Lecture 1. Introduction

In Differential Geometry we study spaces which are smooth enough to do calculus. These include such familiar objects as curves and surfaces in space.

In its earliest manifestation, differential geometry consisted of the study of curves and surfaces in the plane and in space — this goes back at least as far as Newton and Leibniz who applied calculus to the study of curves in the plane, and to Monge and Euler who gave analytic treatments of surfaces. This concentration on geometry in Euclidean space seems quite natural, since curves and surfaces arise very naturally as trajectories or as level sets of functions — for example, in mechanics it is common to consider phase spaces, and within them surfaces on which some preserved quantity (such as energy) is constant. A very substantial body of results was developed in what is now often called 'classical differential geometry', covering such topics as evolutes and involutes, developable surfaces and ruled surfaces, envelopes, minimal surfaces, parallel surfaces, and so on. None of these will feature very much in this course, though there are many fascinating aspects to all of this.

The material covered in this course is almost all of much more recent vintage, but serves as an excellent basis for the treatment of all the subjects mentioned above. The notion of differentiable manifold unifies and simplifies most of the computations involved in these more classical subjects, so that they are now more sensibly treated in a second course, applying the fundamental ideas developed here.

Hermann Weyl

Hassler Whitney

Carl Friedrich Gauss

The notion of a differentiable manifold was not clearly formulated until relatively recently: Hermann Weyl first gave a concise definition in 1912 (in his work making rigorous the theory of Riemann surfaces), but this did not really come into common use until much later, after a series of papers by Hassler Whitney around 1936. However the idea has its roots much earlier. Gauss had a major interest in differential geometry, and published many papers on the subject. His most renowned work in the area was *Disquisitiones generales circa superficies curva* (1828). This paper contained extensive discussion on geodesics and what are now called 'Gaussian coordinates' and 'Gauss curvature', which he called the 'measure of curvature'. The paper also includes the famous theorema egregium:

> *"If a curved surface can be developed (i.e. mapped isometrically) upon another surface, the measure of curvature at every point remains unchanged after the development."*

This result led Gauss to a fundamental insight:

> *"These theorems lead to the consideration of curved surfaces from a new point of view ... where the very nature of the curved surface is given by means of the expression of any linear element in the form* $\sqrt{E\mathrm{d}p^2 + 2F\mathrm{d}p\mathrm{d}q + G\mathrm{d}q^2}$*."*

In other words, he saw that it is posible to consider the geometry of a surface as defined by a metric (here he is locally parametrizing the surface with coordinates p and q), without reference to the way the surface lies in space.

In 1854 Riemann worked with locally defined metrics (now of course known as Riemannian metrics) in any number of dimensions, and defined the object we now call the Riemann curvature. These computations were local, and Riemann only gave a rather imprecise notion of how a curved space can be defined globally.


Bernhard Riemann

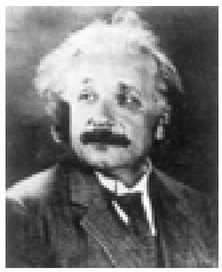

G. Ricci-Curbastro


T. Levi-Civita

Despite the lack of a precise notion of how a manifold should be defined, significant advances were made in differential geometry, particularly in developing the local machinery for computing curvatures — Christoffel and Levi-Civita introduced connections (not precisely in the sense we will define them in this course), and Ricci and Schouten developed the use of tensor calculus in geometric computations. Henri Poincaré worked with manifolds, but never precisely defined them.

Riemann's work surfaced again in spectacular fashion in Einstein's formulation of the theory of general relativity – Einstein had been trying for several years to find a way of expressing mathematically his principle of equivalence, and had attempted to use Gauss's ideas to encapsulate this. In 1912 he learned from his friend Marcel Grossman about Riemann's work and its mathematical development using tensors by Christoffel, Ricci and Levi-Civita, and succeeded in adapting it to his requirements after a further three years of struggle. This development certainly gave great impetus to the further development of the field, but Einstein still worked entirely on the local problem of interpreting the equivalence principle, and never worked in any systematic way on the global spacetime manifold.


E. Christoffel


Albert Einstein

The definition of a manifold encapsulates the idea that there are no preferred coordinates, and therefore that geometric computations must be invariant under coordinate change. Since this is built automatically into our framework, we never have to spend much time checking that things are geometrically well-defined, or invariant under changes of coordinates. In contrast, in the works of Riemann, Ricci and Levi-Civita these computations take some considerable effort.

The abstract notion of a manifold, without reference to any 'background' Euclidean spaces, also arises naturally from several directions. One of these is of course general relativity, where we do not want the unnecessary baggage associated with postulating a larger space in which the physical spacetime should lie. Another comes from the work of Riemann in complex analysis, in what is now called the theory of Riemann surfaces. Consider an analytic function f on a region of the complex plane. This can be defined in a neighbourhood of a point z_0 by a convergent power series. This power series converges in some disk of radius r_0 about z_0. If we now move to some other point z_1 in this disk, we can look at the power series for f about z_1, and this in general converges on a different region, and can be used to extend f beyond the original disk. Using analytic extensions in this way, we can move around the complex plane as long as we avoid poles and singularities of f. However, it may happen that the value of the function that we obtain depends on the path we took to get there, as in the case $f(z) = z^{1/2}$. In this case we can think of the function as defined not on the plane, but instead on an abstract surface which projects onto the plane (if $f(z) = z^{1/2}$, this surface covers $\mathbb{C}\backslash\{0\}$ twice). Another place where it is natural to work with abstract manifolds is in the theory of Lie groups, which are groups with a manifold structure.

1.1 Differentiable Manifolds

Definition 1.1.1 (Manifolds and atlases)

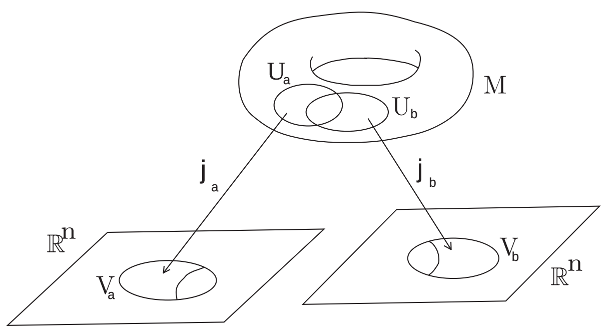
A *manifold* M of dimension n is a (Hausdorff, paracompact) topological space[1] M, such that every point $x \in M$ has a neighbourhood which is homeomorphic to an open set in Euclidean space $\mathbb{R}^n$.

A *chart* for M is a homeomorphism $\varphi : U \to V$ where U is open in M and V is open in $\mathbb{R}^n$.

A collection of charts $\mathcal{A} = \{\varphi_\alpha : U_\alpha \to V_\alpha |\ \alpha \in \mathcal{I}\}$ is called an *atlas* for M if $\cup_{\alpha\in\mathcal{I}} = M$.

Next we want to impose some 'smooth' structure:

Definition 1.1.2 (Differentiable structures) Let M be an n-dimensional manifold. An atlas $\mathcal{A} = \{\varphi_\alpha : U_\alpha \to V_\alpha |\ \alpha \in \mathcal{I}\}$ for M is *differentiable* if for every α and β in $\mathcal{I}$, the map $\varphi_\beta \circ \varphi_\alpha^{-1}$ is differentiable[2] (as a map between open subsets of $\mathbb{R}^n$).



[1] If you are not familiar with these topological notions, it is sufficient to consider metric spaces, or even subsets of Euclidean spaces. If you want to know more, see Appendix A.

[2] I will use the terms 'differentiable' and 'smooth' interchangeably, and both will mean 'infinitely many times differentiable'.

Two differentiable atlases $\mathcal{A}$ and $\mathcal{B}$ are *compatible* if their union is also a differentiable atlas — equivalently, for every chart ϕ in $\mathcal{A}$ and η in $\mathcal{B}$, $\phi \circ \eta^{-1}$ and $\eta \circ \phi^{-1}$ are smooth.

A *differentiable structure* on a manifold M is an equivalence class of differentiable atlases, where two atlases are deemed equivalent if they are compatible.

A *differentiable manifold* is a manifold M together with a differentiable structure on M.

We will usually abuse notation by simply referring to a 'differentiable manifold M' without referring to a differentiable atlas. This is slightly dangerous, because the same manifold can carry many inequivalent differentiable atlases, and each of these defines a different differentiable manifold.

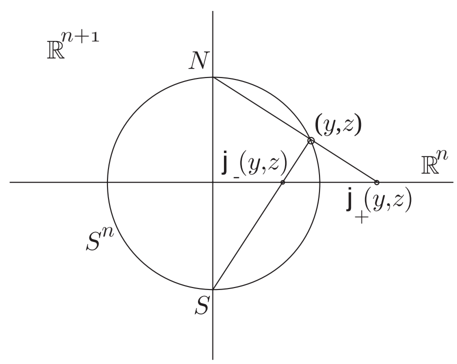
Example 1.1.1 Consider the set $M = \mathbb{R}$ (with the usual topology). This can be made a manifold is many different ways: The obvious way is to take the atlas $\mathcal{A} = \{\mathrm{Id} : \mathbb{R} \to \mathbb{R}\}$. Another way is to take the atlas $\mathcal{B} = \{x \mapsto x^3 : \mathbb{R} \to \mathbb{R}\}$. More generally, any homeomorphism of $\mathbb{R}$ to itself (or to an open subset of itself) can be used to define an atlas. The atlases $\mathcal{A}$ and $\mathcal{B}$ are incompatible because the union is $\{x \mapsto x,\ x \mapsto x^3\}$, which is not an atlas because the first map composed with the inverse of the second is the map $x \mapsto x^{1/3}$, which is not smooth. This example can be extended to show there are infinitely many different differentiable structures on the real line. This seems ridiculously complicated, but it will turn out that these differentiable structures are all equivalent in a sense to be defined later.

Remark. Given a differentiable structure on a manifold M, we can in principle choose a canonical atlas on M, namely the maximal atlas consisting of all those charts which are compatible with some differentiable atlas representing the differentiable structure. This is occasionally useful as a theoretical device but is completely unworkable in practice, as such a maximal atlas necessarily contains uncountable many charts.

Example 1.1.2: Euclidean space. A trivial example of a differentiable manifold is the Euclidean space $\mathbb{R}^n$, equipped with the atlas consisting only of the identity map.

Example 1.1.3: An atlas for S^n. The sphere S^n is the set $\{x \in \mathbb{R}^{n+1} : |x| = 1\}$. Since this is a closed subset of Euclidean space, the topological requirements are satisfied. Define maps $\varphi_+ : S^n \backslash \{N\} \to \mathbb{R}^n$ and $\varphi_- : S^n \backslash \{S\} \to \mathbb{R}^n$ as follows, where N is the "north pole" $(0, \dots, 0, 1)$ and S the "south pole" $(0, \dots, 0, -1)$: Writing $x \in \mathbb{R}^{n+1}$ as (y, z) where $y \in \mathbb{R}^n$ and $z \in \mathbb{R}$, we take $\varphi_+(y, z) = \frac{y}{1-z}$ and $\varphi_-(y, z) = \frac{y}{1+z}$. Then we have $\varphi_+^{-1}(w) = \frac{(2w, |w|^2-1)}{|w|^2+1}$ and $\varphi_-^{-1}(w) = \frac{(2w, 1-|w|^2)}{|w|^2+1}$. It follows that $\varphi_\pm$ is a homeomorphism, since clearly

$\varphi_\pm^{-1}$ is continuous, and $\varphi_\pm$ is the restriction to S^n of a continuous map defined on all of $\mathbb{R}^{n+1}$. Also $\varphi_+ \circ \varphi_-^{-1}(w) = \frac{w}{|w|^2}$ and $\varphi_- \circ \varphi_+^{-1}(w) = \frac{w}{|w|^2}$ for all $w \in \mathbb{R}^n \backslash \{0\}$. Since these maps are differentiable, the two charts $\varphi_\pm$ form a differentiable atlas, and so define a differentiable manifold.



The maps $\varphi_\pm$ in the last example are the *stereographic projections* from the north and south poles. Next we consider an example which does not come from a subset of $\mathbb{R}^{n+1}$:

Example 1.1.4: The real projective spaces. The n-dimensional real projective space $\mathbb{R}P^n$ is the set of lines through the origin in $\mathbb{R}^{n+1}$ (an observer at the origin sees anything along such a line as being at the same position in the field of view – thus real projective space captures the geometry of perspective drawing). Equivalently, $\mathbb{R}P^n = (\mathbb{R}^{n+1}\backslash\{0\})/\sim$ where $x \sim y \iff x = \lambda y$ for some $\lambda \in \mathbb{R}\backslash\{0\}$. A point in $\mathbb{R}P^n$ is denoted $[x_1, x_2, \ldots, x_{n+1}]$, meaning the equivalence class under $\sim$ of the point $(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1}$. We place a topology on $\mathbb{R}P^n$ by taking the open sets to be images of open sets in $\mathbb{R}^{n+1}$ under the projection onto $\mathbb{R}P^n$. This topology is Hausdorff: Take any two non-zero points x and y in $\mathbb{R}^{n+1}$ not lying on the same line. Choose an open set U about x which is disjoint from the line through y and the origin. Then choose an open set V about y which is disjoint from the set $\{\lambda z : \ z \in U, \ \lambda \in \mathbb{R}\}$. Then $U/\sim$ and $V/\sim$ are disjoint open sets in $\mathbb{R}P^n$ with $[x] \in U/\sim$ and $[y] \in V/\sim$. Define subsets V_i of $\mathbb{R}P^n$ for $i = 1, \ldots, n+1$ by $V_i = \{[x_1, \ldots, x_{n+1}] : x_i \neq 0\}$. Note that V_i is well-defined. Define maps $\varphi_i : V_i \to \mathbb{R}^n$ for $i = 1, \ldots, n+1$ by $\varphi_i([x_1, \ldots, x_{n+1}]) = \left(\frac{x_1}{x_i}, \frac{x_2}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_{n+1}}{x_i}\right)$. This has the inverse $\varphi_i^{-1}(x_1, \ldots, x_n) = [x_1, x_2, \ldots, x_{i-1}, 1, x_i, \ldots, x_n]$. The maps φ and their inverses are continuous; the open sets V_i cover $\mathbb{R}P^n$; and (for $i < j$)

$$\varphi_i \circ \varphi_j^{-1}(x_1, \ldots, x_n) = \left(\frac{x_1}{x_i}, \frac{x_2}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_{j-1}}{x_i}, \frac{1}{x_i}, \frac{x_j}{x_i}, \ldots, \frac{x_n}{x_i}\right)$$

which is smooth on the set $\varphi_j(V_i \cap V_j) = \{(x_1, \dots, x_n) : \; x_i \neq 0\}$. The case $i > j$ is similar. Therefore the maps φ_i form an atlas for $\mathbb{R}P^n$.

Exercise 1.1.1 Define $\mathbb{C}P^n$ to be $\left(\mathbb{C}^{n+1}\backslash\{0\}\right) / \sim$, where $x \sim y$ if and only if $x = \lambda y$ for some $\lambda \in \mathbb{C}\backslash\{0\}$. Find a differentiable atlas which makes $\mathbb{C}P^n$ a $2n$-dimensional smooth manifold.

Example 1.1.5: Open subsets. Let M be a differentiable manifold, with atlas $\mathcal{A}$. Let U be any open subset of M. Then U is a differentiable manifold with the atlas $\mathcal{A}_{\mathcal{U}} = \{\varphi\big|_{\mathcal{U}} : \; \varphi \in \mathcal{A}\}$. Any open set in a Euclidean space is trivially a manifold. Other examples of manifolds obtained in this way are:

i). The general linear group $GL(n, \mathbb{R})$ (the set of non-singular $n \times n$ matrices) – this is an open set of the set of $n \times n$ matrices, which is naturally identified with $\mathbb{R}^{n^2}$;

ii). The multiplicative group $\mathbb{C}\backslash\{0\}$, clearly open in $\mathbb{C} \simeq \mathbb{R}^2$;

iii). The complement of a Cantor set (i.e. The set of real numbers which do not have a base 3 expansion consisting only of the digits 0 and 2).

Usually we will find ways to avoid things like the last example — by considering manifolds which are connected, or which satisfy some topological or geometric condition (such as compactness).

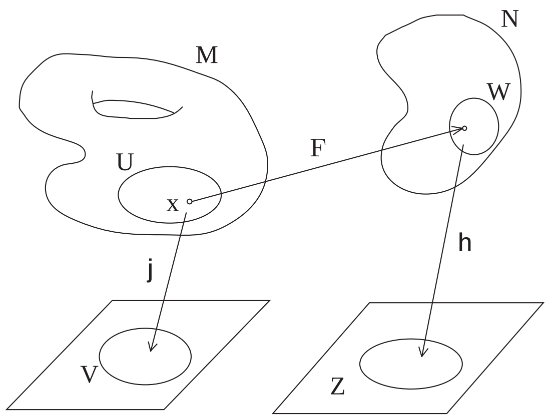
Example 1.1.6: Product manifolds. Let $(M^n, \mathcal{A})$ and $(N^k, \mathcal{B})$ be two manifolds. Then the topological product $M \times N$ can be made a manifold with the atlas $\mathcal{A}\#\mathcal{B} = \{(\varphi, \eta) : \; \varphi \in \mathcal{A}, \; \eta \in \mathcal{B}\}$. Here $(\varphi, \eta)(x, y) = (\varphi(x), \eta(y)) \in \mathbb{R}^{n+k}$ for each $(x, y) \in U \times W$, where $\varphi : U \to V$ is in $\mathcal{A}$ and $\eta : W \to Z$ is in $\mathcal{B}$.

Lecture 2. Smooth functions and maps

2.1 Definition of smooth maps

Given a differentiable manifold, all questions of differentiability are to be reduced to questions about functions between Euclidean spaces, by using charts compatible with the differentiable structure. This principle applies in particular when we decide which maps between manifolds are smooth:

Definition 2.1.1 Let M^n be a differentiable manifold, with an atlas $\mathcal{A}$ representing the differentiable structure on M. A function $f : M \to \mathbb{R}$ is **smooth** if for every chart $\varphi : U \to V$ in $\mathcal{A}$, $f \circ \varphi^{-1}$ is a smooth function on $V \subset \mathbb{R}^n$. Let N^k be another differentiable manifold, with atlas $\mathcal{B}$. Let F be a map from M to N. F is **smooth** if for every $x \in M$ and all charts $\varphi : U \to V$ in $\mathcal{A}$ with $x \in U$ and $\eta : W \to Z$ in $\mathcal{B}$ with $F(x) \in W$, $\eta \circ F \circ \varphi^{-1}$ is a smooth map from $\varphi(F^{-1}(W) \cap U) \subseteq \mathbb{R}^n$ to $Z \subseteq \mathbb{R}^k$.



Remark. Although the definition requires that $f \circ \varphi^{-1}$ be smooth for *every* chart, it is enough to show that this holds for at least one chart around each point: If $\varphi : U \to V$ is a chart with $f \circ \varphi^{-1}$ smooth, and $\eta : W \to Z$ is another chart with W overlapping U, then $f \circ \eta^{-1} = (f \circ \varphi^{-1}) \circ (\varphi \circ \eta^{-1})$ is smooth. A similar argument applies for checking that a map between manifolds is smooth.

Exercise 2.1.1 Show that a map χ between smooth manifolds M and N is smooth if and only if $f \circ \chi$ is a smooth function on M whenever f is a smooth function on N.

Exercise 2.1.2 Show that the map $x \mapsto [x]$ from $\mathbb{R}^{n+1}$ to $\mathbb{R}P^n$ is smooth.

Example 2.1.1 The group $GL(n, \mathbb{R})$. On $GL(n, \mathbb{R})$ we have a natural family of maps: Fix some $\mathsf{M} \in GL(n, \mathbb{R})$, and define $\rho_{\mathsf{M}} : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$ by $\rho_{\mathsf{M}}(\mathsf{A}) = \mathsf{MA}$. The trivial chart ι given by inclusion of $GL(n, \mathbb{R})$ in $\mathbb{R}^{n^2}$ is the map which takes a matrix A to its components $(\mathsf{A}_{11}, \ldots, \mathsf{A}_{1n}, \ldots, \mathsf{A}_{nn})$. To check that ρ_{M} is smooth, we need to check that $\iota \circ \rho_{\mathsf{M}} \circ \iota^{-1}$ is smooth. But this is just the map which takes the components of A to the components of MA, which is

$$\iota^{-1} \circ \rho_{\mathsf{M}} \circ \iota \left(\mathsf{A}_{11}, \mathsf{A}_{12}, \ldots, \mathsf{A}_{nn}\right) = \left(\sum_{j=1}^{n} \mathsf{M}_{1j}\mathsf{A}_{j1}, \sum_{j=1}^{n} \mathsf{M}_{1j}\mathsf{A}_{j2}, \ldots, \sum_{j=1}^{n} \mathsf{M}_{nj}\mathsf{A}_{jn} \right)$$

Each component of this map is linear in the components of A, hence smooth.

Exercise 2.1.3 Show that the multiplication map $\rho : GL(n, \mathbb{R}) \times GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$ given by $\rho(\mathsf{A}, \mathsf{B}) = \mathsf{AB}$ is smooth.

Definition 2.1.2 A **Lie group** G is a group which is also a differentiable manifold such that the multiplication map from $G \times G \to G$ is smooth and the inversion $g \mapsto g^{-1}$ is smooth.

2.2 Further classification of maps

Definition 2.2.1 A map $F : M \to N$ is a **diffeomorphism** if it is smooth and has a smooth inverse. F is a **local diffeomorphism** if for every $x \in M$ there exists a neighbourhood U of x in M such that the restriction of F to U is a diffeomorphism to an open set of N.

A smooth map $F : M \to N$ is an **embedding** if F is a homeomorphism onto its image (with the subspace topology), and for any charts φ and η for M and N respectively, $\eta \circ F \circ \varphi^{-1}$ has derivative of full rank. F is an **immersion** if for every x in M there exist charts $\varphi : U \to V$ for M and $\eta : W \to Z$ for N with $x \in U$ and $F(x) \in Z$, such that the map $\eta \circ F \circ \varphi^{-1}$ has derivative which is injective at $\varphi(x)$. F is a **submersion** if for each $x \in M$ there are charts φ and η such that the derivative of $\eta \circ F \circ \varphi^{-1}$ at $\varphi(x)$ is surjective.

Example 2.2.1 Let G be a Lie group, and $\rho : G \times G \to G$ the multiplication map. For fixed $g \in G$, define $\rho_g : G \to G$ by $\rho_g(h) = \rho(g,h)$. This is a diffeomorphism of G, since ρ_g is smooth and has smooth inverse $\rho_{g^{-1}}$.

Exercise 2.2.1 Show that if M is a manifold, and φ a chart for M, then φ^{-1} is a local diffeomorphism.

Exercise 2.2.2 Consider the map $\pi : S^n \to \mathbb{R}P^n$ defined by $\pi(x) = [x]$ for all $x \in S^n \subset \mathbb{R}^{n+1}$. Show that π is a local diffeomorphism.

Remark. The definitions above reflect the fact that the **rank** of the derivative of a smooth map between manifolds is well-defined: If we change coordinates from φ to $\tilde{\varphi}$ on M and from η to $\tilde{\eta}$ on N, then the chain rule gives:

$$D_{\tilde{\varphi}(x)}\left(\tilde{\eta} \circ F \circ \tilde{\varphi}^{-1}\right) = D_{\eta(x)}\left(\tilde{\eta} \circ \eta^{-1}\right) D_{\varphi(x)}\left(\eta \circ F \circ \varphi^{-1}\right) D_{\tilde{\varphi}(x)}\left(\varphi \circ \tilde{\varphi}^{-1}\right).$$

The first and third matrices on the right are non-singular, so the rank of the matrix on the left is the same as the rank of the second matrix on the right. I will not dwell on this, because it is a simple consequence of the construction of tangent spaces for manifolds, which we will work towards in the next few lectures.

Example 2.2.2 We will show that the projection $[.]$ from $\mathbb{R}^{n+1}\backslash\{0\}$ to $\mathbb{R}P^n$ is a submersion. To see this, fix $x = (x_1, \ldots, x_{n+1})$ in $\mathbb{R}^{n+1}\backslash\{0\}$, and assume that $x_{n+1} \neq 0$. On $\mathbb{R}^{n+1}\backslash\{0\}$ we take the trivial chart ι, and on $\mathbb{R}P^n$ we take $\varphi_{n+1} : V_{n+1} \to \mathbb{R}^n$. Then $\varphi_{n+1} \circ [.] \circ \iota^{-1}(x) = \left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}\right)$, and

$$D_x\left(\varphi_{n+1} \circ [.] \circ \iota^{-1}\right) v = \frac{1}{x_{n+1}} \begin{bmatrix} 1 & 0 & \ldots & 0 & -\frac{x_1}{x_{n+1}} \\ 0 & 1 & \ldots & 0 & -\frac{x_2}{x_{n+1}} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -\frac{x_n}{x_{n+1}} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n+1} \end{bmatrix}$$

This matrix has rank n, so the map is a submersion.

Exercise 2.2.3 Show that the multiplication operator $\rho : G \times G \to G$ is a submersion if G is a Lie group.

Exercise 2.2.4 Let M and N be differentiable manifolds. Show that for any $x \in M$ the map $i_x : N \to M \times N$ given by $y \mapsto (x,y)$ is an embedding. Show that the projection $pi : M \times N \to M$ given by $\pi(x,y) = x$ is a submersion.

Example 2.2.3 We will show that the inclusion map ι of S^n in $\mathbb{R}^{n+1}$ is an embedding. To show this, we need to show that ι is a homeomorphism (this is immediate since we defined S^n by taking the subspace topology induced by inclusion in $\mathbb{R}^{n+1}$) and that the derivative of $\eta \circ \iota \circ \varphi^{-1}$ is injective.

Here we can take $\eta(x) = x$ to be the trivial chart for $\mathbb{R}^{n+1}$, and φ to be one of the stereographic projections defined in Example 1.1.3, say φ_-. Then $\eta \circ \iota \circ \varphi^{-1}(w) = \frac{(2w, 1-|w|^2)}{1+|w|^2}$. Differentiating, we find

$$\frac{\partial(\eta \circ \iota \circ (\varphi^{-1})^i)}{\partial w^j} = \frac{2}{1+|w|^2}\left(\delta_i^j - \frac{2w^i w^j}{1+|w|^2}\right)$$

and

$$\frac{\partial(\eta \circ \iota \circ (\varphi^{-1})^{n+1})}{\partial w^j} = -\frac{4w^j}{(1+|w|^2)^2}.$$

In particular we find for any non-zero vector $v \in \mathbb{R}^n$

$$\left|D_w(\eta \circ \iota \circ \varphi^{-1})(v)\right|^2 = \frac{4}{(1+|w|^2)^2}|v|^2 > 0.$$

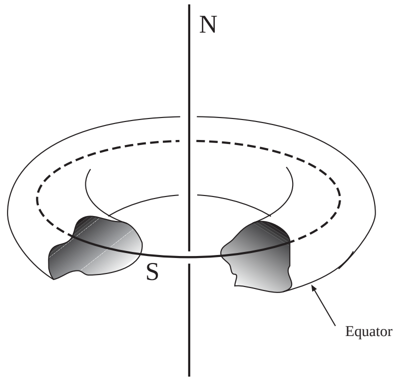
Therefore $D_w(\eta \circ \iota \circ \varphi^{-1})$ has no kernel, hence is injective.

Exercise 2.2.5 Prove that $\mathbb{C}P^1 \simeq S^2$ (Hint: Consider the atlas for S^2 given by the two stereographic projections, and the atlas for $\mathbb{C}P^1$ given by the two projections $[z_1, z_2] \mapsto z_2/z_1 \in \mathbb{C} \simeq \mathbb{R}^2$ and $[z_1, z_2] \mapsto z_1/z_2$. It should be possible to define a map between the two manifolds by defining it between corresponding charts in such a way that it agrees on overlaps).

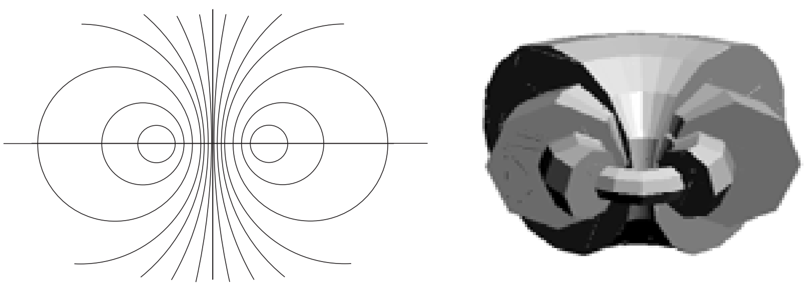
Exercise 2.2.6 (The Hopf map) Consider $S^3 \subseteq \mathbb{R}^4 \simeq \mathbb{C}^2$, and define $\pi : S^3 \to \mathbb{C}P^1 \simeq S^2$ to be the restriction of the canonical projection $(z_1, z_2) \mapsto [z_1, z_2]$. Show that this map is a submersion.

There is a nice way of visualising this map: We can think of S^3 as three-dimensional space (together with infinity). For each point y in $\mathbb{C}P^1$ the set of points in $\mathbb{C}^2$ which project to y is a plane, and the subset of this in S^3 is a circle. Thus we want to decompose space into circles, one corresponding to each point in S^2.

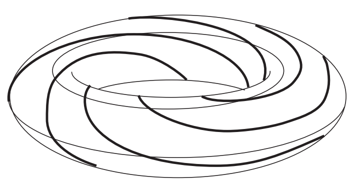
We can start with the north pole, and map this to a circle through infinity (i.e. the z-axis, say). Then we map the south pole to some other circle (say, the unit circle in the x-y plane). Next the equator: Here we have a circle of points, each corresponding to a circle, so this gives us a torus (which is called the Clifford torus):



More generally, consider a circle which moves from south to north pole on S^2. At each 'time' in this sweepout we have a corresponding torus in space, which starts as a thin torus around the unit circle, grows until it agrees with the Clifford torus as our curve passes the equator, and continues growing, becoming larger and larger in size with a smaller and smaller 'tube' down the middle around the z axis as we approach the north pole. This can be obtained by rotating the following picture about the z axis:



Next we should think about how the circles line up on these tori. The crucial thing to keep in mind is that this must be continuous as we move over S^2. In particular, as we approach the south pole the circles have to approach the unit circle, so they have to wind around the torus 'the long way'. On the other hand, as we approach the north pole the circle must start to lie 'along' the z axis. These two things seem to contradict each other, until we realise that we can wind our curves around the tori 'both ways':



Note that each of the circles corresponding to a point in the northern hemisphere intersects the unit disk D_1 in the x-y plane exactly once, and each circle corresponding to a point in the southern hemisphere intersects the right half-plane D_2 in the x-z plane exactly once. We can parametrize each circle through D_1 by an angle θ_1 in such a way that the intersection point with D_1 has $\theta_1 = 0$, and similarly for D_2. Points on the northern hemisphere of S^2 correspond to those circles which pass through a smaller disk $\tilde{D}_1$ in D_1, and those on the southern hemisphere correspond to circles passing through a disk $\tilde{D}_2$ within D_2.

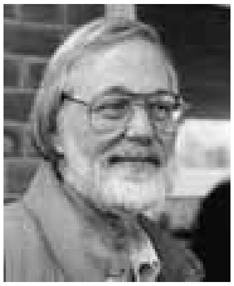
We can parametrize parametrize the boundary of $\tilde{D}_1$ by an angle ϕ_1 (anticlockwise in the x-y plane starting at the right) and that of $\tilde{D}_2$ with an angle ϕ_2 (anticlockwise in the x-z plane starting at the right). Then the angles θ_1 and ϕ_1 give the same point as $\theta_2 = \theta_1 - \phi_1$ and $\phi_2 = \pi - \phi_1$. This describes the 3-sphere as a union of two solid tori $\tilde{D}_1 \times S^1$ and $\tilde{D}_2 \times S^1$, sewn together by identifying points on $\partial\tilde{D}_1 \times S^1$ with points on $\partial\tilde{D}_2 \times S^1$ according to $(\phi, \theta) \in \partial\tilde{D}_1 \times S^1 \sim (\pi - \phi, \theta - \phi) \in \partial\tilde{D}_2 \times S^1$.

Exercise 2.2.7 Suppose we take two solid tori $D_1 \times S^1$ and $D_2 \times S^1$ and glue them together with a different identification, say $(\phi, \theta) \in \partial D_1 \times S^1 \sim (\phi, \theta + k\phi) \in \partial D_2 \times S^1$ for some integer k. How could we decide whether the result is or is not S^3? [This is hard — the answer really belongs to algebraic topology. The idea is to find some quantity which can be associated to a manifold which is invariant under homeomorphism or diffeomorphism (examples are the homotopy groups or homology groups of the manifold). Then if our manifold has a different value of this invariant from S^3, we know it is not diffeomorphic to S^3. Going the other way is much harder, since it is quite possible that non-diffeomorphic manifolds have the same value of any given invariant].

One of the important questions in differential geometry is the study (or classification) of manifolds *up to diffeomorphism*. That is, we introduce an equivalence relation on the space of k-dimensional differentiable manifolds by taking $M \sim N$ if and only if there exists a diffeomorphism from M to N, and study the equivalence classes. Among the results:

- There are only two equivalence classes of connected one-dimensional manifolds, namely those represented by S^1 and by $\mathbb{R}$. In particular the different differentiable structures on $\mathbb{R}$ introduced in Example 1.1.1 are diffeomorphically equivalent (the map $x \mapsto x^{1/3}$ is a diffeomorphism between the two);
- The equivalence class of a compact connected two-dimensional manifold is determined by its genus (an integer representing the number of 'holes' in the surface) and its orientability (which will be defined later).

A more subtle question is the following: Given a topological manifold M, are all differentiable structures on M related by diffeomorphism? If not, how many non-diffeomorphic differentiable structures are there on a given manifold? The answer is that in general there can be more than one differentiable structure on a manifold: It is known that the spheres S^n generally have more than one differentiable structure if $n \geq 7$, and in particular John Milnor proved in 1956 that there are 28 non-diffeomorphic differentiable structures on S^7. Two-dimensional manifolds have unique differentiable structures (this is a consequence of the classification theorem mentioned above), and this is also known for three-dimensional manifolds (although there is no classification of 3-manifolds known). In dimension four it is unknown whether S^4 has more than one differentiable structure, but it is known (through work of Simon Donaldson in the 1980's) that there are four-dimensional manifolds with several (infinitely many) non-diffeomorphic differentiable structures.



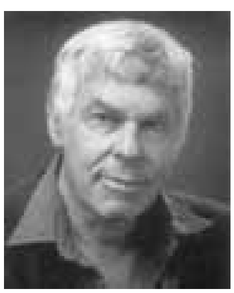
John Milnor

Simon Donaldson

Henri Poincaré

A related question is whether a given topological manifold carries any differentiable structures at all. Again the answer is no, and in particular it is known that there are topological manifolds of dimension 4 which carry no differentiable structure. In dimensions five and higher there are also non-smoothable topological manifolds. In two and three dimensions any manifold carries a differentiable structure.

One further famous problem: Poincaré asked in 1904 whether a 3-manifold which has the same homotopy groups as S^3 must be homeomorphic to S^3 — equivalently, if M is a simply connected compact 3-manifold, is M homeomorphic to S^3? This is the famous Poincaré conjecture, which remains completely open. Remarkably, the analogous problem in higher dimensions — whether an n-manifold with the same homotopy groups as S^n is homeomorphic to S^n — has been solved. Stephen Smale proved this in 1960 for $n \geq 5$ (similar results were also obtained by Stallings and Wallace around the same time). The four-dimensional case was much harder, and was proved by Michael Freedman in 1982 as a result of a truly remarkable theorem classifying closed simply connected four-dimensional topological manifolds. Only the three dimensional case remains unresolved!



Steven Smale

Michael Freedman

Lecture 3. Submanifolds

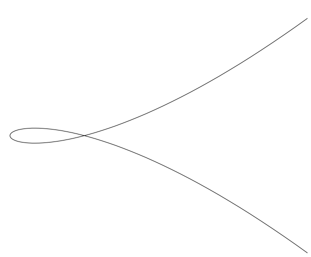
In this lecture we will look at some of the most important examples of manifolds, namely those which arise as subsets of Euclidean space.

2.1 Definition of submanifolds

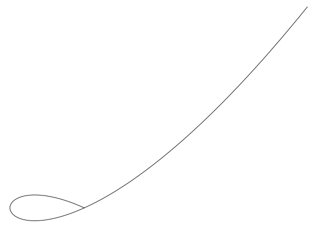
Definition 3.1.1 A subset M of $\mathbb{R}^N$ is a **k-dimensional submanifold** if for every point x in M there exists a neighbourhood V of x in $\mathbb{R}^N$, an open set $U \subseteq \mathbb{R}^k$, and a smooth map $\xi : U \to \mathbb{R}^N$ such that ξ is a homeomorphism onto $M \cap V$, and $D_y\xi$ is injective for every $y \in U$.

Remark. The meaning of 'homeomorphism onto $M \cap V$' in this definition warrants some explanation. This means that the map ξ is continuous and $1:1$, maps U onto $M \cap V$, and the inverse is also continuous. The last statement is to be interpreted as continuity with respect to the 'subspace topology' on M induced from the inclusion into $\mathbb{R}^N$. Since open sets in the subspace topology are given by restrictions of open sets in $\mathbb{R}^N$, this is equivalent to the statement that for every open set $A \subseteq U$ there exists an open set $B \subseteq \mathbb{R}^N$ such that $\xi(A) = M \cap B$.

Example 3.1.2 Consider the following examples of curves in the plane, which illustrate the conditions in the definition of submanifold above: The homeomorphism property can be violated if the function ξ is not injective, for example if $\xi(t) = (t^2, t^3 - t)$ for $t \in (-2, 2)$:

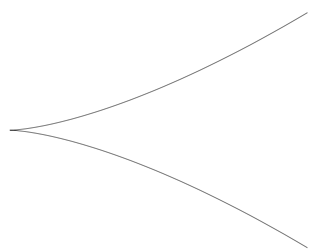


Even if ξ is injective, it might still fail to be a homeomorphism, as in the example $\xi(t) = (t^2, t^3 - t)$ for $t \in (-1, 2)$:



This is not a homeomorphism since $\xi((-0.1, 0.1))$ is not the intersection of the curve with any open set in $\mathbb{R}^2$.

Finally, a smooth map ξ might fail to define a submanifold if the derivative is not injective — in the case of curves, this means the derivative vanishes somewhere. An example is $\xi(t) = (t^2, t^3)$, which describes a cusp:



3.2 Alternative characterisations of submanifolds

Often the definition we have given is not the most convenient way to check whether a given subset of Euclidean space is a submanfold. There are several alternative characterisations of submanifolds, which are given by the following result:

Proposition 3.2.1 *Let M be a subset of Euclidean space $\mathbb{R}^N$. Then the following are equivalent:*

(a). M is a k-dimensional submanifold;

(b). M is a k-dimensional manifold, and can be given a differentiable structure in such a way that the inclusion $i : M \to \mathbb{R}^N$ is an embedding;

(c). *For every $x \in M$ there exists an open set $V \subseteq \mathbb{R}^n$ containing x and an open set $W \subseteq \mathbb{R}^N$ and a diffeomorphism $F : V \to W$ such that $F(M \cap V) = (\mathbb{R} \times \{0\}) \cap W$;*

(d). *M is locally the graph of a smooth function: For every $x \in M$ there exists an open set $V \subseteq \mathbb{R}^N$ containing x, an open set $U \subseteq \mathbb{R}^k$, an permutation $\sigma \in S_N$, and a smooth map $f : U \to \mathbb{R}^{N-k}$ such that*

$$M \cap V = \{(y^1, \ldots, y^N) \in \mathbb{R}^N |\ (y^{\sigma(k+1)}, \ldots, y^{\sigma(N)}) = f(y^{\sigma(1)}, \ldots, y^{\sigma(k)})\}.$$

(e). *M is locally the zero set of a submersion: For every $x \in M$ there exists an open set V containing x and a submersion $G : V \to Z \subseteq \mathbb{R}^{N-k}$ such that $M \cap V = G^{-1}(0)$.*

Our main tool in the proof will be the Inverse function theorem:

Theorem 3.2.2 (Inverse Function Theorem) *Let F be a smooth function from an open neighbourhood of $x \in \mathbb{R}^N$ to $\mathbb{R}^N$, such that the derivative D_xF is an isomorphism. Then there exists an open set A containing x and an open set B containing $F(x)$ such that $F|_A$ is a diffeomorphism from A to B.*

For a proof of Theorem 3.2.2 see Appendix A.

Proof of Proposition 3.2.1: (b)$\Longrightarrow$(a) and (c)$\Longrightarrow$(a) are immediate, as are (d)$\Longrightarrow$(c) and (c)$\Longrightarrow$(e).

Suppose (a) holds, and fix $x \in M$. Choose $\xi_x : U_x \to V_x$ with $x \in M \cap V$ as given in the definition of submanifolds. Let $y = \xi_x^{-1}(x)$. Since $D_y\xi_x$ is injective, we can choose a bijection $\sigma :\ \{1, \ldots, N\} \to \{1, \ldots, N\}$ such that rows $\sigma(1), \ldots, \sigma(k)$ of $D_y\xi_x$ are linearly independent. Define $\pi :\ \mathbb{R}^N \to \mathbb{R}^k$ by $\pi(z^1, \ldots, z^N) = (z^{\sigma(1)}, \ldots, z^{\sigma(k)})$. Then $D_y(\pi \circ \xi_x)$ is an isomorphism, so by the Inverse Function Theorem there are open sets $A \subseteq U$ and $B \subseteq \pi(V) \subset \mathbb{R}^k$ and a smooth map $\eta_x : B \to A$ which is the inverse of $\pi \circ \xi_x|_A$.

Define $\mathcal{A} = \{\varphi_x = \eta_x \circ \pi :\ \ x \in M\}$. Each of these maps is a homeomorphism, since it has a continuous inverse $\xi|_A$. For $x_2 \neq x_1$ we have $\varphi_{x_2} \circ \varphi_{x_1}^{-1} = \varphi_{x_2} \circ \xi_{x_1}$, which is smooth. Therefore $\mathcal{A}$ is a differentiable atlas making M into a differentiable manifold. Finally, the inclusion $i : M \to \mathbb{R}^N$ is a homeomorphism, and for any chart η_x as above, we have $\mathrm{Id} \circ i \circ \eta_x^{-1} = \xi_x$, which has injective derivative. Therefore i is an embedding, and we have established (b). From the proof we can also establish (d), by taking f to be components $\sigma(k+1), \ldots, \sigma(N)$ of $\xi \circ \eta$.

Finally, suppose (e) holds. Let $x \in M$, and let $G : V \to \mathbb{R}^{N-k}$ be a submersion as given in the Proposition. Let $\sigma : \{1, \ldots, N\} \to \{1, \ldots, N\}$ be a bijection, such that columns $\sigma(k+1), \ldots, \sigma(N)$ of D_xG are independent. Define $F : V \to \mathbb{R}^N$ by $F(z) = (z^{\sigma(1)}, \ldots, z^{\sigma(k)}, G(z))$. Then D_xF is an isomorphism, so there exists (locally) an inverse by the Inverse Function Theorem. Define $f(z^1, \ldots, z^k)$ to be components $\sigma(k+1), \ldots, \sigma(N)$ of $F^{-1}(z^1, \ldots, z^k, 0)$. Then (d) holds with this choice of f. $\square$

This proposition gives us a rich supply of manifolds, such as:

(a). The sphere $S^n = \{x \in \mathbb{R}^{n+1} :\ |x| = 1\}$;

(b). The cylinder $(x, y) \in \mathbb{R}^m \oplus \mathbb{R}^n :\ |x| = 1\}$;

(c). The torus $\mathbb{T}^2 = (x, y) \in \mathbb{R}^2 \oplus \mathbb{R}^2 :\ |x| = |y| = 1\}$;

(d). The special linear group $SL(n, \mathbb{R})$ consisting of $n \times n$ matrices with determinant equal to 1 (what dimension is this manifold?);

(e). The orthogonal group $O(n)$ consisting of $n \times n$ matrices A satisfying $\mathsf{A}^T\mathsf{A} = \mathsf{I}$, where I is the $n \times n$ identity matrix (what dimension is this?)

Exercise 3.2.1 Consider the subset of $\mathbb{R}^4$ given by the image of the map $\varphi :\ \mathbb{R} \to \mathbb{R}^4$ defined by

$$\varphi(t) = (\cos t, \sin t, \cos(\sqrt{2}t), \sin(\sqrt{2}t)).$$

Is this a submanifold of $\mathbb{R}^4$?

Definition 3.2.1 A subset M of a manifold N is a k- dimensional **submanifold** of N if for every $x \in M$ and every chart $\varphi : U \to V$ for N with $x \in U$, $\varphi(M \cap U)$ is a k-dimensional submanifold of V.

Exercise 3.2.2 Show that if $M \subset N$ is a submanifold of N then the restriction of every smooth function F on N to M is smooth.

Exercise 3.2.3 Show that the multiplication map $\rho : GL(n, \mathbb{R}) \times GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$ given by $\rho(\mathsf{A}, \mathsf{B}) = \mathsf{AB}$ is smooth.

Exercise 3.2.4 Let M and N be manifolds, and $\chi : M \to N$ a smooth map. Suppose Σ is a submanifold of M, and Γ a submanifold of N.

(i). Show that the restriction of χ to Σ is a smooth map from Σ to N.

(ii). If the $\chi(M) \subset \Gamma$, show that χ is a smooth map from M to Γ.

Example 3.2.1 The groups $SL(n, \mathbb{R})$ and $O(n)$. Each of these groups is contained as a submanifold in $GL(n, \mathbb{R})$. This implies that $SL(n, \mathbb{R}) \times SL(n, \mathbb{R})$ and $O(n) \times O(n)$ are submanifolds of $GL(n, \mathbb{R}) \times GL(n, \mathbb{R})$. Therefore the restriction of the multiplication map ρ on $GL(n, \mathbb{R}) \times GL(n, \mathbb{R})$ to either of these submanifolds is smooth, and has image contained in $SL(n, \mathbb{R})$ or $O(n)$ respectively. Hence by the result of Exercise 3.2.2, the multiplication on $SL(n, \mathbb{R})$ and $O(n)$ are smooth maps.

3.3 Orientability

Definition 3.3.1 An atlas $\mathcal{A}$ for a differentiable manifold M is **orientable** if whenever φ and η in $\mathcal{A}$ have nontrivial common domain of definition, the Jacobian $\det D(\eta \circ \varphi^{-1})$ is positive. A differentiable manifold is **orientable** if there exists such an atlas. An **orientation** on an orientable manifold is an equivalence class of oriented atlases, where two oriented atlases are equivalent if their union is an oriented atlas.

Exercise 3.3.1 Show that every one-dimensional manifold is orientable.

Exercise 3.3.2 Show that every connected manifold has either zero or two orientations.

Example 3.3.1 Hypersurfaces of Euclidean space A submanifold of dimension n in $\mathbb{R}^{n+1}$ is called a *hypersurface*. An orientation on a hypersurface M is equivalent to the choice of a unit normal vector continuously over the whole of M: Given an orientation on the hypersurface, choose the unit normal $\mathbf{N}$ such that for any chart φ in the oriented atlas for M,

$$\det\left[D\varphi^{-1}(e_1), \ldots, D\varphi^{-1}(e_2), \ldots, D_\varphi(e_n), \mathbf{N}\right] > 0. \qquad (+)$$

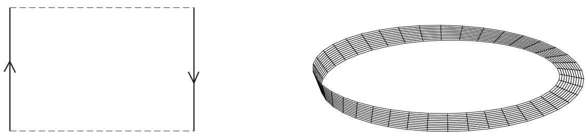
This is continuous on M since it is continuous on overlaps of charts. Conversely, given $\mathbf{N}$ chosen continuously over all of N, we choose an atlas for M consisting of all those charts for which (+) holds.

Exercise 3.3.3 Suppose $F : \mathbb{R}^{n+1} \to \mathbb{R}$ has non-zero derivative everywhere on $M = F^{-1}(0)$. Show that M is orientable.

Example 3.3.2 The Möbius strip and the Klein bottle. The Möbius strip is the topological quotient of $\mathbb{R} \times \mathbb{R}$ by the equivalence relation $\sim$ which identifies (s,t) with $(s+1,-t)$ for every s and t in $\mathbb{R}$. M can be given an atlas as follows: We take a chart $\varphi_1 : (0,1) \times \mathbb{R}/\sim \to (0,1) \times \mathbb{R}$ to be the inverse of the map which takes (s,t) to $[(s,t)]$, and $\varphi_2 : (-1/2,1/2) \times \mathbb{R}/\sim \to (-1/2,1/2) \times \mathbb{R}$ similarly. Then

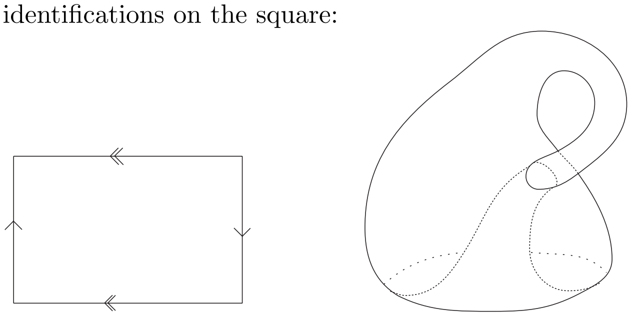
$$\varphi_1 \circ \varphi_2^{-1}(s,t) = \begin{cases} (s,t) & \text{if } s \in (0,1/2) \\ (s+1,t) & \text{if } s \in (-1/2,0) \end{cases}$$

which is smooth on $((-1/2,0) \cup (0,1/2)) \times \mathbb{R}$. The other transition map is similar.



The Möbius strip is not orientable. I will not prove this rigorously yet. Heuristically, the idea is that if we take an oriented pair of vectors at some point $(s, 0)$, and 'slide' them around the Möbius strip to $(s + 1, 0)$, then if there were an oriented atlas it would have to be the case that the vectors obtained in this way were still oriented with respect to the original pair. But this is not the case.

The Klein bottle is another non-orientable surface, given by making the following identifications on the square:



Example 3.3.3 The real projective space $\mathbb{R}P^2$ is another example of a non-orientable surface. One way to visualise $\mathbb{R}P^2$ is as the semisphere with points on the equator identified with their antipodal points.

Some general nonsense: Definition 3.3.1 is one of many similar definitions for special classes of manifolds. More generally, suppose we have a class of maps $\mathcal{M} = \{\varphi : U \to V\}$ between open subsets of a Euclidean space $\mathbb{R}^n$, such that

(*) $\mathcal{M}$ is closed under composition: If $\varphi : U \to V$ and $\eta : W \to Z$ are in $\mathcal{M}$, and $V \cap W$ is non-empty, then $\eta \circ \varphi : \varphi^{-1}(V \cap W) \to Z$ is in $\mathcal{M}$

then one can consider $\mathcal{M}$-manifolds, by requiring the transition maps $\varphi \circ \eta^{-1}$ of an atlas to be in the class $\mathcal{M}$. Some examples are:

(1). The class of continuous maps. This gives rise to **topological manifolds**;

(2). The class of maps which are k times continuously differentiable. The resulting manifolds are C^k **manifolds**;

(3). The class of maps for which each component has a convergent power series (i.e. is a real-analytic function). This gives **real-analytic manifolds**;

(4). The class of maps between open sets of $\mathbb{C}^n$ which are holomorphic – that is, each complex component of the map is given by a convergent power series in the n complex variables. This defines **complex manifolds**;

(5). The class of maps of the form $x \mapsto \mathsf{M}x + v$ where M is in some subgroup G of the general linear group $GL(n, \mathbb{R})$ — such as $SL(n, \mathbb{R})$ (which gives **affine-flat manifolds**), or $O(n)$ (which gives **Euclidean manifolds** or **flat manifolds**);

(6). The class of maps F which have derivative DF in a subgroup G of $GL(n, \mathbb{R})$.

and so on...

Lecture 4. Tangent vectors

4.1 The tangent space to a point

Let M^n be a smooth manifold, and x a point in M. In the special case where M is a submanifold of Euclidean space $\mathbb{R}^N$, there is no difficulty in defining a space of tangent vectors to M at x: Locally M is given as the zero level-set of a submersion $G : U \to \mathbb{R}^{N-n}$ from an open set U of $\mathbb{R}^N$ containing x, and we can define the tangent space to be $\ker(D_xG)$, the subspace of vectors which map to 0 under the derivative of G. Alterntively, if we describe M locally as the image of an embedding $\varphi : U \to \mathbb{R}^N$ from an open set U of $\mathbb{R}^n$, then we can take the tangent space to M at x to be the subspace $\operatorname{rng}\left(D_{\varphi^{-1}(x)}\varphi\right) = \{D_{\varphi^{-1}(x)}\varphi(u) :\ u \in \mathbb{R}^n\}$, the image subspace of the derivative map.

If M is an abstract manifold, however, then we do not have any such convenient notion of a tangent vector.

From calculus on $\mathbb{R}^n$ we have several complementary ways of thinking about tangent vectors: As k-tuples of real numbers; as 'directions' in space, such as the tangent vector of a curve; or as directional derivatives.

I will give three alternative candidates for the tangent space to a smooth manifold M, and then show that they are equivalent:

First, for $x \in M$ we define T_xM to be the set of pairs (φ, u) where $\varphi : U \to V$ is a chart in the atlas for M with $x \in U$, and u is an element of $\mathbb{R}^n$, modulo the equivalence relation which identifies a pair (φ, u) with a pair (η, w) if and only if u maps to w under the derivative of the transition map between the two charts:

$$D_{\varphi(x)}\left(\eta \circ \varphi^{-1}\right)(u) = w. \tag{4.1}$$

Remark. The basic idea is this: We think of a vector as being an 'arrow' telling us which way to move inside the manifold. This information on which way to move is encoded by viewing the motion through a chart φ, and seeing which way we move 'downstairs' in the chart (this corresponds to a vector in n-dimensional space according to the usual notion of a velocity vector). The equivalence relation just removes the ambiguity of a choice of chart through which to follow the motion.

Another way to think about it is the following: We have a local description for M using charts, and we know what a vector is 'downstairs' in each chart.

We want to define a space of vectors T_xM 'upstairs' in such a way that the derivative map $D_x\varphi$ of the chart map φ makes sense as a linear operator between the vector spaces T_xM and $\mathbb{R}^n$, and so that the chain rule continues to hold. Then we would have for any vector $v \in T_pM$ vectors $u = D_x\varphi(v) \in \mathbb{R}^n$, and $w = D_x\eta(v) \in \mathbb{R}^n$. Writing this another way (implicitly assuming the chain rule holds) we have

$$D_{\varphi(x)}\varphi^{-1}(u) = v = D_{\eta(x)}\eta^{-1}(w).$$

The chain rule would then imply $D_{\varphi(x)}\left(\eta \circ \varphi^{-1}\right)(u) = w$.

