Semi-Riemannian Geometry and Rolling

A mathematics bachelor thesis

Juri Sampieri Bjornsson

Studentnumber: 6193358

Written under supervision of

Dr. Alvaro del Pino Gomez

Universiteit Utrecht

January 20, 2021

1 Introduction

1.1 Some Background

Modern **twistor theory** finds its origins in a 1967 paper titled *twistor algebra* written by Roger Penrose, the 2020 nobel prize winner for physics. His original goal with this theory was to unify Einstein's theory of general relativity with quantum mechanics. This theory would have replaced space-time with another more fundamental space called the twistor space. The very basic idea behind this replacement is as follows: In space-time we have three space dimensions and one time dimension. We view this space-time as consisting of an infinite amount on infinitesimal points also called events. Penrose argued that we must not focus on events, but instead study the space of all the possible paths in space-time that light can travel in. To put it in more mathematical terms, instead of looking at the four-dimensional Euclidean space $\mathbb{R}^4$ which models space-time, we must actually focus on the projective 3-space $\mathbb{CP}^3$ which is also refferd to as twistor space. There are reasons why we look at the complex projective 3-space and not at the real one, but explaining that goes beyond the scope of this introduction.

A lot of effort was put into make this a functioning theory for **quantum gravity**, but attempts stranded during the eighties due to technical and philosphical considerations. After a period of slumber, Witten, in a 2003 paper titled *perturbative gauge theory as a string theory in twistor space*, has awoken again interest in twistor theory. Even today this remains still an active field of research with for example a 2020 paper by Woit in which he argues that twistor theory is perfectly compatible with the **Standard Model** for particle physics.

But aside from quantum gravity, twistor theory has found other usefull applications. In physics for example, thanks again to the above paper by Witten, twistor theory has impacted the world of **scattering amplitudes** which, to put it briefly, are the outputs that determine the probability in scattering processes within quantum field theory.

More importantly for this thesis though is the fact that in mathematics too twistor theory has had a lot impact. The first application we discuss is to **integrable systems**. To be brief, an equation is integrable if we can construct an explicit solution and this solution is well-behaved. This is of course not a mathematical precise statement, but offering one is not the goal of this introduction. It appears that integrable systems, at least in low-dimensional cases, appear as a reduction of a set of equations called the anit-self-dual Yang-Mills equations. By reduction we mean that we reduce the number of independent and dependent variables by imposing certain constraints on the Yang-Mills equations. As an example of Yang-Mills equations one can think of the Euler-Lagrange equations in physics. This is the point where twistor theory comes in. It appears that the solutions to the Yang-Mills equations have a correspondence within the realm of twistor theory. So studying twistor theory might give us more insight in the nature of integrable systems. For a more complete description on integrable systems and twistor theory, we refer to [12].

Another application of twistor theory is to **semi-Riemannian geometry** which we will cover chapter 2 of this thesis. Since the essence of twistor theory is to encode differential geometric information on a manifold to another auxiliary space, called the twistor space, one would hope that certain problems facing semi-Riemannian geometry might become solvable. For a more detailed discussion on this topic, we refer to [4].

For more detailed discussion we refer to [8] which have written a paper for the 50-years anniversary of twistor theory. Therein they discuss more applications, but also successess and problems of twistor theory. Another more dated but not less relevant book is [2]. There the authors have collected some of the more important papers on twistor theory which offers a good overview.

The last question which still needs answering is: how does this thesis fit in the context of twistor theory? The answer is that a lot of attention in twistor theory has been given to reformulating conventional physics within the framework of twistor space. This thesis aims to do precisely that. The main result of this thesis is in fact that we can describre rolling without slipping or twisting in terms of semi-Riemannian geometry. The methods we use for this result in fact appear also in more complicated reformulations of conventional physics to twistor theory. So in this sense this thesis can be viewed in the context of twistor theory.

1.2 Brief Overview of this Thesis

In chapter 2 we will start this thesis of by laying the foundations of semi-Riemannian geometry. We will discuss what a semi-Riemannian manifold is and how we define length. Next we will move on to define linear connections. These objects will serve us to generelize straight lines in Euclidean space to straight lines, called geodesics, in our manifolds. The linear connection also allows us to transport vectors in a natural way along a curve. This is what is called parallel transport. Next we will introduce a linear connection with some nice properties called the Levi-Civita connection. To finish that chapter we introduce a new set of coordinates called normal coordinates and take a look at curvature which is a way to classify semi-Riemannian manifolds.

Having done this, we are in a position to take a look at rolling without slipping or twisting. In chapter 3, we define a configuration space which one can view as the analogue of twistor space and is the natural space to describe the rolling problem in. Next we will take a look at the first restraint of our rolling problem, namely no slipping. In essence what this restraint tells us is that we are allowed only two types of movement: 1) we can reorient ourselves or 2) we can keep on moving in the same direction. Next we move on to the no twisting condition. This condition models a different type of question. After rolling starts in the physical world, we see that it defines a unique motion. To put it differently, objects do not suddenly reorient themselves and thus the no twisting condition boils down to Newton's first law.

In the last chapter, we will show that we can recover any vector on our semi-Riemannian manifold from the no slipping condition. To put it more intuitively, two objects rolling on each other in a no slipping kind of way are able just by rolling and turning to reach any potential contact configuration. This underlies a broader statement regarding the way in which we defined the no slipping condition and the curvature of a manifold.

We have also written two appendices. The first one discusses distributions. We discuss there some background knowledge needed to understand the way in which we model the no slipping and no twisting condition. The second one discusses only a short lemma regarding vector spaces and their tagent spaces. Throughout the thesis we will sometimes refer to lemma's there proven.

Contents

2 Fundamentals of Semi-Riemannian Geometry

Before laying the foundations for semi-Riemannian geometry, let us explain why we need it to model rolling without slippng or twisting. The reason is fairly simple: if we have two objects rolling on each other, the rolling speed of object one with respect to object two should be related to the rolling speed of object two with respect to object one. The way to relate them is if we can relate distances on object one with distances on object two, but this implies a notion of length defined on the two objects. This is precisely what semi-Riemannian geometry studies: smooth manifolds on which we have a notion of length.

In this chapter we will start off by introducing tensors and tensor bundles, which will serve to define the cornerstone of any semi-Riemannian manifold: the metric. Thanks to the metric we are able to define infinitesimal distances. In semi-Riemannian geometry, contrary to Riemannian geometry, we allow the metric to give us imaginary distances.

We then move on to the linear connection. This object will allow us to take second order derivatives. More importantly though, since we are not always able to find a global frame for a smooth manifold, introducing the linear connection allows us to talk about a vector field 'pointing in the same direction' at every point along a curve. This is what we call a parallel vector field. Conversely, given a vector at some point on a curve, we can generate from it a vector field parallel to that curve. This is what we call parallel translation and will play a central role when discussing the rolling without slipping or twisting.

Next we move on to geodesics. These are a special class of curves which minimise distance between points on a Riemannian manifold. Yet, we will view geodesics as curves which have zero acceleration as this is easier to define in the context of semi-Riemannian geometry. In [7] chapter 6, it will be shown that these two definitions of geodesic coincide in Riemannian geometry.

Thereafter we introduce the Levi-Civita connection which, as we will show, is the natural choice of linear connection on a semi-Riemannian manifold. We will also prove our second important theorem, namely the existence and uniqueness of the Levi-Civita connection.

We introduce normal coordinates and the curvature tensor. Understanding them is not vital to understand the rolling problem, but they play a significant role in any study in Riemannian geometry and we found it worthwhile to at least mention them in this thesis.

To finish this chapter off, we take a brief look at an example of a space in non-Euclidean geometry, namely the sphere. As main result in this example, we will show that any geodesic is a great circle and vice versa.

This chapter is based on [7] chapter 3 to 6 and on [9] chapter 3 to 5.

2.1 Tensors

2.1.1 Tensors

Definition 2.1. *Let V be a vector space and V^* be the dual of V. A* ***tensor*** *is then a multilinear map:*

$$F : \underbrace{V \times \cdots \times V}_{k \text{ times}} \times \underbrace{V^* \times \cdots \times V^*}_{l \text{ times}} \to \mathbb{R}$$

We call this a k-covariant, l-contravariant tensor, or abbreviated we write $\binom{k}{l}$-tensor. The ***space of all $\binom{k}{l}$-tensors*** *of a certain vector field V will be denoted by $T^k_l(V)$.*

Remark. Suppose $F \in T^k_l(V)$ and given a basis (E_i) for V and (ϕ^j) for V^*. We note that we can write our tensor then as:

$$F = F^{i_1,\dots,i_k}_{j_1,\dots,j_l} E_{i_1} \otimes \cdots \otimes E_{i_k} \otimes \phi^{j_1} \otimes \cdots \otimes \phi^{j_l} \tag{1}$$

Where $F^{i_1,\dots,i_k}_{j_1,\dots,j_l} = F(\phi^{j_1}, \dots, \phi^{j_l}, E_{i_1}, \dots, E_{i_k}) \in \mathbb{R}$. △

There is a characteristic property regarding tensors which allows us to view them as maps into the vector space instead of as map into the real numbers.

Lemma 2.2. *Suppose V a vector space. There exists an isomorphism of vector spaces between $T^k_{l+1}(V)$ and the vector space:*

$$C = \{F : \underbrace{V \times \cdots \times V}_{k\ times} \times \underbrace{V^* \times \cdots \times V^*}_{l\ times} \to V | F \text{ multilinear}\}$$

Proof. We first prove the case $k = l+1 = 1$. In that case our set C is just the set of endomorphisms of V which we will denote by $End(V)$. We define the following map:

$$\Phi : End(V) \to T^1_1(V)$$

where for a given $X \in V$ and $\omega \in V^*$ the function satisfies $\Phi(A)(X, \omega) = \omega(AX)$. Since ω is linear, we have that Φ is linear. Now to show injectivity we suppose $\Phi(A) = \Phi(B)$. This must hold true for any $X \in V$ and $\omega \in V^*$. Hence we have that $\Phi(A)(X, \omega) = \Phi(B)(X, \omega)$. We can write this as $\omega((A - B)X) = 0$. Since this holds true for any X and ω, we must have $A - B = 0$ thus proving our map is injective.
To prove surjectivity we suppose a map $F \in T^1_1(V)$. Suppose we have a basis (E_i) for V and (ϕ^j) for V^*, we write $F = F^i_j E_i \otimes \phi^j$. Now we take A such that its components satisfy $A^i_j = F^i_j$. Now we can write $X = X^i E_i$ and $\omega = \omega_k \phi^k$. Now we look at:

$$\Phi(A)(X, \omega) = \omega(AX) = \omega_k \phi^k(F^i_j X^j E_i) = F^i_j X^j \omega_k \phi^k(E_i) = F^i_j X^j \omega_i$$

Now we note that $F(\omega, X) = F^i_j X^j \omega_i$.
Thus this gives us an isomorphism of vector spaces.
The general case goes along this path. Given a $F \in T^k_{l+1}(V)$. Suppose $(X_i)^k_{i=1}$ elements in V and $(\omega^j)^l_{j=1}$ elements in V^*. This now gives us $F(X_1, \ldots, X_{k-1}, \cdot, \omega^1, \ldots, \omega^l, \cdot) \in T^1_1(V)$. By the above, we can find an $A \in End(V)$ which stands in bijection with $F(X_1, \ldots, X_{k-1}, \cdot, \omega^1, \ldots, \omega^l, \cdot)$. Thus the map $X_k \to A(X_k)$ gives us a isomorphism between $T^k_{l+1}(V)$ and C.

□

We are able to pullback $\binom{k}{0}$-tensors by linear maps.

Definition 2.3. *Suppose V and W are vector spaces and $\Phi : V \to W$ is a linear map. Suppose furthermore k-elements $(v_i)^k_{i=1} \in V$ and we have a tensor $F \in T^k_0(W)$. Then we define the* ***pullback of a tensor*** *by Φ as:*

$$(\Phi^* F)(v_1, \ldots, v_k) := F(\Phi(v_1), \ldots, \Phi(v_k))$$

2.1.2 Tensor Bundles

Definition 2.4. *A $\binom{k}{l}$**-tensor bundle** over a smooth manifold M is a vector bundle whose fibres at some point $p \in M$ are given by $T^k_l(T_pM)$. We denote the $\binom{k}{l}$-tensor bundle over M as $T^k_l M$.*

We will now show that a tensor bundle is a vector bundle.

Lemma 2.5. *The tensor bundle $T^k_l M$ is indeed a vector bundle .*

Proof. Suppose $\pi : T^k_l M \to M$ the projection which sends an element $F \in T^k_l(T_pM)$ to $p \in M$. It should be clear that $\pi^{-1}(p) = T^k_l(T_pM)$ is a vector space. So we are left to construct a smooth local trivialisation for $T^k_l M$. Let U be an open subset of M containing p. Suppose furthermore that we have a coordinate chart $(x^i)^n_{i=1}$ on U. We know that we can write any $F \in T^k_l(T_qM)$ with $q \in U$ as:

$$F = F^{i_1,\ldots,i_k}_{j_1,\ldots,j_l} \partial_{i_1} \otimes \cdots \otimes \partial_{i_k} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_l}$$

Now we take our local trivialisation $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^{n^{k+l}}$ to be the map which sends an element $F \in T^k_l(T_qM)$ to $((x^i), (F^{i_1,\ldots,i_k}_{j_1,\ldots,j_l}))$. This map is smooth and it is easy to check that it satisfies all the demands for the local trivialisation. □

Now since we see that tensor bundles are vector bundles, it is possible to talk about tensor fields.

Definition 2.6. *A* $\binom{k}{l}$***-tensor field*** *is a smooth section of the tensor bundle.*

Remark. A tensor field F can locally be written in the form of (1), but now the components $F^{i_1,\ldots,i_k}_{j_1,\ldots,j_l}$ should be understood as elements of $C^\infty(M)$. △

There is an easy way to determine whether a map is a tensor field.

Lemma 2.7. *A map:*

$$f : \mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M) \times \mathfrak{X}^*(M) \times \cdots \times \mathfrak{X}^*(M) \to C^\infty(M)$$

is induced by a $\binom{k}{l}$*-tensor field if and only if it is multilinear over* $C^\infty(M)$*.*

Proof. Suppose $(X_i)_{i=1}^k$ elements in $\mathfrak{X}(M)$ and $(\omega^j)_{j=1}^l$ elements in $\mathfrak{X}^*(M)$. First we prove that any tensor field is multilinear over $C^\infty(M)$. Let α and β be either vector fields or one forms and let $g, h \in C^\infty(M)$. Lastly we suppose $p \in M$. Let us now look at:

$$\begin{aligned} f(X_1, \ldots, g\alpha + h\beta, \ldots, \omega^l)(p) :=& f_p(X_1|_p, \ldots, (g\alpha + h\beta)|_p, \ldots, \omega^l|_p) \\ =& f_p(X_1|_p, \ldots, g(p)\alpha|_p + h(p)\beta|_p, \ldots, \omega^l|_p) \end{aligned}$$

Where we have that $f_p : (T_pM)^k \times (T_p^*M)^l \to \mathbb{R}$. We see now that locally our tensor field reduces to a regular tensor. We know that a tensor is multilinear over the real numbers. So we get:

$$f(X_1, \ldots, g\alpha + h\beta, \ldots, \omega^l)(p) = g(p)f_p(X_1|_p, \ldots, \alpha|_p, \ldots, \omega^l|_p) + h(p)f_p(X_1|_p, \ldots, \beta|_p, \ldots, \omega^l|_p)$$

Since this holds true for any $p \in M$ we have that our tensor field is multilinear over $C^\infty(M)$.
Let us now show the converse. Suppose f as above and multilinear over $C^\infty(M)$. From the fact that f is multilinear over $C^\infty(M)$, we have that f_p is multilinear over $\mathbb{R}$ and hence f_p is a $\binom{k}{l}$-tensor. Now the map $p \to f_p$ gives us a section of $\pi : T^k_l M \to M$. This section is smooth since f is a map into $C^\infty(M)$ and thus varies smoothly when we vary $p \in M$. By definition our map f is now a $\binom{k}{l}$-tensor field. □

In the same way we are able to pullback $\binom{k}{0}$-tensors, we are able to pullback $\binom{k}{0}$-tensors fields.

Definition 2.8. *Suppose we have two smooth manifolds M and N and a smooth map $\phi : M \to N$. Suppose furthermore we have $f \in T^k_0 N$ and we have k-vector fields $(X_i)_{i=1}^k \in \mathfrak{X}(M)$. We define* ***the pullback of a tensor field*** *f by ϕ as:*

$$(\phi^* f)(X_1, \ldots, X_k) = f(d\phi(X_1), \ldots, d\phi(X_k))$$

Where we have that:

$$f(d\phi(X_1), \ldots, d\phi(X_k))(p) = f(d_p\phi(X_{1p}), \ldots, d_p\phi(X_{kp}))$$

for some $p \in M$.

2.2 The Metric

2.2.1 Semi-Riemannian manifolds

The metric plays a central role in semi-Riemannian geometry. It allows us to measure infinitesimal lengths and on our manifold. The metric can be thought of to be a generalization of the inner product on a vector space. Though, as staid before, in semi-Riemannian geometry our metric need not be positive definite. So our metric should have the same properties as the inner product. This gives rise to our definition.

Definition 2.9. *A* ***metric*** *g is a* $\binom{2}{0}$*-tensor field satifying:*

1. *symmetry: for every two vector field $X, Y \in \mathfrak{X}(M)$, we have $g(X, Y) = g(Y, X)$.*
2. *nondegeneracy: for every vector $v \in T_pM$ in the tangent space at point p in a smooth manifold M, there exists another vector w such that $g(v, w) \neq 0$ holds true.*

Remark. Given any coordinate system (x^i) on an open subset U of our semi-Riemannian manifold, we can write the metric as:

$$g = g_{ij} dx^i dx^j$$

Where $g_{ij} = g(\partial_i, \partial_j)$ and $dx^i dx^j = \frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i)$. △

Now the definition of a semi-Riemannian manifolds follows readily.

Definition 2.10. *A* ***semi-Riemannian manifold*** *is a smooth manifold M endowed with a metric g*

Definition 2.11. *In our definition of the metric, if we replace the demand of nondegeneracy by a more stricter demand of being positive definite, then we would call our semi-Riemannian manifold just a* ***Riemannian manifold****.*

We can also take a look at submanifold in semi-Riemannian sense.

Definition 2.12. *Suppose we have two semi-Riemannian manifolds (M, g_M) and (N, g_N) of dimension m and n where $n \leq m$. If $\iota : N \to M$ is an injective immersion which satisfies $g_N = \iota^* g_M$, then we say that N is a* ***semi-Riemannian submanifold*** *of the ambient manifold M. The metric g_N is then called* ***the induced metric****.*

Remark. If g_N is positive definite, then we say that N is a **Riemannian submanifold**. △

2.2.2 Important Operations on a Semi-Riemannian manifold

Inner products allow us to define the length of a vector in a vector space. Since our metric is in some sense the generelization of the inner product, we can use it to define lengths of vectors in tangent spaces of our semi-Riemannian manifold.

Definition 2.13. *The* ***length of a vector*** *v is defined as:*

$$|v| = (g(v, v))^{1/2}$$

We define a vector to be ***spacelike*** *if $|v|^2 > 0$. We define it* ***null*** *if $|v|^2 = 0$. We define it to be* ***timelike*** *if $|v|^2 < 0$.*

Given some basisvectors for our vector space, it will be usefull to keep track which of them are spacelike, timelike or null. To this end we introduce the metric signature.

Definition 2.14. *The* ***metric signature*** *(p, q, r) of a vector space are three numbers, where p are the quantity of basisvectors with positive length, q the quantity with negative length and r the quantity of null length. If we want to specify that our metric has this signature, we will write it as $g^{(p,q,r)}$.*

Remark (1)*.* Sometimes we will say that our metric has signature (p, q). In that case, we implicitely assume that there are zero basisvectors with length null. Analogously, we then write $g^{(p,q)}$. △

Remark (2)*.* The metric signature does not vary from point to point due to the fact that the metric is non-degenerate. △

An important class of semi-Riemannian manifolds which play a central role in special relativity are Lorentzian manifolds.

Definition 2.15. *A* ***Lorentzian manifold*** *is a semi-Riemannian manifold whose metric signature at the tangent space at every point is $(n-1, 1, 0)$.*

In the coming chapters, we will also need to know what it means for two vectors to be orthogonal.

Definition 2.16. *We say that two vectors v and w are* ***orthogonal*** *if $g(v, w) = 0$.*

Remark. In the case of a Riemannian manifold (M, g), we are able to define angles between vectors in the following way:

$$\cos(\theta) = \frac{g(v, w)}{|v||w|}$$

where we have that $v, w \in T_pM$ for some $p \in M$. Since two vectors are orthogonal when they form an angle of $\frac{\pi}{2}$, it follows that they are orthogonal when $g(v, w) = 0$.

