MA 109a: Introduction to Geometry and geometric analysis

Oliver Knill, Room 172, Tel.: 4325, e-mail: knill@cco.caltech.edu

This is the first part of an introduction into Geometry and geometric analysis, (tought the first term 1995 at Caltech).

Contents

Preliminaries

0.1 Introduction

What is geometry?

Since the classification of mathematical branches is partly fashion and time dependent (the mathematical Review classification is just reorganized these days), it is also not easy to say what part of mathematics is **geometry**.

• The original meaning of geometry goes back to the pre-Greek antiquity: **measurement of the earth**. Moreover, in ancient Greek, all mathematics was considered to be geometry.

• Felix Klein's 100 year-old but still quite modern view (the Erlanger program) was to classify mathematics and especially Geometry in terms of its **group of isomorphisms**.

• Today, **geometry** is mainly used in the sense of differential topology, the study of differentiable manifolds.

Geometry and other parts of Mathematics

It is usually difficult and maybe useless to define the boundaries of a field in a scientific structure. One can take any major field of mathematics and view its branches influencing in all other major fields. An indication of this fact is that one could take any ordered pair of the set of topics {"algebra", "geometry", "analysis", "probability theory","number theory"} and form for example the pair $\{geometric, \ analysis\}$ which defines itself a branch of mathematics like in the example **"geometric analysis"** or **analytic geometry**. There are other branches of geometry like Differential geometry, the study of Riemannian manifolds, Algebraic geometry, the study of varieties=algebraic manifolds, Symplectic geometry, the study of symplectic manifolds, Geometry of Gauge fields, differential geometry on fibre bundles, Spectral Geometry, the spectral theory of differential operators on a manifold. Noncommutative Geometry, geometry on strange spaces like fractals, foliations or discrete manifolds.

Why is geometry important?

One main reason is certainly that many **classical physical theories** can be described in a purely geometrical way: examples are **classical mechanics**, electromagnetism and other **gauge field theories**, **general relativity** or the **standard model in particle physics**.

About this course

This course is the first part of an introduction into geometry and geometric analysis. It does not rely on a specific book. The material treated in this term is covered quite well by the books [14] and [1] (no wonder, these two sources have together almost 2000 pages).

0.2 Some problems in geometry

We give a selected list of problems which belong to the realm of geometric analysis. We hope to give with this list a first glimpse what happens in this subject. It is not so important to understand all the problems now. You will encounter one or the other during the course or in the second or third part of the course.

1) The **Ham-Sandwich-Salad** problem: given a sandwich consisting of bread, salad and ham. Can you cut the sandwhich into two pieces such that both parts have the same amount of ham and the same amount of bread and the same amount of salad? A mathematical reformulation is: given three sets $B, H, S \subset \mathbf{R}^3$. Is there a plane in $\mathbf{R}^3$ such that the volumes of H and B and S are the same on each part separated by the plane?

Remark: the solution of this problem is called the **Stone-Tukey theorem** and is a pleasant application of the **Borsuk-Ulam theorem**, which says that for any continuous mapping $f : S \to \mathbf{R}^2$ from a sphere S to the plane, there is a point p in S such that $f(p) = f(p^*)$, where p^* is the antipode of p.

Example: Borsuk-Ulam's theorem implies for example that there exists always two antipodal points on the earth which have both the same temperature and same pressure.

Explanation. Let us explain, how the more abstract theorem of Borsuk-Ulam gives the solution of the Ham-Sandwich-Salad problem: for any set X and any line l, there exists a point $P(l, X)$ which lies on l and such that the plane perpendicular to l through $P(l, X)$ cuts X into two pieces of the same volume. Define a map f_X from the sphere S to $\mathbf{R}$ by defining $f_X(x) = |x - P(l(x), X)|$, where $l(x)$ is the line through x and the origin. Clearly $f_X(x^*) = 2 - f_X(x)$. Define now the map $g : S \to \mathbf{R}^2$ by $g(x) = (f_B(x) - f_S(x), f_H(x) - f_S(x))$. By Borsuk-Ulam, there exists a point p on the sphere, such that $g(p) = g(p^*)$. Since $g(x) = -g(x^*)$ for all x, we have $g(p) = 0$. This however assures that $f_B(p) = f_H(p) = f_H(p)$.

The proof of the Borsuk-Ulam theorem can be done using the notion of the **degree** of a continuous map f from one manifold to an other manifold.

2) **The annulus problem and the Schönflies problem**: given two centered standard balls B_1, B_2 in $\mathbf{R}^d$ of radius 1 and 2. The region between B_1 and B_2 is called an **annulus**. The problem is to show: for a smooth embedding of a d-dimensional ball into a d-dimensional ball, the region between these two regions is homeomorphic to the annulus. This has been shown in dimensions $d \neq 4$.
The requirement of smoothness was important and we want to illustrate, why it can't be dropped: consider the **Alexander horned sphere** inside the ball in $\mathbf{R}^3$ (I'm not redrawing it here, since I'm sure it can be found on the web). Every closed curve in a three dimensional annulus can be deformed to any other closed curve. This is not true for the region between the horned sphere and the sphere: put a loop around one horn. We can not take it out without breaking it. But the horned sphere is homeomorphic to the three dimensional ball. However, this embedding is not smooth at accumulation points of the horns.
The Alexander horned sphere served also as a counter example of the **Schönflies problem**: is every $d-1$ dimensional sphere which is embedded in $\mathbf{R}^n$ the boundary of an embedded n dimensional disk? In dimension 2 this is true (theorem of Schönflies) and sharpens the Jordan curve theorem.

3) **The knot classification problem**: A **knot** is an embedding of the circle in $\mathbf{R}^3$, which means that it is a closed curve in space. Two knots are called equivalent, if they can be deformed into each other. The problem is to classify all knots. The problem is a part of classifying three dimensional manifolds because the complement of a knot is a three dimensional manifold and these manifolds are homemomorphic if and only if the knots are. Maxwell and others constructed a theory which suggested that elementary particles are made of knots. This motivated the classification even if not much later, Maxwell's idea was given up. The theory of knots has astonishing relations with

topological quantum field theory.

4) **Electromagnetic fields in curved space time**. Consider a four dimensional manifold M. An electromagnetic field on M is a two-form F on M. Without defining it properly now (see tensor fields), the field F is a skew-symmetric 4×4 matrix $F_{ij} = -F_{ji}$ attached at each point $m \in M$ so that the entries of F depend smoothly on the point m. The field is required to satisfy the Maxwell equations $\frac{\partial F_{ij}}{\partial x_k} + \frac{\partial F_{jk}}{\partial x_i} + \frac{\partial F_{ki}}{\partial x_j} = 0$. A class of such fields can be obtained as $F_{ij} = dA_{ij} = \frac{\partial A_i}{\partial x_j} \frac{\partial A_j}{\partial x_i}$, where $A = (A_1, A_2, A_3, A_4)$ is a vector potential: $A_i(m) \in \mathbf{R}$ is depending smoothly on m. Are they all? This problem is a cohomological problem. There exists a group $H^2(M)$ called second cohomology group which is an invariant of the manifold. Isomorphic manifolds have the same group. If this group is trivial, then all the fields F are of the form $F = dA$. For the four dimensional sphere S^4 one has $H^2(S^4) = 0$. The same is true for a smoothly contractible manifold. Consider now the manifold $M = \mathbf{R}^4 \setminus \mathbf{R}_t$, where $\mathbf{R}_t$ is the t axis (one of the four coordinate axis in $\mathbf{R}^4$). This is the place, where the field $F = dA$ a point charge given by the potential $A(x) = (-1/|x|, 0, 0, 0)$ is defined. The cohomology group $H^2(M)$ is one dimensional and generated by $*F$, which is interpreted as the charge of the source.

5) **Fixed points of homeomorphisms and diffeomorphisms**: How many fixed points do there exist for a continuous map $f : M \to M$ or a homeomorphism of a manifold? **Brower's fixed point theorem** states that a continuous map from a closed disk to itself has at least one fixed point. A **theorem of Poincaré-Birkhoff** which is relevant in Hamiltonian mechanics says that an orientation preserving homeomorphisms of an annulus which rotates the boundaries in opposite directions has at least two fixed points. A **theorem of Lefshetz** tells that any continuous map of the two sphere which is homotopic to the identity has at least one fixed point. There are in general not more fixed points as the translation $z \mapsto z + 1$ on the Riemann sphere shows. Nikshin-Simon showed that a homeomorphisms of S^2 homotopic to the identity which preserves the volume measure, has at least two fixed points. Lefshetz introduced also an invariant $\Lambda(f)$ called **Lefshetz number** of a continuous map f from a manifold M to itself. If this number is not vanishing, then the map f has a fixed point. By computing the Lefshetz number one concludes that every Riemann surface of genus > 1 has a fixed point. There are area-preserving diffeomorphisms of the two-torus $\mathbf{T}^2$, which have no fixed points (for example translations). However, every homeomorphism of the two dimensional torus which is homotopic to the identity, which leaves the volume measure and the center of gravity invariant, has at least three fixed points. In many of these examples, global topological effects are important.

6) **The generalized Poincaré conjecture**: This conjecture, which is proven in dimensions $d \geq 5$ by Smale, states that every closed n-dimensional manifold which is homotopy equivalent to the n-sphere is homeomorphic to the n-sphere.
The **Poincaré conjecture** itself states that every simply connected closed manifold of dimension 3 is homeomorphic to the three dimensional sphere. This question is still open.

7) **Classification of Riemann surfaces.** A Riemann surface M is a two dimensional manifold such that coordinate changes on this manifold are holomorphic, when $\mathbf{R}^2$ is identified with $\mathbf{C}$. A choice of an atlas is called a complex structure. If $\mathcal{C}$ is the set of these structures and $\mathcal{D}$ is the set of orientation preserving diffeomorphismson M homotopic to the identiy, then $\mathcal{D}$ acts on $\mathcal{D}$ and the quotient space $\mathcal{C}/\mathcal{D}$ is called the **Teichmüller space** of M. This space is a smooth finite dimensional manifold. This problem is different from the topological classification problem of manifolds and a typical question in geometry.

8) **Can you hear the shape of a drum?** On a Riemannian manifold M, there exists the notion of a *gradient* and the notion of a **divergence**. As in the Euclidean space, one can define the **Laplacian** $\Delta f = \text{div grad} f$. If M is compact, then Δ has eigenvalues $0 \leq \lambda_1 \leq \lambda_2 \ldots$ which

converge to infinity. An interesting problem is to find the set of isospectral manifolds to a given manifold. An other problem is to find the set of metrics on M, which maximize $\det(\Delta)$, where det is a regularized determinant. The question "Can you hear ..." is due to M. Kac and is formulated for manifolds with boundaries, where the Laplace-Beltrami operator is taken with Dirichlet boundary conditions. There are now known two dimensional isospectral domains in the plane, but it is for example not known, if there exist smooth isospectral domains or convex isospectral domains in the plane.

Chapter 1

Manifolds

1.1 Definition of manifolds

Definition. A **locally Euclidean space** M of dimension n is a Hausdorff topological space, for which each point $x \in M$ has a neighborhood U, which is homeomorphic to an open subset $\phi(U)$ of $\mathbf{R}^n$. The pair (U, ϕ) is called a **coordinate system** or a **chart**.

Definition. A C^k **atlas** on a locally Euclidean space M is a collection $\mathcal{F} = \{U_i, \phi_i\}_{i \in I}$ of charts such that $\bigcup_{i \in I} U_i = M$, $\phi_{ij} = \phi_i \circ \phi_j^{-1}$ is in $C^k(\phi_i(U_j \cap U_i), \mathbf{R}^n)$. An atlas is **maximal**, if (U, ϕ) is a chart such that $\phi \circ \phi_i^{-1}$ and $\phi_i \circ \phi^{-1}$ are C^k for all $i \in I$, then $(U, \phi) \in \mathcal{F}$.

Definition. Two atlases $\mathcal{F}, \mathcal{G}$ are called **equivalent** if $\mathcal{F} \cup \mathcal{G}$ is an atlas. Given an atlas $\mathcal{A}$, the union of all atlases equivalent to $\mathcal{A}$ is a maximal and called the **differentiable structure generated by** $\mathcal{A}$.

Definition. A n-dimensional C^k**-differentiable manifold** is a pair $(M, \mathcal{F})$, where M is a n-dimensional second countable locally Euclidean space and where $\mathcal{F}$ is a differentiable structure on M. A manifold is called **smooth** if it is a C^k-differentiable manifold for all $k > 0$.

Remarks.
1) Given a manifold $(M, \mathcal{F})$, the topology of M is recovered as follows: a set $V \subset M$ is open, if and only if $\phi_i(U_i \cap V)$ is open in $\mathbf{R}^n$ for all $i \in I$.
2) In some books, a manifold is defined as a subset of $\mathbf{R}^k, k \geq n$ which looks locally like $\mathbf{R}^n$. We will see later that this is not a restriction of generality since every manifold is actually embedded into a space $\mathbf{R}^k$ (for example with $k = 2n + 1$ if M has dimension n). This definition has the advantage that less topology is involved. It is however not natural and does not work in infinite dimensions.
3) We will always assume that the manifold is smooth. If the vector space $\mathbf{R}^n$ is replaced by $\mathbf{C}^n$ and the coordinate changes are analytic, then, the differentiable structure is called a **complex structure**. If the coordinate changes are only continuous and not differentiable, we have a **topological manifold**. We will use the word manifold always in the sense **smooth differentiable manifold**.
4) A more general notion is a **Banach manifold** which is obtained by replacing $\mathbf{R}^n$ by an arbitrary Banach space E (that is a vector space, which has a norm and such that every Cauchy sequence has a limit). In this case, one drops the requirement of second countability and replaces it with paracompactness (every open cover has a countable locally finite refinement). If E is a Hilbert space (a Banach space, where the norm is given by a scalar product), one calles M a

Hilbert manifold. We will in this course not deal with such infinite dimensional manifolds.

We end this section with a definition of a **manifold with boundary**. We will not use these objects for a while.

Definition. The set $H^n_j = \{x \in \mathbf{R}^n \mid x_j \geq 0\}$ is called a j'th **half space** in $\mathbf{R}^n$. It has the **boundary** $\{x_j = 0\}$ and the **interior** $\{x_j > 0\}$. We write simply H^n for H^n_1.
The topology on H^n is given by the sets $\{U \cap H^n \mid U \in \mathcal{O}\}$, where $\mathcal{O}$ is the topology in $\mathbf{R}^n$. Note that the boundary of H^n is not the topological boundary of H^n.

Definition. A **locally Euclidean space with boundary** M of dimension n is a Hausdorff topological space, for which each point $x \in M$ has a neighborhood U which is either homeomorphic to an open subset $\phi(U)$ of $\mathbf{R}^n$ or to an open subset $\phi(U)$ of H^n. A pair (U, ϕ) is again called a **chart** or a **chart with boundary**.

Definition. A C^k **atlas** on a locally Euclidean space M with boundary is a collection $\mathcal{F} = \{U_i, \phi_i\}_{i \in I}$ of charts and charts with boundary such that $\bigcup_{i \in I} U_i = M$, $\phi_{ij} = \phi_i \circ \phi_j^{-1}$ is in $C^k(\phi_i(U_j \cap U_i))$, (where we require differentiability only at points which have a neighborhood homeomorphic to an open set in $\mathbf{R}^n$). The definition of maximality and differentiable structure is the same as in the case of manifolds without boundary.

Definition. A n-dimensional C^k-**differentiable manifold with boundary** is a pair $(M, \mathcal{F})$, where M is a n-dimensional second countable locally Euclidean space with boundary and where $\mathcal{F}$ is a C^k-differentiable structure on M. The definition of smoothness is as before.

Definition. Let M be an n dimensional manifold, then the **interior** of M is defined as the set of points, which have a neighborhood homeomorphic to an open subset of $\mathbf{R}^n$, The complement is called the **boundary of** M.

Remarks.
1) Note that the boundary of M is not the topological boundary of M but it is the topological boundary of the interior of M.
2) It follows from the definition that the boundary of M is itself a $(n-1)$ dimensional manifold.

1.2 Examples of manifolds

We give some examples of manifolds.

- Any discrete second countable topological space M is a 0-dimensional manifold. The charts are given by $(\{x\}, \phi_x)_{x \in M}$, where $\phi_x(x) = 0$. For example, every finite set M can be made into a manifold.

- The **standard differentiable structure** $\mathcal{F}$ on $M = \mathbf{R}^d$ is the atlas generated by one single chart $(\mathbf{R}^d, \phi)$, where $\phi(x) = x$. $(M, \mathcal{F})$ is a manifold. By choosing a basis, every finite dimensional real vector space M is a manifold. The differentiable structure is basis- independent.

- The $\mathbf{d-sphere}$ is the set $M = S^d = \{x \in \mathbf{R}^{d+1} \mid ||x||^2 = x_1^2 + x_2^2 + \ldots + x_{d+1}^2 = 1\} \subset \mathbf{R}^{d+1}$. Given the points $n = (0, \ldots, 0, 1)$ and $s = (0, \ldots, , -1)$, the **standard differentiable**

structure $\mathcal{F}$ on S^d is generated by $\mathcal{A} = \{(S^d \setminus \{n\}, \phi_n), (S^d \setminus \{s\}, \phi_s)\}$, where ϕ_n, ϕ_s are the **steregraphic projections** from n or s to the d dimensional equatorial plane $\{x_{d+1} = 0\}$. (Given $P \in S^d$, connect n with P and intersect this line with $\{x_n = 0\}$. This point is $\phi_n(P)$).

- Let U be an open subset of a manifold $(M, \mathcal{F} = \{(U_i, \phi_i)\})$ and $\mathcal{G} = \{(U \cap U_i, \phi_i | U \cap U_i) \mid (U_i, \phi_i) \in \mathcal{F}\}$, then $(U, \mathcal{G})$ is a manifold. For example, the set $GL(n, \mathbf{R}) \subset \mathbf{R}^{n^2}$, of $n \times n$ matrices, which are invertible, is a manifold since the determinant is continuous and therefore $\det^{-1}(\mathbf{R} \setminus \{0\})$ is open in $\mathbf{R}^{n^2}$.

- A **Lie group** G be a manifold which is also a group such that the map $G \times G \to G, (g, h) \mapsto gh^{-1}$ is smooth (smoothness of maps from one manifold to an other is defined later). Examples are
$$\mathbf{R}^n, \mathbf{C}^* = \mathbf{C} \setminus \{0\}, S^1, \mathbf{T}^n, GL(n, \mathbf{R}), SO(n, \mathbf{R}), SU(n)$$
and we will see very soon, why the groups $SO(n, \mathbf{R}), SU(n)$ are manifolds.

- Given a Lie group G and a manifold M. If there exists for each $g \in G$ a diffeomorphism T_g on M, one says that M **acts** on M. A Lie group action is called **transitive** if for all $x, y \in M$, there exists $g \in G$ such that $T_g(x) = y$. A manifold M on which a Lie group acts transitively is called a **homogenous space** of the Lie group. An example is any Lie group G itself because G acts onto itself by right multiplication $T_g h = g \circ h$. An other example is the rotation group $G = SO(3, \mathbf{R})$ of all orthogonal matrices with determinant 1. It acts transitively on the manifold $M = S^2$.

- A manifold E is called a **fibre bundle** over a manifold $(M, (U_i, \phi_i))$, if there exists a smooth surjective map $\pi : E \to M$ such that there exist diffeomorphisms $\psi_i : \pi^{-1}(U_i) \to U_i \times F$, where F is a manifold called **fibre**. One also assumes that a Lie group G called **structure group** acts on F by smooth diffeomorphisms. Furthermore, the coordinate changes
$$\psi_i \circ \psi_j^{-1} : (U_j \cap U_i) \times F \to (U_j \times U_i) \times F$$
have the form $(x, y) \mapsto (x, T_{ij}(x)y)$, where $T_{ij}(x) : F \to F$ is an element of G and $T_{ij} : U_i \cap U_j \to G$ is smooth.
If $F = G$ and G acts on F by multiplication on the right, then E is called a **principal bundle**. If the fibre F is discrete, then the fibre bundle E is called a **covering**, if $F = \mathbf{R}^n$ and G is a subgroup of $GL(n, \mathbf{R})$, then E is called a **vector bundle**. If $E = M \times F$, then E is called **trivial**.

- Let $(M_1, \mathcal{F}_1), (M_2, \mathcal{F}_2)$ be manifolds, then $M = M_1 \times M_2, \mathcal{F} = (U_i \times V_j, \phi_i \times \psi_j)_{i \in I, j \in J}$ is a manifold called the **product manifold**.

- We will see that every **one dimensional connected manifold** is diffeomorphic to the circle S^1 or $\mathbf{R}$. One can also show (and we will not do that here) that every two dimensional compact connected orientable manifold is a "sphere with handles".