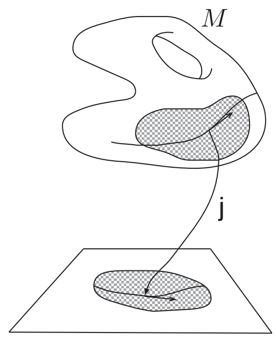


Fig.1: A tangent vector to M at x is implicitly defined by a curve through x

The second definition expresses even more explicitly the idea of a 'velocity vector' in the manifold: We define M_x to be the space of smooth paths in M through x (i.e. smooth maps $\gamma : I \to M$ with $\gamma(0) = x$) modulo the equivalence relation which identifies any two curves if they agree to first order (as measured in some chart): $\gamma \sim \sigma \Leftrightarrow (\varphi\circ\gamma)'(0) = (\varphi\circ\sigma)'(0)$ for some chart $\varphi : U \to V$ with $x \in U$. The equivalence does not depend on the choice of chart: If we change to a chart η, then we have

$$(\eta\circ\gamma)'(0) = \left((\eta\circ\varphi^{-1})\circ(\varphi\circ\gamma)\right)'(0) = D_{\varphi(x)}\left(\eta\circ\varphi^{-1}\right)(\varphi\circ\gamma)'(0)$$

and similarly for sigma, so $(\eta\circ\gamma)'(0) = (\eta\circ\sigma)'(0)$.

Finally, we define D_xM to be the space of derivations at x. Here a derivation is a map v from the space of smooth functions $C^\infty(M)$ to $\mathbb{R}$, such that for any real numbers c_1 and c_2 and any smooth functions f and g on M,

$$\begin{aligned} v(c_1f + c_2g) &= c_1v(f) + c_2v(g) \qquad \text{and} \\ v(fg) &= f(x)v(g) + g(x)v(f) \end{aligned} \tag{4.2}$$

The archetypal example of a derivation is of course the directional derivative of a function along a curve: Given a smooth path $\gamma : I \to M$ with $\gamma(0) = x$, we can define

$$v(f) = \frac{d}{dt}\left(f \circ \gamma\right)\Big|_{t=0}, \tag{4.3}$$

and this defines a derivation at x. We will see below that all derivations are of this form.

Proposition 4.1.1 *There are natural isomorphisms between the three spaces M_x, T_xM, and D_xM.*

$$\begin{tikzcd}
TM_x \arrow[rr, "a"] & & M_x \arrow[dl, "b"] \\
& D_xM \arrow[ul, "c"] &
\end{tikzcd}$$

Proof. First, we write down the isomorphisms: Given an equivalence class $[(\varphi, u)]$ in T_xM, we take $\alpha([(\varphi, u)])$ to be the equivalence class of the smooth path γ defined by

$$\gamma(t) = \varphi^{-1}(\varphi(x) + tu). \tag{4.4}$$

The map α is well-defined, since if (η, w) is another representative of the same equivalence class in T_xM, then α gives the equivalence class of the curve $\sigma(t) = \eta^{-1}(\eta(x) + tw)$, and

$$\begin{aligned}
(\varphi \circ \sigma)'(0) &= \left((\varphi \circ \eta^{-1}) \circ (\eta \circ \sigma)\right)'(0) \\
&= D_{\eta(x)}\left(\varphi \circ \eta^{-1}\right)(\eta \circ \sigma)'(0) \\
&= D_{\eta(x)}\left(\varphi \circ \eta^{-1}\right)(w) \\
&= u \\
&= (\varphi \circ \gamma)'(0),
\end{aligned}$$

so $[\sigma] = [\gamma]$.

Given an element $[\gamma] \in M_x$, we take $\beta([\gamma])$ to be the natural derivation v defined by Eq. (4.3). Again, we need to check that this is well-defined: Suppose $[\sigma] = [\gamma]$. Then

$$\begin{aligned}
\frac{d}{dt}\left(f \circ \gamma\right)\Big|_{t=0} &= D_{\varphi(x)}\left(f \circ \varphi^{-1}\right)(\varphi \circ \gamma)'(0) \\
&= D_{\varphi(x)}\left(f \circ \varphi^{-1}\right)(\varphi \circ \sigma)'(0) \\
&= \frac{d}{dt}\left(f \circ \sigma\right)\Big|_{t=0}
\end{aligned}$$

for any smooth function f.

Finally, given a derivation v, we choose a chart φ containing x, and take $\chi(v)$ to be the element of T_xM given by taking the equivalence class $[(\varphi, u)]$ where $u = \big(v(\varphi^1), \ldots, v(\varphi^n)\big)$. Here φ^i is the ith component function of the chart φ.

There is a technicality involved here: In the definition, derivations were assumed to act on smooth functions defined on all of M. However, φ^i is defined only on an open set U of M. In order to overcome this difficulty, we will extend φ_i (somewhat arbitrarily) to give a smooth map on all of M. For concreteness, we can proceed as follows (the functions we construct here will prove useful later on as well):

$$\xi(z) = \begin{cases} \exp\left(\frac{1}{z^2-1}\right) & \text{for } -1 < z < 1; \\ 0 & \text{for } |z| \geq 1. \end{cases} \tag{4.5}$$

and

$$\rho(z) = \frac{\int_{-1}^{z} \xi(z')dz'}{\int_{-1}^{1} \xi(z')dz'}. \tag{4.6}$$

Then

- ξ is a C^∞ function on all of $\mathbb{R}$, with $\xi(z)$ equal to zero whenever $|z| \geq 1$, and $\xi(z) > 0$ for $|z| < 1$;
- ρ is a C^∞ function on all of $\mathbb{R}$, which is zero whenever $z \leq -1$, identically 1 for $z > 1$, and strictly increasing for $z \in (-1, 1)$.

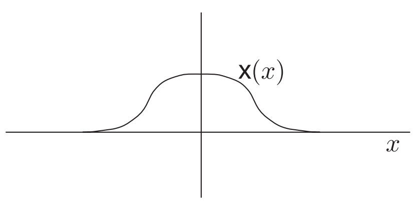


Fig. 2: The 'bump function' or 'cut-off function' ξ.

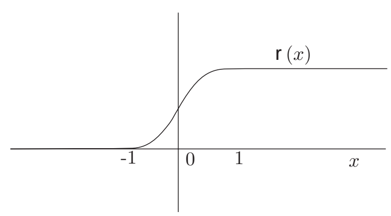


Fig. 3: A C^∞ 'ramp function' ρ.

Now, given the chart $\varphi : U \to V$ with $x \in U$, choose a number r sufficiently small to ensure that the closed ball of radius $4r$ about $\varphi(x)$ is contained in V. Then define a function $\tilde{\rho}$ on M by

$$\tilde{\rho}(y) = \begin{cases} \rho\left(3 - \frac{|\varphi(y)-\varphi(x)|}{r}\right) & \text{for } y \in U; \\ 0 & \text{for all other } y \in M. \end{cases} \tag{4.7}$$

Exercise 4.1 Prove that $\tilde{\rho}$ is a smooth function on M.

Note that this construction gives a function $\tilde{\rho}$ which is identically equal to 1 on a neighbourhood of x, and identically zero in the complement of a larger neighbourhood.

Now we can make sense of our definition above:

Definition 4.1.2 If $f : U \to \mathbb{R}$ is a smooth function on an open set U of M containing x, we define $v(f) = v(\tilde{f})$, where $\tilde{f}$ is any smooth function on M which agrees with f on a neighbourhood of x.

For this to make sense we need to check that there is some smooth function $\tilde{f}$ on M which agrees with f on a neighbourhood of x, and that the definition does not depend on which such function we choose.

Without loss of generality we suppose f is defined on U as above, and define $\tilde{\rho}$ as above. Then we define

$$\tilde{f}(y) = \begin{cases} f(y)\tilde{\rho}(y) & \text{for } y \in U; \\ 0 & \text{for } y \in M \backslash U. \end{cases} \tag{4.8}$$

$\tilde{f}$ is smooth, and agrees with f on the set $\varphi^{-1}(B_{2r}(\varphi(x)))$.

Next we need to check that $v(\tilde{f})$ does not change if we choose a different function agreeing with f on a neighbourhood of x.

Lemma 4.1.3 *Suppose f and g are two smooth functions on M which agree on a neighbourhood of x. Then $v(f) = v(g)$.*

Proof. Without loss of generality, assume that f and g agree on an open set U containing x, and construct a 'bump' function $\tilde{\rho}$ in U as above. Then we observe that $\tilde{\rho}(f - g)$ is identically zero on M, and that $v(0) = v(0.0) = 0.v(0) = 0$. Therefore

$$0 = v(\tilde{\rho}(f - g)) = \tilde{\rho}(x)v(f - g) + (f(x) - g(x))v(\tilde{\rho}) = v(f) - v(g)$$

since $f(x) - g(x) = 0$ and $\tilde{\rho}(x) = 1$. □

This shows that the definition of $v(f)$ makes sense, and so our definition of $\chi(v)$ makes sense. However we still need to check that $\chi(v)$ does not depend on the choice of a chart φ. Suppose we instead use another chart η. Then we have in a small region about x,

$$\eta^i(y) = \eta^i(x) + \sum_{j=1}^{n} G_j^i(y)(\varphi^j(y) - \varphi^j(x)) \tag{4.9}$$

for each $i = 1, \ldots, n$, where G_j^i is a smooth function on a neighbourhood of x for which $G_j^i(x) = \frac{\partial}{\partial z^j}\left(\eta^i \circ \varphi^{-1}\right)\Big|_{\varphi(x)}$. To prove this, consider the Taylor expansion for $\eta^i \circ \varphi^{-1}$ on the set V, where φ is the chart from U to V.

Now apply v to Eq. (4.9). Note that the first term is a constant.

Lemma 4.1.4 *$v(c) = 0$ for any constant c.*

Proof.

$$v(1) = v(1.1) = 1.v(1) = 1.v(1) = 2v(1) \Longrightarrow v(1) = 0 \Longrightarrow v(c) = cv(1) = 0.$$

□

v applied to η^i gives

$$\begin{aligned} v(\eta^i) &= \sum_{j=1}^n G^i_j(x)v(\varphi^j) + (\varphi^j(x) - \varphi^j(x))v(G^i_j) \\ &= \sum_{j=1}^n \frac{\partial}{\partial z^j}\left(\eta^i \circ \varphi^{-1}\right)\Big|_{\varphi(x)} v(\varphi^j). \end{aligned}$$

Therefore we have

$$\begin{aligned} \sum_{i=1}^n v(\eta^i)e_i &= \sum_{i,j=1}^n \frac{\partial}{\partial z^j}\left(\eta^i \circ \varphi^{-1}\right)\Big|_{\varphi(x)} v(\varphi^j)e_i \\ &= D_{\varphi(x)}\left(\eta \circ \varphi^{-1}\right)\left(\sum_{i=1}^n v(\varphi^i)e_i\right) \end{aligned}$$

and so $[(\varphi, \sum_{i=1}^n v(\varphi^i)e_i)] = [(\eta, \sum_{i=1}^n v(\eta^i)e_i)]$ and χ is independent of the choice of chart.

In order to prove the proposition, it is enough to show that the three triple compositions $\chi \circ \beta \circ \alpha$, $\beta \circ \alpha \circ \chi$, and $\alpha \circ \chi \circ \beta$ are just the identity map on each of the three spaces.

We have

$$\begin{aligned} \chi \circ \beta \circ \alpha([(\varphi, u)]) &= \chi \circ \beta\left([t \mapsto \varphi^{-1}\left(\varphi(x) + tu\right)]\right) \\ &= \chi\left(f \mapsto D_{\varphi(x)}\left(f \circ \varphi^{-1}\right)(u)\right) \\ &= \left[\left(\varphi, \sum_{i=1}^n D_{\varphi(x)}\left(\varphi^i \circ \varphi^{-1}\right)(u)e_i\right)\right] \\ &= [(\varphi, u)]. \end{aligned}$$

Similarly we have

$$\begin{aligned} \alpha \circ \chi \circ \beta([\sigma]) &= \alpha \circ \chi\left(f \mapsto \frac{d}{dt}\left(f \circ \gamma\right)\Big|_{t=0}\right) \\ &= \alpha\left(\left[\left(\varphi, \sum_{i=1}^n \left(\varphi \circ \gamma\right)'(0)\right)\right]\right) \\ &= \left[t \mapsto \varphi^{-1}\left(\varphi(x) + t\left(\varphi \circ \gamma\right)'(0)\right)\right], \end{aligned}$$

and this curve is clearly in the same equivalence class as γ.

Finally, we have

$$\begin{aligned}\beta\circ\alpha\circ\chi(v) &= \beta\circ\alpha\left(\left[\left[\left(\varphi,\sum_{i=1}^{n}v(\varphi^i)e_i\right)\right]\right]\right)\\ &= \beta\left(\left[t\mapsto\varphi^{-1}\left(\varphi(x)+t\sum_{i=1}^{n}v(\varphi^i)e_i\right)\right]\right)\\ &= \left(f\mapsto\sum_{i=1}^{n}\frac{\partial}{\partial z^i}\left(f\circ\varphi^{-1}\right)\Big|_{\varphi(x)}v(\varphi^i)\right).\end{aligned}$$

We need to show that this is the same as v. To show this, we note (using the Taylor expansion for $f\circ\varphi^{-1}$) that

$$f(y)=f(x)+\sum_{i=1}^{n}G_i(y)\left(\varphi^i(y)-\varphi^i(x)\right)$$

for y in a sufficiently small neighbourhood of x, where G_i is a smooth function with $G_i(x)=\frac{\partial}{\partial z^i}\left(f\circ\varphi^{-1}\right)|_{\varphi(x)}$. Applying v to this expression, we find

$$v(f)=\sum_{i=1}^{n}G_i(x)v(\varphi^i)=\sum_{i=1}^{n}\frac{\partial}{\partial z^i}\left(f\circ\varphi^{-1}\right)\Big|_{\varphi(x)}v(\varphi^i) \tag{4.10}$$

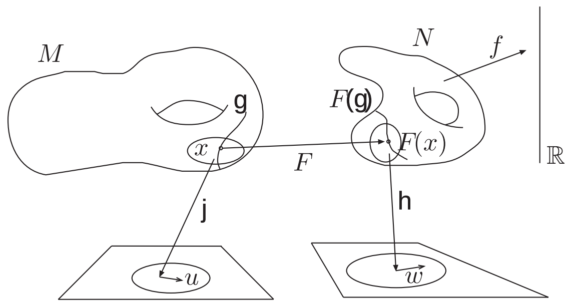
and the right hand side is the same as $\beta\circ\alpha\circ\gamma v(f)$.

Having established the equivalence of the three spaces M_x, T_x, and D_xM, I will from now on keep only the notation T_xM (the tangent space to x at M) while continuing to use all three different notions of a tangent vector.

4.2 The differential of a map

Definition 4.2.1 Let $f:M\to\mathbb{R}$ be a smooth function. Then we define the differential d_xf of f at the point $x\in M$ to be the linear function on the tangent space T_xM given by $(d_xf)(v)=v(f)$ for each $v\in T_xM$ (thinking of v as a derivation). Let $F:M\to N$ be a smooth map between two manifolds. Then we define the differential D_xF of F at $x\in M$ to be the linear map from T_xM to $T_{F(x)}N$ given by $((D_xF)(v))(f)=v(f\circ F)$ for any $v\in T_xM$ and any $f\in C^\infty(M)$.

It is useful to describe the differential of a map in terms of the other representations of tangent vectors. If v is the vector corresponding to the equivalence class $[(\varphi,u)]$, then we have $v:f\mapsto D_{\varphi(x)}(f\circ\varphi^{-1})(u)$, and so by the definition above, $D_xF(v)$ sends a smooth function f on N to $v(f\circ F)$:



$$D_xF(v) : f \mapsto D_{\varphi(x)}\left(f \circ F \circ \varphi^{-1}\right)(u)$$
$$= D_{\eta(F(x))}\left(f \circ \eta^{-1}\right) \circ D_{\varphi(x)}\left(\eta \circ F \circ \varphi^{-1}\right)(u)$$

which is the vector corresponding to $\left[\left(\eta, D_{\varphi(x)}\left(\eta \circ F \circ \varphi^{-1}\right)(u)\right)\right]$.

Alternatively, if we think of a vector v as the tangent vector of a curve γ, then we have $v: \ f \mapsto (f \circ \gamma)'(0)$, and so $D_xF(v): \ f \mapsto (f \circ F \circ \gamma)'(0)$, which is the tangent vector of the curve $F \circ \gamma$. In other words,

$$D_xF([\gamma]) = [F \circ \gamma]. \tag{4.11}$$

In most situations we can use the differential of a map in exactly the same way as we use the derivative for maps between Euclidean spaces. In particular, we have the following results:

Theorem 4.2.2 The Chain Rule *If $F : M \to N$ and $G : N \to P$ are smooth maps between manifolds, then so is $G \circ F$, and*

$$D_x(G \circ F) = D_{F(x)}G \circ D_xF.$$

Proof. By Eq. (4.11),

$$D_x(G \circ F)([\gamma]) = [G \circ F \circ \gamma] = D_{F(x)}G([F \circ \gamma]) = D_{F(x)}G\left(D_xF([\gamma])\right).$$

□

Theorem 4.2.3 The inverse function theorem *Let $F : M \to N$ be a smooth map, and suppose D_xF is an isomorphism for some $x \in M$. Then there exists an open set $U \subset M$ containing x and an open set $V \subset N$ containing $F(x)$ such that $F\big|_U$ is a diffeomorphism from U to V.*

Proof. We have $D_xF([(\varphi, u)]) = [\eta, D_{\varphi(x)}\left(\eta \circ F \circ \varphi^{-1}\right)(u)]$, so D_xF is an isomorphism if and only if $D_{\varphi(x)}\left(\eta \circ F \circ \varphi^{-1}\right)$ is an isomorphism. The result follows by applying the usual inverse function theorem to $\eta \circ F \circ \varphi^{-1}$.

Theorem 4.2.4 The implicit function theorem (surjective form) *Let $F : M \to N$ be a smooth map, with D_xF surjective for some $x \in M$. Then there exists a neighbourhood U of x such that $F^{-1}(F(x)) \cap U$ is a smooth submanifold of M.*

Theorem 4.2.5 The implicit function theorem (injective form) *Let $F : M \to N$ be a smooth map, with D_xF injective for some $x \in M$. Then there exists a neighbourhood U of x such that $F\Big|_U$ is an embedding.*

These two theorems follow directly from the corresponding theorems for smooth maps between Euclidean spaces.

4.3 Coordinate tangent vectors

Given a chart $\varphi : U \to V$ with $x \in U$, we can construct a convenient basis for T_xM: We simply take the vectors corresponding to the equivalence classes $[(\varphi, e_i)]$, where $e_1, \ldots, e_n$ are the standard basis vectors for $\mathbb{R}^n$. We use the notation $\partial_i = [(\varphi, e_i)]$, suppressing explicit mention of the chart φ. As a derivation, this means that $\partial_i f = \frac{\partial}{\partial z^i}\left(f \circ \varphi^{-1}\right)\Big|_{\varphi(x)}$. In other words, ∂_i is just the derivation given by taking the ith partial derivative in the coordinates supplied by φ. It is immediate from Proposition 4.1.1 that $\{\partial_1, \ldots, \partial_n\}$ is a basis for T_xM.

4.4 The tangent bundle

We have just constructed a tangent space at each point of the manifold M. When we put all of these spaces together, we get the *tangent bundle* TM of M:

$$TM = \{(p, v) : p \in M, v \in T_pM\}.$$

If M has dimension n, we can endow TM with the structure of a $2n$-dimensional manifold, as follows: Define $\pi : TM \to M$ to be the projection which sends (p, v) to p. Given a chart $\varphi : U \to V$ for M, we can define a chart $\tilde{\varphi}$ for TM on the set $\pi^{-1}(U) = \{(p, v) \in TM : p \in U\}$, by

$$\tilde{\varphi}(p, v) = \left(\varphi(p), v(\varphi^1), \ldots, v(\varphi^n)\right) \in \mathbb{R}^{2n}.$$

Thus the first n coordinates describe the point p, and the last n give the components of the vector v with respect to the basis of coordinate tangent vectors $\{\partial_i\}_{i=1}^n$, since by Eq. (4.10),

$$v(f) = \sum_{i=1}^n \frac{\partial}{\partial z^i}\left(f\circ\varphi^{-1}\right)\Big|_{\varphi(x)} v(\varphi^i) = \sum_{i=1}^n v(\varphi^i)\partial_i(f) \tag{4.12}$$

for any smooth f, and hence $v = \sum_{i=1}^n v(\varphi^i)\partial_i$. For convenience we will often write the coordinates on TM as $(x^1,\dots x^n,\dot{x}^1,\dots,\dot{x}^n)$

To check that these charts give TM a manifold structure, we need to compute the transition maps. Suppose we have two charts $\varphi : U \to V$ and $\eta : W \to Z$, overlapping non-trivially. Then $\tilde{\eta}\circ\tilde{\varphi}^{-1}$ first takes a $2n$-tuple $(x^1,\dots,x^n,\dot{x}^1,\dots,\dot{x}^n)$ to the element $\left(\varphi^{-1}(x^1,\dots,x^n), \sum_{i=1}^n \dot{x}^i\partial_i^{(\varphi)}\right)$ of TM, then maps this to $\mathbb{R}^{2n}$ by $\tilde{\eta}$. Here we add the superscript (φ) to distinguish the coordinate tangent vectors coming from the chart φ from those given by the chart η. The first n coordinates of the result are then just $\eta\circ\varphi^{-1}(x^1,\dots,x^n)$. To compute the last n coordinates, we need to $\partial_i^{(\varphi)}$ in terms of the coordinate tangent vectors $\partial_j^{(\eta)}$: We have

$$\begin{aligned}
\partial_i^{(\varphi)} f &= D_{\varphi(x)}\left(f\circ\varphi^{-1}\right)(e_i)\\
&= D_{\varphi(x)}\left(\left(f\circ\eta^{-1}\right)\circ\left(\eta\circ\varphi^{-1}\right)\right)(e_i)\\
&= D_{\eta(x)}\left(f\circ\eta^{-1}\right)\circ D_{\varphi(x)}\left(\eta\circ\varphi^{-1}\right)(e_i)\\
&= D_{\eta(x)}\left(f\circ\eta^{-1}\right)\left(\sum_{j=1}^n D_{\varphi(x)}\left(\eta\circ\varphi^{-1}\right)_i^j e_j\right)\\
&= \sum_{j=1}^n D_{\varphi(x)}\left(\eta\circ\varphi^{-1}\right)_i^j D_{\eta(x)}\left(f\circ\eta^{-1}\right)(e_j)\\
&= \sum_{j=1}^n D_{\varphi(x)}\left(\eta\circ\varphi^{-1}\right)_i^j \partial_j^{(\eta)} f
\end{aligned}$$

for every smooth function f. Therefore

$$\sum_{i=1}^n \dot{x}^i\partial_i^{(\varphi)} = \sum_{i,j=1}^n D_{\varphi(x)}\left(\eta\circ\varphi^{-1}\right)_i^j \partial_j^{(\eta)}$$

and

$$\tilde{\eta}\circ\tilde{\varphi}^{-1}(x,\dot{x}) = \left(\eta\circ\varphi^{-1}(x), D_{\varphi(x)}\left(\eta\circ\varphi^{-1}\right)(\dot{x})\right).$$

Since $\eta\circ\varphi^{-1}$ is smooth by assumption, so is its matrix of derivatives $D(\eta\circ\varphi^{-1})$, so $\tilde{\eta}\circ\tilde{\varphi}^{-1}$ is a smooth map on $\mathbb{R}^{2n}$.

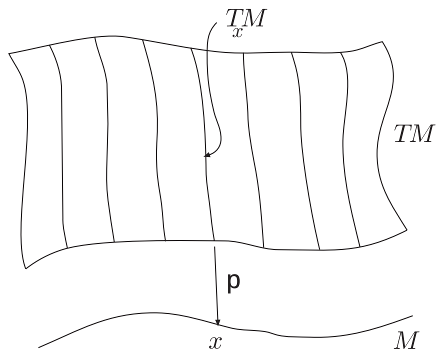


Fig.4: Schematic depiction of the tangent bundle TM as a 'bundle' of the fibres T_xM over the manifold M.

4.5 Vector fields

A vector field on M is given by choosing a vector in each of the tangent spaces T_xM. We require also that the choice of vector varies smoothly, in the following sense: A choice of V_x in each tangent space T_xM gives us a map from M to TM, $x \mapsto V_x$. A smooth vector field is defined to be one for which the map $x \mapsto V_x$ is a smooth map from the manifold M to the manifold TM.

In order to check whether a vector field is smooth, we can work locally. In a chart φ, the vector field can be written as $V_x = \sum_{i=1}^n V_x^i \partial x_i$, and this gives us n functions $V_x^1, \dots, V_x^n$. Then it is easy to show that V is a smooth vector field if and only if these component functions are smooth as functions on M. That is, when viewed through a chart the vector field is smooth, in the usual sense of an n-tuple of smooth functions.

Our notion of a tangent vector as a derivation allows us to think of a vector field in another way:

Proposition 4.5.1 *Smooth vector fields are in one-to-one correspondence with derivations $V : C^\infty(M) \to C^\infty(M)$ satisfying the two conditions*

$$\begin{aligned} V(c_1 f_1 + c_2 f_2) &= c_1 V(f_1) + c_2 V(f_2) \\ V(f_1 f_2) &= f_1 V(f_2) + f_2 V(f_1) \end{aligned}$$

for any constants c_1 and c_2 and any smooth functions f_1 and f_2.

Proof : Given such a derivation V, and $p \in M$, the map $V_p : C^\infty(M) \to \mathbb{R}$ given by $V_p(f) = (V(f))(p)$ is a derivation at p, and hence defines an element of T_pM. The map $p \mapsto V_p$ from M to TM is therefore a vector field, and it remains to check smoothness. The components of V_p with respect to the coordinate tangent basis $\partial_1, \dots, \partial_n$ for a chart $\varphi : U \to V$ is given by

$$V_p^i = V_p(\varphi^i)$$

which is by assumption a smooth function of p for each i (since φ^i is a smooth function and V maps smooth functions to smooth functions – here one should really multiply φ^i by a smooth cut-off function to convert it to a smooth function on the whole of M). Therefore the vector field is smooth.

Conversely, given a smooth vector field $x \mapsto V_x \in T_xM$, the map

$$(V(f))(x) = V_x(f)$$

satisfies the two conditions in the proposition and takes a smooth function to a smooth function. $\square$

A common notation is to refer to the space of smooth vector fields on M as $\mathcal{X}(\mathcal{M})$. Over a small region of a manifold (such as a chart), the space of smooth vector fields is in 1 : 1 correspondence with n-tuples of smooth functions. However, when looked at over the whole manifold things are not so simple. For example, a theorem of algebraic topology says that there are no continuous vector fields on the sphere S^2 which are everywhere non-zero ("the sphere has no hair"). On the other hand there are certainly nonzero functions on S^2 (constants, for example).

Lecture 5. Lie Groups

In this lecture we will make a digression from the development of geometry of manifolds to discuss an very important special case.

5.1 Examples

Recall that a Lie Group is a group with the structure of a smooth manifold such that the composition from $M \times M \to M$ and the inversion from $M \to M$ are smooth maps.

Example 5.1.1 The general linear, special linear, and orthogonal groups. The general linear group $GL(n)$ (or $GL(n, \mathbb{R})$) is the set of non-singular $n \times n$ matrices with real components. The is a Lie group under the usual multiplication of matrices. We showed in Example 2.1.1 that the multiplication on $GL(n)$ is a smooth map from $GL(n) \times GL(n) \to GL(n)$. To show that $GL(n)$ is a Lie group, we nned to show that the inversion is also a smooth map.

$GL(n)$ is an open subset of $M^n \simeq \mathbb{R}^{n^2}$, so is covered by a single chart. With respect to this chart, the inversion is given by the formula

$$\left(M^{-1}\right)_i^j = \frac{1}{n! \det M} \sum_{\sigma,\tau} \operatorname{sgn}\sigma \ \operatorname{sgn}\tau \ M^{\tau(1)}_{\sigma(1)} M^{\tau(2)}_{\sigma(2)} \dots M^{\tau(n-1)}_{\sigma(n-1)} \delta^i_{\sigma(n)} \delta^{\tau(n)}_j$$

where $\delta_i^j = 1$ if $i = j$ and 0 otherwise, and the sum is over pairs of permutations σ and τ of the set $\{1, \dots, n\}$. Each component of M^{-1} is therefore given by sums of productss of components of M, divided by the non-zero smooth function $\det M$:

$$\det M = \frac{1}{n!} \sum_{\sigma,\tau} \operatorname{sgn}\sigma \ \operatorname{sgn}\tau \ M^{\tau(1)}_{\sigma(1)} M^{\tau(2)}_{\sigma(2)} \dots M^{\tau(n-1)}_{\sigma(n-1)} M^{\tau(n)}_{\sigma(n)}$$

and so $\left(M^{-1}\right)_i^j$ is a smooth function of the components of M, and $GL(n)$ is a Lie group.

From Example 3.2.1 we know that the multiplications on $SL(n)$ and $O(n)$ are also smooth; these are both submanifolds of $GL(n)$, so it follows that the restriction of the inversion to each of these is smooth, so $SL(n)$ and $O(n)$ are Lie groups.

Example 5.1.2 The group $O(2)$. Recall that the orthogonal group $O(n)$ (or $O(n,\mathbb{R})$ is the group of $n \times n$ matrices with real components satisfying $M^T M = \mathrm{I}_n$. In particular $O(2)$ is the group of orthogonal 2×2 matrices.

$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is in $O(2)$ if the rows (or columns) of M form an orthonormal basis:

$$a^2 + b^2 = 1;$$
$$c^2 + d^2 = 1;$$
$$ac + bd = 0.$$

In particular M is determined by its first row $[a\ b]$ and its determinant $\delta = \det M = \pm 1$:

$$M = \begin{bmatrix} a & b \\ -\delta b & \delta a \end{bmatrix}.$$

This gives a natural map $\varphi : O(2) \to S^1 \times \{-1, 1\}$ given by

$$\varphi \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ((a,b),\ ad - bc) \in S^1 \times \{-1, 1\}.$$

Then φ is smooth and has the smooth inverse

$$\varphi^{-1}\left((a,b),\ \delta\right) = \begin{bmatrix} a & b \\ -\delta b & \delta a \end{bmatrix}.$$

This shows that $O(2)$ is diffeomorphic to $S^1 \times \{-1, 1\}$. In geometric terms, this map takes an orientation-preserving orthogonal transformation (i.e. a rotation) to its angle of rotation, or an orientation-reversing orthogonal transformation (i.e. a reflection) to twice the angle between the line of reflection and the positive x-axis.

Example 5.1.3 The torus $\mathbb{T}^n$ The Torus $\mathbb{T}^n$ is the product of n copies of $S^1 = \mathbb{R}/\sim$ where $a \sim b \Leftrightarrow a - b \in \mathbb{Z}$, with the group structure given by

$$([a_1], \ldots, [a_n]) + ([b_1], \ldots, [b_n]) = ([a_1 + b_1], \ldots, [a_n + b_n]).$$

This can also be naturally embedded as a subgroup of the group $(\mathbb{C}\backslash\{0\})^n$ (with the usual complex multiplication on each factor) by the map

$$F([x_1], \ldots, [x_n]) = \left(e^{2\pi i x_1}, \ldots, e^{2\pi i x_n}\right).$$

Remark. The previous example is one of an important class of examples of groups (and more generally of manifolds) which arise in the following way: Suppose M is a smooth manifold, and G is a group which acts on M – that is, there is a map $\rho : G \times M \to M$, such that the map $\rho_g : M \to M$ given by $\rho_g(x) = \rho(g, x)$ is a diffeomorphism, and $\rho_g \circ \rho_h = \rho_{hg}$. Equivalently, ρ is a group homomorphism from G to the group of diffeomorphisms of M.

The action of G on M is called *totally discontinuous* if for any $x \in M$ there is a neighbourhood U of x in M such that $\rho_g(U) \cap U = \emptyset$ for $g \neq e$. For example, the action of the group $\mathbb{Z}^n$ on $\mathbb{R}^n$ given by

$$\rho((z_1, \ldots, z_n), (x_1, \ldots, x_n)) = (x_1 + z_1, \ldots, x_n + z_n)$$

is totally discontinuous, but the action of $\mathbb{Z}$ on S^1 given by

$$\rho(k, z) = e^{ik\alpha} z$$

is not totally discontinuous if $\alpha \notin \mathbb{Q}$.

If G acts totally discontinuously on M, then the quotient space $M/\sim$ with $x \sim y \Leftrightarrow y = \rho_g(x)$ for some $g \in G$ can be made a differentiable manifold with the same dimension as M, in a unique way such that the projection from M onto $M/\sim$ is a local diffeomorphism: Let $\mathcal{A}$ be an atlas for M. Construct an atlas $\mathcal{B}$ for $M/\sim$ as follows: For each $[x] \in M/\sim$ choose a representative $x \in M$, and choose a chart $\varphi : W \to Z$ in $\mathcal{A}$ about x. Choose U such that $\rho_g(U) \cap U = \emptyset$ for all $g \neq e$, and take $\tilde{\varphi}([y]) = \varphi(x)$ for each $y \in (U \cap W)/\sim$. $\tilde{\varphi}$ is well-defined since $y \in U \Rightarrow [y] \cap U = \{y\}$, and for the same reason $\tilde{\varphi}$ is $1:1$ on $W \cap U$. The transition maps between these charts are just restrictions of transition maps of charts in $\mathcal{A}$, and so are smooth.

Note that the Klein bottle and the Möbius strip are given by the quotient of the plane by totally discontinuous actions, and the real projective space $\mathbb{R}P^n$ is the quotient of the sphere S^n by the totally discontinuous action of $\mathbb{Z}_2$ generated by the antipodal map.

Example 5.1.4 The group $SO(3)$. $SO(3)$ is the group of orientation-preserving orthogonal transformations of three-dimensional space. Note that an orthogonal transformation is a real unitary transformation, and so is conjugate to a diagonal unitary transformation with spectrum closed under conjugation. In the three-dimensional case this means that the spectrum consists of a pair of conjugate points on the unit circle, together with 1 or -1. In the case of a special orthogonal transformation, the spectrum must be $\{1,\ e^{i\alpha},\ e^{-i\alpha}\}$ for some $\alpha \in [0, \pi]$. In particular such a transformation fixes some vector in $\mathbb{R}^3$, and is a rotation by angle α about this axis. Thus an element in $SO(3)$ is determined by a choice of unit vector v and an angle α.

Example 5.1.5 S^3 as a Lie group. The three-dimensional sphere S^3 can be made into a Lie group as follows: We can think of S^3 as contained in $\mathbb{R}^4$, which we identify with the quaterions $\mathbb{H} = \{x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}\}$ where $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$ and $\mathbf{ij} = \mathbf{k}$, $\mathbf{jk} = \mathbf{i}$, and $\mathbf{ki} = \mathbf{j}$. We think of x in $\mathbb{H}$ as having a real part x_0 and a vector part $x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ corresponding to a vector in $\mathbb{R}^3$. The multiplication on $\mathbb{H}$ can be rewritten in these terms as

$$(x, \mathbf{x}) \cdot (y, \mathbf{y}) = (xy - \mathbf{x} \cdot \mathbf{y},\ x\mathbf{y} + y\mathbf{x} + \mathbf{x} \times \mathbf{y})\,. \tag{5.1}$$

Given $x \in \mathbb{H}$, the *conjugate* $\bar{x}$ of x is given by

$$\overline{(x, \mathbf{x})} = (x, -\mathbf{x}).$$

Then the identity (5.1) gives

$$\overline{(x, \mathbf{x})} \cdot (x, \mathbf{x}) = x^2 + \mathbf{x}^2$$

which is the same as the norm of x as an element of $\mathbb{R}^4$. We also have

$$\begin{aligned} \overline{(x, \mathbf{x}) \cdot (y, \mathbf{y})} &= (xy - \mathbf{x} \cdot \mathbf{y},\ -x\mathbf{y} - y\mathbf{x} - \mathbf{x} \times \mathbf{y}) \\ &= (y, -\mathbf{y}) \cdot (x, -\mathbf{x}) \\ &= \overline{(y, \mathbf{y})} \cdot \overline{(x, \mathbf{x})}. \end{aligned}$$

Therefore

$$|x \cdot y|^2 = \overline{x \cdot y} \cdot x \cdot y = \bar{y} \cdot \bar{x} \cdot x \cdot y = |x|^2 |y|^2.$$

In particular, for x and y in S^3 we have $xy \in S^3$. It follows that the restriction of the multiplication in $\mathbb{H}$ to S^3 makes S^3 into a Lie group. This example will reappear later in a more familiar guise.

Remark. We have seen two examples of spheres which are also Lie groups: S^1 is a Lie group, and so is S^3 (one could also say that $S^0 = \{1, -1\}$ is a 0-dimensional Lie group). This raises the question: Are all spheres Lie groups? If not which ones are? The following proposition, which also introduces some important concepts in the theory of Lie groups, will be used to show that the two-dimensional sphere S^2 can't be made into a Lie group:

Proposition 5.1.1 *Let G be a Lie group. Then there exist n smooth vector fields $E_1, \ldots, E_n$ on G each of which is everywhere non-zero, such that $E_1, \ldots, E_n$ form a basis for T_xM for every $x \in M$.*

Proof. Consider the left and right translation maps from G to itself given by $l_g(g') = gg'$ and $r_g(g') = g'g$ for all $g' \in G$. These are smooth maps, and $(l_g)^{-1} = l_{g^{-1}}$ and $(r_g)^{-1} = r_{g^{-1}}$ are smooth, so l_g and r_g are diffeomorphisms from G to G. The derivative of a diffeomorphism is a linear isomorphism, and in particular this gives us isomorphisms $D_e l_g,\ D_e r_g :\ T_eG \to T_gG$.

Choose a basis $\{e_1, \ldots, e_n\}$ for T_eG. Then define $E_i(g) = D_e l_g(e_i)$ for each i and each $g \in G$. The vector fields $E_1, \ldots, E_n$ have the properties claimed in the proposition.

We say that a vector field V on a Lie group G is **left-invariant** if it satisfies $D_e l_g(V_e) = V_g$ for all $g \in G$. It is straightforward to show that the set of left-invariant vector fields form a vector space with dimension equal to the dimension of G (in fact $\{E_1, \ldots, E_n\}$ are a basis). Similarly one can define right-invariant vector fields.

Now return to the 2-sphere S^2. We have the following result from algebraic topology:

Proposition 5.1.2 ("The sphere has no hair") *There is no continuous non-vanishing vector field on S^2.*

Corollary 5.1.3 *There is no Lie group structure on S^2.*

Proof (Sketch):

Suppose V is a non-vanishing continuous vector field on S^2. Then in the chart given by stereographic projection from the north pole, we have in the neighbourhood of the origin (possibly after rotating the sphere about the polar axis)[3]

$$(D\varphi_+(V))(x) = (1,0) + o(|x|) \quad \text{as } x \to 0.$$

Now look at this vector field through the chart φ_-: The transition map is $(\varphi_- \circ \varphi_+^{-1})(x,y) = (x/(x^2+y^2), y/(x^2+y^2))$. Therefore $\tilde{V} = D\varphi_-(V)$ is given by

$$\begin{aligned}
\tilde{V}(w,z) &= \left(D_{(w,z)}(\varphi_+ \circ \varphi_-^{-1})\right)^{-1}((1,0) + o(|x|)) \\
&= \begin{bmatrix} z^2 - w^2 & -2wz \\ -2wz & w^2 - z^2 \end{bmatrix} \begin{bmatrix} 1 + o(|(w,z)|^{-1}) \\ o(|(w,z)|^{-1}) \end{bmatrix} \\
&= \begin{bmatrix} z^2 - w^2 \\ -2wz \end{bmatrix} + o(|(w,z)|) \quad \text{as } |(z,w)| \to \infty.
\end{aligned}$$

There can be no zeroes of the vector field $\tilde{V}$ on $\mathbb{R}^2$, so we can divide $\tilde{V}$ by its length at each point to obtain a continuous vector field V' which has length 1 everywhere, and such that

$$V'(w,z) = \begin{bmatrix} \frac{z^2-w^2}{w^2+z^2} \\ -\frac{2wz}{w^2+z^2} \end{bmatrix} + o(1) \quad \text{as } |(w,z)| \to \infty$$

For each $r \in \mathbb{R}_+$, consider the restriction of V' to the circle of radius r about the origin. Since V' has length 1 everywhere, this defines for each r a continuous map f_r from the circle S^1 to itself, given by

$$f_r(z) = V'(rz)$$

for each $z \in S^1 \subset \mathbb{R}^2$ and each $r > 0$ in $\mathbb{R}$. If we parametrise S^1 by the standard angle coordinate θ, then we have

$$f_r(\cos\theta, \sin\theta) = (\sin^2\theta - \cos^2\theta, -2\cos\theta\sin\theta) + o(1) = -(\cos 2\theta, \sin 2\theta) + o(1)$$

[3] We use the notation $f(x) = o(g(x))$ as $x \to 0$ to mean $\lim_{x\to 0}\left|\frac{f(x)}{g(x)}\right| = 0$. Another useful notation is $f(x) = O(g(x))$ as $x \to 0$ to mean $\limsup_{x\to 0}\left|\frac{f(x)}{g(x)}\right| < \infty$.

as $r \to \infty$. Thus the limit as $r \to \infty$ of the map f_r is just the map that sends θ to $\pi - 2\theta$. In other words, $f_r(\theta)$ winds twice around the circle backwards as θ traverses the circle forwards once: f_r has winding number -2 for r sufficiently large.

But now consider what happens as $r \to 0$: $f_0(\theta)$ is just the constant vector $V'(0,0) \in S^1$, so $\lim_{r\to 0} f_r(\theta) = \theta_0$, a constant. But this means that for r sufficiently small the function $f_r(\theta)$ does not traverse the circle at all as θ moves around S^1, and f_r has winding number zero for r small. Now we use the result from algebraic topology that the winding number of a map from S^1 to itself is a homotopy invariant. The family of maps $\{f_r\}$ is a continuous deformation from a map of winding number -2 to a map of winding number 0, which is impossible. Therefore there is no such vector field. □

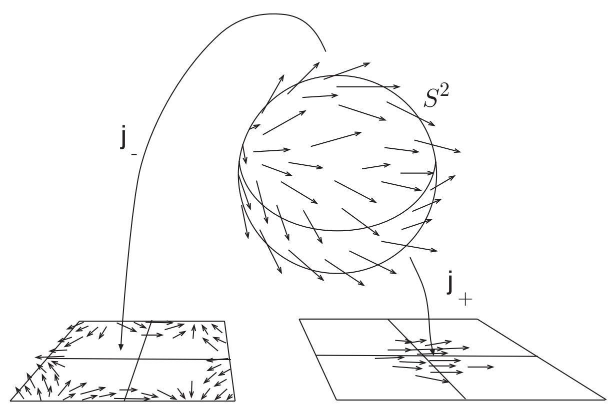


Fig. 5.1: A vector field on S^2, viewed in a region near the centre of the chart φ_+, and in a region near infinity in the chart φ_-.

5.2 Examples of left-invariant vector fields

It is useful to see how to find the left-invariant vector fields for some examples:

Example 5.2.1 More on $\mathbb{R}^n$*.* Let us consider the Lie group $\mathbb{R}^n$ under addition. Then for any $x \in \mathbb{R}^n$ we have $l_x(y) = x + y$, and so $D_0 l_x(v) = v$. Taking the standard basis $\{e_1, \ldots, e_n\}$ for $T_0\mathbb{R}^n \simeq \mathbb{R}^n$, we find that the left-invariant vector fields are given by

$$(E_i)_x = e_i.$$

That is, the left-invariant vector fields are just the constant vector fields.

Example 5.2.2 Matrix groups Let G be a Lie subgroup of $GL(n)$. Then we have $l_M(N) = MN$ for any M and N in G. Differentiating, we get

$$D_{\mathrm{I}} l_M(A) = MA$$

for any $A \in T_{\mathrm{I}}G$. Therefore the left-invariant vector fields have the form

$$A_M = MA$$

where A is a constant matrix in $T_{\mathrm{I}}G$.

Consider the case of the entire group $GL(n)$: In this case, since $GL(n)$ is an open subset of the Euclidean space $M_n \simeq \mathbb{R}^{n^2}$, the tangent space $T_{\mathrm{I}}GL(n)$ is just M_n, and the left-invariant vector fields are given by linear combinations of

$$(E_{ij})_M = Me_{ij}$$

where e_{ij} is the matrix the a 1 in the ith row and the jth column, and zero everywhere else.

Next consider the case $G = SL(n)$: What is the tangent space at the origin? Recall that $SL(n)$ is a submanifold of $GL(n)$, given by the level set $\det^{-1}(1)$ of the determinant function. Therefore we can identify the tangent space of $SL(n)$ with the subspace of M_n given by the kernel of the derivative of the determinant:

$$T_M SL(n) = \left\{ A \in M_n : \ \frac{d}{dt} \det(M + tA)\Big|_{t=0} = 0 \right\}.$$

In particular, we have

$$T_{\mathrm{I}} SL(n) = \{ A \in M_n : \ \mathrm{t}r(A) = 0 \}.$$

Finally, consider $G = O(n)$: This is also a level set of a submersion from M_n, namely the map which sends M to the upper triangular part of $M^T M - I$ (since $M^T M - I$ is symmetric, the entries below the diagonal are superfluous). As before, the tangent space can be identified with the kernel of the derivative of the submersion, and in particular

$$T_{\mathrm{I}} O(n) = \{ A \in M_n : \ B^T = -B \},$$

the vector space of antisymmetric matrices.

Example 5.2.3 The group $\mathbb{C}\backslash\{0\}$. The multiplicative group of non-zero complex numbers is an open set in $\mathbb{C} \simeq \mathbb{R}^2$, so we need not worry about charts. We have

$$l_z(w) = zw,$$

and hence $D_1 l_z(v) = zv$ for all $v \in \mathbb{R}^2$ and $z \in \mathbb{C}\backslash\{0\}$. The left invariant vector fields are linear combinations of the two vector fields $(E_1)_z = z$ and $(E_2)_z = iz$.

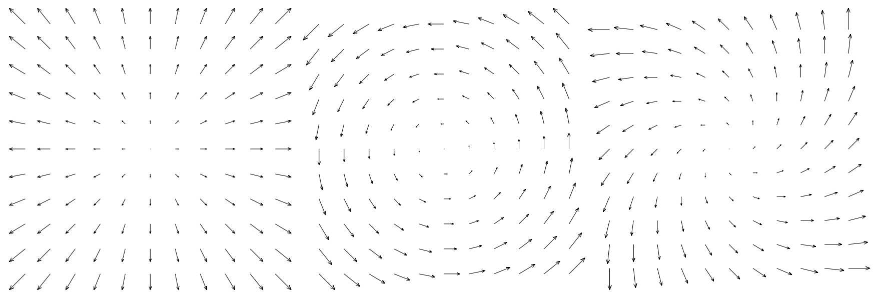


Fig. 5.2: Left-invariant vector fields on $\mathbb{C}\backslash\{0\}$ corresponding to the vectors 1, i, and $1+i$ at the identity.

5.3 1-parameter subgroups

A 1-parameter subgroup Γ of G is a smooth group homomorphism from $\mathbb{R}$ (with addition as the group operation) into G.

Given a one-parameter subgroup $\Gamma : \mathbb{R} \to G$, let $E = \Gamma'(0) \in T_eG$. Then (thinking of a vector as an equivalence class of curves) we have $E = [t \mapsto \Gamma(t)]$. The homomorphism property is that $\Gamma(s)\Gamma(t) = \Gamma(s+t)$. In particular we have $l_{\Gamma(s)}(\Gamma(t)) = \Gamma(s+t)$, and $D_0 I_{\Gamma(s)}(E) = \Gamma'(s)$. In other words, the tangent to the one-parameter subgroup is given by the left-invariant vector field $E_g = D_0 l_g(E)$.

The converse is also true: Suppose E is a left-invariant vector field on G, and $\Gamma : \mathbb{R} \to G$ is a smooth map which satisfies $\Gamma'(t) = E_{\gamma(t)}$. Then Γ defines a one-parameter subgroup.

Proposition 5.3.1 *Let G be a Lie group. To every unit vector $v \in T_eG$ there exists a unique one-parameter subgroup $\Gamma : \mathbb{R} \to G$ with $\Gamma'(0) = v$.*

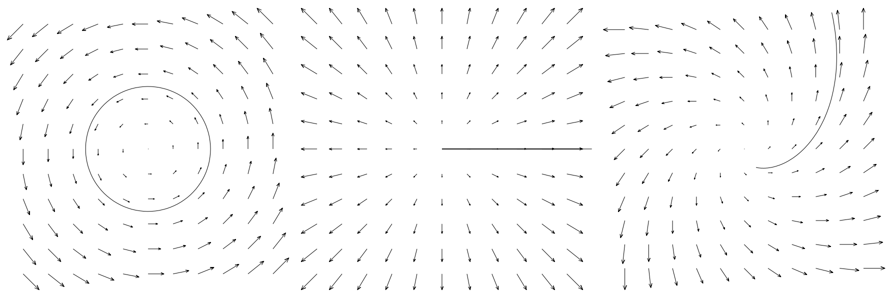
This amounts to solving the ordinary differential equation

$$\frac{d}{dt}\Gamma(t) = D_o l_{\Gamma(t)}(v)$$

with the initial condition $\Gamma(0) = 0$. I will therefore leave the proof until we have seen results on existence and uniqueness of solutions of ordinary differential equations on manifolds in a few lectures from now.

Example 5.3.1: Non-zero complex numbers, revisited. In the case $G = \mathbb{C}\backslash\{0\}$, we have to solve the differential equation

$$\begin{aligned} \dot{z} &= zw \\ z(0) &= 1 \end{aligned}$$



for $w \in \mathbb{C}$ fixed, and this has the solution $z(t) = e^{tw}$. This gives the one-parameter subgroups of $\mathbb{C}\backslash\{0\}$: If $w = 1$ we get the positive real axis, with $\Gamma(t) = e^t$; If $w = i$ we get the unit circle, with $\Gamma(t) = e^{it}$, and for other complex numbers we get logarithmic spirals.

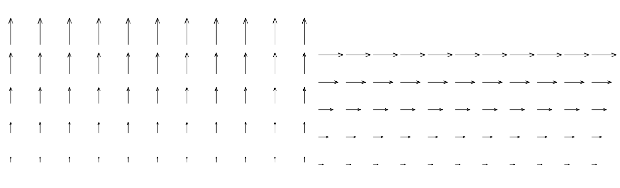
Example 5.3.2: Hyperbolic space. The two-dimensional hyperbolic space $\mathbb{H}^2$ is the upper half-plane in $\mathbb{R}^2$, which we think of as a Lie group by identifying it with the group of matrices

$$\left\{ \begin{bmatrix} \alpha & \beta \\ 0 & 1 \end{bmatrix} : \ \alpha > 0, \ \beta \in \mathbb{R} \right\}$$

by the diffeomorphism $(x, y) \mapsto \begin{bmatrix} y & x \\ 0 & 1 \end{bmatrix}$. The left-invariant vector fields are as follows: To the vector $(1, 0)$ in the tangent space of the identity $(0, 1)$, we associate the vector $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ in the tangent space to the matrix group at the identity. The left-invariant vector field corresponding to this is given by

$$(E_1)_{\begin{bmatrix} \alpha & \beta \\ 0 & 1 \end{bmatrix}} = \frac{d}{dt} \begin{bmatrix} \alpha & \beta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \Bigg|_{t=0} = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}.$$

This corresponds to the vector field $(E_1)_{(\beta,\alpha)} = (\alpha, 0)$. Similarly we get a left invariant vector field $(E_2)_{(\beta,\alpha)} = (0, \alpha)$ corresponding to the vector $(0, 1)$.



The one-parameter subgroups are lines through $(0,1)$, with $\Gamma(t) = (ab^{-1}(e^{bt}-1), e^{bt})$.

Example 5.3.3: One-parameter subgroups of matrix groups We saw in Example 5.2.2 that the left-invariant vector fields in subgroups of $GL(n)$ are given by $A_M = MA$ for some fixed matrix $A \in T_I G$. Therefore the one-parameter subgroups are given by the solutions of the differential equation

$$\dot{M} = MA \tag{5.2}$$

with the initial condition $M(0) = \mathrm{I}$. This is exactly the problem which arises in the process of solving a general linear system of ordinary equations: Recall that to solve the system $\dot{v} = Av$ with any initial condition $v(0) = v_0$, one first solves the system (4.1) to obtain a matrix $M(t)$. Then the required solution is $v(t) = M(t)v_0$. The solution of (4.1) is denoted by e^{At}, and is given by the convergent power series

$$e^{At} = \sum_{n=0}^{\infty} \frac{1}{n!} t^n A^n.$$

It is a consequence of our construction that the matrix exponential of a matrix $A \in T_i G$ is in G – in particular, if A is traceless then $e^{At} \in SL(n)$, and if A is antisymmetric then $e^{At} \in SO(n)$.

Definition 5.3.1 The exponential map $exp : T_e G \to G$ is the map which sends a vector $v \in T_e G$ to $\Gamma(1)$, where $\Gamma : \mathbb{R} \to G$ is the unique one-parameter subgroup satisfying $\Gamma'(0) = v$.

In particular, in the cases $G = \mathbb{R}_+$, $G = \mathbb{C}\backslash\{0\}$, and $G = GL(n)$, $SL(n)$, or $SO(n)$, the exponential map is the same as the usual exponential map. This is also true for the Quaternions, and for the subgroup S^3 of unit length quaternions.

The theory of regularity for ordinary differential equations (which we will discuss later) implies that the exponential map is smooth. Notice also that if Γ is a one-parameter subgroup with $\Gamma'(0) = v$, then $t \mapsto \Gamma(st)$ is also a one-parameter subgroup, with tangent at the origin given by sv. Therefore $\exp(sv) = \Gamma(s)$. This implies

$$D_0 \exp(v) = \frac{d}{ds} \exp(sv) = \frac{d}{ds}\Gamma(s) = \Gamma'(0) = v.$$

In other words, $D_0 \exp = \mathrm{I}$. In particular this implies that the exponential map is a diffeomorphism in the neighbourhood of the identity.

Example 5.3.4: The exponential map on S^3. The exponential map on S^3 can be computed explicitly. First we derive an explicit formula for the exponential of a quaternion $x = x^0 + x^1\mathbf{i} + x^2\mathbf{j} + x^2\mathbf{k}$. We need to solve the differential equation

$$\dot{w} = wx.$$

First note that $\frac{d}{dt}\bar{w} = \bar{x}\bar{w}$, and so

$$\frac{d}{dt}|w|^2 = \bar{x}\bar{w}w + \bar{w}wx = (x + \bar{x})|w|^2 = 2x_0|w|^2.$$

Therefore $|w(t)| = e^{x_0 t}$. Now consider $u(t) = \frac{w(t)}{|w(t)|}$:

$$\frac{d}{dt}u = \frac{wx}{|w|} - \frac{w}{|w|^2}x_0|w| = \frac{w}{|w|}(x - x_0) = u(x - x_0).$$

Therefore $u(t) = \exp(x - x_0)$, the exponential of the 'vector' part of x. So now assume $x_0 = 0$. Then we can write

$$w^1(t) = \frac{1}{2}\left(\bar{w}\mathbf{i} - \mathbf{i}w\right)$$

and

$$\frac{d}{dt}w^1(t) = \frac{1}{2}\left(-\mathbf{i}wx + \bar{x}\bar{w}\mathbf{i}\right).$$

Differentiating again with respect to t, we get

$$\frac{d^2}{dt^2}w^1(t) = \frac{1}{2}\left(-\mathbf{i}wx^2 + \bar{x}^2\bar{w}\mathbf{i}\right).$$

But now x has no real part, so $\bar{x} = -x$, and $x^2 = \bar{x}^2 = -|x|^2$, giving

$$\frac{d^2}{dt^2}w^1(t) = -|x|^2 w^1(t),$$

and so (since $w^1(0) = 0$ and $(w^1)'(0) = x^1$), $w^1(t) = \frac{x^1}{|x|}\sin(|x|t)$. The other components are similar, so we have (in the case $|x| = 1$)

$$\exp(tx) = \cos t + x \sin t. \tag{5.3}$$

Thus in general we have

$$\exp(x^0 + tv) = e^{x^0}(\cos t + v \sin t)$$

whenever x^0 is real and v is a unit length quaternion with no real part. In particular, the tangent space of S^3 is just the quaternions with no real part, so the exponential is given by Eq. (5.3). The one-parameter subgroups of S^3 are exactly the great circles which pass through the identity.

