabla}$ is torsion-free as well. The identity $(*)$ originates from

$$[\iota_* X, \iota_* Y] = \iota_*[X,Y], \tag{6.4.4}$$

which can be shown coordinate-independently using the definition of the commutator,

$$[X,Y][f] = X[Y[f]] - Y[X[f]], \quad f \in C^\infty(\mathcal{M}), \tag{6.4.5}$$

and the fact that $(\iota_* X)[f] = X[f \circ \iota]$.

Alternatively, one can show this on the level of components as well. The vanishing torsion of the ambient connection implies that its connection coefficients $\Gamma^{\lambda}{}_{\mu\nu}$ are symmetric in the lower indices. The induced connection coefficients, given by

$$\bar{\Gamma}^{k}{}_{ij} = E^k_\lambda E^\mu_i E^\nu_j \Gamma^{\lambda}{}_{\mu\nu} + E^k_\lambda \partial_i E^\lambda_j \tag{6.4.6}$$

(cf. eq. (6.1.24)), are then symmetric in the lower indices i, j as well—the first term is symmetric due to the symmetry of $\Gamma^{\lambda}{}_{\mu\nu}$, and the second because of

$$\partial_i E^\lambda_j = \frac{\partial^2 x^\lambda}{\partial y^i \partial y^j}, \tag{6.4.7}$$

which is symmetric as well.

We have now given a coordinate-free and a component-based proof of the fact that an induced connection inherits torsion-freeness from the ambient connection. It remains to show that metric compatibility is preserved as well to allow us to conclude that an ambient Levi-Civita connection induces a Levi-Civita connection on the submanifolds of a foliation.

Metric Compatibility To show that metric compatibility of ∇ is inherited to $\bar{\nabla}$, let us first derive a useful identity—a product rule for the inner product induced by the metric g. Observe that for $X, Y, Z \in \Gamma(T\mathcal{M})$, we have

$$\begin{aligned} &\nabla_X g(Y, Z) = \nabla_X \operatorname{tr}^{3,4}_{1,2}(g \otimes Y \otimes Z) \\ &= \operatorname{tr}^{3,4}_{1,2}[(\nabla_X g) \otimes Y \otimes Z + g \otimes (\nabla_X Y) \otimes Z + g \otimes Y \otimes (\nabla_X Z)] \\ &= (\nabla_X g)(Y, Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \end{aligned} \tag{6.4.8}$$

which is equivalent to

$$\nabla_X g(Y, Z) - (\nabla_X g)(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z). \tag{6.4.9}$$

In particular, a connection ∇ is compatible with a metric g if and only if the first term vanishes, i.e. if

$$\nabla_X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \quad \forall X, Y, Z \in \Gamma(T\mathcal{M}). \tag{6.4.10}$$

We will now assume that this holds for ∇ and g, and show that as a consequence, it is also true for the induced connection $\bar{\nabla}$ and metric γ. This proceeds as follows, for $X, Y, Z \in \Gamma(T\Sigma)$:

$$\begin{aligned} \bar{\nabla}_X \gamma(Y, Z) - \big(\bar{\nabla}_X \gamma\big)(Y, Z) &\overset{1.}{=} \gamma\big(\bar{\nabla}_X Y, Z\big) + (Y \leftrightarrow Z) \\ &\overset{2.}{=} (\iota^* g)\Big((\iota_*)^{-1} \nabla_{\iota_* X}(\iota_* Y), Z\Big) + (Y \leftrightarrow Z) \\ &\overset{3.}{=} g\big(\underbrace{\iota_* \circ (\iota_*)^{-1}}_{=P} \nabla_{\iota_* X}(\iota_* Y), \iota_* Z\big) + (Y \leftrightarrow Z) \\ &\overset{4.}{=} g\Big(\nabla_{\iota_* X}(\iota_* Y), \underbrace{P \iota_* Z}_{=\iota_* Z}\Big) + (Y \leftrightarrow Z) \\ &\overset{5.}{=} \nabla_{\iota_* X} g(\iota_* Y, \iota_* Z) \\ &\overset{6.}{=} \bar{\nabla}_X \gamma(Y, Z) \end{aligned} \tag{6.4.11}$$

which shows that

$$\big(\bar{\nabla}_X \gamma\big)(Y, Z) = 0, \tag{6.4.12}$$

i.e. that $\bar{\nabla}\gamma = 0$ and hence the induced connection is compatible with the induced metric. Since the algebra is quite dense here, let us outline the individual steps in the derivation (6.4.11):

1. Make use of identity (6.4.9);
2. Insert definitions, $\gamma = \iota^* g$ and $\bar{\nabla}_X Y = (\iota_*)^{-1}\nabla_{\iota_* X}(\iota_* Y)$;
3. Employ $(\iota^* g)(X,Y) = g(\iota_* X, \iota_* Y)$
4. Apply orthogonality of P, i.e. $g(PX,Y) = g(X,PY)$;
5. Use identity (6.4.10)—here, the assumption that ∇ is metric-compatible enters;
6. Employ $(\iota^* g)(X,Y) = g(\iota_* X, \iota_* Y)$ as well as $\nabla_{\iota_* X}\varphi = (\iota_* X)[\varphi] = X[\varphi] = \bar{\nabla}_X \varphi$ on functions $\varphi \in C^\infty(\mathcal{M})$.

Though somewhat algebra-heavy and light on geometric intuition, the above derivations yield a powerful result: The induced connection $\bar{\nabla}$ on the foliation Σ—induced by the Levi-Civita connection ∇ on $\mathcal{M}$—is of Levi-Civita type as well. In particular, it follows that the induced connection components are given by the Christoffel symbol expression for γ, i.e.,

$$\bar{\Gamma}^k{}_{ij} = \frac{1}{2}\gamma^{k\ell}\big(\gamma_{\ell i,j} + \gamma_{\ell j,i} - \gamma_{ij,\ell}\big). \tag{6.4.13}$$

This is a direct construction of the connection coefficients that makes no reference to the ambient connection.

6.5 Example: Induced Connection on the Foliation of $\mathbb{R}^3 \setminus \{0\}$ into Spheres

To see the machinery of induced connections as well as the inheritance of the Levi-Civita propery in action, in this section, we reconsider the example of $\mathcal{M} = \mathbb{R}^3 \setminus \{0\}$, foliated into concentric origin-centered spheres from Section 5.5. Let us briefly reestablish the setting.

On $\mathbb{R}^3 \setminus \{0\}$, in Cartesian coordinates $x^\mu = (x,y,z)$, the Euclidean metric reads

$$g = g_{\mu\nu}\, dx^\mu \otimes dx^\nu = dx \otimes dx + dy \otimes dy + dz \otimes dz. \tag{6.5.1}$$

We foliate $\mathbb{R}^3 \setminus \{0\}$ by introducing the function

$$r : \mathcal{M} \to \mathbb{R}, \quad r(p) = \sqrt{x^2 + y^2 + z^2}, \quad p = (x,y,z) \tag{6.5.2}$$

and define the leaves of the foliation $\Sigma = \{\Sigma_r\}$ to be its level sets,

$$\Sigma_{r_0} := \{p \in \mathcal{M} \mid r(p) = r_0\} = r_0 S^2. \tag{6.5.3}$$

In spherical coordinates $(r, y^i) = (r, \theta, \varphi)$ on $\mathbb{R}^3 \setminus \{0\}$, we already derived the induced metric on a leaf Σ_r, with the result

$$\gamma = \gamma_{ij}\, dq^i \otimes dq^j = r^2\, d\theta \otimes d\theta + r^2 \sin^2\theta\, d\varphi \otimes d\varphi, \tag{6.5.4}$$

with nonzero components $\gamma_{\theta\theta} = r^2$, $\gamma_{\varphi\varphi} = r^2 \sin^2\theta$.

We equip the ambient manifold $\mathbb{R}^3 \setminus \{0\}$ with the Levi-Civita connection, which in Cartesian coordinates has vanishing coefficients, i.e.

$$\Gamma^\lambda{}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}\big(g_{\rho\mu,\nu} + g_{\rho\nu,\mu} - g_{\mu\nu,\rho}\big) = 0, \tag{6.5.5}$$

since the metric components $g_{\mu\nu} = \delta_{\mu\nu}$ are constant.

We have two formulae for the components $\bar{\Gamma}^k{}_{ij}$ of the connection $\bar{\nabla}$ induced by ∇ at our disposal; we can either use the projection formula (6.3.17), or calculate them directly using the Christoffel expression (6.4.13). Let us evaluate both and compare.

The evaluation of the projection formula (6.3.17) requires us to compute the components of the pushforward matrix

$$E^\mu_i = \frac{\partial x^\mu}{\partial q^i} \tag{6.5.6}$$

as well as its left-inverse E^i_μ subject to the condition that $P^\mu{}_\nu = E^\mu_i E^i_\nu$ is orthogonal. The components of the pushforward matrix read

$$\begin{aligned} E^x_\theta &= r\cos\theta\cos\varphi, & E^x_\varphi &= -r\sin\theta\sin\varphi, \\ E^y_\theta &= r\cos\theta\sin\varphi, & E^y_\varphi &= r\sin\theta\cos\varphi, \\ E^z_\theta &= -r\sin\theta, & E^z_\varphi &= 0. \end{aligned} \tag{6.5.7}$$

The orthogonal left-inverse can be computed using

$$E^i_\mu = \gamma^{ij} g_{\mu\nu} E^\nu_j, \tag{6.5.8}$$

as then

$$E^i_\mu E^\mu_k = \gamma^{ij} \underbrace{g_{\mu\nu} E^\nu_j E^\mu_k}_{=\gamma_{jk}} = \gamma^{ij}\gamma_{jk} = \delta^i_k, \tag{6.5.9}$$

ensures the left-inverse property and

$$P_{\mu\nu} = g_{\mu\lambda} P^\lambda{}_\nu = g_{\mu\lambda} E^\lambda_i E^i_\nu = g_{\mu\lambda} E^\lambda_i \gamma^{ij} g_{\nu\rho} E^\rho_j \tag{6.5.10}$$

the symmetry of $P_{\mu\nu}$ (which is equivalent to orthogonality).

The components of the orthogonal left-inverse read

$$\begin{aligned} E^\theta_x &= \frac{1}{r}\cos\theta\cos\varphi, & E^\varphi_x &= -\frac{1}{r\sin\theta}\sin\varphi, \\ E^\theta_y &= \frac{1}{r}\cos\theta\sin\varphi, & E^y_\varphi &= \frac{1}{r\sin\theta}\cos\varphi, \\ E^\theta_z &= -\frac{1}{r}\sin\theta, & E^z_\varphi &= 0. \end{aligned} \tag{6.5.11}$$

Since the ambient connection coefficients vanish in Cartesian coordinates, the induced components reduce to the inhomogeneous term in (6.3.17), that is,

$$\bar{\Gamma}^k{}_{ij} = E^k_\lambda \partial_i E^\lambda_j. \tag{6.5.12}$$

While entirely mechanical, the computation of all components is somewhat laborious. We'll carry out one as an illustration and state the remaining results without derivation. As our example, we choose to calculate

$$\begin{aligned} \bar{\Gamma}^\theta{}_{\varphi\varphi} &= E^\theta_\lambda \partial_\varphi E^\lambda_\varphi = E^\theta_x \partial_\varphi E^x_\varphi + E^\theta_y \partial_\varphi E^y_\varphi \\ &= \left(\frac{1}{r}\cos\theta\cos\varphi\right)\partial_\varphi(-r\sin\theta\sin\varphi) + \left(\frac{1}{r}\cos\theta\sin\varphi\right)\partial_\varphi(r\sin\theta\cos\varphi) \\ &= -\cos\theta\sin\theta\cos^2\varphi - \cos\theta\sin\theta\sin^2\varphi = -\cos\theta\sin\theta \\ &= -\frac{1}{2}\sin(2\theta) \end{aligned} \tag{6.5.13}$$

The remaining non-vanishing connection coefficients are given by

$$\bar{\Gamma}^{\varphi}{}_{\theta\varphi} = \bar{\Gamma}^{\varphi}{}_{\varphi\theta} = \cot\theta, \tag{6.5.14}$$

where $\cot\theta = \frac{\cos\theta}{\sin\theta}$.

We now proceed to recompute these coefficients using the standard Christoffel formula. To do so, it is usually advantageous to first compute the components

$$\bar{\Gamma}_{kij} = \frac{1}{2}\big(g_{ki,j} + g_{kj,i} - g_{ij,k}\big) \tag{6.5.15}$$

and to then raise the first index with the inverse metric. Before wildly starting to evaluate all possible index combinations, we should first examine the nature of the components γ_{ij} of the induced metric. There is only one component, $\gamma_{\varphi\varphi} = r^2\sin^2\theta$, which is non-constant[7] and hence has a chance of contributing to $\bar{\Gamma}_{kij}$. It only depends on θ—thus, the only non-zero component of the partial gradient of γ_{ij} is

$$\partial_\theta \gamma_{\varphi\varphi} = 2r^2\sin\theta\cos\theta. \tag{6.5.16}$$

Consequently, a coefficient $\bar{\Gamma}_{kij}$ can only be non-zero if one of the indices is θ and the other two are φ. Due to symmetry in the last two indices, this leaves us with two options,

$$\begin{aligned} \bar{\Gamma}_{\theta\varphi\varphi} &= \frac{1}{2}\Big(\underbrace{\gamma_{\theta\varphi,\varphi}}_{=0} + \underbrace{\gamma_{\theta\varphi,\varphi}}_{=0} - \gamma_{\varphi\varphi,\theta}\Big) = -r^2\sin\theta\cos\theta, \\ \bar{\Gamma}_{\varphi\theta\varphi} = \bar{\Gamma}_{\varphi\varphi\theta} &= \frac{1}{2}\Big(\gamma_{\varphi\varphi,\theta} + \underbrace{\gamma_{\varphi\theta,\varphi}}_{=0} - \underbrace{\gamma_{\varphi\theta,\varphi}}_{=0}\Big) = r^2\sin\theta\cos\theta. \end{aligned} \tag{6.5.17}$$

Given the components of the inverse metric,

$$\gamma^{\theta\theta} = \frac{1}{r^2}, \quad \gamma^{\varphi\varphi} = \frac{1}{r^2\sin^2\theta}, \tag{6.5.18}$$

we can compute the connection coefficients as

$$\begin{aligned} \bar{\Gamma}^{\theta}{}_{\varphi\varphi} &= g^{\theta\theta}\bar{\Gamma}_{\theta\varphi\varphi} = -\sin\theta\cos\theta = -\frac{1}{2}\sin(2\theta), \\ \bar{\Gamma}^{\varphi}{}_{\theta\varphi} = \bar{\Gamma}^{\varphi}{}_{\varphi\theta} &= g^{\varphi\varphi}\bar{\Gamma}_{\varphi\varphi\theta} = \frac{\cos\theta}{\sin\theta} = \cot\theta. \end{aligned} \tag{6.5.19}$$

This reproduces exactly the same result as the projection formula, as we would expect from our derivations made for general manifolds. Nevertheless, it is satisfying to see that the abstract machinery does indeed work when applied to concrete examples.