- A differentiable structure on a manifold is not necessarily unique. This problem is difficult and we only address it here informally: one knows that for $n \leq 3$, every manifold has exactly one differentiable structure. For $n \geq 5$, there are only finitely many different differentiable structures. Milnor has given a first example of so called **exotic structures**, he constructed on S^7 different differentiable structures. It is known that there are there 28 different differentiable structures on S^7. An example of an exotic sphere is the **Brieskorn-Sphere** which is defined in $\mathbf{C}^5$ through the equations
$$z_1^5 + z_2^3 + z_3^2 + z_4^2 + z_5^2 = 1, \; \sum_{i=1}^{5} |z_i|^2 = 1 \; .$$
On $\mathbf{R}^n$, there exist different differentiable structures if and only if $n = 4$. One even knows that there are one-parameter families of differentiable structures on $\mathbf{R}^4$.

1.3 Diffeomorphisms

Definition. Given two manifolds M and N. A function $f : M \to N$ is called a C^k**-map** if for all charts (U_i, ϕ_i) on M and all charts (V_i, ψ_i) on N, the functions $\phi_j \circ f \circ \phi_i^{-1}$ are in $C^k(\phi_i(U_i \cap f^{-1}V_j), \mathbf{R}^n)$. If there exists an inverse of f which is also a C^k map, then f is called a C^k-**diffeomorphism**.

Definition. If a map $f : M \to N$ is a C^k-map for all k, then f is called **smooth**.

If f is a diffeomorphism which is smooth and the inverse of f is also smooth, then f is called a smooth diffeomorphism.

Definition. If there exists a diffeomorphism from M to N, then the two manifolds M, N are called **diffeomorphic**.

Lemma 1.3.1 *Let M, N, P be smooth manifolds of dimension m, n, p. If $f : M \to N$ and $g : N \to P$ are both C^r-maps , then also $g \circ f : M \to P$ is a C^r map.*

Proof. The statement is a local statement. This means that we can take neighborhood charts $(U, \phi), (V, \psi), (W, \theta)$ of the points $x \in M$, $f(x) \in N$, $g(f(x)) \in P$ satisfying $f(U) = V, g(V) = W$ and check the statement there. By replacing f with $\tilde{f} = \psi f \phi^{-1}$ defined on $\tilde{U} = \phi(U) \subset \mathbf{R}^m$ and g with $\tilde{g} = \theta g \theta^{-1}$ defined on $\tilde{V} = \psi(V) \subset \mathbf{R}^n$, we have to check the statement for maps $\tilde{f} : \tilde{U} \subset \mathbf{R}^m \to \tilde{V} \subset \mathbf{R}^n$, $\tilde{g} : \tilde{V} \subset \mathbf{R}^n \to \tilde{W} \subset \mathbf{R}^p$, where we can use the **chain rule** $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$. The k'th derivative $D^k(g \circ f)$ is inductively proven to be a sum of derivatives in g and f which are all of order $\leq k$. □

Corollary 1.3.2 *To be diffeomorphic is an equivalence relation.*

Proof. An equivalence relation has three properties, (i) reflexivity, (ii) symmetry and (iii) transitivity:
(i) $\mathrm{Id} : M \to M$, $\mathrm{Id}(x) = x$ is a smooth diffeomorphism.
(ii) If $f : M \to N$ is a diffeomorphism, then $f^{-1} : N \to M$ is a diffeomorphism.
(iii) If $f : M \to N, g : N \to O$ are diffeomorphisms, then $g \circ f : M \to O$ is a diffeomorphism by the above lemma. □

A natural problem is to classify all n dimensional manifolds, which means to find the set of all equivalence classes of manifolds. A part of this problem is to classify all manifolds, which are homeomorphic to a given manifold.

Definition. A manifold M of dimension m is called **immersed** in a manifold N of dimension $n \geq m$, if there exists a smooth map $f : M \to N$, such that $Df(x)$ has maximal rank $n - m$ for all $x \in M$ (that is $Df(x)$ is surjective). The map is then called an **immersion** of M into N. An immersion is called an **embedding**, if M is diffeomorphic to $f(M)$. In this case, M is called a **submanifold** of N.

1.4 A lemma for constructing manifolds

Definition. Let M, N be manifolds of dimension $m \geq n$ and let $f : M \to N$ be a smooth map. A point $x \in M$ is called a **critical point**, if the rank of $Df(x)$ is not n. The complement of the set of critical points is called the set of **regular points**. The image of a critical point is called a **critical value**. The complement of the set of critical values is the set of **regular values**.

The following lemma allows to define manifolds easily. It is also useful in applications or proofs.

> **Lemma 1.4.1** *Let M and N be manifolds of dimension $m \geq n$. Let $f : M \to N$ be a smooth map. If $y \in N$ is a regular value, then $f^{-1}(y)$ is a manifold of dimension $m - n$ in M.*

Proof. Take a point $x \in f^{-1}(y)$. Since y is a regular value, the derivative $Df(x)$ must map $\mathbf{R}^m$ onto $\mathbf{R}^n$, [1], and the kernel of $Df(x)$ is a $(m-n)$-dimensional vector space. Take a chart (U, ϕ) in M which contains x and a chart (V, ψ) in N which contains y. Defines the smooth map

$$g = \psi \circ f \circ \phi^{-1} : \phi(U) \subset \mathbf{R}^m \to \mathbf{R}^n .$$

Denote the kernel of $Dg(x)$ by H and its orthogonal complement by H' so that $H \oplus H' = \mathbf{R}^m$. We know that H is $(m-n)$-dimensional. Choose the linear map $L : \mathbf{R}^m = H \oplus H' \to \mathbf{R}^{m-n}, (h, h') \mapsto h$ which is nonsingular on H and define $F : \phi(U) \to \mathbf{R}^n \times \mathbf{R}^{m-n}$ by $F(z) = (g(z), Lz)$ which has the derivative $DF(u) = (Dg(u), Lu)$ and which is nonsingular. By the **inverse function theorem**, a neighborhood $\phi(U)$ of $\phi(x)$ is mapped by F diffeomorphically onto a neighborhood $F(\phi(U))$ of $F(\phi(x))$. We get so a chart $U_x = \phi^{-1} \circ F^{-1}(F(\phi(U)) \cap \{\psi(y)\} \times \mathbf{R}^{m-n})$ on $f^{-1}(y)$ which is maped by $\phi_x = F \circ \phi$ into a $(m-n)$-dimensional space. Doing the same construction at any point $x \in M$ produces an atlas for $f^{-1}(y)$ and shows that $f^{-1}(y)$ is a manifold. $\square$

Examples.

- The function $f : \mathbf{R}^m \to \mathbf{R}$, $f(x) = \sum_{i=1}^m x_i^2$ has the set of regular values $\mathbf{R} \setminus \{0\}$. The set $f^{-1}(1)$ is the unit sphere. The lemma assures that the sphere is a manifold.
- Take the $m = k^2$ dimensional vector space $M = M(k, \mathbf{R})$ of all real $k \times k$ matrices. The determinant $f : M \to \mathbf{R}$, $f(x) = \det(x)$ which is a homogenous polynomial of degree k. Every summand in $f(x)$ is a multiplication of k entries $x_{\sigma(1)} \cdots x_{\sigma(k)}$ in x_i and det is therefore smooth. The derivative $df(x)$ is the gradient $\partial f(x)/(\partial x_{ij})$ which is nonvanishing. This shows that $Df(x) : \mathbf{R}^{k^2} \to \mathbf{R}$ has a one dimensional range. The above lemma assures that $f^{-1}(1) =: SL(k, \mathbf{R})$ is a $k^2 - 1$ dimensional manifold. It would here be more difficult to construct an explicit atlas for this manifold.
- The set $M = \{f_i(x) = 0, \{i = 1 \dots n - k\}\} \subset \mathbf{R}^n$ is a manifold of dimension k if $DF(x)$ has rank $n - k$ for all $x \in \mathbf{R}^n$, where $F(x) = (f_1(x), \dots, f_{n-k}(x))$.
- A **Riemann surface** is a two dimensional manifold defined to be a (complex one dimensional) surface in $\mathbf{C}^2$ given by the equation $f(z, w) = 0$, where $f : \mathbf{C}^2 \to \mathbf{C}$ is an analytic function like for example a polynomial in z, w. The Riemann surface is called **nonsingular** if $Df(z, w) = (f_z, f_w) \neq 0$ for all $(z, w) \in \mathbf{Z}^2$. The same proof as in the real case shows that a nonsingular Riemann surface is a one dimensional complex manifold. For example, if $z \mapsto P(z)$ is polynomial without multiple roots, then $P'(z) \neq 0$ if $P(z) = 0$, so that $Df(z, w) \neq 0$ for all z, w if $f(z, w) = w^2 - P(z)$. Such a surface is called a **hyperelliptic curve**.

[1] A better view is that $D(f)(x)$ maps the tangent space $T_x M \sim \mathbf{R}^m$ onto the tangent space $T_y N \sim \mathbf{R}^n$. We will define the tangent space later

1.5 The theorem of Sard

Definition. A subset A of $\mathbf{R}^n$ has **measure zero**, , if there exists for every $\epsilon > 0$ a countable open cover U_i of A such that the sum of the Euclidean volumes of U_i is less than ϵ. A subset A of a finite dimensional manifold M is of **measure zero**, if for each chart $\phi : U \to \mathbf{R}^n$, the set $\phi(U \cap A)$ has measure zero.

Remark. Because the coordinate changes ϕ_{ij} between two charts are smooth, the property of having measure zero is independent of the used chart: if $D\phi_{ij} \leq M$ on the closure of some open set U and C is a cube in U, then the volume of $\phi_{ij}(C)$ is less or equal to $(M\sqrt{n})^n$ times the volume of C so that $\phi_{ij}(C)$ has measure zero if and only if C has measure zero.

Definition. Let $f : M \to N$ be a smooth map. A point $x \in M$ was called **regular**, if the linear map $Df(x) : \mathbf{R}^m \to \mathbf{R}^n$ is surjective. A not regular point is called a **critical point**. The set of critical points is denoted by C. If $x \in C$, then $f(x)$ is called a **critical value**. The complement of the set of critical values is called the set of **regular values**. Note that the image of the set of regular points is only a subset of the set of regular values!

Theorem 1.5.1 (Theorem of Sard) *Let $f : M \to N$ be a smooth map between the m-dimensional manifold M and the n-dimensional manifold N. Then the set of critical values of f has measure zero in N.*

Proof. Because a manifold is by definition second countable, it suffices to prove the statement if M is an open subset U of $\mathbf{R}^m$ and if $f : U \to \mathbf{R}^n$.

The proof goes by induction on the sum $m + n$ of dimensions of M and N. Since the claim is true for $m = 0$ or $n = 0$, we can assume $m, n \geq 1$.

Define the sets $C_i = \{x \in M \mid D^j f(x) = 0,\ j = 1, \dots, i\}$ and the set $C = \{x \in M \mid \mathrm{ran} Df(x) < n\}$ of critical points. For any k, there is the decomposition

$$C = (C \setminus C_1) \cup (C_1 \setminus C_2) \cup \ldots \cup (C_{k-1} \setminus C_k) \cup C_k \ .$$

The proof of the theorem has three steps:

(i) $f(C_k)$ has measure zero if k is large enough.
(ii) $f(C \setminus C_1)$ has measure zero.
(iii) $f(C_j \setminus C_{j+1})$ has measure zero, where $1 \leq j \leq k - 1$.

Proof of (i): Because $k \geq 1$, we have $kn \geq n + k - 1$. If k is large enough, then also $k \geq m - n + 1$ so that $kn \geq m$. It is enough to show that for each closed cube $K := \prod_i [a_i, a_i + e] \subset U$, $f(C_k \cap K)$ has measure zero. Because C_k can be covered with countably many such cubes, this will show that $f(C_k)$ has measure zero. By Taylor's theorem, the compactness of the cube K and the definition of C_k, we have $f(y) = f(x) + R(x, y)$, with $||R(x, y)|| \leq M_0 \cdot ||y - x||^{k+1}$ for $x \in C_k \cap K$ and $y \in K$, where M_0 is some constant.

For $l \in \mathbf{N}$ subdivide the cube K into l^m cubes of length e/l and choose any cube K' in this subdivision which intersects C_k. Clearly any two points $x, y \in K'$ satisfy $||x - y|| \leq \sqrt{m}e/l$. Therefore, $f(K') \subset L$, where L is a cube of length $N_0 l^{-(k+1)}$, with $N_0 := 2M_0(\sqrt{m}l/e)^{k+1}$. The volume of L is $N_0^n l^{-n(k+1)}$. Since there are at most l^m such cubes, the set $f(C_k \cap K)$ is contained in a union of cubes with total volume $V \leq N^n l^{m-n(k+1)} \to 0,\ l \to \infty$.

Proof of (ii): Decompose $C \setminus C_1 = \{x \in U \mid 1 \leq \text{ran}Df(x) < n\} = K_1 \cup \ldots \cup K_{n-1}$, where $K_q = \{x \in U \mid \text{ran}Df(x) = q\}$. It suffices to show that $f(K_q)$ has measure zero for $q \leq m$ because $K_q = \emptyset$ for $q > m$. We can find local coordinates in a neighborhood V of $x \in K_q$ such that $V = V_1 \times V_2$ and so that f has the form $(t,s) \mapsto (t, \eta(t,s)$ with a smooth map $\eta : V_1 \times V_2 \to \mathbf{R}^{n-q}$. Then

$$K_q \cap \{\{t\} \cap V_2)\} = \{t\} \times \{x \in V_2 \mid D_s\eta(t,s) = 0\} \ .$$

Because $q \geq 1$, we see by induction that the set of critical values of $\eta(t,\cdot)$ has measure zero. By Fubini's theorem, also $f(K_q \cap V)$ has measure zero and since K_q can be covered by countably many such V, this shows that $f(K_q)$ has measure zero.

Proof of (iii): Again, it is enough to show that every $x \in C_j \setminus C_{j+1}$ has a neighborhood V such that $f(V \cap (C_j \setminus C_{j+1}))$ has measure zero. Because some partial derivative of order $j+1$ is nonzero, we can assume $D_1 w(x) \neq 0$, where $w(x) := (D_{i(1)} \ldots D_{i(j)} f)(x)$ is a j'th derivative. Define $h : U \to \mathbf{R}^m$ by $x = (x_1, \ldots, x_m) \mapsto h(x) = (w(x), x_2, \ldots, x_m)$. Because dh is nonsingular, there exists by the implicit function theorem a neighborhood V of x and an open set $W \subset \mathbf{R}^m$ such that $h : V \to W$ is a smooth diffeomorphisms. Define $A = C_j \cap V$ and $A' = h(A)$ and $g = h^{-1}$. Consider the smooth function $F : W \to \mathbf{R}^m, F(x) = f \circ g(x)$. Define $W_0 = \{(x_2, \ldots x_m) \in \mathbf{R}^{m-1} \mid (0, x_2, \ldots, x_n) \in W\}$ and $F_0 : W_0 \to \mathbf{R}^m$ by $F_0(x_2, \ldots, x_m) = F(0, x_2, \ldots, x_m)$. By induction, the set $S = \{(x_2, \ldots, x_m) \in W_0 \mid DF(0, x_2, \ldots, x_m) = 0\}$ has measure zero. But $A' = h(C_j \cap V) \subset 0 \times S$ because $x \in A', DF(x) = 0$ and because for $x \in C_j \cap V$, $h(x) = (w(x), x_2, \ldots, x_m) = (0, x_2, \ldots, x_m)$ and w is a j'th derivative of f. Therefore $f(C_j \cap V) = F(h(C_j \cap V)) \subset F(0 \times S) = F_0(S)$ has measure zero. □

Remark. Sard's theorem is also true in the case, when $f \in C^k$ with $k \geq m-n+1, k \geq 1$, but we do not show this here. This smoothness assumption can however not be weakened any more. Whitney has first found C^1-function $f : \mathbf{R}^2 \to \mathbf{R}$ which is not C^2 and which contains $[0,2]$ among its critical values. A simpler example has been found by Grinberg: let C be the standard Cantor set in $[0,1]$. Then $C + C = [0,2]$. Construct a function g which is differentiable and which has C as its critical points. The function $f(x,y) = g(x) + g(y)$ has the set $C + C = [0,2]$ as the set of critical values.

Remark. The Theorem of Sard is wrong in infinite dimensions even in the smooth case (an example of Bonic,Douady,Kupka): consider $M = \{x = (x_1, x_2, \ldots) \mid ||x||^2 = \sum_j x_j^2/j^2 < \infty\}$ and the smooth map $f : M \to \mathbf{R}, f(x) = (3x_j^2 - 2x_j^3)/2^j$ which has the derivative $Df(x)u = \sum_j 6(x_j - x_j^2)2^{-j}u_j$ and $Df(x) = 0$ if and only if $x_j = 0$ or $x_j = 1$. But on the subset C of M, where $x_j \in \{0,1\}$, we have $f(x) = \sum_j x_j/2^j$ which takes all values in $[0,1]$ for $x \in C$.

Corollary 1.5.2 (Theorem of Brown) *The set of regular values of a smooth map $f : M \to N$ is everywhere dense in N.*

Proof. Since the set of critical values has measure zero, it can not contain an open set. □

1.6 Partition of Unity

Definition. Let M be a manifold. A **partition of unity of M with respect to an open cover** $\{U_\alpha\}_{\alpha \in A}$ is a collection of smooth maps $g_\alpha : M \to \mathbf{R}^+ = \{r \in \mathbf{R} \mid r \geq 0\}$ with supports $\text{supp}(g_\alpha) \subset U_\alpha$ satisfying $\sum_\alpha g_\alpha = 1$.

Lemma 1.6.1 (A lemma in topology) *Let $(X, \mathcal{O})$ be a locally compact Hausdorff second countable topological space. Then X is paracompact.*

Proof. (i) There exists a countable basis $\mathcal{C} = \{U_i\}$ such that $\overline{U}_i$ is compact.
Proof. By the property of second countability, there exists a countable basis $\mathcal{B}$. By local compactness and the Hausdorff property, the set $\mathcal{C} = \{U \in \mathcal{B} \mid \overline{U} \text{ compact }\}$ is again a basis.

(ii) There exists a sequence $G_i \in \mathcal{O}$ such that $\overline{G}_i$ is a compact subset of G_{i+1} and such that $\bigcup_i G_i = X$.
Proof. Let $\mathcal{C} = \{U_i\}$ be a countable basis given in (i). Let $G_1 = U_1$ and inductively $G_k = \bigcup_{i=1}^{j_k} U_i$, where j_k the is the smallest number satisfying $j_k > j_{k-1}$ and $\overline{G}_{k-1} \subset \bigcup_{i=1}^{j_k} U_i$.

(iii) X is paracompact.
Proof. Let $\{U_\alpha\}_{\alpha \in A}$ be an arbitrary open cover of X. The set $\overline{G}_i \setminus G_{i-1}$ is compact and contained in the open set $G_{i+1} \setminus \overline{G_{i-2}}$. For each $i \geq 3$, choose a finite subcover of the open cover $\{U_\alpha \cap (G_{i+1} \setminus \overline{G_{i-2}}\}_{\alpha \in A}$ of the compact set $\overline{G_2}$. The union of all these finite covers is a countable, locally finite refinement of the open cover $\{U_\alpha\}_{\alpha \in A}$ and consists of open sets with compact closures. □

Lemma 1.6.2 (A lemma in real analysis) *Let $C(r) = [0, r]^d$ be the cube of length r in $\mathbf{R}^d$. There exists a non-negative smooth function $f_r : \mathbf{R}^d \to \mathbf{R}$ which is 1 on $C(r)$ and 0 on the closure of the complement of $C(r)$.*

Proof. By scaling, it is enough to prove this for $r = 1$.

(i) There exists a nonnegative smooth function h on $\mathbf{R}$ which is 1 on $[-1, 1]$ and zero outside $(-2, 2)$.
Proof. Let $q(t) = 1_{t>0}e^{-1/t}$, which is positive for $t > 0$ and smooth. Define $g(t) = q(t)/(q(t) + q(1-t))$ which is 1 for $t \geq 1$ and 0 for $t \leq -1$. Define $h(t) = g(t+2)g(2-t)$.

(ii) For $x = (x_1, \dots, x_n)$, define $f(x) = h(x_1) \cdot h(x_2) \cdots h(x_n)$. □

```
q[t_]:=If[t>0, Exp[-1/t],0]; g[t_]:=q[t]/(q[t]+q[1-t]); h[t_]:=g[t+2] g[2-t];
f[x_,y_]:=h[x] h[y]; Plot3D[f[x,y],{x,-4,4},{y,-4,4},PlotPoints->40];
```

Theorem 1.6.3 (Existence of a partition of unity) *Given a manifold M and an open cover U_α of M. Then there exists a partition of unity with respect to U_α.*

Proof. A manifold M is a locally compact second countable Hausdorff space. By the topological lemma, it is paracompact.