Exercise 5.3.1 Show that the following map is a group homomorphism and a diffeomorphism from the group S^3 to the group $SU(2)$ of unitary 2×2 matrices with determinant 1:

$$\varphi((x^0, x^1, x^2, x^3)) = \begin{bmatrix} x^0 + ix^3 & x^1 + ix^2 \\ -x^1 + ix^2 & x^0 - ix^3 \end{bmatrix}.$$

Remark. The idea behind this construction is the following: A matrix $U \in SU(2)$ can be chosen by first choosing an eigenvector v (this is given by a choice of any non-zero element of $\mathbb{C}^2$, but is defined only up to multiplication by an arbitrary non-zero complex number, and hence is really a choice of an element of $\mathbb{C}P^1 = \left(\mathbb{C}^2\backslash\{0\}\right)/(\mathbb{C}\backslash\{0\})$. Given this, the other eigenvector is uniquely determined by Gram-schmidt. To complete the specification of U we need only specify the eigenvector of the first eigenvalue, which is an arbitrary element $e^{i\theta}$ of the unit circle S^1. Then the eigenvalue of the other eigenvector is $e^{-i\theta}$, since $\det U = 1$.

Proposition 5.3.2 $\mathbb{C}P^1 \simeq S^2$.

Proof We have natural charts $\varphi_\pm$ for S^2, with transition map defined on $\mathbb{R}^2\backslash\{0\}$ by $(x, y) \mapsto (x, y)/(x^2 + y^2)$. We also have natural charts $\eta_1 : \mathbb{C}P^1\backslash\{[1, 0]\} \to \mathbb{R}^2$ given by $\eta_1([z_1, z_2]) = z_1/z_2 \in \mathbb{C} \simeq \mathbb{R}^2$ and $\eta_2 : \mathbb{C}P^1\backslash\{[0, 1]\} \to \mathbb{R}^2$ given by $\eta_2([z_1, z_2]) = z_2/z_2$. The transition map between these is then $z \mapsto 1/z : \mathbb{C}\backslash\{0\} \to \mathbb{C}\backslash\{0\}$. We define a map χ from S^2 to $\mathbb{C}P^1$ by

$$\chi(x) = \begin{cases} \eta_1^{-1}(\varphi_+(x)), & \text{for } x \neq N; \\ \eta_2^{-1}(\overline{\varphi_-(x)}), & \text{for } x \neq S. \end{cases}$$

This makes sense because in the overlap we have

$$\eta_2^{-1} \circ \bar{\varphi}_- = \eta_1^{-1} \circ (\bar{\varphi}_-^{-1} \circ \varphi_+^{-1}) \circ \varphi_+ = \eta_1^{-1} \circ \eta$$

since $\eta_1 \circ \eta_2^{-1} = \bar{\varphi}_- \circ \varphi_+^{-1}$ and these are involutions. Explicitly, we have

$$\chi(x_1, x_2, x_3) = [x_1 + ix_2, 1 - x_3] = [x_1 - ix_2, 1 + x_3]$$

where the second equality holds since

$$\frac{x_1 + ix_2}{1 - x_3} - \frac{1 + x_3}{x_1 - ix_2} = \frac{x_1^2 + x_2^2 - 1 + x_3^2}{(1 - x_3)(x_1 - ix_2)} = 0$$

for $(x_1, x_2, x_3) \in S^2$. □

We then associate the matrix U with a point in S^3 by starting at the identity $(1, 0, 0, 0)$, and following the great circle in the direction $\chi([v]) \in S^2 \simeq T_{(1,0,0,0)}S^3$ for a distance α, and the resulting map is the map φ given in Exercise 5.3.1.

Exercise 5.3.2 Show that the following map from $SU(2)$ to $SO(3)$ is a group homomorphism, and a *local* diffeomorphism, but not a diffeomorphism:

$$\Psi \begin{bmatrix} a_1 + ia_2 & b_1 + ib_2 \\ c_1 + ic_2 & d_1 + id_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(a_1 + d_1) & -b_2 - c_2 & b_1 - c_1 \\ b_2 + c_2 & \frac{1}{2}(a_1 + d_1) & d_2 - a_2 \\ c_1 - b_1 & a_2 - d_2 & \frac{1}{2}(a_1 + d_1) \end{bmatrix}.$$

Consider the composition of this with the map φ in Exercise 5.3.1: Deduce that the group of rotations in space, $SO(3)$, is diffeomorphic to $\mathbb{R}P^3$.

Remark. There is, similarly, a geometric way to understand this map: An element of $SO(3)$ is just a rotation of three-dimensional space, which can be determined by specifying its axis of rotation v (an element of S^2) together with the angle of rotation α about this axis. As before, this corresponds to a point in S^3 by following a great circle from $(1,0,0,0)$ in the direction $v \in S^2 \subset \mathbb{R}^3 \simeq T_{(1,0,0,0)}S^3$ for a distance $\alpha/2$.

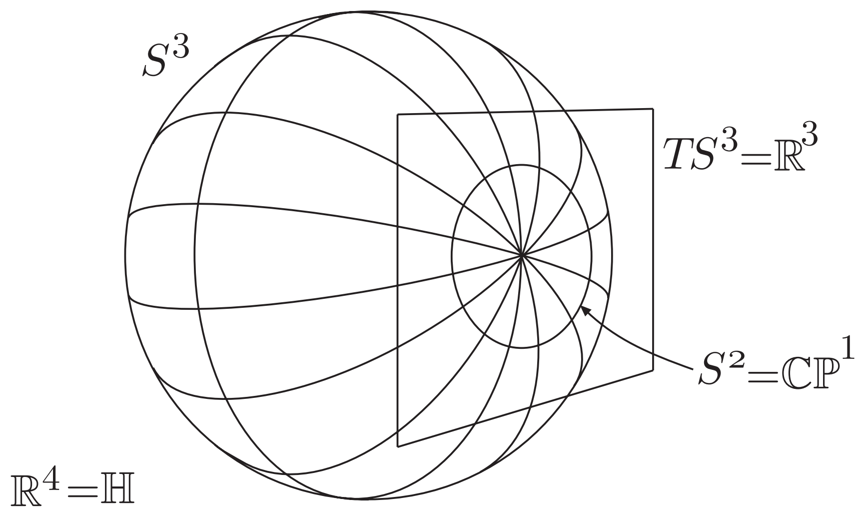


Fig. 5.5: A rotation of $\mathbb{R}^3$ is associated with a point in S^3 by following the great circle through $(1,0,0,0)$ in the direction of the axis of rotation $v \in S^2 \subset TS^3$ for a distance α equal to half the angle of rotation. This is also naturally associated to the matrix $U \in SU(2)$ which sends z to $e^{i\alpha}z$ for any $z \in \mathbb{C}^2$ with $[z] = v \in \mathbb{C}P^1 \simeq S^2 \subset TS^3$.

Notes on Chapter 5:

Lie groups are named after the Norwegian mathematician Sophus Lie (1842–1899). He defined 'infinitesimal transformations' (what are now called Lie algebras) first, as part of his work in describing symmetries of partial differential equations, and then went on to develop his theory of 'continuous transformation groups'.

The classification of compact simple Lie groups was accomplished by Elie Cartan in 1900 (building on earlier work by Killing).


Sophus Lie


Elie Cartan


Wilhelm Killing

Lecture 6. Differential Equations

Our aim is to prove the basic existence, uniqueness and regularity results for ordinary differential equations on a manifold.

6.1 Ordinary differential equations on a manifold.

Any vector field on a differentiable manifold M is naturally associated with a differential equation: If $V \in \mathcal{X}(\mathcal{M})$, and $x \in M$, the basic problem of ODE theory is to find a smooth map $\gamma : I \to M$ for some interval I containing 0, such that

$$\gamma'(t) = V_{\gamma(t)}$$

for all $t \in I$, and

$$\gamma(0) = x.$$

This is the general *initial value problem.* We would like to known several things about this problem: First, that solutions exist; second, that they are unique; and third, that the solutions depend in a smooth way on the initial point $x \in M$. The following proposition incorporates all three aspects:

Proposition 6.1.1 *Let $V \in \mathcal{X}(\mathcal{M})$ and $x \in M$. Then there exists $\delta > 0$, a neighbourhood U of x in M, and a unique smooth map $\Psi : U \times (-\delta, \delta) \to M$ which satisfies* [4]

$$\frac{\partial}{\partial t}\Psi(y,t) = V_{\Psi(y,t)}$$
$$\Psi(y,0) = y$$

for all $y \in U$ and $t \in (-\delta, \delta)$. For each $t \in (\delta, \delta)$, the map $\Psi_t : U \to M$ defined by $\Psi_t(y) = \Psi(y,t)$ is a local diffeomorphism, and

$$\Psi_t \circ \Psi_s = \Psi_{s+t}$$

whenever both sides are defined.

[4] Note that There is a natural vector field ∂_t defined on on $U \times (-\delta, \delta)$ by $(\partial_t f)(x) = \frac{\partial}{\partial s} f(x, t+s)\Big|_{s=0}$. The ODE means that $D_{(y,t)}\Psi(\partial_t) = V_{\Psi(y,t)}$ for all y and t.

The smoothness of Ψ as a function of y amounts to smooth dependence of solutions on their initial conditions, and we get the added bonus that the maps Ψ_t (called the (local) flow of V for time t) are local diffeomorphisms, and they form a "local group": $\Psi_t \circ \Psi_s = \Psi_{s+t}$ as one would expect for a group of diffeomorphisms, but this might hold only on a rather restricted domain.

Example 6.1.1: Problems with V. This example demonstrates why we may only be able to define the flow of a vector field locally, as in the proposition: Take $M = \mathbb{R}$, and take V to be the vector field

$$V_x = x^2 \partial_x.$$

Then for each $x \in \mathbb{R}$ we want to solve the equation

$$\begin{aligned} y'(t) &= y(t)^2; \\ y(0) &= x. \end{aligned}$$

This gives $\Psi(x,t) = y(t) = \frac{x}{1-tx}$, on the time interval $(1/x, \infty)$ if $x < 0$, or $(-\infty, 1/x)$ if $x > 0$ (or $(-\infty,\infty)$ in $x = 0$). Thus the flow of the vector field cannot be defined on $\mathbb{R} \times (\delta, \delta)$ for any $\delta > 0$, and cannot be defined on $U \times \mathbb{R}$ for any open subset U of $\mathbb{R}$. Here the problem seems to arise because the vector field V is not bounded.

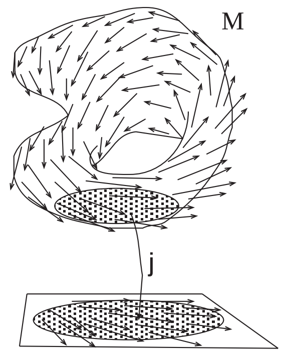
Example 6.1.2: Problems with M. Next consider the example $M = (0, 1)$ with the vector field $V_x = \partial_x$: Now we have the flow

$$\Psi(x,t) = x + t$$

which is defined on the region $\{(x,t) : \ t \in (-x, 1-x)\}$. Again, this flow cannot be defined on $M \times (-\delta, \delta)$ for any $\delta > 0$, or on $U \times \mathbb{R}$ for any $U \subset M$. Here the problem seems to arise because the domain manifold has 'edges'.

Remark. In fact the distinction between these types of difficulties is not so clear. The two situations are in some sense the same, since a diffeomorphism can take a bounded interval to an unbounded region, mapping a bounded vector field to an unbounded one: For example, the map $x \mapsto \cot(\pi x)$ maps $(0, 1)$ to $\mathbb{R}$ and sends the bounded vector field ∂_x to the unbounded vector field $-\pi/\sin^2(\pi x)\partial_x$. When we come to add some further structure to our manifolds in the form of metrics, we will have some notion of when a vector field is unbounded and when a manifold has a boundary, but without this the notions are not meaningful.

To begin the proof of Proposition 5.1.1, we will first reduce the problem to an ordinary differential equation on $\mathbb{R}^n$, by looking at the flow of the vector field through a chart:



Proposition 6.1.2 *Let* $V \in \mathcal{X}(\mathcal{M})$, *and suppose* $\Psi : U \times (-\delta, \delta) \to M$ *satisfies*

$$\begin{aligned} \partial_t \Psi(y,t) &= V_{\Psi(y,t)}; \\ \Psi(y,0) &= y; \end{aligned} \tag{6.1}$$

for each $(y,t) \in U \times (-\delta, \delta)$. *Let* $\varphi : W \to Z \subset \mathbb{R}^n$ *be a chart for* M *such that* $\Psi(U \times (-\delta, \delta)) \subset W$. *Then* $u(z,t) = \varphi \circ \Psi(\varphi^{-1}(z), t)$ *defines a map* $u : \varphi(U) \times (-\delta, \delta) \to Z$ *which satisfies*

$$\begin{aligned} \partial_t u^k(z^1, \ldots, z^n, t) &= \tilde{V}^k(u^1(z^1, \ldots, z^n, t), \ldots, u^n(z^1, \ldots, z^n, t)) \\ u^k(z^1, \ldots, z^n, 0) &= z^k, \end{aligned} \tag{6.2}$$

for $k = 1, \ldots, n$ *and all* $(z^1, \ldots, z^n) \in \varphi(U)$ *and* $t \in (\delta, \delta)$. *Here*

$$V\left(\varphi^{-1}(z^1, \ldots, z^n)\right) = \sum_{k=1}^{n} \tilde{V}^k(z^1, \ldots, z^n)\partial_k.$$

Conversely, if u *satisfies (6.2) then* $\Psi(y,t) = \varphi^{-1}(u(\varphi(y), t))$ *defines a solution of (6.1).*

Proof. We have $D_{(z,t)}u(\partial_t) = D_{\Psi(\varphi^{-1}(z),t)}\varphi \circ D_{(\varphi^{-1}(z),t)}\Psi(\partial_t) = D_\varphi(V)$. By definition $D\varphi(\partial_j) = e_j$ for $j = 1, \ldots, n$, so if $V = \sum \tilde{V}^k \partial_k$, then $D\varphi(V) = \sum \tilde{V}^k e_k$. □

We will prove Proposition 6.1.1 by constructing a unique solution for the ODE (6.2) on a region of $\mathbb{R}^n$.

6.2 Initial value problems.

We begin the proof of Proposition 6.1.1 by constructing solutions of initial value problems:

Proposition 6.2.1 *Let $F : Z \to \mathbb{R}^n$ be smooth, where Z is an open set of $\mathbb{R}^n$. Assume $\sup_Z \|F\| = M_0 < \infty$, and $\sup_Z \|DF\| = M_1 < \infty$. Let $z \in Z$. Then there exists a unique smooth $\gamma : (-\delta, \delta) \to Z$ satisfying*

$$\begin{aligned} \frac{d}{dt}\gamma^i(t) &= F(\gamma^1(t), \dots, \gamma^n(t)); \\ \gamma(0) &= z \end{aligned} \tag{6.3}$$

for $i = 1, \dots, n$. Here $\delta = \frac{d(z, \partial Z)}{M_0}$.

The curves constructed in this proposition are called the *integral curves* of the vector field $F^i \partial_i$.

Proof. We use the method of successive approximations, or 'Picard iteration'. Begin with some approximation to the solution, say $\gamma^{(0)}(t) = z$ for all $t \in (-\delta, \delta)$. Then we try to improve this approximation by iteration: Suppose we have an approximation $\gamma^{(k)}$. Then produce a new approximation by the formula

$$\gamma^{(k+1)}(t) = z + \int_0^t F\left(\gamma^{(k)}(s)\right) ds.$$

This is based on the following observation: The approximation $\gamma^{(k+1)}$ satisfies the differential equation

$$\frac{d}{dt}\gamma^{(k+1)} = F(\gamma^{(k)}),$$

so we are using the kth approximation to tell us the direction of motion for the $(k+1)$st approximation. If this iteration converges to a limit, the required ODE must be satisfied. Note that this iteration makes sense, because we have

$$\|\gamma^{(k)}(t) - z\| \le |t| M_0 < \delta M_0 = d(z, \partial Z)$$

so $\gamma^{(k)}(t)$ is always an element of Z.

To show that the iteration converges, consider the difference between successive approximations: We will prove that

$$\|\gamma^{(k+1)}(t) - \gamma^{(k)}(t)\| \le \frac{M_0 M_1^k |t|^{k+1}}{(n+1)!} \tag{6.4}$$

for all $k \ge 0$ and $t \in (-\delta, \delta)$. This is true for $k = 0$, since $\|\gamma^{(1)} - \gamma^{(0)}\| = \| \int_0^t F(z) ds\| \le |t| M_0$. We proceed by induction: Suppose the inequality holds for $k - 1$. Then

$$
\begin{aligned}
\left\|\gamma^{(k+1)}(t)-\gamma^{(k)}(t)\right\| &= \left\|\int_0^t F(\gamma^{(k)}(s))-F(\gamma^{(k-1)}(s))\,ds\right\| \\
&\le \left|\int_0^t M_1\left\|\gamma^{(k)}(s)-\gamma^{(k-1)}(s)\right\|\,ds\right| \\
&\le M_1\left|\int_0^t \frac{M_0M_1^{k-1}|s|^k}{k!}\,ds\right| \\
&\le \frac{M_0M_1^k|t|^{k+1}}{(k+1)!}.
\end{aligned}
$$

This implies that the sequence $\{\gamma^{(k)}\}$ is a Cauchy sequence in the complete space of continuous maps with respect to uniform convergence, and so converges to a continuous limit γ. The continuity of F and the dominated convergence theorem then imply that

$$
\gamma(t) = z + \int_0^t F(\gamma(s))\,ds \tag{6.5}
$$

for all $t \in (-\delta,\delta)$, so that γ is differentiable and satisfies the equation

$$
\gamma'(t) = F(\gamma(t))
$$

for each t, and the initial condition $\gamma(0) = z$. Smoothness follows, since $\gamma \in C^{(k)}$ implies $\gamma \in C^{k+1}$ by the identity (6.5) and the smoothness of F. This establishes the existence of a solution.

To prove uniqueness, suppose γ and σ are two solutions of the initial value problem. Then

$$
\begin{aligned}
\|\gamma(t)-\sigma(t)\| &= \left\|\int_0^t F(\gamma(s))-F(\sigma(s))\,ds\right\| \\
&\le M_1\left|\int_0^t \|\gamma(s)-\sigma(s)\|\,ds\right|
\end{aligned} \tag{6.6}
$$

Let $C = \sup\|\gamma-\sigma\|$. Then an induction similar to that above shows that

$$
\|\gamma(t)-\sigma(t)\| \le \frac{CM_1^k|t|^k}{k!}
$$

for any k. Taking $k\to\infty$ gives $\gamma\equiv\sigma$. $\square$

6.3 Smooth dependence on initial conditions.

The result of the previous section produced the unique flow Ψ of the vector field V, but did not address the smoothness of Ψ except in the t direction. Since smoothness is measured by reference to charts, it is enough to show smoothness of the map u from Proposition 6.1.2:

Proposition 6.3.1 *The function u which takes a pair (z,t) to the solution $\gamma(t)$ of the initial value problem (6.3) is smooth on the open set*

$$S = \{(z,t) \in Z \times \mathbb{R} : \ |t| < d(z, \partial Z)/M_0\}\,.$$

Proof. I will show that u is the C^k limit of a family of smooth functions for any k: Define $u^{(0)}(z,t) = z$, and successively approximate using

$$u^{(k+1)}(z,t) = z + \int_0^t F\left(u^{(k)}(z,s)\right)\,ds. \tag{6.7}$$

Clearly $u^{(k)}$ maps S to Z for each k, $u^{(k)}$ is smooth for each k, and by (6.4) we have

$$\|u^{(k+1)}(z,t) - u^{(k)}(z,t)\| \le \frac{M_0 M_1^k |t|^{k+1}}{(k+1)!}.$$

Differentiating (6.7) with t fixed gives

$$\left\|D_z u_t^{(k+1)}\right\| = \left\|\mathrm{I} + \int_0^t DF \circ D_z u_s^{(k)}\,ds\right\| \le 1 + M_1\left(\int_0^t \left\|D_z u_s^{(k)}\right\|\,ds\right)$$

which gives by induction

$$\left\|D_z u_t^{(k)}\right\| \le \sum_{j=0}^{k} \frac{M_1^j |t|^j}{j!} \le e^{M_1|t|}$$

independent of k. We also have $\|\partial_t u^{(k)}\| \le M_0$ for all k, so $\{u^{(k)}\}$ is uniformly bounded in C^1. Similar arguments with higher derivatives give uniform bounds on $\{u^{(k)}\}$ in C^j for every j.

Exercise 6.3.2 Show that a sequence of functions $\{u^{(k)}\}$ which converges in C^0 and is bounded in C^j converges in C^m for $m = 1, \ldots, j-1$ [Hint: The key to this is the following *interpolation inequality*: For any C^j function u, and any $l \in \{1, \ldots, j-1\}$, there is a constant C such that

$$\|D^l u\|_{C^0} \le C\|u\|_{C^0}^{1-l/j}\|D^j u\|_{C^0}^{l/j}.$$

Prove this by first proving the case $l = 1$, $j = 2$, and then applying this successively to get the other cases. Then apply the estimate to differences $u^{(k)} - u^{(k')}$].

This completes the proof of Proposition 6.3.1, since Exercise 6.3.2 shows that $u = \lim u^{(k)}$ is in C^j for every j. □

Remark. It is possible to show more explicitly that the approximations $u^{(k)}$ converge in C^j for every j, by differentiating the formula $u^{(k+1)} - u^{(k)}$ j times.

6.4 The local group property.

Next we show that the flow Ψ we have constructed satisfies the local group property

$$\Psi_t \circ \Psi_s = \Psi_{t+s}$$

whenever both sides of this equation make sense.

This is very easy: $\Psi_t \circ \Psi_s$ and Ψ_{t+s} both satisfy the same differential equation

$$\partial_t \Psi = V \circ \Psi,$$

and have the same initial condition at $t = 0$. Hence by the uniqueness part of Proposition 6.2.1, they are the same.

6.5 The diffeomorphism property.

We need to show that $D\Psi_t$ is non-singular. Again, this is very easy: It is true for t small (uniformly in space) since Ψ is smooth and $D\Psi_0 = \mathrm{I}$. But by the local group property, $\Psi_t = \left(\Psi_{t/m}\right)^m$, so $D\Psi_t$ is a composition of m non-singular maps, and hence is non-singular.

This completes the proof of Proposition 6.1.1.

6.6 Global flow.

Proposition 6.1.1 gives the existence of the flow of a vector field *locally*, and we have seen that there are examples which show that one cannot in general expect better than this. However there are some very important situations where we can do better:

Proposition 6.6.1 *Let* $V \in \mathcal{X}(\mathcal{M})$ *be a vector field with compact support – that is, assume that* $\mathrm{supp} V = \overline{\{x \in M : \ V(x) \neq 0\}}$ *is a compact subset of* M*. Then there exists a unique smooth map* $\Psi : M \times \mathbb{R} \to M$ *satisfying*

$$\partial_t \Psi = V \circ \Psi; \qquad \Psi(x, 0) = x.$$

The maps $\{\Psi_t\}$ *form a one-parameter group of diffeomorphisms of* M*.*

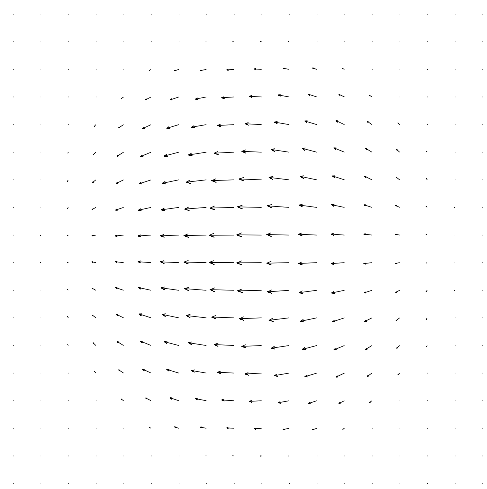


Fig. 6.2: A vector field supported in the unit disk.

Proof. Cover supp$V$ by sets of the form $\varphi_\alpha^{-1}(B_{r_\alpha/2}(0))$ where $\varphi_\alpha : W_\alpha \to Z_\alpha$ is a chart for M and $\overline{B_{r_\alpha}(0)} \subset Z_\alpha$. By compactness there is a finite subcover $\{\varphi_i^{-1}(B_{r_i/2}(0))\}_{i=1}^N$. Let $r = \inf_{i=1,\dots,N} r_i/2 > 0$.

Proposition 6.1.1 gives a local flow Ψ_i on each region $\varphi_i^{-1}(B_{r_i/2}(0)) \times (-\delta, \delta)$, where $\delta = \inf_{i=1,\dots,N} \frac{r}{\sup_{Z_i} \|\tilde{V}_i\|} > 0$. The uniqueness of solutions implies that these local flows agree on the overlaps of these sets, so they combine to give a local flow on the set $M \times (-\delta, \delta)$ (by taking Ψ to be the identity map away from supp$V$). The local group property implies that these maps are diffeomorphisms, since $\Psi_t^{-1} = \Psi_{-t}$ for $|t| < \delta$. Finally, for any $t \in \mathbb{R}$, define $\Psi_t = \left(\Psi_{t/m}\right)^m$, where m is sufficiently large to ensure that $|t|/m < \delta$. Then Ψ satisfies the required differential equation and is defined on $M \times \mathbb{R}$. □

In particular, a smooth vector field on a *compact* manifold M always has a globally defined flow.

Exercise 6.6.1 Show that the exponential map on a Lie group G is a smooth map defined on all of T_eG [Hint: First prove existence in a neighbourhood of the origin. Then use the one-parameter subgroup property to extend to the whole of T_eG].

An important feature of flows of vector fields (i.e. of solving differential equations) is the possibility of substituting new variables in a differential equation. We will formalise this as follows:

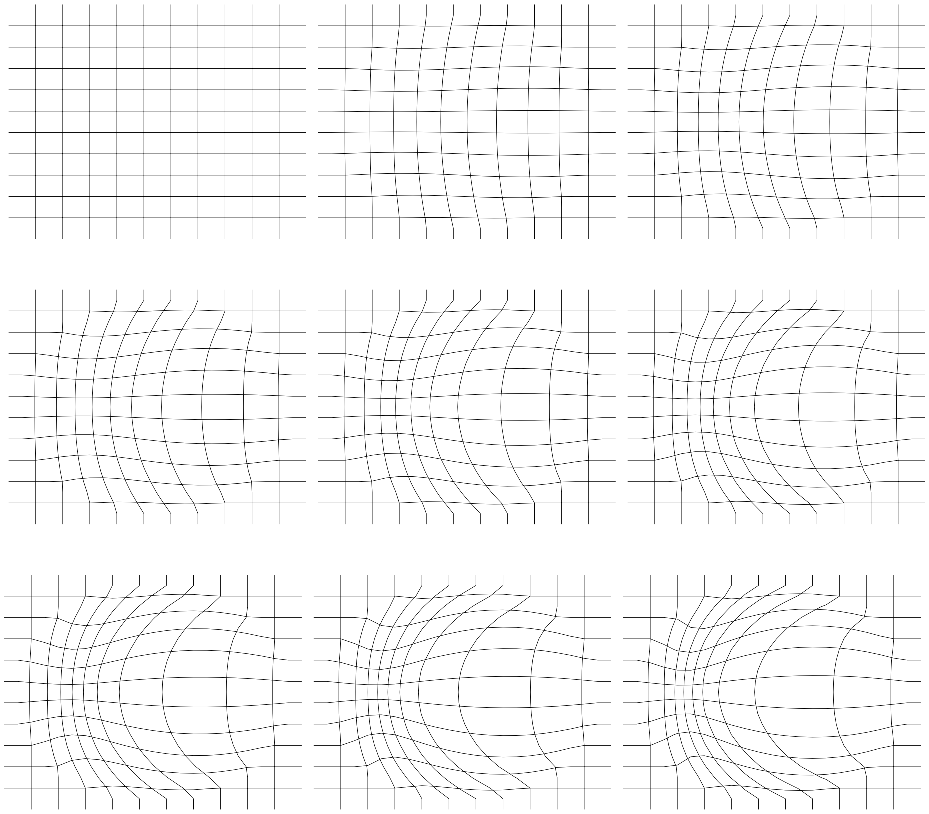


Fig. 6.3: Images of a standard grid at a sequence of times under the flow of the vector field from Fig. 6.2

Definition 6.6.1 Let $F : M \to N$ be smooth. Vector fields $X \in \mathcal{X}(M)$ and $Y \in \mathcal{X}(N)$ are F-related if, for every $x \in X$,

$$D_x F(X_x) = Y_{F(x)}.$$

In particular, if F is a diffeomorphism, every $X \in \mathcal{X}(M)$ has a unique F-related $N \in \mathcal{X}(N)$, which we denote $F_*(X)$, the push forward of X by F.

Proposition 6.6.2 *Let $F : M \to N$ be smooth, and $X \in \mathcal{X}(M)$, $Y \in \mathcal{X}(N)$ F-related. Let Ψ be the local flow of X, and Φ the local flow of Y. Then*

$$\Phi(F(y), t) = F \circ \Psi(y, t)$$

for all y sufficiently close to x and t sufficiently small.

Proof. The uniqueness of solutions of initial value problems applies, since $\partial_t \left(F \circ \Psi\right)(y,t) = D_{\Psi(y,t)} F(X_{\Psi(y,t)}) = Y_{F \circ \Psi(y,t)}$ and $F \circ \Psi(y, 0) = F(y)$. $\square$

Lecture 7. Lie brackets and integrability

In this lecture we will introduce the Lie bracket of two vector fields, and interpret it in several ways.

7.1 The Lie bracket.

Definition 7.1.1 Let X and Y by smooth vector fields on a manifold M. The **Lie bracket** $[X, Y]$ of X and Y is the vector field which acts on a function $f \in C^\infty(M)$ to give $XYf - YXf$.

To make sense of this definition we need to check that this does indeed define a derivation. Linearity is clear, but we need to verify the Leibniz rule:

$$\begin{aligned}
[X,Y](fg) &= X(Y(fg)) - Y(X(fg)) \\
&= X(fY(g) + Y(f)g) - Y(fX(g) + X(f)g) \\
&= f(XYg) + (Xf)(Yg) + (XYf)g + (Yf)(Xg) \\
&\quad - (Yf)(Xg) - f(YXg) - (YXf)g - (Xf)(Yg) \\
&= f(XYg - YXg) + (XYf - YXf)g \\
&= f[X,Y]g + ([X,Y]f)g.
\end{aligned}$$

It is useful to write $[X, Y]$ in terms of its components in some chart: Write $X = X^i \partial_i$ and $Y = Y^j \partial_j$ (here I'm using the summation convention: Sum over repeated indices). Note that

$$[\partial_i, \partial_j]f = \partial_i \partial_j f - \partial_j \partial_i f = \frac{\partial^2}{\partial_i \partial_j}\left(f \circ \varphi^{-1}\right) - \frac{\partial^2}{\partial_j \partial_i}\left(f \circ \varphi^{-1}\right) = 0.$$

Therefore we have

$$[X,Y] = X^i \partial_i Y^j \partial_j - Y^j \partial_j X^i \partial_i = \sum_{i,j=1}^n \left(X^i \partial_i Y^i - Y^i \partial_i X^j\right) \partial_j. \qquad (7.1)$$

The Lie bracket measures the extent to which the derivatives in directions X and Y do not commute. The following proposition makes this more precise:

Proposition 7.1.1 *Let $X, Y \in \mathcal{X}(\mathcal{M})$, and let Ψ and be the local flow of X in some region containing the point $x \in M$. Then*

$$[X,Y]_x = \frac{d}{dt}\left(\left(D_x\Psi_t\right)^{-1} Y_{\Psi_t(x)}\right)\Big|_{t=0}.$$

The idea is this: The flow Ψ_t moves us from x in the direction of the vector field X. We look at the vector field Y in this direction, and use the map $D_x\Psi_t : T_xM \to T_{\Psi_t(x)}M$, which is non-singular, to bring Y back to T_xM. This gives us a family of vectors in the same vector space, so we can compare them by differentiating. In particular, this gives us the nice interpretation: The Lie bracket $[X,Y]$ vanishes if and only if Y is invariant under the flow of X. For this reason, the Lie bracket is also often called the *Lie derivative*, and denoted by $\mathcal{L}_X Y$.

Proof (Method 1): I will first give a proof which is quite illustrative, but which has the drawback that we need to consider two separate cases. Fix $x \in M$.
Case 1: Suppose $X_x \neq 0$. I will construct a special chart about x as follows: First take any chart $\varphi : U \to V$ about x, and by composition with a linear map assume that $X_x = \partial_n(x)$. Let

$$\Sigma = \varphi^{-1}\left(\{(w^1,\dots,w^n) \in V : w^n = \varphi^n(x)\}\right).$$

Then Σ is a smooth $(n-1)$-dimensional submanifold of M which passes through x, and is transverse to the vector field X on some neighbourhood of x (i.e. $T_y\Sigma \oplus \mathbb{R}X_y = T_yM$ for all $y \in \Sigma$ sufficiently close to x). Now consider the map $\tilde{\Psi} : \Sigma \times (-\delta,\delta) \to M$ given by restricting the flow Ψ of X to $\Sigma \times (-\delta,\delta)$: We have

$$D_{(x,0)}\tilde{\Psi} = \begin{bmatrix} \mathrm{I}_{n-1} & 0 \\ 0 & 1 \end{bmatrix}$$

with respect to the natural bases $\{\partial_1,\dots,\partial_{n-1},\partial_t\}$ for $T_{(x,0)}(\Sigma \times \mathbb{R})$ and $\{\partial_1,\dots,\partial_{n-1},\partial_n = X_x\}$ for T_xM. In particular, $D_{(x,0)}\tilde{\Psi}$ is non-singular, and so by the inverse function theorem there is a neighbourhood of $(x,0)$ in $\Sigma\times\mathbb{R}$ on which $\tilde{\Psi}$ is a diffeomorphism. We take $\tilde{\varphi}$ to be the chart obtained by first taking the inverse of $\tilde{\Psi}$ to obtain a point in $\Sigma \times \mathbb{R}$, and then applying φ to the point in Σ to obtain a point in $\mathbb{R}^{n-1} \times \mathbb{R}$. The special feature of this chart is that $X_y = \partial_n$ on the entire chart.

Now we can compute in this chart: The flow of X can be written down immediately: $\tilde{\varphi} \circ \Psi_t \circ \tilde{\varphi}^{-1}(x^1,\dots,x^n) = (x^1,\dots,x^{n-1},x^n+t)$. If we write $Y = Y^k\partial_k$, then we have

$$\left(\left(D_x\Psi_t\right)^{-1} Y_{\Psi_t(x)}\right)^k = Y^k(x^1,\dots,x^n+t),$$

and so

$$\frac{d}{dt}\left((D_x\Psi_t)^{-1}Y_{\Psi_t(x)}\right)^k\Big|_{t=0} = \frac{\partial Y^k}{\partial x^n} = X^j\frac{\partial Y^k}{\partial x^j} = [X,Y]^k$$

by (7.1), since $\partial_i X^j = 0$ for all j.

Case 2: Now suppose $X_x = 0$. Then we have $\Psi_t(x) = x$ for all t, and the local group property implies that the linear maps $\{D_x\Psi_t\}$ form a one-parameter group in the Lie group $GL(T_xM) \simeq GL(n)$. Hence $D_x\Psi_t = e^{tA}$ for some A, and we get

$$(D_x\Psi_t)^{-1}Y_{\Psi_t(x)} = e^{tA}Y_x,$$

and

$$\frac{d}{dt}\left((D_x\Psi_t)^{-1}Y_{\Psi_t(x)}\right)\Big|_{t=0} = -AY_x.$$

It remains to compute A: Working in local coordinates,

$$A(\partial_j) = \partial_t\partial_j\Psi_t = \partial_j\partial_t\Psi_t = \partial_j X.$$

Therefore we get

$$\frac{d}{dt}\left((D_x\Psi_t)^{-1}Y_{\Psi_t(x)}\right)\Big|_{t=0} = -Y^j\partial_j X^k\partial_k = [X,Y]$$

by (7.1), since $X^k = 0$ for all k. □

Proof (Method 2): Here is another proof which is somewhat more direct but perhaps less illuminating: Choose any chart φ for M about x. In this chart we can write uniquely $X = X^j\partial_j$ and $Y = Y^k\partial_k$, and here the coefficients X^j and Y^k are smooth functions on a neighbourhood of x. Then

$$\begin{aligned}
\frac{d}{dt}\left(D_{\Psi_t(x)}\Psi_{-t}\,Y_{\Psi_t(x)}\right)^j\Big|_{t=0} &= \left(\partial_t\partial_k\Psi^j_{-t}\right)\Big|_{t=0}Y^k_x + \left(\partial_k\Psi^j_{-t}\right)\,\partial_t Y^k(\Psi_t x)\Big|_{t=0}\\
&= \left(\partial_k\partial_t\Psi^j_{-t}\right)\Big|_{t=0}Y^k_x + \delta^j_k X^i_x\partial_i Y^k_x\\
&= -Y^k_x\partial_k X^j_x + X^i_x\partial_i Y^j_x\\
&= [X,Y]^j_x.
\end{aligned}$$

□

We could now deduce the following result on the naturality of the Lie bracket directly from Proposition 7.6.2. However I will give a direct proof:

Proposition 7.1.2 *Let $F : M \to N$ be a smooth map, $X, Y \in \mathcal{X}(M)$, and $W, Z \in \mathcal{X}(N)$ with W F-related to X and Z F-related to Y. Then for every $x \in M$,*

$$[W,Z]_{F(x)} = D_xF([X,Y]_x).$$

Think in particular about the case where F is a diffeomorphism: Then the proposition says that the push-forward of the Lie bracket of X and Y is the Lie bracket of the push-forwards of X and Y.

Proof. For any $f : N \to \mathbb{R}$, and $x \in M$, $y = F(x) \in N$,

$$\begin{aligned}
[W,Z]_y f &= W_y Z f - Z_y W f \\
&= D_x F(X_x) Z f - D_x F(Y_x) W f \\
&= X_x((Zf)\circ F) - Y_x((Wf)\circ F) \\
&= X_x(Y(f\circ F)) - Y_x(X(f\circ F)) \\
&= [X,Y]_x(f\circ F) \\
&= D_x F([X,Y]_x) f.
\end{aligned}$$

Here we used the fact that for every $p \in M$,

$$(Zf)\circ F_p = (Zf)_{F(p)} = D_p F(Y_p) f = Y_p(f\circ F),$$

and similarly $(Wf)\circ F = X(f\circ F)$. □

7.2 The Lie algebra of vector fields.

The Lie bracket gives the space of smooth vector fields $\mathcal{X}(M)$ on a manifold M the structure of a *Lie algebra*:

Definition 7.2.1 A **Lie algebra** consists of a (real) vector space E together with a multiplication $[.,.] : E \times E \to E$ which satisfies the three properties

$$\begin{aligned}
[u,v] &= -[v,u]; \\
[au+bv,w] &= a[u,w] + b[v,w]; \\
[[u,v],w] + [[v,w],u] + [[w,u],v] &= 0,
\end{aligned}$$

for all $u, v, w \in E$ and $a, b \in \mathbb{R}$.

The first two properties are immediate. The third (known as the Jacobi identity) can be verified directly:

$$\begin{aligned}
[[X,Y],Z]f &= [X,Y]Zf - Z[X,Y]f \\
&= XYZf - YXZf - ZXYf + ZYXf \\
[[X,Y],Z]f &+ [[Y,Z],X]f + [[Z,X],Y]f \\
&= XYZf - YXZf - ZXYf + ZYXf \\
&\quad + YZXf - ZYXf - XYZf + XZYf \\
&\quad + ZXYf - XZYf - YZXf + YXZf \\
&= 0.
\end{aligned}$$

The following exercise gives another interpretation of the Jacobi identity:

Exercise 7.2.1 Let $X, Y, Z \in \mathcal{X}(M)$. Let Ψ be the local flow of Z near a point $x \in M$. For each t proposition 7.1.2 gives

$$D\Psi_t([X,Y]) = [D\Psi_t(X), D\Psi_t(Y)].$$

Differentiate this with respect to t to get another proof of the Jacobi identity.

Lie algebras play an important role in the theory of Lie groups: Consider the space $\mathfrak{g}$ of left-invariant vector fields on a Lie group G. We have already seen that this is a finite-dimensional vector space isomorphic to the tangent space at the identity T_eG by the natural construction

$$v \in T_eG \mapsto V \in \mathfrak{g} : V_g = D_e l_g(v).$$

We will show that $\mathfrak{g}$ is a Lie algebra. It is sufficient to show that the vector subspace $\mathfrak{g}$ of $\mathcal{X}(M)$ is closed under the Lie bracket operation:

Proposition 7.2.2

$$[\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{g}.$$

Proof. Let $X, Y \in \mathfrak{g}$. Then for any $g \in G$, Proposition 7.1.2 gives

$$[X,Y] = [(l_g)_*X, (l_g)_*Y] = (l_g)_*[X,Y]$$

since X is l_g-related to itself, as is Y. Therefore $[X,Y] \in \mathfrak{g}$. $\square$

The Lie algebra of a group captures the local or infinitesimal structure of a group. It turns out that the group G itself can be (almost) completely recovered just from the Lie algebra $\mathfrak{g}$ – at least, $\mathfrak{g}$ determines the universal cover of G. Different groups with the same universal cover have the same Lie algebra.

Exercise 7.2.2 Show that:

(1). A smooth group homomorphism $\rho : G \to H$ induces a homomorphism from the Lie algebra $\mathfrak{g}$ of G to the Lie algebra $\mathfrak{h}$ of H;

(2). If H is a subgroup of G, then $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$;

(3). A diffeomorphism of manifolds induces a homomorphism between their vector field Lie algebras;

(4). A submersion $F : M \to N$ induces a Lie subalgebra of $\mathcal{X}(M)$ (consisting of those vector fields which are mapped to zero by DF).

7.3 Integrability of families of vector fields

The significance of the Lie bracket, for our purposes, comes from its role in indicating when a family of vector fields can be simultaneously integrated to give a map: Suppose we have k vector fields $E_1, \dots, E_k$ on M, and we want to try to find a map F from $\mathbb{R}^k$ to M for which $\partial_i F = E_i$ for $i = 1, \dots, k$. The naive idea is that we try to construct such a map as follows: Pick some point x in M, and set $F(0) = x$. Then we can arrange $\partial_k F = E_k$, by setting $F(te_k) = \Psi_{E_k}(x, t)$ for each t. The next step should be to follow the integral curves of the vector field E_{k-1} from $F(te_k)$ for time s to get the point $F(te_k + se_{k-1})$:

$$F(te_k + se_{k-1}) = \Psi_{E_{k-1}}\left(\Psi_{E_k}(x,t), s\right),$$

and so on following $E_{k-2}, ..., E_2$, and finally E_1. Unfortunately that doesn't always work:

Example 7.3.1 Take $M = \mathbb{R}^2$, $E_1 = \partial_1$, and $(E_2)_{(x,y)} = (1+x)\partial_2$. Following the recipe outlined above, we set $F(0,0) = (0,0)$, and follow the flow of E_2 to get $F(0,t) = (0,t)$, and then follow E_1 to get $F(s,t) = (s,t)$. However, this gives $DF(\partial_1) = E_1$ but not $DF(\partial_2) = E_2$. Note that if we change our procedure by first integrating along the vector field E_1, then the vector field E_2, then we don't get the same result: Instead we get $F(s,t) = (s, t + st)$.

In fact, there is an easy way to tell this could not have worked: If there was such a map, then we would have

$$[E_1, E_2] = [DF(\partial_1), DF(\partial_2)] = DF([\partial_1, \partial_2]) = 0.$$

But in the example we have $[E_1, E_2] = \partial_2 \neq 0$. It turns out that this is the only obstruction to constructing such a map:

Proposition 7.3.2 *Suppose $E_1, \dots, E_k \in \mathcal{X}(M)$ are vector fields which commute: $[E_i, E_j] = 0$ for $i, j = 1, \dots, k$. Then for each $x \in M$ there exists a neighbourhood U of 0 in $\mathbb{R}^k$ and a unique smooth map $F : U \to M$ satisfying $F(0) = x$ and $D_y F(\partial_i) = (E_i)_{F(y)}$ for every $y \in U$ and $i \in \{1, \dots, k\}$.*

Proof. We construct the map F exactly as outlined above. In other words, we set

$$F(y^1, \dots, y^k) = \Psi_{E_1, y^1} \circ \Psi_{E_2, y^2} \circ \dots \circ \Psi_{E_k, y^k}(x).$$

Where $\Psi_{E_i, t}$ is the flow of the vector field E_i for time t. This gives immediately that $DF(\partial_1) = E_1$ everywhere, but $DF(\partial_2)$ is not so clear.

Lemma 7.3.3 *If $[X, Y] \equiv 0$, then the flows of X and Y commute:*

$$\Psi_{X,t} \circ \Psi_{Y,s} = \Psi_{Y,s} \circ \Psi_{X,t}.$$

Proof. This is clearly true for $t = 0$, since $\Psi_{X,0}(y) = y$ for all y. Therefore it suffices to show that

$$\partial_t \Psi_{X,t} \circ \Psi_{Y,s} = \partial_t \Psi_{Y,s} \circ \Psi_{X,t}.$$

The left-hand side is equal to the vector field X by assumption, while the right is equal to $D\Psi_{Y,s}(X)$. We have

$$\begin{aligned} D\Psi_{Y,s}(X) &= D\Psi_{Y,0}(X) + \int_0^s \frac{d}{dr} D\Psi_{Y,r}(X)\,dr \\ &= X + \int_0^s D\Psi_{Y,r} \frac{d}{dr'} D\Psi_{Y,r'}(X)\Big|_{r'=0} dr, \end{aligned}$$

where I used the local group property of the flow of Y to get $D\Psi_{Y,r+r'} = D\Psi_{Y,r} \circ D\Psi_{Y,r}$. By Proposition 6.1.1, we have

$$D\Psi_{Y,r'}(X)\Big|_{r'=0} = -[Y, X] = 0,$$

and so $D\Psi_{Y,s}(X) = X$. □

Thus we have for every $j \in \{1, \ldots, k\}$ the expression

$$F(y^1, \ldots, y^k) = \Psi_{E^j,y^j} \circ \Psi_{E^1,y^1} \circ \ldots \circ \Psi_{E_{j-1},y^{j-1}} \circ \Psi_{E_{j+1},y^{j+1}} \circ \ldots \circ \Psi_{E_k,y^k}(x)$$

and differentiation in the y^j direction immediately gives $DF(\partial_j) = E_j$. □

A closely related but slightly more general integrability theorem is the theorem of Frobenius, which we will prove next.

Definition 7.3.4 A distribution $\mathcal{D}$ of k-planes on M is a map which associates to each $x \in M$ a k-dimensional subspace $\mathcal{D}_x$ of T_xM, and which is smooth in the sense that for any $V \in \mathcal{X}(M)$ and any chart with coordinate tangent basis $\{\partial_1, \ldots, \partial_n\}$, the vector field $V_{\mathcal{D}}$ given by projecting V onto $\mathcal{D}$ (orthogonally with respect to the given basis) is smooth.

Remark. One can make this definition more natural (i.e. with less explicit reference to charts) as follows: Given a manifold M, there is a natural manifold $G_k(TM)$ constructed from M, called the Grassmannian bundle (of k-planes) over M

$$G_k(TM) = \{(x, E) : \ x \in M, \ E \text{ a } k\text{−dimensional subspace of } T_xM\}.$$

Note that the k-dimensional subspaces of T_xM are in one-to-one correspondence with the space of equivalence classes of rank $(n-k)$ linear maps from T_xM to $\mathbb{R}^{n-k}$, under the equivalence relation

$$M_1 \sim M_2 \Longleftrightarrow M_1 = LM_2, \ L \in GL(n-k):$$

We denote the equivalence class of M by $[M]$. Given a k-dimensional subspace E, we choose any $n-k$ independent linear functions $f_1, \ldots, f_{n-k}$ on T_xM which vanish on E (for example, choose a basis $\{e_1,\dots,e_n\}$ for T_xM such that $\{e_1,\dots,e_k\}$ is a basis for E, and take f_i to be the linear function which takes e_{k+i} to 1 and all the other basis elements to zero. Then $f=(f_1,\dots,f_{n-k})$ defines a rank $n-k$ linear map from T_xM to $\mathbb{R}^{n-k}$ vanishing on E. This correspondence is well-defined as a map into the space of equivalence classes, since if we choose a different basis $\{e'_1,\dots,e'_n\}$ of this form, then it is related to the first by $e'_i=A(e_i)$ for some non-singular matrix, and we have $f'_j=\sum_{m=1}^{n-k}A_{k+j}^{k+m}f_m$, i.e. $f'=Af\sim F$.

Conversely, given a rank $n-k$ linear function M, we associate to it the k dimensional linear subspace $\ker M$. Clearly if $M_1\sim M_2$ then $\ker M_1=\ker M_2$.

Now we can choose charts for $G_k(TM)$: Let $\mathcal{A}$ be an atlas for M, $\varphi: U\to V$ a chart of $\mathcal{A}$. Define an open subset $\tilde{U}$ of $G_k(TM)$ by

$$\tilde{U}=\left\{(x,[F])\in U\times L(T_xM,\mathbb{R}^{n-k}):\ F\Big|_{\mathrm{span}(\partial_1,\dots,\partial_{n-k})}\ \text{nonsingular}\right\}$$

and define $\varphi_\sigma: U_\sigma\to\mathbb{R}^n\times\mathbb{R}^{k(n-k)}$ as follows:

$$\varphi_\sigma(x,[F])=(\varphi,N)$$

where $N\in L_{k,n-k}\simeq\mathbb{R}^{k(n-k)}$ is the $(n-k)\times k$ matrix constructed as follows: Set $F_i^j=F^j(\partial_{\sigma(i)})$. Then by assumption the first $(n-k)$ columns of the matrix F form a non-singular $(n-k)\times(n-k)$ matrix, which we denote by G. Then we have $[F]=[G^{-1}F]$, and $G^{-1}F$ has the form

$$[\mathrm{I}_{n-k}\quad N]$$

for some $(n-k)\times k$ matrix N. N is well-defined as a function of the equivalence class $[F]$, since if $\tilde{F}=AF$, then $\tilde{G}=GA^{-1}$, so $\tilde{G}^{-1}\tilde{F}=G^{-1}F$. Explicitly,

$$N_i^j=\frac{\sum_{\sigma,\tau\in S_{n-k}}\mathrm{sgn}\sigma\mathrm{sgn}\tau F_{\sigma(1)}^{\tau(1)}\dots F_{\sigma(n-k-1)}^{\tau(n-k-1)}F_{n-k+i}^{\tau(n-k)}\delta_{\sigma(n-k)}^j}{\sum_{\sigma,\tau\in S_{n-k}}\mathrm{sgn}\sigma\mathrm{sgn}\tau F_{\sigma(1)}^{\tau(1)}\dots F_{\sigma(n-k-1)}^{\tau(n-k-1)}F_{\sigma(n-k)}^{\tau(n-k)}}$$

for $i=1,\dots,k$ and $j=1,\dots,n-k$, where the sums are over all pairs of permutations of $n-k$ objects.

Note that these charts cover $G_k(TM)$, because even if $[F]\in G_k(T_xM)$ does not have its first $n-k$ columns non-singular with respect to the coordinate tangent basis of φ, we can choose a new chart for which this is true, of the form $A\circ\varphi$ for some $A\in GL(n)$.

Then we can define a k-dimensional distribution to be a smooth map $\mathcal{D}$ from M to $G_k(TM)$ such that $\mathcal{D}_x\in G_k(T_xM)$ for all $x\in M$.

The question which the Frobenius theorem addresses is the following: Given a distribution $\mathcal{D}$, can we find a k-dimensional submanifold Σ through each point, such that $T_x\Sigma=\mathcal{D}_x$ for every $x\in\Sigma$? This is what is meant by *integrating* a distribution.

For convenience we will denote by $\mathcal{X}(\mathcal{D})$ the vector space of smooth vector fields X on M for which $X_x\in\mathcal{D}_x$ for every $x\in M$ (i.e. vector fields tangent to the distribution).

Proposition 7.3.5 (Frobenius' Theorem) *A distribution $\mathcal{D}$ is integrable if and only if $\mathcal{X}(\mathcal{D})$ is closed under the Lie bracket operation.*

Proof. One direction is clear: Suppose there is such a submanifold N, embedded in M by a map F. Then for any smooth vector fields $X, Y \in \mathcal{X}(\mathcal{D})$ there are unique vector fields $\tilde{X}$ and $\tilde{Y}$ in $\mathcal{X}(N)$ such that X is F-related to $\tilde{X}$ and Y is F-related to $\tilde{Y}$. Then Proposition 7.1.2 gives $[X, Y] = DF([\tilde{X}, \tilde{Y}])$ is in $\mathcal{X}(\mathcal{D})$.

The other direction takes a little more work: Given a distribution of k-planes $\mathcal{D}$ which is involutive (closed under Lie brackets), and any point $x \in M$, we want to construct a submanifold through x tangent to the distribution. Choose a chart φ for M near x, and assume that $\mathcal{D}_x$ is the subspace of T_xM generated by the first k coordinate tangent vectors. Then by the smoothness of the distribution we can describe $\mathcal{D}_y$ for y sufficiently close to x as follows:

$$\mathcal{D}_y = \left\{ \sum_{i=1}^{k} c^i \left(\partial_i + \sum_{j=k+1}^{n} a_i^j(y)\partial_j \right) : \ (c^1, \dots, c^k) \in \mathbb{R}^k \right\}.$$

Here the functions a_i^j are smooth functions on a neighbourhood of x, for $i = 1, \dots, k$ and $j = k+1, \dots, n$. In particular we have k non-vanishing vector fields $E_1, \dots, E_k$ tangent to $\mathcal{D}$, given by

$$(E_i)_y = \partial_i + \sum_{j=k+1}^{n} a_i^j(y)\partial_j.$$

Now we compute:

$$\begin{aligned}
[E_i, E_j] &= \left[\partial_i + \sum_{p=k+1}^{n} a_i^p(y)\partial_p, \ \partial_j + \sum_{q=k+1}^{n} a_j^q(y)\partial_q \right] \\
&= \sum_{p=k+1}^{n} \left(\partial_i a_j^p - \partial_j a_i^p + \sum_{q=k+1}^{n} \left(a_i^q \partial_q a_j^p - a_j^q \partial_q a_i^p \right) \right) \partial_p.
\end{aligned}$$

But by assumption $[E_i, E_i] \in \mathcal{X}(\mathcal{D})$, so it is a linear combination of the vector fields $E_1, \dots, E_k$. But $[E_i, E_j]$ has no component in the direction $\partial_1, \dots, \partial_k$. Therefore $[E_i, E_j] = 0$.