Next let us take a look at the mapping between two semi-Riemanian manifolds:

Definition 2.17. *Let (M, g_M) and (N, g_N) be two semi-Riemannian manifolds. A mapping $\psi : M \to N$ is called an* ***isometry*** *if it is a diffeomorphism and pulls the metric on N back to the metric on M, or more formally: $\psi^*(g_N) = g_M$.*

Remark. Again we can generalize isometries further to homothety of coefficient c. A homothety is a diffeomorphism between two semi-Riemannian manifolds, but it satisfies $\psi^*(g_N) = cg_M$ where $c \in \mathbb{R} - \{0\}$ is constant. An isometry is then just a homothety of coefficient 1. △

2.2.3 Flat Map

Definition 2.18. *The* ***flat map*** *is defined as:*

$$.^\flat : TM \to T^*M : X \to X^\flat := g(X, \cdot)$$

The flat map plays a role in the raising and lowering of indices. This operation has no real significant mathematical consequences. It is just a handy way to write stuff.

Definition 2.19. *Suppose we have a local frame $(E_i)_i$ on some open U and its dual $(\phi^i)_i$. For any vector field X, we can write its flat map as:*

$$X^\flat = g(X, \cdot) = g_{ij} X^i \phi^j$$

The ***lowering of an index*** *is then defined as $X_j = g_{ij}X^i$. Analogously the* ***raising of an index*** *is defined as $X^i = g^{ij}X_j$, where g^{ij} satisfies the property $g^{ij}g_{jk} = \delta^i_k$ and δ^i_k is the kronecker delta.*

What we want to show in this subsection is that the flat map induces a $C^\infty(M)$-linear isomorphism from $\mathfrak{X}(M)$ to $\mathfrak{X}^*(M)$. This though underlies a broader mathematical truth. So we will prove it in terms of this broader case and then return as a corollary to this property of the flat map.

Definition 2.20. *Let M_1 and M_2 be two smooth manifolds and let E_1 and E_2 be vector bundles over M_1 and M_2 respectively. Let $\pi_i : E_i \to M_i$ be the projection where $i = 1, 2$. We define a map $F : E_1 \to E_2$ to be a* ***smooth bundle homomorphism*** *if:*

1. *F is smooth;*
2. *there exists a smooth map $f : M_1 \to M_2$ satisfying $\pi_2 \circ F = f \circ \pi_1$;*
3. *at every point $p \in M_1$ we have that $F_p : (E_1)_p \to (E_2)_{f(p)}$ is a $\mathbb{R}$-linear map.*

We say that F is a ***smooth bundle isomorphism*** *if F is a diffeomorphism whose inverse is also a smooth bundle homomorphism.*

Lemma 2.21. *If $F : E_1 \to E_2$ is a smooth bundle homomorphism, then there exists a $C^\infty(M_1)$-linear map $\tilde{F} : \Gamma(E_1) \to \Gamma(E_2)$, where $\Gamma(E_i)$ is the vector space of smooth sections between M_i and E_i.*

Proof. Let $X, Y \in \Gamma(E_1)$ and $f, g \in C^\infty(M_1)$. We define the map $\tilde{F}$ as $\tilde{F}(X) = F \circ X$. To prove that this map is $C^\infty(M_1)$-linear, we work locally. Let $p \in M_1$ be a point. Since the map F_p is $\mathbb{R}$-linear, we get that:

$$F(f(p)X_p + g(p)Y_p) = f(p)F(X_p) + g(p)F(Y_p)$$

Since this holds true for any $p \in M$, it follows that $\tilde{F}$ is $C^\infty(M_1)$-linear. □

As a corllary, we can now also prove the following:

Corollary 2.22. *If $F : E_1 \to E_2$ is a smooth bundle isomorphism, then $\tilde{F} : \Gamma(E_1) \to \Gamma(E_2)$ is a bijective $C^\infty(M_1)$-linear map whose inverse is $C^\infty(M_2)$-linear.*

Proof. Since F is a diffeomorphism, we just define the inverse of $\tilde{F}$ as $\tilde{F}^{-1}(X) = F^{-1} \circ X$, where we have that $X \in \Gamma(E_2)$. Since F^{-1} is also a smooth bundle homomorphism, it follows by the previous lemma that $\tilde{F}^{-1}$ is a $C^\infty(M_2)$-linear map. □

So now we only have to prove that the flat map is a smooth bundle isomorphism to prove the desired result

Lemma 2.23. *Let (M, g) be a semi-Riemannian manifold. The flat map is a smooth bundle isomorphism.*

Proof. Let $\pi : TM \to M$ and $\pi^* : T^*M \to M$ be the projections. It should be clear that the identity map $id : M \to M$ is the smooth map which satisfies that $id \circ \pi = \pi^* \circ .^\flat$. Next it should also be clear that the flat map is $\mathbb{R}$-linear map between T_pM and T_p^*M. This is so because the metric is a linear map.
So what we do now is to show that the flat map is a diffeomorphism between T_pM and T_p^*M. The flat map is smooth because the metric is smooth. Next it is injective, because the metric is linear and non-degenerate. Since we know that the dimension of T_pM and T_p^*M is the same, this must be a bijection. The inverse of the flat map is the **sharp map**:

$$.^\sharp : T^*M \to TM : \omega \to \omega^\sharp$$

where $\omega^\sharp$ is locally given by $\omega^\sharp = \omega^i E_i = g^{ij}\omega_j E_i$, where E_i is a local coordinate frame for some open $U \subseteq M$. Now since $g^{ij}\omega_j$ is smooth for any i and j, we have that the sharp map is smooth and thus the flat map is a diffeomorphism. Since this holds true at any point $p \in M$, it follows also that it is a diffeomorphism between TM and T^*M.
The prove that the sharp map is a vector bundle homomorphism goes along the same way □

Lemma 2.24. *Let M be a semi-Riemannian manifold. The flat map induces a $C^\infty(M)$-linear isomorphism from $\mathfrak{X}(M)$ to $\mathfrak{X}^*(M)$.*

Proof. Just apply lemma 2.23 and corollary 2.22. □

So what this lemma tells us, is that we are free to move from a vector field V to a one-form V^* without losing information.

2.2.4 Euclidean Space

A very important metric is the one in standard Euclidean space.

Definition 2.25. *The* ***Euclidean metric*** *on $\mathbb{R}^n$ is defined as:*

$$\bar{g} = \delta_{ij} dx^i \otimes dx^j$$

Where δ_{ij} is the Kronecker delta.

It is now worthwhile to take a look at how to compute the metric of a given immersed submanifold of the standard Euclidean space.

Example 2.26. Let $U \subseteq \mathbb{R}^n$ be an open subset and let $f : U \to \mathbb{R}^m$ be a parametrization of a submanifold $M \subseteq \mathbb{R}^m$. Suppose furthermore that we are given a metric on U and we want to know how this metric looks like in M. Or to put it differently what is the induced metric on M? If we are given coordinates (x^i) on U, then the metric on M looks like

$$g = \sum_{i=1}^{m} (df^i)^2 = \sum_{i=1}^{m} (\sum_{j=1}^{n} \frac{\partial f^i}{\partial x^j} dx^j)^2$$

As we will see when discussing the sphere, this will be the property that allows us to compute the standard metric on the sphere. ♢

2.3 Linear Connections and The Christoffel symbols

2.3.1 Linear Connection

In these following two subsections we will dedicate ourselves to defining geodesics. The standard way of defining geodesics is by defining them to be the length minimising curves between two points on a manifold. While this definition is intuitively the clearer one, this will not be our approach. The reason for this is that in a Riemannian manifold length minimising is easy to define due to the positive definiteness of the metric, but in a semi-Riemannian manifold we do no longer have that property. Our approach instead will be focused on another aspect of geodesics which holds true in both Riemannian as well as semi-Riemannian manifolds. To this end, let us look at an example.

Example 2.27. Let us look at a n-dimensional Euclidean space $\mathbb{R}^n$ endowed with the Euclidean metric $\bar{g}$. We know that the shortest line connecting any two points $p, q \in \mathbb{R}^n$ is the straight line defined as:

$$y(t) = (q - p)t + p$$

Where we have $t \in [0, 1]$. As we will show further up ahead, these are in fact the geodesics in $\mathbb{R}^n$. Applying basic calculus, we get that the second derivative of y with respect to t equals $\frac{d^2y}{dt^2} = 0$. So the acceleration of the curve is zero. ♢

The fact that the acceleration is zero, will be the property which allows us to generelize the straight line in Euclidean space to geodesics in semi-Riemannian manifolds. But this brings us to our first problem, namely the question: what is the acceleration of a curve in a semi-Riemannian manifold? There are several equivalent ways of defing the velocity of a curve and it is a quite straightforward process (See [5] chapter 4). The problem arises when having to define the acceleration. Let $\gamma : I \to M$ be a curve on the manifold M. Our physical intuition would tell us to look at the velocity $\dot{\gamma}$ at a time t_1, subtract it from the velocity at time t_2 and then divide this by $t_2 - t_1$. This gives us the average acceleration between these two times, but the problem lays in the fact of subtracting the velocities. The time derivative of our curve at time t_1 does not live on our manifold M, but rather in the tangent space $T_{\gamma(t_1)}M$ at point $\gamma(t_1)$. Analogously the time derivative at t_2 lives in $T_{\gamma(t_2)}M$. So we are subtracting two objects which exist in totally seperate worlds. To be able to define the acceleration, we define the linear connection:

Definition 2.28. *The **linear connection** $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ is a map satisfying the following conditions:*

1. *It is linear over the space $C^\infty(M)$ of smooth real-valued functions over M over the first term; $\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z$ with $X, Y, Z \in \mathfrak{X}(M)$ and $f, g \in C^\infty(M)$.*
2. *It is $\mathbb{R}$-linear over the second term; $\nabla_X(aY + bZ) = a\nabla_X Y + b\nabla_X Z$ with $X, Y, Z \in \mathfrak{X}(M)$ and $a, b \in \mathbb{R}$*
3. *it satisfies the Leibniz rule: $\nabla_X(fY) = f\nabla_X Y + X(f)Y$ where $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$. Xf stands for the Lie derivative of f over X.*

A property of the linear connection is the following:

Lemma 2.29. *Given a point $p \in M$, $(\nabla_X Y)_p$ depends only on the values of Y in an neighbourhood U of p and on X only at the point p*

Proof. We start by proving that $(\nabla_X Y)_p$ depend on values of Y in U. Suppose Y and $\tilde{Y}$ two vector fields such that in U we have $Y = \tilde{Y}$. By linearity of the linear connection, we only have to show that $\nabla_X(Y - \tilde{Y})(q) = 0$ for every $q \in U$. For simplicity, we define $Z = Y - \tilde{Y}$. Next we introduce a bump function $\phi \in C^\infty(M)$ which is supported in U and $\phi(p) = 1$. Since $Z = 0$ on U, we have that $\phi Z = 0$ on all of M. This then gives us:

$$X(\phi)Z + \phi\nabla_X Z = \nabla_X(\phi Z) = \nabla_X(0 \cdot \phi Z) = 0\nabla_X(\phi Z) = 0$$

Where for the first step we used the Leibniz rule for the linear connection and for the third step we used the second property of the linear connection. Now since $Z = 0$ on U, we have that $X(\phi)Z = 0$

on U. What this then tells us is that

$$0 = (\nabla_X(\phi Z))_p = \phi(p)(\nabla_X Z)_p = 1 \cdot (\nabla_X Z)_p = (\nabla_X Z)_p$$

From this we conclude that $(\nabla_X Y)_p$ depend on values of Y in U.
Let us now move on to show that $(\nabla_X Y)_p$ depends only on the value of X at p. Suppose again two vector fields X and $\tilde{X}$ such that $X_p = \tilde{X}_p$. Suppose we have a local frame (E_i) for U. We are then able to write $X = X^i E_i$ and $\tilde{X} = \tilde{X}^i E_i$. This tells us that for every i we have $X^i(p) = \tilde{X}^i(p)$. By applying the first property of the linear connection we have:

$$(\nabla_X Y)_p = X^i(p)(\nabla_{E_i} Y)_p = \tilde{X}^i(p)(\nabla_{E_i} Y)_p = (\nabla_{\tilde{X}} Y)_p$$

□

2.3.2 Christoffel symbols

Now given a local frame $(E_i)_{i=1}^n$ for an open U, we can then write any vector field as $X = X^i E_i$, where $X^i : U \to \mathbb{R}$. This elementary fact has an important consequence in the setting of a linear connection:

Definition 2.30. *Given a local frame $(E_i)_{i=1}^n$ for an open U, the* ***Christoffel symbols*** *$\Gamma_{ij}^k : M \to \mathbb{R}$ are the symbols satisfying:*

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

With the definition of the Christoffel symbols, we note:

Lemma 2.31. *As above, given a local frame $\{E_i\}_{i=1}^n$ for an open neighbourhood $U \subseteq M$, we can express our vector field $X, Y \in \mathfrak{X}(M)$ as $X = X^i E_i$ and $Y = Y^j E_j$, with $X^i, Y^i : M \to \mathbb{R}$. Applying this, we can write the linear connection as:*

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

Proof. So we start by writing:

$$\nabla_X Y = \nabla_X (Y^j E_j)$$

Applying the third property of definition 2.28 we get:

$$\nabla_X Y = X(Y^j) E_j + Y^j \nabla_X E_j$$

Next we write:

$$\nabla_X Y = X(Y^j) E_j + Y^j \nabla_{X^i E_i} E_j = X(Y^j) E_j + Y^j X^i \nabla_{E_i} E_j$$

Where in the second step we applied property one of definition 2.28. Now by relabeling and applying the definition of the Christoffel symbols we get the desired result. □

Remark (1). Combining this with lemma 2.29, one should notice that $(\nabla_X Y)(p)$, where $p \in M$ is a point, depends only on the first derivative of Y with respect to X. Or to put it differently, when we fix a local fram E_i, we also have a fixed way of writing $X = X^i E_i$ and $Y = Y^j E_j$. Thus $(\nabla_X Y)(p)$ only depends on the derivative $X(Y^j)(p)$. △

Remark (2). An important consequence of this lemma is that the Christoffel symbols contain the same information as the linear connection except they depend on the choice of the local frame. △

2.4 Parallel translation

In this subsection we will discuss parallel translation. Let us start out with a definition:

Definition 2.32. *Let M be a smooth manifold, endowed with a linear connection ∇, and let $\gamma : I \to M$ be a curve in M. A* ***parallel vector field along a curve*** *γ is a vector field V satisfying that at any time $t \in I$ we have $\nabla_{\dot{\gamma}(t)} V = 0$. We say that V is a* ***parallel vector field*** *if it satisfies this property for any curve.*

The way to understand this definition should be the following: In Euclidean space, we are able to define a global frame. This allows us to decompose any vector field and tells us if our vector field points in the same direction at all points on the manifold. To illustrate this, take as example the vector field $V = x\partial_x + y\partial_y$ defined on $\mathbb{R}^2$. It should be obvious that this vector field does not point in the same direction at all points on the surface, but we are only able to say this because ∂_x and ∂_y are a global frame. In an arbitrary semi-Riemannian manifold, we are not always able to define a global frame. So the next best thing we can do is to say whether or not the vector field points in the same direction at all points along a curve. And that is precisely what it means for a vector field V to be parallel along a curve γ. V does not change if we move it along the curve γ.
This leads us to the following important theorem:

Theorem 2.33. *Given a manifold M, a curve $\gamma : I \to M$ and a vector $v_0 \in T_{\gamma(t_0)}M$, there exists a unique parallel vector field V along γ such that $V(t_0) = v_0$.*

This theorem just tells us that given a vector, we are able to extend it to a vector field along the curve such that, at all points on the curve, the vector field points in "the same direction" as the given vector. In standard Euclidean space, this theorem should be nothing surprising. Given a vector at a point in the Euclidean space and using the fact that we can define a global frame, we just copy that vector to all the points in the Euclidean space. This than gives us the desired vector field. This fact here will be the intuition behind the proof. So let us start:

Proof. The existence of such a vector field V is a consequence of the existence and uniqueness of the solution for ODEs. To make this clear, let us first suppose that $\gamma(I)$ is contained in a single chart. This means that we can find a local frame (E_k) spanning all of $\gamma(I)$. By lemma 2.31 we can then write:

$$\nabla_{\dot{\gamma}(t)}V = (\frac{d}{dt}(V^k \circ \gamma)(t) + \dot{\gamma}^j V^i \Gamma^k_{ij})E_k$$

We know that if such a V exists, it must satisfy $\nabla_{\dot{\gamma}(t)}V = 0$. This reduces our equation above to a linear ODE and we know that this has a solution, which is unique, defined on all of I.
Now suppose that $\gamma(I)$ is not contained in a single chart. Define $\beta \in I$ to be the supremum of all times $t \in I$ such that V is unique. This supremum satisfies $\beta > t_0$ since if we take $t \in [t_0, \beta)$ small enough then $\gamma([t_0, t])$ is contained is a single chart and thus a unique solution exists. So now a unique parallel vector field exists on $[t_0, \beta)$. Again now, we can find an $\epsilon > 0$ small enough such that $\gamma((\beta - \epsilon, \beta + \epsilon))$ is contained in a single chart. Now this means that we can find a unique parallel vector field $\tilde{V}$. We now must have that $\tilde{V} = V$ on $(\beta - \epsilon, \beta)$. But this means that we can extend our parallel vector field V in a unique way contradicting the definition of β. Analogously, one can argue the infimum case showing that a parallel vector field must exist on all of $\gamma(I)$ satisfying the initial condition $V(t_0) = v_0$. ☐

What is also worth mentioning is that parallel translation defines an operator.

Definition 2.34. *Let $\gamma : I \to M$ be a curve on a semi-Riemannian manifold (M, g) endowed with a linear connection ∇. The* ***operator defined by parallel translation along*** *γ is the following map:*

$$P_{t_0 t_1} : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M : V \to P_{t_0 t_1}V = \tilde{V}(\gamma(t_1))$$

where we have that $t_0, t_1 \in I$ and $\tilde{V}$ is the parallel translate of V along γ

So in essence what this map does, is that for every vector $V \in T_{\gamma(t_0)}M$ we construct the unique parallel vector field $\tilde{V}$ from the previous theorem and then evaluate $\tilde{V}$ at the point $\gamma(t_1)$. It appears that this defines an isomorphism of vector spaces between $T_{\gamma(t_0)}M$ and $T_{\gamma(t_1)}M$.

Lemma 2.35. *Let $\gamma : I \to M$ be a curve on a semi-Riemannian manifold (M, g) endowed with a linear connection ∇. For any $t_0, t_1 \in I$ we have that $P_{t_0 t_1}$ is a vector space isomorphism.*

Proof. Let us start of by showing that this is a bijection. Let $V \in T_{\gamma(t_0)}M$ be a vector and let $\tilde{V}$ be the parallel translate of V along γ. We note that $\tilde{V}$ is unique by theorem 2.33. What this means is that for the vector $P_{t_0 t_1}V$ the only vector field which is parallel to γ and satisfies that $W(\gamma(t_1)) = P_{t_0 t_1}V$, is the vector field $\tilde{V}$. Hence from this it follows that $P_{t_1 t_0}P_{t_0 t_1}V = V$ and thus

$P_{t_0t_1}$ is bijective.
Now we show linearity. Let $V, W \in T_{\gamma(t_0)}M$ be vectors and let $\tilde{V}$ and $\tilde{W}$ be their parallel translates. It should be noted that the vector field $\tilde{V}+\tilde{W}$ is the vector field satisfying that $(\tilde{V}+\tilde{W})(\gamma(t_0)) = \tilde{V}(\gamma(t_0)) + \tilde{W}(\gamma(t_0)) = V + W$ and it is parallel along γ by the linearity of the linear connection. Thus by theorem 2.33 this vector field is unique. From it we see that $P_{t_0t_1}(V+W) = (\tilde{V} + \tilde{W})(\gamma(t_1)) = \tilde{V}(\gamma(t_1)) + \tilde{W}(\gamma(t_1)) = P_{t_0t_1}V + P_{t_0t_1}W$. □

2.5 Geodesics

Having defined the linear connection, we are now able to move on to geodesics.

Definition 2.36. *A **geodesic** is a curve $\gamma : I \to M$, where I is an interval, with the property that $\nabla_{\dot{\gamma}(t)}\dot{\gamma} = 0$, where $\dot{\gamma}$ is the speed of the curve.*

Let us remark two things about geodesics:

Remark (1)*.* If $\gamma(0) = p$ and $\dot{\gamma}(0) = v$, we will say that γ is a geodesic starting at p and has initial velocity v. If we want to be specific about the initial velocity, we will also write the geodesic as γ_v. △

Remark (2)*.* We will call a geodesic γ spacelike, timelike or null if for any $t \in I$ we have that $\dot{\gamma}(t)$ is spacelike, timelike or null respectfully. △

This following theorem plays a central role in semi-Riemannian geometry. It tells us that given some initial point and initial velocity, there always exists a geodesic satisfying those conditions.