Let G_i the cover of M defined in the "topological lemma" and set $G_0 = \emptyset$. For $x \in X$, let i_x the largest integer such that $x \in M \setminus \overline{G_{i_x}}$. Choose α_x with $x \in U_{\alpha_x}$, where $\{U_\alpha\}$ is the fixed open cover of M. Let (V, ϕ) be a chart satisfying $x \in V$, $V \subset U_{\alpha_x} \cap (G_{i_x+2} \setminus \overline{G_{i_x}})$ and such that $\phi(V) \subset \mathbf{R}^n$ contains the closed cube $\overline{C(r)}$.

Define the smooth function $h_x(y) = 1_V(y) \cdot f_r \circ \phi(y)$, where f_r is the function defined in the "real analysis lemma" and $1_V(y)$ is the characteristic function of V which is one if $y \in V$ and 0 if $y \notin V$. The function h_x takes the value 1 in some open neighborhood W_x of x and has compact support contained in a set $V \subset U_{\alpha_x} \cap (G_{i_x+2} - \overline{G_{i_x}})$. For every $i \geq 1$, choose a finite set of points $x \in M$, whose corresponding neighborhoods W_x cover the compact set $\overline{G_i} \setminus G_{i-1}$. Number the corresponding h_x functions $\{h_j\}_{j=1}^{\infty}$. The supports of h_j form a locally finite family of subsets of M. Therefore $g(x) = \sum_j h_j(x)$ is a positive finite smooth function on M. The functions $g_j = h_j/h$ have the property that $\sum_j g_j = 1$. Define now g_α to be identically zero if no h_i has support in U_α and otherwise, let g_α be the sum of the g_i with nonempty support in U_α. Then g_α is a partition of unity with respect to the cover U_α.

The support of g_α is contained in U_α because the family of closed sets $\mathrm{supp}(g_\alpha)$ is locally finite which implies that $\bigcup_{i \in I} \mathrm{supp}(g_i)$ is closed so that the support of g_α is the union of the g_i which have nonempty support in U_α. □

1.7 Whitney's embedding theorem

Definition. A smooth map $f : M \to N$ is called an **immersion**, if $Df(x)$ is injective at each point $x \in M$. An immersion is called an **embedding**, if $f : M \to f(M)$ is a homeomorphism .

Theorem 1.7.1 (Whitney's embedding theorem) *A compact manifold of dimension n can be embedded in $\mathbf{R}^{2n+1}$.*

Proof. (i) Because M is compact, we can find an atlas $\mathcal{A} = \{(U_i, \phi_i)\}_{i=1}^{p}$ with a finite number of charts. We choose the charts, so that $\phi_i(U_i)$ contains the disk $\mathbf{D}^n(2) = \{x \in \mathbf{R}^n \mid |x| \leq 2\}$ and such that $\phi^{-1}(\mathbf{D}^n(1))$ is still a cover of M. Take a smooth function λ which is 1 on $\mathbf{D}^2(1) = \{x \in \mathbf{R}^n \mid |x| \leq 1\}$ and 0 outside $\mathbf{D}^n(2)$. Let g_i a partition of unity subordinate to the cover $\phi_i^{-1}(\mathbf{D}(1))$. Define the smooth map $\psi_i : M \to \mathbf{R}^n$ by defining it to be $\lambda(\phi_i) \cdot \phi_i$ on U_i and 0 else. Define $F : M \to \mathbf{R}^{np+p}$ by

$$F(x) = (\psi_1(x), \dots, \psi_p(x), g_1(x), \dots, g_p(x)) .$$

• F is an immersion.
Proof. Given a point $x \in M$. It is contained in a support of one of the maps g_i. The map $y \mapsto \psi_i(y)$ is a homeomorphism of a neighborhood V of x with its image $\psi(V) \subset \mathbf{R}^n$ and $D\psi(x) = D\phi(x)$ is the identity and therefore, $D\psi(x)$ is injective.

• F is an embedding.
Proof. We have to show that F is injective. Given two points $x, y \in M$. If $g_i(x) \neq g_i(y)$ for some y, then $F(x) \neq F(y)$. Assume now $g_i(x) = g_i(y)$ for all i. Let g_i a function in the partition of unity, for which $g_i(x) = g_i(y) \neq 0$. (It is not possible that $g_i(x) = g_i(y) = 0$ for all i). Then x, y are in the support of g_i and so in $\phi_i^{-1}(\mathbf{D}^1)$. Therefore $\psi(x) = \phi(x) \neq \phi(y) = \psi(y)$ so that then also $F(x) \neq F(y)$. Because F is injective, it is a bijection onto its image and since M is compact, the map F is a homeomorphisms between M and $F(M) \subset \mathbf{R}^q$ with $q = np + p$.

(ii) Let S^n be the n-dimensional sphere. In order to build an embedding of M in $\mathbf{R}^{2n+1}$, we proceed inductively as follows to reduce the dimension. Assume M is embedded in $\mathbf{R}^q$, where

$q > 2n+1$. The set of projections π_y onto hyperplanes $H_y \sim \mathbf{R}^{q-1}$ in $\mathbf{R}^q$ is labeled by the sphere $S^{q-1} = \{v \mid |v| = 1\}$.

• There exists a set of measure 1 of S^{q-1} such that for v in this set, the projection $\pi^{(v)} : M \to \mathbf{R}^{q-1}$ is injective, that is

$$v \neq (x-y)/||x-y||, \; x, y \in M \subset \mathbf{R}^q \; .$$

Proof. Consider the map $\sigma : (M \times M \setminus \Delta) \to S^{q-1}$ given by $\sigma(x,y) = |x-y|/||x-y||$, where $\Delta = \{(x,x) \mid x \in M\}$ is the diagonal in $M \times M$. If $2n < q-1$, then σ has by Sard's theorem an image of measure zero. A point v in the complement of the image of σ gives an injective map π_v.

• The exits a set of measure 1 in S^{q-1} such that $\pi^{(v)}$ is an immersion.
Proof. Consider the compact manifold $M \times S^{n-1}$. Define the projection map $\psi : M \times S^{n-1} \subset \mathbf{R}^q \times S^{q-1} \to S^{q-1}$ as the projection onto the second coordinate which is in $S^{n-1} \subset S^{q-1}$, (we identify S^{n-1} as a submanifold of S^{q-1}). Because $\dim(M \times S^{n-1}) = n + (n-1) = 2n-1 < \dim(S^{q-1}) = q-1$, we know by Sard's theorem that the set of critical values of ψ has measure zero. A regular value v gives an immersion π^v because $D\pi(x)u = \pi(u)$ has then maximal rank. □

1.8 Brower's fixed point theorem

Lemma 1.8.1 *Given a continuous map $f : \mathbf{R}^n \to \mathbf{R}^m$ of compact support. For every $\epsilon > 0$, there exists a smooth function $f_\epsilon : \mathbf{R}^n \to \mathbf{R}^m$ of compact support, such that* $\sup_{x \in \mathbf{R}^n} |f_\epsilon(x) - f(x)| \to 0$ *for* $\epsilon \to 0$.

Proof. Define for $\epsilon > 0$ a smooth function $\chi_\epsilon : \mathbf{R}^n \to \mathbf{R}$ with support in $\{|x| < \epsilon\}$ such that $\int_{\mathbf{R}^n} \chi_\epsilon(x) \, dx = 1$ (see the lemma in real analysis proven in the section about the partition of unity). Define

$$g_\epsilon(x) = (\chi_\epsilon \star f)(x) = \int_{\mathbf{R}^n} \chi_\epsilon(y) g(x-y) \, dy = \int_{\mathbf{R}^n} \chi_\epsilon(x+z) g(z) \, dz$$

which is obviously smooth because $D^k g(x) = \int_{\mathbf{R}^n} (D^k \chi_\epsilon(x+z)) g(z) \, dz$. Because g has compact support, there exists a constant C such that $||Dg(x)u|| \leq C \cdot ||u||$ for all $x \in \mathbf{R}^n$. We get

$$\begin{aligned} ||g_\epsilon(x) - g(x)|| &= ||\int_{\mathbf{R}^n} \chi_\epsilon(x+z)(g(z) - g(x)) \, dz|| \\ &\leq \int_{\mathbf{R}^n} \chi_\epsilon(x+z)||g(z) - g(x)|| \, dz \\ &\leq \sup_{x \in \mathbf{R}^n, u \in \mathbf{R}^n, ||u||=1} ||Dg(x)u||\epsilon \leq C\epsilon \; . \end{aligned}$$

□

Definition. The **n-dimensional disk** is the manifold with boundary $\{|x| \leq 1\} \subset \mathbf{R}^n$. It has as the boundary the manifold S^{n-1}.

Proposition 1.8.2 *There exists no continuous map $f : \mathbf{D}^n \to S^{n-1}$ which is the identity on the boundary S^{n-1}.*

Proof. Assume such a function f exists.

(i) There exists then a smooth map g with the same properties.
Proof. The function

$$\tilde{f}(x) = f(x/|x|) = x/|x|,\ |x| \in [1/2, 1], \tilde{f}(x) = f(2x), |x| \in [0, 1/2]$$

is smooth in for $|x| > 1/2$. For $\epsilon < 1/8$, define $g(x) = \chi_\epsilon \star \tilde{f}(x)/||\chi_\epsilon \star \tilde{f}(x)||$ for $|x| < 3/4$ and $g(x) = \tilde{f}(x)$ else. This is a smooth map which maps $\mathbf{D}^n$ to S^{n-1} and which satisfies $g(x) = x$ on S^{n-1}. (Note that $g(x) = \tilde{f}(x)$ for $3/4 < |x| < 7/8$ because $\int \chi_\epsilon(y)(x-y)/|x|\, dy = x/|x|$.

(ii) The theorem of Sard applied to the map $M = \mathrm{int}(\mathbf{D}^n) \to S^{n-1} = N$ implies that there exists a regular value y of g. The set $g^{-1}(y)$ is a one dimensional manifold V in M and a one dimensional manifold with boundary in $\mathbf{D}^n$ because the boundary of V is the intersection of V with S^{n-1} which is nonempty (it contains y). Since $\mathbf{D}^n$ is compact, this manifold is compact. From the classification of one dimensional manifolds (see the next appendix), it must be diffeomorphic to $[0, 1]$ and has on one side the boundary $y \in S^{n-1}$ and on the other side a **different point** on the boundary S^{n-1}. Since $f(z) = y$ and $f(z) = z$, we have $z = y$, which is a contradiction. $\square$

Corollary 1.8.3 (Brower's fixed point theorem) *Every continuous map $f : \mathbf{D}^n \to \mathbf{D}^n$ has a fixed point.*

Proof. Assume $f(x) \neq x$ for all $x \in \mathbf{D}^n$. We can find then a continuous map $g : \mathbf{D}^n \to S^{n-1}$ which is the identity on the boundary and which is given by the point obtained by taking the line connecting x and $f(x)$ and intersecting it with S^{n-1}. By the previous proposition, such a map does not exist. $\square$

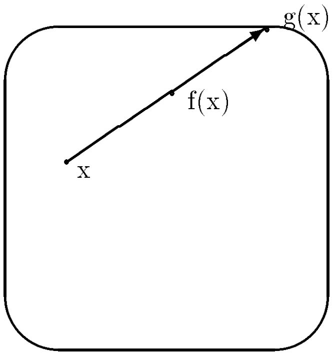


Fig. Proof of Brower's fixed point theorem. Construct a retract g from $\mathbf{D}^n$ to S^{n-1}.

Appendix: classification of one dimensional manifolds

Definition. Let M be a one dimensional manifold. A map $I = (a, b) \to M$ is called a **parametrization by arc-length** if f maps I diffeomorphically onto an open subset U of M and such that $|Df(s)| = 1$ for all $s \in I$.

Remark. Any given parametrization $g : I \to M$ can be changed to a parametrization by arc-length.

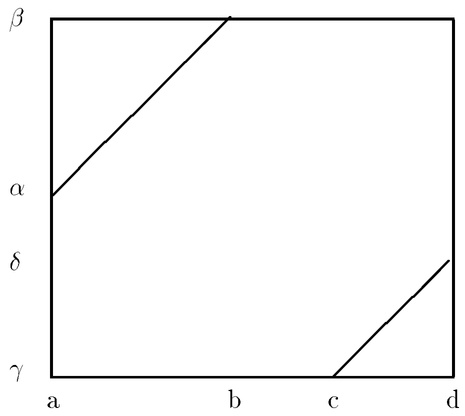
Proposition 1.8.4 *a) Every one dimensional connected manifold without boundary is either diffeomorphic to the circle or diffeomorphic to the real line. b) Every one dimensional connected manifold with boundary is diffeomorphic to either the interval* $[0,1)$ *or the interval* $[0,1]$.

Proof. a)
(i) Given two parametrizations $f : I \to M$, $g : J \to M$ by arc length. Then $f(I) \cap f(J)$ has at most two components.
Proof. The map $g^{-1} \circ f$ maps open sets in I onto open sets in J and $D(f^{-1} \circ g) = \pm 1$ for all $x \in I$. The set $\Gamma = \{(s,t) \in I \times J \mid f(s) = g(t)\}$ is closed in $I \times J$ and made up of line segments of slope ± 1 which can't end in the interiour of $I \times J$ but extend to the boundary. Because $g^{-1} f$ is injective and single valued, at most one of these line segments can end at the four edges of the rectangle $I \times J$. This assures that Γ has at most two components.

(ii) If $f(I) \cap g(J)$ has only one component, then it can be extended to a parametrization of the union $f(I) \cap g(J)$.
If Γ is connected we can extend $g^{-1} \circ h$ to a linear map $L : \mathbf{R} \to \mathbf{R}$. Now, f and $g \circ L$ piece together to yield the extension

$$F : I \cup L^{-1}(J) \mapsto f(I) \cup g(J) .$$

(iii) If $f(I) \cap g(J)$ has two components, then M is diffeomorphic to S^1.
Proof. If Γ has two components, the two must have the same slope say 1. By translating the interval $J = (\gamma, \beta)$ if necessary, we may assume that $\gamma = c$ and $\delta = d$ so that $a < b \leq c < d \leq \alpha < \beta$. The required diffeomorphism $h : S^1 \to M$ can now be constructed by setting $\theta = 2\pi t/(\alpha - a)$ and defining $h : S^1 \to M$ by the formula $h((\cos(\theta), \sin(\theta)) = f(t)$ for $t \in (a,d)$ and $h((\cos(\theta, \sin(\theta))) = g(t)$ for $t \in (c, \beta)$. Since $h(S^1)$ is compact and open in M, it must be M.



(iv) By Zorn's lemma, any parametrization by arc-lengt can be extended to a parametrisation $f : I \to M$ which is maximal.

(v) If M is not diffeomorphic to S^1, then the maximal parametrization is onto, in which case M is diffeomorphic to (a,b) which is diffeomorphic to $\mathbf{R}$.

b) Consider, now a conneced one dimensional manifold M with boundary. The interior $\text{int}(M)$ of M is a one-dimensional connected manifold without boundary. It is by a) either the circle or the real line. Since the circle is compact and has no boundary, it is excluded. The only possibility is that $\text{int}(M)$ is diffeomorphic to the real line which is diffeomorphic to $(0,1)$ which can only be in two ways a subset of a manifold with boundary: adding one point in which case M is diffeomoprhic to $(0,1]$ or adding two points in which case M is diffeomorphic to $[0,1]$. $\square$

Chapter 2

Tensor analysis

2.1 General tensors

Definition. Let E be a linear space. The **dual space** E^* of E is the set of all linear maps $f : E \to \mathbf{R}$. It is a linear space with addition $(f+g)(v) = f(v) + g(v)$ and multiplication with real numbers $(\lambda f)(v) = \lambda f(v)$.

Lemma 2.1.1 *If E has finite dimensions, then the dual space E^* has the same dimension as E. If $\{e_i\}$ is a basis in E, then the set $\{e^i\}_{i=1}^n$ defined by $e^i(v) = v^i$, where $v = \sum_i v^i e_i$, is a basis of E^* called the* **dual basis**.

Proof. (i) The set $\{e^i\}_{i=1}^n$ is linearly independent: if $f = \sum_i f_i e^i = 0$, then $f(e_j) = \sum_i f_i e^i(e_j) = f_j = 0$ for all j so that $f = 0$. Therefore $\dim(E^*) \geq \dim(E) = n$.

(ii) The set $\{e^i\}_{i=1}^n$ spans E^*. Let $f : E \to \mathbf{R}$ be any linear functional, define $f_i = f(e_i)$. We claim that $f = \sum_i f_i e^i$. Given $v = \sum_i v^i e_i \in E$. We have $f(v) = \sum_i v^i f(e_i) = \sum_i v^i f_i$. On the other hand, also $\sum_j f_j e^j(\sum_i v^i e_i) = \sum_j f_j v^j$ because $e^j(e_i) = \delta_i^j$. $\square$

Remark. The dual space of E^* has a natural identification with E. Given $e \in E$, define $e' : E^* \to \mathbf{R}$ by $e'(f) = f(e)$. Since the map $e \mapsto e'$ is injective and linear and by the lemma, the dual space E^{**} of E^* has the same dimension as E^* and so as E, it defines an isomorphism from E to E^{**}.

Definition. Let E be a real finite dimensional vector space and let E^* be the dual space of E. For $p, q \geq 0$, denote by $T_q^p(E)$ the set of all multilinear maps

$$f : (E^*)^p \times E^q \to \mathbf{R},$$

They are called **tensors of type** (p, q). The set $T(E) = \bigcup_{p,q=0}^{\infty} T_q^p(E)$ is the set of **tensors**.

Examples.

- If $f : E = \mathbf{R}^n \to \mathbf{R}$ is a smooth map, then the k'th derivative $D^k f(x)$ is an element in $T_q^0(E)$.
- If $p = 1, q = 0$, then $T_0^1(E)$ is the set of all linear maps from E^* to $\mathbf{R}$ and so $T_0^1(E) = E^{**} \sim E$. It is a custom to write $u \in E$ as $\sum_i u^i e_i = u^i e_i$ (leaving away the sum is the Einstein notation which we will not use here). One says also, an element in $T_0^1(E)$ is a **contravariant vector**.

- If $p=0, q=1$, then $T^0_1(E)$ is the set of linear maps from E to $\mathbf{R}$ and so $T^0_1(E) = E^*$. It is a custom to write $v \in E^*$ as $v = \sum_i v_i e^i$. An element in $T^0_1(E)$ is called a **covariant vector**.
- The map $g : E \times E \to \mathbf{R}, g(v,w) = v \cdot w$ (scalar product) defines a tensor in $T^0_2(E)$. If one defines $g_{ij} = g(e_i, e_j)$, then $g = \sum_{i,j} g_{ij} e^i \otimes e^j = g_{ij} e^i \otimes e^j$.
- Given a linear map $A : E \to E$. Define the map $E^* \times E \to \mathbf{R}$ by $(e^i, e_j) \mapsto a^i_j = e^i \cdot Ae_j$, where $v \cdot w$ is the **standard scalar product** from $E^* \times E$ to $\mathbf{R}$. In components $a = \sum_{i,j} a^j_i e^i \otimes e_j$. A linear map A on E defines so a tensor $a \in T^1_1(E)$.

Given $u_1, \dots u_p \in E$ and $v^1, \dots v^q \in E^*$, then

$$u_1 \otimes \cdots \otimes u_p \otimes v^1 \otimes \cdots \otimes v^q(x^1, \dots, x^p, y_1, \dots, y_q) := x^1(u_1) \cdots x^p(u_p) v^1(y_1) \cdots v^q(y_q)$$

defines a tensor in $T^p_q(E)$.

Especially, if $\{e_i\}^n_{i=1}$ be a basis in E and if $\{e^i\}^n_{j=1}$ be the dual basis in E^*, then

$$e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes e^{j_1} \otimes \cdots \otimes e^{j_q}$$

is a tensor in $T^p_q(E)$.

Notation. Given an index sets $I = (i_1, \dots, i_p)$ and $J = (j_1, \dots, j_q)$, we write for short also $e^J = e^{j_1} \otimes \cdots \otimes e^{j_p} \in T^p_0(E)$ and $e_I = e_{i_1} \otimes \cdots \otimes e_{i_q} \in T^0_q(E)$. and $e^J_I = e^J \otimes e_I \in T^p_q(E)$.

Proposition 2.1.2 *The tensors $e_I \otimes e^J$ form a basis of $T^p_q(E)$. In other words, the dimension of $T^p_q(E)$ is n^{p+q} and every $f \in T^p_q(E)$ can be written uniquely as*

$$f = \sum_{J,I} f^I_J e_I \otimes e^J \ .$$

The real numbers f^I_J are called the **components** *of f.*

Proof. Define for $J = (j_1, \dots, j_p)$ and $I = (i_1, \dots, i_q)$

$$f^I_J = f(e^{i_1}, \dots, e^{i_p}, e_{j_1}, \dots, e_{j_q}) \in \mathbf{R} \ .$$

Given the tensor $f \in T^p_q$. It can be written as

$$f = \sum_{i_1,\dots,i_p,j_1,\dots,j_q} f(e^{i_1}, \dots e^{i_p}, e_{j_1}, \dots, e_{j_q}) e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes e^{j_1} \otimes \cdots \otimes e^{j_q} = \sum_{I,J} f^I_J e_I \otimes e^J$$

so that the vectors $e_I \otimes e^J$ span $T^p_q(E)$.