7 Curvature

In this section, we examine the different kinds of curvature that arise in the study of submanifolds and foliations. We begin by reviewing the definition of the Riemann curvature tensor and its contractions, which will allow us to define both the *ambient curvature* of a manifold $\mathcal{M}$ and the *instrinsic curvature* of the leaves of a foliation Σ on it. The ambient curvature is defined via the ambient connection ∇ on $\mathcal{M}$, whereas the intrinsic curvature is derived from the induced connection ∇, which—as established in the previous section—is the tangential projection of ∇ onto Σ.

This immediately suggests a relationship between the ambient and intrinsic curvatures. One might naïvely expect that—analogous to the connections—the intrinsic curvature is simply the

[7]Recall that on a leaf, r is constant.

projection of the ambient one. A simple counterexample will demonstrate that this cannot be entire picture. This leads us to the notion of *extrinsic curvature*, which—roughly speaking—captures the normal component of the ambient connection ∇ that is discarded when passing to $\bar{\nabla}$.

This will prepare us for the derivation of the Gauss-Codazzi equation in the next section, which relates the intrinsic, extrinsic and projected ambient curvatures in a precise and elegant way.

7.1 Intrinsic Curvature

7.1.1 Review: Parallel Transport

The intuition behind the Riemann curvature tensor is closely tied to *parallel transport.* The idea behind parallel transport is to take a vector and some point p on a manifold and moving it along a curve γ to another point q, keeping it "parallel" to its initial direction (and length) at p. In flat space $\mathbb{R}^n$ with the Levi-Civita connection, the idea of "being parallel" is very clear: the covariant derivative of the coordinate basis vectors vanishes in cartesian coordinates, $\nabla_\mu \partial_\nu = 0$, so a vector remains parallel if its components stay constant along γ.

On a general manifold with an arbitrary connection ∇, this generalises. The connection is simply a measure of how a vector field changes across a manifold—it can thus be used to measure the change of a vector along a curve. The generalised notion of "being parallel" is now replaced by "keeping its direction according to the connection", or, more concisely, being constant with respect to the connection. Let us now make this concrete.

Definition (Pullback Bundle and Connection) Let $\gamma : (a, b) \to \mathcal{M}$ be a smooth curve on a smooth manifold $\mathcal{M}$ equipped with an affine connection ∇. The *pullback bundle* of $T\mathcal{M}$ along γ is the vector bundle

$$\gamma^* T\mathcal{M} = T\mathcal{M}|_\gamma, \tag{7.1.1}$$

or equivalently,

$$\gamma^* T\mathcal{M} = \bigsqcup_{s \in (a,b)} T_{\gamma(s)}\mathcal{M}. \tag{7.1.2}$$

This is a vector bundle with fibres of dimension $\dim \mathcal{M}$ over the curve $\gamma \subset \mathcal{M}$ viewed as a one-dimensional submanifold of $\mathcal{M}$.

The connection ∇ restricts to smooth sections $X \in \Gamma(\gamma^* T\mathcal{M})$ via

$$\nabla_{\dot{\gamma}} X = \nabla_{\tilde{Y}} \tilde{X}, \tag{7.1.3}$$

where $\tilde{X}, \tilde{Y} \in \Gamma(T\mathcal{M})$ extend X and the tangent vector field $\dot{\gamma}$, respectively.

In local coordinates $x^\mu(s)$ on γ, this becomes

$$\nabla_{\dot{\gamma}} X = \frac{dX^\lambda}{ds}\partial_\lambda + X^\mu \nabla_{\dot{\gamma}} \partial_\mu, \tag{7.1.4}$$

with

$$\dot{\gamma} = \frac{d}{ds} = \dot{x}^\mu \partial_\mu, \quad \nabla_{\dot{\gamma}} \partial_\mu = \dot{x}^\nu \nabla_\nu \partial_\mu = \dot{x}^\nu \Gamma^\lambda{}_{\nu\mu} \partial_\lambda \tag{7.1.5}$$

Thus,

$$\nabla_{\dot{\gamma}} X = \left(\frac{dX^\lambda}{ds} + \dot{x}^\nu \Gamma^\lambda{}_{\nu\mu} X^\mu \right) \partial_\lambda, \tag{7.1.6}$$

which describes how a vector field $X \in \Gamma(\gamma^* T\mathcal{M})$ changes along the curve, according to the connection.

Now that we know how to compute how vectors change along curves, we are able to demand this change to be zero—that is, we are ready to define parallel transport.

Definition (Parallel Transport) Let $\mathcal{M}$ be a smooth manifold equipped with a connection ∇, and $\gamma : (a, b) \to \mathcal{M}$ a smooth curve connecting two points $p, q \in \mathcal{M}$, that is, with $\gamma(a) = p, \gamma(b) = q$. Two vectors $X_p \in T_p\mathcal{M}$ and $X_q \in T_q\mathcal{M}$ are said to be *parallel* if there exists a vector field $X \in \Gamma(\gamma^* T\mathcal{M})$ such that

1. $X(p) = X_p$, $X(q) = X_q$.
2. $\nabla_{\dot{\gamma}} X = 0$

We further call X_q the *parallel transport of X_p along γ.*

To unpack this definition: the requirement that X agrees with X_p at p and with X_q at q ensures that X is a smooth interpolation between them—a single smooth field along the curve. The second condition, $\nabla_{\dot{\gamma}} X = 0$, then demands that this field does not "twist" or "deform" relative to the connection as we move along γ. In other words, at each infinitesimal step the vector keeps its direction according to the connection, so that by the time we reach q, the resulting X_q is precisely the parallel transport of X_p.

Notice that parallel transport depends not only on the connection but also on the chosen curve γ. Transporting X_p from p to q along two different paths may yield different vectors at q. Intuitively, this happens because the connection coefficients $\Gamma^{\lambda}{}_{\mu\nu}$ vary across the manifold, so distinct curves encounter different "configurations" of these coefficients. As a result, the differential equation $\nabla_{\dot{\gamma}} X = 0$ may have different solutions depending on the path. If, however, parallel transport from p to q is independent of the path for all choices of $p \in \mathcal{M}$ and $X_p \in T_p\mathcal{M}$, then $(\mathcal{M}, \nabla)$ is called *flat*—excatly as in the case of $\mathbb{R}^n$ with its Levi-Civita connection.

Let us examine the second condition more precisely and characterise the nature of the differential equation and its solutions. In terms of components, it may be expanded as

$$\frac{d}{ds} X^{\lambda} = -\Gamma^{\lambda}{}_{\nu\mu} \dot{x}^{\mu} X^{\nu}. \tag{7.1.7}$$

This is a linear first order differential equation for the vector components X^{λ} with a non-constant coefficient matrix. We may introduce the linear operator $\Gamma_{\dot{\gamma}}(s) : T_{\gamma(s)}\mathcal{M} \to T_{\gamma(s)}\mathcal{M}$, which maps

$$X = X^{\mu} \partial_{\mu} \mapsto \Gamma_{\dot{\gamma}} X = \dot{x}^{\lambda} \Gamma^{\mu}{}_{\lambda\nu} X^{\nu} \partial_{\mu}. \tag{7.1.8}$$

Note that $\Gamma_{\dot{\gamma}}$ is not a tensorial operator—we only introduce it to make notation more compact. In terms of $\Gamma_{\dot{\gamma}}$, equation (7.1.7) can be written as

$$\frac{dX}{ds}(s) = -\Gamma_{\dot{\gamma}}(s) X(s), \tag{7.1.9}$$

where we define

$$\frac{dX}{ds}(s) = \left(\frac{d}{ds} X^{\mu}(s) \right) \partial_{\mu} \tag{7.1.10}$$

to be the component-wise derivative. From the general theory of first order linear differential equations, we know that solutions to equation (7.1.9) above take the form

$$X(s) = U(\gamma, s, t) X(t) \tag{7.1.11}$$

for a propagator $U(s,t)$. Intuitively, the propagator $U(\gamma, s, t)$ takes a vector in $T_{\gamma(t)}\mathcal{M}$ and produces its parallel transport in $T_{\gamma(s)}\mathcal{M}$ via a linear map. It has the following properties:

- It reduces to the identity for equal arguments,
$$U(\gamma, s, s) = \mathrm{id}_{T_{\gamma(s)}\mathcal{M}}, \tag{7.1.12}$$
- Has a group composition law,
$$U(\gamma, s, r)U(\gamma, r, t) = U(\gamma, s, t), \tag{7.1.13}$$
- Is invertible,
$$U(\gamma, t, s)^{-1} = U(\gamma, s, t), \tag{7.1.14}$$
- Satisfies the differential equation in both arguments, that is,
$$\frac{d}{dt}U(\gamma, s, t) = \Gamma_{\dot{\gamma}}(t)U(\gamma, s, t), \quad \frac{d}{ds}U(\gamma, s, t) = -\Gamma_{\dot{\gamma}}(t)U(\gamma, s, t). \tag{7.1.15}$$

Though in general, no expression for $U(\gamma, s, t)$ can be given in closed form, it is possible to derive explicit expressions for *infinitesimal* parallel transport. To do so, we take the following perspective: Suppose we have a path $\gamma : (-\delta, \delta) \to \mathcal{M}$ and a vector field $Z \in \Gamma(\gamma^* T\mathcal{M})$. We are interested in computing the parallel transport $\tilde{Z} \in T_{\gamma(0)}\mathcal{M}$ of the vector $Z(s) := Z(\gamma(s))$ along γ, where s is some infinitesimal parameter distance away from 0. That is, we take a vector at a point infinitesimally away from $\gamma(0)$ and "pull" it onto $\gamma(0)$ while keeping it parallel.

The propagator U gives us an explicit expression for $\tilde{Z}$, namely
$$\tilde{Z} = U(\gamma, 0, s)Z(s). \tag{7.1.16}$$
At this point, we can simply Taylor-expand this to first order in s, making use of
$$U(\gamma, 0, 0) = \mathrm{id}_{T_{\gamma(0)}\mathcal{M}}, \quad \frac{d}{ds}U(\gamma, 0, s)|_{s=0} = \Gamma_{\dot{\gamma}}(0) \tag{7.1.17}$$
leading to
$$\begin{aligned}\tilde{Z} &= \underbrace{U(\gamma, 0, 0)}_{=\mathrm{id}}Z(0) + s\frac{d}{ds}[U(\gamma, 0, s)Z(s)]_{s=0} + \mathcal{O}(s^2) \\ &= Z(0) + s\left[\underbrace{U(\gamma, 0, 0)}_{=\mathrm{id}}\frac{dZ}{ds}(0) + \underbrace{\left(\frac{d}{ds}U(\gamma, 0, s)|_{s=0}\right)Z(0)}_{=\Gamma_{\dot{\gamma}}(0)U(\gamma,0,0)}\right] + \mathcal{O}(s^2) \\ &= Z(0) + s\left[\frac{dZ}{ds}(0) + \Gamma_{\dot{\gamma}}Z(0)\right] + \mathcal{O}(s^2) \\ &= Z(0) + s\nabla_{\dot{\gamma}}Z(0) + \mathcal{O}(s^2).\end{aligned} \tag{7.1.18}$$
Hence, the first order correction of parallel transport is the covariant derivative $\nabla_{\dot{\gamma}}Z(0)$.

Now that we know how to parallel transport a single vector infinitesimally along some curve, we can extend this to vectors fields in $T\mathcal{M}$. After all, a vector field is nothing but a collection of individual vectors, which we can each parallel transport using the constructions above. All we need is a family of curves such that each point $p \in \mathcal{M}$ lies on exactly one curve; such a family is called a *congruence*. Congruences arise naturally as integral curves of a non-vanishing vector field X, obtained by solving $\dot{\gamma}(s) = X(\gamma(s))$.

In particular, for a vector field $Z \in \Gamma(T\mathcal{M})$, the parallel transport $\tilde{Z}$ obtained by transporting each of its constituent vectors $Z_p \in T_p\mathcal{M}$ by some infinitesimal parameter distance s along the flow of another vector field $X \in \Gamma(T\mathcal{M})$ is given by

$$\tilde{Z} = Z + s\nabla_X Z + \mathcal{O}(s^2). \tag{7.1.19}$$

All that has changed compared to equation (7.1.18) is that the curve's tangent field $\dot{\gamma}$ is replaced by X—the tangent field of the congruence—and the point-wise expression has turned into a global one. The above expression will allow us to geometrically motivate *and* derive the Riemann curvature tensor in the next section.

7.1.2 Curvature of Manifolds: the Riemann Tensor

As we have already discussed in the previous section, parallel transport of a vector $Z_p \in T_p\mathcal{M}$ along a curve γ depends both on the connection ∇ and the curve itself. If $\gamma_1, \gamma_2 : [a, b] \to \mathcal{M}$ satisfy $\gamma_1(a) = \gamma_2(a) = p$ and $\gamma_1(b) = \gamma_2(b) = q$, then parallel transport of Z_p along thes two curves may give distinct vectors $\tilde{Z}_1, \tilde{Z}_2 \in T_q\mathcal{M}$. Their difference,

$$\Delta = \tilde{Z}_1 - \tilde{Z}_2, \tag{7.1.20}$$

encodes the so-called *holonomy* of the connection. In particular, Δ vanishes for all such choices if and only if the connection is flat—curvature is precisely the local obstruction to path-independent parallel transport.