Theorem 2.37. *Endow our manifold M with a linear connection. At any point $p \in M$ we can find an interval $I \subseteq \mathbb{R}$ containing 0 and a unique geodesic $\gamma : I \to M$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = V$ for any $V \in T_pM$.*

Again we need an auxiliary lemmas

Lemma 2.38. *Let U be an open neighbourhood of some point p on our semi-Riemannian manifold M. Let I be an interval and $\gamma : I \to U$ be a curve. Choose a coordinate system $(e_i)_{i=1}^n$ on U. Now we can $\gamma(t) = (x^1(t), \dots, x^n(t))$. Now it follows that γ is a geodesic if and only if the components of γ satisfy:*

$$\frac{d^2x^k}{dt^2}(t) + \Gamma^k_{ij}(\gamma(t))\frac{dx^i}{dt}(t)\frac{dx^j}{dt}(t) = 0 \tag{2}$$

The proof of this lemma amounts to applying lemma 2.31 and the definition of geodesics. This equation tells us that we can regard geodesics to be first-order ordinary differential equations (ODEs), namely we could define:

$$\dot{x}^k(t) = v^k(t)$$
$$\dot{v}^k(t) = -v^i(t)v^j(t)\Gamma^k_{ij}(\gamma(t))$$

The domain of definition of these ODEs is the tangent bundle. Now the proof of theorem 2.37 amounts to be an application of the existence and uniqueness theorem for first-order ODEs. For the precise statement of the theorem we refer to [3] theorem IV.4.1.. Now we will prove our theorem.

Proof of Theorem 2.37. Let us start out with uniquess. This amounts to iteratively aplying the uniqueness of first-order ODEs. Let $\gamma, \sigma : I \to M$ be two geodesics satisfying $\gamma(0) = \sigma(0)$ and $\dot{\gamma}(0) = \dot{\sigma}(0)$. Let b be the supremum of all the points in I such that γ and σ agree on the interval $[0, b]$. Such a b must exist by uniqueness of ODEs. Suppose $b \in I$, then by the fact that geodesics are smooth we then have $\gamma(b) = \sigma(b)$ and $\dot{\gamma}(b) = \dot{\sigma}(b)$. But then again by existence and uniqueness of ODEs we can find an open interval around b such that the two geodesics agree on that interval. This contradicts our choice of b. Hence $b \notin I$. Anologously, but working with infimums, we can show that they must agree on whole I.
Now we prove existence. Let $p \in M$ and $U \subseteq M$ be open. Let $V \in \mathbb{R}^n$ be a point. Then by the existence and uniqueness of ODEs, we can find an open interval $I_\epsilon = (-\epsilon, \epsilon)$ with $\epsilon > 0$ and a

unique solution for our first-order ODEs $\eta : I_\epsilon \to U \times \mathbb{R}^n$ satisfying $\eta(0) = (p, V)$. Now if we write η in components we get $\eta(t) = (x^i(t), v^i(t))$. It then readily follows that the curve $\gamma(t) = (x^i(t))$ satisfies the condition and is thus the geodesic we look for. □

2.6 Levi-Civita Connection

2.6.1 Existence and Uniqueness of the Levi-Civita Connection

There are many possible choices for the linear connection on any manifold, but there is a natural choice possible: the Levi-Civita connection.

Theorem 2.39. *On any semi-Riemannian manifold, there exists a unique linear connection, called the* ***Levi-Civita connection****, satisfying conditions 1-3 of definition 2.28, and satisfying the following two conditions:*

1. *It is torsion free:* $[X, Y] = \nabla_X Y - \nabla_Y X$
2. *It is compatible with the metric:* $X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$

This holds true for any $X, Y, Z \in \mathfrak{X}(M)$.

So the first property of the Levi-Civita connection is that it is torsion free. We know that the Lie brackets answers the question: suppose we move for some time t along vector field X. Then for the same time t along Y. And then for the same time we move backwards along X and finally also for Y, would we end up back in the same point? The Levi-Civita connection tells us that this depends on the fact whether or not doing parallel transport of X along Y is the same as doing parallel transport of Y along X.

The second property is a generelization of the product rule for differentiation. To show this, let us look at an example.

Example 2.40. Let $\mathbb{R}^n$ be the Euclidean space with the standard Euclidean metric. Endow this space with the **Euclidean connection** $\bar{\nabla}$ defined as:

$$\bar{\nabla}_X Y = X(Y^i)\partial_i$$

Where X and Y are vector fields on and ∂_i is the global frame for $\mathbb{R}^n$. Now let us look at the following.

$$X(\bar{g}(Y, Z)) = X(\delta_{ij} Y^i Z^j) = \delta_{ij} X(Y^i Z^j)$$

Now we see that Y^i and Z^j are functions and we know that the Lie derivative satisfies the product rule, so we get:

$$X(\bar{g}(Y, Z)) = \delta_{ij} X(Y^i) Z^j + \delta_{ij} Y^i X(Z^j) = \bar{g}(\bar{\nabla}_X Y, Z) + \bar{g}(Y, \bar{\nabla}_X Z)$$

And so we see that this is the generalization of the product rule for differentiation. ♢

This is such a nice property in the case of differentiation that we want the natural choice for the linear connection to also satisfy it.

But now we have to prove that such a connection does exist. For the proof, we will mimic the argument given by O'Neill ([9], page 60 and 61). We will need an auxiliary lemma. This one will serve to prove the uniqueness of the Levi-Civita connection:

Lemma 2.41. *Suppose the Levi-Civita connection exists, then it satisfies the* ***Koszul formula****:*

$$\begin{aligned} 2g(\nabla_X Y, Z) =& X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) \\ &- g(X, [Y, Z]) + g(Y, [Z, X]) + g(Z, [X, Y]) \end{aligned} \tag{3}$$

This holds true for any $X, Y, Z \in \mathfrak{X}(M)$.

Proof. We start from the right hand side. First let us apply condition 1 of theorem 2.8 to the first three terms of the right hand side:

$$\begin{aligned}
X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) + g(\nabla_Y Z, X) + \\
&\quad g(Z, \nabla_Y X) - g(\nabla_Z X, Y) - g(X, \nabla_Z Y) \\
X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) &= g(Z, \nabla_X Y) + g(Z, \nabla_Y X) + g(Y, \nabla_X Z) - \\
&\quad g(Y, \nabla_Z X) + g(X, \nabla_Y Z) - g(X, \nabla_Z Y) \\
X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) &= g(Z, \nabla_X Y + \nabla_Y X) + g(Y, \nabla_X Z - \nabla_Z X) + \\
&\quad g(X, \nabla_Y Z - \nabla_Z Y)
\end{aligned}$$

In the last two steps, we applied the symmetry and linearty condition of the metric. We have written this in the form such that we can apply condition 2 of theorem 2.8. This now gives us:

$$X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) = g(Z, \nabla_X Y + \nabla_Y X) + g(Y, -[Z,X]) + g(X, [Y,Z])$$

Now we write:

$$\begin{aligned}
&X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) - g(X,[Y,Z]) + g(Y,[Z,X]) + g(Z,[X,Y]) \\
&= g(Z, \nabla_X Y + \nabla_Y X) + g(Y, -[Z,X]) + g(X,[Y,Z]) - g(X,[Y,Z]) + g(Y,[Z,X]) + g(Z,[X,Y])
\end{aligned}$$

This now simplifies to:

$$\begin{aligned}
X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) - g(X,[Y,Z]) + g(Y,[Z,X]) + g(Z,[X,Y]) \\
= g(Z, \nabla_X Y + \nabla_Y X) + g(Z,[X,Y])
\end{aligned}$$

Applying condition 2 of theorem 2.39 and applying the linearity of the metric gives us the desired result. □

An interesting result is the following:

Corollary 2.42. *Endowing our semi-Riemannian manifold with the Levi-Civita connection and looking at a local coordinate chart $(U,(x^i))$, we can write the Christoffel symbols as:*

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl} (\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}) \tag{4}$$

Proof. We choose for $X = \partial_i$, $Y = \partial_j$ and $Z = \partial_k$ in the equation (3). Next we note that the Lie bracket of any local coordinate chart is $[\partial_i, \partial_j] = 0$. And lastly noting that any local coordinate chart is a local frame, we get that:

$$\nabla_{\partial_i} \partial_j = \Gamma^k_{ij} \partial_k$$

Using these facts and formula (3) the result follows readily. □

So now let us proof theorem 2.39

Proof of Theorem 2.39. Uniqueness follows trivially from formula (3), because the right hand side does not depend on the linear connection. So the only thing left to prove is existence. Define $F(X,Y,Z)$ to be the right hand side of equation (3). Now take $X, Y \in \mathfrak{X}(M)$ to be fixed. It is possible though tedious to show that the function $Z \to F(X,Y,Z)$ is $C^\infty(M)$-linear. By lemma 2.7 it defines a one-form. In lemma 2.24, we showed that there exists a bijection between the space of vector fields and the space of one-forms. Hence we can find a vector field such that $F(X,Y,Z) = g(V,Z)$. This vector field has to be unique by injectivity, hence $V = \nabla_X Y$. So the linear connection, satisfying the right hand side of formula (3), exists and from it we can deduce all the property of the linear connection plus the two properties of the theorem. □

2.6.2 Nice Properties of the Levi-Civita connection

The first nice property of the Levi-Civita connection is that it preserves the spacelike/lightlike/timelike character of geodesics. In fact this holds true for any linear connection that is compatible with the metric.

Lemma 2.43. *Let (M, g) be a semi-Riemannian manifold endowed with a linear connection ∇ compatible with the metric. Let $\gamma : I \to M$ be a geodesic sucht that at some time $t_0 \in I$ we have that $\dot{\gamma}(t_0)$ is spacelike/timelike/null. Then γ is spacelike/timelike/null for all times $t \in I$*

Proof. For simplicity, we prove the spacelike case. Thr proof for the other cases goes precisely the same way. We know that the geodesic is spacelike at time t_0. This means that $g(\dot{\gamma}(t_0), \dot{\gamma}(t_0)) \geq 0$. Let us look at the time derivative of the metric:

$$\frac{d}{dt} g(\dot{\gamma}(t), \dot{\gamma}(t)) = 2g(\nabla_{\dot{\gamma}} \dot{\gamma}(t), \dot{\gamma}(t)) = 2g(0, \dot{\gamma}(t)) = 0$$

In the first step we used the fact that the linear connection is compatible with the metric. In the second step we used the fact that for geodesics we have $\nabla_{\dot{\gamma}} \dot{\gamma}(t) = 0$. Since we know that $g(\dot{\gamma}(t_0), \dot{\gamma}(t_0)) \geq 0$ and it does not vary in time, we must have that $g(\dot{\gamma}(t), \dot{\gamma}(t)) \geq 0$ for all time $t \in I$. ☐

Secondly, to show that the Levi-Civita connection is the natural choice for Riemannian manifolds, we will prove the following lemma:

Lemma 2.44. *Let $\phi : (M, g_M) \to (N, g_N)$ be an isometry. Let also ∇ be the Levi-Civita connection on M and $\widetilde{\nabla}$ be the Levi-Civita connection on N. Then the following holds true: $\phi_*(\nabla_X(Y)) = \widetilde{\nabla}_{\phi_* X}(\phi_* Y)$.*

Proof. The gist of the proof is the following: Define a new linear connection on M in the following way:

$$\phi^* \widetilde{\nabla} : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M) : (X, Y) \to (\phi^* \widetilde{\nabla})_X Y = \phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}(\phi_* Y))$$

Now what we have to show is that this is satisfies the properties of the Levi-Civita connection and then, by uniqueness of the Levi-Civita connection, it must that connection. To show that the new connection satisfies all the requirements of the Levi-Civita connection is a lot of algebra and not very insightfull, so we will only show two of the properties. First we show that it satisfies property 3 of definition 2.28. So let $f \in C^\infty(M)$, we look at:

$$(\phi^* \widetilde{\nabla})_X (fY) = \phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}(\phi_*(fY)))$$

We note that the pushforward of a function is $\phi_*(f) = f \circ \phi^{-1}$. Next we note $\phi_*(fY) = (f \circ \phi^{-1})\phi_*(Y)$ So now we can apply the fact that $\widetilde{\nabla}$ is the Levi-Civita connection on N to get:

$$\begin{aligned} \phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}(\phi_*(fY))) =& \phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}((f \circ \phi^{-1})\phi_*(Y)) \\ =& \phi_*^{-1}((f \circ \phi^{-1})\widetilde{\nabla}_{\phi_* X}(\phi_*(Y)) + (\phi_* X)(f \circ \phi^{-1})\phi_*(Y)) \end{aligned}$$

So next, by property of the Lie derivative, we have that $(\phi_* X)(f \circ \phi^{-1}) = X(f \circ \phi^{-1} \circ \phi) \circ \phi^{-1}$. Hence we get that $\phi_*^{-1}((\phi_* X)(f \circ \phi^{-1})\phi_*(Y)) = X(f \circ \phi^{-1} \circ \phi) \circ \phi^{-1} \circ \phi(\phi_*^{-1}\phi_*(Y)) = (Xf)Y$. And lastly we note that $\phi_*^{-1}((f \circ \phi^{-1})\widetilde{\nabla}_{\phi_* X}(\phi_*(Y)) = (f \circ \phi^{-1} \circ \phi)(\phi^* \widetilde{\nabla})_X Y = f(\phi^* \widetilde{\nabla})_X Y$ by definition. So now by applying all of this, we get:

$$\phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}(\phi_*(fY))) = f(\phi^* \widetilde{\nabla})_X Y + (Xf)Y$$

So now we will show property 1 of theorem 2.39. To show this, we will look at:

$$\begin{aligned} (\phi^* \widetilde{\nabla})_X Y - (\phi^* \widetilde{\nabla})_Y X =& \phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}(\phi_* Y)) - \phi_*^{-1}(\widetilde{\nabla}_{\phi_* Y}(\phi_* X)) \\ =& \phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}(\phi_* Y) - \widetilde{\nabla}_{\phi_* Y}(\phi_* X)) \end{aligned}$$

The second step is true by the linearity of the pushforward. Next we apply again the fact that $\widetilde{\nabla}$ is the Levi-Civita connection on N and hence $\widetilde{\nabla}_{\phi_* X}(\phi_* Y) - \widetilde{\nabla}_{\phi_* Y}(\phi_* X) = [\phi_* X, \phi_* Y]$. Another property of the pushforward is that it preserves Lie brackets. So it follows that $\phi_*^{-1}([\phi_* X, \phi_* Y]) = [\phi_*^{-1}\phi_* X, \phi_*^{-1}\phi_* Y] = [X, Y]$.
By uniqueness of the Levi-Civita connection, it follows that for any $X, Y \in \mathfrak{X}(M)$ we have:

$$\nabla_X Y = (\phi^* \widetilde{\nabla})_X Y = \phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}(\phi_* Y))$$

Hence it now follows that:

$$\phi_*(\nabla_X Y) = \phi_*(\phi_*^{-1}(\widetilde{\nabla}_{\phi_* X}(\phi_* Y))) = \widetilde{\nabla}_{\phi_* X}(\phi_* Y)$$

□

Another reason why the Levi-Civita connection is the natural choice for the connection is the following:

Lemma 2.45. *Let M and N be two semi-Riemannian manifolds and let $\phi : M \to N$ be an isometry. Endow M and N with the Levi-Civita connection ∇ and $\widetilde{\nabla}$ respectively. If we have that $\gamma : I \to M$ is an geodesic starting at p with initial velocity V, then $\phi \circ \gamma : I \to N$ is an geodesic in N starting at $\phi(p)$ and with initial velocity $\phi_*(V)$.*

Proof. Since γ is a geodesic is satisfies $\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = 0$. Hence by lemma 2.44 we get that $\widetilde{\nabla}_{\phi_*(\dot{\gamma}(t))}(\phi_*(\dot{\gamma}(t))) = 0$. This holds true for any $t \in I$. Lastly note by the chain rule for differentiation that $\frac{\partial}{\partial t}(\phi \circ \gamma) = \phi_*(\dot{\gamma}(t))$. □

To conclude this section, let us look at the Euclidean space:

Example 2.46. As announced, let us show that in standard n-dimensional Euclidean space endowed with the Euclidean metric, the geodesics are straight lines. Let us endow it with the Euclidean connection, defined as:

$$\nabla_X Y = (XY^i)\partial_i$$

Showing that this is the Levi-Civita connection on $\mathbb{R}^n$ just amounts to checking the properties stated in theorem 2.39. Let now $\gamma : I \to \mathbb{R}^n$ be any curve. Now since we endowed $\mathbb{R}^n$ with it's Levi-Civita connection, we are able to use formula (4). Since g_{ij} are constant, it follows that all Christoffel symbols are zero. Using this fact and formula (2), it follows:

$$\frac{d^2 x^k}{dt^2}(t) = 0$$

Hence the straight lines are the geodesics in standard Euclidean space. ♢

2.7 Normal Coordinates

When we work locally in an arbitrary smooth manifold we usually just choose a nice coordinate chart. In a semi-Riemannian manifold there appears to be a natural choice for coordinate charts. These are what we call normal coordinates. In this subsection we introduce what they are and why they are usefull.

2.7.1 Exponential Map

To be able to define normal coordinates, we need the exponential map.

Definition 2.47. *Let (M, g) be a semi-Riemannian manifold. We define the* ***domain of the exponential map*** *to be:*

$$\mathcal{E} := \{(p, v) \in TM | \gamma_{(p,v)} \text{ is defined on } [0, 1]\}$$

where $\gamma_{(p,v)}$ is the geodesic satisfying $\gamma_{(p,v)}(0) = p$ and $\frac{\partial \gamma_{(p,v)}}{\partial t}(0) = v$. Then we define the ***exponential map*** *as:*

$$exp : \mathcal{E} \to M : (p, v) \to exp(p, v) = \gamma_{(p,v)}(1)$$

So what the exponential map tells us is where the geodesics passing through some point $p \in M$ with a directon $v \in T_pM$ ends up after an unit time has passed. To put it in other words, the exponential map tells us how the geodesics depend on their initial conditions.

We can also define the exponential map more locally in the following way:

Definition 2.48. *Let (M, g) be a semi-Riemannian manifold and $p \in M$. We define the* ***exponential map at*** *p as:*

$$exp_p : \mathcal{E}_p \to M : v \to exp_p(v) = \gamma_(p, v)(1)$$

where $\mathcal{E}_p := \{v \in T_pM | \gamma_{(p,v)}$ is defined on $[0, 1]\}$

The exponential map has some nice properties.

Lemma 2.49. *Let (M, g) be a semi-Riemannian manifold an $(p, v) \in TM$. The exponential map satisfies the following properties.*

1. *The exponential map is smooth;*
2. *Let $\gamma_{(p,v)} : I \to M$ be a geodesic. Then for all $t \in I$ we have: $exp_p(tv) = \gamma_{(p,v)}(t)$;*
3. *The domain of the exponential map is an open subset of TM containing the zero section;*
4. *$\mathcal{E}_p$ is star-shaped with respect to the origin;*
5. *There exists an open $U \subseteq M$ containing p and an open $V \subseteq \mathcal{E}_p$ such that $exp_p : V \to U$ is a diffeomorphism.*

Remark. The open subset U from point five of the preceding lemma is also called a **normal neighbourhood** of p. △

For a proof of this lemma we refer to [7] chapter 6 proposition 5.7 and lemma 5.10.

2.7.2 Normal Coordinates

So now we are able to define normal coordinates.

Definition 2.50. *Let (M, g) be a n-dimensional semi-Riemannian manifold and let $p \in M$ be a point. Let furhtermore $(E_i)_{i=1}^n$ be an orthonormal basis for T_pM and $E : \mathbb{R}^n \to T_pM : (x^1, \ldots, x^n) \to x^iE_i$ be an isomorphism of vector spaces. Lastly let U be an normal neighbourhood of p. Then we define the* ***normal coordinates*** *centered at p as:*

$$\phi = E^{-1} \circ exp_p^{-1} : U \to \mathbb{R}^n$$

The reason we introduced normal coordinates is to prove the following lemma which we will need in our discussion in the next chapters.