To show linear independence: assume $f = \sum_{I,J} f^I_J e^J_I = 0$. Take $v = (e^{k_1}, \dots, e^{k_p}, e_{l_1} \dots, e_{l_q})$. Then $f(v) = f^L_K = 0$ for all $K = (k_1, \dots, k_p), L = (l_1, \dots, l_q)$ and so that $f = 0$. □

Definition. The **tensor product** of two tensors $f \in T^p_q(E), g \in T^r_s(E)$ is defined as the tensor

$$f \otimes g(u^1, \dots, u^p, v^1, \dots v^r, u_{p+1}, \dots, u_{p+q}, v_{r+1} \dots v_{r+s}) = f(u_1, \dots, u^{p+q}) g(v_1, \dots, v_{r+s})$$

in $T^{p+r}_{q+s}(E)$. Rewritten shortly, for $u \in (E^*)^p \times E^q$, and $v \in (E^*)^r \times E^s$, then $f \otimes g(u,v) = f(u)g(v)$.

Proposition 2.1.3 *The tensor product is associative and bilinear.*

Proof. The bilinearity

$$(f+g)\otimes h = f\otimes h + g\otimes h,\ (\lambda f)\otimes g = \lambda(f\otimes g) = f\otimes(\lambda g)$$

is obvious from the definition.

The tensor product is associative: Given $f \in T^{p_1}_{q_1}, g \in T^{p_2}_{q_2}$ and $h \in T^{p_3}_{q_3}$. Given $u \in (E^*)^{p_1} \times E^{q_1}$, $v \in (E^*)^{p_2} \times E^{q_2}$ and $w \in (E^*)^{p_3} \times E^{q_3}$. Then

$$\begin{aligned} f\otimes(g\otimes h)(u,v,w) &= f(u)(g\otimes h)(v,w) = f(u)g(v)h(w) \\ (f\otimes g)\otimes h(u,v,w) &= (f\otimes g)(u,v)h(w) = f(u)g(v)h(w)\,. \end{aligned}$$

□

Definition. One sais that the set of all tensors $T(E) = \oplus_{p,q} T^p_q(E)$ (direct product) forms a **graded algebra**.

2.2 Antisymmetric tensors

Definition. Let $\Lambda^q(E)$ be the subspace of $T^0_q(E)$, consisting of tensors f such that $f(x_1,\dots x_q)$ is antisymmetric in $x_1,\dots x_q \in E$:

$$f(x_{\sigma(1)},\dots,x_{\sigma(q)}) = (-1)^\sigma f(x_1,\dots,x_q)$$

for all $i,j = 1,\dots,q$, where $(-1)^\sigma$ is the sign of the permutation σ. One denotes by S_n the set of all permutations of the set $\{1,\dots,n\}$. The tensors in $\Lambda^q(E)$ are called **q-forms** or **exteriour q-forms**.

Corollary 2.2.1 *If E has dimension n, then the space $\Lambda^q(E)$ has dimension* $\left(\begin{array}{c} n \\ q \end{array}\right) = \frac{n!}{(q!(n-q)!)}$.

Proof. We know that the set $\{e^I = e^{i_1}\otimes e^{i_2}\otimes\cdots\otimes e^{i_q} \mid i_j \in \{1,\dots,n\}\ \}$ is a basis in $T^0_q(E)$. Given $f \in \Lambda^q(E)$. Then, for any permutation $\sigma \in S_q$, we have by definition $f_I = (-1)^{\sigma(I)} f_{\sigma(I)}$ and f_J is determined from $f_{\sigma(J)}$, where σ permutes $I = (i_1,\dots,i_q)$ into $i_{\sigma(1)} < \dots < i_{\sigma(q)}$. The subset $\{e^{i_1}\otimes e^{i_2}\otimes\cdots\otimes e^{i_q} \mid i_1 < i_2, \dots < i_q\}$ of the basis of $T^0_q(E)$ is therefore a basis for $\Lambda^q(E)$.

The number of arrangements of q indices $i_1 < \dots < i_q$ with $i_j \in \{1,\dots,n\}$ is $\left(\begin{array}{c} n \\ q \end{array}\right)$. □

Notation. An antisymmetric tensor can be written as

$$f = \sum_{i_1<\dots<i_q} f_{i_1,\dots,i_q} e^{i_1}\wedge\dots\wedge e^{i_q}\ ,$$

where $e^{i_1}\wedge\dots\wedge e^{i_q}$ are a sum of terms $(-1)^\sigma e^{\sigma(i_1)}\wedge\dots\wedge e^{\sigma(i_q)}$, where σ runs over all permutations of $\{1,\dots,n\}$.

Examples.

- $q=0$. $\Lambda^0(E)$ is the set of maps from $\mathbf{R}^0=\{0\}$ to $\mathbf{R}$ and so $\mathbf{R}$.
- $q=1$. Every element in $\Lambda^1(E)$ can be written as $f=\sum_{i=1}^n f_i e^i$. Indeed $\Lambda^1(E)=T_1^0(E)$.
- $q=2$. Every element in $\Lambda^2(E)$ can be written as $f=\sum_{i<j}^n f_{ij}e^i\otimes e^j - e^j\otimes e^i$. We write also $f=\sum_{i<j}^n f_{ij}e^i\wedge e^j$ for that.

Definition. Given $\alpha\in\Lambda^p(E)$ and $\beta\in\Lambda^q(E)$, define $\alpha\wedge\beta\in\Lambda^{p+q}(E)$ by

$$\boxed{\alpha\wedge\beta(u_1,\dots u_{p+q})=\sum_{\sigma\in S_n^*}(-1)^\sigma\alpha(u_{\sigma(1)},\dots,u_{\sigma(p)})\beta(u_{\sigma(p+1)},\dots,u_{\sigma(p+q)})\ ,}$$

where the sum is over the set S_n^* of all permutations σ of the set $\{1,\dots,p+q\}$ such that $\sigma(1)<\sigma(2)<\dots<\sigma(p)$ and $\sigma(p+1)<\dots<\sigma(p+q)$. The product is called **wedge product** or exterior multiplication on the **exterior algebra** or **Grassmann algebra** or **Fermionic algebra** $\Lambda(E)=\oplus_{q=0}^\infty\Lambda^q(E)$.

Remark. We have

$$p!q!\alpha\wedge\beta(u_1,\dots u_{p+q})=\sum_{\sigma\in S_n}(-1)^\sigma\alpha(u_{\sigma(1)},\dots,u_{\sigma(p)})\beta(u_{\sigma(p+1)},\dots,u_{\sigma(p+q)})\ .$$

Examples.

- Given two 1-forms $f\in\Lambda^1(E)$ and $g\in\Lambda^1(E)$, then

$$f\wedge g(u_1,u_2)=f(u_1)g(u_2)-f(u_2)g(u_1)\ .$$

- Given a 2-form $f\in\Lambda^2(E)$ and a 1-form $g\in\Lambda^1(E)$, then

$$(f\wedge g)(u_1,u_2,u_3)=f(u_1,u_2)g(u_3)+f(u_2,u_3)g(u_1)-f(u_1,u_3)g(u_2)\ .$$

- If $f,g,h\in\Lambda^1(E)$, then

$$\begin{aligned} f\wedge g\wedge h(u_1,u_2,u_3) &= f(u_1)g(u_2)h(u_3)+f(u_2)g(u_3)h(u_1)+f(u_3)g(u_1)h(u_2)\\ &\quad -f(u_2)g(u_1)h(u_3)-f(u_3)g(u_2)h(u_1)-f(u_1)g(u_3)h(u_2)\ . \end{aligned}$$

  This is the determinant of the matrix

$$\begin{pmatrix} f(u_1) & f(u_2) & f(u_3)\\ g(u_1) & g(u_2) & g(u_3)\\ h(u_1) & h(u_2) & h(u_3)\end{pmatrix}\ .$$

  See also the Week 3 homework.

In coordinates, we can rewrite the wedge product of $\alpha=\sum_I\alpha_I e^I$, $\beta=\sum_J\beta_J e^J$ as

$$\alpha\wedge\beta=\sum_{I,J}\alpha_I\beta_J e^I\wedge e^J\ =\sum_{K=\sigma(IJ)}\alpha_I\beta_J(-1)^\sigma e^K$$

where σ is the permutation which maps $IJ=(i_1,\dots i_p,j_1,\dots,j_q)$ into $K=(k_1,k_2,\dots,k_{p+q})$, with $k_1<k_2<\dots<k_{p+q})$.

Examples.

- Given two one forms $f = \sum_{i=1}^n f_i e^i$ and $g = \sum_{j=1}^n g_j e^j$, then

$$f \wedge g = \sum_{i<j} (f_i g_j - g_i f_j) e^i \wedge e^j \ .$$

- $E = \mathbf{R}^3$. Given a 1-form $f = f_1 e^1 + f_2 e^2 + f_3 e^3 \in \Lambda^1(E)$ and a 2-form $g = g_{12} e^1 \wedge e^2 + g_{13} e^1 \wedge e^3 + g_{23} e^2 \wedge e^3 \in \Lambda^2(E)$, then

$$g \wedge f = (g_{12} f_3 + g_{23} f_1 - g_{13} f_2) e^1 \wedge e^2 \wedge e^3 \ .$$

Remark. In the literature, there are other definitions of the product, which differ by constants.

Proposition 2.2.2 ($\Lambda(E)$ is a graded associative algebra) *The wedge product is associative, bilinear and satisfies $\alpha \wedge \beta = (-1)^{qp} \beta \wedge \alpha$ if $\alpha \in \Lambda^q$ and $\beta \in \Lambda^p$.*

Proof. We can write

$$\alpha \wedge \beta = \frac{(p+q)!}{p!q!} A(\alpha \otimes \beta) \ ,$$

where A is the **alternation mapping** $A : T_q^0(E) \to \Lambda^q(E)$ defined by

$$Af = \frac{1}{q!} \sum_{\sigma \in S_q} (-1)^\sigma \sigma f \ ,$$

where

$$\sigma f(u_1, \dots, u_q) = f(u_{\sigma(1)}, u_{\sigma(2)}, \dots, u_{\sigma(q)})$$

is a permutation in the set S_q of all permutations. The bilinearity and associativity follows now from the bilinearity and associativity of the tensor product.
The sign of the permutation $\rho : (1, \dots, q, q+1, \dots, q+p) \mapsto (q+1, \dots, q+p, 1, \dots, q)$ is $(-1)^{pq}$. With $\tau = \sigma \circ \rho$, we get

$$p!q! \ \alpha \wedge \beta(u_1, \dots, u_{p+q}) = \sum_{\sigma \in S_n} (-1)^\sigma \alpha(u_{\sigma(1)}, \dots, u_{\sigma(p)}) \beta(u_{\sigma(p+1)}, \dots, u_{\sigma(p+q)})$$

and so

$$\begin{aligned} p!q! \beta \wedge \alpha(u_1, \dots, u_{p+q}) &= \sum_{\tau \in S_n} (-1)^\tau \beta(u_{\tau(1)}, \dots, u_{\tau(p)}) \alpha(u_{\tau(p+1)}, \dots, u_{\tau(p+q)}) \\ &= \sum_{\sigma \in S_n} (-1)^\sigma (-1)^\rho \beta(u_{\sigma(p+1)}, \dots, u_{\sigma(q+p)}) \alpha(u_{\sigma(1)}, \dots, u_{\sigma(p)}) \\ &= (-1)^\rho p!q! \ \alpha \wedge \beta(u_1, \dots, u_{p+q}) \ . \end{aligned}$$

□

2.3 Tangent space and tensor fields

Definition. Given a manifold M. A **curve through** x is a differentiable map $\gamma : (-1, 1) \to M$ with $\gamma(0) = x$.

Definition. Two curves γ_1, γ_2 are called **equivalent** if $D(\gamma_1)(0) = D(\gamma_2)(0)$. An equivalence class of curves through x is called a **tangent vector at** x. The set of all tangent vectors $Y(x) = D(\gamma)(0)$ at x is denoted by $T_x M$.

Lemma 2.3.1 *If M has dimension n, then $\psi : [\gamma] \in T_xM \mapsto D\gamma(0)(1) \in \mathbf{R}^n$ is a bijection between T_xM and $\mathbf{R}^n$.*

Proof. Denote by $[\gamma]$ the equivalence class of a curve γ, that is the set of all curves γ', which are equivalent to γ.

The map ψ is injective because two curves have the same image if and only if they are equivalent.

To show that the map is surjective, choose a chart (U, ϕ) with $x \in U$. Given $v \in \mathbf{R}^n$, define the curve $\gamma_i(t) = \phi^{-1}(\phi(x) + t\epsilon v)$, where ϵ is so small that $\phi(x) + t\epsilon v \in \phi(U)$ for all $t \in (-1, 1)$. By definition of the derivative, $D\phi_i(0)1 = \epsilon v$. Therefore, every $v \in \mathbf{R}^n$ is in the image of ψ. $\square$

This bijection makes T_xM to a vector space by defining $[\gamma_1] + [\gamma_2] = \psi^{-1}(\psi(\gamma_1) + \psi(\gamma_2))$ and $\lambda \cdot [\gamma_1] = \psi^{-1}(\lambda \cdot \psi(\gamma_1))$.

Every tangent vector $Y = [\gamma] \in T_xM$ defines a derivative of $f \in C^\infty(M)$ along Y defined by $Y(f) = \frac{d}{dt} f(\gamma(t))|_{t=0}$. One uses also the notation $Y(f) = Y.f = L_X f$. The possibility to identify a tangent vector as a derivation motivates the next notation.

Notation. A basis in T_xM is denoted by

$$e_i = \frac{\partial}{\partial x^i} .$$

The dual basis is denoted by

$$e^i = dx^i .$$

The reason for this notation becomes clear in the next section.

Definition. $TM = \bigcup_{x \in M} T_xM$ is called the **tangent bundle** of M. A **vector field** is a function $Y : M \to TM$ such that the function $x \mapsto Y_x(f)$ is smooth. The set of all vector fields on M is denoted by $\Gamma(TM)$.

Definition. The space $T^p_q(M) = \bigcup_{x \in M} T^p_q(T_xM)$ is called a (p, q)-**tensor bundle** on M. The set $T^0_1(M)$ is also called the **cotangent bundle** of M because $T^0_1(E)$ can be identified with E^*. There is also a natural identification of $T^1_0(E)$ with E so that $T^1_0(M)$ is identified with the tangent bundle.

Remark. If M is a n-dimensional manifold with atlas (U_i, ϕ_i), then the tangent bundle TM becomes a manifold with atlas $(TU_i, T\phi_i)$, where $T\phi_i : TU_i \to \mathbf{R}^n \times \mathbf{R}^n$ is defined by $T\phi_i(x, [\gamma]) = (\phi_i(x), D(\phi \circ \gamma)(0))$ and where TU_i is identified with $U_i \times \mathbf{R}^n$.

Definition. A 1-**form** on M is a function $\alpha : M \mapsto T^0_1(M)$ such that $\alpha(x) \in T^0_1(T_xM)$ and for any vector field Y, the function $\alpha(Y)$ given by $\alpha(Y)(x) = \alpha(x)(Y(x))$ is in $C^\infty(M)$.

Definition. A **tensor field** of type (p, q) on M is a function $t : M \to T^p_q(M)$ such that $t(x) \in T^p_q(T_xM)$ and for any 1-forms $\alpha^1, \dots, \alpha^p$ and vector fields $Y_1, \dots Y_q$ on M, the function $x \mapsto t(\alpha^1(x), \dots, \alpha^p(x), Y_1(x), \dots Y_q(x))$ is in $C^\infty(M)$. The space of all (p, q)-tensor fields is denoted by $\Gamma^p_q(M)$ or $\Gamma(T^p_q M)$.

Example. Let M, N be two manifolds and let $f : M \to N$ be a smooth map, then $D^k f$ is a tensor field of type $(0, k)$.

Definition. A **k-form** or k-differential form is a $(0, k)$-tensor field α such that $\alpha(x) \in \Lambda^k(T_x M)$. The exterior multiplication is defined pointwise $(\alpha \wedge \beta)(x) = \alpha(x) \wedge \beta(x)$.

It is useful to know, how differential forms transform under diffeomorphisms.

Definition. Let $\phi : M \to N$ be a diffeomorphism and let $\alpha \in \Lambda^p(M)$ be a differential form on N. Then

$$\phi^*\alpha(x)(u_1, \dots, u_p) = \alpha(\phi(x))(D\phi(x)(u_1), D\phi(x)(u_2), \dots, D\phi(x)(u_p))$$

is a differential form on M called the **pull-back** of ϕ.

2.4 Exterior derivative

Definition. For $f \in C^\infty(M)$ we define the **exterior derivative** $df(Y) = Y(f) = Y.f$. For $\alpha \in \Lambda^p(M)$, we define the **exterior derivative** $d\alpha$ as the $(p+1)$-form

$$\boxed{d\alpha = \sum_I d\alpha_I \wedge dx^I = \sum_{I,i} \frac{\partial \alpha_I}{\partial x_i} dx^i \wedge dx^I \, .}$$

The following proposition characterizes d uniquely as a "local, degree one antiderivation satisfying $d^2 = 0$":

Proposition 2.4.1 (Properties of the exterior derivative)
(i) d is of degree 1, that is a map from Λ^p to Λ^{p+1} and $df(Y) = Y(f) = Y.f$.
(ii) d is an antiderivation: it is linear and given $\alpha \in \Lambda^p$ and $\beta \in \Lambda^q$, then

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta \, .$$

(iii)

$$d^2 = d \circ d = 0 \, .$$

(iv) d is local: $d(\alpha|_U) = d\alpha|_U$, where $\alpha|U$ denotes the restriction of α to an open set $U \subset M$.
(v) d is the unique operation on $\Lambda(M)$ with properties (i)-(iv).

Proof. Property (i) holds by definition.
To see (ii), we use the fact that the permutation $iJ \mapsto Ji$ has the sign $(-1)^p$, if $|J| = p$. Using the Leibniz rule, we get

$$\begin{aligned} d(\alpha_I dx^I \wedge \beta_J dx^J) &= (\sum_i \frac{\partial \alpha_I}{\partial x^i} \beta_J + \alpha_I \frac{\partial \beta_J}{\partial x^i}) dx^i \wedge dx^I \wedge dx^J \\ &= (\sum_i \frac{\partial \alpha_I}{\partial x^i} dx^i \wedge dx^I \wedge \beta_J) + (-1)^p \alpha_I dx^I \wedge (\sum_i \frac{\partial \beta_J}{\partial x^i} dx^i \wedge dx^J) \\ &= d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta \, . \end{aligned}$$

(iii)

$$d^2 \alpha_I dx^I = \sum_{i,j} \frac{\partial^2}{\partial x^i \partial x^j} \alpha_I dx^i \wedge dx^j \wedge dx^I$$

is vanishing because partial derivatives commute and $dx^i \wedge dx^j$ anticommute: if we interchange the indices i, j, the sum changes sign, but it is obviously the same sum.

(iv) is clear because taking derivatives is a local operation: $\partial/\partial x^i(f(x)|_U) = \partial/\partial x^i f(x)|_U$.
(v) Consider the 0-form $f(x) = x^i$. From (ii), we know that $ddf = d(dx^i) = 0$. By (i), we have $ddx^I = 0$ for any $I = (i_1, \dots, i_p)$. From (i)-(iii), we know therefore that

$$d(\alpha_I dx^I) = d\alpha_I \wedge dx^I + \alpha_I \wedge d(dx^I) = (d\alpha_I) \wedge dx^I .$$

From the local property (iv), we can compute this in a local chart. From $d\alpha_I(Y) = Y(\alpha_I)$, we get $d\alpha_I(\partial/\partial x^i) = (\partial/(\partial x^i))\alpha_I dx^i$ and so $(d\alpha_I)_i = (\partial/(\partial x^i))\alpha_I$. If we plug in $d\alpha_I = \sum_i (d\alpha_I)_i dx^i$ into the relation $d(\alpha_I dx^I) = (d\alpha_I) \wedge dx^I$, we get the definition at the beginning of this section. $\square$

Reminder. Since $\alpha \in \Lambda^q(M)$ is a map from M to $T^0_q(M)$, $D\alpha(x)$ is a map from E to $T^0_q(E)$, that is $D\alpha(x)(u)$ maps $(u_1, \dots, u_p) \in E^p$ to $\mathbf{R}$.

Corollary 2.4.2 *Given $\alpha \in \Lambda^p(M)$, then*

$$d\alpha(u_0, \dots, u_p) = \sum_{j=0}^{p} (-1)^j D\alpha(x)(u_j)(u_0, \dots, \hat{u}_j, \dots, u_p) ,$$

where $\hat{u}_j$ means that this entry has been left out.