This construction, which relies on transporting a single vector along two curves, is somewhat cumbersome. A more natural and most importantly, global approach is to extend the discussion to the parallel transport of an entire vector field $Z \in \Gamma(T\mathcal{M})$. Since we now have a vector Z_p at each $p \in \mathcal{M}$, we also require two curves that start at each $p \in \mathcal{M}$ and have the same endpoint. To construct these in a systematic way, we introduce two additional vector fields $X, Y \in \Gamma(T\mathcal{M})$, each generating a congruence of curves. Starting from a point p, we may first follow the flow of Y for an infinitesimal parameter distance s, and then that of X, or vice versa. This yields two distinct curves that form a "parallelogram" if X and Y commute. If they do not, i.e. if $[X, Y] \neq 0$, the parallelogram fails to close: for infinitesimal s, the two curves end at points separated by a displacement along $s[X, Y]$. By adjoining this missing segment to the second curve, we obtain two curves γ_1 and γ_2 with identical start- and endpoints. In particular, the common endpoint can be expressed as

$$q = p + sX + sY = p + sY + sX + s^2[X, Y], \tag{7.1.21}$$

where "$+sX$" denotes moving a parameter distance s along the flow of X. Performing this construction for each $p \in \mathcal{M}$ allows us to parallel transport the entire vector field Z along both choices and compare the results.

For infinitesimal s, recall that the parallel transport of a vector field Z along another vector field X is

$$\tilde{Z} = Z + s\nabla_X Z + \mathcal{O}(s^2). \tag{7.1.22}$$

Transporting first along Y and then X therefore gives

$$\begin{aligned} \tilde{Z}_1 &= \Big(\mathrm{id} + s\nabla_X + \mathcal{O}(s^2)\Big)\Big(\mathrm{id} + s\nabla_Y + \mathcal{O}(s^2)\Big)Z \\ &= Z + s(\nabla_X + \nabla_Y)Z + s^2\nabla_X\nabla_Y Z + \mathcal{O}(s^3). \end{aligned} \tag{7.1.23}$$

For the second path—first following X, then Y, and finally the correction $s[X, Y]$—we apply infinitesimal transport three times:

$$\begin{aligned}\tilde{Z}_2 &= \Big(\mathrm{id} + s\nabla_{s[X,Y]} + \mathcal{O}(s^3)\Big)\Big(\mathrm{id} + s\nabla_Y + \mathcal{O}(s^2)\Big)\Big(\mathrm{id} + s\nabla_X + \mathcal{O}(s^2)\Big)Z \\ &= Z + s(\nabla_X + \nabla_Y)Z + s^2\Big(\nabla_Y\nabla_X + \nabla_{[X,Y]}\Big)Z + \mathcal{O}(s^3).\end{aligned} \tag{7.1.24}$$

The difference between the two—the quantity encoding curvature—is

$$\tilde{Z}_1 - \tilde{Z}_2 = s^2\Big(\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}\Big)Z + \mathcal{O}(s^3). \tag{7.1.25}$$

All $\mathcal{O}(s)$ terms cancel, so the leading contribution arises at order s^2. For infinitesimal s, this term dominates, and thus carrise the curvature information we seek. Moreover, since this construction involves only infinitesimal segments, the resulting description is entirely local. Let us now collect this into a formal definition.

Definition (Riemann Curvature Tensor) Let $\mathcal{M}$ be a smooth manifold equipped with an affine connection ∇, and let $X, Y, Z \in \Gamma(T\mathcal{M})$ be vector fields. The *Riemann curvature tensor* is the map

$$R : \Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \to \Gamma(T\mathcal{M}), \tag{7.1.26}$$

given by

$$R(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z. \tag{7.1.27}$$

We define the components of the Riemann tensor through

$$R\big(\partial_\mu, \partial_\nu\big)\partial_\lambda = R^\rho{}_{\lambda\mu\nu}\partial_\rho, \tag{7.1.28}$$

such that on arbitrary vector fields $X, Y, Z \in \Gamma(T\mathcal{M})$ we have

$$R(X,Y)Z = R^\rho{}_{\lambda\mu\nu}X^\mu Y^\nu Z^\lambda \partial_\rho. \tag{7.1.29}$$

In particular, along coordinate directions ∂_μ, we have

$$R\big(\partial_\mu, \partial_\nu\big)Z = \big[\nabla_\mu, \nabla_\nu\big]Z = R^\rho{}_{\lambda\mu\nu}Z^\lambda\partial_\rho \tag{7.1.30}$$

—the $\nabla_{[\partial_\mu, \partial_\nu]}$ term vanishes since $\big[\partial_\mu, \partial_\nu\big] = 0$.

Notice that for fixed X, Y, the expression $R(X,Y)$ can naturally be viewed as a linear operator on $\Gamma(T\mathcal{M})$. Beyond that, its action can be extended to arbitrary tensor fields, by defining

$$R : \Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \times \Gamma\big(T^{(r,s)}\mathcal{M}\big) \to \Gamma\big(T^{(r,s)}\mathcal{M}\big), \tag{7.1.31}$$

$$R(X,Y)T = \nabla_X\nabla_Y T - \nabla_Y\nabla_X T - \nabla_{[X,Y]}T. \tag{7.1.32}$$

Let us now examine some properties of this definition in closer detail.

Remarks

- **Annihilation of Functions** In particular, on functions $\varphi \in C^\infty(\mathcal{M})$, we have

$$\begin{aligned}R(X,Y)\varphi &= \nabla_X\nabla_Y\varphi - \nabla_Y\nabla_X\varphi - \nabla_{[X,Y]}\varphi \\ &= \underbrace{X[Y[\varphi]] - Y[X[\varphi]]}_{=[X,Y][\varphi]} - [X,Y][\varphi] \\ &= 0.\end{aligned} \tag{7.1.33}$$

That is, $R(X,Y)$ annihilates functions.

- **Derivation Property** The linear operator $R(X,Y)$ on $\Gamma\big(T^{(r,s)}\mathcal{M}\big)$ is a *derivation*: for arbitrary tensor fields T and S, it satisfies the Leibniz rule

$$R(X,Y)(T \otimes S) = (R(X,Y)T) \otimes S + T \otimes (R(X,Y)S). \tag{7.1.34}$$

To see this, note that the curvature operator $R(X,Y)$ can be written in the compact form

$$R(X,Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]}. \tag{7.1.35}$$

The second term $\nabla_{[X,Y]}$ is plainly a derivation, as it is but a covariant derivative. It therefore suffices to verify the Leibniz property for the commutator $[\nabla_X, \nabla_Y]$. Consider $\nabla_X \nabla_Y (T \otimes S)$. Using the product rule for ∇, we compute

$$\begin{aligned}
&\nabla_X \nabla_Y (T \otimes S) = \nabla_X((\nabla_Y T) \otimes S + T \otimes \nabla_Y S) \\
&= (\nabla_X \nabla_Y T) \otimes S + \underbrace{(\nabla_Y T) \otimes (\nabla_X S) + (\nabla_X T) \otimes (\nabla_Y S)}_{\text{symmetric in } X \text{ and } Y} + T \otimes (\nabla_X \nabla_Y S),
\end{aligned} \tag{7.1.36}$$

Note that the middle two terms together are symmetric in X and Y, and hence cancel upon antisymmetrisation. Therefore,

$$[\nabla_X, \nabla_Y](T \otimes S) = ([\nabla_X, \nabla_Y]T) \otimes S + T \otimes ([\nabla_X, \nabla_Y]S). \tag{7.1.37}$$

Combining this with the Leibniz property of $\nabla_{[X,Y]}$, we conclude that $R(X,Y)$ is indeed a derivation.

- **Action on 1-Forms** The derivation property of $R(X,Y)$ allows the explicit computation of its action on 1-forms, and thus the extension to arbitrary tensors. Given a 1-form $\omega \in \Gamma(T^*\mathcal{M})$ and a vector field $Z \in \Gamma(T\mathcal{M})$, we have

$$\begin{aligned}
0 &= R(X,Y)\omega(Z) = R(X,Y)\big(\mathrm{tr}_1^1\, \omega \otimes Z\big) = \mathrm{tr}_1^1\, R(X,Y)(\omega \otimes Z) \\
&= \mathrm{tr}_1^1[(R(X,Y)\omega) \otimes Z + \omega \otimes (R(X,Y)Z)] \\
&= (R(X,Y)\omega)(Z) + \omega(R(X,Y)Z).
\end{aligned} \tag{7.1.38}$$

This demonstrates that the 1-form $R(X,Y)\omega$ is given by

$$(R(X,Y)\omega)(Z) = -\omega(R(X,Y)Z), \tag{7.1.39}$$

which characterises the action of curvature on basis 1-forms as

$$R\big(\partial_\mu, \partial_\nu\big)\, dx^\rho = -R^\rho{}_{\lambda\mu\nu}\, dx^\lambda. \tag{7.1.40}$$

This is the dual to the action on basis vectors, identity (7.1.28).

- **Action on Arbitrary Tensors** The action of $R(X,Y)$ on functions, basis vectors ∂_μ and basis 1-forms dx^μ, as well as the Leibniz rule (7.1.34) allow for the evaluation of its action on arbitrary tensors. For example, on a (0,2)-tensor $T \in \Gamma\big(T^{(0,2)}\mathcal{M}\big)$, we have

$$\begin{aligned}
R\big(\partial_\mu, \partial_\nu\big)T &= R\big(\partial_\mu, \partial_\nu\big)\big(T_{\rho\sigma}\, dx^\rho \otimes dx^\sigma\big) \\
&= T_{\rho\sigma}\big(R\big(\partial_\mu, \partial_\nu\big)\, dx^\rho\big) \otimes dx^\sigma + dx^\rho \otimes \big(R\big(\partial_\mu, \partial_\nu\big)\big)\, dx^\sigma\big) \\
&= T_{\rho\sigma}\Big(-R^\rho{}_{\lambda\mu\nu}\, dx^\lambda \otimes dx^\sigma - R^\sigma{}_{\lambda\mu\nu}\, dx^\rho \otimes dx^\sigma\Big) \\
&= \Big(-T_{\lambda\sigma} R^\lambda{}_{\rho\mu\nu} - T_{\rho\lambda} R^\lambda{}_{\sigma\mu\nu}\Big)\, dx^\rho \otimes dx^\sigma.
\end{aligned} \tag{7.1.41}$$

In general, the action of R on tensor components involves contracting each index of the tensor with either the first or second slot of $R^\rho{}_{\lambda\mu\nu}$, depending on the variance: contravariant indices contract against the first index ρ, acquiring a minus sign, while covariant indices contract against the second index λ with a positive sign.

- **Relationship to Connection Coefficients** Since the Riemann tensor is defined through the connection, there exists a relationship between its components and the connection coefficients appearing in

$$\nabla_\mu \partial_\nu = \Gamma^\lambda{}_{\mu\nu} \partial_\lambda. \tag{7.1.42}$$

We may derive this relationship from equation (7.1.30). Before anti-symmetrisation, the second covariant derivatives that appear read

$$\begin{aligned} \nabla_\mu \nabla_\nu \partial_\lambda = \nabla_\mu \left(\Gamma^\rho{}_{\nu\lambda} \partial_\rho \right) &= \left(\partial_\mu \Gamma^\rho{}_{\nu\lambda} \right) \partial_\rho + \Gamma^\rho{}_{\nu\lambda} \nabla_\mu \partial_\rho \\ &= \left(\partial_\mu \Gamma^\rho{}_{\nu\lambda} + \Gamma^\rho{}_{\mu\sigma} \Gamma^\sigma{}_{\nu\lambda} \right) \partial_\rho. \end{aligned} \tag{7.1.43}$$

Anti-symmetrising this expression in μ and ν leads us to

$$R(\partial_\mu, \partial_\nu)\partial_\lambda = [\nabla_\mu, \nabla_\nu]\partial_\lambda = \left(\partial_\mu \Gamma^\rho{}_{\nu\lambda} - \partial_\nu \Gamma^\rho{}_{\mu\lambda} + \Gamma^\rho{}_{\mu\sigma} \Gamma^\sigma{}_{\nu\lambda} - \Gamma^\rho{}_{\nu\sigma} \Gamma^\sigma{}_{\mu\lambda} \right) \partial_\rho, \tag{7.1.44}$$

from which we identify the components of the Riemann tensor as

$$R^\rho{}_{\lambda\mu\nu} = \partial_\mu \Gamma^\rho{}_{\nu\lambda} - \partial_\nu \Gamma^\rho{}_{\mu\lambda} + \Gamma^\rho{}_{\mu\sigma} \Gamma^\sigma{}_{\nu\lambda} - \Gamma^\rho{}_{\nu\sigma} \Gamma^\sigma{}_{\mu\lambda}. \tag{7.1.45}$$

7.1.3 Symmetries and Contractions of the Riemann Tensor

The Riemann curvature tensor introduced in the previous section exhibits a variety of intrinsic algebraic symmetries and contraction properties that depend on the underlying connection's characteristics. In this section, we derive and summarise these symmetries, distinguishing those that hold purely by definition from those that arise under additional assumptions such as metric compatibility and the absence of torsion. This will enable us to identify a canonical contraction of the Riemann tensor—the Ricci tensor—along with its trace, the *Ricci scalar* or *scalar curvature*.