Lemma 2.51. *Suppose (M, g) is a n-dimensional Riemannian manifold and $p \in M$ is a point. Let $U \subseteq M$ be an open neighbourhood of p and let $(x^i)_{i=1}^n$ be normal coordinate charts centered at p. Then the following holds true:*

1. *Let $\gamma_{(p,v)}$ be a geodesic where $v = v^i\partial_i \in T_pM$. Then in normal coordinates this geodesic is represented by:*
$$\gamma_{(p,v)}(t) = (tv^1, \ldots, tv^n)$$
As long as $\gamma_{(t,v)}$ stays in U;
2. *The components of the metrix g_{ij} are equal to the Kronecker delta δ_{ij} at the point p. To put it formally $g_{ij}(p) = \delta_{ij}$.*

Proof. Let $(\partial_i)_{i=1}^n$ be an orhtonormal basis for T_pM. Suppose $E : \mathbb{R}^n \to T_pM : (x^1, \ldots x^n) \to x^i\partial_i$ be the vector space isomorphism such that we can define the normal coordinate chart as $\phi = E^{-1} \circ exp_p^{-1}$. By lemma 2.49 point two, we know that we can write $\gamma_{(p,v)}(t) = exp_p(tv)$. Hence this now gives us:

$$\phi \circ \gamma_{(p,v)}(t) = E^{-1} \circ exp_p^{-1} \circ exp_p(tv) = E^{-1}(tv^i\partial_i) = (tv^1, \ldots, tv^n)$$

Now to prove point two we use that we can write $\gamma_{(p,\partial_i)}(t) = (0, \ldots, 0, t, 0, \ldots, 0)$ where the only non-zero component is the i-th component. Furhtermore since $\frac{\partial \gamma_{(p,\partial_i)}}{\partial t}(0), \frac{\partial \gamma_{(p,\partial_j)}}{\partial t}(0) \in T_pM$ we can talk about them being orthogonal. This of course only holds true at $t = 0$ since at any orhter time t these vectors are in different tangent spaces. It should be clear that:

$$g(\frac{\partial \gamma_{(p,\partial_i)}}{\partial t}(0), \frac{\partial \gamma_{(p,\partial_j)}}{\partial t}(0)) = \delta_{ij}$$

since by the above representation in normal coordinates they are orhtogonal if $i \neq j$. By definition it is also easy to see that $\frac{\partial \gamma_{(p,\partial_i)}}{\partial t}(0) = \partial_i$. Hence this gives us:

$$g(\frac{\partial \gamma_{(p,\partial_i)}}{\partial t}(0), \frac{\partial \gamma_{(p,\partial_j)}}{\partial t}(0)) = g(\partial_i, \partial_j) = g_{ij}$$

The above holds true by definition. This proves that $g_{ij} = \delta_{ij}$ at the point p. □

Remark. We note that the first point of the above lemma holds true for any semi-Riemannian manifold. The second point also holds true in an arbitray semi-Riemannian manifold, but then we have to take into account the metric signature. We would then rewrite the condition as:

$$g_{ij}(p) = \begin{cases} \delta_{ij}, & \text{if } \partial_i \text{ is spacelike} \\ -\delta_{ij}, & \text{if } \partial_i \text{ is timelike} \end{cases}$$

△

2.8 Curvature

Curvature plays a central role in various aspects of Riemannian and semi-Riemannian geometry. It this thesis though, it will not play a significant role. So again in this subsection we will only state some results without offering a proof which we will use later on.

2.8.1 Riemannian Curvature Tensor

A very important way of classifying semi-Riemannian manifolds is through the notion of curvature. To make the intuition of curvature clear, let us take a brief look again at the Euclidean space.

Example 2.52. Suppose we are in $\mathbb{R}^n$ which we have endowed with the standard Euclidean metric and connection. Suppose furthermore that we have a vector field Z. We note that in this case we have:

$$\bar{\nabla}_{\partial_i} \bar{\nabla}_{\partial_j} Z = \partial_i(\partial_j(Z^k))\partial_k = \partial_j(\partial_i(Z^k))\partial_k = \bar{\nabla}_{\partial_j} \bar{\nabla}_{\partial_i} Z$$

This is so since partial derivatives commute. To put this in a more general form, we have:

$$\bar{\nabla}_{\partial_i} \bar{\nabla}_{\partial_j} Z - \bar{\nabla}_{\partial_j} \bar{\nabla}_{\partial_i} Z = \bar{\nabla}_{[\partial_i, \partial_j]} Z$$

♢

The commutation of partial derivatives is such a nice property that we should ask ourselves: Is the first or even the second formula of our example true for any linear connection? The answer is no, but this leads to a very nice property: the curvature.

Definition 2.53. *Given a semi-Riemannian manifold (M, g) endowed with a linear connection, the **Riemannian curvature tensor** is the map:*

$$R : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M) : (X, Y, Z) \to R(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

We have already suggestively written the Riemannian curvature *tensor*. So the following lemma should not come as a surprice.

Lemma 2.54. *The Riemannian curvature tensor is a $\binom{3}{1}$-tensor field.*

We can think about the Riemannian curvature tensor as a way to measure the degree of failure for our linear connection to commute. Phrased in this way, it should be obvious that the curvature is a property intrinsic to our manifold and thus it allows us to classify various semi-Riemannian manifolds. This follows from the following lemma.

Lemma 2.55. *Given two semi-Riemannian manifolds (M, g_M) and (N, g_N) and an isometry $\phi : M \to N$. Suppose furthermore we have $X, Y, Z \in \mathfrak{X}(M)$. It follows that:*

$$R_N(\phi_* X, \phi_* Y)\phi_* Z = \phi_*(R_M(X, Y)Z)$$

Where R_M and R_N are the Riemannian curvature tensors on M and N respectively.

2.8.2 Second Fundamental Form

What should be noted is that the curvature is something intrinsic to the metric on the manifold. We also defined Riemannian submanifolds in definition 2.12. Now since curvature is intrinsic to the metric, we could ask ourselves: What is the relationship between the curvature of the submanifold and the curvature of the surrounding manifold?

To answer this question we start by decomposing part of the tangent bundle over the ambient space in two components: one component which is the tangent bundle over the submanifold and one component orthogonal to this tangent bundle. The idea behind this is that this first components contains the information regarding the intrinsic curvature of the submanifold. The second component contains the relationship between the curvature of the submanifold and the ambient space.

To this end, let us define the restriction of the tangent bundle.

Definition 2.56. *Suppose we have a semi-Riemannian manifolds (M, g_M) and Riemannian submanifold (N, g_N) of M. The **restriction of the tangent bundle** of M to N is defined as:*

$$TM|_N = \sqcup_{p \in N} T_p M$$

Let us now define the normal space.

Definition 2.57. *Suppose we have a semi-Riemannian manifolds (M, g_M) and Riemannian submanifold (N, g_N) of M. And take a point $p \in N$. We decompose the tangent space as*

$$T_p M = T_p N \oplus (T_p N)^\perp$$

*where $(T_pN)^\perp$ is the **normal space** at p which satisfiest that a element $w \in (T_pN)^\perp$ is orthogonal to any $v \in T_pN$ with respect to the metric g_M. The **normal bundle** is then $TN^\perp = \sqcup_{p \in N}(T_pN)^\perp$.*

Remark. The reason we require N to be a Riemannian submanifold has to do with the submaifolds such that for all $p \in N$ we have that T_pN contains only null directions. Since any null direction is orthogonal to itself. This implies that its normal bundle over this submanifold contains its tangent bundle. Hence implying that we cannot split this into a tangential and orthogonal component. To avoid this problem, we jus t let N be a Riemannian submanifold. △

Of course, we can now define the orthogonal and tangent projections which will tell us how to decompose a vector field on M in this way.

Definition 2.58. *Suppose we have two semi-Riemannian manifolds (M, g_M) and (N, g_N), where N is a Riemannian submanifold of M. The **orthogonal projection** is defined as:*

$$\pi^\perp : TM|_N \to TN^\perp$$

*The **tangent projection** is defined as:*

$$\pi^\top : TM|_N \to TN$$

Definition 2.59. *Suppose we have a semi-Riemannian manifolds (M, g_M) and Riemannian manifold (N, g_N), where N is a Riemannian submanifold of M. Suppose also that we have $X \in \mathfrak{X}(N)$ which we extend arbitrarily to all of M. We can decompose this in a **orthogonal component of the vector field** $X^\perp := \pi^\perp(X)$ and a **tangential component** $X^\top = \pi^\top(X)$.*

Now here comes the crux of this all. Since the linear connection on any manifold takes two vector fields and gives us one back, we can decompose it in a orthogonal and a tangent component. The tangent component will coincide with the intrinsic curvature of the submanifold and the orthogonal component will be the relationship between the curvatures of the submanifold and the ambient manifold. This relationship is called the second fundamental form.

Definition 2.60. *Suppose we have a semi-Riemannian manifolds (M, g_M) and Riemannian submanifold (N, g_N) of M. Let M be endowed with a Levi-Civita connection $\tilde{\nabla}$ and let $\pi^{\perp}$ be the orthogonal projection into N. Then we define the* ***second fundamental form*** *of N as:*

$$II : \mathfrak{X}(N) \times \mathfrak{X}(N) \to \mathfrak{X}^{\perp}(N) : (X, Y) \to II(X, Y) := (\tilde{\nabla}_X Y)^{\perp}$$

Where an element of $\mathfrak{X}^{\perp}(N)$ is a smooth secion of $TN^{\perp}$.

Now the second fundamental form has some properties.

Lemma 2.61. *The second fundamental form satisfies:*

1. *independent of the extension of our vector fields;*
2. *bilinear over $C^{\infty}(M)$: $II(fX + gY, Z) = fII(X, Z) + gII(Y, Z)$;*
3. *symmetric in X and Y: $II(X, Y) = II(Y, X)$.*

Proof. We start out by showing the symmetry of the second fundamental form. To this end we look at:

$$II(X, Y) - II(Y, X) = (\tilde{\nabla}_X Y)^{\perp} - (\tilde{\nabla}_Y X)^{\perp} = (\tilde{\nabla}_X Y - \tilde{\nabla}_Y X)^{\perp} = [X, Y]^{\perp}$$

In the second step we made use of the fact that the orthogonal projection is linear. In the third step we made use of a property of the Levi-Civita conneciton. Now since X and Y are tangent to N, it follows by property of the Lie bracket that $[X, Y]$ is tangent to N. Thus $[X, Y]^{\perp} = 0$ proving the symmetry.
Since by lemma 2.29 the vector field $(\nabla_X Y)(p)$ evaluated at some point $p \in N$ depends only on the vector X_p, it does not matter how we extend our vector fields to M. By symmetry this must also hold true for Y.
Lastly noting that $\tilde{\nabla}_X Y$ is $C^{\infty}(M)$-linear in X and the orthogonal projection is $C^{\infty}(M)$-linear, II is linear in the first term. By symmetry, it follows that it is bilinear. □

To conclude this subsection, we state the Gauss Formula.

Theorem 2.62. *Suppose we have a semi-Riemannian manifolds (M, g_M) and Riemannian submanifold (N, g_N) of M and both endowed with the Levi-Civita connection $\tilde{\nabla}$ and ∇ respectively. Suppose furthermore that we have two vector fields $X, Y \in \mathfrak{X}(N)$ which we extend arbitrarily to all of M. Now the following formula holds:*

$$\tilde{\nabla}_X Y = \nabla_X Y + II(X, Y)$$

The way one would go about to prove this is showing that $(\tilde{\nabla}_X Y)^{\top}$ is the Levi-Civita connection on N. By uniqueness of the Levi-Civita connection, the result follows. This formula should not come as a surprise. As stated the goal of the second fundamental form is to relate the curvature of the Riemannian submanifold N to the curvature of the semi-Riemannian manifold M. The curvature is defined by the use of the linear connection and in this case the Levi-Civita connection. So the second fundamental form should relate the Levi-Civita connection of N to the Levi-Civita connection fo M. That is precisely what this formula tells us.

2.9 Computations on $\mathbb{S}^2$

To close this chapter it is instructive to look at $\mathbb{S}^2$. We will first calculate the metric on an open subset of $\mathbb{S}^2$. Then we will calculate the Christoffel symbols. Lastly we will determine how the geodesics look like. For any point $(x, y, z) \in \mathbb{S}^2$ we can find two intervals I and J such that we can parametrize an open neighbourhood of said point by two coordinates (θ, ϕ) in the following way:

$$x = \cos(\theta)\sin(\phi)$$
$$y = \sin(\theta)\sin(\phi)$$
$$z = \cos(\phi)$$

Where $\theta \in I$ and $\phi \in J$. So for ease, let's choose a point $(x, y, z) \in \mathbb{S}^2 - \{(x, y, z) \in \mathbb{S}^2 | y = 0 \text{ and } x > 0\}$. The intervals would then be $I = (0, 2\pi)$ and $J = (0, \pi)$.
We know how to write the metric in standard Euclidean space:

$$g = dx^2 + dy^2 + dz^2$$

And we also know how to write the metric of a local parametrization f:

$$g = \sum_i (\frac{\partial f^i}{\partial x^j} dx^j)^2$$

So then we can write the metric on the open subset spanned by our two intervals on the sphere as:

$$g = (\frac{\partial x}{\partial \theta} d\theta + \frac{\partial x}{\partial \phi} d\phi)^2 + (\frac{\partial y}{\partial \theta} d\theta + \frac{\partial y}{\partial \phi} d\phi)^2 + (\frac{\partial z}{\partial \theta} d\theta + \frac{\partial z}{\partial \phi} d\phi)^2$$

After some calculations we get:

$$g = d\phi^2 + \sin^2(\phi) d\theta^2$$

So let us write this in matrix form:

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(\phi) \end{pmatrix}$$

With inverse:

$$g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sin^2(\phi)} \end{pmatrix}$$

So now using formula (4) we can calculate the Christoffel symbols. We will only compute the matrix of Γ^ϕ. A small note on notation: we will call the element on the upper left hand side of matrix g_{ij} $g_{\phi\phi}$, the one on the upper right hand side $g_{\phi\theta}$, lower left hand side $g_{\theta\phi}$ and lower right hand side $g_{\theta\theta}$.

So now to compute Γ^ϕ we first of all note that the elements which are not on the diagonal of matrix g^{ij}, are zero. Or more formally if $i \neq j$, then $g^{ij} = 0$. Next we note that $\partial_\phi g_{\phi\phi} = 0$. From this we conclude that $\Gamma^\phi_{\phi\phi} = 0$. Next we note that $\Gamma^\phi_{\phi\theta} = \Gamma^\phi_{\theta\phi}$ because of formula (4). We also see that $\partial_\theta g_{\phi\phi} = 0$, this then gives us $\Gamma^\phi_{\theta\phi} = 0$. Lastly we have to compute $\Gamma^\phi_{\theta\theta}$:

$$\Gamma^\phi_{\theta\theta} = \frac{1}{2} g^{\phi\phi}(\partial_\theta g_{\theta\phi} + \partial_\theta g_{\theta\phi} - \partial_\phi g_{\theta\theta})$$
$$\Gamma^\phi_{\theta\theta} = \frac{1}{2} 1(0 + 0 - \partial_\phi(\sin^2(\phi)))$$
$$\Gamma^\phi_{\theta\theta} = -\sin(\phi)\cos(\phi)$$

So this now gives us:

$$\Gamma^\phi = \begin{pmatrix} 0 & 0 \\ 0 & -\sin(\phi)\cos(\phi) \end{pmatrix}$$

Analogously, one can calculate:

$$\Gamma^\theta = \begin{pmatrix} 0 & \frac{\cos(\phi)}{\sin(\phi)} \\ \frac{\cos(\phi)}{\sin(\phi)} & 0 \end{pmatrix}$$

Now we will show that the following lemma.

Lemma 2.63. *A curve on* $(\mathbb{S}^2, g = d\phi^2 + \sin(\phi)^2 d\theta^2)$ *is a geodesic if and only if it is a great circle.*

Proof. To this end we will show that the equator of the sphere is a geodesic. It is easy to see that the equator is characterised by $(\theta(t), \phi(t)) = (t, \frac{\pi}{2})$. We see that $\frac{\partial^2\theta}{\partial t^2} = \frac{\partial^2\phi}{\partial t^2} = 0$ at any time t. And we note that both Γ^θ as Γ^ϕ are zero matrices, because $\cos(\frac{\pi}{2}) = 0$. Using formula (2) we get that the equator is in fact a geodesic. But there is nothing special to the equator. We can rotate our sphere to get an new equator. This shows us that any great circle is a geodesic.
Now we show the converse. Suppose we have a point $p \in \mathbb{S}^2$ on the equator and a direction $V \in T_p\mathbb{S}^2$ pointing along the equator. This is always possible because otherwise we just rotate the sphere such that does. Let γ be a geodesic such that $\gamma(0) = p$ and $\dot{\gamma}(0) = V$. This exists by existence of the geodesics. Now we know by the previous part of the proof that the great circle is a geodesic satisfying these properties. Now by the uniqueness part of theorem 2.37 we get that the great circles are the only geodesics. □

3 Rolling without Slipping and Twisting

In this chapter we will apply the basics of semi-Riemannian geometry to model rolling without slipping or twisting. We must first introduce what the lightcone and the Grassmanian are. After having done that, we introduce the natural space in which to model rolling: the configuration space. This space keeps track of two things: 1) the points of contact between the two objects rolling. 2) the planes of contact which are in essence all the directions in which our two objects can move such that they still roll on each other. We also show that we can phrase configuration space in terms of the Grassmanian of the lightcone.

Having introduced the configuration space, we are able to define the no slipping distribution which models rolling without slipping. We give the physical motivation behind our definition before moving on to the no twisting distribution. The no twisting distribution appears to be a lot harder to define and a lot of attention will be given to the mathematical formalism allowing us to model rolling without twisting. Showing that the no slipping and no twisting distributions are well-defined objects in any dimension will be the main result of both this chapter as well as this thesis.

After having done this, we move on to an example of rolling: the coin rolling on the table. We will compute the no slipping and the no twisting distribution in that case. Having done that, we will show how the no slipping distribution relates to the no twisting distribution. Next we brievely discuss some problems with folliations before generelising the no slipping and the no twisting distribution to arbitrary semi-Riemannian manifolds.

3.1 Lightcone and Grassmanian

Before being able to look at rolling without slipping or twisting, we have to define two central objects which will help us analyze this problem. First of all, it will be usefull to collect all the null vectors in the tangent space at a point on a semi-Riemannian manifold. To this end we introduce the lightcone.

Definition 3.1. *Let (M, g) be a semi-Riemannian manifold and take a $p \in M$. We define the* ***lightcone*** *at p as:*

$$C_p(g) = \{v \in T_pM | |v| = 0\}$$

And we define the ***lightcone space*** *as the disjoint union of the lightcones at all point $p \in M$. More formally we write:*

$$C(g) = \sqcup_{p \in M} C_p(g) = \{(p, v) | p \in M \text{ and } v \in C_p(g)\}$$

As will become apparent throughtout this thesis, the lightcone and especially its subspaces will be the natural objects to look at when modelling the no slipping and no twisting conditions. So secondly, to formalise the discussion regarding the subspaces we also have to introduce the Grassmanian.

Definition 3.2. ***The Grassmanian*** *of a vector space V of dimension k is the collection of all k dimensional linear subspaces of V. More formally we write:*

$$Gr(V, k) = \{W \subseteq V | dim(W) = k \text{ and } W \text{ linear}\}$$

Remark. In the coming chapter we will usually talk about the Grassmanian of the lightcone and the lightcone space. Both of which are not vector spaces. The way one should understand the Grassmanian of the lightcone is as the following set:

$$Gr(C_p(g), k) = \{W \subseteq T_pM | dim(W) = k, \text{ W linear, and } W \subseteq C_p(g)\}$$

where M is a semi-Riemannian manifold and $p \in M$ is a point. To put this into words, we look at the Grassmanian of the vector space T_pM and pick out the linear subspaces which are contained in the lightcone at that point.

The Grassmanian of the lightcone space than just becomes a bundle over M and the construction should be understood to be pointwise. So when we talk about the Grassmanian of the lightcone space, we in essence have a disjoint union of Grassmanians of the lightcone at every point. △

It appears that the Grassmanian is a smooth manifold. We will just state the result without offering proof.

Lemma 3.3. *Let V be a vector space. The Grassmanian $Gr(V,k)$ is a smooth manifold.*

3.2 Rules of the Game

3.2.1 Configuration Space

Before we are able to model two manifolds rolling without slipping or twisting, we must define what it means to roll without slipping of twisting.

When two objects roll on each other, what they always need to have is at least one point of contact. This point of contact can be viewed as a way of identifying a point on the first manifold with a point on the second manifold. But it does actually more than only identifying the two points, it also has to identify some plane of contact between the two objects. This plane of contact tells us along which directions we can move such that our two objects still roll on each other.

More formally, suppose we have two semi-Riemannian manifolds M_1 and M_2 and they have a k-dimensional plane of contact at some point of contact $(p,q) \in M_1 \times M_2$. The way we can think about this plane of contact is that we have a k-dimensional linear subspace of T_pM_1 containing all the directions in which we can move such that M_1 still rolls on M_2. Analogously, we can find a k-dimensional linear subspace of T_qM_2 such that the same holds true for M_2. The additional structure we require is that we can somehow identify the k-dimensional linear subspaces of T_pM_1 and T_qM_2. The way in which we do this is by requiring there exists an isometry between these two k-dimensional subspaces.

This discussion gives rise to the notion of a configuration space.