Proposition 7.3.2 gives the existence of a map F from a region of $\mathbb{R}^k$ into M with $DF(\partial_i) = E_i$ for $i = 1, \dots, k$. In particular DF is of full rank, and F is an immersion, hence locally an embedding, and the image of F is a submanifold with tangent space equal to $\mathcal{D}$ everywhere. □

Example 7.3.2 The Heisenberg group In this example we investigate a situation where a distribution $\mathcal{D}$ is not involutive. Let G be the group of 3×3 matrices of the form

$$\begin{bmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{bmatrix},$$

which we identify with three-dimensional space $\mathbb{R}^3$. The left-invariant vector fields corresponding to the standard basis $\{e_1, e_2, e_3\}$ at the identity $0 \in \mathbb{R}^3$ are given by

$$E_1 = \begin{bmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \simeq e_1;$$
$$E_2 = \begin{bmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & x_1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \simeq e_2 + x_1 e_3;$$
$$E_3 = \begin{bmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \simeq e_3.$$

We take a distribution $\mathcal{D}$ on $\mathbb{R}^3$ to be the subspace at each point spanned by E_1 and E_2. Then $\mathcal{D}$ is not involutive, since we have

$$[E_1, E_2] = [e_1, e_2 + x_1 e_3] = e_3 = E_3.$$

The behaviour in this case turns out to be as far as one could imagine from that in the involutive case : If there were a submanifold tangent to the distribution, then only points in that submaniofld could be reached along curves tangent to the distribution. In contrst, we have

Proposition 7.3.6 *For any $y \in \mathbb{R}^3$ there exists a smooth $\gamma : [0,1] \to \mathbb{R}^3$ with $\gamma'(t) \in \mathcal{D}_{\gamma(t)}$ for all t, $\gamma(0) = 0$, and $\gamma(1) = y$.*

In other words, *every* point can be reached by following curves tangent to $\mathcal{D}$.

Proof. We consider curves tangent to the distribution, given by prescribing the tangent vector at each point:

$$\gamma'(t) = \alpha(t) E_1 + \beta(t) E_2.$$

This gives the system of equations

$$x' = \alpha;$$
$$y' = \beta;$$
$$z' = x\beta;$$

which gives

$$x(t) = \int_0^t \alpha(s)\, ds;$$
$$y(t) = \int_0^t \beta(s)\, ds;$$
$$z(t) = \int_0^t \beta(s) \int_0^s \alpha(s')\, ds'\, ds.$$

Then if $y = (X, Y, Z)$ we make the choice $\alpha(t) = X + af(t)$ and $\beta(t) = Y + bf(t)$, where a and b are constants to be chosen, and f is a smooth function which satisfies $\int_0^1 f(t)dt = 0$ and $\int_0^1 tf(t)ds = 0$, but $\Gamma = \int_0^1 \int_0^t f(t)f(s)ds\,dt \neq 0$ (for example, $f(t) = t^2 - t + 1/6$ will work). Then we have $x(1) = X$, $y(1) = Y$, and $z(1) = XY/2 + ab\Gamma$. Finally, we can choose a and b to ensure that $z(1) = Z$ (e.g. $a = 1$, $b = (Z - XY/2)/\Gamma$). $\square$

Example 7.3.3 Subgroups of Lie groups As an application of the Frobenius theorem, we will consider subgroups of Lie groups. We have seen that the left-invariant vector fields on a Lie group G form a finite-dimenensional Lie algebra $\mathfrak{g}$. We will see that subspaces of $\mathfrak{g}$ which are closed under the Lie bracket correspond exactly to connected smooth subgroups of G. In one direction this is clear: Suppose H is a Lie group which is contained in G, with the inclusion map being an immersion. Then the Lie algebra $\mathfrak{h}$ of H is naturally included in $\mathfrak{g}$, since for any left-invariant Lie algebra V on H, we can extend to a left-invariant Lie algebra $\tilde{V}$ on G by setting

$$\tilde{V}_g = D_0 l_g(V_e).$$

$\tilde{V}$ agrees with V on H. The image of this inclusion is a vector subspace of $\mathfrak{g}$. Furthermore, if $X, Y \in \mathfrak{h}$ then $[X,Y]\tilde{} = [\tilde{X}, \tilde{Y}]$ is in this subspace of $\mathfrak{g}$. So we have a subspace of $\mathfrak{g}$ which is closed under Lie brackets.

Proposition 7.3.7 *Let $\mathfrak{h}$ be any vector subspace of $\mathfrak{g}$ which is closed under Lie brackets. Then there exists a unique connected Lie group H and an inclusion $i : H \to G$ which is an injective immersion, such that $D_h i(T_h H)$ is the subspace of $T_{i(h)}G$ given by the left-invariant vector fields in $\mathfrak{h}$, for every $h \in H$.*

Proof. Since $\mathfrak{h}$ is closed under Lie brackets, the distribution $\mathcal{D}$ defined by the vectors in $\mathfrak{h}$ is involutive: If we take a basis $\{E_1, \ldots, E_k\}$ for $\mathfrak{h}$, then any $X, Y \in \mathcal{X}(M)$ can be written in the form $X = X^i E_i$, $Y = Y^j E_j$ for some smooth functions $X^1, \ldots, X^k$ and $Y^1, \ldots, Y^k$. Then we have

$$[X, Y] = [X^i E_i, Y^j E_j] = (X^i E_i(Y^j) - Y^i E_i(X^j))E_j + X^i Y^j [E_i, E_j]$$

which is in $\mathcal{D}$. By Frobenius' Theorem, there is a submanifold Σ passing through $e \in G$ with tangent space $\mathcal{D}$.

To show that Σ is a subgroup, we show that $xy \in \Sigma$ for all x and y in Σ: Write $y = \exp(sY)$ for $Y \in \mathcal{D}_e$. Then

$$\frac{d}{ds}\left(x \exp(sY)\right) = \frac{d}{dr}\left(x \exp(sY)\exp(rY)\right)\Big|_{r=0} = D_0 l_{x\exp(sY)} Y \in \mathcal{D}$$

for each s. Thus the curve $s \mapsto x\exp(sY)$ starts in Σ and is tangent to $\mathcal{D} = \mathcal{T}\Sigma$, and so stays in Σ. Therefore $xy \in \Sigma$, and Σ is a subgroup. $\square$

Lecture 8. Connections

This lecture introduces connections, which are the machinery required to allow differentiation of vector fields.

8.1 Differentiating vector fields.

The idea of differentiating vector fields in Euclidean space is familiar: If we choose a fixed orthonormal basis $\{e_1, \dots, e_n\}$ for $\mathbb{R}^n$, then any vector field X can be written in the form $\sum_{i=1}^n X^i e_i$ for some smooth functions $X^1, \dots, X^n$. Then the derivative of X in the direction of a vector v is given by

$$D_v X = \sum_{i=1}^n v(X^i) e_i.$$

In other words, we just differentiate the coefficient functions. The key point is that we think of the vectors $\{e_1, \dots, e_n\}$ as being *constant*.

When we are dealing with vector fields on a manifold, this no longer makes sense: There are no natural vector fields which we can take to be 'constant'.

Example 8.1.1 A failed attempt. One might think that a chart would supply natural vector fields that we can think of as being constant – every vector field can be written as a linear combination of the coordinate tangent vectors $\{\partial_1, \dots, \partial_n\}$. Unfortunately, defining derivatives in terms of these does not really make sense: The difficulty is that changing to a different chart would give a different answer for the derivative of a vector field.

To see this, suppose we have two charts, φ and η, and that a vector field X is given by $\sum_{i=1}^n X^i \partial_i^{(\varphi)}$, and a vector $v = \sum_{i=1}^n v^i \partial_i^{(\varphi)}$. Then we can define a derivative $\nabla_v^{(\varphi)} X$ by

$$\nabla_v^{(\varphi)} X = \sum_{i=1}^n \left(v^j \partial_j^{(\varphi)} \left(X^i \right) \right) \partial_i^{(\varphi)}.$$

Now work in the chart η: We have $\partial_i^{(\varphi)} = \sum_{j=1}^n \Lambda_i^j \partial_j^{(\eta)}$, where Λ is the matrix $D\left(\eta \circ \varphi^{-1}\right)$. Therefore we have

$$\nabla_v^{(\varphi)} X = \sum_{i=1}^n v^j \Lambda_j^k \partial_k^{(\eta)} \left(X^i\right) \Lambda_i^l \partial_l^{(\eta)}.$$

However, if we instead worked from the beginning in the chart η, we would have $X = (X^i \Lambda_i^j)\partial_j^{(\eta)}$ and $v = (v^i \Lambda_i^j)\partial_j^{(\eta)}$, and so

$$\nabla_v^{(\eta)} X = v^j \Lambda_j^k \partial_k^{(\eta)} \left(X^i \Lambda_i^l\right) \partial_l^{(\eta)} = \nabla_v^{(\varphi)} X + v^j X^i \Lambda_j^k \partial_k^{(\eta)} \left(\Lambda_i^l\right) \partial_l^{(\eta)}.$$

So these two notions of derivative don't agree if Λ is not constant (i.e. if the second derivatives of the transition map do not vanish).

8.2 Definition of a connection.

Since there does not seem to be any particularly natural way to define derivatives of vector fields on a manifold, we have to introduce extra structure on the manifold to do this. The extra structure required is a *connection* on M.

Definition 8.2.1 Let M be a smooth manifold. A **connection** on M is a map $\nabla : TM \times \mathcal{X}(M) \to TM$ such that $\nabla_v X \in T_x M$ if $v \in T_x M$, and

(1). $\nabla_{c_1 v_1 + c_2 v_2} X = c_1 \nabla_{v_1} X + c_2 \nabla_{v_2} X$ for all v_1 and v_2 in $T_x M$ for any $x \in M$, any real numbers c_1 and c_2, and any $X \in \mathcal{X}(M)$;
(2). $\nabla_v \left(c_1 X_1 + c_2 X_2\right) = c_1 \nabla_v X_1 + c_2 \nabla_v X_2$ for every $v \in TM$, all real numbers c_1 and c_2, and all X_1 and X_2 in $\mathcal{X}(M)$;
(3). $\nabla_v(fX) = v(f)X + f\nabla_v X$ for all $v \in TM$, $f \in C^\infty(M)$, and $X \in \mathcal{X}(M)$.

It is simple to check that the differentiation of vector fields in Eudclidean space satisfies these requirements. Condition (3) is the Leibniz rule for differentiation of vector fields, which requires the differentiation to be consistent with the differentiation of smooth functions.

Example 8.2.2 A natural connection for submanifolds of Euclidean space. Let M be an n-dimensional submanifold of $\mathbb{R}^N$. We will use the inner product structure of $\mathbb{R}^N$ to induce a connection on M: Let $\{e_1, \ldots, e_N\}$ be an orthonormal basis for $\mathbb{R}^N$. Then at a point $x \in M$, $T_x M$ is an n-dimensional subspace of $T_x\mathbb{R}^N \simeq \mathbb{R}^N$. Denote by π the orthogonal projection onto this subspace.

Suppose $v \in T_x M$, and $X \in \mathcal{X}(M)$. Then we can write $X = \sum_{\alpha=1}^N X^\alpha e_\alpha$. In this situation we can regard the vector fields e_α as being constant, so it makes sense to define the derivative $D_v X = \sum_{\alpha=1}^N v(X^\alpha) e_\alpha$. This satisfies all the requirements for a connection, expect that the result may not be a vector in TM, only in $\mathbb{R}^N$. So, to remedy this we can take the orthogonal projection of the result onto TM:

$$\nabla_v X = \pi\left(\sum_{\alpha=1}^{N} v(X^\alpha)e_\alpha\right).$$

Let us check explicitly that this is a connection: The linearity (with constant coefficients) in each argument is immediate. The Leibniz rule also holds:

$$\begin{aligned}
\nabla_v(fX) &= \nabla_v\left(\sum_{\alpha=1}^{N}(fX^\alpha)e_\alpha\right) \\
&= \pi\left(\sum_{\alpha=1}^{N} v(fX^\alpha)e_\alpha\right) \\
&= \pi\left(\sum_{\alpha=1}^{N}\left(v(f)X^\alpha + fv(X^\alpha)\right)e_\alpha\right) \\
&= v(f)\pi\left(\sum_{\alpha=1}^{N} X^\alpha e_\alpha\right) + f\pi\left(\sum_{\alpha=1}^{N} v(X^\alpha)e_\alpha\right) \\
&= v(f)X + f\nabla_v X
\end{aligned}$$

since $\pi X = X$.

Remark. A particle moving subject to a constraint that it lies in the submanifold (with no other external forces) moves as it would in free space, except that any component of its acceleration in direction normal to the surface are automatically cancelled out by the constraint forces. In other words, the motion of the particle is determined by the equation

$$\pi\left(\frac{d^2x}{dt^2}\right) = 0.$$

Example 8.2.3 The left-invariant connection on a Lie group Let G be a Lie group, and $E_1, \ldots, E_n$ a collection of left-invariant vector fields generated by some basis for T_eG. Then any vector field X can be written as a linear combination of these: $X = \sum_{i=1}^n X^i E_i$ for some smooth functions $X^1, \ldots, X^n$. We can define a connection on G, called the *left-invariant connection*, by setting the vector fields E_i to be constant:

$$\nabla_v X = \sum_{i=1}^{n} v(X^i)E_i.$$

This connection is independent of the choice of the basis vector fields $E_1, \ldots, E_n$: If $\gamma : I \to G$ is a path with $\gamma'(0) = v \in T_gG$, then the connection can be written in the form

$$\nabla_v X = D_e l_g\left(\frac{d}{dt}\left(D_e l_{\gamma(t)}\right)^{-1} X_{\gamma(t)}\right).$$

This makes sense because $(D_e l_{\gamma(t)})^{-1} X_{\gamma(t)} \in T_eG$ for all t, and we know how to differentiate vectors in a vector space.

Note that we could just as well have considered right-invariant vector fields to define the connection. This gives the *right-invariant connection* on G, which is generally different from the left-invariant connection if G is a non-commutative Lie group.

8.3 Connection coefficients.

Suppose we are working in a chart $\varphi : U \to V \subset \mathbb{R}^n$ for M, with corresponding coordinate tangent vectors $\partial_1, \ldots, \partial_n$. The *connection coefficients* of a connection ∇ with respect to this chart are defined by

$$\nabla_{\partial_i}\partial_j = \Gamma_{ij}{}^k \partial_k.$$

Proposition 8.3.1 *The connection is determined on U by the connection coefficiants.*

Proof. Write $X = X^i\partial_i$. Then we have for any vector $v = v^k\partial_k$

$$\nabla_v X = \nabla_v \left(X^i\partial_i\right) = v(X^i)\partial_i + X^i\nabla_v\partial_i = v(X^i)\partial_i + v^kX^i\Gamma_{ki}{}^j\partial_j.$$

□

Exercise 8.3.2 Suppose we change to a different chart η. Write down the connection coefficients with respect to η, in terms of the connection coefficients for φ, and the map Λ_i^j defined by $D(\eta \circ \varphi^{-1})(e_i) = \Lambda_i^j e_j$.

8.4 Vector fields along curves

Let $\gamma : I \to M$ be a smooth curve. Then by a *vector field along* γ we will mean a smooth map $V : I \to TM$ with $V(t) \in T_{\gamma(t)}M$ for each t. The set of smooth vector fields alonmg γ will be denoted by $\mathcal{X}_\gamma(M)$.

Proposition 8.4.1 *Let M be a smooth manifold and ∇ a connection on M. Then for any smooth curve $\gamma : I \to M$ there is a natural covariant derivative along γ, denoted ∇_t, which takes a smooth vector field along γ to another. This map satisfies:*

(1). $\nabla_t(V + W) = \nabla_tV + \nabla_tW$ *for all* V *and* W *in* $\mathcal{X}_\gamma(M)$*;*
(2). $\nabla_t(fV) = \frac{df}{dt}V + f\nabla_tV$ *for all* $f \in C^\infty(I)$ *and* $V \in \mathcal{X}_\gamma(M)$*;*
(3). *If V is given by the restriction of a vector field $\tilde{V} \in \mathcal{X}(M)$ to γ, i.e. $V_t = \tilde{V}_{\gamma(t)}$, then $\nabla_tV = \nabla_{\dot\gamma}\tilde{V}$, where $\dot\gamma$ is the tangent vector to γ.*

The main point of this proposition is that the derivative of a vector field V in the direction of a vector v can be computed if one only knows the values of V along some curve with tangent vector v.

The covariant derivative along γ is defined by

$$\nabla_t \left(V^i(t)\partial_i\right) = \frac{dV^i}{dt}\partial_i + \dot{\gamma}^j V^i {\Gamma_{ji}}^k(\gamma(t))\partial_k$$

where $\dot{\gamma} = \dot{\gamma}^k\partial_k$. The first two statements of the proposition follow immediately, and it remains to check that our definition is consistent with the third.

Let $\tilde{V}$ be a smooth vector field in M. Then the induced vector field along a curve γ is given by $V(t) = \tilde{V}_{\gamma(t)}$, or in terms of the coordinate tangent basis, $V(t) = V^i(t)\partial_i$, where $\tilde{V}_x = \tilde{V}^i(x)\partial_i$, and $V^i(t) = \tilde{V}^i(\gamma(t))$.

Then

$$\begin{aligned}
\nabla_{\dot{\gamma}}\tilde{V} &= \nabla_{\dot{\gamma}}\left(\tilde{V}^i\partial_i\right) \\
&= \left(D\gamma(\partial_t)\right)(\tilde{V}^i)\partial_i + \dot{\gamma}^k\tilde{V}^i {\Gamma_{ki}}^l\partial_l \\
&= \partial_t\left(\tilde{V}^i\circ\gamma\right)\partial_i + \dot{\gamma}^k V^i {\Gamma_{ki}}^l(\gamma(t))\partial_i \\
&= \nabla_t V.
\end{aligned}$$

8.5 Parallel transport.

Definition 8.5.1 Let $\gamma : I \to M$ be a smooth curve. A vector field $V \in \mathcal{X}_\gamma(M)$ is said to be **parallel** along γ if $\nabla_t V = 0$.

For example, the constant vector fields in $\mathbb{R}^n$ are parallel with respect to the standard connection, and the left-invariant vector fields on a Lie group are parallel with respect to the left-invariant connection.

Proposition 8.5.2 *Let $\gamma : I \to M$ be a piecewise smooth curve. Then there exists a unique family of linear isomorphisms $P_t : T_{\gamma(0)}M \to T_{\gamma(t)}M$ such that a vector field V along γ is parallel if and only if $V(t) = P_t(V_0)$ for all t.*

The maps P_t are the *parallel transport* operators along γ.

To prove the proposition, consider the differential equation that defines when a vector field along γ is parallel, on an interval I where γ remains inside some chart with corresponding connection coefficients ${\Gamma_{kl}}^m$:

$$0 = \nabla_t V = \left(\frac{dV^i}{dt} + {\Gamma_{kl}}^i\dot{\gamma}^k V^l\right)\partial_i.$$

This means that for $i = 1, \ldots, n$,

$$0 = \frac{dV^i}{dt} + \Gamma_{kl}{}^i \dot\gamma^k V^l. \tag{8.5.1}$$

This is a linear system of n first order ordinary differential equations for the n functions $V^1 \ldots, V^n$ along I, and so (since the coefficient functions $\dot\gamma^k \Gamma_{kl}{}^i$ are bounded and piecewise continuous) there exist unique solutions with arbitrary initial values $V^1(0), \ldots, V^n(0)$. We define P_t by $P_t(V^i(0)\partial_i) = V^i(t)\partial_i$, where $V^i(t)$ satisfy the system (8.5.1). We leave it to the reader to check that this defines a linear isomorphism for each t.

Proposition 8.5.3 *The covariant derivative of a vector field $V \in \mathcal{X}_\gamma(M)$ along γ can be written as*

$$\nabla_t V = P_t\left(\frac{d}{dt}\left((P_t)^{-1}V(t)\right)\right).$$

Proof. Choose a basis $\{e_1, \ldots, e_n\}$ for $T_{\gamma(0)}M$, and define $E_i(t) = P_t(e_i) \in T_{\gamma(t)}M$. Then we can write $V(t) = V^i(t)E_i(t) = V^i(t)P_t(e_i) = P_t(V^i(t)e_i)$, or in other words

$$P_t^{-1}(V) = V^i(t)e_i.$$

Therefore we have

$$\frac{d}{dt}\left(P_t^{-1}V(t)\right) = \left(\frac{d}{dt}V^i\right)e_i,$$

and

$$P_t\frac{d}{dt}\left(P_t^{-1}V(t)\right) = \left(\frac{d}{dt}V^i(t)\right)E_i(t).$$

On the other hand we have

$$\nabla_t V = \nabla_t\left(V^i(t)E_i(t)\right) = \left(\frac{d}{dt}V^i(t)\right)E_i(t) + V^i(t)\nabla_t E_i = \left(\frac{d}{dt}V^i(t)\right)E_i(t)$$

since E_i is parallel. □

This tells us that the connection is determined by the parallel transport operators. The parallel transport operators give a convenient way to identify the tangent spaces to M at different points along a smooth curve, and in some ways this is analogous to the left-shift maps $D_e l_g$ on a Lie group. However, it is important to note that the parallel transport operators depend on the curve. In particular the parallel transport operators cannot be extended to give canonical identifications of all the tangent spaces to each other (if we could do this, we could construct non-vanishing vector fields, but we already know this is impossible in some situations). Also, it does not really makes sense to think of a parallel vector field as being "constant", as the following example illustrates:

Example 8.5.4 Parallel transport on the sphere. Let $M = S^2$, and take the submanifold connection given by Example 8.2.2. Then a vector field along a curve in S^2 is parallel if and only if its rate of change (as a vector in $\mathbb{R}^3$) is always normal to the surface of S^2.

Consider the path γ on S^2 which starts at the north pole, follows a line of longitude to the equator, follows the equator for some distance (say, a quarter of the way around) and then follows another line of longitude back to the north pole. Note that each of the three segments of this curve is a geodesic. We compute the vector field given by parallel transport along γ of a vector which is orthogonal to the initial velocity vector at the north pole.

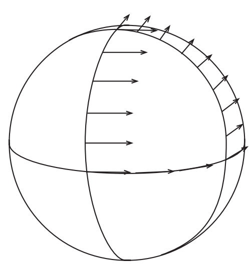
On the first segment, the parallel transport keeps the vector constant as a vector in $\mathbb{R}^3$ (this must be the parallel transport since it remains tangent to S^2, and has zero rate of change, so certainly the tangential component of its rate of change is zero).

On the segment around the equator, we start with a vector tangent to the equator, and parallel transport will give us the tangent vector to the equator of the same length as we move around the equator: Then we have $V = L\frac{dx}{dt}$ for some constant L (the length), and so

$$\dot{V} = L\frac{d^2x}{dt^2} = -Lx$$

since $x(t) = (\cos t, \sin t, 0)$. In particular, $\dot{V}$ is normal to S^2.

On the final segment, the situation is the same as the first segment: We can take V to be constant as a vector in $\mathbb{R}^3$.



So parallel transport around the entire loop has the effect of rotating the vector through $\pi/2$.

8.6 The connection as a projection.

Another way of looking at connections is the following: Let us return to the original problem, of defining a derivative for a vector field. Recall that a vector field $V \in \mathcal{X}(M)$ can be regarded as a smooth map $V : M \to TM$ for which $V_x \in T_xM$ for all $x \in M$. Since this is just a smooth map between manifolds, we can differentiate it!

This gives the derivative map $DV : TM \to TTM$. In other words, we can think of the derivative of a vector field on M in a direction $v \in TM$ as a vector tangent to the $2n$-dimensional manifold $T(TM)$. To get a better understanding of this, look at the situation geometrically: We think of TM as being a union of fibres T_xM. Thus a tangent vector to TM will have some component tangent to the fibre T_xM, and some component transverse to the fibre. In local coordinates $x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n$, the tangent vectors $\partial_1, \ldots, \partial_n$ corresponding to the first n coordinates represent change in position in M, which means that motion in these directions amounts to moving across a family of 'fibres' in TM; the tangent vectors $\dot{\partial}_1, \ldots, \dot{\partial}_n$ corresponding to the last n coordinates (which are just the components of the vector in TM) are tangent to the fibres.

A vector field V can be written in these coordinates as $V(x^1, \ldots, x^n) = (x^1, \ldots, x^n, V^1, \ldots, V^n)$. Then

$$D_vV = v^i\partial_i + v(V^i)\dot{\partial}_i.$$

The idea of the connection is to project this onto the subspace of $T(TM)$ tangent to the fibre, which we can naturally identify with the fibre itself (the fibre is a vector space, so we can identify it with its tangent space at each point). We will denote by $\mathcal{V}$ the subspace of $T(TM)$ spanned by $\dot{\partial}_1, \ldots, \dot{\partial}_n$ (note that this is independent of the choice of local coordinates), and we call this the *vertical subspace* of $T(TM)$. This is naturally identified with TM by the map ι which sends $v^i\dot{\partial}_i$ to $v^i\partial_i$.

Definition 8.6.1 A **vertical projection** on TM is a map $\xi \mapsto \pi(\xi)$, where $\xi = (p, v) \in TM$ and $\pi(\xi)$ is a linear map from $T_\xi(TM) \to T_pM$ such that

(1). $\pi(\xi) = \iota$ on $\mathcal{V}$;

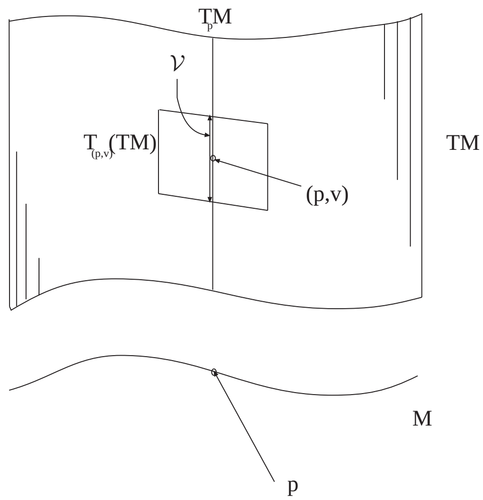
(2). π is consistent with the additive structure on TM: If we take γ_1 and γ_2 to be paths in TM of the form $\gamma_i(t) = (p(t), v_i(t))$, then

$$\pi_{(p,v_1)}(\gamma_1') + \pi_{(p,v_2)}(\gamma_2') = \pi_{(p,v_1+v_2)}(\gamma')$$

where $\gamma(t) = (p(t), v_1(t) + v_2(t))$.

Given a vertical projection π, we can produce a connection as follows: If $X \in \mathcal{X}(M)$ and $v \in T_pM$, define

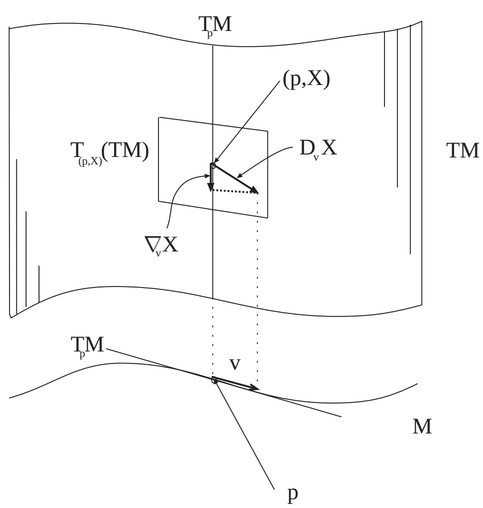
$$\nabla_vX = \pi_{(p,X_p)}(D_vX).$$



Conversely, given a connection, we can produce a corresponding vertical projection by taking

$$\pi_{(p,X)}(v^i\partial_i + \dot{v}^i\dot{\partial}_i) = \nabla_t V,$$

the covariant derivative along the curve $\gamma(t) = \varphi^{-1}\left(\varphi(p) + tv^i e_i\right)$ of the vector field $V(t) = \left(X^i + t\dot{v}^i\right)\partial_i$.



It is interesting to consider the parallel transport operators in terms of the vertical projections: Since π maps a $2n$-dimensional vector space to an n-dimensional vector space, and is non-degenerate on the n-dimensional 'vertical' subspace tangent to the fiber of TM, the kernel of π is an n-dimensional subspace of $T(TM)$ which is complementary to the vertical subspace. We call this the *horizontal subspace* $\mathcal{H}$ of $T(TM)$. A vector field X is parallel along a curve γ if and only if D_tX lies in $\mathcal{H}$ at every point. We will come back to this description when we discuss curvature. A vertical projection is uniquely determined by the choice of a horizontal subspace at each point, complementary to the vertical subspace and consistent with the linear structure.

8.7 Existence of connections.

We will show that every smooth manifold can be equipped with a connection – in fact, there are many connections on any manifold, and no preferred or canonical one (later, when we introduce Riemannian metrics, we will have a way of producing a canonical connection).

Definition 8.7.1 Let M be a smooth manifold. A (smooth) **partition of unity** on M is a collection of smooth functions $\{\rho_\alpha\}_{\alpha\in I}$ which are non-negative, locally finite (i.e. for every $x \in M$ there exists a neighbourhood U of x on which only finitely many of the functions ρ_α are non-zero), and sum to unity: $\sum_\alpha \rho_\alpha(x) = 1$ for all $x \in M$ (note that this is a finite sum for each x). Let $\{U_\beta\}_{\beta\in J}$ be a locally finite covering of M by open sets. Then a partition of unity $\{\rho_\alpha\}$ is **subordinate** to this cover if each of the functions ρ_α has support contained in one of the sets U_β.

We assume that our manifolds are paracompact – that is, every covering of M by open sets has a refinement which is locally finite. In particular, we have an atlas $\mathcal{A}$ for M which is locally finite (i.e. each point of M lies in only finite many of the coordinate charts of the atlas).

Proposition 8.7.2 *For any locally finite cover of M by coordinate charts there exists a subordinate partition of unity.*

Proof. By refining if necessary, assume that the coordinate charts have images in $\mathbb{R}^n$ which are bounded open sets which have smooth boundary. Consider one of these charts, $\varphi_\alpha : U_\alpha \to V_\alpha$. We can choose a number $\epsilon > 0$ such that the distance function $d(y) = d(y, \partial V_\alpha)$ to the boundary of V_α is smooth at points where $d(y) < \varepsilon$. Then define

$$\psi_\alpha = F_\varepsilon \circ d \circ \varphi_\alpha,$$

where $F_\varepsilon : \mathbb{R} \to \mathbb{R}$ is defined by $F_\varepsilon(t) = F_1(\varepsilon t)$, and F_1 is a smooth function which is identically zero for $t \leq 0$, monotone increasing for $t \in (0, 1)$, and identically 1 for $t \geq 1$. It follows that ψ_α is smooth on U_α, extends smoothly to M by taking zero values away from U_α, and is strictly positive on U_α.

Finally, define

$$\rho_\alpha(x) = \frac{\psi_\alpha(x)}{\sum_{\alpha'} \psi_{\alpha'}(x)}.$$

□

Proposition 8.7.3 *For any smooth manifold M there exists a connection on M.*

Proof. Choose a cover of M by coordinate charts $\{U_\alpha\}$, and a subordinate partition of unity $\{\rho_\alpha\}$. We will use this partition of unity to patch together the connections on each chart given by coordinate differentiation: Recall that on each coordinate chart U_α we have a derivative $\nabla^{(\alpha)}$ which is well-defined on vector fields with support inside U_α. Then we define

$$\nabla_v X = \sum_\alpha \nabla_v^{(\alpha)} \left(\rho_\alpha X\right).$$

This makes sense because the sum is actually finite at each point of M. The result is clearly linear (over $\mathbb{R}$) in both v and X, and

$$\begin{aligned}
\nabla_v(fX) &= \sum_\alpha \nabla_v^{(\alpha)} \left(f\rho_\alpha X\right) \\
&= \sum_\alpha \Big(v(f)\rho_\alpha X + f\nabla_v^{(\alpha)}\left(\rho_\alpha X\right)\Big) \\
&= v(f)X + f\nabla_v X
\end{aligned}$$

where I used the fact that $\sum_\alpha \rho_\alpha = 1$ to get the first term in the last equality.

□

The following exercise is a another application of partitions of unity, to show the existence of F-related vector fields. Recall that if $F : M \to N$ is an smooth map, then $X \in \mathcal{X}(M)$ and $Y \in \mathcal{X}(N)$ are called F-related if $D_x F(X_x) = Y_{F(x)}$ for every $x \in M$.

Exercise 8.7.2 Suppose $F : M \to N$ is an embedding. Show that for any vector field $X \in \mathcal{X}(M)$ there exists a (generally not unique) F-related vector field $Y \in \mathcal{X}(Y)$ [Hint: Cover N by coordinate charts, such that those containing a portion of $F(M)$ are submanifold charts (i.e. $\varphi_\alpha : U_\alpha \to V_\alpha$ such that $F(M) \cap U_\alpha = \varphi_\alpha^{-1}(\mathbb{R}^k \times \{0\})$). Choose a partition of unity subordinate to this cover. Use each submanifold chart to extend X from $F(M)$ to a vector field on the chart in N, and take a zero vector field in all the other charts on N. Patch these together using the partition of unity and check that it gives a vector field with the required properties].

8.9 The difference of two connections

Suppose $\nabla^{(1)}$ and $\nabla^{(2)}$ are two connections on a smooth manifold M. Then we have the following important result:

Proposition 8.9.1 *For any vector $v \in T_xM$ and any vector field $X \in \mathcal{X}(M)$, the difference*

$$\nabla^{(2)}_v X - \nabla^{(1)}_v X$$

depends only on the value of X at the point x.

Proof. In local coordinates, writing $v = v^i\partial_i$ and $X = X^j\partial_j$,

$$\begin{aligned}
\nabla^{(2)}_v X - \nabla^{(1)}_v X &= v(X^j)\partial_j + v^j X^i {\Gamma^{(1)}_{ji}}^k \partial_k - v(X^j)\partial_v - v^j X^i {\Gamma^{(2)}_{ji}}^k \partial_k \\
&= v^j X^i \left({\Gamma^{(2)}_{ji}}^k - {\Gamma^{(1)}_{ji}}^k \right) \partial_k .
\end{aligned}$$

□

It follows that the difference of two connections defines a map for each $x \in M$ from $T_xM \times T_xM$ to T_xM, which varies smoothly in x and is linear in each argument. We also have a kind of converse to this:

Proposition 8.9.2 *Let M be a smooth manifold, and ∇ a connection on M. Suppose $A_x : T_xM \times T_xM \to T_xM$ is linear in each argument for each $x \in M$, and varies smoothly over M (in the sense that A applied to two smooth vector fields gives a smooth vector field). Then*

$$x, X \mapsto \nabla_v X + A(v, X)$$

is also a connection on M.

The proof is a simple calculation to check that the Leibniz rule still holds. It is easy to show that there are many such suitable maps A (for example, one can write in a local coordinate chart $A(u^i\partial_i, v^j\partial_j) = u^i v^j {A_{ij}}^k \partial_k$ where the coefficient functions ${A_{ij}}^k$ are smooth and have support inside the coordinate chart. Then sums and linear combinations of such things are also suitable. Thus there are many different connections on a manifold.

8.10 Symmetry and torsion

Given a connection ∇, we define its torsion in the following way: If X and Y are two smooth vector fields, then we observe that

$$\nabla_X Y - \nabla_Y X - [X, Y],$$

evaluated at a point $x \in M$, depends only on the values of X and Y at the point x. This follows because

$$\begin{aligned}
\nabla_{fX}(gY) - \nabla_{gY}(fX) - [fX, gY] &= fg\nabla_X Y + fX(g)Y \\
&\quad - fg\nabla_Y X - gY(f)X \\
&\quad - fX(g)Y + gY(f)X - fg[X, Y] \\
&= fg\left(\nabla_X Y - \nabla_Y X - [X, Y]\right).
\end{aligned}$$

This map is called the *torsion* $T(X, Y)$ of ∇ applied to X and Y. This can be written in terms of its components ${T_{ij}}^k$, defined by $\nabla_{\partial_i}\partial_j - \nabla_{\partial_j}\partial_i = {T_{ij}}^k\partial_k$.

Definition 8.10.1 A connection ∇ is called **symmetric** if the torsion of ∇ vanishes — i.e. $\nabla_X Y - \nabla_Y X = [X, Y]$ for every pair of vector fields X and Y on M.

The standard connection on $\mathbb{R}^n$ is symmetric, as are submanifold connections. However, the left-invariant connection on a Lie group is typically not symmetric: If X and Y are left-invariant vector fields, then $\nabla_X Y = \nabla_Y X = 0$, but if G is non-commutative then usually $[X, Y] \neq 0$.

8.11 Geodesics and the exponential map

Definition 8.11.1 Let M be a smooth manifold, and ∇ a connection. A curve $\gamma : [0, 1] \to M$ is a **geodesic** of ∇ if the tangent vector $\dot\gamma$ is parallel along γ:

$$\nabla_t \dot\gamma = 0.$$

We can think of a path γ as the trajectory of a particle moving in M. Then γ is a geodesic if the acceleration of the particle (measured using ∇) is zero, so γ is in this sense the trajectory of a 'free' particle. This picture is particularly relevant in the case where M is a submanifold of Eulidean space, since then the geodesics are exactly the trajectories of particles which are constrained to lie in M but are otherwise free of external forces.

Proposition 8.11.2 *For any $x \in M$ and any $v \in T_x M$ there exists $\delta > 0$ and a unique geodesic $\gamma : (-\delta, \delta) \to M$ such that $\gamma(0) = x$ and $\dot\gamma(0) = v$.*

Proof. Let $\varphi : U \to V$ be a chart for M about x. Then for any path $\gamma : I \to M$, we have

$$\nabla_t\dot\gamma = \nabla_t\left(\dot\gamma^i\partial_i\right) = \left(\ddot\gamma^i + \Gamma_{kl}{}^i\dot\gamma^k\dot\gamma^l\right)\partial_i,$$

so a geodesic must satisfy the system of second order equations

$$\ddot\gamma^i + \Gamma_{kl}{}^i\dot\gamma^k\dot\gamma^l = 0$$

for $i = 1, \ldots, n$. Existence and uniqueness then follow from the standard theory of second order ODE. Alternatively, we can rewrite as a system of $2n$ first-order equations by setting $v^i = \dot\gamma^i$, which gives

$$\frac{d}{dt}\gamma^i(t) = v^i(t);$$
$$\frac{d}{dt}v^i(t) = -\Gamma_{kl}{}^i(\gamma^1(t), \ldots, \gamma^n(t))v^k(t)v^l(t).$$

Then apply the ODE existence and uniqueness theorem from Lecture 6 to deduce that there is a unique geodesic for some time starting from any point x and any vector v (i.e. any values of $\gamma^1, \ldots, \gamma^n, v^1, \ldots, v^n$). $\square$

Using Proposition 8.11.2 we can define the *exponential map* on a smooth manifold with a connection, by analogy with the exponential map on a Lie group: Given $x \in M$ and $v \in T_xM$, we define

$$\exp_p(v) = \exp(p, v) = \gamma(1),$$

where γ is a geodesic with initial position x and initial velocity v. Then we know that the exponential map is defined on a neighbourhood of the zero section $\{(x, 0) : x \in M\}$ within TM. In order to determine the smoothness propoerties of the exponential map, it is useful to interpret it using the flow of a vector field so that we can apply the methods of Lecture 6. As we saw above, the geodesic equation corresponds to a second-order system of equations, and it turns out that this arises not from the flow of a vector field on M, but from the flow of a vector field on the tangent bundle TM: Recall that the tangent bundle can be given local coordinates $x^1, \ldots, x^n, \dot x^1, \ldots, \dot x^n$. Then let $\partial_1, \ldots, \partial_n, \dot\partial_1, \dot\partial_n$ be the corresponding coordinate tangent vectors, and define a vector field $\mathcal{G}$ on TM by

$$\mathcal{G}_{(x,v)} = \dot x^i\partial_i - \Gamma_{kl}{}^i(x)\dot x^k\dot x^l\dot\partial_i.$$

Then our local calculations above show that the geodesic equation is given exactly by the flow for time 1 of the vector field $\mathcal{G}$. More precisely, it is given by the projection onto M of this flow.

Exercise 8.11.3 Show that $\mathcal{G}$, defined as above in local coordinates, is well-defined (i.e. does not depend on the choice of coordinates).

It follows immediately that the exponential map is smooth, and that the map

$$(x, v) \mapsto \left(\exp_x(v), \frac{d}{dt} \exp_x(tv)\Big|_{t=1} \right)$$

from TM to itself is a local diffeomorphism where defined.

Proposition 8.11.4 *For any $x \in M$, the exponential map $\exp_x : T_xM \to M$ from x is a local diffeomorphism near $0 \in T_xM$, onto a neighbourhood of x in M.*

Proof. The smoothness of the map is immediate since it is the composition of the smooth flow of the vector field $\mathcal{G}$ on TM with the smooth projection onto M. Next we calculate the derivative of this map at the origin in the direction of some vector v: We have $\exp_x(v) = \gamma(1)$, where γ is the geodesic through x with velocity v. But then also $\exp_x(tv) = \gamma(t)$, since $s \mapsto \gamma(st)$ is a geodesic through x with velocity tv. Therefore

$$D_0 \exp_x(v) = \frac{d}{dt} \exp_x(tv)\Big|_{t=0} = \gamma'(0) = v$$

Therefore $D_0 \exp_x$ is just the identity map (here we are identifying the tangent space to T_xM at 0 with T_xM itself, in the usual way). In particular, this derivative is non-degenerate, so $\exp_x$ is a local diffeomorphism. □

We can get a slightly stronger statement:

Proposition 8.11.5 *For any $x \in M$ there exists an open neighbourhood U of $x \in M$, and an open neighbourhood $\mathcal{O}$ of the zero section $\{(y, 0) : y \in U\}$ such that for every pair of points p and q in U, there exists a unique geodesic γ_{qr} with $\gamma_{qr}(0) = 1$, $\gamma_{qr}(1) = r$, and $\gamma'(0) \in \mathcal{O}$.*

Roughly speaking, this says that points which are sufficiently close to each other can be joined by a unique "short" geodesic (we cannot expect to drop the shortness condition – consider for example the case where M is a cylinder or a sphere, when there are generally many geodesics joining any pair of points.

Proof. Consider the map $\tilde{\exp} : TM \to M \times M$ defined by

$$\tilde{\exp}(p, v) = (p, \exp_p(v)).$$

Let us compute the derivative of this map at a point $(p, 0)$: Breaking up the $2n \times 2n$ derivative matrix into $n \times n$ blocks, the top left corresponds to the derivative of the map $(p, v) \mapsto p$ with respect to p, which is just the identity map, and the bottom left block is the derivative of the same with respect to v, which is zero. The top right block might be quite complicated, but as we calculated in the proof of proposition 8.11.4, the bottom right block is the derivative of the exponential map from p with respect to v at the origin, which is again just the identity map:

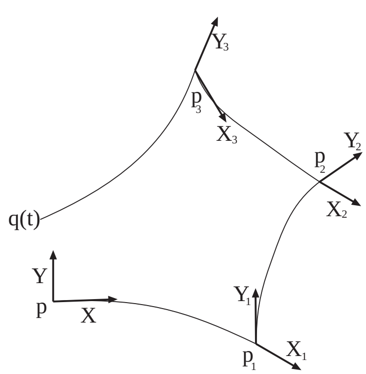
$$D_{(p,0)}\tilde{\exp} = \begin{bmatrix} I & * \\ 0 & I \end{bmatrix}.$$

In particular, this is non-degenerate, so $\tilde{\exp}$ is a local diffeomorphism near $(p,0)$ for any p. Choose a small neighbourhood $\mathcal{O}$ of $(x,0)$ in TM, and a neighbourhood U of x in M sufficiently small that $U \times U \subset \tilde{\exp}(\mathcal{O})$. These then satisfy the requirements of the proposition. □

The following exercise gives an interpretation of the torsion of a connection in terms of geodesics and parallel transport:

Exercise 8.11.6 Let M be a smooth manifold with a connection ∇. Fix $p \in M$, and vectors X and Y in T_pM. Let $q(t)$ be the point of M constructed as follows: First, follow the geodesic from p in direction X until time $\sqrt{t}$, to get a point p_1. Then parallel transport Y along that geodesic to get a vector Y_1 at p_1. Also, let X_1 be tangent vector to the geodesic at time t (i.e. at the point p_1). Then follow the geodesic from p_1 in direction Y_1 until time $\sqrt{t}$, to get a point p_2, and parallel transport X_1 and Y_1 along this geodesic to get vectors X_2 and Y_2 at p_2 (Thus, Y_2 is the tangent vector of the geodesic at p_2). Then take the geodesic from p_2 in direction $-X_2$ until time $\sqrt{t}$, parallel transporting X_2 and Y_2 along this to get vectors X_3 and Y_3 at a point p_3. Finally, follow the geodesic from p_3 in direction $-Y_3$ until time $\sqrt{t}$, and call the resulting point $q(t)$.

Show that the tangent vector to the curve $t \mapsto q(t)$ is given by the torsion of ∇ applied to the vectors X and Y. In other words, the torsion measures the extent to which geodesic rectangles fail to close up.



Lecture 9. Riemannian metrics

This lecture introduces Riemannian metrics, which define lengths of vectors and curves in the manifold.

9.1 Definition

There is one more crucial ingredient which we need to introduce for dealing with manifolds: Lengths and angles. Given a smooth manifold, since we know what it means for a curve in the manifold to be smooth, and we have a well-defined notion of the tangent vector to a curve, all we need in order to have a notion of distance on the manifold is a way of defining the *speed* of a curve — that is, the length of its tangent vector.

Definition 9.1.1 A **Riemannian metric** g on a smooth manifold M is a smoothly chosen inner product $g_x : T_xM \times T_xM \to \mathbb{R}$ on each of the tangent spaces T_xM of M. In other words, for each $x \in M$, $g = g_x$ satisfies
(1). $g(u,v) = g(v,u)$ for all $u, v \in T_xM$;
(2). $g(u,u) \geq 0$ for all $u \in T_xM$;
(3). $g(u,u) = 0$ if and only if $u = 0$.

Furthermore, g is smooth in the sense that for any smooth vector fields X and Y, the function $x \mapsto g_x(X_x, Y_x)$ is smooth.

Locally, a metric can be described in terms of its coefficients in a local chart, defined by $g_{ij} = g(\partial_i, \partial_j)$. The smoothness of g is equivalent to the smoothness of all the coefficient functions g_{ij} in some chart.

Example 9.1.2 The standard inner product on Euclidean space is a special example of a Riemannian metric. $\mathbb{R}^n$ can be made a Riemannian manifold in many ways: Let f_{ij} be a bounded, smooth function for each i and j in $\{1, \ldots, n\}$, with $f_{ij} = f_{ji}$. Then for C sufficiently large, the functions $g_{ij} = C\delta_{ij} + f_{ij}$ are positive definite everywhere, and so define a Riemannian metric.

9.2 Existence of Riemannian metrics

Every smooth manifold carries a Riemannian metric (in fact, many of them). We will prove this using an argument very similar to that used in showing the existence of connections.

Choose an atlas $\{\varphi_\alpha : U_\alpha \to V_\alpha\}$, and a subordinate partition of unity $\{\rho_\alpha\}$. On each of the regions V_α in $\mathbb{R}^n$, choose a Riemannian metric $g^{(\alpha)}$ (as in example 9.1.2). Then define

$$g(u,v) = \sum_\alpha \rho_\alpha g^{(\alpha)}\left(D\varphi_\alpha(u), D\varphi_\alpha(v)\right).$$

This is clearly symmetric; $g(u,u) \geq 0$; and $g(u,u)=0$ iff $u=0$. Furthermore it is smooth, and so defines a Riemannian metric on M.

9.3 Length and distance

Definition 9.3.1 Let $\gamma : [a,b] \to M$ be a (piecewise) smooth curve. Then the **length** $L[\gamma]$ of γ is defined by $L[\gamma] = \int_a^b g_{\gamma(t)}(\gamma'(t),\gamma'(t))^{1/2}\,dt$. Given two points p and q in M, we define the **distance** from p to q by

$$d(p,q) = \inf\left\{L[\gamma]\ \middle|\ \gamma : [a,b] \to M \text{ piecewise smooth}, \gamma(a)=p, \gamma(b)=q\right\}.$$

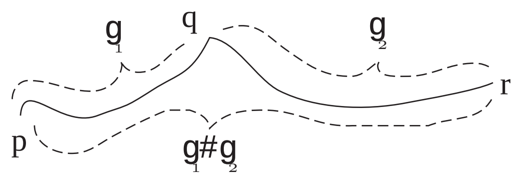
Proposition 9.3.2 *If M is a Riemannian manifold with metric g, then M is a metric space with the distance function d defined above. The metric topology agrees with the manifold topology.*

Proof. The symmetry of the distance function is immediate, as is its non-negativity. The triangle inequality is also easily established: For any curves $\gamma_1 : [a,b] \to M$ and $\gamma_2 : [a',b'] \to M$ with $\gamma_1(b) = \gamma_2(a')$, we can define the concatenation $\gamma_1\#\gamma_2 : [0, b+b'-a-a'] \to M$ by

$$\gamma_1\#\gamma_2(t) = \begin{cases} \gamma_1(t+a) & \text{for } 0 \leq t \leq b-a; \\ \gamma_2(t+a-b+a') & \text{for } b-a \leq t \leq b+b'-a-a'. \end{cases}$$

Thus $\gamma_1\#\gamma_2$ is a curve with length $L[\gamma_1]+L[\gamma_2]$. Given points p, q, and r in M, for any $\varepsilon > 0$ we can choose curves γ_1 joining p to q and γ_2 joining q to r, such that $L[\gamma_1] < d(p,q)+\varepsilon$ and $L[\gamma_2], d(q,r)+\varepsilon$. Therefore

$$d(p,r) \leq L[\gamma_1\#\gamma_2] = L[\gamma_1]+L[\gamma_2] < d(p,q)+d(q,r)+2\varepsilon,$$



which gives the triangle inequality when $\varepsilon \to 0$.

We still need to check that $d(x, y) = 0$ only when $x = y$. Suppose we have $x \neq y$ with $d(x, y) = 0$. Choose a chart $\varphi : U \to V$ around x. Then we can choose $\delta > 0$ and $C > 0$ such that on $B_\delta(\varphi(x)) \subset V$, $g(u, u) \geq C|D\varphi(u)|^2$. Therefore for points z in $\varphi^{-1}(B_\delta(\varphi(x))$, we have $d(x, z) \geq C|\varphi(x) - \varphi(z)|$. So y cannot be in this set. But for y outside this set, any curve from x to y must first pass through $\varphi^{-1}(\partial B_\delta(\varphi(x)))$, and so has length at least $C\delta$, contradicting $d(x, y) = 0$.

The claim that the metric topology is equivalent to the manifold topology follows similarly: The metric restricted to charts is comparable to the Euclidean distance on the chart. □

9.4 Submanifolds

An important situation where a manifold can be given a Riemannian metric is when it is a submanifold of some Euclidean space $\mathbb{R}^N$. The inclusion into $\mathbb{R}^N$ gives a natural identification of tangent vectors to M with vectors in $\mathbb{R}^N$: Explicitly, if we write the inclusion as a map $F : M \to \mathbb{R}^N$, then we identify a vector $u \in TM$ with its image $DF(u)$ under the differential of the inclusion.

The inner product on $\mathbb{R}^N$ can then be used to induce a Riemannian metric on M, by defining

$$g(u, v) = \langle DF(u), DF(v)\rangle.$$

Similarly, if N is a Riemannian manifold with a metric h, and $F : M \to N$ is an immersion, then we can define the induced Riemannian metric on M by

$$g(u, v) = h(DF(u), DF(v)).$$

Many important Riemannian manifolds can be produced in this way, including the standard metrics on the spheres S^n (induced by the standard embedding in $\mathbb{R}^{n+1}$), and on cylinders.

9.5 Left-invariant metrics

Let G be a Lie group, and choose an inner product h on $T_eG \simeq \mathfrak{g}$. This can be extended to give a unique *left-invariant* Riemannian metric on G, by defining

$$\langle u, v\rangle_g = h\left((D_e l_g)^{-1}(u), (D_e l_g)^{-1}(v)\right).$$

Similarly, one can define right-invariant metrics; in general these are not the same.

Example 9.5.1 A metric on hyperbolic space. Recall that the hyperbolic plane $\mathbb{H}^2$ is upper half-plane, identified with the group of matrices of the form $\begin{bmatrix} y & x \\ 0 & 1 \end{bmatrix}$ for $y > 0$. If we choose an inner product at the identity $(0, 1)$ such that $(0, 1)$ and $(1, 0)$ are orthonormal, then the corresponding left-invariant Riemannian metric on $\mathbb{H}^2$ is the one for which the left-invariant vector fields $E_1 = (y, 0)$ and $E_2 = (0, y)$ are orthonormal. Thus in terms of the basis of coordinate tangent vectors $e_1 = (1, 0)$ and $e_2 = (0, 1)$, the metric has the form $g_{ij} = y^{-2}\delta_{ij}$.

Example 9.5.2 Left-invariant metrics on S^3 Recall that S^3 is the group of unit length elements of the quaternions. The tangent space at the identity is the three- dimensional space spanned by i, j, and k, and any inner product on this space gives rise to a left-invariant metric on S^3. If i, j, and k are chosen to be orthonormal, the resulting metric is the standard metric on S^3 (i.e. it agrees with the induced metric from the inclusion of S^3 in $\mathbb{R}^4$). Other choices of inner product give deformed spheres called the *Berger spheres*.

9.6 Bi-invariant metrics

A more stringent requirement on a Riemannian metric on a Lie group is that it should be invariant under both left and right translations. Such a metric is called *bi-invariant*.

Every bi-invariant metric is left-invariant, and so can be constructed in a unique way from an inner product for T_eG. This raises the question: Which inner products on T_eG give rise to bi-invariant metrics?

For any g, and any $u, v \in T_eG$, we must have

$$\langle u, v\rangle_e = \langle D_e l_g(u), D_e l_g(v)\rangle_g = \langle (D_e r_g)^{-1} D_e l_g(u), (D_e r_g)^{-1} D_e l_g(v)\rangle_e.$$

The maps $\mathrm{Ad}_g = (D_e r_g)^{-1} D_e l_g : T_eG \to T_eG$ are isomorphisms of T_eG for each $g \in G$, and give a representation of G on the vector space T_eG, since $\mathrm{Ad}_g \mathrm{Ad}_h = \mathrm{Ad}_{gh}$ for all g and h in G. Then the requirement that the inner product give rise to a bi-invariant metric is the same as requiring that it be invariant under the representation Ad of G. If G is commutative, then Ad_g is the identity map for every g, so this requirement is vacuous. In some cases there may be no Ad-invariant inner product on T_eG, but it can be shown that any compact Lie group carries at least one.

Exercise 9.6.1 Show that the adjoint action of S^3 on its Lie algebra $\mathbb{R}^3$ gives a homomorphism $\rho : S^3 \to SO(3)$ (compare Exercise 5.3.2 and the remark following it). Deduce that the only bi-invariant metric on S^3 is the standard one.