To discuss algebraic symmetries, i.e. relations between permutations of the tensor slots, we introduce the fully covariant $(0,4)$-tensor associated to the Riemann curvature. It is defined by lowering the upper index of the $(1,3)$-curvature operator using the metric by defining

$$R(W, Z, X, Y) := g(R(X, Y)Z, W). \tag{7.1.46}$$

In components, this corresponds to

$$R_{\rho\lambda\mu\nu} W^\rho Z^\lambda X^\mu Y^\nu = g_{\rho\sigma} \left(R^\rho{}_{\lambda\mu\nu} X^\mu Y^\nu Z^\lambda \right) W^\sigma, \tag{7.1.47}$$

or equivalently,

$$R_{\rho\lambda\mu\nu} = g_{\rho\sigma} R^\rho{}_{\lambda\mu\nu}. \tag{7.1.48}$$

Thus, the fully covariant tensor arises simply by lowering the vector index of the curvature operator via the metric—one could also write

$$R(\cdot, Z, X, Y) = (R(X, Y)Z)^\flat. \tag{7.1.49}$$

We are now ready to discuss the symmetries of $R(W, Z, X, Y)$:

1. **Anti-Symmetry in Second Pair** By the anti-symmetry of the curvature operator, $R(X,Y) = -R(Y,X)$, which follows directly from the definition and the skew-symmetry of the commutator,

$$R(X,Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]} = -R(Y,X), \tag{7.1.50}$$

we obtain anti-symmetry in the last two slots of the fully covariant tensor, i.e.

$$R(W,Z,X,Y) = -R(W,Z,Y,X). \tag{7.1.51}$$

This holds for any connection. In components, we have

$$R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu} \quad \Leftrightarrow \quad R^{\rho}{}_{\sigma\mu\nu} = -R^{\rho}{}_{\sigma\nu\mu}. \tag{7.1.52}$$

2. **Anti-Symmetry in the First Pair** If the connection is metric-compatible (i.e. $\nabla g = 0$), the curvature operator acts on the metric as $R(X,Y)g = 0$. By the derivation propery of $R(X,Y)$, for any vector fields $W, Z \in \Gamma(T\mathcal{M})$, we have[8]

$$\begin{aligned} 0 &= R(X,Y)g(Z,W) \\ &= \underbrace{(R(X,Y)g)}_{=0}(Z,W) + \underbrace{g(R(X,Y)Z,W)}_{=R(W,Z,X,Y)} + \underbrace{g(R(X,Y)W,Z)}_{=R(Z,W,X,Y)} \\ &= R(W,Z,X,Y) + R(Z,W,X,Y) \end{aligned} \tag{7.1.53}$$

where the first equality with zero holds because $R(X,Y)$ is acting on a function. The above can be rearranged for

$$R(W,Z,X,Y) = -R(Z,W,X,Y), \tag{7.1.54}$$

which proves anti-symmetry in the first pair of arguments of the covariant tensor. In components, this reads

$$R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}. \tag{7.1.55}$$

3. **Bianchi Identity** We first derive an identity relating the sum of the cyclic permutations of $R(X,Y)Z$ with $X,Y,Z \in \Gamma(T\mathcal{M})$ and the torsion tensor. This identity will then produce a symmetry in the case of zero torsion. Note that for the torsion, we have

$$\nabla_X Y - \nabla_Y X = T(X,Y) + [X,Y] \tag{7.1.56}$$

using which we can begin writing out

$$\begin{aligned} &R(X,Y)Z + R(Y,Z)X + R(Z,X)Y \\ &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \\ &\quad + \nabla_Y \nabla_Z X - \nabla_Z \nabla_Y X - \nabla_{[Y,Z]} X \\ &\quad + \nabla_Z \nabla_X Y - \nabla_X \nabla_Z Y - \nabla_{[Z,X]} Y \\ &= \nabla_X (\nabla_Y Z - \nabla_Z Y) - \nabla_{[Y,Z]} X + \circlearrowleft_{X,Y,Z} \\ &= \nabla_X T(Y,Z) + \underbrace{\nabla_X [Y,Z] - \nabla_{[Y,Z]} X}_{=T(X,[Y,Z]) + [X,[Y,Z]]} \\ &= \nabla_X T(Y,Z) + T(X,[Y,Z]) + [X,[Y,Z]] + \circlearrowleft_{X,Y,Z} \end{aligned} \tag{7.1.57}$$

When expanding the sum over cyclic permutations of X,Y,Z indicated by $\circlearrowleft$, the last term drops out by the Jacobi identity for the commutator,

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0. \tag{7.1.58}$$

[8]This makes use of the property that $Dg(X,Y) = (Dg)(X,Y) + g(DX,Y) + g(X,DY)$ for any derivation D. This is a generalisation of what we have shown before in eq. (6.4.8) for the particular case of $D = \nabla$.

We hence arrive at the identity

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = \nabla_X T(Y,Z) + T(X,[Y,Z]) + \underset{X,Y,Z}{\circlearrowleft} . \quad (7.1.59)$$

In the case of vanishing torsion, $T = 0$, this turns into the symmetry

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0, \quad (7.1.60)$$

and equivalently,

$$R(W,Z,X,Y) + R(W,X,Y,Z) + R(W,Y,Z,X) = 0. \quad (7.1.61)$$

or in terms of components,

$$R^{\rho}{}_{[\sigma\mu\nu]} = R_{\rho[\sigma\mu\nu]} = 0. \quad (7.1.62)$$

Here, the cyclic permutation of the last three indices/slots is proportional to their anti-symmetrisation since we have anti-symmetry in the last two indices/slots.

4. **Symmetry in First and Second Pair** For a connection that is both metric-compatible and has vanishing torsion—i.e., the Levi-Civita connection—the above symmetries imply a further symmetry,

$$R(W,Z,X,Y) = R(X,Y,W,Z). \quad (7.1.63)$$

That is, the expression is symmetric under the exchange of the first and second pair of slots. This is shown by repeatedly applying the symmetries 1., 2. and 3. to the left-hand side:

$$\begin{aligned}
R(W,Z,X,Y) &\overset{2.}{=} -R(Z,W,X,Y) \\
&\overset{3.}{=} R(Z,X,Y,W) + R(Z,Y,W,X) \\
&\overset{2.}{=} -R(X,Z,Y,W) - R(Y,Z,W,X) \\
&\overset{3.}{=} R(X,Y,W,Z) + R(X,W,Z,Y) + R(Y,W,X,Z) + \underbrace{R(Y,X,Z,W)}_{\overset{1.\,\&\,2.}{=} R(X,Y,W,Z)} \\
&\overset{2.}{=} 2R(X,Y,W,Z) - R(W,X,Z,Y) - R(W,Y,X,Z) \\
&\overset{3.}{=} 2R(X,Y,W,Z) + R(W,Z,Y,X) \\
&\overset{1.}{=} 2R(X,Y,W,Z) - R(W,Z,X,Y).
\end{aligned} \quad (7.1.64)$$

The claim now follows when adding $R(W,Z,X,Y)$ to both sides. In terms of the tensor components, the symmetry reads

$$R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}. \quad (7.1.65)$$

This is certainly not the most elegant or efficient way to establish this symmetry. However, since the result follows directly from the symmetries imposed by metric compatibility and vanishing torsion, and offers little in terms of geometric insight, a somewhat brute-force algebraic proof suffices for our purposes.

The Riemann curvature, being a rank 4 tensor, allows for contractions between its slots. In principle, for 4 slots, there are twelve combinations for contractions—however, symmetries reduce these significantly. For a metric-compatible connection, we have anti-symmetry in the first and second pairs. This means that the contractions over these pairs vanish,

$$\operatorname{tr}^1_2 R(\cdot,\cdot,X,Y) = \operatorname{tr}^1_2 R(W,Z,\cdot,\cdot) = 0. \quad (7.1.66)$$

Hence, for a contraction not to vanish, it must contract over one slot in the first and one in the second pair. Without loss of generality, this contraction can be performed over the first an third slot, as all others are related by signs (due to anti-symmetry in the first and second pair). This contraction defines a new rank 2 tensor called the *Ricci tensor*, denoted by[9]

$$\mathrm{Ric}(X, Y) = \mathrm{tr}_3^1\, R(\cdot, X, \cdot, Y). \tag{7.1.67}$$

In components, this reads

$$\mathrm{Ric}(X, Y) = R_{\mu\nu} X^\mu Y^\nu, \tag{7.1.68}$$

where the Ricci tensor components $R_{\mu\nu}$ emerge as a contraction of the Riemann tensor,

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}. \tag{7.1.69}$$

If further, the connection is torsion-free, we have symmetry of the covariant Riemann tensor under the exchange of the first and second pairs, making the Ricci tensor symmetric, i.e.,

$$\mathrm{Ric}(X, Y) = \mathrm{Ric}(Y, X). \tag{7.1.70}$$

At this point, we can perform a second contraction to obtain the so-called Ricci scalar

$$\mathcal{R} = \mathrm{tr}_2^1\, \mathrm{Ric}(\cdot, \cdot) = g^{\mu\nu} R_{\mu\nu}. \tag{7.1.71}$$

For a Levi-Civita connection, it is the unique scalar contraction of the Riemann tensor[10].

7.1.4 Curvature of the Induced Connection

Previously, we have defined the induced connection $\bar{\nabla}$ on the submanifolds of a foliation Σ of a smooth manifold $\mathcal{M}$ equipped with the ambient connection ∇. The ambient connection gives rise to the Riemann curvature tensor

$$R : \Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \to \Gamma(T\mathcal{M}), \tag{7.1.72}$$

with

$$R(X, Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z. \tag{7.1.73}$$

In this context, it is referred to as the *ambient curvature*.

Since each of the foliation's submanifolds are equipped with the induced connection $\bar{\nabla}$—the projection of the ambient connection ∇ onto the tangent bundle $T\Sigma$—we can define a Riemann curvature tensor with respect to it on each of the leaves of the foliation, leading to

$$\bar{R} : \Gamma(T\Sigma) \times \Gamma(T\Sigma) \times \Gamma(T\Sigma) \to \Gamma(T\Sigma) \tag{7.1.74}$$

with

$$\bar{R}(X, Y)Z = \left[\bar{\nabla}_X, \bar{\nabla}_Y\right]Z - \bar{\nabla}_{[X,Y]}Z. \tag{7.1.75}$$

We refer to $\bar{R}$ as the *intrinsic (Riemann) curvature* of the foliation Σ. On any individual leaf Σ_t, $\bar{R}$ is simply the Riemann tensor of the induced connection—that is, the curvature one would assign having access only to the intrinsic geometry, i.e., without any knowledge of the ambient geometry.

[9]Here, the contraction over two covariant slots is to be interpreted as with respect to the inverse metric.

[10]The unique scalar linear in the Riemann tensor. Of course, there are other contractions possible at higher orders.

Since both metric compatibility and absence of torsion are properties $\bar{\nabla}$ inherits from ∇, the fully covariant tensors $R(W,Z,X,Y)$ and $\bar{R}(W,Z,X,Y)$ exhibit the same algebraic symmetries (cf. previous section). In the case that ∇ is the Levi-Civita connection associated to the metric g on $\mathcal{M}$, then $\bar{\nabla}$ is the Levi-Civita connection associated to the induced metric $\gamma = \iota^* g$ on the leaves of the foliation. This makes the components of $\bar{R}$ computable entirely from the components of the induced metric, due to the relationships

$$\begin{aligned}\bar{R}^k{}_{\ell ij} &= \partial_i \bar{\Gamma}^k{}_{j\ell} - \partial_j \bar{\Gamma}^k{}_{i\ell} + \bar{\Gamma}^k{}_{im}\bar{\Gamma}^m{}_{j\ell} - \bar{\Gamma}^k{}_{jm}\bar{\Gamma}^m{}_{i\ell}, \\ \bar{\Gamma}^k{}_{ij} &= \frac{1}{2}\gamma^{k\ell}\big(\gamma_{\ell i,j} + \gamma_{\ell j,i} - \gamma_{ij,\ell}\big)\end{aligned} \tag{7.1.76}$$

we have established in preceding sections. We may define an intrinsic Ricci curvature tensor as

$$\overline{\mathrm{Ric}}(X,Y) := \mathrm{tr}^1_3\, \bar{R}(\cdot, X, \cdot, Y), \tag{7.1.77}$$

and an intrinsic scalar curvature,

$$\bar{\mathcal{R}} := \mathrm{tr}^1_2\, \overline{\mathrm{Ric}}(\cdot,\cdot) = \gamma^{ij}\bar{R}_{ij}, \tag{7.1.78}$$

where $\bar{R}_{ij}$ are the components of the induced Ricci tensor, given by

$$\bar{R}_{ij} = \bar{R}^k{}_{ikj}. \tag{7.1.79}$$

7.1.5 Intrinsic vs Projected Ambient Curvature

A natural question following the above introduction of the intrinsic curvature $\bar{R}$ on a foliation Σ is how it relates to the ambient curvature R. Since $\bar{\nabla}$ is defined as the projection of ∇ onto $T\Sigma$, one might expect a similar relationship between the intrinsic and ambient curvature—something like

$$\bar{R}(X,Y)Z = (\iota_*)^{-1}\Big(R(\iota_* X, \iota_* Y)(\iota_* Z)\Big), \quad X,Y,Z \in \Gamma(T\Sigma) \subset \Gamma(T\mathcal{M}). \tag{7.1.80}$$

where $(\iota_*)^{-1}$ denotes the orthogonal left-inverse of the pushforward ι_*. This would follow the same pattern we have for the connection: the action of $R(X,Y)Z$ is simply followed by a projection onto the tangent bundle $T\Sigma$, removing any normal contributions that might emerge.

However, equation (7.1.80) is *not* correct. There are (at least) two ways to see this—first, through a concrete counterexample; and second, via an algebraic derivation that, while less geometrically intuitive, reveals the deeper structure behind the failure of this naïve guess. Let us consider both perspectives in turn.

- **Curved Submanifolds of Flat Manifolds** Let us reconsider a recurring example from these notes: The foliation by origin-centered spheres of the ambient manifold $\mathbb{R}^3 \setminus \{0\}$, equipped with the Euclidean metric

  $$\begin{aligned} g &= dx \otimes dx + dy \otimes dy + dz \otimes dz \\ &= dr \otimes dr + r^2\big(d\theta \otimes d\theta + \sin^2\theta\, d\varphi \otimes d\varphi\big), \end{aligned} \tag{7.1.81}$$

  —where $x^\mu = (x,y,z)$ are Cartesian and $y^\alpha = (r, y^i) = (r,\theta,\varphi)$ spherical coordinates. The connection under consideration is the Levi-Civita connection associated with this metric. Our goal is to compute and compare the ambient and induced Riemann tensor components.