Definition 3.4. *Let (M_1,g_1) and (M_2,g_2) be two Riemannian manifolds of dimension d_1 and d_2 respectively. The k-dimensional **configuration space** is defined as:*

$$Con_k(M_1,M_2) := \{(p,q,A) | p \in M_1, q \in M_2, L_1 \subseteq T_pM_1, L_2 \subseteq T_qM_2, A : L_1 \to L_2 \text{ isometry}\}$$

Where L_1 and L_2 are arbitrary k-dimensional subspaces and $k \leq \min\{d_1,d_2\}$.

Remark. Sometimes we will omit writing the manifolds when talking about the configuration space. In that case we write Con_k instead of $Con_k(M_1,M_2)$. △

There is an alternative way to talk about the configuration space, namely as the Grassmanian of the lightcone. This alternative approach will appear to be usefull when defining the no twisting distribution, but also when we generelize our discussion to arbitrary semi-Riemannian manifolds.

Lemma 3.5. *Let (M_1,g_1) and (M_2,g_2) be Riemannian manifolds. We define a semi-Riemannian manifold $(M_1 \times M_2, g = g_1 \oplus (-g_2))$. Then there exists a bijection between $Gr(C(g),k)$ and $Con_k(M_1,M_2)$.*

Proof. To this end we must show that given apoint $(p,q,A : L_1 \to L_2) \in Con_k(M_1,M_2)$ we can construct an associated element in $Gr(C(g),k)$. We must also show that given an element in $Gr(C(g),k)$ we can construct an isometry.
Let us start with the first one. Define the following set $(L_1, A(L_1)) := \{(w, A(w)) | w \in L_1\}$. Since A is an isometry, it is easy to see that the length of any element in $(L_1, A(L_1))$ is null. That this is k-dimensional should also be obvious becausee L_1 is k-dimensional. The linearity of the space follows from the linearity of the isometry. So given two elements $(v, A(v)), (w, A(w)) \in (L_1, A(L_1))$, we have that $(v, A(v)) + (w, A(w)) = (v + w, A(v) + A(w)) = (v + w, A(v + w))$. Since $v + w \in L_1$, we get the desired result.
Now we must show the converse. Let $\pi_1 : M_1 \times M_2 \to M_1$ and $\pi_2 : M_1 \times M_2 \to M_2$ be the projective maps to respectively the first and second component. Let $H \in Gr(C(g),k)$. What we will show is that the map $A : \pi_1(H) \to TM_2 : \pi_1(h) \to \pi_2(h)$ is an isometry. What this A does, is, given an $h = (h_1,h_2) \in H$, it maps the first element onto the second one. So we have $A(\pi_1(h)) = A(h_1) = h_2$. This is linear, because both projections are linear. So the only thing left

to show is isometry. To this end, note that we can write any element $h = h_1 + A(h_1)$. Since any elemt $h \in H$ is null we now have:

$$0 = g(h, \tilde{h}) = g(h_1 + A(h_1), \tilde{h}_1 + A(\tilde{h}_1)) = g_1(h_1, \tilde{h}_1) - g_2(A(h_1), A(\tilde{h}_1))$$

This shows us that $g_1(h_1, \tilde{h}_1) = g_2(A(h_1), A(\tilde{h}_1))$ and thus it is an isometry. □

3.2.2 No Slipping Distribution

Let us now move on to defining the no slipping condition. The picture one should have in mind when thinking about the no slipping condition is that of a runner on a treadmill. When increasing the speed of the treadmill, the runner should also increase his or her speed if they do not want to be launched off.

The essence of this all is that when two manifolds roll on each other and have a plane of contact of dimension k, we should be able to identify movements of the one manifold along the plane of contact with the movements of the other manifold along the plane of contact. This information is contained in the tangent bundle over the configuration space or $TCon_k$.

To be more precise, $TCon_k$ keeps track of two things. Firstly it contains all the information also contained in the tangent bundle over $M_1 \times M_2$. Secondly, it contains the information regarding how our A varies as the manifolds start rolling on each other. We will discuss this second point when talking about the no twisting condition. But the main problem is that if we take all of $TCon_k$ there is no relation between the rolling speed of manifold M_1 on manifold M_2 and the rolling speed of M_2 on M_1. The way to fix this is by letting the distribution be the preimage of $(p, q, graph(A))$.

By the use of lemma 3.5 we will define the no slipping distribution in terms of the Grassmanian of the lightcone space. Afther this we will show that it agrees with the above discussion where we used the more intuitive approach through the configuration space.

Definition 3.6. *Let (M_1, g_1) and (M_2, g_2) be two Riemannian manifolds and we construct its Cartesian product as $(M_1 \times M_2, g = g_1 \oplus (-g_2))$. We take a point $(p, q, H) \in Gr(C(g), k)$ in the Grassmanian of the lightcone space. Now we define the* ***no slipping distribution*** *at (p, q, H) to be:*

$$\mathcal{D}^{NS}(p, q, H) := d\pi^{-1}(p, q, H) \subseteq T_{(p,q,H)}Gr(C(g), k)$$

where $\pi : Gr(C(g), k) \to M_1 \times M_2$ is the projection.

Lemma 3.7. *In configuration space, the no slipping distribution at a point $(p, q, A : L_1 \to L_2) \in Con_k(M_1, M_2)$ is defined as:*

$$\mathcal{D}^{NS}(p, q, A) := d\pi^{-1}(p, q, graph(A)) \subseteq T_{(p,q,A)}Con_k$$

where $graph(A) = \{(v, A(v))|v \in L_1\}$ and where now the projection is $\pi : Con_k \to M_1 \times M_2$.

Proof. This lemma just amounts to an application of lemma 3.5. Since by the proof of that lemma we can view the graph of A as a null plane and we can view any null plane as the graph of some isometry, the result readily follows. □

3.2.3 No Twisting Distribution

Remark. To make the definitions and lemmata shorter, throughout this subsection we assume that (M_1, g_1) and (M_2, g_2) are Riemannian manifolds and from them we construct a semi-Riemannian manifold $(M_1 \times M_2, g = g_1 \oplus (-g_2))$ which we endow with the Levi-Civita connection ∇. Furthermore we suppose that H is a k-dimensional null plane at the point (p, q) or to put it formally $H \in Gr(C_{(p,q)}(g), k)$. Lastly we let $v \in H$ be the rolling direction of the two manifolds and we let $\gamma : I \to M_1 \times M_2$ be the geodesic such that we have $\gamma(0) = (p, q)$ and $\dot{\gamma}(0) = v$. △

Let us now move on to the no twisting condition. The reason for this condition is given to us by physical consideration. When looking at two objects rolling on each other, the rolling of one object determines uniquely the motion of the second object. Think about two gears for example.

When one moves clockwise, the other has to move counterclockwise. Another way to think about the no twisting condition is as saying that in nature, if no external force is applied, objects do not suddenly reorient themselves. Stated in this way, the no twisting condition boils down to Newton's first law.

In the context of distributions, what this means is that we want the evolution of the plane of contact to be uniquely determined. This is not the case in the no slipping distribution since there is no relationship between the rolling direction and the time evolution of the plane of contact.

As stated, in the real physical world we know that this is not the case. So we need to construct a distribution which relates the time evolution of the plane of contact with the rolling direction of the manifold. But before we are able to do this, we must develop some mathematical formalism.

So what we want to do now is define what it means precisely for our plane of contact to evolve as the two manifolds roll on top of each other.

Definition 3.8. *Let $(E_i)_{i=1}^k$ be a basis for our null plane $H \in Gr(C_{(p,q)}(g), k)$. By the* ***parallel transport of H*** *along γ, we mean the following linear subspace:*

$$H_v(t) = span\{\tilde{E}_i(\gamma(t)) | i \in \mathbb{N}_k\} \subseteq T_{\gamma(t)}(M_1 \times M_2)$$

Where $\tilde{E}_i$ is the parallel vector field along γ satisfying $\tilde{E}_i(\gamma(0)) = E_i$.

So what this definition tells us is suppose our two manifolds $M_1 \times M_2$ touch each other at some point (p, q) and have a rolling direction v along the plane of contact H. We are able to find a unique geodesic γ such that $\gamma(0) = (p, q)$ and $\dot{\gamma}(0) = v$. What we want is that our plane of contact H evolves in some nice way along the geodesic. Well, what is the nicest way possible? The answer is by parallel transport. So we parallel transport all the vectors E_i spanning H along γ and we define this as $\tilde{E}_i$. Then we define the parallel transport of our plane H at some time t to be the span of vector fields $\tilde{E}_i(\gamma(t))$ evaluated at the point $\gamma(t)$

A very nice and important property is that the parallel transport preserves the fact that our plane of contact is null. The reason that this is nice, is that it allows us to view $H_v(t)$ as a curve in $Gr(C(g), k)$. So first let us prove that this is indeed the case.

Lemma 3.9. *The parallel transport of H along γ is also a null plane. To be more formal we have that $H_v(t) \in Gr(C_{\gamma(t)}(g), k)$ for all times $t \in I$.*

Proof. We know that we can write $H_v(t) = span\{\tilde{E}_i(\gamma(t)) | i \in \mathbb{N}_k\}$. Since the Levi-Civita connection is compatible with the metric, we get:

$$\nabla_{\dot{\gamma}} g(\tilde{E}_i, \tilde{E}_i) = 2g(\nabla_{\dot{\gamma}} \tilde{E}_i, \tilde{E}_i)$$

But we know that $\nabla_{\dot{\gamma}} \tilde{E}_i = 0$ because $\tilde{E}_i$ is a parallel vector field along γ. Thus we now get that:

$$\nabla_{\dot{\gamma}} g(\tilde{E}_i, \tilde{E}_i) = 2g(0, \tilde{E}_i) = 0$$

Since $g(E_i, E_i) = 0$, we have show that $g(\tilde{E}_i, \tilde{E}_i) = 0$. This proves that $H_v(t)$ is a null planes for all times $t \in I$. □

As said, what this lemma tells us is that for all time t on which our geodesic is defined, we have that $H_v(t)$ remains a null plane. Thus this allows us to lift the geodesic to a curve in $Gr(C(g), k)$.

Definition 3.10. *We define* ***the lift of*** *γ by H to be the curve:*

$$\gamma \times H_v : I \to Gr(C(g), k) : t \to (\gamma(t), H_\gamma(t))$$

We are now in a position to define what it means to take the time-derivative of our lift at $t = 0$.

Definition 3.11. *We define the* ***time derivative at*** *$t = 0$* ***of the lift of*** *γ* ***by*** *H as:*

$$\frac{\partial}{\partial t}(\gamma \times H_v)(0) := (\dot{\gamma}(0), \frac{\partial}{\partial t} H_v(0)) \in T_{(p,q,H)} Gr(C(g), k)$$

Lemma 3.12. *The time derivative at $t = 0$ of the lift of γ by H depends linearly only on the value of $\dot{\gamma}(0)$.*

Proof. By lemma 2.35 it follows that parallel transporting sends linear subspaces to linear subspaces. This also tells us that the value for $H_v(t)$ does not depend on the basis we pick for H. Hence it derivative will also not depend on the choice of basis.
The way in which we construct the parallel transport of the plane, is by parallel transporting all of its basis elements. So the vector fields $\tilde{E}_i$ satisfy that $\nabla_{\dot{\gamma}}\tilde{E}_i = 0$. We are looking at this at the point $\gamma(0)$ i.e. we are looking at $(\nabla_{\dot{\gamma}}\tilde{E}_i)(\gamma(0))$. Now we know by lemma 2.29 that this only depends on the value $\dot{\gamma}(0)$. Now since the linear connection is linear, we see that this depends too is linear. $\square$

As a corollary to this lemma, we can show the following:

Corollary 3.13. *For any two curves γ, ν, we have that the time derivative of their lift by H satisfies:*

$$\frac{\partial}{\partial t}(\gamma \times H_v)(0) + \frac{\partial}{\partial t}(\nu \times H_w)(0) = \frac{\partial}{\partial t}((\gamma + \nu) \times H_{v+w})(0)$$

where $\dot{\gamma}(0) = v$ and $\dot{\nu}(0) = w$.

Proof. Since the time derivative at $t = 0$ of any curve γ by H depends linearly only on the value $\dot{\gamma}(0)$, the result follows trivially. $\square$

So what we wanted to do was to construct an object which allows us to relate the rolling direction of the manifold with the time evolution of the plane of contact. In other words, we want to contruct a lift from elements in our nullplane (p, q, H) to elements in $T_{(p,q,H)}Gr(G(g), k)$. So let us construct this lift:

Definition 3.14. *The **lift at** (p, q) **of** H **to** $TGr(C(g), k)$ is defined as the map:*

$$B_{(p,q,H)} : H \to T_{(p,q,H)}Gr(C(g), k) : v \to B_{(p,q,H)}(v) = (\dot{\gamma}(0), \frac{\partial}{\partial t}H_\gamma(0))$$

where $\dot{\gamma}(0) = v$.

Having defined all the ingredients necessary, we are finally able to define the no twisting distribution.

Definition 3.15. *Let $B_{(p,q,H)}$ be the lift of H to $TGr(C(g), k)$. Then we define the **no twisting distribution** at the point (p, q, H) as:*

$$\mathcal{D}^{NT}(p, q, H) = im(B_{(p,q,H)})$$

where $im(B_{(p,q,H)})$ is the image of $B_{(p,q,H)}$.

To show that the no twisting distribution is indeed a distribution we have to show the following lemma.

Lemma 3.16. *Suppose $(p, q, H) \in Gr(C(g), k)$. The lift $B_{(p,q,H)}$ is linear and injective.*

Proof. Suppose two elements $v, w \in H$. Let us look at:

$$\begin{aligned}
B_{(p,q,H)}(v) + B_{(p,q,H)}(w) &= (v, \frac{\partial}{\partial t}H_v(0)) + (w, \frac{\partial}{\partial t}H_w(0)) \\
&= (v + w, \frac{\partial}{\partial t}H_v(0) + \frac{\partial}{\partial t}H_w(0)) \\
&= (v + w, \frac{\partial}{\partial t}H_{v+w}(0)) \\
&= B_{(p,q,H)}(v + w)
\end{aligned}$$

That it is injective follows trivially from the fact that is $B_{(p,q,H)}(v) = B_{(p,q,H)}(w)$, then $(v, \frac{\partial}{\partial t}H_v(0)) = (w, \frac{\partial}{\partial t}H_w(0))$ and thus $v = w$. $\square$

Now we are able to prove that the no twisting distribution is a distribution.

Lemma 3.17. *The no twisting distribution is in fact a distribution.*

Proof. By lemma 3.16 we have that the distribution is linear. Since, by the same lemma, the map is linear and injective, it follows that the kernel of the lift is trivial and hence the no twisting distribution is non-trivial.
So now the only thing left to prove is that the distribution is smooth. Now since the no twisting distribution is defined as the image of the lift, we have that the restriction of the lift to $B_{(p,q,H)} : (p,q,H) \to \mathcal{D}^{NT}(p,q,H)$ is a linear bijection and hence smooth. Since we have a global frame for H, we can lift of this frame to form a global frame for $\mathcal{D}^{NT}(p,q,H)$. Since our lift is smooth, it follows that the distribution is smooth. □

As a corollary to this lemma one should note now that the rank of the no slipping distribution is equal to the dimension of the plane of contact.

Corollary 3.18. *The rank of the no slipping distribution at some point $(p,q,H) \in Gr(C(g),k)$ is equal to the dimension of H i.e. $rank(D^{NT}(p,q,H)) = dim(H) = k$.*

Proof. Since $B_{(p,q,H)} : H \to \mathcal{D}^{NT}(p,q,H)$ is a bijection by the above lemma, the result follows. □

More intuitively, the no twisting distribution tells us that, while rolling, we do not allow for reorientation. As an example think about biking. What normal people do while biking and they have to make a right or left turn, is they smoothly turn their steering wheel in the direction they have to go in. The no twisting distribution would not allow this. You can only move in a straight line.

3.2.4 An alternative approach to the no twisting distribution

In broad strokes, what we did to define the no twisting distribution is the following: we were given a plane of contact H and a rolling direction $v \in H$. We integrated the rolling direction to a geodesic γ. Then we parallel transported H along γ. Lastly we took the time derivative of the parallel transport of H along γ at the time $t = 0$ to lift the rolling direction v to a direction in $T_{(p,q,H)}Gr(C(g),k)$.

We could also approach this differently. We are given an ODE for the parallel transport of the plane H, namely:

$$\dot{\gamma}(\tilde{E}_i^l) = -\dot{\gamma}^j \tilde{E}_i^k \Gamma_{jk}^l \tag{5}$$

Now what we could also have done, is solve this above ODE formally up to order one. This is possible due to lemma 2.29. So what this lemma tells us is that, firstly, the time derivative of the parallel transport at $t = 0$ does not depend on the choice of curve. We could have picked any curve γ, not necessarily a geodesic, satisfying that $\dot{\gamma}(0) = v$ where v is the rolling direction.

Secondly, as remarked after lemma 2.31, the information contained in $\frac{\partial}{\partial t} H_v(0)$ depends only on first order data. This information we are given since we know that all $\tilde{E}_i(p,q)$ together span the plane of contact H. This means that this is a linear ODE which we can solve formally up to first order. Let us make it clear what we mean by solving an ODE formally up to first order through the use of an example:

Example 3.19. Suppose we have a function $f \in C^\infty(\mathbb{R})$ which satisfies $f' = f$. Now we can make an ansatz by taking $f = \sum_{n=0}^{\infty} a_n x^n$ where $a_n \in \mathbb{R}$ are constants. This would then give us $\sum_{n=0}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} a_n x^n$. This then gives us $na_n = a_{n-1}$ and hence we get the solution $a_n = \frac{1}{n!} a_0$. ♢

We can apply the same process to the equation (5) to get an explicit value for $\dot{\gamma}(\tilde{E}_i^l)$. We write the formal solution as f_i^l which then satisifes:

$$\dot{\gamma}(f_i^l)(p,q) = -f_i^l(p,q)\dot{\gamma}^k(p)\Gamma_{jk}^l(p)$$

Now of course f_i^l are explicitely given and so we do not need to introduce the lift to define the no twisting distribution. We can let the no twisting distribution at some point be just the planes spanned by these explicit solutions. Let us write this formally:

Definition 3.20. *Alternatively, we can define the* ***no twisting distribution*** *at some point* $(p, q, H) \in Gr(C(g), k)$ *as:*

$$\mathcal{D}^{NT}(p, q, H) = \{(v, \frac{\partial}{\partial t} H_v(0)) | v \in H\}$$

where now $\frac{\partial}{\partial t} H_v(0)$ *are given by their formal solution.*

Let us prove that this definition is equivalent.

Lemma 3.21. *The no twisting distribution defined by definition 3.15 and definition 3.20 are equivalent*

Proof. The formal solutions do not give us different planes $\frac{\partial}{\partial t} H_v(0)$. It just gives them in a more explicit way. □

3.3 Example: The Rolling Coin on a Table

Let us now turn our attention to an example: a coin rolling without slipping and twisting on a table. The easiest model for this situration is $\mathbb{S}^1 \times \mathbb{R}^2$. In this case $\mathbb{S}^1$ tells us which point of the coin is in touch with the surface and $\mathbb{R}^2$ tells us where it is on the table. We furthermore endow it with the metric $g = d\alpha^2 - dx^2 - dy^2$, where we take α to be the coordinate of $\mathbb{S}^1$ and (x, y) the coordinates on $\mathbb{R}^2$. So moving an infinitesimal distant along the surface will be a timelike movement and along the coin will be a spacelike movement.

We note that in this situation the plane of contact is one-dimensional. This should be clear due to the fact that the path of the coin is one-dimensional. So we are interested in the one-dimensional configuration space. We have show in the chapter regarding the no twisting distribution that there exists an isomorphism between the configuration space and the one-dimensional Grassmanian of the lightcone space. Let us state it again explicitely for this case.

Lemma 3.22. *There exists a bijection between one-dimensional Grassmanian of the lightcone space* $Gr(C(g), 1)$ *and the one-dimensional configuration space* $Con_1(\mathbb{S}^1, \mathbb{R}^2)$. .

To be able to analyze $Gr(C(g), 1)$, we must first know how the lightcone looks like.

Lemma 3.23. *The lightcone* $C_p(g)$ *is isomorphic to* $\{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 = z^2\}$.