Proof. This is an exercice in Week 5. $\square$

2.5 Integration on manifolds

Definition. Let $\alpha = f dx^1 \wedge dx^2 \wedge \ldots \wedge dx^n$ be a n-form which is bounded on an open set $M \subset \mathbf{R}^n$. Define $\int_M \alpha = \int_M f(x)\, dx$.

Recall the (Change of variables): if ϕ is a diffeomorphism from $A \subset \mathbf{R}^n$ to $\phi(A) \subset \mathbf{R}^n$. Then

$$\int_{\phi(A)} f(x)\, dx = \int_A f(\phi(x)) |\det D\phi(x)|\, dx ,$$

where $D\phi$ is the **Jacobean** of ϕ.

Definition. The set

$$\Delta^n = \{x \in \mathbf{R}^{n+1} \mid \sum_{i=1}^{n+1} x^i = 1, x^i \geq 0\}$$

is called a **simplex**. We consider it also as a subset of $\mathbf{R}^n$, the plane in $\mathbf{R}^{n+1}$ spanned by Δ^n. For $n = 0$, the convention is that $\Delta^0 = \{0\}$.

Examples. A one dimensional simplex is homeomorphic to a closed interval. A two dimensional simplex is homeomorphic to a closed filled triangle, a three dimensional simplex is homeomorphic to a closed filled tetrahedron which simplies some proofs.

Remark. We take here the more convenient approach of homogeneous coordinates. A simplex can also be defined directly as a subset of $\mathbf{R}^n$. The advantage of considering the simplex also as a subset of $\mathbf{R}^{n+1}$ is that there is more symmetry in the coordinates.

Definition. Let M be a n dimensional manifold. A **singular p-simplex** or shortly **p-simplex** σ is a map $\sigma : \Delta^p \to M$ which extends to a smooth diffeomorphism from a neighborhood of $\Delta^p \subset \mathbf{R}^n$ into M. A 0-simplex is just a map from the point $\{0\}$ to a point in M.

Example. The change of variable formula at the beginning of this section can be reformulated as

$$\int_{\phi(A)} \omega = \int_A \phi^* \omega \; dx \; ,$$

because for a n-form ω, $\phi^*(\omega)(x) = \det(D\phi(x))\omega(\phi(x))$.

Definition. Let α be a p-form on M and σ a p-simplex. We can pull back the form α to a neighborhood of Σ_p in $\mathbf{R}^n$. Define the **integration of the p-form on a p-simplex** as

$$\int_\sigma \alpha = \int_{\Delta^p} \sigma^* \alpha \; .$$

2.6 Chains and boundaries

Definition. A **p-chain** in M is a finite linear combination $c = \sum_i c_i \sigma_i$, where σ_i are p-simplices and c_i are real numbers.

Definition. The integral of a p-form α

$$\int_\sigma \alpha = \int_{\Delta^p} \sigma^* \alpha \; .$$

can be extended from simplices to chains by linearity:

$$\int_c \alpha = \int_{\sum_i c_i \sigma_i} \alpha = \sum_i c_i \int_{\sigma_i} \alpha \; .$$

Definition. Define for $i = 0, \dots, p+1$ the maps $k_i^p : \Delta^p \to \Delta^{p+1}$ by

$$k_i^p(x) = (x^1, \dots, x^{i-1}, 0, x^i, \dots, x^p), i \geq 1 \; .$$

The sets $k_i^p(\Delta^p)$ are called the **faces** of the simplex Δ^{p+1}.

Definition. Given a manifold M and a p-simplex σ, we define the ***i*'th face** of σ to be the $(p-1)$-simplex $\sigma^i = \sigma \circ k_i^{p-1}$. The **boundary** of σ is defined as the $(p-1)$-chain $\delta\sigma = \sum_{i=0}^p (-1)^i \sigma^i$. The boundary map can linearly extended to all p-chains by

$$\delta(\sum_j a_j \sigma_j) = \sum_{i=0}^p \sum_j a_j (-1)^i \sigma_i^j \; .$$

Lemma 2.6.1 $\delta \circ \delta(c) = 0$ *for all chains c.*

Proof. For if $i \leq j$, then

$$k_i^{p+1} \circ k_j^p(x) = k_{j+1}^{p+1} \circ k_i^p(x) = (x^1, \dots, x^{i-1}, 0, x^i, \dots, x^{j-1}, 0, x^j, \dots, x^p) \; .$$

With $\delta\sigma = \sum_{i=0}^{p}(-1)^i \sigma \circ k_i^{p-1}$, we get

$$\begin{aligned}
\delta \circ \delta\sigma &= \sum_{j=0}^{p-1}\sum_{i=0}^{p}(-1)^{i+j}\sigma \circ k_i^{p-1} \circ k_j^{p-2} \\
&= \sum_{j<i}(-1)^{i+j}\sigma \circ k_i^{p-1} \circ k_j^{p-2} + \sum_{i\le j}(-1)^{i+j}\sigma \circ k_i^{p-1} \circ k_j^{p-2} \\
&= \sum_{j<i}(-1)^{i+j}\sigma \circ k_i^{p-1} \circ k_j^{p-2} + \sum_{i\le j}(-1)^{i+j}\sigma \circ k_{j+1}^{p-1} \circ k_i^{p-2} \\
&= \sum_{j<i}(-1)^{i+j}\sigma \circ k_i^{p-1} \circ k_j^{p-2} + \sum_{i<l} -(-1)^{i+l}\sigma \circ k_l^{p-1} \circ k_i^{p-2} = 0 \\
&= \sum_{j<i}(-1)^{i+j}\sigma \circ k_i^{p-1} \circ k_j^{p-2} + \sum_{j<i} -(-1)^{i+j}\sigma \circ k_i^{p-1} \circ k_j^{p-2} = 0\,.
\end{aligned}$$

□

Preview. This fact leads to **singular homology groups**. If S_p is the vector space of all singular chains on M, then $H_p(M, \mathbf{R}) = \ker\delta_p/\mathrm{im}\delta_{p-1}$ is a vector space called the p'th singular homology group of M with real coefficients. It turns out to be isomorphic to the dual space of the de Rham cohomology groups.

2.7 Theorem of Stokes for chains

The following theorem is a generalisation of the fundamental theorem of calculus, which says $\int_{\delta\sigma} f = \int_\sigma Df$ for a function on a one dimensional manifold. It also generalizes the theorem of Green in the plane or the theorems of Stokes or Gauss in the three dimensional space as you see partly in a homework.

Theorem 2.7.1 (Theorem of Stokes for chains) *Let c be a p-chain on a manifold M and let α be a $(p-1)$-form on a neighborhood of c in M. Then $\int_c d\alpha = \int_{\delta c} \alpha$.*

Proof. By linearity, it is enough to consider the case, where the p-chain consists of a single p-simplex. The claim $\int_c d\alpha = \int_{\delta c}\alpha$ is by definition of the integral equivalent to

$$\int_{\Delta^p} d(\sigma^*\alpha) = \sum_{i=0}^{p}(-1)^i \int_{\Delta^{p-1}} (\sigma^i)^*\alpha \quad . \tag{2.1}$$

We can assume $p \ge 2$, because for $p = 1$

$$\int_0^1 d\alpha(\sigma(t))(D\sigma(t))\,dt = \alpha(\sigma(1)) - \alpha(\sigma(0))$$

is the fundamental theorem of calculus. (See also the Week 6 homework set).

Consider $x^1, \dots, x^{p-1}$, which are coordinates of Δ^{p-1} because $x^0 = 1 - \sum_{i=0}^{p-1} x^i$. In this coordinates, the maps $k_i^{p-1} : \Delta^{p-1} \to \Delta^p$ are given by

$$\begin{aligned}
k_0^{p-1}(x^1, \dots, x^{p-1}) &= (1 - \sum_{k=1}^{p-1} x^k, x^1, \dots, x^{p-1}),\ i = 0 \\
k_i^{p-1}(x^1, \dots, x^{p-1}) &= (x^1, \dots, x^{i-1}, 0, x^i \dots x^{p-1}),\ i \ne 0\,.
\end{aligned}$$

The $(p-1)$-form $\beta = \sigma^*\alpha$ defined in a neighborhood of Δ^p can uniquely be written as $\beta = \sum_{j=1}^p \beta_j dx^1 \wedge \ldots \wedge \hat{dx^j} \ldots \wedge dx^p$, where $\hat{dx^j}$ means that this factor has been taken away. By linearity, we can assume that only one β_j is not vanishing so that

$$\beta = \beta_j dx^1 \wedge \ldots \wedge \hat{dx^j} \wedge \ldots \wedge dx^p$$

The left hand side of Equation (2.1) is therefore

$$(-1)^{j-1} \int_{\Delta^p} \frac{\partial \beta_j}{\partial x^j} dx^1 \wedge \ldots \wedge dx^p \; .$$

Let $y^1, \ldots, y^{p-1}$ be coordinates in the neighborhood of Δ^{p-1}. The pull backs of β under the boundary maps k_i^{p-1} are

$$\begin{aligned}
(k_0^{p-1})^*\beta(y) &= \beta_j(1-\sum_{k=1}^{p-1} y^k, y^1, \ldots, y^{p-1}) dy^1 \wedge \ldots \wedge dy^{p-1} \; , \\
(k_i^{p-1})^*\beta(y) &= 0, \; i \neq j \; , \\
(k_i^{p-1})^*\beta(y) &= \beta_j(y^1, \ldots, y^{j-1}, 0, y^j, \ldots, y^{p-1}) dy^1 \wedge \ldots \wedge dy^{p-1} \; , i = j \; .
\end{aligned}$$

We have therefore to prove

$$\boxed{\begin{aligned}
\int_{\Delta^p} \frac{\partial \beta_j}{\partial x^j}(x) dx^1 \ldots dx^p &= (-1)^{j-1} \int_{\Delta^{p-1}} \beta_j(1-\sum_k y^k, y^1, \ldots, y^{p-1}) \\
&- \beta_j(y^1, \ldots, y^{j-1}, 0, y^{j+1}, \ldots y^{p-1}) dy^1 \ldots dy^{p-1} \; .
\end{aligned}}$$

The left hand side is

$$\int_{\Delta^{p-1}} \left(\int_0^{1-\sum_{k \neq j} x^k} \frac{\partial \beta_j}{\partial x^j} dx^j \right) dx^1 \ldots \hat{dx^j} \ldots dx^p$$

which is, after performing the inner integral, equal to

$$\begin{aligned}
&\int_{\Delta^{p-1}} \beta_j(x^1, \ldots, x^{j-1}, 1-\sum_{k \neq j} x^k, x^{j+1}, \ldots, x^p) dx^1 \ldots \hat{dx^j} \ldots dx^p \\
-&\int_{\Delta^{p-1}} \beta_j(x^1, \ldots, x^{j-1}, 0, x^{j+1}, \ldots, x^p) dx^1 \ldots \hat{dx^j} \ldots dx^p \; .
\end{aligned}$$

In this expression, do in the first integral the change of variables

$$\begin{aligned}
(x^1, \ldots, \hat{x}^j, \ldots, x^p) \mapsto (y^1, \ldots, y^{p-1}) &= (x^2, \ldots, x^{j-1}, 1-\sum_{k \neq j} x^k, x^{j+1}, \ldots, x^p), \; j \neq 2 \\
(x^1, \ldots, \hat{x}^j, \ldots, x^p) \mapsto (y^1, \ldots, y^{p-1}) &= (1-\sum_{k \neq 2} x^k, x^3, \ldots, x^p), \; j = 2
\end{aligned}$$

which has the Jacobean determinant $(-1)^{j-1}$ and in the second integral the change of variables

$$(x^1, \ldots, \hat{x}^j, \ldots, x^p) \mapsto (y^1, \ldots, y^{p-1}) = (x^1, \ldots, x^{j-1}, x^{j+1}, \ldots, x^p)$$

which has the Jacobean determinant 1. We get

$$\begin{aligned}
&\int_{\Delta^{p-1}} (-1)^{j-1} \beta_j(1-\sum_{k=1}^{p-1} y^k, y^1, \ldots, y^{p-1}) dy^1 \ldots dy^{p-1} \\
-&\int_{\Delta^{p-1}} \beta_j(y^1, \ldots, y^{j-1}, 0, y^j, \ldots, y^{p-1})) dy^1 \ldots dy^{p-1} \; ,
\end{aligned}$$

which is the right hand side of the equation in the box. $\square$

2.8 Theorem of Stokes for oriented manifolds

Definition. A n-dimensional manifold M is called **orientable**, if there exists a nowhere vanishing n-form on M.

Remark. Denote by $\phi_{ij} : x \mapsto y = \phi_{ij}(x)$ the coordinate transformations from two overlapping charts (U_i, ϕ_i) and (V_i, ϕ_j). The pull back of a n-form $\omega = dy^1 \wedge \ldots \wedge dy^n$ is $\phi_{ij}^* dx^1 \wedge \ldots \wedge dx^n = \det D\phi_{ij}\omega$. A manifold is orientable, if and only there exists an atlas $\{U_i, \phi_i\}$ of M for which all coordinate changes ϕ_{ij} satisfy $\det(D\phi_{ij}) > 0$.

Definition. A subset D of M is called a **regular domain**, if for all points $x \in \delta D \subset M$, there exists a chart (U, ϕ) with $x \in U$ such that $\phi(U \cap D) = \phi(U) \cap \mathbf{H}^n$, where $\mathbf{H}^n = \{x \in \mathbf{R}^n \mid x_1 > 0\}$ is the half space in $\mathbf{R}^n$.

Remark. A regular domain D is just a n dimensional manifold with boundary inside the manifold M.

Definition. A n-form ω has **compact support**, if the closure of the set, where ω is nonvanishing, is compact.

Definition. Let ω be a n-form of compact support and let D be a regular domain in the orientable manifold M. Define $U_0 = M \setminus (\mathrm{supp}(\omega) \cap D)$. Cover the compact set $\mathrm{supp}(\omega) \cap D$ with a finite cover $\{U_i\}_{i=1}^k$. Assume, there exist simplices σ_i such that $U_i \subset \sigma_i(\Delta^n)$ and either $\sigma(\Delta^n) \subset \mathrm{int}(D)$ or $U_i = \sigma(V \cap \Delta^n)$, where $V \subset \mathbf{R}^n$ is open and such that $\delta\Delta^n \cap V$ is an open subset of a face which is mapped by σ_i into $U \cap \delta(D)$. Let $\{g_i\}_{i=0}^n$ be a partition of unity subordinate to the cover $\{U_i\}_{i=0}^n$. Define

$$\int_D \omega = \sum_{i=1}^k \int_{\sigma_i} g_i \omega \; .$$

Lemma 2.8.1 *The integral $\int_D \omega$ does not depend on the chosen cover and the chosen partition of unity.*

Proof. Let $\{V_i\}_{i=0}^l$ be an other such cover and $h_0, \ldots, h_l$ an other such partition of unity and denote by τ_i the underlying simplices. We have $\sum_{j=1}^l h_j = 1$ on $\mathrm{supp}(\omega) \cap D$ because $h_0 = 0$ on $\mathrm{supp}(\omega) \cap D$. We get

$$\sum_{i=1}^k \int_{\sigma_i} g_i \omega = \sum_{i=1}^k \int_{\sigma_i} \sum_{j=1}^l h_j g_i \omega = \sum_{i,j} \int_{\sigma_i} h_j g_i \omega \; .$$

Similarly we get

$$\sum_{j=1}^l \int_{\tau_j} h_j \omega = \sum_{i,j} \int_{\tau_j} h_j g_i \omega \; .$$

Since $\sigma_i^{-1} \circ \tau_j$ is an orientation -preserving diffeomorphism on an open possibly empty set where it is defined and since $\mathrm{supp}(h_j g_i \omega) \cap \sigma_i \Delta^n = \mathrm{supp}(h_j g_i \omega) \cap \tau_j \Delta^n$, it follows that

$$\begin{aligned} \int_{\sigma_i} (h_j g_i \omega) &= \int_{\Delta^n} \det D(\sigma_i) h_j g_i \omega(\sigma_i) \\ &= \int_{\Delta^n} \det D(\sigma_i) \det D(\sigma_i^{-1}\tau_j) h_j g_i \omega(\tau_j) \end{aligned}$$

$$= \int_{\Delta^n} \det D(\tau_j)(h_j g_i \omega(\tau_j)) = \int_{\tau_j} (h_j g_i \omega) \ .$$

□

Theorem 2.8.2 (Theorem of Stokes for orientable manifolds) *Let D be a regular domain in an oriented n-dimensional manifold M and let ω be a $(n-1)$-form of compact support. Then*

$$\int_{\delta D} \omega = \int_D d\omega \ .$$

Proof. Let $g_1, \dots, g_k$ and $\sigma_1, \dots, \sigma_k$ be chosen as in the definition of integration. Since $\sum_{i=1}^k g_i = 1$ on a neighborhood U of $\mathrm{supp}(\omega) \cap D$, we have $d(\sum_{i=1}^k g_i) = 0$ on U and

$$\sum_{i=1}^k d(g_i\omega) = \sum_{i=1}^k (dg_i) \wedge \omega + \sum_{i=1}^k g_i d\omega = d\omega \ .$$

If σ_i is a n-simplex which maped into the interiour of D, then

$$\int_{\delta\sigma_i} g_i\omega = 0 = \int_{\delta D} g_i\omega$$

because $\mathrm{supp}(g_i)\omega \subset \mathrm{int}D$. Assume now that σ_i is a n-simplex which has an image intersecting the boundary of D. In this case, $g_i\omega$ is vanishing on the boundary of $\sigma_i(\Delta^n)$ except possibly at the n'th face σ_i^n. Now σ_i^n is an orientation-preserving regular $(n-1)$-simplex in δD if n is even and orientation-reversing if n is odd. It follows that

$$\int_{\delta\sigma_i} g_i\omega = (-1)^n \int_{\sigma_i^n} g_i\omega = (-1)^n(-1)^n \int_{\delta D} g_i\omega = \int_{\delta D} g_i\omega \ .$$

Using the Stokes theorem for chains, we conclude

$$\int_D d\omega = \sum_i \int_D d(g_i\omega) = \sum_i \int_{\sigma_i} d(g_i\omega) = \sum_i \int_{\delta\sigma_i} g_i\omega = \sum_i \int_{\delta D} g_i\omega = \int_{\delta D} \omega \ .$$

□

The theorem of Stokes already proven for chains and for orientable manifolds with orientable boundary has different generalisations.

1) Stokes for **nonorientable manifolds**.
We would like to have for any n-dimensional manifold M and $(n-1)$-form α the property $\int_M d\alpha = \int_{\delta M} \alpha$. The later integral should more precicely be written as $\int_{\delta M} i^*\alpha$, where $i : \delta M \to M, x \mapsto x$ is the embedding of δM into M. It is not possible to make sense of the integrals in a global way. One defines an orientable vector bundle $\tilde{M}$ over M and considers so called **twisted k-forms** over M. The theorem of Stokes can so be defined also in nonorientable situations like on a Moebius strip or Klein bottle. There is no indication that space time is not orientable (you could make a tour in our universe and when you return, a right glove had turned into a left glove).

2) Stokes for manifolds with **piecewise smooth boundaries**.
Manifolds with boundary can be generalized to manifolds with piecewise smooth boundaries. An

example is the cube $M = \{|x| + |y| \leq 1 \subset \mathbf{R}^2\}$. As long as the set of singular points in the boundary, (that is points in which the coordinate maps ϕ_i fail to be smooth) has measure zero, the theorem of Stokes stays true.

3) Stokes for **Riemannian manifolds**. We will discuss later this generalisation. It is a reformulation of Stokes for vector fields. In a "curved" manifold, the identification of vectors with 1-forms is space dependent.

Chapter 3

Riemannian geometry

3.1 Metric tensor

Definition. Let E be a finite dimensional vector space. A **metric tensor** on E is a symmetric tensor $g \in T_2^0(E)$, which is **nondegenerate**: that is

$$g(u, v) = 0, \forall v \in E \Rightarrow u = 0 \ .$$

Remark.
The symmetric tensor g is nondegenerate if and only if the map $v \mapsto g(v, \cdot)$ from E to E^* is an isomorphism.
(This is no more true if E would be infinite dimensional. The above definition of degeneracy is then called **weakly nondegenerate** and the property that $v \mapsto g(v, \cdot)$ from E to E^* is an isomorphism is called **strongly nondegenerate**).

Remark. Let e_i be a basis in E. Let u^i, v^i be the coordinates of u, v, so that $u = \sum_i u^i e_i$ and $v = \sum_i v^i e_i$. Let g_{ij} be the coordinates of g defined by $g_{ij} = g(e_i, e_j)$. Then $g(u, v) = \sum_{ij} g_{ij} u^i v^j$. The symmetry implies that $g_{ij} = g_{ji}$. If we consider g_{ij} as a $n \times n$ matrix, then the nondegeneracy is equivalent with the invertibility of the matrix g.

Definition. An **orthonormal basis** of E relative to g is a basis e_i in E which satisfies $g(e_i, e_j) = \pm 1$ for $i = j$ and $g(e_i, e_j) = 0$ for $i \neq j$.