  Observe that, in Cartesian coordinates, the Levi-Civita has vanishing coefficients, i.e.

  $$\Gamma^\lambda{}_{\mu\nu} = 0, \tag{7.1.82}$$

  as the metric components $g_{\mu\nu} = \delta_{\mu\nu}$ are constant. From this, it follows immediately that the Riemann curvature tensor also vanishes in these coordinates,

$$R^{\rho}{}_{\lambda\mu\nu} = 0, \tag{7.1.83}$$

due to the formula (7.1.45). In contrast to equation (7.1.82), the above equation is tensorial—that is, it holds in any coordinate system—and we conclude that the ambient manifold is flat:

$$R(X,Y)Z = 0, \quad \forall X, Y, Z \in \Gamma(T\mathcal{M}), \tag{7.1.84}$$

If our earlier naïve guess (7.1.80) were correct, this would imply that

$$\bar{R}(X,Y)Z = P(R(X,Y)Z) = 0, \tag{7.1.85}$$

and thus that the intrinsic curvature vanishes as well.

This already casts doubt on our naïve guess—after all, the word "sphere" does not readily evoke "flatness". Let us indulge this suspicion and compute the components of the intrinsic curvature explicitly to see where the discrepancy lies. We already derived the induced connection coefficients for this foliation in Section 6.5, arriving at the following non-zero Christoffel symbols:

$$\bar{\Gamma}^{\theta}{}_{\varphi\varphi} = -\sin\theta\cos\theta, \qquad \bar{\Gamma}^{\varphi}{}_{\theta\varphi} = \bar{\Gamma}^{\varphi}{}_{\varphi\theta} = \cot\theta. \tag{7.1.86}$$

At first glance, computing the components of a rank-4 tensor like the Riemann curvature might seem daunting. However, in two dimensions—and with the Levi-Civita connection—the symmetries of the Riemann tensor drastically reduce the number of independent components. Specifically, the antisymmetry in both the first and second pair of indices,

$$\bar{R}_{\rho\sigma\mu\nu} = -\bar{R}_{\rho\sigma\nu\mu} = -\bar{R}_{\sigma\rho\mu\nu}, \tag{7.1.87}$$

implies that the indices θ and φ must each appear exactly once in both pairs. All valid permutations are then related by symmetry. Moreover, since the metric is diagonal, we only need to compute a single nontrivial component, say $\bar{R}^{\theta}{}_{\varphi\theta\varphi}$. We proceed by applying the standard formula:

$$\begin{aligned}
\bar{R}^{\theta}{}_{\varphi\theta\varphi} &= \partial_\theta \bar{\Gamma}^{\theta}{}_{\varphi\varphi} - \underbrace{\partial_\varphi \bar{\Gamma}^{\theta}{}_{\theta\varphi}}_{=0} + \underbrace{\bar{\Gamma}^{\theta}{}_{\theta i}\bar{\Gamma}^{i}{}_{\varphi\varphi}}_{=0} - \bar{\Gamma}^{\theta}{}_{\varphi i}\bar{\Gamma}^{i}{}_{\theta\varphi} \\
&= -\partial_\theta(\sin\theta\cos\theta) - \bar{\Gamma}^{\theta}{}_{\varphi\varphi}\bar{\Gamma}^{\varphi}{}_{\theta\varphi} \\
&= -\cos^2\theta + \sin^2\theta + \underbrace{\sin\theta\cos\theta\cot\theta}_{=\cos^2\theta} \\
&= \sin^2\theta.
\end{aligned} \tag{7.1.88}$$

This is very clearly *not* zero. We have thus found a counterexample to our naïve guess (7.1.80). We conclude that the intrinsic curvature does not, in general, arise from a simple projection of the ambient curvature.

This is a good point to take a step back and generalise the insight, in order to build further intuition for why our guess cannot be correct. What we have done is the following: we took a flat manifold, $\mathbb{R}^3 \setminus \{0\}$, and foliated it into surfaces that are scaled copies of the 2-sphere. Intuitively, spheres possess curvature—this is evident from the fact that their normal vector field varies as one moves along their surface. Our guess (7.1.80), however, attempted to capture something quite different: it projected the ambient curvature tensor (which vanishes in this case) onto the tangent bundle of the foliation (where it still vanishes). The projection of the ambient curvature captures only the part of the it that is tangential to the foliation; it entirely neglects how the surface itself bends within the ambient space. In other words, this projection measures the curvature of the background in which the leaves of the foliation live, but not how those leaves curve within it. The change of the normal field vector plays no role in this projection. Hence, while the projected ambient curvature may contribute to the intrinsic curvature $\bar{R}$, it clearly does not suffice to determine it completely: the way in which the surface curves relative to the background also generates intrinsic curvature.

- **Algebraic Argument** In Section 6.3, we introduced the induced connection on a foliation Σ by defining its action on vector fields as

$$\bar{\nabla}_X Y = (\iota_*)^{-1} \nabla_{\iota_* X}(\iota_* Y), \quad X, Y \in \Gamma(T\Sigma), \tag{7.1.89}$$

where ι_* is the pushforward of the inclusion map $\iota : \Sigma \to \mathcal{M}$, and $(\iota_*)^{-1}$ is the unique left-inverse with the property that

$$P = \iota_* \circ (\iota_*)^{-1} \tag{7.1.90}$$

is the orthogonal projection from $T\mathcal{M}$ onto $T\Sigma$. This is a very precise definition, as it distinguishes $T\Sigma$ and $\mathrm{im}(\iota_*) \subset T\mathcal{M}$ as separate objects.

In practice, however—particularly when working with foliations—it is often more convenient to treat $T\Sigma$ as a proper subbundle of $T\mathcal{M}$, which we may do via the linear embedding map ι_*. From this perspective, the induced connection takes a simpler and more direct form:

$$\bar{\nabla}_X Y = P \nabla_X Y, \quad X, Y \in T\Sigma. \tag{7.1.91}$$

Since $T\Sigma \subset T\mathcal{M}$, both X and Y are valid inputs for the ambient connection ∇. The appearance of the full projector P on the right-hand side is then a result of pushing forward the image of $(\iota_*)^{-1}$ into $T\mathcal{M}$, which gives rise to the combination (7.1.90). We will adopt this more algebraic, embedded viewpoint for the remainder of the discussion, as it makes many derivations more transparent: equation (7.1.91) makes clear that the induced connection is simply the projection of the ambient connection onto the tangent bundle of the foliation.

With this notational preface in place, we are now ready to give an algebraic argument for why the intrinsic curvature $\bar{R}$ cannot, in general, be written as the orthogonal projection of the ambient curvature R onto $T\Sigma$.

Inserting into the definition of $\bar{R}$, we find that for $X, Y, Z \in \Gamma(T\Sigma)$,

$$\begin{aligned} \bar{R}(X,Y)Z &= \left[\bar{\nabla}_X, \bar{\nabla}_Y\right] Z - \bar{\nabla}_{[X,Y]} Z \\ &= [P\nabla_X, P\nabla_Y] Z - P\nabla_{[X,Y]} Z. \end{aligned} \tag{7.1.92}$$

The projection of the ambient curvature, on the other hand, reads

$$PR(X,Y)Z = P[\nabla_X, \nabla_Y]Z - P\nabla_{[X,Y]}Z, \tag{7.1.93}$$

so the difference lies in the first term. Let us now expand the commutator appearing in $\bar{R}$ as

$$\begin{aligned} [P\nabla_X, P\nabla_Y]Z &= P\nabla_X(P\nabla_Y Z) - (X \leftrightarrow Y) \\ &= \underbrace{P^2}_{=P} \nabla_X \nabla_Y Z + P(\nabla_X P)\nabla_Y Z - (X \leftrightarrow Y) \\ &= P[\nabla_X, \nabla_Y]Z + P\big((\nabla_X P)\nabla_Y Z - (\nabla_Y P)\nabla_X Z\big) \end{aligned} \tag{7.1.94}$$

We can see that there is an additional term involving covariant derivatives of the projector P —we hence obtain the identity

$$\bar{R}(X,Y)Z = PR(X,Y)Z + P\big((\nabla_X P)\nabla_Y Z - (\nabla_Y P)\nabla_X Z\big). \tag{7.1.95}$$

This demonstrates that the intrinsic curvature contains more than just the projection of the ambient curvature. In fact, the above is the algebraic seed to the *Gauss-Codazzi equation*, a result we will steadily work towards throughout the following sections.

To properly interpret the additional term, we must however first introduce the concept of *extrinsic curvature*. As it stands, the expression

$$P\big((\nabla_X P)\nabla_Y Z - (\nabla_Y P)\nabla_X Z\big) \tag{7.1.96}$$

clearly introduces a dependence on how the submanifold is situated within the ambient space—it involves the derivative of the projector onto the tangent bundle, and hence reflects how that bundle varies from point to point. This suggests that the *shape* of the submanifold is encoded here—the part that we found to be missing in the concrete example above. However, the precise geometric content of this term is not transparent in its current algebraic form. The introduction of extrinsic curvature will allow us to isolate and interpret this shape-dependence more clearly, ultimately leading to the Gauss-Codazzi equation. This equation links intrinsic, extrinsic and ambient curvature in a precise way.

7.2 Extrinsic Curvature

7.2.1 Algebraic Motivation for Extrinsic Curvature

In this section, we introduce the concept of *extrinsic curvature*. Its connection to the additional term (7.1.96) appearing in the expression for $\bar{R}$ will not be immediate—we will establish that link later. For now, we treat it as a concept in its own right, motivated by the structure behind the ambient and induced connections.

Recall from the previous section that the induced connection can be written as

$$\bar{\nabla}_X Y = P\nabla_X Y, \tag{7.2.1}$$

when we regard $T\Sigma \subset T\mathcal{M}$ as an actual subbundle via the embedding ι_*. This form makes it clear that $\bar{\nabla}$ is simply the tangential projection of the ambient connection. Crucially, this means that certain components of $\nabla_X Y$ are *discarded* in the process. Which components are lost? The short answer is: the *normal* ones—but this deserves a more precise formulation.

To that end, recall that the metric g on $\mathcal{M}$ allows a decomposition into tangential and normal projectors,

$$g = P + Q, \tag{7.2.2}$$

when viewed as a $(0,2)$-tensor. Interpreted as $(1,1)$-tensors, this becomes the orthogonal identity decomposition

$$\mathrm{id} = P + Q, \tag{7.2.3}$$

where

$$\mathrm{id} = \delta^\mu_\nu \partial_\mu \otimes dx^\nu, \quad P = P^\mu{}_\nu \partial_\mu \otimes dx^\nu, \quad Q = Q^\mu{}_\nu \partial_\mu \otimes dx^\nu. \tag{7.2.4}$$

This follows from raising an index of the metric,

$$g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu_\nu. \tag{7.2.5}$$

Now consider the action of the ambient connection on two tangent vector fields $X, Y \in \Gamma(T\Sigma)$. Since $\mathrm{id} = P + Q$, we may decompose

$$\nabla_X Y = P\nabla_X Y + Q\nabla_X Y = \bar{\nabla}_X Y + Q\nabla_X Y. \tag{7.2.6}$$

This leads us to make a definition: we now introduce the *extrinsic curvature* (also called the *second fundamental form*[11]) as

$$K(X,Y) := Q\nabla_X Y. \tag{7.2.7}$$

It captures the normal part of the ambient covariant derivative of one tangent vector along another. We can thus write the clean decomposition

$$\nabla_X Y = \bar{\nabla}_X Y + K(X,Y), \tag{7.2.8}$$

which splits the ambient connection into its tangential and normal components relative to the foliation Σ.

Let us now examine why this object might be referred to as the extrinsic curvature. Firstly, we note that the above relationship can be turned around:

$$\bar{\nabla}_X Y = \nabla_X Y - K(X,Y). \tag{7.2.9}$$

While this may appear trivial at first glance, it offers a useful geometric perspective: $K(X,Y)$ is the *correction* necessary to make $\nabla_X Y$ tangent to Σ. Such a correction becomes necessary when the tangent plane "tilts"—for instance, on a sphere, moving from one point to another causes the tangent plane to rotate. The full derivative $\nabla_X Y$ will generally not remain in the new tangent plane, and thus must have parts of it "chopped off" to become tangent. That "chopped-off" part is $K(X,Y)$. It hence encodes the change in orientation of the tangent spaces—or equivalently, the change in the normal spaces—as one moves along the submanifold. This makes it clear why K is rightly referred to as the *extrinsic curvature*: it measures how the submanifold is shaped *within* the ambient manifold. This is exactly what we found to be missing when guessing that $\bar{R} = PR|_{T\Sigma}$: a measure of how the leaves themselves curve on top of the background. How $K(X,Y)$ relates to $\bar{R}$ precisely remains to be established—but it certainly feels like a step in the right direction.

The idea of K encoding the change in orientation of the tangent/normal spaces is also reflected in an alternative expression one may derive for it. By partial integration, we obtain

$$K(X,Y) = Q\nabla_X Y = \nabla_X(QY) - (\nabla_X Q)Y. \tag{7.2.10}$$

Since $QY = 0$ for $Y \in \Gamma(T\Sigma)$, this reduces to

$$K(X,Y) = -(\nabla_X Q)Y, \tag{7.2.11}$$

an equivalent but different characterisation of the extrinsic curvature. This version highlights another aspect of the same idea: extrinsic curvature measures how the normal projection operator changes along X, with the change being probed by acting on Y. In other words, it tracks how the normal bundle twists and turns along the leaves of the foliation—consistent with our earlier interpretation in terms of tilting tangent spaces.

7.2.2 Geometric Interpretation via the Normal Vector

The definition of extrinsic curvature we presented in the previous section is valid for any foliation Σ of a smooth manifold $\mathcal{M}$, regardless of the codimension of the leaves. In this section, we specialise to a hypersurface foliation, where the expression simplifies considerably.