Proof. Take a vector $v \in C_p(g)$. Let $\partial_\alpha, \partial_x, \partial_y$ be a local basis for $T_p(\mathbb{S}^1 \times \mathbb{R}^2)$. This means that we can write $v = v_\alpha \partial_\alpha + v_x \partial_x + v_y \partial_y$. Now we know that points on the lightcone are parametrized by having length 0. This implies that $g(v, v) = d\alpha^2(v, v) - dx^2(v, v) - dy^2(v, v) = 0$. Hence we get that $v_\alpha^2 = v_x^2 + v_y^2$. □

What this lemma tells us is that the lightcone, when viewed in $\mathbb{R}^3$, is the regular cone. Now it should be clear that the only one-dimensional linear subspaces of the lightcone are the straight lines through the origin. In other words an element $L \in Gr(C(g), 1)$ can be represented by taking the span of some vector $v \in C_p(g)$, or more formally $L = \langle v \rangle$. What should also be clear is that we can write any element $v \in C_p(g)$ as $v = v_\alpha cos(\theta)\partial_x + v_\alpha sin(\theta)\partial_y + v_\alpha \partial_z$. So then we can write $L = \langle cos(\theta)\partial_x + sin(\theta)\partial_y + \partial_z \rangle$

The intuitive way to think about it is as follows: if we fix an element $L \in Gr(C_p(g), 1)$, we actually fix in which direction on the table the coin is it going to move. The speed of this movement is not fixed since we can multiply any element $v \in L$ by a scaler and still get an element in L back.
So now we can phrase the main question of this subsection: After our coin starts rolling along a line L_1, is it allowed to reorient itself and roll along another line L_2? Or in other words, can our coin change rolling direction? The next theorem tells us yes.

Theorem 3.24. *Let $\pi : T(\mathbb{S}^1 \times \mathbb{R}^2) \to \mathbb{S}^1 \times \mathbb{R}^2$ be the projection and $\rho : C(g) \to T(\mathbb{S}^1 \times \mathbb{R}^2)$ be the inclusion. The distribution $\mathcal{D}(p, v) := d(\pi \circ \rho)^{-1}(p, \langle v\rangle)$, where $(p, v) \in \rho(C(g))$, models the no slipping condition and can be written as:*

$$\mathcal{D}(p, v) = (p, L, L, 0) \oplus (p, L, 0, \langle -sin(\theta)\partial_x + cos(\theta)\partial_y\rangle) \oplus (p, L, 0, L)$$

where $L = \langle v\rangle = \langle cos(\theta)\partial_x + sin(\theta)\partial_y + \partial_z\rangle$ for some fixed $\theta \in [0, 2\pi\rangle$.

Remark. First of all we note that $(p, L, L, 0)$ tells us is that we keep on moving in the direction we are moving in without acceleration.
Secondly we note that the second term accounts for a reorientation of the coin. What it tells us is that the null line L is moving in a direction orthogonal to L and it is not rolling.
The last term is just an acceleration along L. So what this tells us is that the coin is speeding up or slowing down.

Proof. First of all, we note by lemma 3.22 we have that we can realise $\langle v\rangle$ as a graph of some isometry A. So by definition we get that $\mathcal{D}(p, v)$ models no slipping.
Secondly it should be clear that $(p, L, L, 0)\oplus(p, L, 0, \langle -sin(\theta)\partial_x+cos(\theta)\partial_y\rangle)\oplus(p, L, 0, L) \subseteq \mathcal{D}(p, v)$.
Let us now look at an element $x \in \mathcal{D}(p, v)$. This can locally be written in the shape $x = (p, w, u, b)$. $p \in \mathbb{S}^1 \times \mathbb{R}^2$ is the point of contact between the coin and the table. $w \in L \subseteq C_p(g)$ is an element in the plane of contact which in our case is just the line L. Now these two points are in some sense fixed since the point and plane of contact are given to us. The next element is $u \in T_pM$. This element tells us how p evolves. This element is actually also fixed due to the fact that we are only allowed to move along the plane of contact. It has actually even less degrees of freedom than that since we have already picked a point $(p, w) \in \rho(C(g))$ and hence know how p is going to evolve namely along w. So this gives us $u = w$. The last element has no restrictions. We are free to accelerate any way we want tangent to $TC(g)$.
What this discussion tells us is that $\mathcal{D}(p, v)$ has three degrees of freedom. The first one is in its choice for $w \in L$ and the other two are in the choice of acceleration. Since $(p, L, L, 0) \oplus (p, L, 0, \langle -sin(\theta)\partial_x + cos(\theta)\partial_y\rangle) \oplus (p, L, 0, L)$ has three dimensions, it follows that $\mathcal{D}(p, v) = (p, L, L, 0) \oplus (p, L, 0, \langle -sin(\theta)\partial_x + cos(\theta)\partial_y\rangle) \oplus (p, L, 0, L)$. □

What is also worthwhile looking at is the no twisting distribution.

Theorem 3.25. *The no twisting distribution in the case of the rolling coin is given by:*

$$\mathcal{D}^{NT}(p, L) = (p, L, L, 0)$$

where again we write $L = \langle v\rangle = \langle cos(\theta)\partial_x + sin(\theta)\partial_y + \partial_z\rangle$ for some fixed $\theta \in [0, 2\pi\rangle$.

Proof. Let $B_{(p,L)}$ be the lift such that $\mathcal{D}^{NT}(p, L) = im(B_{(p,L)})$. Now let $\gamma : I \to \mathbb{S}^1 \times \mathbb{R}^2$ be the geodesic such that $\gamma(0) = p$ and $\dot\gamma(0) = v$. Since this is a geodesic we have that $\nabla_{\dot\gamma}\dot\gamma = 0$. So by theorem 2.33 $\dot\gamma$ is the unique vector field parallel to γ and satisfying $\dot\gamma(0) = v$. What this tells us is that we can realise the parallel translate of L as $L_v(t) = \langle\dot\gamma(t)\rangle$. Now since γ is a geodesic, we have that its acceleration is 0 and hence we have that $\frac{\partial}{\partial t}L_v(0) = 0$. So this gives us for the lift $B_{(p,L)}(L) = (L, 0)$. With this we have proven that $\mathcal{D}^{NT}(p, L) = (p, L, L, 0)$ □

3.4 The Cartesian Product

Remark. Throughout this subsection we have two Riemannian manifolds (M_i, g_i) of dimension d_i, where we take $d_1 \le d_2$, and we look at the cartesian product of these two manifolds $(M_1 \times M_2, g = g_1 \oplus (-g_2))$. Since the dimension of the plane of contact k can only be as big as the dimension of the smaller manifold, we assume $k \le d_1$. △

We should now be in a position to produce the distribution of rolling without slipping and twisting of a cartesian product of two Riemannian manifolds. This subsection will mimic the previous subsection regarding the coin rolling on a table.

By lemma 3.5 we know that we can identify a point in the configuration space with a linear subspace of the lightcone.

Let us now prove the main result of this thesis.

Theorem 3.26. *Suppose we have a point $(p, q) \in M_1 \times M_2$. Let $\pi : C(g) \to M_1 \times M_2$ be the projection and take $(p, q, H) \in Gr(C(g), k)$. Then the distribution $\mathcal{D}^{NS}(p, q, H) := d\pi_p^{-1}(p, q, H)$ models the no slipping condition and can be written as:*

$$\mathcal{D}^{NS}(p, q, H) = \mathcal{D}^{NT}(p, q, H) \oplus T_H Gr(C_{(p,q)}(g), k)$$

Where $\mathcal{D}^{NT}(p, H)$ is the no twisting distribution.

Remark. This notation is a bit problematic since $\mathcal{D}^{NT}(p, q, H)$ lives in $T_{(p,q,H)}Gr(C(g), k)$ which is a larger space than $T_H Gr(C_{(p,q)}(g), k)$. It is the same difference as between the tangent bundle and the tangent space. The way to understand this notation is as follows: given a rolling direction $v \in H$, the way in which we defined the no twisting distribution is that we lift this rolling direction to an element in $T_{(p,q,H)}Gr(C(g), k)$. To put it in other words, what we do is $B_{(p,q,H)}(v) = (v, \frac{\partial}{\partial t} H_v(0))$. As discussed, what this does is that it establishes a relationship between the time evolution of the plane of contact and the rolling direction. What this notation aims to communicate is that this relationship gets broken in the no slippin distribution. To put it differently, an element $x \in \mathcal{D}^{NS}$ can be written as $(v, \frac{\partial}{\partial t} H_w(0))$ where $v, w \in H$ are not necessarily the same. △

Proof. We have by definition that $\mathcal{D}^{NS}$ models no slipping. Next it should also be clear that $\mathcal{D}^{NT}(p, q, H) \oplus T_H Gr(C_{(p,q)}(g), k) \subseteq d\pi_p^{-1}(p, q, H) = \mathcal{D}^{NS}(p, q, H)$.
Now we are going to make a similar argument as in the case of the rolling coin. We know that the no slipping distribution is a subbundle of $TGr(C(g), k)$. This means that an element in $TGr(C(g), k)$ has five components. The first three tell us at which point we are in $Gr(C(g), k)$. These are given to us since we are modelling the no slipping distribution at the point (p, q, H). The forth element is the rolling direction. This is restricted due to the fact that it must be tangent to the null plance H. The last component is the acceleration of H. There we have no restrictions on. By the above remark, this is precisely the information contained in $\mathcal{D}^{NT}(p, q, H) \oplus T_H Gr(C_{(p,q)}(g), k)$ and hence the result follows. □

3.5 Troubles with Foliations

Now that we have discussed the case of two Riemannian manifolds rolling on each other, we can move on to the general case of a semi-Riemannian manifold rolling without slipping and twisting. Now one could ask: what does it mean if one manifold is rolling without slipping and twisting? And we see that this is a place where intuition leaves us, but one should have the idea of two Riemannian manifolds in mind. There it should be clear that we have one manifold which defines some positive distance and the other defines the negative distance and they roll on top of each other.

In the setting of semi-Riemannian geometry, this might not always be the case. The problem lies in the fact that in a arbitrary semi-Riemannian manifold it is not always possible to choose coordinates such that the timelike foliation is orthogonal to the spacelike foliation. To illustrate this point we will look at an example.

Example 3.27. Let us look at $(\mathbb{D}^2_1 \times \mathbb{D}^2_2, g = dx_1^2 + dx_3^2 + e^{x_1} dx_2^2 - e^{x_3} dx_4^2 - dx_1 dx_4)$, where $\mathbb{D}^2$ is the two-dimensional disk. Furthermore we have that (x_1, x_2) are the coordinates on $\mathbb{D}^2_1$ and (x_3, x_4) are the coordinates on $\mathbb{D}^2_2$.
So it should be clear that $\langle \partial_1, \partial_3 \rangle$ spans a spacelike foliation at every point on our disks. It is possible to calculate that the orthogonal to this foliation is $\langle \partial_2, \partial_1 + \partial_4 \rangle$. This though is not a timelike foliation since clearly $g(\partial_2, \partial_2) = e^{x_1}$ which is positive for all x_1. ♢

What this example aims to make clear is that rolling does not happen between a timelike part rolling on a spacelike part, but that two orthogonal spaces roll on each other. In the context of the Cartesian product, this simply was the timelike component and the spacelike component.

Now it should be noted that this is not exactly a counterexample. It is possible that there exists some other choice for coordinates which do give us a timelike foliation with a spacelike foliation as its orthogonal. But even if such a coordinate choice exists, it is not guaranteed that we can totally seperate the spacelike directions of our manifold from time which was done by construction in our discussion of the cartesian product of two Riemannian manifold. This next example should make this clear.

Example 3.28. Suppose we have $(\mathbb{S}^1 \times \mathbb{R}, g = e^t ds^2 - dt^2)$ a semi-Riemannian manifold where $\mathbb{S}^1$ is spacelike and $\mathbb{R}$ is timelike. We will calculate the curvature tensor. Locally we can have a coordinate frame ∂_t, ∂_s. Now the metric in matrix form is given by:

$$g_{ij} = \begin{pmatrix} e^t & 0 \\ 0 & -1 \end{pmatrix} \qquad g^{ij} = \begin{pmatrix} e^{-t} & 0 \\ 0 & -1 \end{pmatrix}$$

Where the first matrix is the metric and the second matrix is the inverse of the first. The i and j can be replaced for s or t. So $g_{ss} = e^t$ for example. Now this allows us to compute the Christoffel symbols using equation (4). This now gives us the following:

$$\Gamma^s = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \qquad \Gamma^t = \begin{pmatrix} e^t & 0 \\ 0 & 0 \end{pmatrix}$$

Now we are in a position to compute the curvature tensor $R(\partial_s, \partial_t, \partial_s)$. We first of all note that the Lie bracket of any coordinate charts is zero. This simplifies the computations to:

$$R(\partial_s, \partial_t, \partial_s) = \nabla_{\partial_s} \nabla_{\partial_t} \partial_s - \nabla_{\partial_t} \nabla_{\partial_s} \partial_s$$

Applying the definition of the Christoffel symbols gives us:

$$\nabla_{\partial_s} \nabla_{\partial_t} \partial_s = \frac{1}{4} e^t \partial_t, \qquad \nabla_{\partial_t} \nabla_{\partial_s} \partial_s = \frac{1}{2} e^t \partial_t$$

So this gives us that $R(\partial_s, \partial_t, \partial_s) = -\frac{1}{4} e^t \partial_t$. This does not vanish for any $t \in \mathbb{R}$ and so our semi-Riemannian manifold is not locally isometric to the euclidean space. This implies that it is impossible for us to find coordinates such that the spacelike component of the metric does not depend on the time. ♢

The takeaway should be this: The intuition from the case of the cartesian product of two Riemannian manifold, where we have that the manifold which defines the positive distance rolls on the manifold which defines the negative distance, is not in general true since it might not always be possible to seperate our semi-Riemannian manifold in this way.

3.6 The General Case

So in the previous subsections we have discussed the case where two objects are rolling on top of each other. Of course, when looking at rolling, it is intuitively more clear what is happening when two objects roll on each other, but actually there is nothing which stops us from saying that an arbitrary semi-Riemannian manifold is "rolling on itself". So in this subsection, we will discuss what we mean by this.

We can actually reinterpret the Cartesian product case as follows: instead of talking about two Riemannian manifolds rolling on each other, we can talk about an internal motion in the space of cartesian product of these two Riemannian manifold. To put it differently, the motion taking place is not so much between two objects, but actually within the one greater space containing these two objects. In that sense one can talk about the Cartesian product of these two Riemannian manifolds rolling on itself. In the same way, we can extend this definition to any semi-Riemannian manifold rolling on itself.

This is the intuition behind the rolling motion of an arbitrary semi-Riemannian manifold. To make this mathematically rigorous, we actually have to change nothing in the above discussion. The point of contact just becomes an arbitrary point in the semi-Riemannian manifold. Furthermore as discussed, we can represent configuration space in terms of the Grassmanian of the lightcone $Gr(C(g), k)$. This object is well defined in the case of an arbitrary semi-Riemannian manifold. So what this tells us is that the definition of the no slipping distribution and the no twisting distribution, which we have already framed in terms of null planes, actually still holds true in the case of an arbitrary semi-Riemannian manifold. As a formality, let us state the definitions.

Definition 3.29. *Let (M, g) be a semi-Riemannian manifolds. We take a point $(p, H) \in Gr(C(g), k)$ in the Grassmanian of the lightcone space. Now we define the* ***no slipping distribution*** *at (p, H) to be:*

$$\mathcal{D}^{NS}(p, H) := d\pi^{-1}(p, H) \subseteq T_{(p,H)}Gr(C(g), k)$$

where $\pi : Gr(C(g), k) \to M$ is the projection.

Definition 3.30. *Let $B_{(p,H)}$ be the lift of H to $TGr(C(g), k)$. Then we define the* ***no twisting distribution*** *at the point (p, H) as:*

$$\mathcal{D}^{NT}(p, H) = im(B_{(p,H)})$$

where $im(B_{(p,H)})$ is the image of $B_{(p,q,H)}$.

The main theorem of this chapter and the proof offered for it, actually still holds true in this case. So we just it again without offering proof.

Theorem 3.31. *Suppose we have a point $p \in M$. Let $\pi : C(g) \to M$ be the projection and take $(p, H) \in Gr(C(g), k)$. Then the distribution $\mathcal{D}^{NS}(p, H) := d\pi_p^{-1}(p, H)$ models the no slipping condition and can be written as:*

$$\mathcal{D}^{NS}(p, H) = \mathcal{D}^{NT}(p, H) \oplus T_H Gr(C_p(g), k)$$

Where $\mathcal{D}^{NT}(p, H)$ is the no twisting distribution.

4 Control Systems and The No Slipping Distribution

In this chapter, we again take a look at the no slipping distribution. The main goal is to show that the no slipping distribution is bracket generating. Being bracket generating is of course a property inherent to the distribution. Yet, in the case of the no slipping distribution being bracket generating is a consequence of the geometric properties of the lightcone. The no slipping distribution over the lightcone space is an example of a tautological distribution over a control system. Control systems are (not necessarily linear) subbundles of the tangent bundle. The main result of this chapter is showing that if a control system is suffciently curved, then the tautological distribution over that control system is bracket generating.

To this end we first must define what it means for a control system to be sufficiently curved. This definition is based on higher order derivatives of a smooth map between a manifold and a vector space. So we start out by defining this higher order derivatives before moving on to formally define control systems, tautological distributions and sufficiently curved control systems. To finish this thesis, we show the main result of this chapter and prove that the lightcone space is sufficiently curved implying that the no slipping distribution is bracket generating.

4.1 Hessian and Higher Order Derivatives

It is possible to generalize the Hessian to smooth manifolds.

Definition 4.1. *Let M be an n-dimensional smooth manifold and V a n-dimensional vector space. Let furthermore $p \in M$ be a point and $(x^i)_{i=1}^n$ a local coordinate chart for some open U containing p. Lastly let $\rho : M \to V$ be a smooth map. Then we define the **Hessian** at p as the map:*

$$H_p\rho : T_pM \times T_pM \to V/(d_p\rho(T_pM)) : (X_1, X_2) \to H_p\rho(X_1, X_2) = \frac{\partial^2 \rho}{\partial x^i \partial x^j}(p)(dx^i \otimes dx^j)(X_1, X_2)$$

*We also define the **global Hessian** as the map:*

$$H\rho : p \in M \to H_p\rho$$

Now one could ask: Why do we quotient out $(d_p\rho(T_pM))$ from V? This has to do with the fact that we want the Hessian to be independent of our choice of parametrization. Let us make this more formal.

Lemma 4.2. *The Hessian at a point $p \in M$ is well-defined in the sense that it does not depend on the parametrization of both M and V.*

Proof. Let $\phi : M \to M$ be a diffeomorphism and $f : V \to V$ an isomorphism of vector spaces. Let $(x^i)_{i=1}^n$ be a local coordinate chart for some open U containing p. We define $\tilde{x}_i = \phi^{-1}(x_i)$ to form a coordinate chart for $\phi^{-1}(U)$ containing $\phi^{-1}(p)$. Let us look at:

$$H_{\phi^{-1}(p)}(f \circ \rho \circ \phi) = \frac{\partial^2 (f \circ \rho \circ \phi)}{\partial \tilde{x}^i \partial \tilde{x}^j}(\phi^{-1}(p)) d\tilde{x}^i \otimes d\tilde{x}^j$$

So first of all, we note that f does not depend on and $\tilde{x}^i$ and hence we can take it out of the partial derivation. This proves that the Hessian does not depend on the choice of parametrization as a vector space of V.
Secondly, what we are going to do now is a simple application of the chain rule for differentiation:

$$\frac{\partial^2 (f \circ \rho \circ \phi)}{\partial \tilde{x}^i \partial \tilde{x}^j}(\phi^{-1}(p)) = (f \circ \frac{\partial^2 \rho}{\partial x^l \partial x^k})(p) \frac{\partial x^l}{\partial \tilde{x}^i}(\phi^{-1}(p)) \frac{\partial x^k}{\partial \tilde{x}^j}(\phi^{-1}(p)) + (f \circ \frac{\partial \rho}{\partial x^l})(p) \frac{\partial^2 x^l}{\partial \tilde{x}^i \partial \tilde{x}^j}(\phi^{-1}(p))$$

We notice that $\frac{\partial \rho}{\partial x^l}(p) = 0$ since we quotient out the image of the differential from the range of the Hessian. Hence we get:

$$H_{\phi^{-1}(p)}(f \circ \rho \circ \phi) = (f \circ \frac{\partial^2 \rho}{\partial x^l \partial x^k})(p) \frac{\partial x^l}{\partial \tilde{x}^i}(\phi^{-1}(p)) \frac{\partial x^k}{\partial \tilde{x}^j}(\phi^{-1}(p)) d\tilde{x}^i \otimes d\tilde{x}^j$$

What we also note is:

$$d\tilde{x}^i = \frac{\partial \tilde{x}^i}{\partial x^l}(\phi^{-1}(p))dx^l$$

This then gives us:

$$H_{\phi^{-1}(p)}(f \circ \rho \circ \phi) = (f \circ \frac{\partial^2 \rho}{\partial x^l \partial x^k})(p)dx^l \otimes dx^k = f \circ H_p\rho$$

□

Remark. We have defined the Hessian at a point by the use of coordinates. It is also possible to avoid this. To this end we define $\Phi_v : T_vV \to V$ and $\Psi_{(p,w)} : T_wT_pM \to T_pM$ to be the isomorphisms form lemma 6.1. We would then define the Hessian at p as:

$$H_p\rho : (X_1, X_2) \to H_p\rho(X_1, X_2) = \Phi_{\rho(p)} \circ d_{X_2}(\Phi_{\rho(p)} \circ d_p\rho)(\Psi^{-1}_{(p,X_2)}(X_1))$$

We have not changed the domain and the range of the Hessian.