9.7 Semi-Riemannian metrics

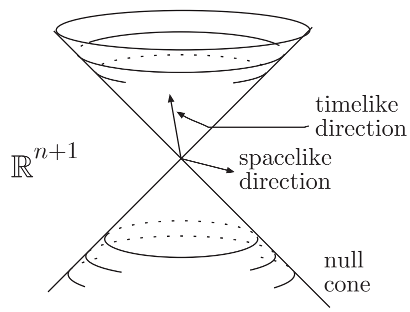
It is sometimes useful to consider a generalisation of Riemannian manifolds which drops the requirement of positivity: A semi- Riemannian manifold is a smooth manifold together with a smoothly defined symmetric bilinear form g_x on each tangent space T_xM, which is non-singular: $g_x(u, v) = 0$ for all v implies $u = 0$. If M is connected, then the signature of M is constant: One can choose a basis $\{e_1, \ldots, e_n\}$ at each point of M such that g_x takes the form

$$g_x = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & -1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & -1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & -1 \end{bmatrix}$$

with 1 occuring k times and -1 occuring $(n-k)$ times, independent of x. Of particular interest is the case of *Lorenzian* manifolds, in which $k = n-1$; when $n = 4$ these arise as spacetimes in general relativity.

In a semi-Riemannian manifold one can divide vectors into three categories: Spacelike vectors, for which $g(v, v) > 0$; timelike vectors, which have $g(v, v) < 0$; and null vectors, for which $g(v, v) = 0$. A submanifold M of a Lorenzian manifold is similarly called spacelike, timelike, or null if all of its tangent vectors are spacelike, timelike, or null respectively. If M is a spacelike submanifold of a semi-Riemannian manifold, then the induced metric makes M a Riemannian manifold.

Example 9.7.1 Minkowski space. Let $M = \mathbb{R}^{n+1}$ with variables $x^1, \ldots, x^n, t$, with the constant semi-Riemannian metric $g_{ij} = \mathrm{diag}(1, \ldots, 1, -1)$. The is the Minkowski space $\mathbb{R}^{n,1}$. The geometry of the Minkowski space $\mathbb{R}^{3,1}$ is the subject of special relativity. The *null cone* of directions with zero length is the right circular cone $|x|^2 = t^2$. Directions which point from the origin above the upper surface or below the lower surface of this cone are timelike, while those that point between the two surfaces are spacelike. In particular, a hypersurface given by the graph of a function u over $\mathbb{R}^n \times \{0\}$ is spacelike if the gradient of u is everywhere less than 1 in magnitude.



Example 9.7.2 Hyperbolic space. The n-dimensional hyperbolic space arises as a spacelike hypersurface in the Minkowski space $\mathbb{R}^{n,1}$: Let $\mathbb{H}^n = \{(x,t) \in \mathbb{R}^{n,1} : \; t = \sqrt{1+|x|^2}\}$. Thus $\mathbb{H}^n$ is one component of the set of vectors of length -1 in $\mathbb{R}^{n,1}$; thus it is an analogue of the sphere in this setting. We can also describe $\mathbb{H}^n$ as the open unit ball in $\mathbb{R}^n$ with a particular Riemannian metric: We identify $\mathbb{H}^n$ with the unit ball by stereographic projection – given a point $z \in \mathbb{H}^n$, we consider the line from z to the point $(0,-1)$, and let $\xi(z)$ be the point of intersection of this line with the plane $\mathbb{R}^n \times \{0\}$. with the unique point on the line from the origin to z which has $t = 1$. As z ranges over $\mathbb{H}^n$, $\xi(z)$ ranges over the unit ball in $\mathbb{R}^n \simeq \mathbb{R}^n \times \{0\}$. This should be though of as being analogous to stereographic projection from the north pole in the usual sphere; here we are doing stereographic projection from the point $(0,-1)$. Explicitly, this gives $z = \left(\frac{2x}{1-|x|^2}, \frac{1+|x|^2}{1-|x|^2}\right)$.

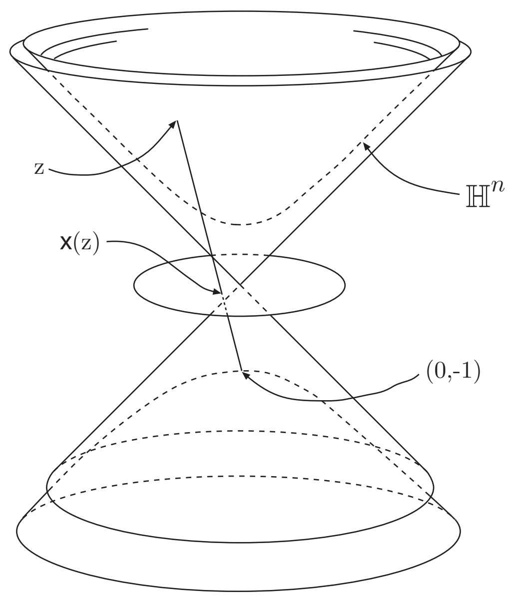
Exercise 9.7.3 Show that the induced Riemannian metric on the unit ball in $\mathbb{R}^n$ by the stereographic projection described above is

$$g_{ij} = \frac{4}{(1-|x|^2)^2}\delta_{ij}.$$

Exercise 9.7.4 Show that the map $(x,y) \mapsto \left(\frac{2x}{|x|^2+(y-1)^2}, \frac{1-|x|^2-y^2}{|x|^2+(y-1)^2}\right)$ for $(x,y) \in \mathbb{R}^{n-1} \times \mathbb{R} \simeq \mathbb{R}^n$ diffeomorphically maps the open unit ball to the upper half space, and induces from the metric on the unit ball in exercise 9.7.3 the following metric on the upper half-space:

$$g_{ij} = \frac{1}{y^2}\delta_{ij}.$$

In the case $n = 2$, show that this is the same as the left-invariant metric on $\mathbb{H}^2$ from example 9.5.1.



Lecture 10. The Levi-Civita connection

In this lecture we will show that a Riemannian metric on a smooth manifold induces a unique connection.

10.1 Compatibility of a connection with the metric

Let M be a smooth Riemannian manifold with metric g. A connection ∇ on M is said to be *compatible with the metric* on M if for every pair of vector fields X and Y on M, and every vector $v \in T_xM$,

$$v\left(g(X,Y)\right) = g\left(\nabla_v X, Y\right) + g\left(X, \nabla_v Y\right).$$

Here on the left hand side we are applying the vector v (as a derivation) to the smooth function $x \mapsto g_x(X_x, Y_x)$. This is something which is well-defined without reference to any connection. The right-hand side does depend on the connection.

Example 10.1.1 Differentiation of vector fields on a vector space is compatible with any inner product on the vector space: We have

$$D_v\left(X \cdot Y\right) = D_v X \cdot Y + X \cdot D_v Y.$$

Compatibility of the connection with the metric can be expressed in terms of parallel transport: Suppose γ is a smooth curve in M, and E_1 and E_2 are smooth vector fields along γ. Then

$$\frac{d}{dt} g(E_1, E_2) = g(\nabla_t E_1, E_2) + g(E_1, \nabla_t E_2).$$

In the special case where E_1 and E_2 are parallel along γ, this implies that $\frac{d}{dt} g(E_1, E_2) = 0$. Therefore vector fields which are parallel with respect to a compatible connection have constant length (take $E_1 = E_2$) and make a constant angle to each other. A particular consequence is that geodesics have tangent vectors of constant length.

Exercise 10.1.2 Show that any connection for which the lengths and angles between parallel vector fields are constant must be compatible with the metric.

In particular it is very easy to parallel transport a vector along a geodesic for a compatible connection on a two-dimensional manifold, since there is a unique vector at each point which makes the same angle with the tangent vector of the geodesic.

10.2 Submanifolds

Let M be a submanifold of Euclidean space $\mathbb{R}^N$, and induce on M the submanifold connection (given by projecting the derivatives of vector fields onto the tangent space of M) and the submanifold metric (where the lengths of tangent vectors to M are given by the lengths of their image in $\mathbb{R}^N$ under the inclusion map).

Proposition 10.2.1 *The submanifold connection ∇ is compatible with the induced metric g.*

Proof. We compute directly:

$$\begin{aligned} vg(X,Y) &= v\langle X^\alpha e_\alpha, Y^\beta e_\beta\rangle \\ &= \langle (D_v X^\alpha)\, e_\alpha, Y^\beta e_\beta\rangle + \langle X^\alpha e_\alpha, \left(D_v Y^\beta\right) e_\beta\rangle \\ &= \langle \pi\,(D_v X^\alpha)\, e_\alpha, Y^\beta e_\beta\rangle + \langle X^\alpha e_\alpha, \pi\left(D_v Y^\beta\right) e_\beta\rangle \\ &= g\,(\nabla_v X, Y) + g\,(X, \nabla_v Y)\,. \end{aligned}$$

□

10.3 The Levi-Civita Theorem

Proposition 10.3.1 *Let M be a smooth Riemannian manifold with metric g. Then there exists a unique connection ∇ on M whicj is symmetric and compatible with g.*

The connection given by this proposition is called the Levi-Civita connection, or sometimes the Riemannian connection. Note that the Levi-Civita connection on a submanifold of Euclidean space (with the metric induced by the standard inner product) is just the submanifold connection.

Proof. First we show uniqueness: Let X, Y, and Z be three smooth vector fields on M. Then we must have the symmetry conditions

$$\begin{aligned}\nabla_X Y - \nabla_Y X &= [X, Y];\\ \nabla_Y Z - \nabla_Z Y &= [Y, Z];\\ \nabla_Z X - \nabla_X Z &= [Z, X],\end{aligned}$$

and the compatibility conditions

$$\begin{aligned}g(\nabla_X Y, Z) + g(Y, \nabla_X Z) &= Xg(Y, Z);\\ g(\nabla_Y Z, X) + g(Z, \nabla_Y X) &= Yg(Z, X);\\ g(\nabla_Z X, Y) + g(X, \nabla_Z Y) &= Zg(X, Y).\end{aligned}$$

Take the sum of the first two of the latter equations, and subtract the third. Then apply the symmetry conditions, yielding:

$$\begin{aligned}2g(\nabla_X Y, Z) = {} & Xg(Y, Z) + Yg(Z, X) - Zg(X, Y)\\ & + g(Z, [X, Y]) + g(Y, [Z, X]) + g(X, [Z, Y]).\end{aligned}$$

This determines the inner product of $\nabla_x Y$ with any vector field Z purely in terms of the metric, and so implicitly determines $\nabla_x Y$. This completes the proof of uniqueness. To prove existence, it is only necessary to check that the formula above does indeed define a connection with the desired properties.

Denote the right-hand side of the formula by $C(X, Y, Z)$. C is C^∞ in X and in Z:

$$\begin{aligned}C(fX, Y, hZ) &= fX\,(hg(Y, Z)) + Y\,(fhg(Z, X)) - hZ\,(fg(X, Y))\\ &\quad + hg(Z, [fX, Y]) + g(Y, [hZ, fX]) + fg(X, [hZ, Y])\\ &= fhXg(Y, Z) + fhYg(Z, X) - fhZg(X, Y)\\ &\quad + f(Xh)g(Y, Z) + fY(h)g(Z, X)\\ &\quad + hY(f)g(Z, X) - hZ(f)g(X, Y)\\ &\quad + hfg(Z, [X, Y]) + hfg(Y, [Z, X]) + fhg(X, [Z, Y])\\ &\quad - hY(f)g(Z, X) - fX(h)g(Y, Z)\\ &\quad + hZ(f)g(Y, X) - fY(h)g(X, Z)\\ &= fhC(X, Y, Z)\end{aligned}$$

It follows that the map $x \mapsto C(X, Y, Z)_x$ depends only on Y and the values X_x and Z_x of the vector fields X and Z at x In the second slot we have instead:

$$\begin{aligned}C(X, fY, Z) &= fC(X, Y, Z) + X(f)g(Y, Z) - Z(f)g(X, Y)\\ &\quad + X(f)g(Z, Y) + Z(f)g(X, Y)\\ &= fC(X, Y, Z) + 2X(f)g(Y, Z).\end{aligned}$$

Therefore if we define $\nabla_X Y$ by requiring that $g(\nabla_X Y, Z) = \frac{1}{2}C(X, Y, Z)$, then $(\nabla_X Y)_x$ depends only on X_x and Y, and

$$g(\nabla_X(fY), Z) = fg(\nabla_X Y, Z) + X(f)g(Y, Z) = g(f\nabla_X Y + X(f)Y, Z),$$

so ∇ satisfies the Leibniz rule and defines a connection. $\square$

In working with the Levi-Civita connection, it is often convenient to look at the formula which defines it in terms of a local coordinate tangent basis: Choose a local chart $\varphi : U \to V$ for M. Then the formula above, applied with X, Y, and Z given by coordinate tangent vector fields ∂_i, ∂_j and ∂_k, gives the following expression for the connection coefficients:

$$\Gamma_{ij}{}^{k} = \frac{1}{2} g^{kl} \left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij} \right),$$

where we denote by g^{ij} the *inverse* of the metric g_{ij}.

Exercise 10.3.1 Suppose M is a smooth manifold, and N is a smooth Riemannian manifold with metric h. Let $F : M \to N$ be an immersion. Then h induces on M a metric g. For each $x \in M$ let $\pi_x : T_F(x)N \to T_xM$ be the map given by orthogonal projection onto $D_xF(T_xM)$, followed by application of $(D_xF)^{-1}$. Define

$$\nabla_v X = \pi_x \left(\nabla^{(h)}_{DF(v)} DF(X) \right),$$

where $\nabla^{(h)}$ is the Levi-Civita connection of the metric h on N. Show that this is well-defined (i.e. makes sense even though $DF(X)$ is only defined on the image $F(M) \subset N$) and defines a connection on M, which is the Levi-Civita connection on M.

As we have already observed, the geodesics of the Levi-Civita connection have tangent vectors of constant speed, and parallel transport along any curve preserves inner products between parallel vector fields.

10.4 Left-invariant metrics

Let G be a Lie group. Then we have a connection on G defined by taking the left-invariant vector fields to be parallel. We have seen that this connection does not give a symmetric connection, and so does not give the Levi-Civita connection for a left-invariant metric. However, this connection is compatible with any left-invariant metric.

Applying the formula above in the case of a left-invariant metric, we have the following expression for the Levi-Civita connection: Let $\{E_1, \dots, E_n\}$ be left-invariant vector fields which are orthonormal for the metric. Then by assumption the inner products $g(E_i, E_j)$ are constant, and the first three terms become zero. This gives

$$\nabla_{E_i} E_j = \frac{1}{2} \sum_{k=1}^{n} \left(c_{ijk} + c_{kij} + c_{kji} \right) E_k,$$

where the *structure constants* of G are defined by

$$[E_i, E_j] = \sum_{k=1}^{n} c_{ijk} E_k.$$

In contrast the left-invariant connection would have $\nabla_{E_i} E_j = 0$.

10.5 Exponential normal coordinates

It is often convenient to use the exponential map of the Levi-Civita connection to produce a chart for a Riemannian manifold.

Proposition 10.5.1 *Let (M, g) be a smooth Riemannian manifold, and ∇ the Levi-Civita connection. Fix $x \in M$. Choose an orthonormal basis $\{e_1, \dots, e_n\}$ for $T_x M$, and define a chart φ on a neighbourhood U of x by*

$$\varphi^{-1}\left(x^1, \dots, x^n\right) = \exp_x\left(x^j e_j\right).$$

In these coordinates (called exponential normal coordinates) we have

$$g_{ij}(0) = \delta_{ij};$$

and

$$\Gamma_{ij}{}^{k}(0) = 0.$$

Proof. Recall that the derivative of the exponential map at the origin is just the identity map. Therefore

$$g_{ij}(0) = g(\partial_i, \partial_j) = g(D_0 \exp_x(e_i), D_0 \exp_x(e_j)) = \delta_{ij}$$

since the $\{e_i\}$ were chosen to be orthonormal.

We also have

$$\nabla_{a^i \partial_i}\left(a^j \partial_j\right) = 0$$

at the origin for any constants $a^1, \dots, a^n$, since $a^i \partial_i$ is the tangent vector to the geodesic through 0 in that direction. By symmetry we have at the origin

$$\begin{aligned} 0 &= \nabla_{\partial_i + \partial_j}\left(\partial_i + \partial_j\right) \\ &= \nabla_{\partial_i} \partial_i + \nabla_{\partial_j} \partial_j + \nabla_{\partial_i} \partial_j + \nabla_{\partial_j} \partial_i \\ &= 2 \nabla_{\partial_i} \partial_j. \end{aligned}$$

for every i and j. Therefore all the connection coefficients vanish at the origin. $\square$

Remark. By the same proof, the connection coefficients at the origin vanish with respect to exponential coordinates for a connection (not necessarily a Levi-Civita connection) if and only if the connection is symmetric. This gives another interpretation of the torsion of a connection.

Lecture 11. Geodesics and completeness

In this lecture we will investigate further the metric properties of geodesics of the Levi-Civita connection, and use this to characterise completeness of a Riemannian manifold in terms of the exponential map.

11.1 Geodesic polar coordinates and the Gauss Lemma

Let (M, g) be a Riemannian manifold, and $x \in M$. Choose an orthonormal basis $\{e_1, \ldots, e_n\}$ for T_xM, and induce an isomorphism from $\mathbb{R}^n$ to T_xM. This in turn induces a map from $S^{n-1} \times (0, \infty) \to M$, by sending (ω, r) to $\exp_x(r\omega)$. This is called *geodesic polar coordinates from x*.

Proposition 11.1.1 *In geodesic polar coordinates the metric takes the form*

$$\begin{aligned} g(\partial_r, \partial_r) &= 1; \\ g(\partial_r, u) &= 0, \end{aligned}$$

for any $u \in TS^{n-1}$.

In other words, the image under the exponential map of the unit radial vector in T_xM is always a unit vector, and the image of a vector tangent to a sphere about the origin is always orthogonal to the image of the radial vector. Think of a polar 'grid' of radial lines and spheres in T_xM, mapped onto M by the exponential map T_xM. Then this map says that the images of the spheres are everywhere orthogonal to the images of the radial lines (which are of course geodesics).

Proposition 11.1.1 is often called the Gauss Lemma.

Proof. The first part of the claim follows from the fact that ∂_r is the tangent vector to a geodesic, so $\nabla_r \partial_r = 0$. Hence by compatibility,

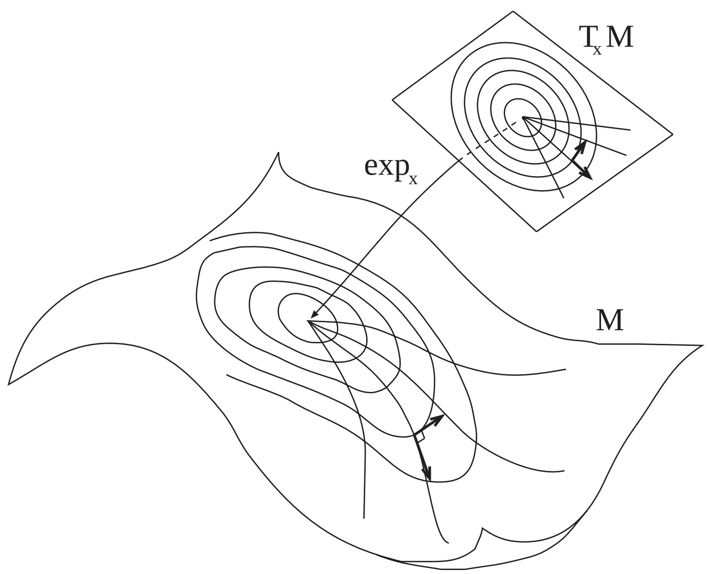
$$\partial_r g(\partial_r, \partial_r) = 2g(\partial_r, \nabla_r \partial_r) = 0,$$

and the length of ∂_r is constant in the r direction. But when $r = 0$ we have $|\partial_r| = 1$.

To prove the second part of the proposition, let $\gamma : [0, r] \times (-\varepsilon, \varepsilon) \to M$ be given by $\gamma(r, s) = \exp_x(r\omega(s))$, where $\omega(s) \in S^{n-1}(0) \subset T_xM$. Then in T_xM, $\omega'(s)$ is tangent to a sphere about the origin, and we need to show that the image ∂_s of this vector is orthogonal to the radial vector ∂_r:

$$\begin{aligned}
\partial_r g(\partial_r, \partial_s) &= g(\nabla_r \partial_r, \partial_s) + g(\partial_r, \nabla_r \partial_s) \\
&= g(\partial_r, \nabla_s \partial_r) \\
&= \frac{1}{2}\partial_s g(\partial_r, \partial_r) \\
&= 0,
\end{aligned}$$

since we know $g(\partial_r, \partial_r) = 1$ everywhere. But as before, we have $g(\partial_r, \partial_s) = 0$ when $r = 0$. $\square$

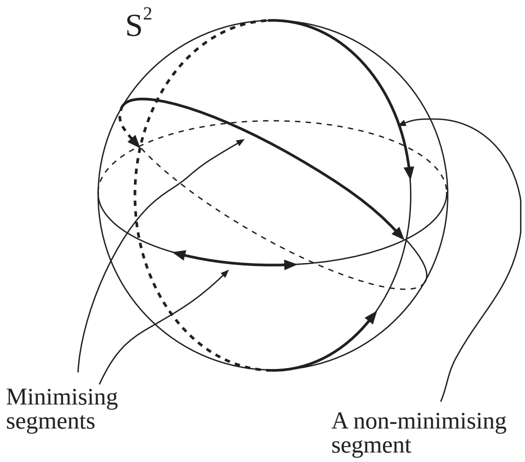


11.2 Minimising properties of geodesics

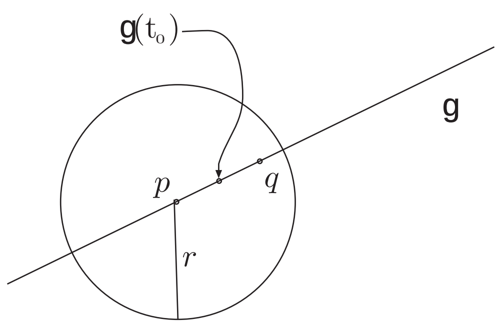
An important consequence of the Gauss Lemma is the fact that geodesics of the Levi-Civita connection, restricted to sufficiently short intervals, have smaller length than any other path between their endpoints.

Proposition 11.2.1 *Let (M, g) be a Riemannian manifold, and ∇ the Levi-Civita connection of g. Let $\gamma : I \to M$ be a ∇-geodesic. Then for any $t \in I$ there exists $\delta > 0$ such that* $L\left[\gamma\big|_{[t_0-\delta, t_0+\delta]}\right] = d(\gamma(t_0 - \delta), \gamma(t_0 + \delta))$.

Note that we cannot expect that geodesics minimise length on long intervals – consider the example of the sphere S^2: Geodesics are great circles, and these achieve the distance between their endpoints on intervals of length no greater then π, but not on longer intervals.



Proof. Choose ε sufficiently small so that $\exp_{\gamma(t_0-\varepsilon)}$ is a diffeomorphism on a ball of radius $r > 2\varepsilon$ about the origin. For convenience we denote by p the point $\gamma(t_0 - \varepsilon)$, and by q the point $\gamma(t_0 + \varepsilon)$.



Now let σ be any other curve joining the points p and q. Suppose first that σ remains in the set $B_r(p) = \exp_p(B_r(0))$. Then we can write $\sigma(t) = r(t)\omega(t)$ where $r > 0$ and $|\omega(t)| = 1$. The squared length of the tangent vector σ' is then

$$\begin{aligned} |\sigma'(t)|^2 &= (r')^2 g(\partial_r, \partial_r) + 2rr'g(\partial_r, \omega') + r^2 g(\omega', \omega') \\ &= (r')^2 + r^2|\omega'|^2 \\ &\geq (r')^2, \end{aligned}$$

with equality if and only if ω' vanishes. Therefore we have

$$L[\sigma] \geq \int |r'| dt \geq 2\varepsilon,$$

with equality if and only if ω is constant and r is monotone, in which case σ is simply a reparametrisation on γ.

The other possibility is that σ leaves the ball. But then the same argument applies on the portion of sigma joining p to the boundary of the ball, giving $L[\sigma] \geq r > 2\varepsilon$.

Therefore $L[\sigma] \geq L[\gamma]$, with equality if and only if σ is a reparametrisation of σ. □

Proposition 11.2.2 *Suppose $\gamma : [0,1] \to M$ is a piecewise smooth path for which $L[\gamma] = d(\gamma(0), \gamma(1))$. Then $\gamma = \sigma \circ f$, where $f : [0,1] \to [0,1]$ is a piecewise smooth monotone increasing function, and σ is a geodesic.*

Proof. First observe that γ achieves the distance between any pair of its points: If there were some subinterval on which this were not true, then replacing γ by a shorter path on that subinterval would also yield a shorter path from $\gamma(0)$ to $\gamma(1)$.

For any $t \in (0,1)$, there is a sufficiently small neighbourhood J of t in $[0,1]$ such that $\gamma\big|_J$ is contained in a diffeomorphic image of the exponential map from one of its endpoints, and so by the Gauss Lemma $\gamma\big|_J$ is a reparametrised geodesic. □

This is a nice feature of the Levi-Civita connection: The geodesics of ∇ are precisely the length-minimising paths, re-parametrised to have constant speed.

11.3 Convex neighbourhoods

We know that points which are sufficiently close to each other can be connected by a unique 'short' geodesic. This result can be stated somewhat more cleany in the special case of the Levi-Civita connection than it can in the general case:

Proposition 11.3.1 *Let (M, g) be a Riemannian manifold, and ∇ the Levi-Civita connection of g. Then for every $p \in M$ there exist constants $0 < \varepsilon \leq \eta$ such that for every pair of points q and r in $B_\varepsilon(p)$ there exists a unique geodesic γ_{qr} of length $L[\gamma_{qr}] < \eta$ such that $\gamma_{qr}(0) = q$ and $\gamma_{qr}(1) = r$. Furthermore, $L[\gamma_{qr}] = d(q, r)$.*

Proof. The idea is exactly the same as the proof of Proposition 8.11.5: The map $\tilde{\exp}$ is a local diffeomorphism from a neighbourhood of $(p, 0)$ in TM to a neighbourhood of (p, p) in $M \times M$. Choose η sufficiently small to ensure that

$\tilde{\exp}$ is a diffeomorphism on $\mathcal{O} = \{(q,v)|\ q \in B_\eta,\ |v| \leq \eta\}$, and then choose ε sufficiently small to ensure that $B_\varepsilon(p) \times B_\varepsilon(p) \subset \tilde{\exp}\mathcal{O}$.

This gives the existence of a geodesic γ_{qr} of length less than η joining q to r. Also γ_{qr} achieves the distance between its endpoints, by the Gauss Lemma. □

We will improve the result slightly:

Proposition 11.3.2 *For any $p \in M$ there exists a constant $\varepsilon > 0$ such that for every pair of points q and r in $B_\varepsilon(p)$, there exists a unique geodesic γ_{qr} for which $\gamma(0) = q$, $\gamma(1) = r$, and $d(p,\gamma(t)) \leq \max\{d(p,q), d(p,r)\}$ for all $t \in [0,1]$. Furthermore $L[\gamma_{qr}] = d(q,r)$.*

Proof. By Proposition 11.5.1, ${\mathit{\Gamma}_{ij}}^k(0) = 0$ in exponential normal coordinates. Therefore there is $\eta > 0$ such that $\left|\sum_{k,i,j} x^k {\mathit{\Gamma}_{ij}}^k \xi^i \xi^j\right| < \frac{1}{2}\sum_k \left(\xi^k\right)^2$ provided $\sum_k \left(x^k\right)^2 < 2\eta$, and such that for some $\varepsilon \in (0,\eta)$ the conditions of Proposition 11.3.1 hold.

Then by Proposition 11.3.1 there is a geodesic γ_{qr} of length less than η joining q to r; since p and q are within distance ε of p, the entire geodesic γ_{qr} stays within distance 2η of p.

Now along γ_{qr} we compute:

$$\begin{aligned}
\frac{d^2}{dt^2} d(p,\gamma(t))^2 &= \frac{d^2}{dt^2}\sum_k \left(x^k\right)^2 \\
&= 2\frac{d}{dt}\sum_k \left(x^k \dot{x}^k\right) \\
&= 2\sum_k \left(\dot{x}^k\right)^2 + 2\sum_k x^k \frac{d^2}{dt^2} x^k \\
&= 2\sum_k \left(\dot{x}^k\right)^2 - 2\sum_k x^k {\mathit{\Gamma}_{ij}}^k \dot{x}^i \dot{x}^j \\
&\geq 2\left(\sum_k \left(\dot{x}^k\right)^2 - \frac{1}{2}\sum_{[} \left(\dot{x}^k\right)^2\right) \\
&\geq 0,
\end{aligned}$$

since we are within the ball of radius 2η about p. Therefore $d(p,\gamma(t))^2$ is a convex function along γ_{qr}, so the maximum value is attained at the endpoints. □

11.5 Completeness and the Hopf-Rinow Theorem

Now we are ready to prove the main result of this section:

Theorem 11.5.1 The Hopf-Rinow Theorem. *Let (M, g) be a connected Riemannian manifold. The following are equivalent:*

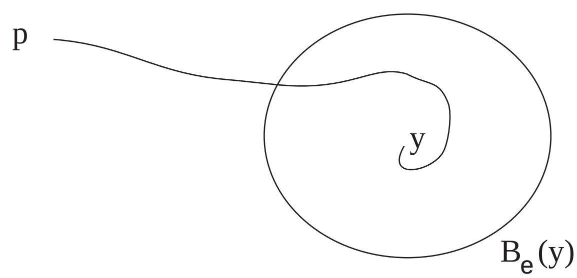
(1). M is a complete metric space with the distance function d;
(2). $\exp$ is defined on all of TM (i.e. all geodesics can be extended indefinitely);
(3). There exists $p \in M$ for which $\exp_p$ is definied on all of T_pM.

Furthermore, each of these conditions implies

(*). For every p and q in M, there exists a geodesic $\gamma : [0, 1] \to M$ for which $\gamma(0) = p$, $\gamma(1) = q$, and $L[\gamma] = d(p, q)$.

Proof. (2) $\Longrightarrow$ (3) trivially. We prove (1) $\Longrightarrow$ (2), and then (3) $\Longrightarrow$ $(*_p)$ and ((3) and $(*_p)$) $\Longrightarrow$ (1), where $(*_p)$ is the statement that every point $q \in M$ can be connected to p by a length-minimising geodesic.

Suppose (M, d) is a complete metric space. If (2) does not hold, then there is $p \in M$, $v \in T_pM$ with $|v| = 1$, and $T < \infty$ such that $\gamma(t) = \exp_p(tv)$ exists for $t < T$ but not for $t = T$.



But $d(\gamma(s), \gamma(t)) \le |s - t|$, so $\gamma(t)$ is Cauchy as t approaches T. By completeness, $\gamma(t)$ converges to a limit $y \in M$ as $t \to T$. Choose $\varepsilon > 0$ such that the ball $B_\varepsilon(y)$ is convex in the sense of Proposition 10.3.2. Then for $s, t > T - \varepsilon$ we have $\gamma(s)$ and $\gamma(t)$ contained in this geodesically convex set, and γ a geodesic joining them; therefore γ achieves the distance between $\gamma(s)$ and $\gamma(t)$:

$$d(\gamma(s), \gamma(t)) = |t - s|, \qquad \text{for } t, s > T - \varepsilon.$$

Then we have by continuity of the distance function,

$$d(\gamma(t), y) = \lim_{s \to T} d(\gamma(t), \gamma(s)) = \lim_{s \to T} (s - t) = T - t = L[\gamma\big|_{(t,T)}].$$

Therefore $\gamma|_{(t,T)}$ is a minimising path, and must be a radial geodesic from y:

$$\gamma(t) = \exp_y(w(T-t)).$$

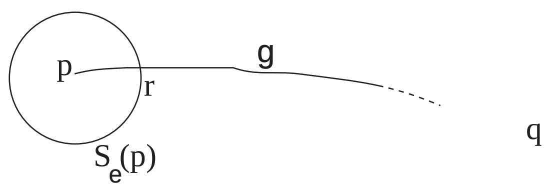
But then γ can be extended beyond y by taking $t > T$ in this formula. This is a contradiction, so condition (2) must hold.

Next we prove (3) $\implies$ $(*_p)$. Let $p \in M$ be such that (3) holds, and $q \in M$ any other point. Choose $\varepsilon > 0$ such that $\exp_p$ is a diffeomorphism on a set containing the closure of $B_\varepsilon(0)$. If $q \in B_\varepsilon(p)$ then we are done, so assume not. Then $S_\varepsilon(p) = \exp_p(S_\varepsilon(0))$ is the image of a compact set under a continuous map, and so the function $d(.,q)$ attains its minimum on this set. In other words, there exists $r = \exp_p(\varepsilon v)$ such that $d(r,q) = d(S_\varepsilon(p),q) = \inf\{d(r',q) :\ r' \in S_\varepsilon(p)\}$. Then we must have

$$d(p,q) = \varepsilon + d(r,q).$$

The inequality $d(p,q) \leq \varepsilon + d(r,q)$ follows by the triangle inequality since $d(p,r) = \varepsilon$, and the other inequality follows because any path from p to q must pass through $S_\varepsilon(p)$.

Define $\gamma(t) = \exp_p(tv)$.



Let $A \subset [0, d(p,q)]$ be the set of t for which

$$d(p,q) = t + d(\gamma(t), q).$$

A is non-empty, as we have just shown; it is closed, by continuity of the distance function and because γ can be extended indefinitely. We also have that $t \in A$ implies $[0,t] \subset A$: If $s < t$ we have by the triangle inequality and the fact that $\gamma|_{[0,t]}$ is a minimising path

$$\begin{aligned} d(p,q) &= t + d(\gamma(t),q) \\ &= s + (t-s) + d(\gamma(t),q) \\ &\geq s + d(\gamma(s),q) \\ &\geq d(p,q), \end{aligned}$$

so equality must hold throughout, and $s \in A$.

Finally, we will prove that A is open in $[0, d(p,q)]$. This will imply $A = [0, d(p,q)]$, so in particular $d(\gamma(d(p,q)),q) = 0$, and $\gamma(d(p,q)) = q$.

Suppose $T \in A$, and write $p' = \gamma(T)$. Choose $\delta > 0$ such that $\exp_{p'}$ is a diffeomorphism on $B_\delta(0)$. If $q \in B_\delta(p')$, then write $q = \exp_{p'}(d(p',q)w)$, and define

$$\sigma(t) = \begin{cases} \gamma(t), & \text{for } t \leq T; \\ \exp_{p'}((t-T)w) & \text{for } t \geq T. \end{cases}$$

Then σ is a curve joining p to q, and by assumption

$$d(p,q) = T + d(p',q) = L[\sigma].$$

Therefore σ is a minimising curve, hence a geodesic, so $\sigma(t) = \exp_p(t)$ for all $t \in [0, d(p,q)]$ and we are done.

Otherwise, $q \notin B_\delta(p')$, and we can choose $r' \in S_\delta(p')$ such that $d(r',q) = d(B_\delta(p'),q)$. Then as before, we have

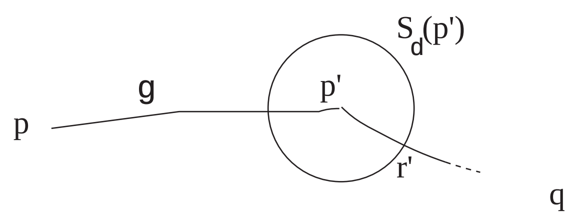
$$d(p',q) = \delta + d(r',q),$$

and so by assumption

$$d(p,q) = d(p,p') + d(p',q) = T + \delta + d(r',q).$$

Let σ be the unit speed curve given by following γ from p to p', then following the radial geodesic from p' to r'. Then $L[\sigma] = d(p,p') + \delta$. Therefore,

$$d(p,q) = L[\sigma] + d(r',q).$$



It follows that $L[\sigma] = d(p,r')$, since if there were any shorter path σ' from p to r', we would have

$$d(p,q) \leq L[\sigma'] + d(r',q) < L[\sigma'] + d(r',q) = d(p,q)$$

which is a contradiction. Therefore σ is a geodesic, and $\sigma(t) = \gamma(t)$; and $[0, T+\delta] \subset A$. Therefore A is open as claimed, and we have proved $(*_p)$.

Now we complete the proof by showing that (3) and $(*_p)$ together imply condition (1). For p satisfying (3) and $(*_p)$, let $M_k = \exp_p\left(\overline{B_k(0)}\right)$. This is the image of a compact set under a continuous map, and so is compact. By $(*_p)$ we have $\cup_{k=1}^{\infty} M_k = \infty$.

Suppose $\{z_i\}$ is a Cauchy sequence in M. Then in particular $\{z_i\}$ is contained in some M_k, and hence converges by the compactness of M_k.

This completes the proof of the Theorem. $\square$

Lecture 12. Tensors

In this lecture we define tensors on a manifold, and the associated bundles, and operations on tensors.

12.1 Basic definitions

We have already seen several examples of the idea we are about to introduce, namely linear (or multilinear) operators acting on vectors on M.

For example, the metric is a bilinear operator which takes two vectors to give a real number, i.e. $g_x : T_xM \times T_xM \to \mathbb{R}$ for each x is defined by $u, v \mapsto g_x(u, v)$.

The difference between two connections $\nabla^{(1)}$ and $\nabla^{(2)}$ is a bilinear operator which takes two vectors and gives a vector, i.e. a bilinear operator $S_x : T_xM \times T_xM \to T_xM$ for each $x \in M$. Similarly, the torsion of a connection has this form.

Definition 12.1.1 A **covector** ω at $x \in M$ is a linear map from T_xM to $\mathbb{R}$. The set of covectors at x forms an n-dimensional vector space, which we denote T_x^*M. A **tensor** of type (k, l) at x is a multilinear map which takes k vectors and l covectors and gives a real number

$$T_x : \underbrace{T_xM \times \ldots \times T_xM}_{k \text{ times}} \times \underbrace{T_x^*M \times \ldots \times T_x^*M}_{l \text{ times}} \to \mathbb{R}.$$

Note that a covector is just a tensor of type $(1, 0)$, and a vector is a tensor of type $(0, 1)$, since a vector v acts linearly on a covector ω by $v(\omega) := \omega(v)$.

Multilinearity means that

$$\begin{aligned} &T\left(\sum_{i_1} c^{i_1} v_{i_1}, \ldots, \sum_{i_k} c^{i_k} v_{i_k}, \sum_{j_1} a_{j_1}\omega^{j_1}, \ldots, \sum_{j_l} a_{j_l}\omega^{j_l}\right) \\ &\quad = \sum_{i_1,\ldots,i_k,j_1,\ldots,j_l} c^{i_1} \ldots c^{i_k} a_{j_1} \ldots a_{j_l} T(v_{i_1}, \ldots, v_{i_k}, \omega^{j_1}, \ldots, \omega^{j_l}) \end{aligned}.$$

To elucidate this definition, let us work in a chart, so that we have a basis of coordinate tangent vector fields $\partial_1, \ldots, \partial_n$ at each point. Then we can define a convenient basis for the cotangent space T_x^*M by defining

$$dx^i : \; T_xM \to \mathbb{R}$$

by

$$dx^i(\partial_j) = \delta_j^i.$$

The notation here comes from the fact that dx^i is the differential of the smooth function x^i, in other words

$$dx^i(v) = v(x^i).$$

Similarly, for any $f \in C^\infty(M)$ we have a covector $d_xf \in T_x^*M$ defined by

$$df(v) = v(f).$$

12.2 Tensor products

Definition 12.2.1 Let T and S be two tensors at x of types (k, l) and (p, q) respectively. Then the **tensor product** $T \otimes S$ is the tensor at x of type $(k+p, l+q)$ defined by

$$\begin{aligned} T \otimes S(v_1, \ldots, v_{k+p}, \omega_1, \ldots, \omega_{l+q}) = T(&v_1, \ldots, v_k, \omega_1, \ldots, \omega_l) \\ &\times S(v_{k+1}, \ldots, v_{k+p}, \omega_{l+1}, \ldots, \omega_{l+q}) \end{aligned}$$

for all vectors $v_1, \ldots, v_{k+p} \in T_xM$ and all covectors $\omega_1, \ldots, \omega_{l+q} \in T_x^*M$.

Proposition 12.2.2 *The set of tensors of type (k, l) at x is a vector space of dimension n^{k+l}, with a basis given by*

$$\{dx^{i_1} \otimes \ldots \otimes dx^{i_k} \otimes \partial_{j_1} \otimes \ldots \otimes \partial_{j_l} : \; 1 \le i_1, \ldots, i_k, j_1, \ldots, j_l \le n.\}$$

Proof. The vector space structure of the space of (k, l)-tensors is clear.

Note that a tensor T is completely determined by its action on the basis vectors and covectors:

$$\begin{aligned} &T(v_1^i\partial_i, \ldots, v_k^j\partial_j, \omega_a^1 dx^a, \ldots, \omega_b^l dx^b) \\ &\quad = v_1^i \ldots v_k^j \ldots \omega_a^1 \ldots \omega_b^l T(\partial_i, \ldots, \partial_j, dx^a, \ldots, dx^b), \end{aligned}$$

by multilinearity.

Then observe that this allows T to be written as a linear combination of the proposed basis elements: Define

$$T_{i_1 \ldots i_k}{}^{j_1 \ldots j_l} = T(\partial_{i_1}, \ldots, \partial_{i_k}, dx^{j_1}, \ldots, dx^{j_l})$$

for each $i_1, \ldots, i_k, j_1, \ldots, j_l \in \{1, \ldots, n\}$. Then define $\hat{T}$ by

$$\hat{T} = \sum_{i_1,\ldots,i_k,j_1,\ldots,j_l=1}^{n} T_{i_1\ldots i_k}{}^{j_1\ldots j_l} dx^{i_1} \otimes \ldots dx^{i_k} \otimes \partial_{j_1} \otimes \ldots \partial_{j_l}.$$

We claim that $\hat{T} = T$. To check this, we compute the result of acting with $\hat{T}$ on k elementary tangent vectors and l elementary covectors:

$$\begin{aligned}
&\hat{T}(\partial_{i_1}, \ldots, \partial_{i_k}, dx^{j_1}, \ldots, dx^{j_l}) \\
&= \sum_{a_1,\ldots,a_k,b_1,\ldots,b_l=1}^{n} T_{a_1\ldots a_k}{}^{b_1\ldots b_l} dx^{a_1} \otimes \ldots dx^{a_k} \otimes \partial_{b_1} \otimes \ldots \partial_{b_l} \left(\partial_{i_1}, \ldots, \partial_{i_k}, dx^{j_1}, \ldots, dx^{j_l}\right) \\
&= \sum_{a_1,\ldots,a_k,b_1,\ldots,b_l=1}^{n} T_{a_1\ldots a_k}{}^{b_1\ldots b_l} dx^{a_1}(\partial_{i_1}) \ldots dx^{a_k}(\partial_{i_k}) \partial_{b_1}(dx^{j_1}) \ldots \partial_{b_l}(dx^{j_l}) \\
&= T_{i_1\ldots i_k}{}^{j_1\ldots j_l} \\
&= T\left(\partial_{i_1}, \ldots, \partial_{i_k}, dx^{j_1}, \ldots, dx^{j_l}\right).
\end{aligned}$$

Therefore the proposed basis does generate the entire space.

It remains to show that the basis elements are linearly independent. Suppose we have some linear combination of them which gives zero:

$$A_{i_1\ldots i_k}{}^{j_1\ldots j_l} dx^{i_1} \otimes \ldots \otimes dx^{i_k} \otimes \partial_{j_1} \otimes \ldots \otimes \partial_{j_l} = 0.$$

Then in particular the value of this linear combination on any set of vectors $\partial_{a_1}, \ldots, \partial_{a_k}$ and covectors $dx^{b_1}, \ldots, dx^{j_l}$ must give zero. But this value is exactly the coefficient:

$$A_{a_1\ldots a_k}{}^{b_1\ldots b_l} = 0.$$

This establishes the linear independence, and show that we do indeed have a basis. □

For example, we can write the metric as

$$g = g_{ij} dx^i \otimes dx^j,$$

and the torsion T of a connection as

$$T = T_{ij}{}^k dx^i \otimes dx^j \otimes \partial_k.$$

We denote the space of (k, l)-tensors at x by

$$\bigotimes^k T_x M \otimes \bigotimes^l T_x^* M.$$

12.3 Contractions

Let T be a tensor of type (k,l) at x, with k and l at least 1. Then T has components $T_{i_1,\dots,i_k}{}^{j_1,\dots j_l}$ as before. Then there is a tensor of type $(k-1,l-1)$ which has components

$$\sum_{a=1}^{n} T_{i_1\dots i_{k-1}a}{}^{j_1\dots j_{l-1}a}.$$

This tensor is called a *contraction* of T (If k and l are large then there will be many such contractions, depending on which indices we choose to sum over).

A special case is where T is a tensor of type $(1,1)$. This takes a vector and gives another vector, and so is nothing other than a linear operator on TM. There is a unique contraction in this case, which is just the *trace* of the operator:

$$\mathrm{tr}A = \sum_{i=1}^{n} A_i{}^i.$$

12.4 Musical isomorphisms

If we have an inner product on TM, then there are natural isomorphisms between tensors of types (k,l) and (m,n) whenever $k+l=m+n$. In particular, there is an isomorphism between the tangent space and the dual tangent space, given by sending a vector v to the covector ω which acts on any vector u to give the number $\langle u,v\rangle$.

It is useful to have an explicit expression for this isomorphism in terms of the standard bases for T_xM and T_x^*M: A vector ∂_i gets sent to the covector ω_i which acts on the vector ∂_j to give the number $g(\partial_i,\partial_j)=g_{ij}$. Therefore $\omega_i = \sum_{j=1}^n g_{ij}dx^j$. More generally, a vector $v^i\partial_i$ is sent to the covector $v^ig_{ij}dx^j$.

More generally, the isomorphism from (k,l) tensors to $(k+l,0)$-tensors is given as follows: A (k,l)-tensor T with coefficients $T_{i_1\dots i_k}{}^{j_1\dots j_l}$ becomes the $(k+l,0)$-tensor $\tilde{T}$ with coefficients

$$\tilde{T}_{i_1\dots i_k i_{k+1}\dots i_{k+l}} = T_{i_1\dots i_k}{}^{j_1\dots j_l} g_{j_1 i_{k+1}}\dots g_{j_l i_{k+l}}.$$

Thus the metric acts to 'lower' an index by a tensor product followed by a contraction. The inverse of this operation ('rasing indices') is given by multiplying by the inverse matrix of the metric, $(g^{-1})^{ij}$, which defines a tensor of type $(0,2)$. For convenience we will denote this by g^{ij}, suppressing the inversion. This is consistent with the fact that is we use the metric to raise its own indices, then we get the tensor

$$g^{ij} = g_{kl}(g^{-1})^{ik}(g^{-1})^{jl} = (g^{-1})^{ij}.$$

Example 12.4.1 If T is a tensor of type $(k,1)$, then it takes k vectors and gives another one, which we denote by $T(v_1,\dots,v_k)$. The index-lowering map should produce from this a map which takes $(k+1)$ vectors and gives a number, which we denote by $T(v_1,\dots,v_{k+1})$. Following the procedure above, we see that this isomorphism is just given by

$$T(v_1,\dots,v_{k+1}) = g(T(v_1,\dots,v_k),v_{k+1}).$$

12.5 Tensor fields

A *tensor field* T of type (k,l) on M is a smooth choice of a tensor T_x of type (k,l) at x for each $x \in M$. In particular a tensor field of type $(0,1)$ is just a vector field, and a tensor field ω of type $(1,0)$ is given by a covector ω_x at each point. In this case smoothness is interpreted in the sense that for every smooth vector field X on M, the function $x \mapsto \omega_x(X_x)$ is smooth. A smooth tensor field of type $(1,0)$ is also called a 1-form. The space of 1-forms is denoted $\Omega(M)$.

More generally, smoothness of a tensor field T of type (k,l) is to be interpreted in the sense that the function obtained by acting with T on k smooth vector fields and l 1-forms is a smooth function on M.

Equivalently, a (k,l)-tensor field T is smooth if the component functions ${T_{i_1\dots i_k}}^{j_1\dots j_l}$ with respect to any chart are smooth functions.

It is straightforward to check that tensor products, contractions, and index raising and powering operators go across from tensors to smooth tensor fields.

12.6 Tensor bundles

In analogy with the definition of the tangent bundle, we can define (k,l)-tensor bundles:

$$\bigotimes^k TM \otimes \bigotimes^l T^*M = \bigcup_{x\in M} \bigotimes^k T_xM \otimes \bigotimes^l T_x^*M.$$

This is made into a manifold as follows: Given a chart φ for a region U of M, we define a chart $\tilde\varphi$ for the region

$$\bigcup_{x\in U} \bigotimes^k T_xM \otimes \bigotimes^l T_x^*M$$

by taking for $p \in U$ and T a (k,l)-tensor at p,

$$\tilde\varphi(p,T) = (\varphi^1,\dots,\varphi^n, {T_{1\dots 1}}^{1\dots 1},\dots,{T_{n\dots n}}^{n\dots n}) \in \mathbb{R}^{n+n^{k+l}}.$$

If we have two such charts φ and η, the transition maps between them are given by $\eta\circ\varphi^{-1}$ on the first n components, and by the map

$$T_{i_1\dots i_k}{}^{j_1\dots j_l} \mapsto \Lambda^{a_1}_{i_1}\dots\Lambda^{a_k}_{i_k}(\Lambda^{-1})^{j_1}_{b_1}\dots(\Lambda^{-1})^{j_l}_{b_l}T_{a_1\dots a_k}{}^{b_1\dots b_l}$$

in the remaining components, where $\Lambda = D(\eta\circ\varphi^{-1})$ is the derivative matrix of the transition map on M. In particular, these transition maps are smooth, so the tensor bundle is a smooth manifold.

It is common to denote the space of smooth (k,l)-tensor fields on M by $\Gamma\left(\bigotimes^k TM\otimes\bigotimes^l T^*M\right)$, read as "sections of the (k,l)-tensor bundle".

12.7 A test for tensoriality

Often the tensors which we work with will arise in the following way: We are given an operator T which takes k vector fields and l 1-forms, and gives a smooth function. Suppose that the operator is also multilinear (over $\mathbb{R}$), so that multiplying any of the vector fields or 1-forms by a constant just changes the result by the same factor, and taking one of the vector fields or 1-forms to be given by a sum of two such, gives the sum of the results applied to the individual vector fields or 1-forms.

Proposition 12.7.1 *The operator T is a tensor field on M if and only if it is C^∞-linear in each argument:*

$$T(f_1X_1,\dots,f_kX_k,g_1\omega^1,\dots,g_l\omega^l) = f_1\dots f_kg_1\dots g_lT(X_1,\dots X_k,\omega^1,\dots,\omega^l)$$

for any smooth functions $f_1,\dots f_k,g_1,\dots,g_l$, vector fields $X_1,\dots,X_k$, and 1-forms $\omega^1,\dots,\omega^l$.

We have already seen this principle in action, in our discussion of the difference of two connections and of the torsion of a connection.

Proof. One direction is immediate: A (k,l)-tensor field is always C^∞-linear in each argument.

Conversely, suppose $T:\mathcal{X}(M)^k\times\Omega(M)^l\to C^\infty(M)$ is multilinear over $C^\infty(M)$. Choose a chart φ in a neighbourhood of a point $x\in M$. Now compute the action of T on vector fields $X_{(j)} = X^i_{(j)}\partial_i$, $j=1,\dots,k$, and 1-forms $\omega^{j)} = \omega^{(j)}_i dx^i$, $j=1,\dots,l$, evaluated at the point x:

$$\begin{aligned}
&\Big(T(X_{(1)},\dots,X_{(k)},\omega^{(1)},\dots,\omega^{(l)})\Big)_x\\
&\qquad = X^{i_1}_{(1)}(x)\dots X^{i_k}_{(k)}(x)\omega^{(1)}_{j_1}(x)\dots\omega^{(l)}_{j_l}(x)\left(T(\partial_{i_1},\dots,\partial_{i_k},dx^{j_1},\dots,dx^{j_l}\right)_x\\
&\qquad = T_x\left(X_{(1)}(x),\dots,X_{(k)}(x),\omega^{(1)}(x),\dots,\omega^{(l)}(x)\right)
\end{aligned}$$

where T_x is the (k,l)-tensor at x with coefficients given by

$$(T_x)_{i_1 \dots i_k}{}^{j_1 \dots j_l} = \left(T(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l})\right)_x.$$

□

Example 12.7.2 The Lie bracket of two vector fields takes two vector fields and gives another vector field. This is not a tensor, because

$$[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X \neq fg[X, Y]$$

in general. Similarly, the connection ∇ takes two vector fields and gives a vector field, but is not a tensor, because

$$\nabla_X(fY) = f\nabla_X Y + X(f)Y \neq f\nabla_X Y$$

in general.

12.8 Metrics on tensor bundles

A Riemannian metric on M is given by an inner product on each tangent space T_xM. This induces a natural inner product on each of the tensor spaces at x. This is uniquely defined by the requirement that tensor products of basis elements of TM are orthonormal, and that the metric by invariant under the raising and lowering operators.

Explicitly, the inner product of two tensors S and T of type (k, l) at x is given by

$$\langle S, T \rangle = g^{a_1 b_1} \dots g^{a_k b_k} g_{i_1 j_1} \dots g_{i_l j_l} S_{a_1 \dots a_k}{}^{i_1 \dots i_l} T_{b_1 \dots b_k}{}^{j_1 \dots j_l}.$$

12.9 Differentiation of tensors

We can also extend the connection (defined to give derivatives of vector fields) to give a connection on each tensor bundle – i.e. to allow the definition of derivatives for any tensor field on M.

We want the following properties for the connection: ∇ should take a tensor field T of type (k, l), and give a tensor field ∇T of type $(k+1, l)$, such that

(1). The Leibniz rule holds for tensor products: If S and T are tensor fields of type (k, l) and type (p, q) respectively, then

$$\begin{aligned}
&\nabla(S \otimes T)(X_0, X_1, \dots, X_{k+p}, \omega_1, \dots, \omega_{l+q}) \\
&= \nabla S(X_0, X_1, \dots, X_k, \omega_1, \dots, \omega_l) T(X_{k+1}, \dots, X_{k+p}, \omega_{l+1}, \dots, \omega_{l+q}) \\
&\quad + S(X_1, \dots, X_k, \omega_1, \dots, \omega_l) \nabla T(X_0, X_{k+1}, \dots, X_{k+p}, \omega_{l+1}, \dots, \omega_{l+q})
\end{aligned}$$

for any vector fields $X_0, \ldots, X_{k+p}$ and 1-forms $\omega_1, \ldots, \omega_{l+q}$.

(2). ∇ applied to a contraction of a tensor T is just the contraction of ∇T: More precisely, if T is the tensor with components $T_{i_1 \ldots i_k}{}^{j_1 \ldots j_l}$, and CT is the contracted tensor with components $CT_{i_1 \ldots i_{k-1}}{}^{j_1 \ldots j_{l-1}} = \sum_{i=1}^n T_{i_1 \ldots i_{k-1} i}{}^{j_1 \ldots j_{l-1} i}$, then

$$C\left(\nabla T\right) = \nabla\left(CT\right).$$

(3). If T is a tensor of type $(0,1)$ (i.e. a vector field) then ∇ is the same as out previous definition:

$$\nabla T(X, \omega) = (\nabla_X T)(\omega)$$

for any vector field X and 1-form ω.