In the case where $\Sigma = \Sigma_t$ is a hypersurface foliation generated by a scalar function $t \in C^\infty(\mathcal{M})$, the normal bundle has one-dimensional fibres. From Section 5.2, we recall that in this scenario, we are provided with a normal vector field $n^\sharp$, associated with the normal 1-form $n = \alpha\, dt$, which satisfies the relations

[11] The first fundamental form is the induced metric, but I will not use this terminology in these notes.

$$g(n^\sharp, n^\sharp) = \varepsilon = \pm 1, \quad g(n^\sharp, X) = 0, \quad \forall X \in T\Sigma. \tag{7.2.12}$$

Thus any normal vector field $N \in \Gamma(N\Sigma)$ can be written as a scalar multiple of $n^\sharp$,

$$N = \lambda n^\sharp, \quad \lambda \in C^\infty(\mathcal{M}). \tag{7.2.13}$$

This means that for any hypersurface foliation, we may express $K(X,Y)$—because it is normal to Σ—as

$$K(X,Y) = k(X,Y)n^\sharp, \tag{7.2.14}$$

where

$$k(X,Y) := \varepsilon g\big(K(X,Y), n^\sharp\big) \tag{7.2.15}$$

substitutes as the proportionality factor λ from the previous equation. The map $k(X,Y)$ is a scalar-valued bilinear map,

$$k : \Gamma(T\Sigma) \times \Gamma(T\Sigma) \to \mathbb{R}, \tag{7.2.16}$$

which—in components—can be written as

$$k = k_{\mu\nu}\, dx^\mu \otimes dx^\nu. \tag{7.2.17}$$

Since k only takes vectors tangent to the foliation as inputs, in adapted coordinates (t, y^i), we can also write it in terms of the cobasis associated with the transverse coordinates y^i as

$$k = k_{ij}\, dy^i \otimes dy^j = k_{ij} E^i_\mu E^j_\nu\, dx^\mu \otimes dx^\nu, \tag{7.2.18}$$

implying that k is its own pullback; we have $k = \iota^* k$. Its action on vector fields $X = X^\mu \partial_\mu = X^i \partial_i$ and $Y = Y^\mu \partial_\mu = Y^i \partial_i$, both in $\Gamma(T\Sigma)$, is given by

$$k(X,Y) = k_{\mu\nu} X^\mu Y^\nu = k_{ij} X^i Y^j, \tag{7.2.19}$$

making use of the fact that $T\Sigma \subset T\mathcal{M}$ is a proper subspace via the embedding ι_*.

We now examine the proportionality factor $k(X,Y)$ in more detail. Explicitly—by inserting the definition of extrinsic curvature $K(X,Y)$—it expands to

$$k(X,Y) = \varepsilon g\big(K(X,Y), n^\sharp\big) = \varepsilon g\big(Q\nabla_X Y, n^\sharp\big) = \varepsilon g\big(\nabla_X Y, Q n^\sharp\big) = \varepsilon g\big(\nabla_X Y, n^\sharp\big), \tag{7.2.20}$$

since Q is orthogonal and acts on $n^\sharp$ as the identity. For a metric-compatible connection (i.e., when $\nabla g = 0$, an assumption we make from hereon out), we may further rewrite this as

$$k(X,Y) = \varepsilon g\big(\nabla_X Y, n^\sharp\big) = \varepsilon g\big(Y, -\nabla_X n^\sharp\big). \tag{7.2.21}$$

Upon defining the *shape operator*[12] $S : \Gamma(T\Sigma) \to \Gamma(T\Sigma)$

$$S(X) = -\nabla_X n^\sharp, \tag{7.2.22}$$

this becomes

$$k(X,Y) = \varepsilon g(S(X), Y). \tag{7.2.23}$$

Comment: Correct this, should involve an additional projection or at least a comment since ∂_μ is the basis of $T\mathcal{M}$ and k is not defined on all of it. Explicitly, the components of k are thus given by

[12] A priori, it is unclear that the domain is $\Gamma(T\Sigma)$ and not $\Gamma(T\mathcal{M})$. However, due to normalisation of $n^\sharp$, we have $0 = \nabla_X \varepsilon = \nabla_X g(n^\sharp, n^\sharp) = -2g\big(S(X), n^\sharp\big)$ which shows that $S(X)$ has no normal component.

$$k_{\mu\nu} = -\varepsilon g_{\nu\lambda}(\nabla_\mu n)^\lambda = -\varepsilon(\nabla_\mu n)_\nu = -\varepsilon\nabla_\mu(\alpha\, dt)_\nu, \tag{7.2.24}$$

or equivalently,

$$k_{ij} = -\varepsilon E_i^\mu E_j^\nu (\nabla_\mu n)_\nu, \tag{7.2.25}$$

with respect to the transverse coordinates y^i. As a tensor, k is hence nothing but

$$k = -\varepsilon\nabla n. \tag{7.2.26}$$

We have now derived a rich set of equations. Let us conclude by reflecting on the geometric meaning behind these expressions. The central identity

$$k(X,Y) = \varepsilon g(S(X),Y), \quad \text{where} \quad S(X) = -\nabla_X n^\sharp, \tag{7.2.27}$$

and the relationship

$$K(X,Y) = k(X,Y)n^\sharp, \tag{7.2.28}$$

show that the extrinsic curvature is closely linked to the change in the normal vector field as one moves along the foliation. The shape operator $S(X)$ encodes how the normal vector evolves when traversing a leaf, revealing the curvature of the hypersurface within the ambient manifold. The scalar $k(X,Y)$ then measures the projection of this change along a tangent vector Y, providing a clear geometric interpretation of the curvature in the case of hypersurface foliations.

In summary, the extrinsic curvature provides a measure of how the leaves are embedded and shaped within the ambient manifold, with $k(X,Y)$ quantifying the degree of this deformation. In the specific case of hypersurfaces, this bending is described by the change of the normal vector field $n^\sharp$.

7.2.3 Symmetry of the Extrinsic Curvature

We have previously observed that, under certain conditions such as metric compatibility or vanishing torsion, the Riemann curvature tensor acquires additional symmetries. A similar result holds for the extrinsic curvature $K(X,Y)$, which becomes a (vector-valued) symmetric bilinear form when the connection is torsion-free.

Recall that the torsion tensor is defined as

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y], \quad X,Y \in \Gamma(T\mathcal{M}), \tag{7.2.29}$$

such that the requirement of vanishing torsion is equivalent to

$$\nabla_X Y = \nabla_Y X + [X,Y]. \tag{7.2.30}$$

Moreover, since the Lie bracket is a closed operation on the tangent spaces of (sub)manifolds, for $X, Y \in T\Sigma$ we also have $[X,Y] \in T\Sigma$.

With this in mind, it is rather straightforward to show that the extrinsic curvature tensor is symmetric. We simply derive

$$K(X,Y) = Q\nabla_X Y = \underbrace{Q\nabla_Y X}_{=K(Y,X)} + \underbrace{Q[X,Y]}_{=0} = K(Y,X). \tag{7.2.31}$$

Here, we made use of the fact that Q is the normal projection onto $N\Sigma$, whereas $[X,Y] \in T\Sigma$, so that Q annihilates it. We have thus shown that the extrinsic curvature is symmetric, i.e.,

$$K(X,Y) = K(Y,X). \tag{7.2.32}$$

For the case of a hypersurface foliation, we also obtain symmetry for the scalar factor $k(X, Y)$ in the expression $K(X, Y) = k(X, Y)n^\sharp$. Specifically, we find

$$k(X, Y) = k(Y, X), \tag{7.2.33}$$

showing that the scalar function $k(X, Y)$, which encodes the magnitude of the extrinsic curvature, is symmetric as well.

7.2.4 Extrinsic Curvature as Lie Derivative of the Metric

In the case that the ambient manifold $\mathcal{M}$ is endowed with a metric-compatible and torsion-free connection ∇, and Σ is a hypersurface foliation of it, one can derive a further identity relating the magnitude factor $k(X, Y)$ of the extrinsic curvature to the Lie derivative $\mathcal{L}_{n^\sharp}$ of the metric along the normal flow generated by $n^\sharp$.

To derive this, consider $X, Y \in \Gamma(T\Sigma)$. We begin by expanding

$$(\mathcal{L}_{n^\sharp} g)(X, Y) = n^\sharp[g(X, Y)] - g\big(\big[n^\sharp, X\big], Y\big) - g\big(X, \big[n^\sharp, Y\big]\big), \quad X, Y \in \Gamma(T\Sigma), \tag{7.2.34}$$

where we made use of the fact that $\mathcal{L}$ fulfils a Leibniz rule, that it acts as a directional derivative on scalars, and $\mathcal{L}_X Y = [X, Y]$ on vectors. We now treat the individual terms separately.

For the first term, we may replace the directional with a covariant derivative since it is acting on a scalar. After applying the product rule and metric compatibility, we get

$$n^\sharp[g(X, Y)] = \nabla_{n^\sharp} g(X, Y) = g(\nabla_{n^\sharp} X, Y) + g(X, \nabla_{n^\sharp} Y). \tag{7.2.35}$$

The second (and third term, analogously), may be rewritten by making use of the identity

$$[X, Y] = \nabla_X Y - \nabla_Y X \tag{7.2.36}$$

that holds for torsion-free connections. Inserting it, we find

$$g\big(\big[n^\sharp, X\big], Y\big) = g(\nabla_{n^\sharp} X, Y) - g\big(\nabla_X n^\sharp, Y\big). \tag{7.2.37}$$

Putting everything back together, we arrive at

$$\begin{aligned}(\mathcal{L}_{n^\sharp} g)(X, Y) &= \cancel{g(\nabla_{n^\sharp} X, Y)} - \cancel{g(\nabla_{n^\sharp} X, Y)} + \underbrace{g\big(\nabla_X n^\sharp, Y\big)}_{=-\varepsilon k(X,Y)} + (X \leftrightarrow Y) \\ &= -\varepsilon k(X, Y) - \varepsilon k(Y, X) \\ &= -2\varepsilon k(X, Y),\end{aligned} \tag{7.2.38}$$

implying the identity

$$k(X, Y) = -\frac{\varepsilon}{2}(\mathcal{L}_{n^\sharp} g)(X, Y), \quad X, Y \in \Gamma(T\Sigma) \tag{7.2.39}$$

or equivalently,

$$k = -\frac{\varepsilon}{2}\mathcal{L}_{n^\sharp} g|_{\Gamma(T\Sigma)}. \tag{7.2.40}$$

This gives a direct relationship between the extrinsic curvature and the metric; more precisely, it tells us that its magnitude corresponds to the change of the metric along the normal flow of the foliation.

We can make this more explicit in terms of the lapse α, the shift β and the induced metric γ that appear in the ADM-decomposed form of the metric we derived in Section 5.4.2. Recall that in terms of these quantities, the metric components with respect to the adapted coordinates (t, y^i) are given by

$$g_{tt} = \varepsilon\alpha^2 + \gamma_{ij}\beta^i\beta^j, \quad g_{ti} = \gamma_{ij}\beta^j, \quad g_{ij} = \gamma_{ij}. \tag{7.2.41}$$

Moreover, by the definition of α and β, the normal vector $n^\sharp$ can be written as

$$n^\sharp = \frac{\varepsilon}{\alpha}(\partial_t - \beta). \tag{7.2.42}$$

Since the right-hand side of equation (7.2.40) is restricted to $\Gamma(T\Sigma)$, the only nonzero components[13] of k are given by

$$k_{ij} = -\frac{\varepsilon}{2}(\mathcal{L}_{n^\sharp}g)_{ij} = -\frac{\varepsilon}{2}\Big(n^\sharp[g_{ij}] + (\partial_i n^\mu)g_{\mu j} + (\partial_j n^\mu)g_{i\mu}\Big). \tag{7.2.43}$$

The second equality simply carries out the Lie derivative explicitly. Let us now examine the terms individually by inserting the metric components (7.2.41).

We begin with the first term. Here, we find

$$n^\sharp[g_{ij}] = \frac{\varepsilon}{\alpha}\partial_t\gamma_{ij} - \frac{\varepsilon}{\alpha}\beta[\gamma_{ij}]. \tag{7.2.44}$$

This is simply a directional derivative of the induced metric, encoding both the change along the time direction ∂_t as well as the shift β.

Let us proceed to the second term. This one expands to

$$\begin{aligned}(\partial_i n^\mu)g_{\mu j} &= (\partial_i n^t)g_{tj} + (\partial_i n^k)g_{kj} = \Big(\partial_i\frac{\varepsilon}{\alpha}\Big)\gamma_{jk}\beta^k - \Big(\partial_i\Big(\frac{\varepsilon}{\alpha}\beta^k\Big)\Big)\gamma_{kj} \\ &= -\frac{\varepsilon}{\alpha}(\partial_i\beta^k)\gamma_{kj}.\end{aligned} \tag{7.2.45}$$

By analogy, the last term turns into

$$(\partial_j n^\mu)g_{i\mu} = -\frac{\varepsilon}{\alpha}(\partial_j\beta^k)\gamma_{ik}. \tag{7.2.46}$$

Collecting the results and recombining the individual terms, we arrive at

$$k_{ij} = -\frac{1}{2\alpha}\Big(\partial_t\gamma_{ij} - \Big(\beta[\gamma_{ij}] + (\partial_i\beta^k)\gamma_{kj} + (\partial_j\beta^k)\gamma_{ik}\Big)\Big). \tag{7.2.47}$$

We recognise the inner parentheses as the Lie derivative $\bar{\mathcal{L}}_\beta\gamma$. Since both γ and β are intrinsic to the foliation, this really is a Lie derivative on the leaves, hence the notation $\bar{\mathcal{L}}$—the bar indicates a foliation-intrinsic object, i.e., an object that can be computed from only the submanifold geometry. In particular, this means that this term can be rewritten in terms of the induced covariant derivative $\bar\nabla$:

$$\begin{aligned}(\bar{\mathcal{L}}_\beta\gamma)_{ij} &= (\beta^k\underbrace{\bar\nabla_k\gamma}_{=0})_{ij} + (\bar\nabla_i\beta)^k\gamma_{kj} + (\bar\nabla_j\beta)^k\gamma_{ik} \\ &= (\bar\nabla_i\beta)_j + (\bar\nabla_j\beta)_i\end{aligned} \tag{7.2.48}$$

Plugging this back into equation (7.2.47), we conclude that

$$k_{ij} = -\frac{1}{2\alpha}\Big(\partial_t\gamma_{ij} - (\bar\nabla_i\beta)_j - (\bar\nabla_j\beta)_i\Big). \tag{7.2.49}$$

[13]Notice that $\mathcal{L}_{n^\sharp}g$ generally also has ti and tt components. However, these encode what $(\mathcal{L}_{n^\sharp}g)(X,Y)$ evaluates to for general vectors $X, Y \in \Gamma(T\mathcal{M})$, which do not concern us here—k is a bilinear form on $\Gamma(T\Sigma)$.