Following the steps in the above proof, but writing them in the coordinate free way, gives us as a result that:

$$H_{\phi^{-1}(p)}(f \circ \rho \circ \phi)(X_1, X_2) = f \circ (H_p\rho)(d_{\phi^{-1}(p)}\phi(X_1), d_{\phi^{-1}(p)}\phi(X_2))$$

where again $\phi : M \to M$ is a diffeomorphism and $f : V \to V$ is a isomorphism of vector spaces. △

There is no reason why we should limit ourselves to the Hessian. We could also look at some higher order derivatives.

Definition 4.3. *Let M be an n-dimensional smooth manifold and V a n-dimensional vector space. Let furthermore $p \in M$ be a point and $(x^i)_{i=1}^n$ a local coordinate chart for some open U containing p. Lastly let $\rho : M \to V$ be a smooth map. Then we define the **derivative of order r** at p as the map:*

$$H^r_p\rho : (T_pM)^{r+1} \to V/(im(H^{k<r}_p\rho)) : (X_1, \dots, X_{r+1}) \to H^r_p(X_1, \dots, X_{r+1})$$

Where we have that $im(H^{k<r}_p\rho) = d_p\rho(T_pM) \oplus im(H_p\rho) \oplus \dots \oplus im(H^{r-1}_p\rho)$ and $im(H^k_p\rho)$ is just the image of the Hessian of order k. Furthermore we have that H^r_p satisfies:

$$H^r_p(X_1, \dots, X_r) = \frac{\partial^r \rho}{\partial x^{i_1} \partial x^{i_2} \dots \partial x^{i_{r+1}}}(dx^{i_1} \otimes dx^{i_2} \otimes \dots \otimes dx^{i_{r+1}})(X_1, \dots, X_{r+1})$$

*Of course we can also define the **global derivative of order r** as the map:*

$$H^r\rho : p \in M \to H^r_p\rho$$

Remark. As in the case of the remark after definition 4.1, it is possible after quotienting to define the higher order derivative without the use of coordinates. △

Analogously, we can show that the higher order derivatives are independent of choice of parametrization.

Lemma 4.4. *Any higher order derivative at a point $p \in M$ is well-defined in the sense that it does not depend on the parametrization of both M and V.*

Proof. That it does not depend on the parametrization of V should be clear from the proof in the case of the Hessian.
Let $\phi : M \to M$ be a diffeomorphism. If we write the higher order derivatives in their coordinate free way and if we would follow the steps laid out in the proof in the case of the Hessian, we would notice a pattern in the formulas for the higher order derivatives. For example in the third order

derivative we would get the following formula:

$$\begin{aligned} H^2_{\phi^{-1}(p)}(\rho\circ\phi)(X_1,X_2,X_3) = \; & H^2_p(\rho)(d_{\phi^{-1}(p)}\phi(X_1), d_{\phi^{-1}(p)}\phi(X_2), d_{\phi^{-1}(p)}\phi(X_3)) + \\ & H_p\rho(d_{\phi^{-1}(p)}\phi(X_1), H_{\phi^{-1}(p)}\phi(X_2,X_3)) + \\ & H_p\rho(d_{\phi^{-1}(p)}\phi(X_2), H_{\phi^{-1}(p)}\phi(X_3,X_1)) + \\ & H_p\rho(d_{\phi^{-1}(p)}\phi(X_3), H_{\phi^{-1}(p)}\phi(X_1,X_2)) + \\ & d_p\rho(H^2_{\phi^{-1}(p)}\phi(X_1,X_2,X_3)) \end{aligned}$$

where $X_1, X_2, X_3 \in T_pM$ are vectors and $H_{\phi^{-1}(p)}\phi : T_pM \times T_pM \to T_pM$ is the Hessian of ϕ. The formulas for the even higher order derivatives will not look any nicer but they follow the same pattern. We would get all the lower order derivates where we would cyclically permute the vectors. The most important part to notice is that since we quotient out all the images of the lower order derivativer, all terms on the right hand side except the first one are zero. Thus we get that the higher order derivatives do not depend on the parametrization of M. □

4.2 Control systems

One of the main results of this thesis will be the fact that the no slipping distribution is bracket generating. Underlying this is a broader statement. In this subsection we will show that in the special case of control systems being bracket generating is linked to the curvature of that control system.

4.2.1 Control System and the Tautological Distribution

Let us begin by defining what a control system is.

Definition 4.5. *Suppose M is a smooth manifold. A* ***rank k control system*** *C is a subbundle of TM whose fibres are k-dimensional smooth manifolds. We denote the* ***fibre of a control system*** *by $C_p \subseteq T_pM$.*

The other central object in this chapter is the tautological distribution.

Definition 4.6. *Let M be a smooth manifold and C control system. Let $\pi : TM \to M$ the projection and $\rho : C \to TM$ the inclusion. We define the* ***tautological distribution*** *over a control system at $(p, v) \in C$ as $\mathcal{D}(p, v) = d(\pi \circ \rho)^{-1}(p, \langle v\rangle) \subseteq TTM$ where $\langle v\rangle \subseteq T_pM$ is the span of v.*

4.2.2 Sufficiently Curved

Remark. To simplify definitions and lemmata, we assume throughout this chapter that M is a n-dimensional smooth manifold and C a rank k control system. Furhtermore we take $\pi : TM \to M$ to be the projection and $\rho : C \to TM$ to be the inclusion. We will abbreviate the composition $\pi \circ \rho$ by π_C. △

We know that if a distribution is bracket generating, then it is possible to acquire all the vector fields on a smooth manifold from the Lie flag of the distribution. When formulated in this way, being bracket generating is something intrinsic to the distribution itself. Yet when we look at the way in which we defined the tautological distribution, we see that we construct it from the control system over which it is a distribution. One can argue from this fact, as we will do now, that if the control system satisfies some nice geometrical properties, then our tautological distribution must be bracket generating. The nice property we will focus on is the case that our control system is sufficiently curved. At the end of this subsection we will also briefly discuss another property which makes sure that the tautological distribution is bracket generating, namely that the control system itself is bracket generating.

Let us start out by defining what it means to be sufficiently curved.

Definition 4.7. *Let M be a smooth manifold and C a control system. Suppose we have a point $p \in M$ and let $\rho_p : C_p \to T_pM$ be the inclusion. We say that our control system is* ***sufficiently***

curved *if there exists some* $k \in \mathbb{N}$ *such at every point* $p \in M$ *and at any* $v \in C_p$ *we can realize:*

$$T_pM = im(H_v^{n\leq k}\rho_p)$$

where we now have that $H_v^n \rho_p : (T_vC_p)^n \to T_pM/(im(H_v^{l<n}\rho(p))$

Now what this tells us it that we can generate the entire tangent space T_pM of a smooth manifold M at a point p by taking at every vector $v \in C_p$ in the fibre of the control system the derivatives from elements in the tangent space T_vC_p and identifying them with an element in T_pM if the control system is sufficiently curved.

The rest of this chapter will be dedicated to show that the tautological distribution over a suffi-iciently curved control system is bracket generating. To this end we need to introduce two auxiliary lemmata. The first one will give us an explicit way to write the tautological distribution.

Lemma 4.8. *Suppose we have a point* $(p, v) \in D$*. Suppose we have a coordinate chart* $(x^i)_{i=1}^n$ *for some open* $U \subseteq M$ *containing* p*. Choose furthermore a coordinate chart* $(y_j)_{j=1}^k$ *for some open* $V \subseteq C_p$ *containing* v*. Let* $\mathcal{D}(p, v)$ *be the tautological distribution over* C *at the point* (p, v)*. Then we can write this distribution as:*

$$\mathcal{D}(p, v) = \sqcup_{w\in\langle v\rangle} a_{(p,w)} \oplus \{(p, w, 0, b)|b \in T_vC_p\}$$

where $a : C \to TC$ *is a locally defined smooth section which at the point* $p \in M$ *is defined as* $a_{(p,w)} = (p, w, w, 0)$.

Proof. Firstly we note that it should be clear that any element in $\sqcup_{w\in\langle v\rangle} a_{(p,w)} \oplus \{(p, w, 0, b)|b \in T_wC_p\}$ indeed projects down to an element $\langle\rho(p, v)\rangle$. Thus we have $\sqcup_{w\in\langle v\rangle} a_{(p,w)} \oplus \{(p, w, 0, b)|b \in T_wC_p\} \subseteq \mathcal{D}(p, v)$.
Now we note that any element $x \in TC$ can locally at p be written in the form (p, w, u, c), where $p \in M$, $w \in C_p$, $u \in T_pM$ and $c \in T_wC_p$. We furthermore note that $d\pi_C(x)$ kills of the two elements of x in the following way $d\pi_C(p, w, u, c) = (p, u)$.
What we should also note is that our choice of u is limited by the fact that our direction of movement is fixed due to the fact that we only allow movement along the line $\langle v\rangle$. Hence we have that $u \in \langle v\rangle$. Note also that there are no restraints on our choice of b. Hence this gives us that $\mathcal{D}(p, v)$ is a $k + 1$-dimensional object. Since $\sqcup_{w\in\langle v\rangle} a_{(p,w)} \oplus \{(p, w, 0, b)|b \in T_wC_p\}$ is a $k + 1$-dimensional object satisfying $\sqcup_{w\in\langle v\rangle} a_{(p,w)} \oplus \{(p, w, 0, b)|b \in T_wC_p\} \subseteq \mathcal{D}(p, v)$, this gives us that $\sqcup_{w\in\langle v\rangle} a_{(p,w)} \oplus \{(p, w, 0, b)|b \in T_wC_p\} = \mathcal{D}(p, v)$. □

The following is an easy to show corollary which we will need in the proof of our theorem.

Corollary 4.9. *Let the situation be such as in lemma 4.8. It then follows that for all* $w \in C_p$ $d\pi_C((\phi_X^t)^*(a_{(p,w)})) = \rho \circ \phi_X^t(p, w)$ *where* X *is a vector field tanget to the fibre* C_p *and* ϕ_X^t *is the flow of* X *at time* t.

Proof. The first thing to note is that since the vector field X is tangent to the fibre, $\phi_X^t(p, w)$ always gives us back an element in the fibre. To put it formally, for any element $(p, w) \in C$ and at any time t, $\phi_X^t(p, w)$ is of the shape $\phi_X^t(p, w) = (p, u)$ where $u \in C_p$. From this it follows that for any time t and any $(p, w) \in C$ we have that $\pi_C \circ \phi_X^t(p, w) = \pi_C(p, u) = p = \pi_C(p, w)$ and thus $\pi_C \circ \phi_X^t = \pi_C$. Furthermore, we know that by definition $(\phi_X^t)^*(a_{(p,w)})$ we can write as $(\phi_X^t)^*(a_{(p,w)}) = (d\phi_X^{-t})_{\phi_X^t(p,w)}(a_{\phi_X^t(p,w)})$. Thus we get that

$$\begin{aligned}
d\pi_C((\phi_X^t)^*(a_{(p,w)})) &= d\pi_C((d\phi_X^{-t})_{\phi_X^t(p,w)}(a_{\phi_X^t(p,w)})) \\
&= d(\pi_C \circ \phi_X^{-t})_{\phi_X^t(p,w)}(a_{\phi_X^t(p,w)})) \\
&= d(\pi_C)_{\phi_X^t(p,w)}(a_{\phi_X^t(p,w)})
\end{aligned}$$

Now by the way in which a is defined it follows $(d\pi_C)_{\phi_X^t(p,w)}(a_{\phi_X^t(p,w)}) = \rho(\phi_X^t(p, w))$. □

The second lemma we will prove gives us a relationship between the Lie flag of the tautological distribution and the Hessians of any order.

Lemma 4.10. *Suppose we have a point $(p, v) \in C$. Let $\mathcal{D}(p, v)$ be the tautological distribution over C at the point (p, v). Then there exists a relationship between the Lie brackets of the tautological distribution of length r and the derivative of order $r - 1$ in the following way:*

$$H_v^{r-1}\rho_p(X_1, \dots, X_r) = d\pi_C([X_r, [\dots, [X_1, a] \dots]](p, v))$$

where $a, X_i : C \to TC$ are smooth sections where a is defined as in lemma 4.8. Furthermore we assume that any X_i is tangent to the fibre C_p.

Proof. We start of by showing the relationship between the differential and the Lie brackets of length one. Let $(p, w) \in \langle \rho(p, v) \rangle$. First of all we note that we are able to write the Lie brackets of two vector fields in terms of its Lie derivatives as:

$$[X, a](p, w) = \mathcal{L}_X(a)(p, w) = \frac{d}{dt}|_{t=0}(\phi_X^t)^*(a)(p, w)$$

where $X_{(p,w)} \in \{(p, w, 0, b) | b \in T_w C_p\}$. We also have that X is tangent to the fibres of the control system and ϕ_X^t is the flow of X for a time t. Now we look at:

$$d\pi_C([X, a])(p, w) = d(\pi_C)(\frac{d}{dt}|_{t=0}(\phi_X^t)^*(a)(p, w))$$

We have that $d\pi_C \circ \frac{d}{dt}|_{t=0} = \frac{d}{dt}|_{t=0} \circ d\pi_C$. This now gives us:

$$d\pi_C([X, a])(p, w) = \frac{d}{dt}|_{t=0} \circ d\pi_C((\phi_X^t)^*(a)(p, w))$$

Now we apply corollary 4.9 to get:

$$d\pi_C([X, a])(p, w) = \frac{d}{dt}|_{t=0}\rho(\phi_X^t(p, w))$$

Now again by definition of the Lie derivative we get that:

$$d\pi_C([X, a])(p, w) = (d\rho)_{(p,w)}(X_{(p,w)})$$

Now we move on to the Hessian. The proof for any higher order derivative goes the same but is just a lot more work to write out.
Again let $(p, w) \in \langle \rho(p, v) \rangle$ and let X, Y be two vector fields tangent to the fibres of the control system. We now look again at the following Lie brackets:

$$d\pi_C([Y, [X, a]])(p, w) = \frac{d}{ds}|_{s=0}\frac{d}{dt}|_{t=0} d\pi_C((\phi_Y^s)^*(\phi_X^t)^*(a)(p, w))$$

We immediatly applied the fact that the time derivatives commute with $d\pi_C$. Furthermore we note again that ϕ_X^t and ϕ_Y^s are the flow of X and Y respectively.
Now as in the case of the regular differential we again have that the formula $(\phi_Y^s)^*(\phi_X^t)^*(a)(p, w) = (d\phi_X^{-t})_{\phi_X^t(p,w)}((d\phi_Y^{-s})_{\phi_Y^s \circ \phi_X^t(p,w)}(a_{\phi_Y^s \circ \phi_X^t(p,w)}))$ is satisfied. Also as in the case of the regular differential, we are still allowed to use corollary 4.9. This gives us:

$$\begin{aligned} d\pi_C([Y, [X, a]])(p, w) &= \frac{d}{ds}|_{s=0}\frac{d}{dt}|_{t=0} d\pi_C((d\phi_X^{-t})_{\phi_X^t(p,w)}((d\phi_Y^{-s})_{\phi_Y^s \circ \phi_X^t(p,w)}(a_{\phi_Y^s \circ \phi_X^t(p,w)}))) \\ &= \frac{d}{ds}|_{s=0}\frac{d}{dt}|_{t=0}\rho(\phi_Y^s \circ \phi_X^t(p, w)) \end{aligned}$$

By applying the definition of the Lie derivative we see that:

$$d\pi_C([Y, [X, a]])(p, w) = \frac{d}{ds}|_{s=0}(d\rho)_{\phi_Y^s(p,w)}(X_{\phi_Y^s(p,w)})$$

Now the right hand side is nothing else but the second order derivative of ρ and hence the Hessian. □

So now we are able to prove the main result of this chapter.

Theorem 4.11. *Suppose we have that $\mathcal{D}(p,v)$ is the tautological distribution at the point (p,v) over the control system C. Now if C is sufficiently curved, then $\mathcal{D}$ is bracket generating.*

Proof. Lemma 4.8 tells us how we can write our tautological distribution and lemma 4.2.2 tells us that there exists a relationship between the elements of the distribution and the Hessians of any order. Since our control system is sufficiently curved, this tells us that we can generate any element in T_pM from our distribution $\mathcal{D}(p,v)$. □

Remark (1). Note that if the Hessian or any higher order derivative is zero at some point $q \in M$, but not at nearby points, then the Lie flag will not vary smoothly as a function of $p \in M$ since the Lie flag then is not of pointwise constant rank. Since we have assumed that the control system is sufficiently curved, there must be some higher order derivative at that point $q \in M$ which is non-zero. So we are still able to generete all of the tangent plane at that one point q. The take-home message should be that some higher order derivatives might produce more direction at one point in M and less at another. △

Remark (2). It is perfectly possible to have a tautological distribution which is bracket generating, but its control system is not sufficiently curved. For example, look at example 5.6. There we looked at a distribution $\mathcal{D}$ over $\mathbb{R}^3$ spanned by $X = y\partial_x + z\partial_y$ and $Y = \partial_z$.
Let $p = (p_1, p_2, p_3) \in \mathbb{R}^3$ be a point. It should be clear that the fibres of the distribution $\mathcal{D}(p)$ are planes and thus smooth manifolds. This means that $\mathcal{D}$ is a control system. Let $\rho_p : \mathcal{D}(p) \to T_p\mathbb{R}^3$ be the inclusion. If we were to compute the Hessian or any higher order dervative, we would see that they always equal zero. What this tells us is that the distribution $\mathcal{D}$ is not sufficiently curved. Yet the tautological distribution over this control system $\mathcal{D}$ is bracket generating since by example 5.6 we have that $\mathcal{D}$ itself is bracket generating.
What this example tells us is that a control system being sufficiently curved is a sufficient condition, but not a necessary one, for the tautological distribution to be bracket generating. △

4.3 No Slipping Distribution Revisited

In the previous chapter we introduced the notion of control systems and proved the statement that if a control system is sufficiently curved, then the tautological distribution over that control system is bracket generating. What we want to do now is show that the no slipping distribution is bracket generating. When one looks at the no slipping distribution we see that this is indeed a tautological distribution over the Grassmanian of the lightcone $Gr(C(g), k)$.

So now we have to determine how the lightcone looks like. There is a problem in this case which we did not discuss before. It is namely possible that the metric on our semi-Riemannian manifold has crossterms. For example in the case of the rolling coin, we could have had that the metric was given by $g = d\alpha^2 - dx^2 - dy^2 - dxdy$. In that case the proof of lemma 3.23 would not have been so straighforward.

Luckily for us, this problem is not insurmountable. Since we are looking at the lightcone at some point, we can just take the normal coordinates centered at that point and then we know that the metric at that point has no crossterms. This then gives rise to the following lemma.

Lemma 4.12. *Let $(M, g^{(d_1,d_2)})$ be a n-dimensional semi-Riemannian manifold with metric signature $(d, n-d)$ and let $p \in M$ be a point. Then there exists a vector space isomorphism between $C_p(g)$ and $\{(x_1, \ldots, x_n) \in \mathbb{R}^n | \sum_{i=1}^{d} x_i^2 = \sum_{i=d+1}^{n} x_i^2\}$*

Proof. Let $U \subseteq M$ be the normal neighbourhood of p and let $(x^i)_{i=1}^n$ be normal coordinates on that U centered at p. Then by lemma 2.51 we can write the metric at p as:

$$g_p = \sum_{i=1}^{d} dx^i \otimes dx^i - \sum_{j=d+1}^{n} dx^j \otimes dx^j$$

Now since any point $v \in C_p(g)$ satisfies $g_p(v,v) = 0$, we can just follow the steps from the proof of lemma 3.23 and the result follows. □

So now we know that the lightcone is isomorphic to the regular cone viewed in the Euclidean space. So we just have to show that the cone in Euclidean space is sufficiently curved. We can also make another simplification, namely we can include the two-dimensional cone in any higher-dimensional cone. So we first of all show that the two-dimensional cone is sufficiently curved in $\mathbb{R}^3$.

Lemma 4.13. *Let $C = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 = z^2\}$ be the cone in $\mathbb{R}^3$. Let $\rho : C \to \mathbb{R}^3$ be the inclusion. Then the Grassmanian $Gr(C, 1)$ is sufficiently curved in the sense that for every line $l \in Gr(C, 1)$ and at every point $p \in l - \{0\}$ the cone is sufficiently curved in the following way:*

$$T_{\rho(p)}\mathbb{R}^3 = d_p\rho(T_pC) \oplus im(H_p\rho)$$

where $H_p\rho : T_pC \times T_pC \to T_{\rho(p)}\mathbb{R}^3/(d_p\rho(T_pC))$ is the Hessian.