Remark. With respect to an orthonormal basis $\{e_i\}_{i=1}^n$, the metric tensor g is $g = \sum_{i,j} c_i e^i \otimes e^i$, where $c_i = g(e_i, e_i) \in \{-1, 1\}$.

Lemma 3.1.1 (Gramm-Schmidt) *For every metric tensor g, there exists an orthonormal basis relative to g.*

Proof. Let e_i any basis. Put $f_i = e_1$. Assume $h_1, \ldots, h_k$ is an orthogonal set already constructed, define inductively

$$h_{k+1} = e_{k+1} - \sum_{j=1}^{k} \frac{g(h_j, e_{k+1})}{g(h_j, h_j)} h_j$$

which is nonzero: (else $e_1, \dots, e_{k+1}$ would be linearly independent). We verify $g(h_{k+1}, h_i) = 0$ for $i \neq k+1$:

$$\begin{aligned} g(h_i, h_{k+1}) &= g(h_i, e_{k+1}) - \sum_{j=1}^{k} \frac{g(h_j, e_{k+1})}{g(h_j, h_j)} g(h_i, h_j) \\ &= g(h_i, e_{k+1}) - g(h_i, e_{k+1}) = 0 \; . \end{aligned}$$

Like this, we get an orthogonal set $h_1, \dots, h_{k+1}$. The orthogonal basis $f_k = h_{k+1}/\sqrt{|g(h_{k+1}, h_{k+1})|}$ satisfies $g(f_{k+1}, f_{k+1}) = \pm 1$ and is therefore also orthonormal. □

Remark. Note that an orthonormal basis is **not unique** because we can replace the basis $\{e_i\}_{i=1}^n$ with the basis $\{\pm e_i\}_{i=1}^n$.

Definition. Assume E is finite dimensional of dimension n. Let $e_1, \dots, e_n$ be an orthonormal basis with respect to g and let $e^1, \dots, e^n$ be the dual basis. The tensor $\mu \in \Lambda^n(E)$ defined by

$$\mu = e^1 \wedge e^2 \wedge \dots \wedge e^n$$

is called the **volume element** of E with respect to g.

Lemma 3.1.2 *There are only two volume elements on E with respect to a given metric g.*

Proof. Let $e_1, \dots, e_n$ be an orthonormal basis and μ the corresponding volume form. We have by definition $\mu(e_1, \dots, e_n) = 1$, if $e_1, \dots, e_n)$ is a basis in E because

$$e^1 \wedge \dots \wedge e^n = \sum_{\sigma \in S_n} e^{\sigma(1)} \otimes \dots \otimes e^{\sigma(n)} \; .$$

Any other basis $f_1, \dots, f_n$ is obtained from $e_1, \dots, e_n$ by a orthogonal transformation $f_i = Ue_i$. If the determinant of $U : E \to E$ is 1, then the corresponding volume forms are the same. (If $f_i = Ue_i$ is a change of the basis, then the corresponding volume forms satisfy $f^1 \wedge ... \wedge f^n = \det(U) e^1 \wedge ... \wedge e^n$.)

If we had constructed μ with an orthogonal transformation U of determinant -1 like the basis $f_1 = -e_1, f_i = e_i, i = 2, \dots, n$, then we had $\mu(f_1, \dots, f_n) = -1$.) Because the only ambiguity in the choice of the basis was the sign of the determinant of U, the volume element is determined up to a sign. □

Definition. A choice of the two possible volume elements in E is called an **orientation** in E. If μ is a volume element in E and $v_1, \dots, v_n$ is a basis, then it is called **positively oriented**, if $\mu(v_1, \dots, v_n) = 1$.

3.2 Hodge star operation

Notation. We write for

$$\alpha = \sum_{i_1 < i_2 \dots < i_q} \alpha_{i_1, \dots, i_q} e^{i_1} \wedge \dots \wedge e^{i_q}$$

also shortly $\alpha = \sum_I \alpha_I e^I$, where I runs over all index sets $I = \{i_1, \dots, i_q\}$ with $i_1 < i_2, \dots < i_q$.

Definition. Given $\alpha = \sum_I \alpha_I e^I, \beta = \sum_I \beta_J e^J \in \Lambda^q(E)$, where $\{e_i\}$ is an orthonormal basis with respect to the metric tensor $g \in T_2^0(E)$. Define $g_I = \prod_{i \in I} g(e_i, e_i) \in \{-1, 1\}$. Define

$$g(\alpha, \beta) = \sum_I g_I \alpha_I \beta_I \ .$$

Examples.
1) Let $\alpha = \alpha_1 e^1 + \alpha_2 e^2$ and $\beta = \beta_2 e^1 + \beta_2 e^2$. Then $g(\alpha, \beta) = \alpha_1 \beta_1 g_{11} + \alpha_2 \beta_2 g_{22}$.
2) Let $\alpha = e_1 \wedge e_2$ and $\beta = e_1 \wedge e_2$. Then $g(\alpha, \beta) = g(e_1, e_1) \cdot g(e_2, e_2)$.

Proposition 3.2.1 (Lifting of a metric to $\Lambda^q(E)$.) *A metric tensor g on E induces a metric tensor on $\Lambda^q(E)$ by $g(\alpha, \beta) = \sum_I g_I \alpha_I \beta_I$.*

Proof. g is obviously a symmetric tensor in $T_2^0(\Lambda^q(E))$. The nondegeneracy follows from the nondegeneracy of g on E: assume $g(\alpha, \beta) = 0$ for all β. Then $g(\alpha, dx^I) = g_I \alpha_I = 0$ for all I, so that $\alpha_I = 0$ for all I and so $\alpha = 0$. □

Example. Let $e_1, e_2, \dots, e_n$ be an orthonormal basis in E. In this basis, we have $g = \sum_i c_i e^i \otimes e^i$, where $c_i \in \{-1, 1\}$. Then

$$g(e^I, e^I) = g(e^{i_1} \wedge \cdots \wedge e^{i_q}, e^{i_1} \wedge \cdots \wedge e^{i_q}) = c_{i_1} \dots c_{i_q}$$

and $g(e^J, e^I) = 0$ for $I \neq J$. So, e^I is an orthonormal basis with respect to the metric g.

Proposition 3.2.2 (Existence of a Hodge operation) *Assume E is a finite dimensional vector space of dimension n. Let g be a metric on E and ω a volume element on E with respect to E. There exists a unique isomorphism*

$$* : \Lambda^q(E) \to \Lambda^{n-q}(E)$$

such that $\alpha \wedge *\beta = g(\alpha, \beta)\omega$ for all $\alpha, \beta \in \Lambda^q(E)$.

Proof. For $\gamma \in \Lambda^{n-q}(E)$, define the linear map $\phi_\gamma : \Lambda^q(E) \to \mathbf{R}$ by

$$\phi_\gamma(\alpha)\omega = \alpha \wedge \gamma \ .$$

This map is injective. A comparison of the dimension shows that it is an isomorphism. Therefore, for all $\beta \in \Lambda^q(E)$, there exists a unique $\gamma \in \Lambda^{n-q}(E)$ such that $\phi_\gamma(\alpha) = g(\alpha, \beta)$. Define $*\beta$ to be this γ.

The uniqueness (and also existence) of $*$ is best seen using coordinates: let $e_1, \dots, e_n$ be an orthonormal basis in E. Given $e^I = e^{i_1} \wedge \ldots \wedge e^{i_q} \in \Lambda^q(E)$. Then, with $c_i = g(e_i, e_i)$, we have

$$*e^I = (c_{i_1} \cdots c_{i_q})(-1)^{\sigma(I, I')} e^{I'} \ ,$$

where $I \cup I' = \{1, \dots, n\}$ and $\sigma(I, I')$ is the permutation which brings $I \cup I'$ into $\{1, \dots, n\}$, with $I = \{i_1, \dots, i_q\}$ with $i_1 < \ldots < i_q$ and $I' = \{i'_1, \dots, i'_q\}$. □

Notation. If e_i is an orthonormal basis with respect to g, then $g = \sum_{i,j} c_i e^i \otimes e^i$ with $c_i = g(e_i, e_i)$ as in the above proof. Define $(-1)^g = \prod_{i=1}^n c_i$.

Proposition 3.2.3 (Properties of the Hodge star operator) .
a) For $\alpha \in \Lambda^q(E)$, *we have*

$$* * \alpha = (-1)^g (-1)^{q(n-q)} \alpha \ .$$

b) $g(\alpha, \beta) = (-1)^g g(\alpha^*, \beta^*)$.
c) $*1 = \omega,\ *\omega = (-1)^g$.

Proof. a) Take $\alpha = e^{\sigma(q+1)} \wedge \ldots \wedge e^{\sigma(n)}$. We know that $*\alpha = C e^{\sigma(1)} \wedge \ldots \wedge e^{\sigma(q)}$ for some constant C. Because $\alpha \wedge *\alpha = (-1)^q (c_{\sigma(q+1)} \cdots c_{\sigma(n)}) \omega$

$$* e^{\sigma(1)} \wedge \cdots \wedge e^{\sigma(q)} = c_{\sigma(1)} \ldots c_{\sigma(q)} (-1)^{\sigma} e^{\sigma(q+1)} \wedge \cdots \wedge e^{\sigma(n)} \ .$$

Therefore,

$$\begin{aligned} * * e^{\sigma(1)} \wedge \cdots \wedge e^{\sigma(q)} &= (-1)^{q(n-q)} (c_{\sigma(1)} \cdots c_{\sigma(n)}) ((-1)^{\sigma})^2 e^{\sigma(1)} \wedge \cdots \wedge e^{\sigma(q)} \\ &= (-1)^{q(n-q)} (-1)^g e^{\sigma(1)} \wedge \cdots \wedge e^{\sigma(q)} \ . \end{aligned}$$

b) We have, using a),

$$g(*\alpha, *\beta)\omega = *\alpha \wedge * * \beta = (-1)^g (-1)^{q(n-q)} * \alpha \wedge \beta = (-1)^g \beta \wedge *\alpha = (-1)^g g(\alpha, \beta)\omega \ .$$

c) follows directly from the definition and b). □

Examples.
1) Let $M = \mathbf{R}^3$ and let g be the **flat standard metric** in M. Let e^1, e^2, e^3 be the standard basis in $\Lambda^1(\mathbf{R}^3)$. The 3-form $\omega = dx^1 \wedge dx^2 \wedge dx^3$ is the volume form on M associated to g. Then $*e^1 = e^2 \wedge e^3, *e^2 = e^3 \wedge e^1 = -e^1 \wedge e^3$ and $*e^3 = -e^1 \wedge e^2$ because in each case $e^i \wedge *e^i = \omega$. From the above proposition a), we see that $* * \alpha = \alpha$ so that $*$ is an involution.

2) Let $M = \mathbf{R}^4$ and g the **Lorentz inner product**, which is defined by $g_{11} = -1, g_{22} = g_{33} = g_{44} = 1$ and $g_{ij} = g(e_i, e_j) = 0$ for $i \neq j$. The Hodge star operation from $\Lambda^1(\mathbf{R}^4)$ to $\Lambda^3(\mathbf{R}^4)$ is an involution and given by

$$\begin{aligned} *e^1 &= -e^2 \wedge e^3 \wedge e^4 \ , \\ *e^2 &= -e^1 \wedge e^3 \wedge e^4 \ , \\ *e^3 &= e^1 \wedge e^2 \wedge e^4 \ , \\ *e^4 &= -e^1 \wedge e^2 \wedge e^3 \end{aligned}$$

because

$$\begin{aligned} e^1 \wedge (*e^1) &= -e^1 \wedge e^2 \wedge e^3 \wedge e^4 = g_{11}\omega \\ e^2 \wedge (*e^2) &= -e^2 \wedge e^1 \wedge e^3 \wedge e^4 = g_{22}\omega \\ e^3 \wedge (*e^3) &= e^3 \wedge e^1 \wedge e^2 \wedge e^4 = g_{33}\omega \\ e^4 \wedge (*e^4) &= -e^4 \wedge e^1 \wedge e^2 \wedge e^3 = g_{44}\omega \ . \end{aligned}$$

The Hodge star operation from $\Lambda^2(\mathbf{R}^4$ to $\Lambda^2(\mathbf{R}^4)$ satisfies $** = -1$ and is given by

$$\begin{aligned} *(e^1 \wedge e^2) &= -e^3 \wedge e^4 \ , \\ *(e^2 \wedge e^3) &= e^1 \wedge e^4 \ , \\ *(e^1 \wedge e^3) &= e^2 \wedge e^4 \end{aligned}$$

because

$$\begin{aligned}(e^1 \wedge e^2) \wedge (-e^3 \wedge e^4) &= -g_{11} g_{22} \omega = \omega \\ (e^2 \wedge e^3) \wedge (e^1 \wedge e^4) &= g_{22} g_{33} \omega = \omega \\ (e^1 \wedge e^3) \wedge (e^2 \wedge e^4) &= -g_{11} g_{33} \omega = \omega \, .\end{aligned}$$

3.3 Riemannian manifolds

Definition. A **metric tensor field** g is a tensor field $g \in T_2^0(M)$ such that $g(x)$ is a metric tensor in $T_2^0(T_x M)$. Recall that this means that there exist bilinear symmetric nondegenerate functions $g_x(\cdot,\cdot) : T_x M \times T_x M \to \mathbf{R}$, so that for all vector fields $X, Y \in \Gamma(TM)$, the function $x \mapsto g(x)(X(x), Y(x))$ is smooth.

Definition. A pair (M, g), where M is a manifold and g is a metric tensor field on M, is called a **pseudo Riemannian manifold**.

Definition. A pseudo Riemannian manifold (M, g) is called a **Riemannian manifold**, if the metric tensor field g is positive definite, that is if $g(x)(v, v) \geq 0$ for all $v \in T_x(M)$.

Definition. Given an n-dimensional Pseudo Riemannian manifold M. A smooth map $x \to TM^n, x \mapsto e(x) = (e_1(x), \dots, e_n(x))$ such that for all x, the set $\{e_1(x), \dots, e_n(x)\}$ is an orthonormal basis on $T_x(M)$, is called an **orthonormal frame field** on M. Denote by $e^1, \dots, e^n$ the dual basis of $e_1, \dots, e_n$. The map $x \to T^*M^n, x \mapsto e(x) = (e^1(x), \dots, e^n(x))$ is called an **orthonormal coframe field**. It defines a volume form $\omega = e^1 \wedge \dots \wedge e^n \in \Lambda^n(M)$.

Remark. By Gramm-Schmidt orthonomalisation at each point $x \in M$, as seen in the last section, a volume form exists in every chart U. Since coordinate changes ϕ_{ij} from one chart to an other chart have an orthogonal Jacobean $D\phi_{ij}$, the determiant is ± 1. If M is oriented, we can choose ω so that all coordinate changes ϕ_{ij} satisfy $\det(D\phi_{ij}) = 1$. Such a volume form ω is called the **volume** on M.

Definition. An **orientation** on an orientable manifold is an equivalence class of volume forms, where $\mu_1 \sim \mu_2$ if and only if there exists a function $f > 0$ such that $\mu_1 = f\mu_2$.

Lemma 3.3.1 *Assume M is orientable. Then M is connected if and only if there exist exactly two orientations on M.*

Proof. Let μ and ν be two volume forms. Then $\mu = f\nu$ with $f(x) \neq 0$ for all $x \in M$. Because M is connected, we have $f^{-1}((0,\infty)) = M$ or $f^{-1}((-\infty, 0)) = M$ so that either $f > 0$ or $f < 0$ and μ, ν are equivalent.

If M is not connected and is a union of two disjoint manifolds M_1, M_2. Let μ be a volume on M and denote by μ_i the restriction on M_i. Let ν be a volume on M which is μ_1 on M_1 and $-\mu_2$ on M_2. Then, $\mu, -\mu$ and $\nu, -\nu$ are not equivalent. □

We come now to some examples of Pseudo Riemannian manifolds. We use the notation to denote by $[g_{ij}]$ the matrix which has the entry g_{ij} in the i'th row and the j'th column.

Examples.
1) (**Flat three dimensional space in spherical coordinates**). Let (M, g) be the flat three dimenensional space. With respect to the standard basis e_1, e_2, e_3, the metric g is given by $g_{ij}(x) = 1$ if $i = j$ and $g_{ij}(x) = 0$ if $i \neq j$.
Let $r = x^1, \theta = x^2, \phi = x^3$ be the spherical coordinates in $\mathbf{R}^3$. A point in $\mathbf{R}^3$ can be written as

$$x = r\cos(\phi)\sin(\theta)e_1 + r\sin(\phi)\sin(\theta)e_2 + r\cos(\theta))e_3 .$$

Consider the three orthogonal basis vectors

$$\begin{aligned} e_r &= \partial x/\partial x^1 = \cos(\phi)\sin(\theta)e_1 + \sin(\phi)\sin(\theta)e_2 + \cos(\theta)e_3 \\ e_\theta &= \partial x/\partial x^3 = r\cos(\phi)\cos(\theta)e_1 + r\sin(\phi)\cos(\theta)e_2 - r\sin(\theta)e_3 , \\ e_\phi &= \partial x/\partial x^2 = -r\sin(\phi)\sin(\theta)e_1 + r\cos(\phi)\sin(\theta)e_2 . \end{aligned}$$

It is an exercice in the seven'th week homework to show that with respect to this basis, the metric is given by

$$[g_{ij}](x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2(\theta) \end{pmatrix} .$$

2) (Flat Lorentz manifold) Take on the manifold $M = \mathbf{R}^4$ the metric

$$[g_{ij}] = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ,$$

where c is the speed of light.

3) (Schwarzschild manifold). Take on the manifold $M = \mathbf{R}^4 \setminus \{r = 0\}$ the coordinates $t = x^0, r = x^1, \theta = x^2, \phi = x^3$ (that is x^1, x^2, x^3 are the polar coordinates in space and x^0 is the time coordinate. For $m/r < 1$, the **Schwarzschild metric** is given by

$$[g_{ij}] = \begin{pmatrix} -c^2(1 - 2m/r) & 0 & 0 & 0 \\ 0 & (1 - 2m/r)^{-1} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\sin^2(\theta)) \end{pmatrix} .$$

The Schwarzschild manifold is important in **general relativity**. It is a model for the space-time in presence of a single massive object of mass M. (It is according to a **theorem of Birkhoff** the only solutions of Einstein's equations (see later) in the vaccuum which is spherically symmetric). The constant m is defined as $m = GM/c^2$, where G is the gravitational constant. ($2m$ is called the **Schwarzschild radius**). The Schwarschild model is the basis for the known five tests of general relativity like the perihelion advance of planets, the bending of light near the sun, the time delay in radar sounding of planets, the spectral shift of light emanating a massive object or the precession of a gyroscope which is freely orbiting the earth.

Definition. Given a Riemannian manifold (M, g) and let u be a vector in the tangent space T_xM. The **length** of u is defined as

$$\sqrt{g(x)(u, u)} = \sqrt{|\sum_{ij} g_{ij}u^iu^j|} .$$

A vector, which has zero length is called **null**. One calles vectors u for which $\sum_{ij} g_{ij}u^iu^j < 0$ **time like**, vectors u with $\sum_{ij} g_{ij}u^iu^j < 0$ **space like**.

Examples. 1) Take example 2) from above. The vectors on the **light cone**

$$\{u \mid \sqrt{|g(u,u)|} = \sqrt{|-c^2(u^0)^2 + (u^1)^2 + (u^2)^2 + (u^3)^2| = 0}\}$$

are the null vectors.
2) In a Lorentzian manifold (M, g) with $[g] = Diag(-1, 1, 1, 1)$, the vector $(1, 0, 0, 0)$ is time like and for example $(0, 1, 0, 0)$ is space like. Particles with space like velocity vectors have velocity smaller than the speed of light, particles with time like velocity vectors have velocity larger than the speed of light (tachions).

Definition. The **angle** θ between two non-null vectors u, v is defined as

$$\cos(\theta) = g(u,v)/(\sqrt{g(u,u)}\sqrt{g(v,v)}) .$$

Definition. Let $\gamma : [a, b] \to M$ be a curve on the Riemannian manifold M. The **length** of the curve is defined by

$$\int_a^b \sqrt{|g(\gamma(\tau))(D\gamma(\tau)(1), D\gamma(t)(\tau))|}\, d\tau .$$

Note that $D\gamma(\tau)$ is a linear map from $\mathbf{R}$ to $\mathbf{R}^n$ and that $D\gamma(t)(1)$ is therefore a vector in $\mathbf{R}^n \sim T_{\gamma(\tau)}M$.

Question. Does every manifold allow to be a Riemannian manifold (g positive definite) or a pseudo Riemannian manifold of a certain signature?