This characterises the extrinsic curvature's components entirely in terms of the time derivative of the induced metric, the symmetrised covariant derivative of the shift, and a scaling factor involving the lapse.

7.2.5 Example: Extrinsic Curvature of the Foliation of $\mathbb{R}^3 \setminus \{0\}$ into Spheres

In this section, we compute the extrinsic curvature of (once again) the hypersurface foliation of $\mathbb{R}^3 \setminus \{0\}$ into concentric spheres. We assume the standard metric

$$g = dr \otimes dr + r^2\, d\theta \otimes d\theta + r^2 \sin^2\theta\, d\varphi \otimes d\varphi \tag{7.2.50}$$

in spherical coordinates $x^\mu = (r, y^i) = (r, \theta, \varphi)$. Recall that in Section 5.5, we found this to be a canonical choice to get the metric into an ADM-decomposed form with zero shift and unit lapse. The leaves of the foliation are given by scaled copies of S^2,

$$\Sigma_{r_0} = r^{-1}(\{r_0\}) = r_0 S^2. \tag{7.2.51}$$

We endow the ambient manifold with the Levi-Civita connection (which has vanishing components in Cartesian coordinates). In spherical coordinates, the nonzero connection coefficients (up to symmetry in the lower indices) read

$$\begin{aligned}
\Gamma^r{}_{\theta\theta} &= -r, & \Gamma^r{}_{\varphi\varphi} &= -r\sin^2\theta, \\
\Gamma^\theta{}_{r\theta} &= \frac{1}{r}, & \Gamma^\theta{}_{\varphi\varphi} &= -\sin\theta\cos\theta, \\
\Gamma^\varphi{}_{\theta\varphi} &= \cot\theta, & \Gamma^\varphi{}_{r\varphi} &= \frac{1}{r}
\end{aligned} \tag{7.2.52}$$

Recall from Section 5.5 that the normal 1-form and vector field of this foliation are given by

$$n = dr, \quad n^\sharp = \partial_r. \tag{7.2.53}$$

According to equation (7.2.26) (with $\varepsilon = 1$) the extrinsic curvature's magnitude $k(X, Y)$ is thus the bilinear map defined by the tensor

$$\begin{aligned}
k = -\nabla n = -\nabla\, dr &= \Gamma^r{}_{\mu\nu}\, dx^\mu \otimes dx^\nu \\
&= \Gamma^r{}_{\theta\theta}\, d\theta \otimes d\theta + \Gamma^r{}_{\varphi\varphi}\, d\varphi \otimes d\varphi \\
&= -r\, d\theta \otimes d\theta - r\sin^2\theta\, d\varphi \otimes d\varphi.
\end{aligned} \tag{7.2.54}$$

We identify the nonzero components of k as

$$k_{\mu\nu} = \Gamma^r{}_{\mu\nu} \quad \Leftrightarrow \quad k_{\theta\theta} = -r, \quad k_{\varphi\varphi} = -r\sin^2\theta. \tag{7.2.55}$$

The equivalence between symmetry of the extrinsic curvature and absence of torsion is manifest here; the connection coefficients are symmetric in their lower indices if and only if the connection is torsion-free.

This result also gives us an opportunity to also verify the identity (7.2.40) which relates k to the Lie derivative of the metric along the normal flow. Since $n^\sharp = \partial_r$ has constant coefficients in spherical coordinates, the Lie derivative reduces to a regular derivative. Explicitly deriving leads to,

$$\begin{aligned}
-\frac{1}{2}\mathcal{L}_n^\sharp g = -\frac{1}{2}(\partial_r g_{\mu\nu})\, dx^\mu \otimes dx^\nu &= -\frac{1}{2}\Big[\underbrace{(\partial_r g_{\theta\theta})}_{=2r}\, d\theta \otimes d\theta + \underbrace{(\partial_r g_{\varphi\varphi})}_{=2r\sin^2\theta}\, d\varphi \otimes d\varphi\Big] \\
&= -r\, d\theta \otimes d\theta - r\sin^2\theta\, d\varphi \otimes d\varphi \\
&= k,
\end{aligned} \tag{7.2.56}$$

establishing the expected result.

8 Gauss-Codazzi Equation

Now that we have established the concepts of ambient, intrinsic and extrinsic curvature, we can finally deduce relationships between them. In this section we derive expressions for the tangential part $PR(X,Y)Z$ of the ambient curvature R acting with arguments $X, Y, Z \in \Gamma(T\Sigma)$. Specifically, this result is known as the Gauss-Codazzi equation, which we already touched upon very lightly at the end of Section 7.1.5. We first treat the case of a general foliation, and then specialise to hypersurface foliations, where terms simplify somewhat. In the following, $\mathcal{M}$ will denote a (pseudo-)Riemannian manifold equipped with a metric g, an affine connection ∇, and a foliation Σ. The leaves of the foliation are endowed with the induced metric $\gamma = \iota^* g$ as well as the induced connection $\bar{\nabla}$. Moreover, if not stated otherwise, X, Y, Z, W will denote foliation-tangent vector fields in $\Gamma(T\Sigma)$.

8.1 General Case

To derive the Gauss-Codazzi equation, let us first collect some definitions and identities we have derived in the previous sections. The starting point of the derivation will be the equation (7.1.95),

$$\bar{R}(X,Y)Z = PR(X,Y)Z + P\big((\nabla_X P)\nabla_Y Z - (\nabla_Y P)\nabla_X Z\big), \tag{8.1.1}$$

Further, we will need the definition and alternate characterisation of the extrinsic curvature,

$$K(X,Y) = Q\nabla_X Y = -(\nabla_X Q)Y. \tag{8.1.2}$$

Lastly, the decomposition of the identity into the orthogonal projectors P and Q,

$$\mathrm{id} = P + Q, \tag{8.1.3}$$

will be of use as well.

The derivation is best performed using the fully covariant Riemann tensor

$$\bar{R}(W,Z,X,Y) = g\big(\bar{R}(X,Y)Z, W\big). \tag{8.1.4}$$

We begin by inserting the precursor (8.1.1) into this, finding

$$\begin{aligned}
\bar{R}(W,Z,X,Y) &= g\big(\bar{R}(X,Y)Z, W\big) \\
&= g\Big(P\Big(R(X,Y)Z + (\nabla_X P)\nabla_Y Z - (\nabla_Y P)\nabla_X Z\Big), W\Big) \\
&= g\Big(R(X,Y)Z + (\nabla_X P)\nabla_Y Z - (\nabla_Y P)\nabla_X Z, \underbrace{PW}_{=W}\Big) \\
&= R(W,Z,X,Y) + \Big(g\Big((\nabla_X P)\nabla_Y Z, W\Big) - (X \leftrightarrow Y)\Big)
\end{aligned} \tag{8.1.5}$$

The first term is clear; it is nothing but the ambient Riemann tensor applied to the vector tuple (W,Z,X,Y). The tangential projection P drops out due to the fact that it is orthogonal and $PW = W$, as $W \in \Gamma(T\Sigma)$. The second term requires a more detailed examination. It is an anti-symmetrisation in X, Y, so it is sufficient to begin by considering

$$g\Big((\nabla_X P)\nabla_Y Z, W\Big) \tag{8.1.6}$$

and perform the anti-symmetrisation afterwards. The treatment of this expression requires a series of operations, the steps behind which we list for clarity below:

$$
\begin{aligned}
g\Big((\nabla_X P)\nabla_Y Z, W\Big) &\overset{1.}{=} -g\Big((\nabla_X Q)\nabla_Y Z, W\Big) \\
&\overset{2.}{=} -g\Big(\nabla_X \underbrace{(Q\nabla_Y Z)}_{=K(Y,Z)}, W\Big) + \underbrace{g(Q\nabla_X\nabla_Y Z, W)}_{=g(\nabla_X\nabla_Y, \underbrace{QZ}_{=0})=0} \\
&\overset{3.}{=} \underbrace{-\nabla_X g\Big(\underbrace{K(Y,Z)}_{\in\Gamma(N\Sigma)}, \underbrace{W}_{\in\Gamma(T\Sigma)}\Big)}_{=0} + g(K(Y,Z), \nabla_X W) \\
&\overset{4.}{=} g(K(Y,Z), K(X,W)).
\end{aligned}
\tag{8.1.7}
$$

The steps are as follows:

1. Replace P using $P = \mathrm{id} - Q$ and note that

$$
\nabla_X P = \underbrace{\nabla_X \mathrm{id}}_{=0} - \nabla_X Q = -\nabla_X Q \tag{8.1.8}
$$

2. Integrate the first argument by parts,

$$
(\nabla_X Q)\nabla_Y Z = \nabla_X(Q\nabla_Y Z) - Q\nabla_X\nabla_Y Z \tag{8.1.9}
$$

3. Integrate by parts using

$$
g(\nabla_X A, B) = \nabla_X g(A, B) - g(A, \nabla_X B) \tag{8.1.10}
$$

   This step assumes metric compatibility of the connection.

4. First term vanishes since $K(Y,Z) \perp W$. In the second term, the first argument lies in $\Gamma(N\Sigma)$ such that only the normal part $Q\nabla_X W = K(X,W)$ of the second argument contributes.

Inserting this partial result back into equation (8.1.5) above and carrying out the anti-symmetrisation yields the *Gauss-Codazzi equation*

$$
\bar{R}(W,Z,X,Y) = R(W,Z,X,Y) + g(K(X,W),K(Y,Z)) - g(K(Y,W),K(X,Z)). \tag{8.1.11}
$$

This is the result we anticipated in Section 7.1.5, now fully worked out in terms of the extrinsic curvature. It relates the intrinsic curvature R of the leaves of the foliation to the (pullback of) the ambient curvature R and the extrinsic curvature K.

Remarks:

- **Metric Compatibility** The derivation of the Gauss-Codazzi equation made use of the assumption that the connection is metric-compatible. This reduces its generality, but also introduces anti-symmetry in the W, Z-pair for both $\bar{R}(W,Z,X,Y)$ and $R(W,Z,X,Y)$. Consequently, the terms involving the extrinsic curvature must satisfy this anti-symmetry as well —to remain consistent. It is straightforward to see that this is the case; the exchange $(W \leftrightarrow Z)$ yields a negative sign.

- **Vanishing Torsion** Though we did not have to assume vanishing torsion in the derivation, we may add it as a further requirement. This introduces further symmetries of both the ambient and intrinsic curvature tensors, namely symmetry under the exchange of the first and second pair (since the connection is also metric-compatible),

$$
R(W,Z,X,Y) = R(X,Y,Z,W), \tag{8.1.12}
$$

  as well as the Bianchi identity

$$R(W,Z,X,Y) + R(W,X,Y,Z) + R(W,Y,Z,X) = 0. \tag{8.1.13}$$

The terms involving extrinsic curvature must satisfy these as well to remain consistent. Let us verify that this is the case for the former symmetry. Recall that for a torsion-free connection, K is symmetric, i.e.

$$K(X,Y) = K(Y,X). \tag{8.1.14}$$

This implies that

$$\begin{aligned}\bar{R}(W,Z,X,Y) - R(W,Z,X,Y) &= g(K(X,W),K(Y,Z)) - g(K(Y,W),K(X,Z)) \\ &= g(K(W,X),K(Z,Y)) - g(K(W,Y),K(Z,X)) \\ &= \bar{R}(X,Y,W,Z) - R(X,Y,W,Z),\end{aligned} \tag{8.1.15}$$

as required.

- **Curvature Operator** Our derivation of the Gauss-Codazzi equation started with its precursor (7.1.95). This is an expression for the action of the intrinsic curvature operator $\bar{R}(X,Y)$ on a vector, rather than the fully covariant curvature tensor $\bar{R}(W,Z,X,Y)$ which appears in the final Gauss-Codazzi equation (8.1.11). Naturally, the Gauss-Codazzi equation can be brought back into this form, turning it into

$$\bar{R}(X,Y)Z = PR(X,Y)Z + g(K(X,\cdot),K(Y,Z))^\sharp - g(K(Y,\cdot),K(X,Z))^\sharp. \tag{8.1.16}$$

In particular, this allows us to solve for $PR(X,Y)Z$, the foliation-tangent part of the ambient curvature, as

$$PR(X,Y)Z = \bar{R}(X,Y)Z - g(K(X,\cdot),K(Y,Z))^\sharp + g(K(Y,\cdot),K(X,Z))^\sharp. \tag{8.1.17}$$

This expresses the ambient curvature in terms of the intrinsic and extrinsic curvatures of the foliation—up to two limitations. The first one is rather obvious; we only recover the tangential part of the ambient curvature—the normal part, $QR(X,Y)Z$, does not emerge from the Gauss-Codazzi equation alone. For this, there exists another result; the *Codazzi-Mainardi equation*, which relates the normal part of the ambient curvature to derivatives of the extrinsic curvature. We will not derive it in these notes.

The second limitation is a bit more subtle. The objects on the right-hand side of equation (8.1.17) are defined for $X,Y,Z \in \Gamma(T\Sigma)$ only. Hence, we obtain no expression for $PR(X,Y)Z$ on arbitrary vector fields in $\Gamma(T\mathcal{M})$, as these generally contain a normal component.