(4). If f is a tensor of type $(0,0)$ (i.e. a smooth function) then ∇ is just the usual derivative:

$$\nabla f(X) = X(f).$$

Let us investigate what this means for the derivative of a 1-form: If X is a vector field and ω a 1-form, then the contraction of $X \otimes \omega$ is just the function $\omega(X)$. By conditions 1, 2, and 4, we then have

$$\nabla_v(\omega(X)) = C\left((\nabla_v \omega) \otimes X + \omega \otimes \nabla_v X\right).$$

Taking the special case $X = \partial_k$ and $v = \partial_i$, we get $\omega(X) = \omega_k$, where $\omega = \omega_k dx^k$, and so

$$\partial_i \omega_k = (\nabla_i \omega_k) + \Gamma_{ik}{}^j \omega_j$$

where I used condition (3) to get $\nabla_i \partial_k = \Gamma_{ik}{}^j \partial_j$. This gives

$$\nabla_i \omega_j = \partial_i \omega_j - \Gamma_{ij}{}^k \omega_k.$$

We know already that the derivative of a vector field is given by

$$\nabla_i X^j = \partial_i X^j + \Gamma_{ik}{}^j X^k.$$

Using these two rules and condition (1), and the fact that any tensor can be expressed as a linear combination of tensor products of vector fields and 1-forms, we get for a tensor T of type (k,l),

$$\begin{aligned}
\nabla_i T_{i_1 \ldots i_k}{}^{j_1 \ldots j_l} = \partial_i T_{i_1 \ldots i_k}{}^{j_1 \ldots j_l} \\
&- \Gamma_{ii_1}{}^p T_{pi_2 \ldots i_k}{}^{j_1 \ldots j_l} - \ldots - \Gamma_{ii_k}{}^p T_{i_1 \ldots i_{k-1} p}{}^{j_1 \ldots j_l} \\
&+ \Gamma_{ip}{}^{j_1} T_{i_1 \ldots i_k}{}^{p j_2 \ldots j_l} + \ldots + \Gamma_{ip}{}^{j_l} T_{i_1 \ldots i_k}{}^{j_1 \ldots j_{l-1} p}.
\end{aligned}$$

In particular, we note that the compatibility of the connection with the metric implies that the derivative of the connection is always zero:

$$\nabla_i g_{jk} = \partial_i g_{jk} - \Gamma_{ij}{}^p g_{pk} - \Gamma_{ik}{}^p g_{jp} = \Gamma_{ijk} + \Gamma_{ikj} - \Gamma_{ijk} - \Gamma_{ikj} = 0.$$

This implies in particular that the covarant derivative commutes with the index-raising and lowering operators.

12.10 Parallel transport

As with the case of vector fields, the connection allows us to define derivatives of tensors aloong curves, and parallel transport of tensors along curves (by solving the first order linear ODE system corresponding to $\nabla_t T = 0$. As in the case of vector fields, parallel transport preserves inner products between tensors.

12.11 Computing derivatives of tensors

Since the covariant derivative of a tensor is again a tensor, it is often convenient to use the fact that it is determined by its components with respect to any coordinate tangent basis. In particular, to compute the covariant derivative of a tensor at a point x, one can work in local coordinates in which the connection coefficients ${\Gamma_{ij}}^k$ vanish at the point x (for example, use exponential normal coordinates from x).

Then we have

$$\left(\nabla_i {T_{i_1\dots i_k}}^{j_1\dots j_l}\right)_x = \partial_i {T_{i_1\dots i_k}}^{j_1\dots j_l}(x).$$

One must take care, however, in computing second derivatives of tensors: The expression obtained using geodesic coordinates at x is simple at the point x itself, but not at neighbouring points; thus it cannot be used to get simple expressions for second derivatives.

Lecture 13. Differential forms

In the last few lectures we have seen how a connection can be used to differentiate tensors, and how the introduction of a Riemannian metric gives a canonical choice of connection. Before exploring the properties of Riemannian spaces more thoroughly, we will first look at a special class of tensors for which there is a notion of differentiation that makes sense even without a connection or a metric. These are called differential forms, and they play an extremely important role in differential geometry.

13.1 Alternating tensors

We will first look a little more at the linear algebra of tensors at a point. We will consider a natural subspace of the space of k-tensors, namely the **alternating tensors**.

Definition 13.1.1 A k-tensor $\omega \in \otimes^k T_x^* M$ is **alternating** if it is antisymmetric under interchange of any two of its arguments. Equivalently, for any k vectors $v_1, \ldots, v_k \in T_x M$, and any permutation $\sigma \in S_k$,

$$\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \mathrm{sgn}\sigma \omega(v_1, \ldots, v_k),$$

where $\mathrm{sgn}\sigma = 1$ if σ is an even permutation, $\mathrm{sgn}\sigma = -1$ is σ is an odd permutation.

The space of alternating k-tensors at x is denoted $\Lambda^k T^* M$. Note that $\Lambda^1 T_x^* M = T_x^* M$, so alternating 1-tensors are just covectors.

There is a natural projection $\mathcal{A} : \otimes^k T_x^* M \to \Lambda^k T_x^* M$ defined as follows:

$$\mathcal{A}T(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}\sigma T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

Then T is alternating if and only if $\mathcal{A}T = T$.

Example 13.1.1 The geometric meaning of this definition is probably not clear at this stage. An illustrative example is the following: Choose an orthonormal basis $\{\phi^1, \ldots, \phi^n\}$ for T_x^*M. Then we can define an alternating n-tensor A by taking

$$A(v_1, \ldots, v_n) = \det[\phi^i(v_j)].$$

This is antisymmetric since the determinant is antisymmetric under interchange of any pair of columns. Geometrically, the result $A(v_1, \ldots, v_n)$ is the (signed) n-dimensional volume of the parallellopiped generated by $v_1, \ldots, v_n$.

Given a basis $\{\partial_1, \ldots, \partial_n\}$ for T_xM, we can define a natural basis for $\Lambda^k T_x^*M$: For each k-tuple $i_1, \ldots, i_k$, we define

$$dx^{i_1} \wedge dx^{i_2} \wedge \ldots dx^{i_k} = k!\mathcal{A}\left(dx^{i_1} \otimes \ldots \otimes dx^{i_k}\right).$$

Note that this is zero if the k-tuple is not distinct, and that if we change the order of the k-tuple then the result merely changes sign (depending whether the k-tuple is rearranged by an even or an odd permutation).

The factor $k!$ is included in our definition for the following reason: If we apply the alternating k-tensor $dx^{i_1} \wedge \ldots \wedge dx^{i_k}$ (where $i_1, \ldots, i_k$ are distinct) to the k vectors $\partial_{i_1}, \ldots, \partial_{i_k}$, the result is 1. If we apply it to the same k vectors in a different order (say, rearranged by some permutation σ), then the result is just the sign of σ. Any other k vectors yield zero.

These 'elementary alternating k-tensors' have a geometric interpretation similar to that in Example 13.1.1: The value of $dx^I(v_1, \ldots, v_k)$ is the determinant of the matrix with (m, n) coefficient $dx^{i_m}(v_n)$, and this gives the signed k-dimensional volume of the projection of the parallelopiped generated by $v_1, \ldots, v_k$ onto the subspace generated by $\partial_{i_1}, \ldots, \partial_{i_k}$. This relationship between alternating forms and volumes will be central in the next lecture when we define integration of differential forms and prove Stokes' theorem.

Proposition 13.1.1
(1). $dx^{i_1} \wedge \ldots \wedge dx^{i_k} = \sum_{\sigma \in S_k} \operatorname{sgn}\sigma \, dx^{i_{\sigma(1)}} \otimes \ldots \otimes dx^{i_{\sigma(k)}} = k!\mathcal{A}(dx^{i_1} \otimes \ldots \otimes dx^{i_k})$.
(2). For each k, $\{dx^{i_1} \wedge \ldots \wedge dx^{i_k} : 1 \leq i_1 < \ldots < i_k \leq n\}$ *is a basis for* $\Lambda^k T_x^*M$. *In particular the space of alternating* k*-tensors at* x *has dimension* $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Proof. (1) is immediate, since the value on any k-tuple of coordinate basis vectors agrees. To prove (2), we note that by Proposition 12.2.2, any alternating tensor can be written as a linear combination of the basis elements $dx^{i_1} \otimes \ldots \otimes dx^{i_k}$. Invariance under $\mathcal{A}$ shows that this is the same as a linear combination of k-forms of the form $\mathcal{A}(dx^{i_1} \otimes \ldots \otimes dx^{i_k})$, and these are all of the form given. It remains to show the supposed basis is linearly independent, but this is also immediate since if $I = (i_1, \ldots, i_k)$ then $dx^I(\partial_{i_1}, \ldots, \partial i_k) = 1$, but $dx^J(\partial_{i_1}, \ldots, \partial i_k) = 0$ for any increasing k-tuple $J \neq I$. $\square$

It follows that any alternating k-tensor T can be written in the form

$$T = \sum_{1 \leq i_1 < \ldots < i_k \leq n} T_{i_1 \ldots i_k} dx^{i_1} \wedge \ldots \wedge dx^{i_k}$$

for some coefficients $T_{i_1 \ldots i_k}$. Some caution is required here, because T can also be written in the form

$$T = \frac{1}{k!} \sum_{i_1, \ldots, i_k} T_{i_1 \ldots i_k} dx^{i_1} \wedge \ldots \wedge dx^{i_k}$$

where the coefficients are the same as above for increasing k-tuples, and to be given from these by antisymmetry in other cases. Thus the coefficients in this expression differ by a factor $k!$ from those given in the expression after Proposition 12.2.2.

13.2 The wedge product

The projection $\mathcal{A}$ immediately gives us a notion of product on alternating tensors, which we have already implicity built into our notation for the basis elements for $\Lambda^k T_x^* M$:

Definition 13.2.1 Let $S \in \Lambda^k T_x^* M$ and $T \in \Lambda^l T_x^* M$ be alternating tensors. Then the **wedge product** $S \wedge T$ of S and T is the alternating $k + l$-tensor given by

$$S \wedge T = \frac{(k+l)!}{k!l!} \mathcal{A}(S \otimes T).$$

This may not seem the obvious definition, because of the factor on the right. This is chosen to make our notation consistent with that in our definition of the basis elements: Take an increasing k-tuple $i_1, \ldots, i_k$ and an increasing l-tuple $j_1, \ldots, j_l$, and assume for simplicity that $i_k < j_1$. Then we can form the alternating tensors $dx^{i_1} \wedge \ldots \wedge dx^{i_k}$, $dx^{j_1} \wedge \ldots \wedge dx^{j_l}$ and $dx^{j_1} \wedge \ldots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_l}$, and we would like to know that the third of these is the wedge product of the first two.

Proposition 13.2.1
The wedge product is characterized by the following properties:
(i). *Associativity:* $f \wedge (g \wedge h) = (f \wedge g) \wedge h$;
(ii). *Homogeneity:* $(cf) \wedge g = c(f \wedge g) = f \wedge (cg)$;
(iii). *Distributivity: If* f *and* g *are in* $\Lambda^k T_x^* M$ *then*

$$(f + g) \wedge h = (f \wedge h) + (g \wedge h);$$

(iv). *Anticommutativity: If* $f \in \Lambda^k T_x^* M$ *and* $g \in \Lambda^l T_x^* M$*, then*

$$g \wedge f = (-1)^{kl} f \wedge g;$$

(v). In any chart,

$$(dx^{i_1}\wedge\ldots\wedge dx^{i_k})\wedge(dx^{j_1}\wedge\ldots\wedge dx^{j_l}) = dx^{i_1}\wedge\ldots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\ldots\wedge dx^{j_l}.$$

Proof. We start by proving (v). Choose a chart about x. Then

$$\begin{aligned}
&(dx^{i_1}\wedge\ldots\wedge dx^{i_k})\wedge(dx^{j_1}\wedge\ldots\wedge dx^{j_l})\\
&\quad=\frac{(k+l)!}{k!l!}\mathcal{A}\left((dx^{i_1}\wedge\ldots\wedge dx^{i_k})\otimes(dx^{j_1}\wedge\ldots\wedge dx^{j_l})\right)\\
&\quad=\frac{(k+l)!}{k!l!}\mathcal{A}\left(\sum_{\sigma\in S_k,\tau\in S_l}\mathrm{sgn}\sigma\,\mathrm{sgn}\tau\, dx^{i_{\sigma(1)}}\otimes\ldots\otimes dx^{i_{\sigma(k)}}\otimes dx^{j_{\tau(1)}}\otimes\ldots\otimes dx^{j_{\tau(l)}}\right)\\
&\quad=\frac{1}{k!l!}\sum_{\sigma\in S_k,\tau\in S_l}\mathrm{sgn}\sigma\,\mathrm{sgn}\tau\, dx^{i_{\sigma(1)}}\wedge\ldots\wedge dx^{i_{\sigma(k)}}\wedge dx^{j_{\tau(1)}}\wedge\ldots\wedge dx^{j_{\tau(l)}}\\
&\quad=dx^{i_1}\wedge\ldots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\ldots\wedge dx^{j_l}.
\end{aligned}$$

The homogeneity and distributivity properties of the wedge product are immediate from the definition. From this we can deduce the following expression for the wedge product in local coordinates: For $S = S_{i_1\ldots i_k}dx^{i_1}\wedge\ldots\wedge dx^{i_k}$ and $T = T_{j_1\ldots j_l}dx^{j_1}\wedge\ldots\wedge dx^{j_l}$ (summing over increasing k-tuples and l-tuples respectively)

$$S\wedge T = \frac{1}{k!l!}\sum_{i_1,\ldots,i_k,j_1,\ldots,j_l} S_{i_1\ldots i_k}T_{j_1\ldots j_l}dx^{i_1}\wedge\ldots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\ldots\wedge dx^{j_l}.$$

The associativity property can now be checked straightforwardly. Finally, we derive the anticommutativity property (iv): If $g = g_{i_1\ldots i_k}dx^{i_1}\wedge\ldots\wedge dx^{i_k}$ and $f = f_{j_1\ldots j_l}dx^{j_1}\wedge\ldots\wedge dx^{j_l}$, then

$$\begin{aligned}
g\wedge f &= \frac{1}{k!l!}\sum_{i_1,\ldots,i_k,j_1,\ldots,j_l} g_{i_1\ldots i_k}f_{j_1\ldots j_l}dx^{i_1}\wedge\ldots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\ldots\wedge dx^{j_l}\\
&= \frac{(-1)^k}{k!l!}\sum_{i_1,\ldots,i_k,j_1,\ldots,j_l} g_{i_1\ldots i_k}f_{j_1\ldots j_l}dx^{j_1}\wedge dx^{i_1}\wedge\ldots\wedge dx^{i_k}\wedge dx^{j_2}\wedge\ldots\wedge dx^{j_l}\\
&\ldots\\
&= \frac{(-1)^{kl}}{k!l!}\sum_{i_1,\ldots,i_k,j_1,\ldots,j_l} g_{i_1\ldots i_k}f_{j_1\ldots j_l}dx^{j_1}\wedge\ldots\wedge dx^{j_l}\wedge dx^{i_1}\wedge\ldots\wedge dx^{i_k}\\
&= (-1)^{kl} f\wedge g.
\end{aligned}$$

□

13.3 Differential forms

Definition 13.3.1 A k-form ω on a differentiable manifold M is a smooth section of the bundle of alternating k-tensors on M. Equivalently, ω associates to each $x \in M$ an alternating k-tensor ω_x, in such a way that in any chart for M, the coefficients $\omega_{i_1 \dots i_k}$ are smooth functions. The space of k-forms on M is denoted $\Omega^k(M)$.

In particular, a 1-form is a covector field. We will also interpret a 0-form as being a smooth function on M, so $\Omega^0(M) = C^\infty(M)$.

By using the local definition in section 13.2, we can make sense of the wedge product as an operator which takes a k-form and an l-form to a $k + l$-form, which is associative, C^∞-linear in each argument, distributive and anticommutative.

13.4 The exterior derivative

Now we will define a differential operator on differential k-forms.

Proposition 13.4.1 *There exists a unique linear operator* $\mathrm{d} : \Omega^k(M) \to \Omega^{k+1}(M)$ *such that*

(i). *If* $f \in \Omega^0(M) = C^\infty(M)$*, then* $\mathrm{d}f$ *agrees with the differential of* f *(Definition 4.2.1);*

(ii). *If* $\omega \in \Omega^k(M)$ *and* $\eta \in \Omega^l(M)$*, then*

$$\mathrm{d}(\omega \wedge \eta) = (\mathrm{d}\omega) \wedge \eta + (-1)^k \omega \wedge (\mathrm{d}\eta);$$

(iii). $\mathrm{d}^2 = 0$.

Proof. Choose a chart $\varphi : U \to V$ with coordinate tangent vector fields $\partial_1, \dots, \partial_n$.

We will first produce the operator d acting on differential forms on $U \subseteq M$. On this region we have the smooth functions $x^1, \dots, x^n$ given by the components of the map φ. The differentials of these are the one-forms $dx^1, \dots, dx^n$, and in agreement with condition (iii) we assume that $d(dx^i) = 0$ for each i.

By induction and condition (ii), we deduce that $\mathrm{d}(dx^{i_1} \wedge \dots \wedge dx^{i_k}) = 0$ for any k-tuple $i_1, \dots, i_k$.

Now let $f = \frac{1}{k!} \sum_{i_1, \dots, i_k} f_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k}$. The linearity of d, together with condition (ii) and condition (i), imply

$$\mathrm{d}f = \frac{1}{k!} \sum_{i_0, i_1, \dots, i_k} \partial_{i_0} f_{i_1 \dots i_k} dx^{i_0} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

One can easily check that this formula defines an operator which satisfies the required conditions. In particular we can compute d^2 to check that it vanishes:

$$\begin{aligned}
&\mathrm{d}^2\left(\omega_{i_1\dots i_k}dx^{i_1}\wedge\dots\wedge dx^{i_k}\right)\\
&\quad=\mathrm{d}\left(\partial_i\omega_{i_1\dots i_k}dx^i\wedge dx^{i_1}\wedge\dots\wedge dx^{i_k}\right)\\
&\quad=\partial_j\partial_i\omega_{i_1\dots i_k}dx^j\wedge dx^i\wedge dx^{i_1}\wedge\dots\wedge dx^{i_k}\\
&\quad=\frac{1}{2}\left(\frac{\partial^2\omega_{i_1\dots i_k}}{\partial x^j\partial x^i}-\frac{\partial^2\omega_{i_1\dots i_k}}{\partial x^i\partial x^j}\right)dx^j\wedge dx^i\wedge dx^{i_1}\wedge\dots\wedge dx^{i_k}\\
&\quad=0.
\end{aligned}$$

To extend this definition to all of M we need to check that it does not depend on the choice of coordinate chart. Let η be any other chart, with components $y^1,\dots,y^n$. On the common domain of η and φ, we have $x^i = F^i(y)$, where $F=\varphi\circ\eta^{-1}$, and

$$dx^i=\frac{\partial F^i}{\partial y^j}dy^j.$$

This implies that

$$dx^{i_1}\wedge\dots\wedge dx^{i_k}=\sum_{i_1,\dots,i_k,j_1,\dots,j_k}\frac{\partial F^{i_1}}{\partial y^{j_1}}\cdots\frac{\partial F^{i_k}}{\partial y^{j_k}}dy^{j_1}\wedge\dots\wedge dy^{j_k}.$$

Now we can check that if we define the operator d in the y coordinates, then $\mathrm{d}(dx^{i_1}\wedge\dots\wedge dx^{i_k})=0$:

$$\begin{aligned}
\mathrm{d}\left(dx^{i_1}\wedge\dots\wedge dx^{i_k}\right)&=\sum_{I,J}\mathrm{d}\left(\frac{\partial F^{i_1}}{\partial y^{j_1}}\cdots\frac{\partial F^{i_k}}{\partial y^{j_k}}dy^{j_1}\wedge\dots\wedge dy^{j_k}\right)\\
&=\sum_{I,J,j_0}\sum_{m=1}^{k}\frac{\partial^2F^{i_m}}{\partial y^{j_m}\partial y^{j_0}}\prod_{p\neq m}\frac{\partial F^{i_p}}{\partial y^{j_p}}dy^{j_0}\wedge\dots\wedge dy^{j_k}\\
&=\frac{1}{2}\sum_{I,J,j_0}\sum_{m=1}^{k}\left(\frac{\partial^2F^{i_m}}{\partial y^{j_m}\partial y^{j_0}}-\frac{\partial^2F^{i_m}}{\partial y^{j_0}\partial y^{j_m}}\right)\\
&\qquad\qquad\times\left(\prod_{p\neq m}\frac{\partial F^{i_p}}{\partial y^{j_p}}\right)dy^{j_0}\wedge\dots\wedge dy^{j_k}\\
&=0
\end{aligned}$$

It follows (by linearity and distributivity) that the differential operators defined in the two charts agree. □

The differential operator may seem somewhat mysterious. The following example may help:

Example 13.4.1 (The exterior derivative on $\mathbb{R}^3$) The exterior derivative in $\mathbb{R}^3$ captures the differential operators which are normally defined as part of vector calculus: First, the differential operator of a 0-form (i.e. a function f) is just the differential of the function, which we can identify with the gradient vector field ∇f.

Next, consider d applied to a 1-form: For purposes of visualisation, we can identify a 1-form with a vector field by duality: The 1-form $\omega = \omega_1 dx^1 \wedge \omega_2 dx^2 + \omega_3 dx^3$ is identified with the vector field $(\omega_1, \omega_2, \omega_3)$. Applying d to ω, we obtain

$$\begin{aligned}
\mathrm{d}\omega &= \mathrm{d}(\omega_i dx^i) \\
&= \partial_j \omega_i dx^j \wedge dx^i \\
&= (\partial_1\omega_2 - \partial_2\omega_1) dx_1 \wedge dx^2 + (\partial_2\omega_3 - \partial_3 dx^2) dx^2 \wedge dx^3 \\
&\qquad + (\partial_3\omega_1 - \partial_1\omega_3) dx^3 \wedge dx^1.
\end{aligned}$$

The result is a 2-form. We identify 2-forms with vector fields again, by sending $adx^1 \wedge dx^2 + bdx^2 \wedge dx^3 + cdx^3 \wedge dx^1$ to the vector field (b, c, a). With this identification, the exterior derivative on 1-forms is equivalent to the curl operator on vector fields.

Finally, consider d applied to a 2-form (which we again associate to a vector field $V = (V_1, V_2, V_3)$). We find

$$\begin{aligned}
&\mathrm{d}\left(V_3 dx^1 \wedge dx^2 + V_1 dx^2 \wedge dx^3 + V_2 dx^3 \wedge dx^1\right) \\
&\qquad = \left(\partial_3 V_3 + \partial_1 V_1 + \partial_2 V_2) dx^1 \wedge dx^2 \wedge dx^3\right) \\
&\qquad = (\mathrm{div} V)\ dx^1 \wedge dx^2 \wedge dx^3.
\end{aligned}$$

Thus the exterior derivative acting on 2-forms is equivalent to the divergence operator acting on vector fields. The familiar identities from vector calculus that the curl of a gradient is zero and that the divergence of a curl is zero are therefore special cases of the identity $\mathrm{d}^2 = 0$.

13.5 Pull-back invariance

Now we will prove a remarkable result which really makes the theory of differential forms work:

Proposition 13.5.1 *Suppose M and N are differentiable manifolds, and $F : M \to N$ is a smooth map. Then for any $\omega \in \Omega^k(N)$ and $\eta \in \Omega^l(N)$,*

$$F_*(\omega \wedge \eta) = F_*(\omega) \wedge F_*\eta$$

and

$$\mathrm{d}(F_*\omega) = F_*(\mathrm{d}\omega).$$

Proof. The proof of the first statement is immediate from the definition. The second statement is proved by an argument identical to that used to prove that the definition of the exterior derivative does not depend on the chart in Proposition 13.4.1, except that the map F may be a smooth map between Euclidean spaces of different dimension. $\square$

13.6 Differential forms and orientability

There is a useful relationship between orientability of a differentiable manifold M^n and the space of n-forms $\Omega^n(M)$:

Proposition 13.6.1 *A differentiable manifold M is orientable if and only if there exists an n-form $\omega \in \Omega^n(M)$ which is nowhere vanishing on M.*

Proof. Suppose there exists such an n-form ω. Let $\mathcal{A}$ be the set of charts $\varphi : U \to V$ for M for which $\omega(\partial_1, \ldots, \partial_n) > 0$. Then $\mathcal{A}$ is an atlas for M, since any chart for M is either in $\mathcal{A}$ or has its composition with a reflection in $\mathcal{A}$ — in particular charts in $\mathcal{A}$ cover M. Furthermore $\mathcal{A}$ is an oriented atlas: For any pair of charts φ and η in $\mathcal{A}$ with non-trivial common domain of definition in M, we have

$$\partial_i^{(\eta)} = \left(D_i\eta \circ \varphi^{-1}\right)_i^j \partial_j^{(\varphi)},$$

and therefore by linearity and antisymmetry of ω,

$$\omega(\partial_1^{(\eta)}, \ldots, \partial_n^{(\eta)}) = \det D(\eta \circ \varphi^{-1})\omega(\partial_1^{(\varphi)}, \ldots, \partial_n^{(\varphi)}).$$

By assumption, $\omega(\partial_1^{(\eta)}, \ldots, \partial_n^{(\eta)})$ and $\omega(\partial_1^{(\varphi)}, \ldots, \partial_n^{(\varphi)})$ are positive and non-zero. It follows that $\det D(\eta \circ \varphi^{-1}) > 0$.

Conversely, suppose M has an oriented atlas $\mathcal{A} = \{\varphi_\alpha : \ U_\alpha \to V_\alpha\}_{\alpha \in \mathcal{A}}$. Let $\{\rho_\beta\}_{\beta \in \mathcal{J}}$ be a partition of unity subordinate to the cover $\{U_\alpha : \ \alpha \in \mathcal{I}\}$, so that for each $\beta \in \mathcal{J}$ there exists $\alpha(\beta) \in \mathcal{I}$ such that $\mathrm{supp}\rho_\beta \subseteq U_{\alpha(\beta)}$. Define

$$\omega = \sum_{\beta \in \mathcal{J}} \rho_\beta \, dx^1_{\varphi_{\alpha(\beta)}} \wedge \ldots \wedge dx^n_{\varphi_{\alpha(\beta)}}.$$

Then ω is everywhere non-zero, since $dx^1_{\varphi_{\alpha(\beta_1)}} \wedge \ldots \wedge dx^n_{\varphi_{\alpha(\beta_1)}}$ is a positive multiple of $dx^1_{\varphi_{\alpha(\beta_2)}} \wedge \ldots \wedge dx^n_{\varphi_{\alpha(\beta_2)}}$ for $\beta_1 \neq \beta_2$. $\square$

We can interpret this in a slightly different way: For each $x \in M$, let $\mathrm{Or}_x M$ be the set of equivalence classes of non-zero alternating n-tensors at x, where $\omega \sim \eta$ if ω is a positive multiple of η. $\mathrm{Or}_x M$ has exactly two elements for each $x \in M$. Then we take $\mathrm{Or}M = \cup_{x \in M} \mathrm{Or}_x M$, which is the **orientation bundle** of M. On any chart $\varphi : U \to V$ for M, the restriction of this bundle to U is diffeomorphic to $U \times \mathbb{Z}_2$, but this is not necessarily true globally.

A slight modification of the proof of Proposition 13.6.1 gives the result that M is orientable if and only if the orientation bundle of M is trivial (that is, diffeomorphic to $M \times \mathbb{Z}_2$).

13.7 Frobenius' Theorem revisited

Differential forms allow an alternative formulation of the Theorem of Frobenius that we proved in Lecture 7 (Proposition 7.3.5). In order to formulate this, let $\mathcal{D}$ be a k-dimensional distribution on M. We relate this distribution to differential forms by considering the subspace $\Omega_0(\mathcal{D})$ of $\Omega(M)$ consisting of differential forms which yield zero when applied to vectors in the distribution $\mathcal{D}$. This subspace is closed under C^∞ scalar multiplication and under wedge products.

Proposition 13.7.1 *The distribution $\mathcal{D}$ is integrable if and only if the subspace $\Omega_0(\mathcal{D})$ is closed under exterior differentiation.*

Proof. First suppose $\mathcal{D}$ is integrable. Then locally we can choose charts $\varphi : U \to V \subseteq \mathbb{R}^k \times \mathbb{R}^{n-k}$ where the first k directions are tangent to $\mathcal{D}$.

In such a chart, forms in $\Omega_0^l(\mathcal{D})$ have the form

$$\omega_{i_1\dots i_l} dx^{i_1} \wedge \dots \wedge dx^{i_l}$$

where $\omega_{i_1\dots i_l} = 0$ whenever $i_1, \dots, i_l \leq k$. This implies that $\partial_m \omega_{i_1\dots i_l} = 0$ for $i_1, \dots, i_l \leq k$ and arbitrary m. Applying the exterior derivative, we find

$$\mathrm{d}\omega = \partial_m \omega_{i_1\dots i_k} dx^m \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k}$$

which is clearly in $\Omega_0^{l+1}(\mathcal{D})$.

Next suppose $\mathcal{D}$ is not integrable. Then by Frobenius's theorem we can find vector fields X and Y in $\mathcal{X}(\mathcal{D})$ such that $[X, Y] \notin \mathcal{X}(\mathcal{D})$, say in particular $[X, Y]_x \notin \mathcal{D}_x$ for some $x \in M$. Choose a 1-form $\omega \in \Omega_0^1(\mathcal{D})$ such that $\omega_x([X, Y]_x) \neq 0$ (How would you construct such a 1-form?)

Then we compute:

$$\begin{aligned}
\mathrm{d}\omega(X, Y) &= X^i Y^j (\partial_i \omega_j - \partial_j \omega_i) \\
&= X^i \partial_i (Y^j \omega_j) - Y^j \partial_j (X^i \omega_i) - X^i (\partial_i Y_j) \omega_j + Y^j (\partial_j X_i) \omega_i \\
&= X\omega(Y) - Y\omega(X) - \omega([X, Y]) \\
&\neq 0 \quad \text{at } x,
\end{aligned}$$

since $\omega(Y) = 0$ and $\omega(X) = 0$ everywhere, and $\omega([X, Y]) \neq 0$ at x by assumption. Therefore $\Omega_0(\mathcal{D})$ is not closed under exterior differentiation. $\square$

Exercise 13.7.1 The identity $d\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y])$ for the exterior derivative of a one-form generalises to an expression for exterior derivatives of k-forms: If $\omega \in \Omega^k(M)$, then

$$\begin{aligned}\mathrm{d}\omega(X_0,\dots,X_k) &= \sum_{i=0}^{k}(-1)^i X_i\omega(X_0,\dots,\hat{X}_i,\dots,X_k)\\ &\quad+\sum_{0\le i<j\le k}(-1)^{i+j}\omega([X_i,X_j],X_0,\dots,\hat{X}_i,\dots,\hat{X}_j,\dots,X_k).\end{aligned}$$

Prove this identity.

Lecture 14. Stokes' Theorem

In this section we will define what is meant by integration of differential forms on manifolds, and prove Stokes' theorem, which relates this to the exterior differential operator.

14.1 Manifolds with boundary

In defining integration of differential forms, it will be convenient to introduce a slightly more general notion of manifold, allowing for the possibility of a boundary.

Definition 14.1.1 A *boundary chart* $\varphi : U \to V$ for a topological space M about a point $x \in M$ is a continuous map from an open set $U \subseteq M$ to a (relatively) open subset V of $\mathbb{R}^n_+ = \{(x^1, \ldots, x^n) : \ x^n \geq 0$ with $\varphi(x) \in \mathbb{R}^{n-1} \times \{0\}$.

Here the open subsets of $\mathbb{R}^n_+$ are the sets of the form $W \cap \mathbb{R}^n_+$ where $W \subseteq \mathbb{R}^n$ is open.

Definition 14.1.2 A smooth *boundary atlas* $\mathcal{A}$ for a topological space M is a collection of maps $\varphi_\alpha : U_\alpha \to V_\alpha$ each of which is either a chart or a boundary chart for M, such that $\cup U_\alpha = M$ and such that $\varphi_\alpha \circ \varphi_\beta^{-1}$ is a smooth map between open sets of $\mathbb{R}^n_+$ for each α and β.

Definition 14.1.3 A smooth *manifold with boundary* is a topological space M equipped with an equivalence class of smooth boundary atlases, where two boundary atlases are equivalent if their union is again a boundary atlas.

If M is a manifold with boundary, then the *boundary* ∂M of M is the subset of M consisting of all those points $x \in M$ for which there is a boundary chart about x.

Proposition 14.1.1 *If M is a smooth manifold with boundary of dimension n, then ∂M is a smooth manifold of dimension $n-1$, with atlas given by the restriction to ∂M of all boundary charts for M.*

Proof. Let $\varphi : U \to V$ and $\eta : W \to Z$ be boundary charts for M, and let $U_0 = \varphi^{-1}(\mathbb{R}^n \times \{0\} \cap V)$ and $W_0 = \eta^{-1}(\mathbb{R}^n \times \{0\} \cap Z)$. Assume that $U_0 \cap W_0$ is non-empty. Then the associated charts for ∂M are $\varphi_0 = \varphi|_{U_0}$ and $\eta_0 = \eta|_{W_0}$. The transition map $\eta_0 \circ \varphi_0^{-1}$ is given by the restriction of the smooth map $\eta \circ \varphi^{-1}$ to $\mathbb{R}^n \times \{0\}$, and is therefore smooth. □

14.2 Induced orientation on the boundary

Suppose M^n is an oriented manifold with boundary — that is, M is equipped with a smooth boundary atlas $\mathcal{A}$ such that $\varphi_\alpha \circ \varphi_\beta^{-1}$ is an orientation-preserving map for all α and β.

Proposition 14.2.1 *∂M is an orientable manifold of dimension $n-1$.*

Proof. Let $\mathcal{A}$ be an oriented boundary atlas for M. Then the corresponding atlas for ∂M is automatically oriented: Any pair of overlapping oriented boundary charts for M map Rls^{n+} to $\mathbb{R}^n_+$, and the derivative of a transition map on the boundary must have the form

$$D\left(\eta \circ \varphi^{-1}\right) = \begin{bmatrix} D\left(\eta_0 \circ \varphi_0^{-1}\right) & * \\ 0 & a \end{bmatrix}$$

where $a = \langle D(\eta \circ \varphi^{-1})(e_{n+1}), e_{n+1}\rangle > 0$. Therefore $\eta_0 \circ \varphi_0^{-1}$ is orientation-preserving, and the atlas is oriented. □

Note that the orientation can be understood geometrically as follows: An n-tuple of linearly independent vectors $u_1, \ldots, u_n$ tangent to ∂M is called positively oriented if the $(n+1)$-tuple $u_1, \ldots, u_n, \partial_{n+1}$ is oriented in M for any boundary chart.

The orientation constructed on ∂M in the proof is called the *induced orientation* on ∂M from the orientation on M (this is a matter of convention — we could just as well have chosen the opposite orientation).

In the proof of the Proposition above we ignored the case $n = 1$ — the boundary of a 1-dimensional manifold is a 0-dimensional manifold (i.e. a collection of points). What does it mean to define an orientation on a zero-dimensional manifold? Our original definition clearly makes no sense in that case. However the equivalent definition in Proposition 13.6.1 does make sense: We will say a 0-manifold N is oriented if it is equipped with a function (i.e. a 0-form) from N to $\mathbb{Z}_2 = \{-1, 1\}$ (this corresponds to the remarks after the proof of Proposition 13.6.1: The orientation bundle in this case is just $N \times \mathbb{Z}_2$. In this case we also have to allow boundary charts for 1-manifolds which map to $(-\infty, 0]$ as well as $[0, \infty)$ (in higher dimensions we can always transform charts into any half-plane via an orientation-preserving map to map into the upper half-plane, but not if $n = 1$).

14.3 More on partitions of unity

Now we want to extend our results on partitions of unity on manifolds to the slightly more general setting of manifolds with boundary.

Proposition 14.3.1 *Let M be a differentiable manifold with boundary. Then there exists a partition of unity on M subordinate to any boundary atlas for M.*

The proof of this result is identical to our result on existence of partitions of unity on differentiable manifolds, except that we have to include functions with support $B^n_r(0)\times[0,r)$. These can be constructed easily from the smooth compactly supported functions we already know.

14.4 Integration of forms on oriented manifolds

Now we can give some further meaning to differential forms by defining what is meant by integration of differential forms on oriented manifolds. A key point to keep in mind here is that none of our definitions depend on us having a metric on the manifold, so we do not in general have any notion of volume of surface area or length. Nevertheless the structure of differential forms is exactly what is required to produce a well-defined notion of integration.

Let M^n be a compact, oriented differentiable manifold with boundary, and let $\omega\in\Omega^n(M)$. Then we define the integral of ω over M, denoted $\int_M\omega$, as follows: Let $\{\rho_\alpha:\ \alpha\in\mathcal{I}\}$ be a partition of unity subordinate to an oriented boundary atlas for M, so that for each α there exists an oriented chart (either a regular chart or a boundary chart) $\varphi_\alpha:U_\alpha\to V_\alpha$ for M, such that $\mathrm{supp}\rho_\alpha\subset U_\alpha$.

Define

$$\int_M\omega=\sum_{\alpha\in\mathcal{I}}\int_{V_\alpha}\left((\varphi_\alpha^{-1})_*(\rho_\alpha\omega)\right)(e_1,\dots,e_n)\,dx^1\dots dx^n.$$

To put this into words: We write ω as a sum $\sum_\alpha\rho_\alpha\omega$ of forms which are supported in charts. Each of these can be integrated over the chart by integrating the smooth function obtained by plugging in the coordinate tangent vectors for that chart. Then we add the resulting numbers together to get the integral of ω.

We need to check that this does not depend on the choice of partition of unity. Suppose $\{\phi_\beta:\beta\in\mathcal{J}\}$ is any other partition of unity for M, with corresponding oriented coordinate charts $\eta_\beta:W_\beta\to Z_\beta$. Then we have

$$\sum_\alpha \int_{V_\alpha} \left((\varphi_\alpha^{-1})_*(\rho_\alpha\omega)\right)(e_1,\ldots,e_n)\,dx^1\ldots dx^n$$
$$= \sum_{\alpha,\beta} \int_{\varphi_\alpha(U_\alpha\cap W_\beta)} \left((\varphi_\alpha^{-1})_*(\rho_\alpha\phi_\beta\omega)\right)(e_1,\ldots,e_n)\,dx^1\ldots dx^n.$$

Fix α and β. Then by definition of the pull-back, we have for any form σ

$$\left((\varphi_\alpha^{-1})_*\sigma\right)(e_1,\ldots,e_n) = \left((\eta_\beta^{-1})_*\sigma\right)(D(\eta_\beta\circ\varphi_\alpha^{-1})(e_1),\ldots,D(\eta_\beta\circ\varphi_\alpha^{-1})(e_n)).$$

We apply the following useful Lemma:

Lemma 14.4.1 *Let ω be an alternating n-tensor, and L a linear map. Then*

$$\omega(Le_1,\ldots,Le_n) = (\det L)\omega(e_1,\ldots,e_n).$$

Proof. We can assume that ω is non-zero. Consider the map from $GL(n)$ to $\mathbb{R}$ defined by

$$\tilde{\det} : L \mapsto \frac{\omega(Le_1,\ldots,Le_n)}{\omega(e_1,\ldots,e_n)}.$$

The denominator is non-zero by assumption.

The multilinearity and antisymmetry of ω imply $\tilde{\det}$ is linear in each row, is unchanged by adding one row to another, and has the value 1 if $L = I$. These are the axioms that define the determinant, so $\tilde{\det} = \det$. □

It follows that

$$\left((\varphi_\alpha^{-1})_*\sigma\right)(e_1,\ldots,e_n) = \det(D(\eta_\beta\circ\varphi_\alpha^{-1}))\left((\eta_\beta^{-1})_*\sigma\right)(e_1,\ldots,e_n).$$

We also know, since the charts are oriented, that the determinant on the right-hand side is positive. The change of variables formula therefore gives

$$\int_{\varphi_\alpha(U_\alpha\cap W_\beta)} \left((\varphi_\alpha^{-1})_*(\rho_\alpha\phi_\beta\omega)\right)(e_1,\ldots,e_n)\,dx^1\ldots dx^n$$
$$= \int_{\varphi_\alpha(U_\alpha\cap W_\beta)} |\det(D(\eta_\beta\circ\varphi_\alpha^{-1}))|\left((\eta_\beta^{-1})_*\sigma\right)(e_1,\ldots,e_n)\,dx^1\ldots dx^n$$
$$= \int_{\eta_\beta(U_\alpha\cap W_\beta)} \left((\eta_\beta^{-1})_*\sigma\right)(e_1,\ldots,e_n)\,dx^1\ldots dx^n.$$

Consequently

$$\sum_\alpha \int_{V_\alpha} \left((\varphi_\alpha^{-1})_*(\rho_\alpha\omega)\right)(e_1,\ldots,e_n)\,dx^1\ldots dx^n$$
$$= \sum_\beta \int_{Z_\beta} \left((\eta_\beta^{-1})_*(\phi_\beta\omega)\right)(e_1,\ldots,e_n)\,dx^1\ldots dx^n,$$

so the integral of ω is well-defined.

14.5 Stokes' theorem

Now we are in a position to prove the fundamental result concerning integration of forms on manifolds, namely Stokes' theorem. This will also give us a geometric interpretation of the exterior derivative.

Proposition 14.5.1 *Let M^n be a compact differentiable manifold with boundary, and let $\omega \in \Omega^{n-1}(M)$. Then*

$$\int_M d\omega = \int_{\partial M} \omega$$

where the integral on the right-hand side is taken using the induced orientation on $\partial\Omega$, integrating the restriction of ω to ∂M (i.e. the pull-back of ω by the inclusion map).

In particular, if M is a compact manifold without boundary, then the integral of the exterior derivative of any $(n-1)$-form is zero.

Proof. Let $\{\rho_\alpha\}$ be a partition of unity on M, with each ρ_α supported in a chart $\varphi_\alpha : U_\alpha \to V_\alpha$. We can write

$$\begin{aligned}
\int_M d\omega &= \sum_\alpha \int_M d(\rho_\alpha \omega) \\
&= \sum_\alpha \int_{V_\alpha} ((\varphi_\alpha^{-1})_*(d(\rho_\alpha\omega)))(e_1, \ldots, e_n) dx^1 \ldots dx^n \\
&= \sum_\alpha \int_{V_\alpha} d\left((\varphi_\alpha^{-1})_*(\rho_\alpha\omega)\right)(e_1, \ldots, e_n)\, dx^1 \ldots dx^n.
\end{aligned}$$

For each α there are two possibilities: The chart φ_α is either a regular chart or a boundary chart.

In the first case, V_α is an open set in $\mathbb{R}^n$. Write ω in components in the chart φ_α:

$$\omega = \sum_{j=1}^n \omega_j dx^1 \wedge \hat{dx^j} \wedge \ldots \wedge dx^n.$$

Then the integrand in the corresponding integral becomes

$$\sum_{j=1}^n \frac{\partial(\rho_\alpha \omega_j)}{\partial x^j}.$$

Applying Fubini's theorem and the fundamental theorem of calculus, and noting that $\rho_\alpha = 0$ on the boundary of our domain, we find that the resulting integral is zero.

In the second case the value of the corresponding integral is

$$\begin{aligned}\sum_{j=1}^{n}\int_{V_\alpha}\frac{\partial(\rho_\alpha\omega_j)}{\partial x^j}dx^1\dots dx^n &= \int_{\mathbb{R}^{n-1}\times\{0\}}\int_{-\infty}^{0}\frac{\partial(\rho_\alpha\omega_n)}{\partial x^n}dx^n\,dx^1\dots dx^{n-1}\\ &= \int_{\mathbb{R}^{n-1}\times\{0\}}\rho_\alpha\omega_n dx^1\dots dx^{n-1}\\ &= \int_{\partial M}\rho_\alpha\omega.\end{aligned}$$

Note that verifying the last line here involves checking that the orientation on ∂M is correct. Summing over α and noting that $\sum_\alpha \rho_\alpha = 1$, we obtain the result. □

Turning the result of Stokes' theorem around, we can interpret the exterior derivative in the following way: Let ω be a k-form in a manifold M. Fix linearly independent vectors $v_1,\dots,v_{k+1}$ in T_xM, and choose any chart φ about x. Write $v_j = v_j^l\partial_l$ in this chart. For r small we can define smooth maps x_r from the k-dimensional sphere S^k into M, by

$$x_r(z^ie_i) = \varphi^{-1}(\varphi(x) + rz^iv_i^ke_k).$$

Using Stokes' theorem, we can deduce

$$d\omega(v_1,\dots,v_{k+1}) = \lim_{r\to 0}\frac{1}{r^{k+1}|B^{k+1}|}\int_{S^k}(x_r)_*\omega$$

where $|B^{k+1}|$ is the volume of the unit ball in $\mathbb{R}^{k+1}$. In this sense the exterior derivative measures the 'boundary integral per unit volume' of a form (where 'volume' is measured in comparison to that of the parallelepiped generated by $v_1,\dots,v_{k+1}$, not using any notion of measure or metric on the manifold).

This is easy to understand in the case of a 0-form (i.e. a function). Then the 'boundary integral' becomes 'difference in values at the endpoints', while 'per unit volume' means 'per unit time along a curve with velocity v_1'. So this just recaptures the usual notion of the directional derivative of a function in terms of difference quotients.

Example 14.5.2 (The case of 1-manifolds)

Let f be a 0-form on a compact 1-manifold M. Note that M is a union of circles $\{S_i^1\}$ and closed intervals I_i with endpoints x_i^+ and x_-^i. An orientation on M amounts to choosing a direction ('left' or 'right') on each component of M, and the induced orientation on ∂M (i.e. the endpoints $x_i^\pm$) is given by assigning an endpoint the value 1 if the orientation direction of M points out of M there, -1 if it points inwards. Each of the closed interval components I_i of M therefore has one endpoint with orientation $+1$ (say x_i^+) and the other with orientation -1. Stokes' theorem becomes

$$\int_M df = \sum_i f(x_i^+) - f(x_i^-).$$

Example 14.5.3 (Regions in $\mathbb{R}^2$)

Let V be a vector field on an bounded open set U in $\mathbb{R}^2$ with smooth boundary curves. We can write $V = V^i e_i$. There is a corresponding 1-form ω defined by

$$\omega(v) = \langle V, v\rangle$$

for all vectors v. Explicitly, this means $\omega = \omega_i dx^i$ where $\omega_i = \omega(e_i) = \langle V, e_i\rangle = V^i$. The exterior derivative is then

$$d\omega = \frac{\partial V^i}{\partial x^j} dx^j \wedge dx^i = (\frac{\partial V^2}{\partial x^1} - \frac{\partial V^1}{\partial x^2}) dx^1 \wedge dx^2,$$

which we recognize as the curl of the vector field V times $dx^1 \wedge dx^2$. The integral of $d\omega$ over U is then

$$\int_U d\omega = \int_U \mathrm{curl} V \, dx^1 dx^2$$

and the integral of ω around the boundary is

$$\int_{\partial U} \omega = \int_{\partial U} \langle V, T\rangle$$

where T is the unit tangent vector to ∂U, taken to run anticlockwise on those parts of the boundary that lie on the 'outside' of U, and clockwise on parts that are 'inside' U. Stokes' theorem tells us that these two are equal. This recaptures the classical Stokes' theorem in the plane.

Example 14.5.4 (Vector fields in space).

The same argument as above applies if V is a vector field in $\mathbb{R}^3$ and M is a two-dimensional submanifold with boundary: There is a corresponding 1-form ω defined as above, and this can be restricted (i.e. pulled back by the inclusion map) to M. The exterior derivative of the resulting form is the curl of V in the normal direction to M, times the volume form on M. So applying out general Stokes' theorem in this case gives that the flux of the curl of the vectgor field V through the surface M is equal to the circulation of V around the boundary of M, which is the classical Stokes' theorem.

There is another way to associate the vector field V with a form: If $V = V^i e_i$ is a vector field, then we can take ω to be a 2-form defined by

$$\omega_{ij} = \varepsilon_{ijk} V^k$$

where ε is the alternating tensor, defined by

$$\varepsilon_{ijk} = \begin{cases} 1 \text{ if } (i,j,k) \text{ is an even permutation of } (1,2,3), \\ -1 \text{ if } (i,j,k) \text{ is an odd permutation of } (1,2,3), \\ 0 \text{ otherwise.} \end{cases}$$

Explicitly this gives

$$\omega = V^3 dx^1 \wedge dx^2 - V^2 dx^1 \wedge dx^3 + V^1 dx^2 \wedge dx^3.$$

Taking the exterior derivative gives

$$d\omega = \sum_{i=1}^{3} \frac{\partial V^i}{\partial x^i} dx^1 \wedge dx^2 \wedge dx^3.$$

This is just the divergence of the vector field V times the volume form. Stokes' theorem then says that the integral of $d\omega$ over a region U (i.e. the integral of the divergence of V over U) is equal to the integral of the 2-form ω over ∂U. The latter is equal to the integral of $\langle V, \mathbf{n} \rangle$ over ∂U, where $\mathbf{n}$ is the outward-pointing unit normal vector. This is just the classical Gauss theorem (or divergence theorem).

Lecture 15. de Rham cohomology

In this lecture we will show how differential forms can be used to define topological invariants of manifolds. This is closely related to other constructions in algebraic topology such as simplicial homology and cohomology, singular homology and cohomology, and Čech cohomology.

15.1 Cocycles and coboundaries

Let us first note some applications of Stokes' theorem: Let ω be a k-form on a differentiable manifold M. For any oriented k-dimensional compact submanifold Σ of M, this gives us a real number by integration:

$$\omega : \Sigma \mapsto \int_\Sigma \omega.$$

(Here we really mean the integral over Σ of the form obtained by pulling back ω under the inclusion map).

Now suppose we have two such submanifolds, Σ_0 and Σ_1, which are (smoothly) homotopic. That is, we have a smooth map $F : \Sigma \times [0,1] \to M$ with $F|_{\Sigma\times\{i\}}$ an immersion describing Σ_i for $i = 0, 1$. Then $d(F_*\omega)$ is a $(k+1)$-form on the $(k+1)$-dimensional oriented manifold with boundary $\Sigma \times [0,1]$, and Stokes' theorem gives

$$\int_{\Sigma\times[0,1]} d(F_*\omega) = \int_{\Sigma_1} \omega - \int_{\Sigma_1} \omega.$$

In particular, if $d\omega = 0$, then $d(F_*\omega) = F_*(d\omega) = 0$, and we deduce that $\int_{\Sigma_1} \omega = \int_{\Sigma_0} \omega$.

This says that k-forms with exterior derivative zero give a well-defined functional on homotopy classes of compact oriented k-dimensional submanifolds of M.

We know some examples of k-forms with exterior derivative zero, namely those of the form $\omega = d\eta$ for some $(k-1)$-form η. But Stokes' theorem then gives that $\int_\Sigma \omega = \int_\Sigma d\eta = 0$, so in these cases the functional we defined on homotopy classes of submanifolds is trivial.

This leads us to consider the space of 'non-trivial' functionals on homotopy classes of submanifolds: Each of these is defined by a k-form ω with exterior derivative zero, but is unchanged if we add the exterior derivative of an arbitrary $(k-1)$-form to ω.

We call a k-form with exterior derivative zero a k-cocycle, and a k-form which is an exterior derivative of a form is called a k-coboundary. The space of k-cocycles on M is a vector space, denoted $Z^k(M)$, and the space of k-coboundaries is then $d\Omega^{k-1}(M)$, which is contained in $Z^k(M)$.

15.2 Cohomology groups and Betti numbers

We define the k-th de Rham cohomology group of M, denoted $H^k(M)$, to be

$$H^k(M) = \frac{Z^k(M)}{d\Omega^{k-1}(M)}.$$

Thus an element of $H^k(M)$ is defined by any k-cocycle ω, but is unchanged by changing ω to $\omega + d\eta$ for any $(k-1)$-form η, which agrees with the notion we produced before of a 'nontrivial' functional on homotopy classes of submanifolds.

An element of $H^k(M)$ is called a *cohomology class*, and the cohomology class containing a k-cocycle ω is denoted $[\omega]$. Thus

$$[\omega] = \{\omega + d\eta : \ \eta \in \Omega^{k-1}(M)\}.$$

Since the exterior derivative and Stokes' theorem do not depend in any way on the presence of a Riemannian metric on M, the cohomology groups of M depend only on the differentiable structure on M. It turns out that they in fact depend only on the topological structure of M, and not on the differentiable structure at all — any two homeomorphic manifolds have the same cohomology groups[5]. The groups $H^k(M)$ are therefore topological invariants, which can be used to distinguish manifolds from each other: If two manifolds have different cohomology groups, they cannot be homeomorphic (let alone diffeomorphic).

The k-the cohomology group $H^k(M)$ is a real vector space. The dimension of this vector space is called the kth *Betti number* of M, and denoted $b_k(M)$.

---

[5] The de Rham theorem states that the de Rham cohomology groups are isomorphic to the singular or Čech cohomology groups with real coefficients, and these are defined in purely topological terms. It is also a consequence of this theorem that the cohomology groups are finite dimensional.

15.3 The group $H^0(M)$

The group $H^0(M)$ is relatively easy to understand: The space $Z^0(M)$ is just the space of functions on M with derivative zero, which is the space of locally constant functions. We interpret Ω^{-1} as the trivial vector space. Therefore $H^0(M) \simeq Z^0(M) = \mathbb{R}^N$ where N is the number of connected components of M. Thefore $b_0(M)$ is equal to the number of connected components of M.

15.4 The group $H^1(M)$

The group $H^1(M)$ is closely related to the fundamental group $\pi_1(M)$. We will examine some aspects of this relationship:

Proposition 15.4.1 *Suppose M is connected. If $\omega \in Z^1(M)$ and $[\omega] \neq 0$ in $H^1(M)$, then there exists a smooth curve $\gamma : S^1 \to M$ such that*

$$\omega(\gamma) := \int_{S^1} \gamma^* \omega \neq 0.$$

Proof. We will prove that if $\omega(\gamma) = 0$ for every smooth map $\gamma : S^1 \to M$, then $[\omega] = 0$ in $H^1(M)$.

On each connected component of M choose a 'base point' x_0. We define a function $f \in C^\infty(M) \simeq \Omega^0(M)$ by setting $f(x_0) = 0$ and extending to other points of M according to

$$f(x) = \int_{[0,1]} \gamma^* \omega$$

for any $\gamma : [0,1] \to M$ with $\gamma(0) = x_0$ and $\gamma(1) = x$. This is well-defined, since if γ_1 and γ_2 are two such curves, then the curve $\gamma_1 \#(-\gamma_2)$ obtained by concatenating γ_1 and $-\gamma_2$ (i.e. γ_2 with orientation reversed) gives a map from S^1 to M, so by assumption

$$0 = \int_{S^1} (\gamma_1 \#(-\gamma_2))^* \omega = \int_{S^1} \gamma_1^* \omega - \int_{S^1} \gamma_2^* \omega,$$

and so the value of $f(x)$ is independent of the choice of γ. Finally, $df = \omega$, since in a chart φ about x,

$$df(\partial_i) = \partial_i f = \frac{d}{dt} \int_0^{1+t} \gamma^* \omega = \omega(\partial_i)$$

where $\gamma(1+t) = \varphi^{-1}(\varphi(x) + te_i)$.