Proposition 3.3.2 *Given a manifold M. Then there exists a Riemannian metric g on M, that is: there exists a tensor field $g \in T_2^0(M)$ which is symmetric, nondegenerate and positive definite.*

Proof. Leg $\{U_i, \phi_i\}$ be an atlas for M and let g_i be a partition of unity subordinate to the cover U_i. Let q be a Riemannian metric on $\mathbf{R}^n$. For example $[q] = \mathrm{Diag}(1, 1, 1 \dots, 1)$. Let $q_i = \phi_i^* q$ be the pull back metrics on U_i. Define $g(x) = \sum_i g_i(x)q_i(x)$. This is smooth and positive definite because for $x \in M$ and $u \in T_xM$, we have $g(x)(u,u) = \sum_i g_i q_i(u,u) > 0$. □

Remark. It is not always possible to build on a given manifold a metric of a given signature. For example, on the sphere $M = S^2$, there exists no Lorentzian metric, that is a metric of signature $(-1, 1)$ because we could else find a vector field on the sphere, which is nowhere vanishing. This does not exist.

3.4 Theorem of Stokes for Riemannian manifolds

Definition. Assume (M, g) is an oriented Riemannian manifold with volume form ω. The point-wise Hodge operation $* : \Lambda^k(T_xM) \to \Lambda^{n-k}(T_xM)$ defines a map $\Lambda^k(M) \to \Lambda^{n-k}(M)$ called the **Hodge star** operation on $\Lambda(M)$.

Definition. Let M be an oriented Riemannian manifold with volume form ω. For a continuous function f on M with compact support, define

$$\int_M f := \int_M *f = \int_M f\omega$$

Notation. Let M be a Riemannian manifold. The metric g defines an isomorphism from T_xM to T_x^*M. A common musical notion for this isomorphism is

$$\flat : T_xM \to T_x^*M,\ \sharp : T_x^*M \to T_xM$$

which comes from the fact that $\flat$ is **index raising** and $\sharp$ is **index lowering**.

Recall. Given a vector field $v \in \Gamma(TM)$. Define the function

$$\mathrm{div}(v) = *d * v^\flat .$$

Theorem 3.4.1 (Stokes for Riemannian manifolds) *Given $v \in \Gamma(TM)$ of compact support and let D be a regular domain in M and n the outer normal vector field on the boundary δD. Then*

$$\int_D \mathrm{div}(v) = \int_{\delta D} g(v, n) .$$

Proof. Using Stokes theorem, we get by definition of the integral and Stokes theorem for oriented manifolds that

$$\int_D \mathrm{div}(v) = \int_D *d * v^\flat = \int_D d * v^\flat = \int_{\delta D} *v^\flat$$

which reduces the claim to the identity

$$\int_{\delta D} *v^\flat = \int_{\delta D} g(v, n) .$$

Choose at a point in δD a chart so that $n(x) = e_1$ is the outer unit normal vector. Let $e_1, \dots, e_n$ be an orthonormal basis and let $e^1, \dots, e^n$ be the dual basis. Using a suitable partition of unity, we can assume that v has its support in a chart U which is contained in a simplex σ, such that σ_1 is mapped into δD. With $v^\flat = \sum_i v_i e^i$, where $v_i = g_{ii} v^i$, we get

$$\begin{aligned} *v^\flat &= \sum_i v_{1,\dots,\hat{i},\dots,n} e^1 \wedge \dots \wedge \hat{e^i} \wedge \dots \wedge e^n \\ &= \sum_i g_{ii} (-1)^i v_i e^1 \wedge \dots \wedge \hat{e^i} \wedge \dots \wedge e^n \end{aligned}$$

so that $\int_{\delta D} *v^\flat = \int_{\sigma_1} \det(D\sigma_1) g_{11} v_{2,\dots,n}$. On the other hand, since $g(v, n) = v_1$, we obtain by the definition of the integration on δD also $\int_{\sigma_1} \det(D\sigma_1) v_1 = \int_{\sigma_1} \det(D\sigma_1) g_{11} v_{2,\dots,n}$. □

3.5 Connections

Definition. A curve ,$\phi : I \to M, t \mapsto \phi_t$ is called an **integral curve** of a vector field X, if $\frac{d}{dt}\phi_t(x) = X(\phi_t(x))$.
Every differentiable vector field defines so a set of diffeomorphism $\phi_t : U \to \phi_t(U)$ on M, if $|t|$ is small enough. This set of diffeomorphisms is also called a **flow**.

Definition. Let $X \in \Gamma(TM)$ be a vector field. Denote by ϕ_t the flow defined by the vector field X. Given an other vector field Y, define $L_X Y(x) = \frac{d}{dt} Y(\phi_t(x))_{t=0}$. It is called the **Lie derivative** of Y with respect to the vector field X.

Definition. Let X, Y be vector fields on M. Define the vector field $[X, Y]$ called **Lie bracket** as

$$[X, Y] = L_X Y - L_Y X .$$

Remark. Write $\partial_i = \partial/\partial x^i$. In local coordinates, $X = \sum_i X^i \partial_i, Y = \sum_i Y^i \partial_i$, we have $Z = [X,Y] = \sum_i Z^i \partial_i$ with

$$Z^i = \sum_{j=1}^n X^j \partial_j Y^i - \sum_{j=1}^n Y^j \partial_j X^i \,.$$

Definition. A **connection** on a manifold M is a bilinear map $\nabla : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ which satisfies

$$\nabla_{fX}(Y) = f\nabla_X(Y),\ \nabla_X(fY) = (Xf)Y + f\nabla_X(Y)$$

for all $f \in C^\infty(M)$ and which has vanishing **torsion**

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] = 0 \,.$$

Example. The canonical conection on Euclidean space is $\nabla_X Y = DY(X) = \sum_j X^j \partial_j Y^i$.

Definition. A connection on a pseudo Riemannian manifold (M,g) is called **compatible with the metric** if for all vector fields X, Y, Z, the identity

$$Xg(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

holds. (This is motivated by the Leibniz rule).

Theorem 3.5.1 (Fundamental theorem of Riemannian geometry)
On every pseudo Riemannian manifold (M,g), there exists a unique connection ∇ which is compatible with the metric.

Proof. Because ∇ is compatible with the metric, we obtain the three identities

$$\begin{aligned} Xg(Y,Z) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \,, \\ Yg(Z,X) &= g(\nabla_Y Z, X) + g(Z, \nabla_Y X) \,, \\ Zg(X,Y) &= g(\nabla_Z X, Y) + g(X, \nabla_X Y) \,. \end{aligned}$$

Adding the first two equalities and subtracting the third one and using that ∇ is torsion free, this gives

$$\boxed{\begin{aligned} 2g(\nabla_X Y, Z) &= Xg(Y,Z) + Yg(Z,X) - Zg(X,Y) \\ &+ g([X,Y],Z) - g([X,Z],Y) - g([Y,Z],X) \end{aligned}}$$

which shows uniqueness. The right hand side is linear and is for fixed X, Y a tensor in Z. It satisfies especially $g(\nabla_X Y, fZ) = f(\nabla_Y, Z)$ for all $f \in C^\infty(M)$. The formula for $g(\nabla_X Y, Z)$ defines ∇_X.

(i) $\nabla_{fX}(Y) = f\nabla_X(Y)$.
It is enough to show $g(\nabla_{fX}(Y), Z) = g(f\nabla_X(Y), Z)$ for all vector fields Z. This is verified by putting in the formula in the box.
(ii) $\nabla_X(fY) = (Xf)Y + f\nabla_X(Y)$.
Also here, it is enough to show $g(\nabla_X(fY), Z) = g((Xf)Y, Z) + g(f\nabla_X(Y), Z)$ for all vector fields Z which is verified by inserting the formula in the box.
(iii) $\nabla_X Y - \nabla_Y X - [X,Y] = 0$
Dito.
(iv) $Xg(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$.
Substitute the formula on the right hand side of the formula in the box.

□

Definition. The unique connection compatible with the Pseudo Riemannian metric is called the **Levi-Civita connection** or **canonical connection**.

Definition. In a local coordinate system, take $X_i = \partial_i$, $g_{ij} = g(X_i, X_j)$, and $[g^{ij}] = [g_{ij}]^{-1}$, so that $g = \sum_{ij} g_{ij} dx^i \otimes dx^j$. Use the short hand notation ∇_i for ∇_{X_i}, define

$$\Gamma^i_{jk} = \sum_l g^{il} g(\nabla_j \partial_k, \partial_l) \ ,$$

the **Christoffel symbols** of the connection. Alternatively, the Christoffel symbols can be defined by the relation

$$\nabla_i \partial_j = \sum_k \Gamma^k_{ij} \partial_k \ .$$

It is an exercice in Homework 8, to see that these two definitions are equivalent.

Remark. Because the vector fields $\frac{\partial}{\partial x^i}$ commute, we have $\Gamma^i_{jk} = \Gamma^i_{kj}$.

Remarks.
1) The equality $\nabla_{fX} Y = f \nabla_X Y$ for $f \in C^\infty(M)$ implies that for a given Y, $\nabla_{fX}(Y)(x)$ depends only on $X(x)$. The connection can therefore by viewed as a family of bilinear maps from $T_x M \times \Gamma(TM) \to T_x M$ indexed by $x \in M$. However, $\nabla_X Y(x)$ depends not only on $Y(x)$ but also on its first derivative and is so not a tensor (see later).

Proposition 3.5.2 (The canonical connection in local coordinates) .
a) In a local coordinate system, we have with $X = \sum_j X^j \partial_j$ and $Y = \sum_k Y^k \partial_k$,

$$\nabla_X Y = \sum_i \left(\sum_j X^j \frac{\partial Y^i}{\partial x^j} + \sum_{j,k} \Gamma^i_{jk} X^j Y^k \right) \frac{\partial}{\partial x^i}$$

b)

$$2\Gamma^i_{jk} = \sum_l g^{il} (\partial_j g_{kl} + \partial_k g_{lj} - \partial_l g_{jk}) \ .$$

Proof. a) Since ∇ is bilinear in X, Y, we can assume $X = X^j \delta_j, Y = Y^i \delta_i$. From the two properties in the definition of a connection, we get

$$\begin{aligned}
\nabla_X Y &= \nabla_{X^j \delta_j}(Y^i \delta_i) = X^j \nabla_{\delta_j}(Y^i \delta_i) \\
&= X^j \delta_j Y^i \delta_i + X^j Y^i \nabla_{\delta_j}(\delta_i) \\
&= X^j \delta_j Y^i \delta_i + X^j Y^i \sum_k \Gamma^k_{ij} \delta_k \\
&= X^j \frac{\partial Y^i}{\partial x^j} \delta_i + \sum_k \Gamma^k_{ij} X^j Y^i \delta_k \ .
\end{aligned}$$

By taking linear combinates $X = \sum_j X^j \delta_j, Y = \sum_i Y^i \partial_i$ (or switching on the Einstein convention), we get from this the formula we have to prove.
b) The boxed formula in the last proof together with the definition of the Christoffel symbols gives together with $[\partial_i, \partial_j] = 0$ for all i, j

$$2\Gamma_{jkl} = 2g(\nabla_{\partial_j} \partial_k, \partial_l) = \partial_j g_{kl} + \partial_k g_{lj} - \partial_l g_{jk} \ .$$

The formula for $\Gamma^i_{jk} = \sum_l g^{il} \Gamma_{jkl}$ follows from that. $\square$

Examples.
1) Let $M = \mathbf{R}^n$ be equiped with a flat metric $[g_{ij}] = \mathrm{Diag}(c_1, \dots, c_n)$. From the formula for Γ^k_{ij} in local coordinates, we see that $\Gamma^k_{ij} = 0$ because $\delta_j g_{kl} = \delta_k g_{lj} = \delta_l g_{jk} = 0$. This is also true, if g is constant.

2) Let $M = \{(x, y) \in \mathbf{R}^2 \mid y > 0\ \}$ be the upper half plane and $[g_{ij}] = \mathrm{Diag}(1/y^2, 1/y^2)$. This Riemannian manifold is the **Lobachevsky plane**.
We compute

$$\Gamma^2_{11} = 1/y, \Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -1/y \ .$$

All other coefficients are vanishing.

One of the main main reasons for defining connections is the notion of parallelism and the fundamental concepts of geodesics and curvature to which we come in the following sections.

3.6 Covariant derivative

In this section is explained, how a connection defines a a derivative on tensors which has the property that it maps tensors into tensors and replaces the partial derivative (which does not map tensors into tensors in general).

Definition. Let X be a vector field on a manifold M with connection ∇. The linear map ∇_X on $\Gamma(TM)$ defines a linear map on the space of tensor fields $\Gamma(T^0_1(M))$ (= differential 1-forms = covector fields) by

$$\nabla_X t(Y) = Xt(Y) - t(\nabla_X Y) \ .$$

Having defined the connection on 1-forms, we can define ∇_X on tensors $t \in T^p_q(M)$ by

$$\begin{aligned}(\nabla_X t)(Y^1, \dots, Y^p, X_1, \dots, X_q) &= Xt(Y^1, \dots, Y^p, X_1, \dots, X_q) \\ &- \sum_{i=1}^{p} t(Y^1, \dots, \nabla_X Y^i, \dots, Y^p, X_1, \dots, X_q) \\ &- \sum_{i=1}^{q} t(Y^1, \dots, \dots, \dots, Y^p, X_1, \dots, \nabla_X X_i, \dots, X_q) \ ,\end{aligned}$$

where $Y^1, \dots, Y^p$ are covector fields and $X_1, \dots, X_q$ are vector fields.
Remark.
If $t \in \Gamma(T^p_q(M))$ and $s \in \Gamma(T^{p'}_{q'}(M))$, then

$$\nabla_X(t \otimes s) = (\nabla_X t) \otimes s + (t \otimes \nabla_X s) \ .$$

(see the exercice in Homework 8).

Example.
If M is equiped with a Riemannian metric $g \in \Gamma(T^0_2(M))$ and ∇ is compatible with g, then $\nabla_X g = 0$ because this equation is exactly paraphrasing that ∇ is compatible with the metric:

$$\nabla_X g(X_1, X_2) = Xg(X_1, X_2) - g(\nabla_X X_1, X_2) - g(X_1, \nabla_X X_2) = 0 \ .$$

Definition. The **covariant derivative** is the linear map from $\Gamma(T^p_q M)$ to $\Gamma(T^p_{q+1} M)$ defined by

$$Dt(Y^1, \dots, Y^p, X_1, \dots, X_q, X) = \nabla_X t(Y^1, \dots, Y^p, X_1, \dots, X_q) \ .$$

Remark. A connection defines a differentiation on the space of all tensors similar as the exterior derivative d did it on skew-symmetric tensors. The naive partial differentiation would not lead to a tensor in general. Covariant differentiation is the right modification.

Notation. If the coordinates of t are $t^{j_1,\dots,j_p}_{i_1,\dots,i_q}$, we write

$$t^{j_1,\dots,j_p}_{i_1,\dots,i_q;i} = D_{\partial i} t^{j_1,\dots,j_p}_{i_1,\dots,i_q}$$

for the coordinates of $D_{\partial_i} t$.

Examples.

1) The covariant derivative of a function f is $f_{;i} = \frac{\partial f}{\partial x^i}$ and is so identical to the partial derivative.

2) We have seen in a Proposition of the last section, that if $Y = \sum Y^j \partial_j$ is a vector field, then

$$\boxed{Y^j_{;i} = \frac{\partial Y^j}{\partial x^i} + \sum_k \Gamma^j_{ki} Y^k \, .}$$

If $\Gamma^j_{ki} = 0$, then this is partial differentiation with respect to the i'th coordinate.

3) If $t = \sum_k t_k dx^k$ is in $\Gamma(T^0_1(M))$, then

$$\boxed{t_{k;i} = \frac{\partial t_k}{\partial x^i} - \sum_j \Gamma^j_{ki} t_j \, .}$$

Proof. By the definition at the beginning of this section,

$$\begin{aligned} t_{k;i} &= \nabla_i t(\partial_k) = \frac{\partial t_k}{\partial x^i} - t(\nabla_i \partial_k) \\ &= \frac{\partial t_k}{\partial x^i} - \sum_j \Gamma^j_{ki} t(\partial_j) \\ &= \frac{\partial t_k}{\partial x^i} - \sum_j \Gamma^j_{ki} t_j \, . \end{aligned}$$

4) If $t = \sum_{k,l} t_{kl} dx^k \otimes dx^l$ is in $\Gamma(T^0_2)(M)$, then

$$\boxed{t_{kl;i} = \frac{\partial t_{kl}}{\partial x^i} - \sum_j \Gamma^j_{ki} t_{jl} - \sum_j \Gamma^j_{il} t_{kj} \, .}$$

And it is clear, how this will be generalized to general tensors (see the exercice). For every upper index, there will be a correction term with positive sign and for every lower index a correction term with negative sign.

3.7 Parallel transport

Definition. A vector field V is **parallel transported** along a vector field X, if $\nabla_X V = 0$.

Definition. A **vector field along a curve** $c : I \to M$ is a smooth curve $X : I \to TM$ in the tangent bundle of M such that $X(x) \in T_x M$. For example, the **velocity field** $Dc(t)$ is such a vector field.

Definition. The covariant derivative along a curve is a map on the space of all vector fields along the curve c defined in a neighborhood U of $x = c(t)$ by

$$\nabla_c X = \nabla_Y \tilde{X} \; .$$

where $\tilde{X}$ is a vector field on U which is X when restricted to $c(I) \cap U$ and where Y is a vector field on U which is Dc when restricted to $c(I) \cap U$.

Notation. We write $X(t)$ for a vector field on the curve c instead of $X(c(t))$.

Remark. From the properties of the connection, we have

$$\nabla_c (fY)(t) = (\frac{d}{dt} f(c(t)))Y(t) + f(t)\nabla_c Y(t) \; .$$

Definition. A vector field X along a curve c is called **parallel**, if $\nabla_c X = 0$.

Proposition 3.7.1 (Parallel transport) *Given a curve $c : I \to M$ and given a vector $v \in T_x M$ where $x = c(t_0)$. There exists a unique parallel vector field X along c such that $X(t_0) = v$.*

Proof. By splitting the interval I in many little pieces, we can assume that $c(I)$ is contained in a single chart of M. In order to find the parallel vector field $X(c(t)) = v(t)$ on the curve, one has to solve the equation

$$\nabla_c X = 0 = \frac{d}{dt} v^k(t) + \sum_{i,j} \Gamma^k_{ij} v^i(t) Dc^j(t) \; .$$

This is a differential equation of the form $\dot{v} = F_t(v)$ in $\mathbf{R}^n$ with smooth function $F_t : \mathbf{R}^n \to \mathbf{R}^n$ and has a solution for small t. □

Definition. The parallel transport from $c(0)$ to $c(t)$ along a curve c in a Riemannian manifold (M, g) is defined as the map

$$v(0) \in T_{c(0)} M \mapsto v(t) \in T_{c(t)} M$$

by constructing a parallel vector field $v(t)$ on $c(t)$.

Proposition 3.7.2 *The parallel transport defines an isometry from $T_{c(0)}M$ to $T_{c(t)}(M)$. That is, if $v(0), w(0) \in T_{c(0)}M$ are given, then $g(c(t))(v(t), w(t)) = g(c(0))(v(0), w(0))$ and especially, angles between two vectors are preserved.*

Proof. For any two vector fields X, Y on the curve

$$\frac{d}{dt} g(X(t), Y(t)) = g(\nabla_c X(t), Y(t)) + g(X(t), \nabla_c Y(t)) \ .$$

The right hand side is vanishing if $X(t), Y(t)$ are parallel transported vector fields on c because then $\nabla_c X = 0$ and $\nabla_c Y = 0$. □

Remark. Parallel transport from one point x to an other point y in the manifold depends in general on the path from x to y.

3.8 Geodesics

Definition. A curve c is called a **geodesic**, if its velocity vector field Dc on c is parallel.

Remark. The length of $Dc(t)$ is constant because this is a property of parallel transport of any vector field.

Proposition 3.8.1 *Let $c(t)$ be a geodesic in the Riemannian manifold (M, g). In local coordinates, if $c(t) = (c^1(t), \dots, c^n(t))$, then*

$$\frac{d^2}{dt^2} c^i + \sum_{j,k} \Gamma^i_{jk}(c(t)) \dot{c}^j(t) \dot{c}^k(t) = 0 \ .$$

Proof. This is the formula of the parallel transport of the vector $v(t) = \dot{c}(t)$. □

Corollary 3.8.2 (Local existence of geodesics) *Given a point $x \in M$ and $v \in T_x M$, there exists an interval $I = [-\epsilon, \epsilon]$ and a geodesic $c : I \to M$, which satisfies $Dc(0) = v$ and $c(0) = x$.*

Proof. Also this result follows from an existence result of solutions of ordinary differential equations. The differential equation

$$\ddot{c}^i = -\sum_{j,k} \Gamma^i_{jk}(c(t)) \dot{c}^j(t) \dot{c}^k(t)$$

is of the form $d/dt(c, d) = d(F_t(c, d)$ which can be rewritten as $\frac{d}{dt} z(t) = G(z)$ using $z = (t, c, d)$. □

The explicit solutions of the differential equations giving the geodesics can in general not be found. Even on a sphere, the flow can be chaotic and come arbitrarily close to any point in the sphere. Nevertheless, there are examples like the Lobachevsky plane, the sphere or the plane, where the geodesics are known.