8.2 Special Case: Hypersurface Foliation

The Gauss-Codazzi equation (8.1.11) can be specialised to hypersurface foliations, where it simplifies a bit. This is done by making use of the relationship

$$K(X,Y) = k(X,Y)n^\sharp \tag{8.2.1}$$

that we derived in Section 7.2.2. Inserting it into the Gauss-Codazzi equation yields

$$\begin{aligned}&\bar{R}(W,Z,X,Y) \\ &= R(W,Z,X,Y) + g\big(k(X,W)n^\sharp, k(Y,Z)n^\sharp\big) - g\big(k(Y,W)n^\sharp, K(X,Z)n^\sharp\big) \\ &= R(W,Z,X,Y) + \underbrace{g\big(n^\sharp,n^\sharp\big)}_{=\varepsilon}(k(X,W)k(Y,Z) - k(Y,W)k(X,Z)) \\ &= R(W,Z,X,Y) + \varepsilon(k(X,W)k(Y,Z) - k(Y,W)k(X,Z))\end{aligned} \tag{8.2.2}$$

We have thus arrived at the Gauss-Codazzi equation for hypersurface foliations,

$$\bar{R}(W,Z,X,Y) = R(W,Z,X,Y) + \varepsilon(k(X,W)k(Y,Z) - k(Y,W)k(X,Z)). \tag{8.2.3}$$

In terms of components in the adapted coordinates (t, y^i), this reads

$$\bar{R}_{k\ell ij} = R_{k\ell ij} + \varepsilon\big(k_{ik}k_{j\ell} - k_{jk}k_{i\ell}\big), \tag{8.2.4}$$

where

$$R_{k\ell ij} = (\iota^* R)_{k\ell ij} = E^\rho_k E^\sigma_\ell E^\mu_i E^\nu_j R_{\rho\sigma\mu\nu} = R\big(\partial_k, \partial_\ell, \partial_i, \partial_j\big) \tag{8.2.5}$$

are the components of the pushforward of the ambient Riemann tensor.

8.3 Example: Foliation of $\mathbb{R}^3 \setminus \{0\}$ into Spheres

Let us once again revisit the example of the foliation of $\mathbb{R} \setminus \{0\}$ into spheres. In Section 7.1.5, we computed the intrinsic curvature component

$$\bar{R}^\theta{}_{\varphi\theta\varphi} = \sin^2\theta, \tag{8.3.1}$$

to which all others are related by symmetry. Further, from Section 7.2.5 we know that the extrinsic curvature's magnitude $k(X,Y)$ has the components

$$k_{\theta\theta} = -r, \quad k_{\varphi\varphi} = -r\sin^2\theta. \tag{8.3.2}$$

Equipped with this—as well as the induced metric components

$$\gamma_{\theta\theta} = r^2, \quad \gamma_{\varphi\varphi} = r^2\sin^2\theta, \tag{8.3.3}$$

—we have all we need to explicitly check the Gauss-Codazzi equation (8.2.4) for this example.

For this, we compare the component $\bar{R}_{\theta\varphi\theta\varphi}$ of the fully covariant intrinsic curvature, which is given by

$$\bar{R}_{\theta\varphi\theta\varphi} = \gamma_{\theta\theta}\bar{R}^\theta{}_{\varphi\theta\varphi} = r^2\sin^2\theta, \tag{8.3.4}$$

to the right-hand side of the Gauss-Codazzi equation (8.2.4). Recalling that the ambient curvature is zero, we compute the right-hand side by insertion,

$$\underbrace{R_{\theta\varphi\theta\varphi}}_{=0} + \Big(k_{\theta\theta}k_{\varphi\varphi} - \underbrace{k_{\varphi\theta}}_{=0}\underbrace{k_{\varphi\theta}}_{=0}\Big) = (-r)(-r\sin^2\theta) = r^2\sin^2\theta. \tag{8.3.5}$$

Note that this matches the anticipated result (8.3.4) precisely—we have thus verified the Gauss-Codazzi equation explicitly for a specific example foliation. Although this case is geometrically simple, it provides a concrete reference point: the relations between intrinsic curvature, extrinsic curvature, and the ambient geometry is entirely transparent here. More intricate ambient spacetimes and foliations will follow the same structure, but without the luxury of vanishing ambient curvature or such straightforward coordinate expressions.

9 ADM Decomposition of the Scalar Ambient Curvature

In this final section, we apply what we've learned to derive a core result in the ADM formalism, the Hamiltonian formulation of general relativity. In this formalism, the foliation parameter of a hypersurface foliation serves as the evolution parameter for the Hamiltonian, with the induced metric (and its conjugate momentum) evolving along it. To derive the canonical equations of motion, it is necessary to perform a Legendre transform of the Einstein-Hilbert action,

$$S[g] = \int_{\mathcal{M}} d^m x \sqrt{g} \mathcal{R}. \tag{9.0.1}$$

To do this, we require a decomposition of the Ricci scalar $\mathcal{R}$ (associated to the Levi-Civita connection) in terms of the metric functions $\alpha, \beta_i, \gamma_{ij}$—these are the degrees of freedom of an ADM-decomposed metric. The goal of this section is to derive such a decomposition of the Ricci scalar, known as the ADM Lagrangian density.

To carry out this derivation, it is instructive to use foliation-adapted coordinates $x^\mu = (t, y^i)$ and to express components in the basis $E_A = (E_\perp, E_i)$ with

$$E_\perp = n^\sharp, \quad E_i = \partial_i. \tag{9.0.2}$$

which clearly separates tangential from normal components. In particular, we have

$$\operatorname{span}\{E_\perp\} = \Gamma(N\Sigma), \quad \operatorname{span}\{E_i\} = \Gamma(T\Sigma). \tag{9.0.3}$$

From Section 5.4.2 we know that the inverse metric can be written as

$$\begin{aligned} g^{-1} &= \varepsilon n^\sharp \otimes n^\sharp + \gamma^{ij} \partial_i \otimes \partial_j \\ &= \varepsilon E_\perp \otimes E_\perp + \gamma^{ij} E_i \otimes E_j. \end{aligned} \tag{9.0.4}$$

This is a block-diagonal matrix with components

$$g^{\perp\perp} = \varepsilon, \quad g^{ij} = \gamma^{ij}, \quad g^{\perp i} = 0. \tag{9.0.5}$$

Hence, the Ricci scalar may be written as

$$\begin{aligned} \mathcal{R} = g^{AC} g^{BD} R_{ABCD} &= g^{ki} g^{\ell j} R_{k\ell ij} + 2 g^{\perp\perp} g^{AB} R_{\perp A \perp B} \\ &= \gamma^{ki} \gamma^{\ell j} R_{k\ell ij} + 2 R^{\perp A}{}_{\perp A}. \end{aligned} \tag{9.0.6}$$

Let us now examine the two terms separately.

For the first term, we may refer back to the Gauss-Codazzi equation (8.2.4) we derived in the previous section. It states that the tangential components of the fully covariant ambient curvature, $R_{k\ell ij}$, are given by

$$R_{k\ell ij} = \bar{R}_{k\ell ij} - \varepsilon \big(k_{ik} k_{j\ell} - k_{jk} k_{i\ell} \big) \tag{9.0.7}$$

Thus, the first term in the expansion (9.0.6) may be rewritten as

$$\begin{aligned} \gamma^{ki} \gamma^{\ell j} R_{k\ell ij} &= \underbrace{\gamma^{ki} \gamma^{\ell j} \bar{R}_{k\ell ij}}_{=\bar{\mathcal{R}}} - \varepsilon \Big(\underbrace{\gamma^{ik} k_{ik}}_{=:k} \underbrace{\gamma^{\ell j} k_{j\ell}}_{=k} - \underbrace{k_{jk} \gamma^{ki} k_{i\ell} \gamma^{\ell j}}_{=k_{ij} k^{ij}} \Big) \\ &= \bar{\mathcal{R}} - \varepsilon \big(k^2 - k_{ij} k^{ij} \big). \end{aligned} \tag{9.0.8}$$

Here, $\bar{\mathcal{R}}$ denotes the Ricci scalar of the intrinsic curvature, k the trace of the extrinsic curvature with respect to the induced metric, and $k^{ij} = \gamma^{ik} \gamma^{j\ell} k_{k\ell}$.

We now turn our attention to the second term in (9.0.6). Unfortunately, this one requires a bit more work, as we do not have a compact result like the Gauss-Codazzi equation to lean on. We work in a mix of coordinate and E_A-basis components, with the goal of computing $R^{\perp A}{}_{\perp A} = R^{\perp\mu}{}_{\perp\mu}$. To compute this trace, we start with

$$\begin{aligned} R^\perp{}_{\lambda\mu\nu} = E^\perp_\rho R^\rho{}_{\lambda\mu\nu} &= \varepsilon n_\rho R^\rho{}_{\lambda\mu\nu} = -\varepsilon \big(R(\partial_\mu, \partial_\nu) n \big)_\lambda \\ &= -\varepsilon \big((\nabla_\mu \nabla_\nu n)_\lambda - (\nabla_\nu \nabla_\mu n)_\lambda \big) \end{aligned} \tag{9.0.9}$$

where $n = n_\mu dx^\mu$ is the normal 1-form. To derive $R^{\perp\mu}{}_{\perp\mu}$, it is useful to introduce the shorthand notation

$$\operatorname{div} X = \operatorname{tr}_1^1(\nabla X) = (\nabla_\mu X)^\mu \tag{9.0.10}$$

for the covariant divergence of a vector field $X \in \Gamma(T\mathcal{M})$. With this definition, we may now continue deriving:

$$\begin{aligned}
-\varepsilon R^{\perp\nu}{}_{\perp\nu} &= -\varepsilon E^\mu_\perp R^{\perp\nu}{}_{\mu\nu} = -\varepsilon n^\mu R^{\perp\nu}{}_{\mu\nu} \\
&= n^\mu(\nabla_\mu\nabla_\nu n^\sharp)^\nu - n^\mu(\nabla_\nu\nabla_\mu n^\sharp)^\nu \\
&= \underbrace{\nabla_\mu\big(n^\sharp \otimes \underbrace{(\nabla_\nu n^\sharp)^\nu}_{=\operatorname{div} n^\sharp}\big)^\mu}_{=\operatorname{div}(n^\sharp\cdot\operatorname{div} n^\sharp)} - (\nabla_\mu n^\sharp)^\mu(\nabla_\nu n^\sharp)^\nu - \underbrace{\nabla_\nu\big(\underbrace{n^\mu\nabla_\mu n^\sharp}_{=\nabla_{n^\sharp}n^\sharp}\big)^\nu}_{=\operatorname{div}(\nabla_{n^\sharp}n^\sharp)} + (\nabla_\nu n^\sharp)^\mu(\nabla_\mu n^\sharp)^\nu \\
&= \operatorname{div}\big(n^\sharp\cdot\operatorname{div} n^\sharp - \nabla_{n^\sharp}n^\sharp\big) - (\nabla_\mu n^\sharp)^\mu(\nabla_\nu n^\sharp)^\nu + (\nabla_\nu n^\sharp)^\mu(\nabla_\mu n^\sharp)^\nu
\end{aligned} \tag{9.0.11}$$

The first term is a total derivative/divergence, and will hence only contribute to the action as a boundary term.

The other two terms are more interesting; we recall that for hypersurface foliations, the extrinsic curvature may be written as

$$k(X,Y) = \varepsilon g(Y, -\nabla_X n^\sharp). \tag{9.0.12}$$

In components, this reads **Comment:** Correct normal components claim, not entirely correct or at least needs justification

$$k_{\mu\nu} = \varepsilon g(\partial_\nu, -\nabla_\mu n^\sharp) = -\varepsilon(\nabla_\mu n^\sharp)_\nu \quad\Leftrightarrow\quad (\nabla_\mu n^\sharp)_\nu = -\varepsilon k_{\mu\nu}. \tag{9.0.13}$$

This is precisely the tensor appearing in the remaining terms of equation (9.0.11). Thus, we find

$$\begin{aligned}
-(\nabla_\mu n^\sharp)^\mu(\nabla_\nu n^\sharp) + (\nabla_\nu n^\sharp)^\mu(\nabla_\mu n^\sharp)^\nu &= -(-\varepsilon)^2 k_\mu{}^\mu k_\nu{}^\nu + (-\varepsilon)^2 k_{\mu\nu}k^{\mu\nu} \\
&= -k^2 + k_{ij}k^{ij},
\end{aligned} \tag{9.0.14}$$

where we made use of the fact that $k_{\mu\nu}$ has no normal components and that $k = k^\mu{}_\mu = k^i{}_i$. Putting everything together, we find

$$R^{\perp\nu}{}_{\perp\nu} = \varepsilon\operatorname{div}\big(\nabla_{n^\sharp}n^\sharp - n^\sharp\cdot\operatorname{div} n^\sharp\big) + \varepsilon\big(k^2 - k_{ij}k^{ij}\big). \tag{9.0.15}$$

Hence, the ambient Ricci scalar is given by

$$\begin{aligned}
\mathcal{R} &= \bar{\mathcal{R}} - \varepsilon\big(k^2 - k_{ij}k^{ij}\big) + 2\varepsilon\operatorname{div}\big(\nabla_{n^\sharp}n^\sharp - n^\sharp\cdot\operatorname{div} n^\sharp\big) + 2\varepsilon\big(k^2 - k_{ij}k^{ij}\big) \\
&= \bar{\mathcal{R}} + \varepsilon\big(k^2 - k_{ij}k^{ij}\big) + 2\varepsilon\operatorname{div}\big(\nabla_{n^\sharp}n^\sharp - n^\sharp\cdot\operatorname{div} n^\sharp\big)
\end{aligned} \tag{9.0.16}$$

This—together with $\sqrt{g} = \alpha\sqrt{\gamma}$—allows us to write down the full ADM action,

$$S_{\text{ADM}}[\alpha,\beta,\gamma] = \int_{\mathcal{M}} d^m x\,\alpha\sqrt{\gamma}\Big[\bar{\mathcal{R}} + \varepsilon\big(k^2 - k_{ij}k^{ij}\big)\Big] + (\text{boundary terms}). \tag{9.0.17}$$

This is a significant result—given a foliation, it splits the Einstein-Hilbert action into purely foliation-intrinsic parts (the intrinsic Ricci scalar $\bar{\mathcal{R}}$, associated to the Levi-Civita connection of the induced metric γ) and extrinsic parts (the extrinsic curvature $k(X,Y)$). In particular, this makes the dependenc of the action on α,β,γ manifest. The intrinsic Ricci scalar only depends on γ, whereas the dependence of k on α,β,γ is provided by the compact result (7.2.49).

The ADM action is the starting point of the ADM formalism. At this point, the next step would be to compute the canonical momenta associated to the degrees of freedom of the metric, i.e. the lapse α, the shift components β_i and the induced metric components γ_{ij}.