This shows that $\omega = df$, so that $[\omega] = 0$ in $H^1(M)$. □

Corollary 15.4.2 *If M has finite fundamental group then $H^1(M) = 0$. In particular if M is simply connected, then $H^1(M) = 0$.*

Proof. Let $\omega \in \Omega^1(M)$ with $d\omega = 0$. Then for any closed loop $\gamma : S^1 \to M$, we have $[\gamma]^n = 0$ in $\pi^1(M)$ for some integer n. Therefore we have a homotopy $F : S^1 \times [0,1] \to M$ from $\gamma\#\gamma\ldots\#\gamma$ to the constant loop c, and Stokes' theorem gives

$$0 = \int_{S^1\times[0,1]} F^* d\omega = \int_{S^1} (\gamma\#\ldots\#\gamma)^*\omega - \int_{S^1} c^*\omega = n\omega(\gamma).$$

Since this is true for all closed loops γ, Proposition 15.4.1 applies to show $[\omega] = 0$ in $H^1(M)$, and so $H^1(M) = 0$. □

The same argument tells us something more: In fact $H^1(M)$ is a subspace of the dual space of the vector space $G \otimes \mathbb{R}$, where G is the abelianisation of $\pi_1(M)$, which is the abelian group given by taking $\pi_1(M)$ and imposing the extra relations $aba^{-1}b^{-1} = 1$ for all elements a and b. In particular, $b_1(M)$ is no greater than the smallest number of generators of $\pi_1(M)$. In fact it turns out (at least for compact manifolds) that $H^1(M)$ is isomorphic to the torsion-free part of the abelianisation of $\pi_1(M)$, as described above. We will not prove this here.

15.5 Homotopy invariance

In this section we will prove a remarkable topological invariance property of cohomology groups: They do not change when the space is continuously deformed.

More precisely, suppose M and N are two manifolds, and F is a smooth map from M to N. Then the pullback of forms induces a homomorphism of cohomology groups: If $\omega \in \Omega^k(N)$ is a cocycle, then so is $F^*\omega \in \Omega^k(M)$, since $d(F^*\omega) = F^*(d\omega)$. Also, is $\omega = d\eta$ then $F^*\omega = F^*d\eta = d(F^*\eta)$, so this map is well-defined on cohomology.

Proposition 15.5.1 *Let $F : M \times [0,1] \to N$ be a smooth map, and set $f_t(x) = F(x,t)$ for each $t \in [0,1]$. Then f_t^* is independent of t.*

Proof. Let $\omega \in \Omega^k(N)$ be a cocycle. Then we can write

$$F^*\omega = \omega_0 + dt \wedge \omega_1$$

where $\omega_0 \in \Omega^k(M)$ and $\omega_1 \in \Omega^{k-1}(M)$ for each t. Then $f_t^*\omega = \omega_0$ for each t. Since $F^*\omega$ is a cocycle, we have

$$0 = dF^*\omega = dt \wedge \left(\frac{\partial\omega_0}{\partial t} - d_M\omega_1\right) + \ldots$$

and therefore

$$f_1^*\omega - f_0^*\omega = \omega_0(1) - \omega_0(0) = \int_0^1 \frac{\partial \omega_0}{\partial t} dt = \int_0^1 d_M \omega_1 dt = d_M \int_0^1 \omega_1 dt.$$

Therefore $f_1^*\omega$ and $f_0^*\omega$ represent the same cohomology class. □

Corollary 15.5.2 *If the smooth map $f : M \to N$ is a homotopy equivalence (that is, there exists a continuous map $g : N \to M$ such that $f \circ g$ and $g \circ f$ are both homotopic to the identity) then f^* is an isomorphism.*

15.6 The Poincaré Lemma

We will compute the cohomology groups for a simple example: A subset B in $\mathbb{R}^n$ is star-shaped (with respect to the origin) if for every point $y \in B$, the interval $\{ty : t \in [0, 1]\}$ is in B.

Proposition 15.6.1 *(The Poincaré Lemma). Let B be a star-shaped open set in $\mathbb{R}^n$. Then $H^k(B) = \{0\}$ for $k = 1, \ldots, n$.*

Proof. We need to show that for $k > 0$ every k-cocycle is a k-coboundary. In other words, given a k-form ω on B with $d\omega = 0$, we need to find a $(k-1)$-form η such that $\omega = d\eta$. We will do this with k replaced by $k+1$.

Write $\omega = \omega_{i_0 \ldots i_k} dx^{i_0} \wedge \ldots \wedge dx^{i_k}$. For $y \in B$ write $y = y^j e_j$ and define

$$\eta_{i_1 \ldots i_k} = y^j \int_0^1 t^k (\omega_{ty})_{j i_1 \ldots i_k} \, dt.$$

This defines a k-form η on B. We compute the exterior derivative of η at y:

$$\begin{aligned}
(d\eta)_{i_1 \ldots i_k} &= \sum_{p=0}^{k} (-1)^p \partial_{i_p} \eta_{i_0 \ldots \hat{i_p} \ldots i_k} \\
&= \sum_{p=0}^{k} (-1)^p \int_0^1 t^k (\omega_{ty})_{i_p i_0 \ldots \hat{i_p} \ldots i_k} dt \\
&\quad + \sum_{p=0}^{k} (-1)^p y^j \int_0^1 t^{k+1} ((\partial_{i_p} \omega)_{ty})_{j i_0 \ldots \hat{i_p} \ldots i_k} dt.
\end{aligned}$$

We rewrite the last term using the fact that $d\omega = 0$: This means

$$\begin{aligned}
0 &= (d\omega)_{j i_0 \ldots i_k} \\
&= \partial_j \omega_{i_0 \ldots i_k} - \sum_{p=0}^{k} (-1)^p \partial_{i_p} \omega_{j i_0 \ldots \hat{i_p} \ldots i_k}.
\end{aligned}$$

This gives (by the antisymmetry of the components of ω)

$$
\begin{aligned}
(d\eta)_{i_1\ldots i_k} &= (k+1)\int_0^1 t^k(\omega_{ty})_{i_0\ldots i_k}\,dt \\
&\quad + y^j\int_0^1 t^{k+1}((\partial_j\omega)_{ty})_{i_0\ldots i_k}\,dt \\
&= (k+1)\int_0^1 t^k(\omega_{ty})_{i_0\ldots i_k}\,dt \\
&\quad + \int_0^1 t^{k+1}\frac{\partial}{\partial t}(\omega_{ty})_{i_0\ldots i_k}\,dt \\
&= (\omega_y)_{i_0\ldots i_k}
\end{aligned}
$$

by the fundamental theorem of calculus. Thus $d\eta = \omega$, and $H^{k+1}(B) = \{0\}$. □

Remark. Now that we have seen the explicit proof of the Poincaré Lemma, I remark that there is a very simple proof using the homotopy invariance result of Proposition 15.4.1: First, the cohomology of $\mathbb{R}^0$ is trivial to compute. Second, there is a smooth homotopy equivalence $f : B \to \mathbb{R}^0 = \{0\}$ defined by $f(x) = 0$: If we take $g : \mathbb{R}^0 \to B$ to be given by $g(0) = 0$, then we have $f \circ g$ equal to the identity on $\mathbb{R}^0$, and $g \circ f(x) = 0$ on B. The latter is homotopic to the identity under the homotopy $F : B \times [0,1] \to B$ given by $F(x,t) = (1-t)x$. Corollary 15.5.2 applies.

15.7 Chain complexes and exact sequences

In this section we will discuss some algebraic aspects of cohomology.

The algebraic situation we are dealing with is the following: We have a complex Ω^* consisting of a sequence of real vector spaces Ω^k, together with linear operators $d : \Omega^k \to \Omega^{k+1}$ satisfying $d^2 = 0$. In any such situation we can define the cohomology groups of the complex as $H^k(\Omega) = \mathrm{ker}d_k/\mathrm{im}d_{k-1}$. We will call such a complex a co-chain complex, the elements of the complex as co-chains, co-chains in the kernel of d as cocycles, and those in the image of d as coboundaries.

Suppose we have two chain complexes A^* and B^*. A chain map f from A^* to B^* is given by a sequence of linear maps f^k from A^k to B^k such that $df^k = f^{k+1}d$ for any k.

A chain map induces a homomorphism of cohomology groups: If $\omega \in A^k$ with $d\omega = 0$, then $f^k\omega \in B^k$ with $df^k\omega = 0$. If we take another representative of the same cohomology class, say $\eta = \omega + d\mu$, then $f^k\eta = f^k\omega + df^{k-1}\mu$ is in the same cohomology class as $f^k\omega$. Therefore we have a well-defined homomorphism of cohomology groups, which we also denote by f.

Of particular interest here is the situation where we have three chain complexes, say A^*, B^* and C^*, with chain maps $f : A^* \to B^*$ and $g : B^* \to C^*$ forming a short exact sequence — that is, for each k, f^k is injective, g^k is surjective, and the kernel of the map g^k coincides with the image of the map f^k.

It follows that we have a sequence of maps f from $H^k(A)$ to $H^k(B)$ and g from $H^k(B)$ to $H^k(C)$, and that the kernel of g coincides with the image of f. Let us consider those cohomology classes in $H^k(C)$ which are in the image of g. If ω is a C-cocycle, then we know that $\omega = g\eta$ for some cochain η in B^k, by the assumption of surjectivity of g. However, we cannot deduce that η is a cocycle. However we can deduce that $gd\eta = df\eta = d\omega = 0$, and since the kernel of g coincides with the image of f it follows that $d\eta = f\mu$ for some cochain $\mu \in A^{k+1}$. Then we have $fd\mu = df\mu = dd\eta = 0$, and the injectivity of f implies $d\mu = 0$. Therefore μ represents a cohomology class in $H^{k+1}(A)$. In the case where ω is the image of a cocycle in B under g, we have $d\eta = 0$ and hence $\mu = 0$. Conversely, if $[\mu] = 0$ then $\mu = d\sigma$, hence $d(\eta - f\sigma) = 0$, and $\omega = g(\eta - f\sigma)$, so $[\omega] = g[\eta - f\sigma]$ is in the image of g.

This suggests that we have a homomorphism from $H^k(C)$ to $H^{k+1}(A)$ with kernel coinciding with the image of g. To verify this we need to show that the cohomology class of μ does not depend on our choice of η or on our choice of representative of the cohomology class of ω.

Independence of the choice of η is easy to check: η can be replaced by $\eta + f\sigma$ for arbitrary $\sigma \in A^k$. Therefore $d\eta$ is replaced by $d\eta + df\sigma = d\eta + fd\sigma$, and μ is placed by $\mu + d\sigma$ which is in the same cohomology class as μ.

Independence of the choice of representative in the cohomology class of ω also follows easily: If we replace ω by $\omega + d\alpha$, then η is replaced by $\eta + d\beta$, and $d\eta$ is unchanged, so μ is unchanged.

The homomorphism we have constructed is called the connecting homomorphism. Finally, we note that the image of the connecting homomorphism coincides with the kernel of f: If μ arises from some cohomology class ω, then we have by construction $f\mu = d\eta$, so $[f\mu] = 0$ in $H^{k+1}(B)$. Conversely, if $[f\mu] = 0$, then $f\mu = d\eta$ for some η, and then μ is given by applying the connecting homomorphism to $[g\eta]$ (note that $dg\eta = gd\eta = gf\mu = 0$, so $g\eta$ does represent a cohomology class).

We have therefore produced from the short exact sequence of chain complexes a long exact sequence of cohomology groups:

$$\ldots \to H^k(A) \to H^k(B) \to H^k(C) \to H^{k+1}(A) \to \ldots$$

In the next few sections we will see some example of these long exact sequences in cohomology and their applications.

15.8 The Meyer-Vietoris sequence

Next we want to discuss a way to compute the cohomology of complicated manifolds by cutting them up into simpler pieces. Suppose M is a manifold which is the union of two open subsets U and V, and suppose that we know the cohomology groups of U, V and the intersection $U \cap V$. We want to relate the cohomology groups of M to these. We will do this by constructing an exact sequence relating the cohomology groups of M, U, V and $U \cap V$.

Let ω be a k-cochain on M. Then the restriction of ω to U and to V are also k-cochains. This defines a chain map i from $\Omega^k(M)$ to $\Omega^k(U) \oplus \Omega^k(V)$, given by

$$i(\omega) = (\omega|_U, \omega|_V).$$

Similarly, if α and β are k-cochains on U and V respectively, then we can consider their restrictions to the intersection $U \cap V$, and these are again k-cocycles. We consider the map j from $\Omega^k(U) \oplus \Omega^k(V) \to \Omega^k(U \cap V)$ given by

$$(\alpha, \beta) \mapsto \alpha|_{U\cap V} - \beta|_{U\cap V}.$$

This is again a chain map.

These two chain maps define a short exact sequence: The map i is injective, j is surjective, and the image of i coincides with the kernel of j.

By the argument of the previous section, this short exact sequence of cochain complexes gives rise to a long exact sequence of cohomology groups. This long exact sequence of cohomology groups is called the *Meyer-Vietoris sequence* for de Rham cohomology:

$$H^0(M) \xrightarrow{i} H^0(U) \oplus H^0(V) \xrightarrow{j} H^0(U \cap V) \xrightarrow{\eta} H^1(M) \xrightarrow{i} \dots$$
$$\dots H^k(M) \xrightarrow{i} H^k(U) \oplus H^k(V) \xrightarrow{j} H^k(U \cap V) \xrightarrow{\eta} H^{k+1}(M) \xrightarrow{i} \dots$$

15.9 Compactly supported cohomology

The algebraic discussions of section 15.7 allow us to extend our notion of cohomology to more general situations where we have chain complexes. Here we introduce the notion of compactly supported cohomology:

Let M be a smooth manifold. Then we denote by $\Omega^k_c(M)$ the space of differential k-forms on M with compact support. This is a subspace of $\Omega^k(M)$ which is closed under exterior differentiation, and hence forms a cochain complex. The cohomology of this complex is called the compactly supported cohomology of M, and denoted $H^k_c(M)$.

Next we will give another useful example of a long exact sequence relating compactly supported cohomology to the usual cohomology.

Suppose M is a smooth manifold, and Σ is a submanifold within M. Then we have a natural chain map from $H^*(M)$ to $H^*(\Sigma)$ given by pulling back forms via the inclusion map i. This map is surjective. The kernel consists of those forms ω on M which have $i^*\omega = 0$. This is again a chain complex, since $i^*(d\omega) = di^*\omega = 0$ if $i^*\omega = 0$. This gives us a short exact sequence relating the cohomologies of M, Σ, and the chain complex $\Omega^k_0(M,\Sigma) = \{\omega \in \Omega^*(M) : \ i^*\omega = 0\}$.

We will now show that the cohomology of the latter is isomorphic to the compactly supported cohomology of $M\backslash\Sigma$.

To see this, we note that $\Omega^k_c(M\backslash\Sigma) \subset \Omega^k_0(M,\Sigma)$. We denote by C^k the quotient space $\Omega^k_0(M,\Sigma)/\Omega^k_c(M\backslash\Sigma)$. We define an operator $d : C^k \to C^{k+1}$ by $d[\omega] = [d\omega]$. If $\eta \in \Omega^k_c(M\backslash\Sigma)$, then $d(\omega+\eta) = d\omega+d\eta \in d\omega+\Omega^{k+1}_c(M\backslash\Sigma)$, so this operator is well defined and satisfies $d^2 = 0$. Therefore the complex C is a cochain complex, and we have a short exact sequence of chain complexes

$$0 \to \Omega^*_c(M\backslash\Sigma) \to \Omega^*_0(M,\Sigma) \to C \mapsto 0.$$

This induces a long exact sequence in cohomology:

$$\ldots \to H^k_c(M\backslash\Sigma) \to H^k_0(M,\Sigma) \to H^k(C) \to H^{k+1}_c(M\backslash\Sigma)\ldots$$

We will prove that $H^k(C) = 0$ for all k, and the long exact sequence above then implies that $H^k_c(M\backslash\Sigma) \simeq H^k_0(M,\Sigma)$.

Suppose $\omega \in \Omega^k_0(M,\Sigma)$ satisfies $d[\omega] = 0$, that is,

$$d\omega = \eta$$

for some $\eta \in \Omega^{k+1}_c(M\backslash\Sigma)$. We want to show that $[\omega] = d[\sigma]$ for some $[\sigma] \in C^{k-1}$, which means we want to show that $\omega - d\sigma \in \Omega^k_c(M\backslash\Sigma)$.

Since Σ is a smooth compact submanifold, the nearest-point projection p (defined using any Riemannian metric on M) is a smooth map from a neighbourhood T of Σ in M to Σ, and is homotopic to the identity map on T. We can assume that $du = 0$ on T since $du \in \Omega^{k+1}_c(M\backslash\Sigma)$. Therefore by the homotopy invariance we have

$$\omega - p^*\omega = dv$$

for some $v \in \Omega^{k-1}_0(T,\Sigma)$. But we also have $p = i \circ p$ and so $p^*\omega = p^*i^*\omega = 0$ since $\omega \in \Omega^k_0(M,\Sigma)$, and we have $\omega = dv$. Now let φ be a smooth function on M which is identically 1 in a neighbourhood of Σ, but identically zero in a neighbourhood of $M\backslash T$. Then $\varphi v \in \Omega^{k-1}_0(M,\Sigma)$, and $\omega - d(\varphi v) \in \Omega^k_c(M\backslash\Sigma)$. Therefore $[\omega] - d[\varphi v] = [\omega - d(\varphi v)] = 0$ in C^k, and $H^k(C) = 0$ as claimed.

Therefore we have a long exact sequence in cohomology:

$$\ldots H^k_c(M\backslash\Sigma) \to H^k(M) \to H^k(\Sigma) \to H^{k+1}_c(M\backslash\Sigma) \to \ldots$$

15.10 Cohomology of spheres

We will use the Meyer-Vietoris sequence to deduce the cohomology groups of the spheres S^n for any n. We start with the circle S^1: We can think of this as a union of two intervals U and V, such that $U \cap V$ is a union of two disjoint intervals.

Now we can apply the Meyer-Vietoris sequence to compute the cohomology of $S^1 = U \cup V$: The sequence becomes

$$0 \mapsto \mathbb{R} \mapsto \mathbb{R}^2 \mapsto \mathbb{R}^2 \mapsto H^1(S^1) \mapsto 0$$

which implies that $H^1(S^1) = \mathbb{R}$ (this could be computed directly by seeing what the cocycles are on S^1 explicitly). Clearly $H^k(S^1) = \{0\}$ for $k > 1$ because S^1 is a 1-manifold.

Now consider the cohomology of the sphere S^2: We observe that $S^2 = U \cup V$ where U and V are diffeomorphic to disks and $U \cap V$ is diffeomorphic to $S^1 \times (0,1)$. By homotopy invariance $S^1 \times (0,1)$ has the same cohomology as S^1. So the sequence in this case becomes:

$$0 \mapsto \mathbb{R} \mapsto \mathbb{R}^2 \mapsto \mathbb{R} \mapsto H^1(S^2) \mapsto 0 \mapsto \mathbb{R} \mapsto H^2(M) \mapsto 0.$$

It follows that $H^1(S^2) = \{0\}$ and $H^2(S^2) = \mathbb{R}$.

Proceeding in the same way for higher dimensions, we find $H^k(S^n) = \mathbb{R}$ if $k = 0$ or $k = n$ and $H^k(S^n) = 0$ otherwise.

15.11 Compactly supported cohomology of $\mathbb{R}^n$

Proposition 15.11.1

$$H_c^k(\mathbb{R}^n) = \begin{cases} 0, & k < n \\ \mathbb{R}, & k = n. \end{cases}$$

Proof. For $k = 0$ the result is immediate because constants are not compactly supported in $\mathbb{R}^n$.

For $k = 1$ and $n = 1$ the result is also immediate: If $\omega = \omega_1 dx^1$, then $\omega = df$ implies $\int_{\mathbb{R}} \omega_1 = 0$, and conversely.

We will use the long exact sequence from section 15.9 together with the results on cohomology groups of spheres from section 15.10: The sphere S^n contains an equatorial S^{n-1} as a submanifold, and the complement $S^n \backslash S^{n-1}$ is diffeomorphic to two copies of $\mathbb{R}^n$. Therefore the long exact sequence becomes (for $n > 1$)

$$\begin{aligned} &0 \to \mathbb{R} \to \mathbb{R} \\ &\to H_c^1(\mathbb{R}^n)^2 \to 0 \to 0 \\ &\dots \\ &\to H_c^{n-1}(\mathbb{R}^n)^2 \to 0 \to \mathbb{R} \\ &\to H_c^n(\mathbb{R}^n)^2 \to \mathbb{R} \to 0 \end{aligned}$$

It follows that $H_c^k(\mathbb{R}^n) = 0$ for $k = 1, \dots, n-1$ and $H_c^n(\mathbb{R}^n) = \mathbb{R}$. □

Furthermore, it is immediate that the n-coboundaries of compactly supported cohomology are precisely those that have integral zero: This contains the n-coboundaries (by Stokes' theorem), and has codimension 1.

15.12 The group $H^n(M)$

If M is a compact manifold of dimension n, then we always know what the nth cohomology group is:

Proposition 15.8.1 *If M is a compact connected manifold of dimension n, then $H^n(M) = \mathbb{R}$ if M is orientable, and $H^n(M) = \{0\}$ if M is not orientable.*

Proof. First suppose M is oriented. Choose an atlas for M consisting of coordinate regions U_α, $\alpha = 1, \dots, N$, each of which is diffeomorphic to $\mathbb{R}^n$. Let $\{\rho_\alpha\}$ be a smooth partition of unity with ρ_α supported in U_α.

Define a map $\xi : \Omega^n(M) \to \mathbb{R}^N$ by

$$\xi(\omega) = (\int_M \rho_1 \omega, \dots, \int_M \rho_N \omega).$$

Now consider the subspace X of $\mathbb{R}^N$ defined by

$$X = \{\xi(dv) |\ v \in \Omega^{n-1}(M)\}.$$

If ω is exact, then clearly $\xi(\omega) \in X$. Conversely, if $\xi(\omega) \in X$ then we have $v \in \Omega^{n-1}(M)$ such that $\int_M \rho_\alpha(\omega - dv) = 0$ for every α. Now $\rho_\alpha(\omega - dv)$ is a compactly supported form in $\mathbb{R}^n$, with integral zero, and hence by Proposition 15.11.1 there exists $v_\alpha \in \Omega_c^{n-1}(U_\alpha)$ such that $\rho_\alpha(\omega - dv) = dv_\alpha$. Summing over α, we find

$$\omega - dv = \sum_\alpha \rho_\alpha(\omega - dv) = \sum_\alpha dv_\alpha$$

and hence

$$\omega = d(v + \sum_\alpha v_\alpha)$$

and ω is exact.

The subspace X is defined by a finite collection of equations $c_{jk}x_k = 0$, $j = 1, \ldots, K$. Therefore an n-form ω on M is exact if and only if

$$\int_M (c_{jk}\rho_k)\omega = 0$$

for $j = 1, \ldots, K$. Suppose that $c_{jk}\rho_k$ is non-constant for some j. Then in one of the regions U_α we can find an n-form ω supported in U_α with integral zero such that $\int_{U_\alpha} c_{jk}\rho_k\omega$ is non-zero. But then it follows that ω is not exact, contradicting Proposition 15.11.1. Therefore $c_{jk}\rho_k$ is constant for each j, and $\omega \in \Omega^n(M)$ is exact if and only if $\int_M \omega = 0$.

Next suppose M is not orientable. Let $\tilde{M}$ be the double cover of M, which we can define as follows: Define an equivalence relation on $\Lambda^n T_x^* M \backslash \{0\}$ for each $x \in M$ by taking $\omega \sim \eta$ iff $\omega = \lambda\eta$ for some $\lambda > 0$. The quotient space at each point consists of two points, and the quotient bundle $P\Lambda^n T^* M = \{(x, [\omega]) : \ \omega \in \Lambda^n T_x^* M\}$ is a $\mathbb{Z}_2$-bundle over M. Fix $x \in M$ and $\omega \neq 0$ in $\Lambda^n T_x^* M$. Then we take $\tilde{M}$ to be the connected component of $(x, [\omega])$ in $P\Lambda^n T^* M$. If M is orientable then there is a global non-vanishing section of $\Lambda^n T^* M$, so $\tilde{M}$ is diffeomorphic to M, while if M is not orientable then $\tilde{M}$ covers M twice (if M is connected), and there is a natural projection π from $\tilde{M}$ to M given by $\pi(x, [\omega]) = x$. $\tilde{M}$ is always orientable, since $T_(x, [\omega])\tilde{M} \simeq T_x M$, hence $\Lambda^n T^*_{(x,[\omega])}\tilde{M} \simeq \Lambda^n T_x^* M$, and so $[\omega] \in P\Lambda^n T_x^* M$ gives a global section of $P\Lambda^n T^* \tilde{M}$.

In the case where M is not orientable, there is a natural involution i of $\tilde{M}$ induced by the map $\omega \mapsto -\omega$ of $\Lambda^n T^* M$, and this is orientation-reversing. Since $\pi \circ i = \pi$, a differential n-form $\tilde{\omega}$ on $\tilde{M}$ arises from pull-back by π of a differential form ω on M if and only if $i^*\tilde{\omega} = \tilde{\omega}$. But then we have

$$\int_{\tilde{M}} \tilde{\omega} = -\int_{\tilde{M}} i^*\tilde{\omega} = -\int_{\tilde{M}} \tilde{\omega}$$

since i is orientation-reversing, and hence $\int_{\tilde{M}} \tilde{\omega} = 0$. It follows from the case we have already considered that $\tilde{\omega} = d\eta$ for some $\eta \in \Omega^{n-1}(\tilde{M})$. Then let $\tilde{\eta} = (\eta + i^*\eta)/2$. Then we have $i^*\tilde{\eta} = \tilde{\eta}$, so $\tilde{\eta} = \pi^*\eta'$ for some $\eta' \in \Omega^{n-1}(M)$, and $d\tilde{\eta} = (d\eta + di^*\eta)/2 = (\tilde{\omega} + i^*\tilde{\omega})/2 = \tilde{\omega}$. It follows that $d\eta' = \omega$, so ω is exact. Therefore $H^n(M) = 0$, as claimed. $\square$

15.13 Cohomology of surfaces

In this section we will use the results we have developed above about cohomology groups to compute the cohomology of compact surfaces.

We already know the cohomology groups of S^2. Next we will compute the cohomology groups of the torus $\mathbb{T}^2 = S^1 \times S^1$. This can be written as the union of open sets U and V, where U and V are each diffeomorphic to $S^1 \times \mathbb{R}$, and $U \cap V$ is diffeomorphic to two copies of $S^1 \times \mathbb{R}$. We therefore know $H^0(\mathbb{T}^2) = H^0(U) = H^0(V) = \mathbb{R}$, $H^0(U \cap V) = \mathbb{R}^2$, $H^1(U) = H^1(V) = \mathbb{R}$ and $H^1(U \cap V) = \mathbb{R}^2$, and $H^2(U) = H^2(V) = H^2(U \cap V) = 0$, $H^2(\mathbb{T}^2) = \mathbb{R}$. Thus we have the long exact sequence

$$0 \to \mathbb{R} \to \mathbb{R}^2 \to \mathbb{R}^2 \to H^1(\mathbb{T}^2) \to \mathbb{R}^2 \to \mathbb{R}^2 \to \mathbb{R} \to 0 \to 0$$

which implies that $H^1(\mathbb{T}^2) = \mathbb{R}^2$.

We will proceed by induction on the genus of the surface: A surface M_{g+1} of genus $g+1$ can be written as the union of sets A and B, where A is diffeomorphic to $\mathbb{T}^2 \backslash \{p\}$, B is diffeomorphic to $M_g \backslash \{q\}$, and $A \cap B$ is diffeomorphic to $S^1 \times \mathbb{R}$. To use this we first need to find the cohomology groups of $\mathbb{T}^2 \backslash \{p\}$ and $M_g \backslash \{p\}$.

Proposition 15.13.1 *Let M be a compact oriented manifold of dimension $n > 1$, $p \in M$. Then $H^n(M \backslash \{p\}) = 0$.*

Proof. Let ω be an n-form on $M \backslash \{p\}$. Then we can write $\omega = \omega_0 + \omega_1$, where ω_0 is compactly supported in $M \backslash \{p\}$ and has integral equal to zero, and ω_1 is supported in a region diffeomorphic to $S^1 \times (0,1)$, and is identically zero on $S^1 \times (0, 1/2)$.

ω_0 extends to a form on M with integral zero, so there exists $\eta_0 \in \Omega^{n-1}(M)$ such that $\omega_0 = d\eta_0$.

By the proof of the Poincaré Lemma, there also exists a form $\eta_1 \in \Omega^{n-1}(S^1 \times (0,1))$, vanishing on $S^1 \times (0, 1/2)$, such that $\omega_1 = d\eta_1$.

Therefore $\omega = d\eta_0 + d\eta_1$ is exact, and $H^n(M \backslash \{p\}) = 0$. □

From this we can deduce the cohomology of $M_g \backslash \{p\}$ as follows: M_g is the union of U and V, where $U \simeq M_g \backslash \{p\}$, $V \simeq \mathbb{R}^2$, and $U \cap V \simeq S^1 \times \mathbb{R}$.

From this we obtain the long exact sequence

$$0 \to \mathbb{R} \to \mathbb{R}^2 \to \mathbb{R} \to H^1(M_g) \to H^1(M_g \backslash \{p\}) \to \mathbb{R} \to \mathbb{R} \to 0$$

which implies that $H^1(M_g \backslash \{p\}) \simeq H^1(M_g)$.

Finally, we can apply the Meyer-Vietoris sequence to A and B as above, obtaining

$$0 \to \mathbb{R} \to \mathbb{R}^2 \to \mathbb{R} \to H^1(M_{g+1}) \to H^1(M_g) \oplus \mathbb{R}^2 \to \mathbb{R} \to \mathbb{R} \to 0$$

which implies that $H^1(M_{g+1}) \simeq H^1(M_g) \oplus \mathbb{R}^2$.

By induction, we deduce that the cohomology groups of the surface of genus g are given by $H^0(M_g) = \mathbb{R}$, $H^1(M_g) = \mathbb{R}^{2g}$, and $H^2(M_g) = \mathbb{R}$.

Lecture 16. Curvature

In this lecture we introduce the curvature tensor of a Riemannian manifold, and investigate its algebraic structure.

16.1 The curvature tensor

We first introduce the curvature tensor, as a purely algebraic object: If X, Y, and Z are three smooth vector fields, we define another vector field $R(X,Y)Z$ by

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

Proposition 16.1.1 *$R(X,Y)Z$ is a tensor of type $(3,1)$.*

Proof. R is tensorial in the first two arguments, because we can write

$$R(X,Y)Z = \left(\nabla\nabla Z\right)(Y,X) - \left(\nabla\nabla Z\right)(X,Y),$$

and each of the terms of the right is a tensor in X and Y. This leaves one further calculation:

$$\begin{aligned}
R(X,Y)(fZ) &= \nabla_Y\left(f\nabla_X Z + X(f)Z\right) - \nabla_X\left(f\nabla_Y Z + Y(f)Z\right)\\
&\quad - [X,Y](f)Z - f\nabla_{[X,Y]}Z\\
&= f\nabla_Y\nabla_X Z + Y(f)\nabla_X Z + YX(f)Z + X(f)\nabla_Y Z\\
&\quad - f\nabla_X\nabla_Y Z - X(f)\nabla_Y Z - XY(f)Z - Y(f)\nabla_X Z\\
&\quad - [X,Y](f)Z - f\nabla_{[X,Y]}Z\\
&= f\left(\nabla_Y\nabla_X Z - \nabla_X\nabla_Y Z - \nabla_{[X,Y]}Z\right)\\
&\quad + \left(YX(f) - XY(f) - [X,Y](f)\right)Z\\
&= fR(X,Y)Z.
\end{aligned}$$

□

Remark. Note that this calculation does not use the compatibility of the connection with the metric, only the symmetry of the connection. Thus any (symmetric) connection gives rise to a curvature tensor. However, we will only be interested in the case of the Levi-Civita connection from now on.

As usual we can write the curvature tensor in terms of its components in any coordinate tangent basis:

$$R = R_{ikj}{}^l dx^i \otimes dx^k \otimes dx^j \otimes \partial_l.$$

Then an application of the metric index-lowering operator gives a tensor of type $(4,0)$ defined by

$$R(u,v,w,z) = g(\nabla_v\nabla_u w - \nabla_u\nabla_v w - \nabla_{[v,u]}w, z).$$

The components of this are $R_{ijkl} = R_{ijk}{}^p g_{pl}$.

Proposition 16.1.2 (Symmetries of the curvature tensor)

(1). $R_{ikjl} + R_{kijl} = 0$;
(2). $R_{ikjl} + R_{kjil} + R_{jikl} = 0$;
(3). $R_{ikjl} + R_{iklj} = 0$;
(4). $R_{ikjl} = R_{jlik}$.

The second identity here is called the *first Bianchi identity*.

Proof. The first symmetry is immediate from the definition of curvature. For the second, work in a coordinate tangent basis:

$$\begin{aligned}
R_{ikjl} + R_{kjil} + R_{jikl} &= g\left(\nabla_k\nabla_i\partial_j - \nabla_i\nabla_k\partial_j, \partial_l\right)\\
&\quad + g\left(\nabla_j\nabla_k\partial_i - \nabla_k\nabla_j\partial_i, \partial_l\right)\\
&\quad + g\left(\nabla_i\nabla_j\partial_k - \nabla_j\nabla_i\partial_k, \partial_l\right)\\
&= g\left(\nabla_k\left(\nabla_i\partial_j - \nabla_j\partial_i\right), \partial_l\right)\\
&\quad + g\left(\nabla_j\left(\nabla_k\partial_i - \nabla_i\partial_k\right), \partial_l\right)\\
&\quad + g((\nabla_i\left(\nabla_j\partial_k - \nabla_k\partial_j\right), \partial_l)\\
&= 0
\end{aligned}$$

by the symmetry of the connection.

The third symmetry is a consequence of the compatibility of the connection with the metric:

$$\begin{aligned}
0 &= \partial_i\partial_j g_{kl} - \partial_j\partial_i g_{kl}\\
&= \partial_i\left(g(\nabla_j\partial_k, \partial_l) + g(\partial_k, \nabla_j\partial_l)\right)\\
&\quad - \partial_j\left(g(\nabla_i\partial_k, \partial_l) + g(\partial_k, \nabla_i\partial_l)\right)\\
&= g(\nabla_i\nabla_j\partial_k, \partial_l) + g(\nabla_i\partial_k, \nabla_j\partial_l) + g(\nabla_j\partial_k, \nabla_i\partial_l) + g(\partial_k, \nabla_i\nabla_j\partial_l)\\
&\quad - g(\nabla_j\nabla_i\partial_k, \partial_l) - g(\nabla_j\partial_k, \nabla_i\partial_l) - g(\nabla_i\partial_k, \nabla_j\partial_l) - g(\partial_k, \nabla_j\nabla_i\partial_l)\\
&= R_{jikl} + R_{jilk}.
\end{aligned}$$

Finally, the last symmetry follows from the previous ones:

$$\begin{aligned}
R_{ikjl} &=^{(2)} -R_{kjil} - R_{jikl} \\
&=^{(3)} R_{kjli} + R_{jilk} \\
&=^{(2)} -R_{jlki} - R_{lkji} - R_{iljk} - R_{ljik} \\
&=^{(3),(1)} 2R_{jlik} + R_{lkij} + R_{ilkj} \\
&=^{(2)} 2R_{jlik} - R_{kilj} \\
&=^{(1),(3)} 2R_{jlik} - R_{ikjl}.
\end{aligned}$$

□

Note that if M is a one-dimensional Riemannian manifold, then the curvature is zero (since it is antisymmetric). This reflects the fact that any one-dimensional manifold can be locally parametrised by arc length, and so is locally isometric to any other one-dimensional manifold. The curvature tensor is invariant under isometries.

Next consider the two-dimensional case: Any component of R in which the first two or the last two indices are the same must vanish, by symmetries (1) and (3). There is therefore only one independent component of the curvature: If we take $\{e_1, e_2\}$ to be an orthonormal basis for T_xM, then we define the *Gauss curvature* of M at x to be $K = R_{1212}$. This is independent of the choice of basis: Any other one is given by $e'_1 = \cos\theta e_1 + \sin\theta e_2$ and $e'_2 = -\sin\theta e_1 + \cos\theta e_2$, so we have

$$\begin{aligned}
R_{1'2'1'2'} &= R(\cos\theta e_1 + \sin\theta e_2, -\sin\theta e_1 + \cos\theta e_2, e'_1, e'_2) \\
&= \cos^2\theta R(e_1, e_2, e'_1, e'_2) - \sin^2\theta R(e_2, e_1, e'_1, e'_2) \\
&= R(e_1, e_2, e'_1, e'_2) \\
&= R(e'_1, e'_2, e_1, e_2) \\
&= R(e_1, e_2, e_1, e_2) \\
&= R_{1212}.
\end{aligned}$$

More generally, we see that (in any dimension), if $\{e_i\}$ are orthonormal, then R_{ijkl} depends only the (oriented) two-dimensional plane generated by e_i and e_j, and the one generated by e_k and e_l.

16.2 Sectional curvatures

The last observation motivates the following definition:

If Σ is a two-dimensional subspace of T_xM, then the *sectional curvature* of Σ is $K(\sigma) = R(e_1, e_2, e_1, e_2)$, where e_1 and e_2 are any orthonormal basis for Σ. This is indepedent of basis, by the calculation above.

Proposition 16.2.1 *The curvature tensor is determined by the sectional curvatures.*

Proof. We will give an explicit expression for a component R_{ijkl} of the curvature tensor, in terms of sectional curvatures. We work with an orthonormal basis $\{e_1, \ldots, e_n\}$ at a point of M.

For convenience we will refer to the oriented plane generated by e_1 and e_j by the notation $e_i \wedge e_j$. We compute the sectional curvature of the plane $\frac{1}{2}(e_i + e_k) \wedge (e_j + e_l)$:

$$\begin{aligned}
K\left(\frac{(e_i+e_k)\wedge(e_j+e_l)}{2}\right) &= \frac{1}{4}R(e_i+e_k, e_j+e_l, e_i+e_k, e_j+e_l)\\
&= \frac{1}{4}K(e_i\wedge e_j) + \frac{1}{4}K(e_i\wedge e_l) + \frac{1}{4}K(e_j\wedge e_k) + \frac{1}{4}K(e_k\wedge e_l)\\
&\quad + \frac{1}{2}R_{ijil} + \frac{1}{2}R_{ijkj} + \frac{1}{2}R_{ilkl} + \frac{1}{2}R_{klkj}\\
&\quad + \frac{1}{2}R_{ijkl} + \frac{1}{2}R_{kjil}.
\end{aligned}$$

Now add the same expression with e_k and e_l replaced by $-e_k$ and $-e_l$:

$$\begin{aligned}
R_{ijkl} + R_{kjil} &= K\left(\frac{(e_i+e_k)\wedge(e_j+e_l)}{2}\right) + K\left(\frac{(e_i-e_k)\wedge(e_j-e_l)}{2}\right)\\
&\quad - \frac{1}{2}K(e_i\wedge e_j) - \frac{1}{2}K(e_i\wedge e_l) - \frac{1}{2}K(e_j\wedge e_k) - \frac{1}{2}K(e_k\wedge e_l).
\end{aligned}$$

Finally, subtract the same expression with e_i and e_j interchanged: On the left-hand side we get

$$R_{ijkl} + R_{kjil} - R_{jikl} - R_{kijl} = 2R_{ijkl} - R_{jkil} - R_{kijl} = 3R_{ijkl}$$

by virtue of the Bianchi identity. Thus we have

$$\begin{aligned}
R_{ijkl} &= \frac{1}{3}K\left(\frac{(e_i+e_k)\wedge(e_j+e_l)}{2}\right) + \frac{1}{3}K\left(\frac{(e_i-e_k)\wedge(e_j-e_l)}{2}\right)\\
&\quad - \frac{1}{3}K\left(\frac{(e_j+e_k)\wedge(e_i+e_l)}{2}\right) - \frac{1}{3}K\left(\frac{(e_j-e_k)\wedge(e_i-e_l)}{2}\right)\\
&\quad - \frac{1}{6}K(e_j\wedge e_l) - \frac{1}{6}K(e_i\wedge e_k) + \frac{1}{6}K(e_i\wedge e_l) + \frac{1}{6}K(e_j\wedge e_k).
\end{aligned}$$

□

16.3 Ricci curvature

The Ricci curvature is the symmetric $(2,0)$-tensor defined by contraction of the curvature tensor:

$$R_{ij} = \delta_l^k R_{ikj}{}^l = g^{kl} R_{ikjl}.$$

This can be interpreted in terms of the sectional curvatures: Given a unit vector v, choose an orthonormal basis for TM with $e_n = v$. Then we have

$$R(v,v) = \sum_{i=1}^{n} R(e_i, v, e_i, v) = \sum_{i=1}^{n=1} R_{inin} = \sum_{i=1}^{n} K(v \wedge e_i).$$

Thus the Ricci curvature in direction v is an average of the sectional curvatures in 2-planes containing v.

16.4 Scalar curvature

The scalar curvature is given by a further contraction of the curvature:

$$R = g^{ij} R_{ij} = g^{ij} g^{kl} R_{ikjl}.$$

$R(x)$ then (except for a constant factor depending on n) the average of the sectional curvatures over all 2-planes in $T_x M$.

16.5 The curvature operator

The full algebraic structure of the curvature tensor is elucidated by constructing a vector space on which it acts as a bilinear form.

At each point x of M we let $\Lambda^2 T_x M$ be the vector space obtained by dividing the space $T_x M \otimes T_x M$ by the relation

$$u \otimes v \sim -v \otimes u.$$

This is a vector space of dimension $n(n-1)/2$, with basis elements

$$e_i \wedge e_j = [e_i \otimes e_j]$$

for $i < j$. More generally, if u and v are any two vectors in $T_x M$, we denote

$$u \wedge v = [u \otimes v].$$

This is called the wedge product of the vectors u and v.

In particular, if u and v are orthogonal and have unit length, then we identify $u \wedge v \in \Lambda^2 T_x M$ with the two dimensional oriented plane in $T_x M$

generated by u and v. The construction of $\Lambda^2 T_x M$ simply extends the set of two-dimensional planes in $T_x M$ to a vector space, to allow formal sums and scalar multiples of them. We refer to the space $\Lambda^2 T_x M$ as the space of 2-planes at x (even though not everything can be interpreted as a plane in $T_x M$), and the corresponding bundle is the 2-plane bundle of M. We extend the metric to $\Lambda^2 TM$ by taking $\{e_i \wedge e_j | \ 1 \leq i < j \leq n\}$ to be an orthonormal basis for $\Lambda^2 T_x M$ whenever $\{e_1, \ldots, e_n\}$ is an orthonormal basis for $T_x M$.

A 2-plane which can be expressed in the form $u \wedge v$ for some vectors u and v is called a *simple* 2-plane, and such a plane corresponds to a subspace of $T_x M$.

Exercise 16.5.1 Show that every 2-plane is simple if $n = 2$ or $n = 3$, but not if $n \geq 4$.

The importance of the 2-plane bundle is the following:

Proposition 16.5.2 *The curvature tensor defines a symmetric bilinear form on the space of 2-planes $\Lambda^2 T_x M$, by*

$$R(A^{ij} e_i \wedge e_j, B^{kl} e_k \wedge e_l) = A^{ij} B^{kl} R_{ijkl}.$$

Here the sum is over all i and j with $i < j$, and all k and l with $k < l$.

In particular, this curvature operator, since symmetric, can be diagonalised. It is important to note that the eigenvalues of the curvature operator need not be sectional curvatures! The sectional curvatures are the values of the curvature operator on simple 2-planes, but there is no reason why the eigen-vectors of the curvature operator should be simple 2-planes. In particular, it is possible to have all the sectional curvatures positive (or negative) at a point, while not having all of the eigenvalues of the curvature operator positive (negative).

In the special case of three dimensions, however, every 2-plane is simple, and so the eigenvalues of the curvature operator are sectional curvatures. In this case we refer to the eigenvectors of the curvature operator as the principal 2-planes, and the eigenvalues the principal sectional curvatures.

16.6 Calculating curvature

Suppose we are given a metric g and wish to calculate the curvature. In principle we have all the ingredients to do this, but in practice this can get very messy:

First, in local coordinates we can write down the connection ceofficients, which are smooth functions on the coordinate domain, such that

$$\nabla_{\partial_i}\partial_j = \Gamma_{ij}{}^k\partial_k.$$

From these we can calculate the second derivatives:

$$\begin{aligned}\nabla_k\nabla_i\partial_j &= \nabla_k\left(\Gamma_{ij}{}^l\partial_l\right)\\ &= \partial_k\Gamma_{ij}{}^l\partial_l + \Gamma_{ij}{}^l\Gamma_{kl}{}^p\partial_p.\end{aligned}$$

This gives the expression for the curvature:

$$R_{ikj}{}^l = \partial_k\Gamma_{ij}{}^l - \partial_i\Gamma_{kj}{}^l + \Gamma_{ij}{}^q\Gamma_{kq}{}^l - \Gamma_{kj}{}^q\Gamma_{iq}{}^l.$$

Let us consider a simple case, namely when the parametrisation is conformal, so that the metric takes the very simple form

$$g_{ij} = f\delta_{ij}$$

for some function f. Then the inverse metric is also easy to compute:

$$g^{ij} = f^{-1}\delta^{ij}.$$

Therefore a connection coefficient is given by

$$\begin{aligned}\Gamma_{ij}{}^k &= \frac{1}{2}g^{kl}\left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right)\\ &= \frac{1}{2}f^{-1}\delta^{kl}\left(\partial_i f\delta_{jl} + \partial_j f\delta_{il} - \partial_l f\delta_{ij}\right)\\ &= \frac{1}{2}\left(\delta_j^k\partial_i\log f + \delta_i^k\partial_j\log f - \delta_{ij}\delta^{kl}\partial_l\log f\right).\end{aligned}$$

This gives the following expression for the curvature tensor components, where we write $u = \log\sqrt{f}$:

$$\begin{aligned}R_{ikj}{}^l &= \delta_j^l\partial_k\partial_i u + \delta_i^l\partial_k\partial_j u - \delta_{ij}\delta^{lp}\partial_k\partial_p u\\ &\quad - \delta_j^l\partial_i\partial_k u - \delta_k^l\partial_i\partial_j u + \delta_{kj}\delta^{lp}\partial_i\partial_p u\\ &\quad + \left(\delta_j^q\partial_i u + \delta_i^q\partial_j u - \delta_{ij}\delta^{qp}\partial_p u\right)\left(\delta_q^l\partial_k u + \delta_k^l\partial_q u - \delta_{kq}\delta^{lm}\partial_m u\right)\\ &\quad - \left(\delta_j^q\partial_k u + \delta_k^q\partial_j u - \delta_{kj}\delta^{qp}\partial_p u\right)\left(\delta_q^l\partial_i u + \delta_i^l\partial_q u - \delta_{iq}\delta^{lm}\partial_m u\right)\\ &= \delta_i^l\partial_k\partial_j u - \delta_{ij}\delta^{lp}\partial_k\partial_p u - \delta_k^l\partial_i\partial_j u + \delta_{kj}\delta^{lp}\partial_i\partial_p u\\ &\quad + |Du|^2\left(\delta_{jk}\delta_i^l - \delta_{ij}\delta_k^l\right)\\ &\quad + \delta_k^l\partial_i u\partial_j u + \delta_{ij}\delta^{lq}\partial_q u\partial_k u - \delta_i^l\partial_j u\partial_k u - \delta_j k\delta^{lq}\partial_q u\partial_i u.\end{aligned}$$

Taking the trace over k and l gives the Ricci curvature:

$$R_{ij} = -\delta_{ij}\left(\Delta u + (n-2)|Du|^2\right) - (n-2)\left(\partial_i\partial_j u - \partial_i u\partial_j u\right),$$

where $\Delta u = \delta^{kl}\partial_k\partial_l u$ is the Laplacian of u. Finally, multiplying by $f^{-1}\delta^{ij}$ gives the scalar curvature:

$$R = -(n-1)f^{-1}\left(2\Delta u + (n-2)|Du|^2\right).$$

Example 16.6.1 We will compute the curvature of the left-invariant metric $g_{ij} = y^{-2}\delta_{ij}$ on the upper half-plane $\mathbb{H}^n$ in $\mathbb{R}^n$, with coordinates $(x_1, \ldots, x_{n-1}, y)$, $y > 0$. If this case we have $f = y^{-2}$, so $u = -\log y$. Therefore $\partial_i u = -y^{-1}\delta_i^n$, and $\partial_i\partial_j u = y^{-2}\delta_i^n\delta_j^n$. Also $|Du|^2 = y^{-2}$. The equation above therefore gives

$$\begin{aligned} R_{ikjl} = y^{-4} \big(&\delta_{il}\delta_j^n\delta_k^n + \delta_{jk}\delta_i^n\delta_l^n - \delta_{ij}\delta_k^n\delta_l^n - \delta_{kl}\delta_i^n\delta_j^n \\ &+ (\delta_{jk}\delta_{il} - \delta_{ij}\delta_{kl}) \\ &+ \delta_{kl}\delta_i^n\delta_j^n + \delta_{ij}\delta_k^n\delta_l^n - \delta_{jk}\delta_i^n\delta_l^n - \delta_{il}\delta_k^n\delta_l^n \big) \\ = y^{-4}\left(\delta_{jk}\delta_{il} - \delta_{ij}\delta_{kl}\right). \end{aligned}$$

An orthonormal basis is given by $\{ye_i\}$, which gives for any sectional curvature

$$K = -1.$$

Thus hyperbolic space has constant negative curvature.

16.7 Left-invariant metrics

Another situation in which is it possible to conveniently write down the curvature of a metric is when it arises as a left-invariant metric for a Lie group.

Let G be a Lie group, with a left-invariant metric g, and an orthonormal basis of left-invariant vector fields $E_1, \ldots, E_n$. Write $[E_i, E_j] = c_{ij}{}^k E_k$. Then we have

$$\nabla_{E_i} E_j = \frac{1}{2}\left(c_{ij}{}^l + c^l{}_{ij} + c^l{}_{ji}\right) E_l,$$

and so

$$\nabla_k \nabla_i E_j = \frac{1}{4}\left(c_{ij}{}^l + c^l{}_{ij} + c^l{}_{ji}\right)\left(c_{kl}{}^p + c^p{}_{kl} + c^p{}_{lk}\right) E_p.$$

Also, we have

$$\nabla_{[E_k, E_i]} E_j = c_{ki}{}^l \nabla_l E_j = \frac{1}{2} c_{ki}{}^l \left(c_{lj}{}^p + c^p{}_{lj} + c^p{}_{jl}\right) E_p.$$

Combining these:

$$\begin{aligned} R_{ikj}{}^p = &\frac{1}{4}\left(c_{ij}{}^l + c^l{}_{ij} + c^l{}_{ji}\right)\left(c_{kl}{}^p + c^p{}_{kl} + c^p{}_{lk}\right) \\ &- \frac{1}{4}\left(c_{kj}{}^l + c^l{}_{kj} + c^l{}_{jk}\right)\left(c_{il}{}^p + c^p{}_{il} + c^p{}_{li}\right) \\ &- \frac{1}{2} c_{ki}{}^l \left(c_{lj}{}^p + c^p{}_{lj} + c^p{}_{jl}\right). \end{aligned}$$

Example 16.7.1 This gives us another method to compute the curvature of the metric on the upper half-plane: We think of this as the Lie group G of matrices of the form

$$\begin{bmatrix} v_n & v_1 & v_2 & \dots & v_{n-1} \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

with $v_n > 0$. We identify this with the upper half-space of $\mathbb{R}^n$ by associating the above matrix with the point $(v_1, \dots, v_n)$. We take the left-invariant metric for which the usual basis at the identity $(0, \dots, 0, 1)$ is orthonormal. The corresponding left-invariant vector fields are:

$$(E_i)_v = \begin{bmatrix} 0 & 0 & \dots & 0 & v_n & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \sim v_n e_i$$

for $1 \le i \le n-1$, and

$$(E_n)_v = \begin{bmatrix} v_n & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \sim v_n e_n.$$

Therefore the corresponding metric is $g_{ij} = x_n^{-2}\delta_{ij}$. The structure coefficients ${c_{ij}}^k$ can be computed as follows: If $1 \le i, j \le n-1$, then $[E_i, E_j] = 0$; If $1 \le i \le n-1$, then

$$[E_i, E_n] = -v_n E_i.$$

Therefore

$${c_{ij}}^k = -\delta_j^n \delta_i^k + \delta_i^n \delta_j^k.$$

This gives (since with respect to the basis $\{E_i\}$, $g_{ij} = \delta_{ij}$)

$$\begin{aligned} R_{ikjl} = &-\frac{1}{4}\left(\delta_j^n\delta_{kp} - \delta_k^n\delta_{jp} + \delta_k^n\delta_{jp} - \delta_p^n\delta_{jk} + \delta_j^n\delta_{kp} - \delta_p^n\delta_{jk}\right) \\ &\times \left(\delta_p^n\delta_{il} - \delta_i^n\delta_{pl} + \delta_i^n\delta_{pl} - \delta_l^n\delta_{ip} + \delta_p^n\delta_{il} - \delta_l^n\delta_{ip}\right) \\ &+ \frac{1}{4}\left(\delta_j^n\delta_{ip} - \delta_i^n\delta_{jp} + \delta_i^n\delta_{jp} - \delta_p^n\delta_{ji} + \delta_j^n\delta_{ip} - \delta_p^n\delta_{ji}\right) \\ &\times \left(\delta_p^n\delta_{kl} - \delta_k^n\delta_{pl} + \delta_k^n\delta_{pl} - \delta_l^n\delta_{kp} + \delta_p^n\delta_{kl} - \delta_l^n\delta_{kp}\right) \\ &+ \frac{1}{2}\left(\delta_k^n\delta_{ip} - \delta_i^n\delta_{kp}\right)\left(\delta_j^n\delta_{lp} - \delta_p^n\delta_{jl} + \delta_p^n\delta_{jl} - \delta_l^n\delta_{jp} + \delta_j^n\delta_{jp}\right) \\ = &-\delta_j^n\delta_k^n\delta_{il} + \delta_j^n\delta_l^n\delta_{ik} + \delta_{il}\delta_{jk} - \delta_i^n\delta_l^n\delta_{jk} \\ &+ \delta_j^n\delta_i^n\delta_{kl} - \delta_j^n\delta_l^n\delta_{ik} - \delta_{kl}\delta_{ij} + \delta_k^n\delta_l^n\delta_{ij} \\ &+ \delta_k^n\delta_j^n\delta_{il} - \delta_k^n\delta_l^n\delta_{ij} - \delta_i^n\delta_j^n\delta_{kl} + \delta_i^n\delta_l^n\delta_{kj} \\ = &-\delta_{kl}\delta_{ij} + \delta_{jk}\delta_{il}. \end{aligned}$$

Hence the curvature operator has all eigenvalues equal to -1, and all of the sectional curvatures are -1.

16.8 Bi-invariant metrics

The situation becomes even simpler if we have a bi-invariant metric:

Proposition 16.8.1 *If g is a bi-invariant metric on a Lie group G, then for any left-invariant vector fields X, Y, and Z,*

$$g([X,Y],Z) = g([Z,X],Y).$$

Proof. g must be left-invariant, so that for any $h \in G$,

$$g_h(D_e l_h(X), D_e l_h(Y)) = g_e(X,Y).$$

Similarly, g must be right-invariant, so

$$g_h(D_e r_h(X), D_e r_h(Y)) = g_e(X,Y).$$

Together these imply that

$$g_e((D_e r_h)^{-1} D_e l_h(X), (D_e r_h)^{-1} D_e l_h(Y)) = g_e(X,Y).$$

We will denote by Ad_h the isomorphism of T_eG given by $(D_e r_h)^{-1} D_e l_h$. Then Ad is a representation of the group G on the vector space T_eG, called the *adjoint* representation: If γ has tangent vector X, then

$$\begin{aligned} Ad_k Ad_h(X) &= \frac{d}{dt} (r_k)^{-1} l_k (r_h)^{-1} l_h(\gamma(t))\Big|_{t=0} \\ &= \frac{d}{dt} \left(kh\gamma(t)h^{-1}k^{-1}\right)\Big|_{t=0} \\ &= Ad_{kh}(X). \end{aligned}$$

So we have $g(Ad_h(X), Ad_h(Y)) = g(X,Y)$, or g is Ad-invariant. But now differentiate this equation with $h = e^{tZ}$, at $t = 0$. In doing so we differentiate the maps Ad_h.