Proof. Let $l \in Gr(C, 1)$ and $p \in l$ be a point. Since we work in $\mathbb{R}^3$ we can find a global frame such that we can write $p = (p_1, 0, p_3)$. In other words, we choose the x- and z-axis such that p lies on the plane spanned by the x- and z-axis. So now we note that the tangent space T_pC can be written as the span of $X_p = p_1\partial_x + p_3\partial_z$ and $Y_p = p_3\partial_y$ or to put it more formally $T_pC = span\{X_p, Y_p\}$. We furthermore also assume that $p \in C$ is in the upper half of the cone or in other words $p_3 > 0$. This gives us then $p_3 = p_1$. The reason we do this is so that we can explicitely compute the Hessian. It is tedious work to show that it equals:

$$H_p\rho = (\frac{p_1^2}{p_3^3}\partial_z)dy \otimes dy = (\frac{1}{p_1}\partial_z)dy \otimes dy$$

In the second step we made use of the fact that $p_1 = p_3$. Now let us compute one case explicitely to make this more clear. We know that $\rho(x, y, z) = (x, y, z)$. Now we compute the first partial derivative with respect to x. This gives us:

$$\frac{\partial \rho}{\partial x} = \partial_x + \frac{\partial z}{\partial x}\partial_z = \partial_x + \frac{x}{\sqrt{x^2 + y^2}}\partial_z$$

Taking again the partial derivative with respect to x gives us:

$$\frac{\partial^2 \rho}{\partial x^2} = \frac{\partial}{\partial x}(\frac{x}{\sqrt{x^2 + y^2}})\partial_z = \frac{y^2}{\sqrt{x^2 + y^2}}\partial_z$$

Now plugging in the coordinates $\frac{\partial^2\rho}{\partial x^2}(p_1, 0, p_1) = 0$. Analogously, one can compute all the other cases and see that only the second order derivates with respect to y does not vanish.
Now the following equation follows trivially:

$$H_p\rho(Y_p, Y_p) = p_3\partial_z$$

Now it should also be easy to see that T_pC and $im(H_p\rho)$ together span all of $T_{\rho(p)}\mathbb{R}^3$. Since we can always choose our coordinates such that we can write $p = (p_1, 0, p_3)$, the above must hold true for any point $p \in l - \{0\}$ proving that $Gr(C, 1)$ is sufficiently curved. □

Remark. It should be noted that the cone is not smooth at the origin because the rank changes, but we just disregard that point. We are allowed to do this since we could have also looked at the sphere bundle over the cone. This would still have preserved the tautological distribution, but would have elimited the origin. Then the same but slightly modified argument is true. △

As a corollary to this lemma we can prove the same but then for a cone of arbitrary dimension.

Corollary 4.14. *Let $C^{m,n} = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n | \sum_{i=1}^m x_i^2 = \sum_{j=1}^n y_j^2\}$, where x_i and y_j are the components of x and y respectively. Then the Grassmanian $Gr(C^{m,n}, 1)$ is sufficiently curved in the sense that for every line $l \in Gr(C^{m,n}, 1)$ and at every point $p \in l - \{0\}$ the cone is sufficiently curved in the following way:*

$$T_{\rho(p)}\mathbb{R}^{m+n} = d_p\rho(T_pC^{m,n}) \oplus im(H_p\rho)$$

where $H_p\rho : T_pC^{m,n} \times T_pC^{m,n} \to T_{\rho(p)}\mathbb{R}^{m+n}/(d_p\rho(T_pC^{m,n}))$ is the Hessian.

Proof. Let $i : C^{2,1} \to C^{m,n}$ be the inclusion. We know by lemma 4.13 that $Gr(C^{2,1}, 1)$ is sufficiently curved. We note now that the dimension of $C^{m,n}$ is one dimension lower than that of $\mathbb{R}^{m+n}$. Since we can include the cone $C^{2,1}$ into the cone $C^{m,n}$, we can generate this missing direction from $C^{2,1}$. This shows that $Gr(C^{m,n}, 1)$ is sufficiently curved. □

So now we can show that the no slipping distribution over the space $Gr(C(g), 1)$ is bracket generating.

Lemma 4.15. *Let (M, g) be a semi-Riemannian manifold and let $(p, l) \in Gr(C(g), 1)$. Then the no slipping distribution $\mathcal{D}^{NS}(p, l) = d\pi^{-1}(p, l)$ is bracket generating.*

Proof. Since $Gr(C(g), 1)$ is sufficiently curved control system by lemma 4.14 and $\mathcal{D}^{NS}$ is the tautological distribution over $Gr(C(g), 1)$, we get by theorem 4.11 the desired result. □

As a corollary to this lemma we can show the following:

Corollary 4.16. *Let (M, g) be a semi-Riemannian manifold and let $(p, H) \in Gr(C(g), k)$. Then the no slipping distribution $\mathcal{D}^{NS}(p, H) = d\pi^{-1}(p, H)$ is bracket generating.*

Proof. Let $l \in Gr(C(g), k)$ be a line satisfying $l \subseteq H$. We see that $\mathcal{D}^{NS}(p, l) = d\pi^{-1}(p, l) \subseteq d\pi^{-1}(p, H) = \mathcal{D}^{NS}(p, H)$. Since $\mathcal{D}^{NS}(p, l)$ is bracket generating, we have by lemma 5.8 that also $\mathcal{D}^{NS}(p, H)$ is bracket generating. □

Remark. We will not show this here, but it is worth noting that when we look at the no slipping distribution at a point $(p, H) \in Gr(C(g), k)$ where $k \geq 2$, the no slipping distribution bracket generates in one step. On the other hand, in the case of a point $(p, l) \in Gr(C(g), 1)$ the no slipping distribution bracket generates in two steps. △

5 Appendix A: Distributions

To define the no slipping and the no twisting conditions we made use of distributions. In this appendix we will give some background information regarding distributions and prove some properties which we use in the thesis.

5.1 Distributions

Distributions will play a major role throughout this thesis. The easiest way to think about them is that, at a given point in the manifold and given restraints on our movements, a distributions contains all the directions in which we are allowed to move in.

Definition 5.1. *Suppose M is a smooth manifold and TM the tangent bundle over the manifold. A* ***rank*** *k* ***smooth distribution*** *$\mathcal{D}$ satisfies that, at every $p \in M$, $\mathcal{D}(p)$ is a linear subspace of T_pM and there exists an open neighbourhood U and k smooth vector fields $X_i : U \to TM$ such that at every point $q \in U$, $(X_i(q))_{i=1}^k$ spans $\mathcal{D}(q)$.*

Remark. Throughout this thesis, when we say distribution we mean a smooth distribution. △

5.2 Integrable and Involutive Distributions

A question one could ask now is: does a distribution have any real world representations? One should think about driving a car. When driving you have to stay on your side of the road. Certainly you are able to drive in any direction you like, but the only legal one is the one which keeps you on the road. The distribution will coincide with this legal direction. But this distribution has a real world representation, namely the road. This gives rise to the notion of a integral manifold.

Definition 5.2. *Let $\mathcal{D} \subseteq TM$ be a rank k distribution and $N \subseteq M$ a k-dimensional immersed submanifold. We say that N is an* ***integral manifold*** *of $\mathcal{D}$ if at every point $p \in N$ we have that $T_pN = \mathcal{D}(p)$. If at every point there exists such a manifols, we say that the distribution is* ***integrable****.*

A celebrated theorem in differential geometry called Frobenius Theorem, whose proof falls outside the scope of this thesis, tells us that there is an equivalent definition to $\mathcal{D}$ being integrable, namely being involutive. We will discuss this definition and at the end of this subsection we will state Frobenius Theorem

To give this equivalent definition we must introduce the notion of modules associated to a distribution.

Definition 5.3. *Let $\mathcal{D}$ be a distribution The* ***module associated to this distribution*** *$\Gamma(\mathcal{D})$ is the C^∞-module of vector fields of pointwise constant rank tangent to the distribution.*

While the distribution tells us at every point which direction is permitted, the flag tells us globally which vector fields are permitted. In terms of integral curves of a vector field we could also say that the flag tells us along which curves we are allowed to move.

Now we introduce the Lie flag.

Definition 5.4. *A* ***bracket expression*** *of length zero is the variable x itself. Inductively we define the bracket expression of length $i + j + 1$ as:*

$$[A(a_1, \dots, a_i), B(a_{i+1}, \dots, a_{i+j})]$$

Where A and B are bracket expressions of length i and j.

Definition 5.5. *Let M be a manifold and given a distribution $\mathcal{D}$ on that manifold, the* ***Lie flag*** *is a increasing squence of C^∞-modules $(\xi^{(i)})_{i\in\mathbb{N}}$, where every $\xi^{(i)}$ is given by:*

$$\xi^{(i)} = \langle A(X_1, \dots, X_i) | X_1, \dots, X_i \in \Gamma(\mathcal{D})\rangle$$

Where A is a bracket expression of lenght at most i and the span should be taken over all $C^\infty(M)$ instead of over the real numbers. Note that $\xi^{(0)} = \Gamma(\mathcal{D})$.

The Lie flag tells us which vector fields we can generate from the permitted vector fields. We will make this clear with an example.

Example 5.6. Suppose we are in $\mathbb{R}^3$ and we have a distribution $\mathcal{D}$ spanned by $X = y\partial_x + z\partial_y$ and $Y = \partial_z$. At first look, it is not entirely clear to us that we are allowed to move in the ∂_x direction. But when computing the Lie brackets, we see:

$$[X, Y] = \partial_y$$
$$[X, [X, Y]] = \partial_x$$

Hence we have that the second element in the Lie flag equals $\xi^{(2)} = \mathfrak{X}(\mathbb{R}^3)$. $\diamondsuit$

As this example makes clear, after generating one non-permissible motion through our Lie flag, there's nothing preventing us from taking the Lie bracket of the non-permissible motion with a permissible motion to generate another motion. This example illustrates also another interesting property, namely that we can generate all the vector fields from our distribution. Let us define this formally.

Definition 5.7. *A distribution is **bracket generating** if there exists some number $n \in \mathbb{N}$ such that the Lie flag $(\xi^{(i)})_{i\in\mathbb{N}}$ associated to this distribution satisfies $\xi^{(n)} = \mathfrak{X}(M)$.*

A lemma which we will need later on regaring bracket generating distributions is the following:

Lemma 5.8. *Suppose M a smooth manifold. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be two distributions on M of rank k and l respectively where $k \leq l$. Let furthermore $\mathcal{D}_1 \subseteq \mathcal{D}_2$ and let $\mathcal{D}_1$ be bracket generating. Then also $\mathcal{D}_2$ must be bracket generating.*

Proof. Let $(\xi^{(i)})_{i\in\mathbb{N}}$ and $(\psi^{(j)})_{j\in\mathbb{N}}$ be the Lie flags associated to $\mathcal{D}_1$ and $\mathcal{D}_2$ respectively. It is easy to see that for any $n \in \mathbb{N}$ it holds true that $\xi^{(n)} \subseteq \psi^{(n)}$. Since we can find a $k \in \mathbb{N}$ such that $\xi^{(k)} = \mathfrak{X}(M)$ and we have that $\psi^{(k)} \subseteq \mathfrak{X}(M)$, it follows that $\psi^{(k)} = \mathfrak{X}(M)$. $\square$

Yet it is also imaginable that the Lie flag does not give us any new direction in which we can move. This is precisely what it means to be involutive.

Definition 5.9. *An **involutive** distribution is a distribution whose Lie flag satisfies $\xi^{(i)} = \xi^{(0)}$ for any $i \in \mathbb{N}$.*

Let us now state Frobenius Theorem.

Theorem 5.10. *(Frobenius) A distribution is involutive if and only if it is integrable.*

We will now show that integrable implies involutive, but we will only give a brief sketch of why involutive implies integrable.

Lemma 5.11. *Suppose $\mathcal{D}$ is a rank k integrable distribution. Then $\mathcal{D}$ is also involutive.*

Proof. Since $\mathcal{D}$ is integrable, there exists a k-dimensional immersed submanifold N. Furthermore suppose we have an open subset $U \subseteq M$ around a point $p \in N$ and $(X_i)_{i=1}^k$ smooth vector field such that $(X_i(q))_{i=1}^k$ span $\mathcal{D}(q)$ for every $q \in U$. Since $\mathcal{D}$ is integrable, we have that each of these vector fields are tangent to N. We know that the Lie bracket of any two vector fields which are tangent to an immersed submanifold, is itself tangent to that submanifold. Hence we have that for any $i, j \leq k$ that $[X_i, X_j](p)$ is tangent to N and thus $[X_i, X_j](p) \in \mathcal{D}(p)$. We can do this for every point in N. This then implies that any element of the Lie flag of this distribution does not generate a new element and hence $\xi^{(i)} = \xi^{(0)}$. $\square$

5.3 Foliations

Arguing the inverse relies on a notion of foliation. To this end we introduce the foliation chart.

Definition 5.12. *Let M be a n-dimensional smooth manifold, $p \in M$ and $\mathcal{D}$ a rank k distribution on that manifold. A **foliation chart** ψ of $\mathcal{D}$ at a point p is a smooth embedding from the n-*

dimensional hypercube to M or formally:

$$\psi : [0,1]^n \to M$$

Furthermore ψ satisfies that p is in its image and that the pullback of $\mathcal{D}$ is given by $\psi^*\mathcal{D} = ker(dx_{k+1}, \dots, dx_n)$, where $(x_i)_{i=1}^n$ are coordinates on $[0,1]^n$.

The next lemma plays a major role in the proof for Frobenius theorem.

Lemma 5.13. *Suppose M a smooth manifold and $\mathcal{D}$ a rank k distribution. If there exists a foliation chart of $\mathcal{D}$ at every point p in M, then $\mathcal{D}$ is integrable*

Proof. Let ψ be the foliation chart of $\mathcal{D}$ at some point p. Since ψ is a smooth embedding, it follows that also its restriction is a smooth embedding. So $N_1 := \psi|_{[0,1]^k}([0,1]^k)$ is a k-dimensional embedded manifold of M. Since by property of the foliation chart $\psi^*\mathcal{D} = ker(dx_{k+1}, \dots, dx_n)$, we have that $(d\psi(\partial_i))_{i=1}^k$ span $\mathcal{D}(q)$ for all $q \in N$. But since they also span the tangent space T_qN_1 for every $q \in N_1$, we have that $\mathcal{D}(q) = T_qN_1$ and hence $\mathcal{D}$ is integrable.
The only thing left to check is that $\mathcal{D}$ agrees on the overlap of the image of two foliation charts. Let ψ and χ be two foliation charts such that $N_2 := \psi|_{[0,1]^k}([0,1]^k) \cap \chi|_{[0,1]^k}([0,1]^k) \neq \emptyset$. Now since both $(d\psi(\partial_i))_{i=1}^k$ and $(d\chi(\partial_i))_{i=1}^k$ form a coordinate chart for every $q \in N_2$ and coordinate charts should be smoothly compatible, we now have that $\mathcal{D}$ agrees on the overlap. $\square$

So now we are able to define foliations.

Definition 5.14. *Let M be a n-dimensional smooth manifold. A **foliation** $\mathcal{F}$ of dimension k is a disjoint collection $(N_i)_{i \in I}$, which we call the **leaves** of the foliation, of k-dimensional immersed submanifolds of M whose union is M. Furthermore we require that for every $p \in M$ there exists a foliation chart of $T\mathcal{F} := \{T_pN_i | p \in M, i \in I\}$ at p.*

Example 5.15. All linear subspaces parallel to $\mathbb{R}^k \times \{0\}$ form a k-dimensional foliation of $\mathbb{R}^n$. $\diamondsuit$

More generally, a foliation should be thought of as creating some sort of layering in the manifold.

The proof of Frobenius theorem goes along this way, but for more detail we refer to [6] theorem 19.12.

6 Appendix B: Vector Spaces

In this appendix we prove a lemma which is important in the discussion regarding the fact that the no slipping distribution is bracket generating. What this lemma tells us, is that in cases of vector spaces we can identify a direction in the tangent space at a point with the vector space itself. This is evident when thought about in the Euclidean space. There when we take a partial derivative, we don't view the direction as an element of the tangent space, but immediately identify it with a direction in the Euclidean space. Let us state this lemma.

Lemma 6.1. *Suppose V a n-dimensional vector space and $p \in V$. Let $(x^i)_{i=1}^n$ be a basis for V. There exists an isomorphism of vector spaces between T_pV and V defined by the map:*

$$\Phi_p : V \to T_pV : v \to [t \to p + tv](0)$$

where $[t \to p + tv](0)$ is the equivalence class of curves at $t = 0$ by which we mean all the curves $\gamma : I \to M$ which are equivalent to the curve $t \to p + tv$ at time $t = 0$.

Proof. We can define the inverse of Φ_p is by the map:

$$\Phi_p^{-1} : T_pV \to V : [t \to p + tv](0) \to v$$

This map is well-defined since any two curves γ, ν are in the equivalence class of $[t \to p + tv](0)$ if they satisfy that $\gamma(0) = \nu(0) = p$ and $\frac{\partial \gamma}{\partial t}(0) = \frac{\partial \nu}{\partial t}(0) = v$. Hence the value of Φ_p^{-1} does not depend on the choice of element in $[t \to p + tv](0)$. So the inverse of this map is well-defined and hence our map is bijective.
Furthermore, Φ_p is linear since we can write $\Phi_p(v + w) = [t \to p + t(v + w)](0)$. Hence a curve $\gamma \in [t \to p + t(v + w)](0)$ has to satisfy $\frac{\partial \gamma}{\partial t}(0) = v + w$. Now we take two curve $\nu_1 \in [t \to p + tv](0)$ and $\nu_2 \in [t \to p + tw](0)$. This tells us that $\frac{\partial \gamma}{\partial t}(0) = \frac{\partial \nu_1}{\partial t}(0) + \frac{\partial \nu_2}{\partial t}(0)$ and hence $\Phi_p(v + w) = \Phi_p(v) + \Phi_p(w)$.
The inverse is also linear since $\Phi_p^{-1}([t \to p+t(v+w)](0)) = v+w = \Phi_p^{-1}([t \to p+tv](0))+\Phi_p^{-1}([t \to p + tw](0))$. □

Remark. In this thesis, we will often state that there exists an isomorphism Φ of vector spaces between the tangent bundle TM and the double tangent bundle TTM over a smooth manifold M. The way to understand this is to let Φ be the followig map:

$$\Phi : (p, v) \in TM \to \Phi_{(p,v)}$$

where $\Phi_{(p,v)} : T_vT_pM \to T_pM$ is the vector space isomorphism from the above lemma. △

References

[1] D. An and P. Nurowski. "Twistor Space for Rolling Bodies". In: *Communications in Mathematical Physics* 326 (2013), pp. 393–414.

[2] T.N. Bailey and R.J. Baston. *Twistors in Mathematics and Physics*. 1990. DOI: https://doi-org.proxy.library.uu.nl/10.1017/CBO9781107325821.

[3] W.M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry (Second edition)*. Academic Press, 1986. ISBN: 0121160505.

[4] F. Burstall. "Integrable Systems in Twistor Theory". In: *Twistors in Mathematics and Physics* (1990). Ed. by T. Bailey and R. Baston, pp. 53–70.

[5] M. Crainic. *Manifolds 2019-2020*. DOI: https://webspace.science.uu.nl/~crain101/manifolds-2019/.

[6] J.M. Lee. *Introduction to Smooth Manifolds (Second Edition)*. Springer, 2003. ISBN: 9780387954950.

[7] J.M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Springer, 1997. ISBN: 9780387982717.

[8] M. Dunajski M. Atiyah and L.J. Mason. "Twistor Theory at Fifty: From Contour integrals to Twistor Strings". In: *Proceedings of the Royal Society A* 473 (Oct. 2017). DOI: https://royalsocietypublishing.org/doi/10.1098/rspa.2017.0530#d3e2896.

[9] B. O'Neill. *Semi-Riemannian Geometry with Applications to Relativity*. Academic Press, 1983. ISBN: 0125267401.

[10] R. Penrose. "Twistor Algebra". In: *Journal of Mathematical Physics* 8 (1967), pp. 345–366.

[11] Á. Del Pino. *Lecture Notes: Topological Aspects in the Study of Tangent Distributions*. 2018. DOI: https://alvarodelpino.files.wordpress.com/2019/10/coimbra2018notes.pdf.

[12] R.S. Wars. "Integrable Systems in Twistor Theory". In: *Twistors in Mathematics and Physics* (1990). Ed. by T. Bailey and R. Baston, pp. 246–259.

[13] E. Witten. "Perturbative Gauge Theory as a String Theory in Twistor Space". In: (2003). DOI: https://arxiv.org/pdf/hep-th/0312171.pdf.

[14] P. Woit. "Twistor Geometry and the Standard Model in Euclidean Space". In: (2020). DOI: https://www.math.columbia.edu/~woit/twistors.pdf.