Definition 16.8.2 The derivative at the identity of the map $Ad : G \to GL(T_eG)$ is denoted $ad : T_eG \to L(T_eG, T_eG)$.

The map ad is a representation of the Lie algebra $\mathfrak{g}$ of G: It is a linear map such that

$$[ad(X), ad(Y)] = ad([X,Y]),$$

where on the left hand side the Lie bracket is to be interpreted as the commutator of the linear transformations $Ad(X)$ and $Ad(Y)$.

Now when we differentiate the Ad-invariance condition, we get:

$$0 = g(ad(Z)(X), Y) + g(X, ad(Z)(Y)).$$

Lemma 16.8.3

$$ad(X)Y = [X, Y].$$

Proof. By definition we have

$$ad(X)Y = \frac{d}{dt}\left(Ad_{e^{tX}}Y\right)|_{t=0} = \frac{d}{dt}\frac{d}{ds}\left(e^{tX}e^{sY}e^{-tX}\right)\Big|_{s=t=0}.$$

This makes sense because $\frac{d}{ds}\left(e^{tX}e^{sY}e^{-tX}\right)|_{s=0}$ is a vector in T_eG for each t, and so can be differentiated with respect to t. The identification between this and the Lie bracket comes from a more general statement:

Lemma 16.8.4 *Let M be a smooth manifold, X and Y in $\mathcal{X}(M)$, and $x \in M$. Then if $\Psi_{X,t}$ and $\Psi_{Y,t}$ are local flows of the vector fields X and Y near x,*

$$\frac{\partial}{\partial t}\frac{\partial}{\partial s}\left(\Psi_{Y,t}\Psi_{X,s}\Psi_{Y,-t}(x)\right)\Big|_{s=0}\Big|_{t=0} = [X, Y](x).$$

Proof. In local coordinates near x (say, with $x^k = 0$), we have

$$X^k(z) = X^k(x) + z^i\partial_i X^k(x) + O(z^2).$$

Therefore by definition of the local flow,

$$\begin{aligned}
(\Psi_{X,t}(z))^k &= z^k + t(X^k(x) + z^i\partial_i X^k(x) + O(z^2)) + O(t^2)\\
&= z^k + tX^k(x) + tz^i\partial_i X^k(x) + O(tz^2, t^2).
\end{aligned}$$

From this we can compute:

$$(\Psi_{Y,-t}(x))^k = -tY^k(x) + O(t^2),$$

and

$$(\Psi_{X,s}\Psi_{Y,-t}(x))^k = -tY^k(x) + sX^k(x) - stY^i(x)\partial_i X^k(x) + O(st^2, s^2)$$

and finally,

$$\begin{aligned}
(\Psi_{Y,t}\Psi_{X,s}\Psi_{Y,-t}(x))^k &= -tY^k(x) + sX^k(x) - stY^i(x)\partial_i X^k(x) + O(st^2, s^2)\\
&\quad + tY^k(x) + t(-tY^i(x) + sX^i(x))\partial_i Y^k(x) + O(st^2, t^2)\\
&= sX^k(x) - st\left(Y^i(x)\partial_i X^k(x) - X^i(x)\partial_i Y^k(x)\right)\\
&\quad + O(s^2, t^2).
\end{aligned}$$

Differentiating with respect to s and t when $s = t = 0$ gives

$$\partial_t \partial_s \left(\Psi_{Y,t}\Psi_{X,s}\Psi_{Y,-t}(x)\right)^k \Big|_{s=t=0} = X^i(x)\partial_i Y^k(x) - Y^i(x)\partial_i X^k(x) = \left([X,Y](x)\right)^k .$$

□

In the present case, we have

$$e^{tX} = \Psi_{X,t}(e);$$

And by left-invariance,

$$\frac{d}{dt}\Psi_{X,t}(h) = X_h = D_e l_h X_e = \frac{d}{dt}\left(he^{tX}\right),$$

so that

$$\Psi_{X,t}(h) = he^{tX}.$$

Therefore

$$e^{tX}e^{sY}e^{-tX} = \Psi_{X,-t}\left(e^{tX}e^{sY}\right) = \Psi_{X,-t}\Psi_{Y,s}\Psi_{X,t}(e).$$

So Lemma 16.8.4 gives the result. □

This completes the proof of Proposition 16.8.1. □

Corollary 16.8.5 *If G is a Lie group and g a bi-invariant Riemannian metric, then*

$$R_{ijkl} = \frac{1}{4}c_{ijp}c_{klp} = \frac{1}{4}\left(c_{ikp}c_{jlp} - c_{jkp}c_{ilp}\right).$$

Proof. By Proposition 16.8.1 we have $c_{ijk} = c_{jki}$, and we also have the symmetry $c_{ijk} = -c_{jik}$. Therefore

$$\begin{aligned}\nabla_{E_i}E_j &= \frac{1}{2}g^{kl}\left(c_{ijl} + c_{lij} + c_{lji}\right)E_k\\ &= \frac{1}{2}g^{kl}\left(c_{ijl} + c_{ijl} + c_{jil}\right)E_k\\ &= \frac{1}{2}g^{kl}c_{ijl}E_k.\end{aligned}$$

Therefore

$$\begin{aligned}R_{ikjl} &= g\left(\nabla_k\nabla_i E_j - \nabla_i\nabla_k E_j - \nabla_{[E_k,E_i]}E_j, E_l\right)\\ &= \frac{1}{4}c_{kpl}c_{ijp} - \frac{1}{4}c_{ipl}c_{kjp} - \frac{1}{2}c_{kip}c_{pjl}\end{aligned}$$

Now we note that the Jacobi identity (cf. Lecture 6) gives

$$
\begin{aligned}
0 &= [E_i, [E_j, E_k]] + [E_j, [E_k, E_i]] + [E_k, [E_i, E_j]] \\
&= c_{ipl}c_{jkp} + c_{jpl}c_{kip} + c_{kpl}c_{ijp}
\end{aligned}
$$

which gives the result, including the equality between the two expressions.

□

Example 16.8.6 We will apply this to a simple example, namely the Lie group S^3: Here the metric for which i, j and k are orthonormal at the identity is easily seen to be bi-invariant, and we have the structure coefficients

$$[i, j] = ij - ji = 2k;$$

and similarly $[j, k] = 2i$ and $[k, i] = 2j$. This gives, if we label $E_1 = i$, $E_2 = j$, and $E_3 = k$,

$$c_{123} = c_{231} = c_{312} = 2$$

and

$$c_{213} = c_{321} = c_{132} = -2,$$

and all others are zero. Thus we have

$$R_{1212} = \frac{1}{4}c_{12p}c_{12p} = \frac{1}{4}c_{123}^2 = 1;$$

and similarly $R_{1313} = R_{2323} = 1$. Also

$$R_{1213} = \frac{1}{4}c_{12p}c_{13p} = 0,$$

and similarly $R_{1223} = R_{1323} = 0$. Therefore the curvature operator is just the identity matrix with respect to the basis $E_1 \wedge E_2$, $E_2 \wedge E_3$, $E_3 \wedge E_1$, and all the eigenvalues are equal to 1. In particular all the sectional curvatures are equal to 1.

Lecture 17. Extrinsic curvature of submanifolds

In this lecture we define the extrinsic curvature of submanifolds in Euclidean space.

17.1 Immersed submanifolds

By an *immersed submanifold* of Euclidean space $\mathbb{R}^N$ I will mean a differentiable manifold M together with an immersion $X : M \to \mathbb{R}^N$. Note that for any $x \in M$ there is a neighbourhood U of x such that $X|_U$ is an embedding. A particular case of an immersed submanifold is an embedded submanifold.

The inner product $\langle .,. \rangle$ on $\mathbb{R}^N$ induces a metric g and corresponding Levi-Civita connection ∇ on M, defined by

$$g(u,v) = \langle DX(u), DX(v) \rangle$$

and

$$\nabla_u v = \pi_{TM} \left(D_u(DX(v)) \right).$$

A particular case of this is an *immersed hypersurface*, which is the case where M is of dimension $N-1$. We will develop the theory of extrinsic curvature first for the simpler case of hypersurfaces, and then extend this to the more general case of immersed submanifolds.

17.2 The Gauss map of an immersed hypersurface

Let M^n be an oriented immersed hypersurface in $\mathbb{R}^{n+1}$. Then for each point $x \in M$ there is a well-defined unit normal $\mathbf{n}$ to M (more precisely, to $X(M)$) at x. This is defined by the requirements $\langle \mathbf{n}, \mathbf{n} \rangle = 1$, $\langle n, DX(u) \rangle = 0$ for all $u \in T_xM$, and if $e_1, \ldots, e_n$ are an oriented basis for T_xM then $DX(e_1), \ldots, DX(e_n), \mathbf{n}$ is an oriented basis for $\mathbb{R}^{n+1}$.

This defines a smooth map $\mathbf{n} : M \to S^n \subset \mathbb{R}^{n+1}$, called the Gauss map of M.

17.3 The second fundamental form of a hypersurface

Having defined the Gauss map of an oriented immersed hypersurface, we can define a tensor as follows:

$$h(u,v) = \langle D_u \mathbf{n}, DX(v)\rangle.$$

This is called the second fundamental form on M, and is a tensor of type $(2,0)$.

The second fundamental form has an alternative expression, which we can deduce as follows: Let U and V be smooth vector fields on M. Since $\langle \mathbf{n}, DX(V) = 0$, we have

$$\begin{aligned} 0 &= U\langle \mathbf{n}, DX(V)\rangle \\ &= \langle D_u \mathbf{n}, DX(V)\rangle + \langle \mathbf{n}, D_U DX(V)\rangle \\ &= h(U,V) + \langle \mathbf{n}, D_U DX(V)\rangle \end{aligned}$$

and therefore

$$h(U,V) = -\langle D_U D_V X, \mathbf{n}\rangle.$$

From this we can deduce a useful symmetry:

$$h(U,V) = -\langle D_U D_V X, \mathbf{n}\rangle = -\langle D_V D_U X + D_{[U,V]} X, \mathbf{n}\rangle = -h(V,U)$$

since $D_{[U,V]}X = DX([U,V])$ is tangential to M, hence orthogonal to $\mathbf{n}$. Therefore the second fundamental form is a symmetric bilinear form on the tangent space T_xM at each point.

Since h is symmetric, it can be diagonalized with respect to the metric g — that is, we can find a basis $e_1, \ldots, e_n$ for T_xM and real numbers $\lambda_1, \ldots, \lambda_n$ such that $h(e_i, u) = \lambda_i g(e_i, u)$ for all vectors $u \in T_xM$. The numbers $\lambda_1, \ldots, \lambda_n$ are called the *principal curvatures* of M at x.

The *mean curvature* H is the trace of h with respect to g: $H = g^{ij}h_{ij}$. This can also be expressed in terms of the principal curvatures: $H = \lambda_1 + \ldots + \lambda_n$.

The *Gauss curvature* K is the determinant of h with respect to g, which is therefore also equal to $\prod_{i=1}^n \lambda_i$.

In the case where M is not orientable, it is not possible to choose a unit normal vector continuously on M, and so $\mathbf{n}$, and hence h and the principal curvatures λ_i are defined only up to sign.

Remark. It is very easy to get a geometric understanding of the second fundamental form of a hypersurface: Fix $z \in M$. Assume that the origin of $\mathbb{R}^{n+1}$ is at $X(z)$ and choose an orthonormal basis $e_1, \ldots, e_{n+1}$ for $\mathbb{R}^{n+1}$ such that $DX(T_zM) = \mathrm{span}\{e_1, \ldots, e_n\}$. By the implicit function theorem, $X(M)$ can be written locally in the form $\{x^i e_i : \ x^{n+1} = u(x^1, \ldots, x^n)\}$. Then near z we have in the coordinates $x^1, \ldots, x^n$

$$\partial_i X = e_i + \frac{\partial u}{\partial x^i} e_{n+1}$$

and

$$n(z) = e_{n+1}.$$

Therefore

$$\partial_i \partial_j X = \frac{\partial^2 u}{\partial x^i \partial x^j} e_{n+1}$$

and so at z,

$$h_{ij} = -\frac{\partial^2 u}{\partial x^i \partial x^j}.$$

To put this another way, we have

$$u(y) = -\frac{1}{2} h_{ij}(z) y^i y^j + O(y^3)$$

as $y \to 0$. This says that the second fundamental form gives the best approximation of the hypersurface by a paraboloid defined over its tangent plane.

17.4 The normal bundle of an immersed submanifold

Now we go on to the general case of an immersed submanifold M^n in $\mathbb{R}^N$. Then at each point of M, rather than having a single unit normal vector, we have a normal subspace $N_x M = \{v \in \mathbb{R}^N : v \perp DX(T_x M)\}$. This defines the *normal bundle* NM of M: $NM = \{(p, v) : p \in M, v \perp DX(T_p M)\}$. This is a differentiable manifold of dimension N.

17.5 Vector Bundles

The normal bundle (and indeed the tangent bundle and the tensor bundles we have already defined) is an example of a more general object called a *vector bundle*. A vector bundle E of dimension k over M is defined by associating to each $x \in M$ a vector space E_x (often called the *fibre* at x), and taking $E = \{(p, v) : p \in M, v \in E_p\}$. We require that E be a smooth manifold, and that for each $x \in M$ there is a neighbourhood U of x in M such that there are k smooth sections $\phi_1, \ldots, \phi_k$ of E (i.e. smooth maps ϕ_i from M to E such that $\pi \circ \phi_i = \mathrm{id}$) such that $\phi_1(y), \ldots, \phi_k(y)$ form a basis for E_y for each $y \in U$ (it follows that the restriction of the bundle E to U is diffeomorphic to $U \times \mathbb{R}^k$).

We denote the space of smooth sections of E (i.e. smooth maps from M to E which take each $x \in M$ to the fibre E_x at x) by $\Gamma(E)$.

A connection on a vector bundle E is a map which takes a vector $u \in T_x M$ and section $\phi \in \Gamma(E)$ and gives an element $\nabla_u \phi \in E_x$, smoothly in the sense that if $U \in \mathcal{X}(M)$ and $\phi \in \Gamma(E)$ then $\nabla_U \phi \in \Gamma(E)$, which is linear in the first argument and satisfies a Leibniz rule in the second:

$$\nabla_u(f\phi) = f\nabla_u\phi + u(f)\phi$$

for all $f \in C^\infty(M)$, $u \in T_xM$ and $\phi \in \Gamma(E)$.

We can also define tensors which either act on E or take their values in E, to be C^∞-multilinear functions acting on sections of E or its dual E^*, and the connection extends to such tensors.

17.6 Curvature of a vector bundle

If E is a vector bundle over M with a metric $\langle .,. \rangle$ and a connection ∇ which is compatible with the metric:

$$\nabla_u\langle \phi, \psi \rangle = \langle \nabla_u\phi, \psi \rangle + \langle \phi, \nabla_u\psi \rangle.$$

Then we can define the curvature of the bundle E as follows: If $X, Y \in \mathcal{X}(M)$ and ϕ, $\psi \in \Gamma(E)$, then we take

$$R(X, Y, \phi, \psi) = \langle \nabla_Y\nabla_X\phi - \nabla_X\nabla_Y\phi - \nabla_{[Y,X]}\phi, \psi \rangle.$$

This is tensorial in all arguments — that is, the value of the resulting function when evaluated at any point $x \in M$ depends only on the values of X, Y, ϕ and ψ at x. The proof of this is identical to the proof that the curvature of M is a tensor (Lecture 16). This can be considered as an operator which takes Λ^2T_xM to Λ^2E_x, since it is antisymmetric in the first two and the last two arguments.

17.7 Connection on the normal bundle

We can define a connection on the normal bundle as follows: If V is a section of the normal bundle, and U is a smooth vector field on M, then we define

$$\nabla_UV\big|_x = \pi_{N_xM}\left(D_UV\right).$$

This is a connection: For any $f \in C^\infty(M)$, we have

$$\begin{aligned}
\nabla_U(fV) &= \pi_{NM}\left((Uf)V + fD_UV\right) \\
&= U(f)\pi_{NM}V + f\pi_{NM}D_UV \\
&= U(f)V + f\nabla_UV
\end{aligned}$$

so the Leibniz rule holds. This connection is compatible with the metric induced on NM by the inner product on $\mathbb{R}^N$. By the argument above, this defines a curvature tensor acting on $\Lambda^2TM \otimes \Lambda^2E$, which we denote by $R^\perp$ and call the *normal curvature* of M.

17.8 Second fundamental form of a submanifold

The second fundamental form is defined in an analogous way to that for the hypersurface case: Given $U, V \in \mathcal{X}(M)$ define

$$h(U, V) = -\pi_{N_x M}(D_U D_V X) = -D_U D_V X + DX(\nabla_U V).$$

This does in fact define a tensor field, since

$$\begin{aligned} h(fU, gV) &= -\pi_{N_x M}(D_{fU} D_{gV} X) \\ &= -\pi_{N_x M}(fg D_U D_V X + f(Ug) D_V X) \\ &= fgh(U, V) \end{aligned}$$

since $D_V X \perp N_x M$. h therefore defines at each $x \in M$ a bilinear map from $T_x M \times T_x M$ to $N_x M$.

We can also define an operator $\mathcal{W}$ from $T_x M \times N_x M$ to $T_x M$ as follows:

$$\mathcal{W}(u, \phi) = \pi_{T_x M}(D_u \phi)$$

for $u \in T_x M$ and $\phi \in \Gamma(NM)$. This is again tensorial, since

$$\mathcal{W}(u, f\phi) = \pi_{T_x M}(D_u f\phi) = \pi_{T_x M}(f D_u \phi + (uf)\phi) = f\mathcal{W}(u, \phi).$$

This is related to the second fundamental form as follows:

$$0 = v\langle \phi, D_u X\rangle = \langle D_v \phi, D_u X\rangle + \langle \phi, D_v D_u X\rangle = \langle \mathcal{W}(v, \phi), D_u X\rangle - \langle h(v, u), \phi\rangle$$

and so $\langle \mathcal{W}(v, \phi), D_u X\rangle = \langle h(v, u), \phi\rangle$ for any u and v in $T_x M$ and ϕ in $N_x M$.

The second fundamental form of a submanifold can be interpreted in a similar way to the hypersurface case: If we fix $z \in M$, then $X(M)$ can be written locally as the graph of a smooth function from $T_x M$ to $N_x M$ — that is, if we choose a basis $e_1, \dots, e_N$ such that $DX(T_z M) = \mathrm{span}\{e_1, \dots, e_n\}$ and $N_z M = \mathrm{span}\{e_{n+1}, \dots, e_N\}$, then for some open set U containing z,

$$X(M) = \{X(z) + x^i e_i : \ x^j = f^j(x^1, \dots, x^n), \ j = n+1, \dots, N\}.$$

Then we find

$$f^j(x^1, \dots, x^n) = -\frac{1}{2}\langle h_{kl}(z), e_j\rangle x^k x^l + O(x^3)$$

as $x \to 0$. Thus the second fundamental form at z defines the best approximation to $X(M)$ as the graph of a quadratic function over $DX(T_z M)$.

Lecture 18. The Gauss and Codazzi equations

In this lecture we will prove the fundamental identities which hold for the extrinsic curvature, including the Gauss identity which relates the extrinsic curvature defined via the second fundamental form to the intrinsic curvature defined using the Riemann tensor.

18.1 The fundamental identities

The definitions (from the last lecture) of the connection on the normal and tangent bundles, and the second fundamental form h and the associated operator $\mathcal{W}$, can be combined into the following two useful identities: First, for any pair of vector fields U and V on M,

$$D_U D_V X = -h(U,V) + DX(\nabla_U V). \tag{18.1}$$

This tells us how to differentiate an arbitrary tangential vector field $D_V X$, considered as a vector field in $\mathbb{R}^N$ (i.e. an N-tuple of smooth functions). Then we have a corresponding identity which tells us how to differentiate sections of the normal bundle, again thinking of them as N-tuples of smooth functions: For any vector field U and section ϕ of NM,

$$D_U \phi = DX(\mathcal{W}(U,\phi)) + \nabla_U \phi. \tag{18.2}$$

Since we can think of vector fields in this way as N-tuples of smooth functions, we can deduce useful identities in the following way: Take a pair of vector fields U and V. Applying the combination $UV - VU - [U,V]$ to any function gives zero, by definition of the Lie bracket. In particular, we can applying this to the position vector X:

$$\begin{aligned} 0 &= (UV - VU - [U,V])X \\ &= -h(U,V) + DX(\nabla_U V) + h(V,U) - DX(\nabla_V U) - DX([U,V]). \end{aligned}$$

Since the right-hand side vanishes, both the normal and tangential components must vanish. The normal component is $h(V,U) - h(U,V)$, so this establishes the fact we already knew that the second fundamental form is symmetric. The tangential component is $DX(\nabla_U V - \nabla_V U - [U,V])$, so the vanishing of this tells us that the connection is symmetric (as we proved before).

18.2 The Gauss and Codazzi equations

We will use the same method as above to deduce further important identities, by applying $UV - VU - [U,V]$ to an arbitrary tangential vector field.

Let W be a smooth vector field on M. Then we have

$$
\begin{aligned}
0 &= (UV - VU - [U,V])WX \\
&= U(VWX) - V(UWX) - [U,V]Wx \\
&\overset{(18.1)}{=} U\left(-h(V,W) + (\nabla_V W)X\right) - V\left(-h(U,W) + (\nabla_U W)X\right) \\
&\quad - \big(-h([U,V],W) + (\nabla_{[U,V]}W)X\big) \\
&\overset{(18.2)}{=} -DX(\mathcal{W}(U,h(V,W))) - (\nabla h(U,V,W) + h(\nabla_U V,W) + h(V,\nabla_U W)) \\
&\quad + DX(\mathcal{W}(V,h(U,W))) - (\nabla h(V,U,W) + h(\nabla_V U,W) + h(U,\nabla_V W)) \\
&\quad + U(\nabla_V W)X - V(\nabla_U W)X + h([U,V],W) - (\nabla_{[U,V]}W)X \\
&\overset{(18.1)}{=} -DX(\mathcal{W}(U,h(V,W))) - (\nabla h(U,V,W) + h(\nabla_U V,W) + h(V,\nabla_U W)) \\
&\quad + DX(\mathcal{W}(V,h(U,W))) - (\nabla h(V,U,W) + h(\nabla_V U,W) + h(U,\nabla_V W)) \\
&\quad - h(U,\nabla_V W) + (\nabla_U\nabla_V W)X + h(V,\nabla_U W) - (\nabla_V\nabla_U W)X \\
&\quad + h([U,V],W) - (\nabla_{[U,V]}W)X \\
&= DX\left(R(V,U)W - \mathcal{W}(U,h(V,W)) + \mathcal{W}(V,h(U,W))\right) \\
&\quad + \nabla h(U,V,W) - \nabla h(V,U,W).
\end{aligned}
$$

In deriving this we used the definition of curvature in the last step, and used the symmetry of the connection to note that several of the terms cancel out. The tangential and normal components of the resulting identity are the following:

$$R(U,V)W = \mathcal{W}(U,h(V,W)) - \mathcal{W}(V,h(U,W)) \tag{18.3}$$

and

$$\nabla h(U,V,W) = \nabla h(V,U,W). \tag{18.4}$$

Note that the tensor ∇h appearing here is the covariant derivative of the tensor h, defined by

$$\nabla h(U,V,W) = \nabla_U(h(V,W)) - h(\nabla_U V,W) - h(V,\nabla_U W)$$

for any vector fields U, V and W. Here the ∇ appearing in the first term on the right-hand side is the connection on the normal bundle, and the other two terms involve the connection on TM.

Equation (18.3) is called the Gauss equation, and Equation (18.4) the Codazzi equation. Note that the Gauss equation gives us a formula for the intrinsic curvature of M in terms of the extrinsic curvature h.

It is sometimes convenient to write these identities in local coordinates: Given a local chart with coordinate tangent vectors $\partial_1, \ldots, \partial_n$, choose also a collection of smooth sections e_α of the normal bundle, $\alpha = 1, \ldots, N-n$, which are linearly independent at each point. Let g be the metric on TM and $\tilde{g}$ the metric on NM, and write $g_{ij} = g(\partial_i, \partial_j)$ and $\tilde{g}_{\alpha\beta} = \tilde{g}(e_\alpha, e_\beta)$. Then we can write

$$h(\partial_i, \partial_j) = h_{ij}{}^\alpha e_\alpha$$

and

$$\mathcal{W}(\partial_i, e_\beta) = \mathcal{W}_{i\beta}{}^j \partial_j.$$

The relation between h and $\mathcal{W}$ then tells us that $\mathcal{W}_{i\beta}{}^j = g^{jk} \tilde{g}_{\beta\alpha} h_{ik}{}^\alpha$.

The Gauss identity then becomes

$$R_{ijkl} = \left(h_{jk}{}^\alpha h_{il}{}^\beta - h_{jk}{}^\beta h_{il}{}^\alpha \right) \tilde{g}_{\alpha\beta} \tag{18.5}$$

If we write $\nabla h = \nabla_i h_{jk}{}^\alpha dx^i \otimes dx^j \otimes dx^k \otimes e_\alpha$, then the Codazzi identity becomes

$$\nabla_i h_{jk}{}^\alpha = \nabla_j h_{ik}{}^\alpha. \tag{18.6}$$

Since we already know that h_{ij} is symmetric in j and k, this implies that $\nabla_k h_{ij}$ is totally symmetric in i, j and k.

18.3 The Ricci equations

We will complete our suite of identities by applying $UV - VU - [U,V]$ to an arbitrary section ϕ of the normal bundle:

$$\begin{aligned}
0 &= (UV - VU - [U,V])\phi \\
&\overset{(18.2)}{=} U(\mathcal{W}(V,\phi)X + \nabla_V \phi) - V(\mathcal{W}(U,\phi)X + \nabla_U \phi) \\
&\quad - (\mathcal{W}([U,V],\phi)X + \nabla_{[U,V]}\phi) \\
&\overset{(18.1)}{=} -h(U, \mathcal{W}(V,\phi)) + (\nabla\mathcal{W}(U,V,\phi) + \mathcal{W}(\nabla_U V, \phi) + \mathcal{W}(V, \nabla_U \phi))X \\
&\quad + h(V, \mathcal{W}(U,\phi)) - (\nabla\mathcal{W}(V,U,\phi) + \mathcal{W}(\nabla_V U, \phi) + \mathcal{W}(U, \nabla_V \phi))X \\
&\quad + U\nabla_V \phi - V\nabla_U \phi - \mathcal{W}([U,V],\phi)X - \nabla_{[U,V]}\phi \\
&\overset{(18.2)}{=} -h(U, \mathcal{W}(V,\phi)) + (\nabla\mathcal{W}(U,V,\phi) + \mathcal{W}(V, \nabla_U \phi))X \\
&\quad + h(V, \mathcal{W}(U,\phi)) - (\nabla\mathcal{W}(V,U,\phi) + \mathcal{W}(U, \nabla_V \phi))X \\
&\quad + \mathcal{W}(U, \nabla_V \phi)X + \nabla_U \nabla_V \phi - \mathcal{W}(V, \nabla_U \phi)X - \nabla_V \nabla_U \phi - \nabla_{[U,V]}\phi \\
&= R^\perp(V,U)\phi - h(U, \mathcal{W}(V,\phi)) + h(V, \mathcal{W}(U,\phi)) \\
&\quad + \left(\nabla\mathcal{W}(U,V,\phi) - \nabla\mathcal{W}(V,U,\phi)\right) X.
\end{aligned}$$

As before, this gives us two sets of identities, one from the tangential component and one from the normal component. In fact the tangential component is just the Codazzi identities again, but the normal component gives a new identity, called the *Ricci identity*, which expresses the curvature of the normal bundle in terms of the second fundamental form:

$$R^{\perp}(U,V)\phi = h(V,\mathcal{W}(U,\phi)) - h(U,\mathcal{W}(V,\phi)). \tag{18.7}$$

In local coordinates, with a local basis for the normal bundle as above, this identity can be written as follows: If we write $R^{\perp}_{ij\alpha\beta} = \tilde{g}(R^{\perp}(\partial_i,\partial_j)e_\alpha, e_\beta)$ and $h_{ij\alpha} = \tilde{g}(h(\partial_i,\partial_j),e_\alpha)$, then we have

$$R^{\perp}_{ij\alpha\beta} = g^{kl}(h_{ik\alpha}h_{jl\beta} - h_{jk\alpha}h_{il\beta}). \tag{18.8}$$

18.4 Hypersurfaces

In the case of hypersurfaces the identities we have proved simplify somewhat: First, since the normal bundle is one-dimensional, the normal curvature vanishes and the Ricci equations become vacuous.

Also, the basis $\{e_\alpha\}$ for NM can be taken to consist of the single unit normal vector $\mathbf{n}$, and the Gauss and Codazzi equations become

$$R_{ijkl} = h_{ik}h_{jl} - h_{jk}h_{il}$$

and

$$\nabla_i h_{jk} = \nabla_j h_{ik}.$$

The curvature tensor becomes rather simple in this setting: At any point $x \in M$ we can choose local coordinates such that $\partial_1,\dots,\partial_n$ are orthonormal at x and diagonalize the second fundamental form, so that

$$h_{ij} = \begin{cases} \lambda_i, & i = j \\ 0, & i \neq j \end{cases}$$

Then we have an orthonormal basis for the space of 2-planes $\Lambda^2 T_xM$, given by $\{e_i \wedge e_j :\ i < j\}$. The Gauss equation gives for all $i<j$ and $k<l$

$$\mathrm{Rm}(e_i\wedge e_j, e_k \wedge e_l) = \begin{cases} \lambda_i\lambda_j, & i = k,\ j = l \\ 0, & \text{otherwise.} \end{cases}$$

In particular, this basis diagonalizes the curvature operator, and the eigenvalues of the curvature operator are precisely $\lambda_i\lambda_j$ for $i<j$. Note that all of the eigenvectors of the curvature operator are simple planes in $\Lambda^2 TM$. It follows that if M is any Riemannian manifold which has a non-simples 2-plane as a eigenvector of the curvature operator at any point, then M cannot be immersed (even locally) as a hypersurface in $\mathbb{R}^{n+1}$.

Example 18.4.1 (Curvature of the unit sphere). Consider the unit sphere S^n, which is a hypersurface of Euclidean space $\mathbb{R}^{n+1}$ (so we take the immersion X to be the inclusion). With a suitable choice of orientation, we find that the Gauss map is the identity map on S^n — that is, we have $\mathbf{n}(z) = X(z)$ for all $z \in S^n$. Differentiating this gives

$$DX(\mathcal{W}(u)) = D_u\mathbf{n} = DX(u)$$

so that $\mathcal{W}$ is the identity map on TS^n, and $h_{ij} = g_{ij}$. It follows that all of the principal curvatures are equal to 1 at every point, and that all of the sectional curvatures are equal to 1. Therefore the unit sphere has constant sectional curvatures equal to 1.

Example 18.4.2 (Totally umbillic hypersurfaces). A hypersurface M^n in Euclidean space is called *totally umbillic* if for every $x \in M$ the principal curvatures $\lambda_1(x), \ldots, \lambda_n(x)$ are equal — that is, the second fundamental form has the form $h_{ij} = \lambda(x)g_{ij}$. The Codazzi identity implies that a connected totally umbillic hypersurface in fact has constant principal curvatures (hence also constant sectional curvatures by the Gauss identity):

$$(\nabla_k\lambda)g_{ij} = \nabla_k h_{ij} = \nabla_i h_{kj} = (\nabla_i\lambda)g_{kj}$$

for any i, j and k. Fix k, and choose $j = i \neq k$ (we assume $n \geq 2$, since otherwise the totally umbillic condition is vacuous). This gives $\nabla_k\lambda = 0$. Since k is arbitrary, this implies $\nabla\lambda = 0$, hence λ is constant. There are two possibilities: $\lambda = 0$ (in which case M is a subset of a plane), or $\lambda \neq 0$ (in which case M is a subset of a sphere).

Example 18.4.3 Spacelike hypersurfaces in Minkowski space The definitions we have made for the second fundamental form were given for submanifolds of Euclidean space. However the same definitions work with very minor modifications for certain hypersurfaces in the Minkowski space $\mathbb{R}^{n,1}$: A hypersurface M^n in $\mathbb{R}^{n,1}$ is called *spacelike* if the metric induced on M from the Minkowski metric is Riemannian — equivalently, if every non-zero tangent vector u of M has $\langle u, u\rangle_{\mathbb{R}^{n,1}} > 0$. This is equivalent to the statement that M is given as the graph of a smooth function with slope less than 1 over the plane $\mathbb{R}^n \times \{0\}$.

Given a spacelike hypersurface, we can choose at each $x \in M$ a unit normal by taking the unique future-pointing vector $\mathbf{n}$ which is orthogonal to T_xM with respect to the Minkowski metric, normalized so that $\langle \mathbf{n}, \mathbf{n}\rangle = -1$.

The definitions are now identical to those for hypersurfaces in Euclidean space, except that the metric on the normal bundle is now negative definite. The Codazzi identity is unchanged, and the Gauss identity is almost unchanged: The one difference arises from the presence of the $\tilde{g}$ term in Equation (18.5), which is now negative instead of positive. This gives

$$R_{ijkl} = -\left(h_{ik}h_{jl} - h_{jk}h_{il}\right). \tag{18.9}$$

We can now compute the second fundamental form for a very special example, namely Hyperbolic space $\mathbb{H}^n$, which is the set of unit future timelike vectors in Minkowski space. In this case we have $\mathbf{n}(z) = z$ for every $z \in \mathbb{H}^n$, since $0 = D_u\langle z, z\rangle = 2\langle D_u z, z\rangle = 2\langle u, z\rangle$. Differentiating, we find (exactly as in the calculation for the sphere in Example 18.4.1 above) that $h_{ij} = g_{ij}$, so that the principal curvatures are all equal to 1. The Gauss equation therefore gives that the sectional curvatures are identically equal to -1.

More generally, a hypersurface with all principal curvatures of the same sign (i.e. a convex hypersurface) in Minkowski space has negative curvature operator, while a hypersurface with all principal curvatures of the same sign in Euclidean space has positive curvature operator.

Lecture 19. The Cartan Moving Frame Method

In this lecture we will explore the use of differential forms to compute connection coefficients and curvatures for given Riemannian metrics.

We start with a technical Lemma (Cartan's Lemma):

Lemma 19.1 *Let ω_i, $i = 1, \dots, n$ be a collection of 1-forms on a region U which form a basis for the space of covectors at each point. Suppose η_{ij}, $1 \leq i, j \leq n$ be a collection of 1-forms which satisfy*

$$\sum_{j=1}^{n} \eta_{ij} \wedge \omega_j = 0$$

for each i, and

$$\eta_{ij} + \eta_{ji} = 0$$

for each i and j. Then $\eta_{ij} = 0$ for all i and j at every point of U.

Proof. We can write uniquely

$$\eta_{ij} = \sum_k a_{ijk} \omega_k.$$

The identity $\eta_{ij} \wedge \omega_j$ then becomes

$$a_{ijk} - a_{ikj} = 0$$

for each i, j and k. The second identity gives

$$a_{ijk} + a_{jik} = 0.$$

Thus we have

$$a_{ijk} = -a_{jik} = -a_{jki} = a_{kji} = a_{kij} = -a_{ikj} = -a_{ijk},$$

so that $a_{ijk} = 0$ and $\eta_{ij} = 0$. □

To proceed, suppose we have a manifold M equipped with a Riemannian metric g, such that there is a smoothly defined orthonormal collection of vector fields $e_1, \dots, e_n$. Let $\omega_1, \dots, \omega_n$ be the dual basis of 1-forms.

Then we have the following result:

Proposition 19.2 *There exists a unique collection of 1-forms ω_{ij} for $1 \leq i, j \leq n$ such that*

$$d\omega_i = \omega_{ij} \wedge \omega_j$$

and

$$\omega_{ij} + \omega_{ji} = 0.$$

Proof. We start by proving uniqueness: Suppose that $\bar{\omega}_{ij}$ is any other collection of one-forms satisfying the same conditions. Then let $\eta_{ij} = \bar{\omega}_{ij} - \omega_{ij}$. Then we have

$$\eta_{ij} \wedge \omega_j = \bar{\omega}_{ij} \wedge \omega_j - \omega_{ij} \wedge \omega_j = d\omega_i - d\omega_i = 0$$

and

$$\eta_{ij} + \eta_{ji} = 0.$$

By Lemma 19.1, we have $\eta_{ij} = 0$, and therefore $\omega_{ij} = \bar{\omega}_{ij}$.

Now we prove existence. We set $\omega_{ij} = g(\nabla_{e_k} e_i, e_j)\omega_k$, where ∇ is the Levi-Civita connection of g. We use the identity

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

from section 13.7. Applying this with $\omega = \omega_i$, $X = e_j$ and $Y = e_k$, we find (noting $\omega_i(e_j) = \delta_{ij}$) that

$$d\omega_i(e_k, e_l) = -\omega_i([e_k, e_l]) = -\omega_i(\nabla_{e_k} e_l - \nabla_{e_l} e_k) = -\Gamma_{kli} + \Gamma_{lki}$$

and

$$\omega_{ij} \wedge \omega_j(e_k, e_l) = g(\nabla_{e_k} e_i, e_j)\delta_{jl} - g(\nabla_{e_l} e_i, e_j)\delta_{jk} = \Gamma_{kil} - \Gamma_{lik}.$$

The fact that these two agree follows from the compatibility of the connection with the metric, which gives

$$0 = e_k \delta_{ij} = \Gamma_{kij} + \Gamma_{kji}.$$

The same identity shows that $\omega_{ij} + \omega_{ji} = 0$. □

The 1-forms ω_{ij} are called the *connection 1-forms* corresponding to the frame $\{e_i\}$. Once these have been computed, we can compute the curvature as follows:

Proposition 19.3 *For any orthonormal frame $\{e_i\}$, the curvature 2-form*

$$\Omega_{ij} = -\frac{1}{2} R_{ijkl} \omega_k \wedge \omega_l$$

can be computed from the connection 1-forms as follows:

$$\Omega_{ij} = d\omega_{ij} - \omega_{ik} \wedge \omega_{kj}.$$

Proof. From the proof of the previous proposition, we have $\omega_{ij}(e_k) = g(\nabla_{e_k} e_i, e_j)$. Taking the exterior derivative, we find

$$\begin{aligned}
d\omega_{ij}(e_k, e_l) &= e_k \omega_{ij}(e_l) - e_l \omega_{ij}(e_k) - \omega_{ij}([e_k, e_l]) \\
&= e_k g(\nabla_{e_l} e_i, e_j) - e_l g(\nabla_{e_k} e_i, e_j) - g(\nabla_{[e_k, e_l]} e_i, e_j) \\
&= g(\nabla_{e_k} \nabla_{e_l} e_i, e_j) + g(\nabla_{e_l} e_i, \nabla_{e_k} e_j) \\
&\quad - g(\nabla_{e_l} \nabla_{e_k} e_i, e_j) - g(\nabla_{e_k} e_i, \nabla_{e_l} e_j) \\
&\quad - g(\nabla_{[e_k, e_l]} e_i, e_j) \\
&= -R_{ijkl} + \omega_{ip}(e_l)\omega_{jp}(e_k) - \omega_{ip}(e_k)\omega_{jp}(e_l) \\
&= (\Omega_{ij} + \omega_{ip} \wedge \omega_{pj})(e_k, e_l)
\end{aligned}$$

□

An important special case to keep in mind is the following: If $n = 2$, then we have the simple set of equations

$$\begin{aligned}
d\omega_1 &= \omega_{12} \wedge \omega_2 \\
d\omega_2 &= -\omega_{12} \wedge \omega_1 \\
\Omega_{12} &= d\omega_{12}.
\end{aligned}$$

In this case it is very easy to find the connection 1-form ω_{12}, since if we write $\omega_{12} = a\omega_1 + b\omega_2$, then

$$d\omega_1(e_1, e_2) = a$$

and

$$d\omega_2(e_1, e_2) = -b$$

so that

$$\omega_{12} = d\omega_1(e_1, e_2)\omega_1 - d\omega_2(e_1, e_2)\omega_2.$$

The connection 2-form is also particularly simple in this case, since the curvature tensor has only one component up to symmetries: $R_{1212} = K$, the Gauss curvature. Thus we have

$$\Omega_{12} = -K\omega_1 \wedge \omega_2.$$

Example 19.4 We will compute the curvatures of the Riemannian metric $g_{ij} = f^2 \delta_{ij}$ on a region of $\mathbb{R}^n$, where $f = f(x^n)$.

Here we have an obvious orthonormal frame given by $e_i = f^{-1}\partial_i$, and the corresponding basis of 1-forms $\omega_i = f dx^i$. Computing exterior derivatives, we find

$$d\omega_i = d(f dx^i) = f' dx^n \wedge dx^i.$$

This gives the equations

$$\omega_{ij} \wedge \omega_j = f'/f^2 \omega_n \wedge \omega_i$$

for $i < n$, and

$$\omega_{nj} \wedge \omega_j = 0.$$

Now define 1-forms η_{ij} by taking $\eta_{ij} = \omega_{ij}$ for $i < j < n$, $\eta_{in} = \omega_{in} + f'/f^2 \omega_i$ for $i < n$, and requiring $\eta_{ij} + \eta_{ji} = 0$. Then the equations read:

$$\eta_{ij} \wedge \omega_j = 0.$$

Therefore by Cartan's Lemma, $\eta_{ij} = 0$ everywhere, and we deduce that $\omega_{ij} = 0$ for $i < j < n$ and $\omega_{in} = -f'/f^2 \omega_i$.

Taking exterior derivatives, we find:

$$d\omega_{ij} = 0$$

for $i < j < n$, and

$$d\omega_{in} = -(f'/f)'/f^2 \omega_n \wedge \omega_i.$$

Also we have

$$\omega_{ik} \wedge \omega_{kn} = 0$$

since the sum over k has either $k = n$, hence $\omega_{kn} = 0$, or $k < n$, hence $\omega_{ik} = 0$. On the other hand we have

$$\omega_{ik} \wedge \omega_{kj} = \omega_{in} \wedge \omega_{nj} = -(f')^2/f^4 \omega_i \wedge \omega_j.$$

Combining these identities, we find

$$\Omega_{ij} = (f')^2/f^2 \omega_i \wedge \omega_j$$

for $i < j < n$, and

$$\Omega_{in} = (f'/f)'/f^2 \omega_i \wedge \omega_n.$$

This shows that

$$R_{inin} = -(f'/f)'/f^2,$$

for $1 \leq i \leq n-1$, and and

$$R_{ijij} = -(f')^2/f^4$$

for $1 < i < j < n$, while (except for symmetries) all other curvature components are zero.

An important special case of this example is where $f(x) = x^{-1}$. Then we find $(f'/f)'/f^2 = 1$ and $(f')^2/f^4 = 1$, and therefore all sectional curvatures of this metric are equal to -1.

Lecture 20. The Gauss-Bonnet Theorem

In this lecture we will prove two important global theorems about the geometry and topology of two-dimensional manifolds. These are the Gauss-Bonnet theorem and the Poincaré-Hopf theorem.

Let us begin with a special case:

Suppose M is a compact oriented 2-dimensional manifold, and assume that there exists a vector field V on M which is nowhere zero. Now let g be any Riemannian metric on M. Then we can produce an orthonormal frame globally on M by taking $e_1 = V/g(V,V)^{1/2}$, and taking e_2 to be the unique unit vector orthogonal to e_1 such that e_1 and e_2 are an oriented basis at each point. The dual basis ω_1, ω_2 is then also globally defined, and we have a globally defined 1-form ω_{12} on M such that $d\omega_1 = \omega_{12} \wedge \omega_2$ and $d\omega_2 = -\omega_{12} \wedge \omega_1$. Then we have $-K(g)\omega_1 \wedge \omega_2 = d\omega_{12}$.

Stokes' Theorem then applies to give

$$-\int_M K(g) dVol(g) = \int_M \Omega_{12} = 0.$$

So we have the remarkable conclusion that the integral of the Gauss curvature over M is zero for every Riemannian metric on M.

Remark. In fact the existence of a non-vanishing vector field implies M is diffeomorphic to a torus $S^1 \times S^1$. A corollary of the computation above is that S^2 does not admit a nonvanishing vector field (as we have already seen), since if it did the integral of Gauss curvature would be zero for any metric, but we know that the standard metric on S^2 has Gauss curvature 1.

.

The result we proved above is a special case of the famous Gauss-Bonnet theorem. The general case is as follows:

Theorem 20.1 *The Gauss-Bonnet Theorem Let M be a compact oriented two-dimensional manifold. Then for any Riemannian metric g on M,*

$$\int_M K(g)\, dVol(g) = 2\pi\chi(M)$$

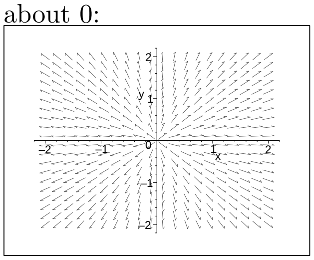
where $\xi(M)$ is the Euler characteristic of M.

Here the Euler characteristic is an integer associated to M which depends only on the topological type of M. If M admits a triangulation (i.e. a decomposition into diffeomorphic images of closed triangles, such that the intersection of any two is either empty, a common vertex, or a common edge) then the Euler characteristic is equal to $V+F-E$, where V is the number of vertices in the triangulation, F is the number of faces, and E is the number of edges.

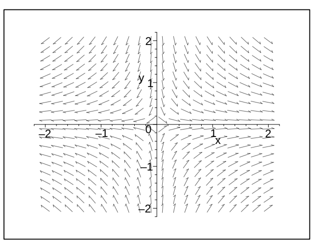
The second important theorem we will prove gives an alternative method of computing the Euler characteristic. To state the result we first need to introduce some new definitions.

Definition 20.2 Let M be a two-dimensional oriented manifold, and V a vector field on M. Suppose $y \in M$, and assume that there exists an open set U containing y such that there are no zeroes of V in $U\backslash\{y\}$. Then the index $I(V,y)$ of V about y is equal to the winding number of V around the boundary of a simply connected neighbourhood of y in U.

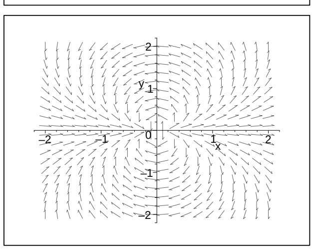
Example 20.3 Here are some vector fields in the plane with nontrivial index about 0:



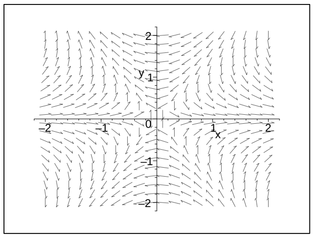
$V=(x,y)$, $I(V,0)=+1$



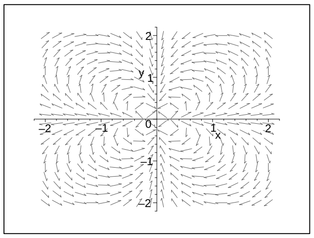
$V=(x,-y)$, $I(V,0)=-1$



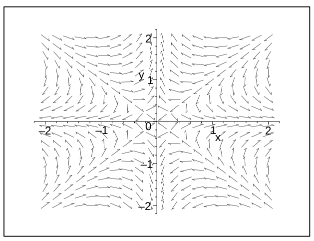
$V=(x^2-y^2,2xy)$, $I(V,0)=+2$



$V = (x^2 - y^2, -2xy)$, $I(V, 0) = -2$



$V = (x^3 - 3xy^2, 3x^2y - y^3)$, $I(V, 0) = +3$



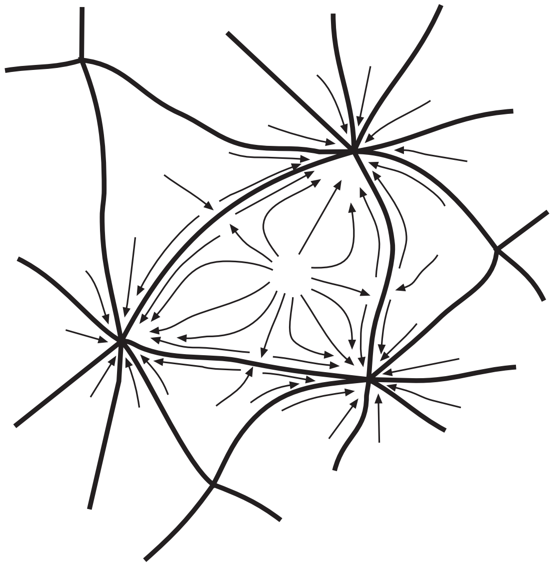
$V = (x^3 - 3xy^2, y^3 - 3x^2y)$, $I(V, 0) = -3$

Now we can state the Poincaré-Hopf theorem:

Theorem 20.4 *Let M be a compact oriented two-dimensional manifold, and let V be any vector field on M which has all zeroes isolated — that is, if y is any zero of V then there is a neithbourhood U of y in M such that V is nonzero on $U \backslash \{y\}$. Label the zeroes of V as $x_1, \dots, x_N$. Then*

$$\sum_{i=1}^{N} I(V, x_i) = \chi(M).$$

We will prove that the sum of the indices of the vector field V at its zeroes is independent of V, and the result can then be taken to define the Euler characteristic. To see that this definition of the Euler characteristic agrees with the one given above in terms of triangulations, suppose that we are given a triangulation of M. Then choose a vector field as sketched below with zeroes at the centre of each face, the middle of each edge, and at each vertex. This vector field points outwards from the centre of each face, so has index 1 there, and points inwards at each vertex, so also has index 1 there. At the middle of each edge the vector field has index -1. Therefore the sum of the indices equals $F + V - E$, agreeing with our previous definition of the Euler characteristic.



We will prove the Gauss-Bonnet theorem and the Poincaré-Hopf theorem at the same time, by showing that for any Riemannian metric g on M and any vector field V with isolated zeroes, we have

$$\int_M K(g)\, dVol(g) = 2\pi \sum_{i=1}^{N} I(V, x_i).$$

If we keep V fixed and vary g, we deduce that the left-hand side is independent of g, and if we keep g fixed and vary V, we deduce that the right-hand side is independent of V, so we can define the Euler characteristic and deduce both results.

So let V and g be given, and let $x_1, \dots, x_N$ be the zeroes of V. About each of these zeroes we can choose a small neighbourhood U_i of x_i (say, given by the image of a ball under some chart) such that V is nonvanishing on $\tilde{M} = M \setminus \cup_{i=1}^{N} U_i$. Let γ_i be the boundary of U_i, parametrised anticlockwise in some oriented chart.

Now $\tilde{M}$ is a compact oriented manifold with boundary, and V is nonvanishing on $\tilde{M}$. Define $e_1 = V/g(V,V)^{1/2}$, and take e_2 to be the unit vector orthogonal to e_1 which is given by a positive right-angle rotation of e_1. Denote the dual frame by ω_1 and ω_2, and let ω_{12} be the corresponding connection one-form. Then we have

$$\Omega_{12} = d\omega_{12},$$

and Stokes' theorem gives

$$-\int_{\tilde{M}} K(g)\, dVol(g) = \int_{\tilde{M}} \Omega_{12} = \int_{\partial \tilde{M}} \omega_{12}.$$

The boundary of $\tilde{M}$ is the union of the circles γ_i, parametrized clockwise. Therefore

$$-\int_{\tilde{M}} K(g)\, dVol(g) = -\sum_{i=1}^{N} \int_{\gamma_i} \omega_{12}.$$

Now on each of the regions U_i we can choose a non-vanishing frame $\bar{e}_1^{(i)}$, $\bar{e}_2^{(i)}$. This gives a corresponding connection one-form $\bar{\omega}_{12}^{(i)}$, and we have for each i

$$-\int_{U_i} K(g)\, dVol(g) = \int_{U_i} \Omega_{12} = \int_{U_i} d\bar{\omega}_{12}^{(i)} = \int_{\gamma_i} \bar{\omega}_{12}^{(i)}.$$

Combining these results, we get

$$-\int_{M} K(g)\, dVol(g) = \sum_{i=1}^{N} \int_{\gamma_i} \bar{\omega}_{12}^{(i)} - \omega_{12}.$$

To interpret the integral of $\bar{\omega}_{12}^{(i)} - \omega_{12}$ around γ_i, we need to consider the relationship between the frames e_1 and $\bar{e}_1^{(i)}$. Since these are both unit vectors, they are related by a rotation. Therefore there exists a map from γ_i to $SO(1) \simeq S^1 \subset \mathbb{R}^2$, say $x \mapsto z = (\alpha, \beta) \in S^1$, such that

$$\begin{aligned}
\bar{e}_1^{(i)} &= \alpha e_1 + \beta e_2 \\
\bar{e}_2^{(i)} &= -\beta e_1 + \alpha e_2.
\end{aligned}$$

Then we have

$$\begin{aligned}
\bar{\omega}_1^{(i)} &= \alpha \omega_1 + \beta \omega_2 \\
\bar{\omega}_2^{(i)} &= -\beta \omega_1 + \alpha \omega_2.
\end{aligned}$$

Since $\alpha^2 + \beta^2 = 1$, we have $\beta^{-1} d\alpha = -\alpha^{-1} d\beta = d\theta$. Applying the exterior derivative we find

$$\begin{aligned}
d\bar{\omega}_1^{(i)} &= d\alpha \wedge \omega_1 + \alpha d\omega_1 + d\beta \wedge \omega_2 + \beta d\omega_2 \\
&= -(\omega_{12} - \beta^{-1} d\alpha) \wedge \beta\omega_1 + (\omega_{12} + \alpha^{-1} d\beta) \wedge \alpha\omega_2 \\
&= (\omega_{12} + d\theta) \wedge \bar{\omega}_2^{(i)}.
\end{aligned}$$

and similarly

$$d\bar{\omega}_2^{(i)} = -(\omega_{12} + d\theta) \wedge \bar{\omega}_1^{(i)}.$$

It follows that $\bar{\omega}_{12}^{(i)} - \omega_{12} = d\theta$, and so

$$\int_{\gamma_i} \bar{\omega}_{12}^{(i)} - \omega_{12} = \int_{\gamma_i} d\theta$$

is 2π times the winding number of the map $z : \gamma_i \to S^1$. This is equal to the difference betweeen $I(\bar{e}_1^{(i)}, x_i)$ and $I(e_1, x_i)$. But $\bar{e}_1^{(i)}$ is nonvanishing in U_i, so $I(\bar{e}_1^{(i)}, x_i) = 0$, and by construction $I(e_1, x_i) = I(V, x_i)$. Thus

$$\int_{\gamma_i} \bar{\omega}_{12}^{(i)} - \omega_{12} = -I(V, x_i).$$

We have proved

$$-\int_M K(g)\, dVol(g) = -2\pi \sum_{i=1}^{N} I(V, x_i),$$

and the proofs of the Gauss Bonnet and Poincaré-Hopf theorems are complete.