Examples.
1) On the flat plane (M, g), where $[g]_{ij} = \mathrm{Diag}(1, 1)$, the geodesic flow is easy to find since $\Gamma^i_{kl}((x, y)) = 0$ for all $x \in M$. The differential equations are therefore

$$\ddot{x} = 0 \ , \ddot{y} = 0 \ .$$

That is the geodesics are the straight lines.
2) On the Lobachevsky plane, with metric $[g]_{ij} = \mathrm{Diag}(1/y^2, 1/y^2)$, we have already computed the Christoffel symbols. This gives the differential equations

$$\begin{aligned} \ddot{x} &= -\frac{2\dot{x}\dot{y}}{y}, \\ \ddot{y} &= \frac{(\dot{x}^2 - \dot{y}^2)}{y}. \end{aligned}$$

One can show that the solutions of these differential equations are circles through (x, y) which have in (x, y) the tangent direction $(\dot{x}, \dot{y})$ and which hit the line $y = 0$ in a right angle.
3) On the sphere, you compute the Christoffel symbols in the homework. It is convenient to make such computations with the computer

```
g= {{          1    ,       0       },
   {          0    ,  Sin[theta]^2  }};
ginverse= {{       1 , 0                 },
          {        0 , Csc[theta]^2       }};
x={theta,phi};
GDown[j_,k_,i_]:=1/2(D[g[[k,i]],x[[j]]] + D[g[[i,j]],x[[k]]] - D[g[[j,k]],x[[i]]]);
GDown=Table[GDown[j,k,i],{j,2},{k,2},{i,2}]
G[j_,k_,i_]:=Sum[ginverse[[i,l]] GDown[[j,k,l]],{l,2}];
G=Table[G[j,k,i],{j,2},{k,2},{i,2}]
```

Definition. A Riemannian manifold, for which a geodesic can be defined for all times $t \in \mathbf{R}$ is called **geodesically complete**.

Remark. Not every manifold is complete. Take from any manifold a point away, then the new manifold is not complete because there exists a geodesic moving into the point which has been taken away.

Definition. Assume (M, g) is a geodesically complete manifold. Define the **exponential map** from TM to M by $(x, v) \mapsto c(1)$, where c is a geodesic satisfying $c(0) = x$ and $Dc(0) = v$.

Remark. The reason for this name is that for $SO(3)$, the tangent space at zero is a space $so(3)$ of skew symmetric matrices and that $\exp(0, v) = e^v = 1 + v + v^2/2! + v^3/3! + \ldots$ is the usual exponential map. (without proof).

Theorem 3.8.3 (Hopf-Rinow) *Let (M, g) be a connected Riemannian manifold (g is assumed to be positive definite). Then M is a complete metric space if and only if M is geodesically complete. Then, every two points $x, y \in M$ can be connected by geodesic. This path minimizes the length between all paths from x to y.*

The proof is quite lengthy and we skip it here.

3.9 Riemannian Curvature

Definition. The combination $R(X,Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}$ is called the **curvature** of the connection. It defines the **Riemann curvature tensor**

$$K(X,Y,Z,W) = g(R(X,Y)Z,W) .$$

Recall. If a function $t : \Gamma(T^*M)^p \times \Gamma(TM)^q \to \mathbf{R}$ is a tensor field, then the value of $t(Y^1,\dots,Y^q,X_1,\dots,X_p)(x)$ depends only on $Y^i(x)$ and $X_k(x)$.

Definition. A function $t : \Gamma(T^*M)^p \times \Gamma(TM)^q \to \mathbf{R}$ is $C^\infty(M)$- **multi-linear** if for all $f \in C^\infty(M)$,

$$\begin{aligned} t(Y^1,\dots,fY^i,\dots,Y^p,X_1,\dots,X_q) &= t(Y^1,\dots,Y^i,\dots,Y^p,X_1,\dots,fX_i,\dots,X_q) \qquad (3.1)\\ &= ft(Y^1,\dots,Y^p,X_1,\dots,X_q) . \end{aligned}$$

Proposition 3.9.1 (Convenient criterium for tensorality) *A map $t : \Gamma(T^*M)^p \times \Gamma(TM)^q \to \mathbf{R}$ is a tensor field if and only if it is $C^\infty(M)$-multilinear.*

Proof. If t is $C^\infty(M)$-multi linear, then $t(Y^1,\dots,Y^p,X_1,\dots,X_q)(x)$ depends only on $X_i(x), Y^i(x)$ because we can choose in formula 3.1 functions f with support in an arbitrary small neighborhood of a point $x \in M$. That is

$$t(Y^1,\dots,Y^p,X_1,\dots,X_q)(x) = t(x)(Y^1(x),\dots,Y^p(x),X_1(x),\dots,X_q(x)) .$$

If on the other hand $t(Y^1,\dots,Y^p,X_1,\dots,X_q)(x) = t(x)(Y^1(x),\dots,Y^p(x),X_1(x),\dots,X_q(x))$ is a tensor field, then 3.1 holds. $\square$

Examples.
a) The map $(X,Y,Z) \mapsto g(\nabla_X Y, Z)$ is not a tensor. It is C^∞-linear in X and Z but not in Y, because by a defining property of the connection

$$g(\nabla_X fY, Z) = g(f\nabla_X Y, Z) + (X.f)g(Y,Z)$$

for all vector fields X, W.
b) The map $X,Y \mapsto g(X,Y)$ is by definition a tensor. That is $g(fX,Y) = g(X,fY) = fg(X,Y)$.
c) The map $X,Y,Z \mapsto g([X,Y],Z)$ is not a tensor, because in general

$$\begin{aligned} g([X,fY],Z) &= g(L_X fY - L_{fY}X, Z) = X.fg(Y,Z) + fg(L_X Y, Z) - fg(L_Y X, Z)\\ &\neq fg(L_X Y,Z) - fg(L_Y X,Z) = fg([X,Y],Z) . \end{aligned}$$

As the name suggests:

Lemma 3.9.2 *The Rieman curvature tensor K is a tensor.*

Proof. We have to show that

$$fK(X,Y,Z,W) = K(fX,Y,Z,W) = K(X,fY,Z,W) = K(X,Y,fZ,W) = K(X,Y,Z,fW) .$$

a) The equality $fK(X,Y,Z,W) = K(X,Y,Z,fW)$ is true because $g(X,fW) = fg(X,W)$.
b)

$$\begin{aligned}
K(fX,Y,Z,W) &= g(\nabla_{fX}\nabla_Y Z,W) - g(\nabla_Y\nabla_{fX}Z,W) - g(\nabla_{[fX,Y]}Z,W)\\
&= g(f\nabla_X\nabla_Y Z,W) - g(\nabla_Y f\nabla_X Z,W) - g(\nabla_{f[X,Y]-(Y.f)X}Z,W)\\
&= g(f\nabla_X\nabla_Y Z,W) - g((Y.f)\nabla_X Z,W) - fg(\nabla_Y\nabla_X Z,W)\\
&\quad -f\nabla_{[X,Y]}Z,W) + (Y.f)g(\nabla_X Z,W)\\
&= f(\nabla_X\nabla_Y Z,W) - fg(\nabla_Y\nabla_X Z,W) - fg(\nabla_{[X,Y]}Z,W)\\
&= fK(X,Y,Z,W)\ .
\end{aligned}$$

c) The proof of $K(X,fY,Z,W) = fK(X,Y,Z,W)$ is the same as in b).
d) We have $K(X,Y,fZ,W) = fK(X,Y,Z,W) + X.Y.f - Y.X.f - [X,Y].f = fK(X,Y,Z,W)$ since $X.f = L_X f$ and $[X,Y]$ was defined as $L_X L_Y - L_Y L_X$. □

In order to express K in local coordinates, take $X_i = \delta/\delta x^i$ and define

$$R_{ijkl} = K(X_i, X_j, X_k, X_l)\ .$$

(It is custom to take here the letter R and not K).
In the following theorem, we need the identity

Lemma 3.9.3 (The Jacobi identity)

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]]\ .$$

Proof. By definition, we have

$$L_{[X,Y]} = L_X L_Y - L_Y L_X$$

so that

$$\begin{aligned}
L_{[X,[Y,Z]]} &= L_X L_{[Y,Z]} - L_{[Y,Z]} L_X\\
&= L_X L_Y L_Z - L_X L_Z L_Y - L_Y L_Z L_X + L_Z L_Y L_X\\
L_{[Y,[Z,X]]} &= L_Y L_{[Z,X]} - L_{[Z,X]} L_Y\\
&= L_Y L_Z L_X - L_Y L_X L_Z - L_Z L_X L_Y + L_X L_Z L_Y\\
L_{[Z,[X,Y]]} &= L_Z L_{[X,Y]} - L_{[X,Y]} L_Z\\
&= L_Z L_X L_Y - L_Z L_Y L_X - L_X L_Y L_Z + L_Y L_X L_Z\ .
\end{aligned}$$

Adding up the right hand sides gives zero. □

Definition. A vector space X with a bilinear multiplication $X \times X \to X, a,b \mapsto [a,b]$ is called a **Lie algebra** if $[a,b] = -[b,a]$ for all $a,b \in X$ and $[a,[b,c]] + [b,[c,a]] + [c,[a,b]] = 0$.

Examples.
1) We have just seen, that the space $\Gamma(TM)$ of all vector fields on a manifolds M is a Lie algebra if $[X,Y] = L_X L_Y - L_Y L_X$ is the commutator.
2) The vector space $M(n,\mathbf{R})$ of all $n \times n$ matrices is a Lie algebra with multiplication $[a,b] \mapsto ab - ba$.

Remark. Lie algebras are important in subjects like **particle physics**, the **theory of Lie groups** or in **dynamical systems**.

Theorem 3.9.4 (First Bianchi identity and other symmetries of K.)
a) $R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$ *(Bianchi identity)*

b) $K(X,Y,Z,W) = -K(Y,X,Z,W)$, $K(X,Y,W,Z) = -K(X,Y,Z,W)$.

c) $K(X,Y,Z,W) = K(Z,W,X,Y)$.

Proof. a) Adding the three lines

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

gives on the right hand side

$$\nabla_X(\nabla_Y Z - \nabla_Z Y) + \nabla_Y(\nabla_Z X - \nabla_X Z) + \nabla_Z(\nabla_X Y - \nabla_Y X) - \nabla_{[X,Y]}Z - \nabla_{[Y,Z]}X - \nabla_{[Z,X]}Y \,.$$

This is, because the connection is torsion free,

$$\nabla_X[Y,Z] + \nabla_Y[Z,X] + \nabla_Z[X,Y] - \nabla_{[X,Y]}Z - \nabla_{[Y,Z]}X - \nabla_{[Z,X]}Y$$

and again because the connection is torsion free

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] \,.$$

The Jacobi identity assures that this is vanishing.

b) The first identity is clear from the definition of R. From the compatibility of the connection with the metric, we get

$$\begin{aligned} g(\nabla_X\nabla_Y Z, W) &= Xg(\nabla_Y Z, W) - g(\nabla_Y Z, \nabla_X W) \\ &= XYg(Z,W) - Xg(Z,\nabla_Y W) \\ &- Yg(Z,\nabla_X W) + g(Z,\nabla_Y\nabla_X W)\,, \end{aligned}$$

$$\begin{aligned} g(\nabla_Y\nabla_X Z, W) &= Yg(\nabla_X Z, W) - g(\nabla_X Z, \nabla_Y W) \\ &= YXg(Z,W) - Yg(Z,\nabla_X W) \\ &- Xg(Z,\nabla_Y W) + g(Z,\nabla_X\nabla_Y W)\,, \end{aligned}$$

$$g(\nabla_{[X,Y]}Z, W) = (XY - YX)g(Z,W) - g(Z,\nabla_{[X,Y]}W)\,.$$

We see that $g(R(X,Y)Z,W) = -g(Z,R(X,Y)W)$ which establishes the second identity.

c) Take X_i so that $R_{ijkl} = K(X_i,X_j,X_k,X_l)$.

$$\begin{aligned} R_{ijkl} &= -R_{jkil} - R_{kijl} \ \ (a) \\ &= R_{jkli} + R_{kilj} \ \ (b) \\ &= -R_{klji} - R_{ljki} - R_{ilkj} - R_{lkij} \ \ (a) \\ &= R_{klij} + R_{ljik} + R_{iljk} + R_{klij} \ \ (b) \\ &= 2R_{klij} - R_{jilk} \ \ (a) \\ &= 2R_{klij} - R_{ijkl} \ \ (b) \end{aligned}$$

which gives $R_{ijkl} = R_{klij}$. ∎

Definition. If P is a two dimensional plane in T_xM and v, w are a basis in P, the expression

$$K(v,w) = \frac{K(x)(v,w,v,w)}{g(v^\flat \wedge w^\flat, v^\flat \wedge w^\flat)}$$

is called the **sectional curvature** at x in the plane P. (It can be interpreted as half the scalar curvature of the surface obtained by taking all geodesics starting at x in directions of the plane P. The value of $K(v,w)$ is independent of the basis. If at every point, the sectional curvature is constant, then the manifold is has **constant curvature**.

Remark. If v, w is an orthonormal basis in the plane $P \subset T_xM$, then

$$K(v,w) = K(v,w,v,w)/(g(v,v)g(w,w)) = \pm K(v,w,v,w) .$$

In the Riemannian case, when g is positive definite, then $g(v,v) = g(w,w) = 1$ if v, w are orthonormal in P so that then the sectional curvature is given from the Riemannian curvature by $K(v,w) = K(v,w,v,w$.

We will introduce in the next section the scalar curvature $S(x)$ at a point x, which is a number and related to the sectional curvature by $S(x) = \sum_{i \neq j} K(e_i, e_j)$, where $e_1, \dots, e_n$ is an orthonormal basis at x.

3.10 Ricci tensor and scalar curvature

Let us first compute the Riemann curvature tensor in cordinates using $X_i = \delta/\delta x^i$.

Define $R_{k,m,n,j} = R(X_k, X_m, X_n, X_j)$, so that

$$R(X_k, X_m)X_n = \sum_j R^j_{kmn} X_j = \sum_j R^j_{nkm} .$$

(The second expression is often used also and is the same like the middle expression because of the symmetry of R).

Lemma 3.10.1 (Riemann curvature in coordinates)

$$R^i_{kmn} = \partial_m \Gamma^i_{kn} - \partial_n \Gamma^i_{km} + \sum_j \Gamma^j_{kn}\Gamma^i_{jm} - \sum_j \Gamma^j_{km}\Gamma^i_{jn} .$$

Proof. This is a straightforward calculation using $\nabla_{X_m} X_k = \sum_j \Gamma^j_{mk} X_j$ and $[X_k, X_l] = 0$.
To make the proof more transparent, let Γ_m be the matrix with entries $[\Gamma_m]^j_k = \Gamma^j_{km}$. Like this, the connection can be written more comfortably as $\nabla_m = \partial_m + \Gamma_m$. For any m, n, let R_{mn} be the symmetric matrix with entries R^i_{kmn}. The above equation is now the matrix equality

$$R_{mn} = \partial_m \Gamma_n - \partial_n \Gamma_m + \Gamma_m \Gamma_n - \Gamma_m \Gamma_n .$$

But the left hand side was defined as

$$\begin{aligned} \nabla_m \nabla_n - \nabla_n \nabla_m &= [\nabla_m, \nabla_n] \\ &= [\partial_m + \Gamma_m, \partial_n + \Gamma_n] \\ &= \partial_m \Gamma_n - \partial_n \Gamma_n + \Gamma_n \Gamma_m - \Gamma_m \Gamma_n . \end{aligned}$$

(Note that in a product of two matrices, the matrix, with which one multiplies first is standing to the right). $\square$

Definition. Let (M, g) be a Riemannian manifold. The **Ricci tensor** in $T_2^0(M)$ is in local coordinates defined as

$$\text{Ricci}_{ik} = \sum_j R^j_{ijk} \ .$$

Note. There are different signs around for the sign of the curvature and the Ricci tensor! For the scalar curvature, there is no disagreement. The sphere should have scalar curvature $+1$.

Definition. If $A : TM \to TM$ is a linear map, then the trace of A is defined as $\text{trace}(A(x)) = \sum_i A^i_i(x)$, where $Ae_i = \sum_j A^j_i e_j$ for a basis e_i in T_xM. This trace is independent of the basis because the trace of a matrix is basis independent (it is a coefficient of the characteristic polynomial which has this property).

Remark. The Ricci tensor $Ricci(X, Y)(x)$ can in a coordinate-free way be defined as the trace of the map $v \mapsto R(X(x), v)Y(x)$, where $R = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}$, $\Gamma(TM) \to \Gamma(TM)$ was defined at the beginning of the last section.

> **Proposition 3.10.2** *The Ricci tensor is a symmetric tensor field in* $\Gamma(T_2^0(M))$.

Proof. The trace of a tensor field is again a tensor field. The symmetry follows from the symmetry of the Riemannian tensor: use Bianchi's first identity and the anti-symmetry of $R_{ijkl} = -R_{jikl}$ to obtain

$$\begin{aligned} \text{Ricci}_{ik} &= \sum_j R^j_{ijk} \\ &= -\sum_j R^j_{jki} - \sum_j R^j_{kij} \\ &= 0 - \sum_j R^j_{kij} \\ &= \sum_j R^j_{kji} = \text{Ricci}_{ki} \ . \end{aligned}$$

□

Definition.
The **scalar curvature** $\text{Scal}(x)$ of M at a point x is defined as the trace of the Ricci tensor:

$$\text{Scal}(x) = \sum_i \text{Ricci}^i_i(x) = \sum_{i,j} g^{ij} Ricci_{ij}; .$$

Note. In some textbooks, the definition of Ricci is multiplied with a factor $1/(n-1)$ and the definition of Scal is multiplied with a factor $1/n(n-1)$.

Reminder. Given two vectors v, w in the tangent space T_xM, the sectional curvature $K(x)(v, w)$ was defined as

$$K(x)(v, w) = \frac{R(v, w, v, w)}{g(v^\flat \wedge w^\flat, g(v^\flat \wedge w^\flat)} \ ,$$

where g denotes also the metric lifted to $\Lambda^2(M)$. If $e_1, \dots, e_n$ is an orthonormal basis in T_xM, then this simplifies to

$$K(x)(e_i, e_j) = R(e_i, e_j, e_i, e_j)/(c_i c_j)$$

where $c_i = g_{ii} \in \{-1, 1\}$.

Remarks.
1) If the metric is positive definite (Riemannian case) and $\{e_i\}$ is an orthonormal basis at $T_x M$, then
$$\mathrm{Scal}(x) = \sum_{i,j} R(e_i, e_j, e_i, e_j) ,$$
because $\mathrm{Scal}(x) = \sum_{i,j} g^{ii} g^{jj} R(e_i, e_j, e_i, e_j)$ and $g^{ii} = 1$ for a positive definite metric.

2) If the manifold (M, g) has constant sectional curvature K, positive definite g and dimension n, we have
$$\mathrm{Ricci} = (n-1)Kg \, , \mathrm{Scal} = n(n-1)K \, .$$
Proof. By definition of the sectional curvature, we have using an orthonormal basis e_i at $T_x M$
$$K c_i c_j = R(e_i, e_j, e_i, e_j) \, .$$
Summing over i gives, since $R(e_j, e_j, e_j, e_j) = 0$
$$K(\sum_{i \neq j} c_i) g_{jj} = \mathrm{Ricci}_{jj} \, .$$
The left hand side is $K(n-1)g_{jj}$ if the metric g is positive definite. Taking again the trace gives
$$K(n-1) \sum_j g_{jj} g^{jj} = \sum_j \mathrm{Ricci}_j^j = \mathrm{Scal} \, .$$
That is
$$K(n-1)n = \mathrm{Scal} \, .$$

3) If g is multiplied with a factor $\lambda \in \mathbf{R}$, that is if $g' = \lambda g$, then the Rieman curvature tensors of g and g' are related by
$$R' = \lambda \cdot R$$
and so
$$\begin{aligned} \mathrm{Ricci}' &= \mathrm{Ricci} \\ \mathrm{Scal}' &= \lambda^{-1} \mathrm{Scal} \, . \end{aligned}$$

Example. Consider the sphere again. Using the notation at the beginning of this section, we compute
$$\begin{aligned} \Gamma_1 &= \begin{pmatrix} 0 & 0 \\ 0 & \cot(\theta) \end{pmatrix} \\ \Gamma_2 &= \begin{pmatrix} 0 & \cot(\theta) \\ -\cos(\theta)\sin(\theta) & 0 \end{pmatrix} \, . \end{aligned}$$
Since $R_{11} = R_{12} = 0$, we have only to determine the matrix $R_{12} = -R_{21}$ and it is
$$R_{12} = \begin{pmatrix} 0 & -\cot(\theta)^2 \\ \cos(\theta)^{-2} & 0 \end{pmatrix} \, .$$
We get
$$[\mathrm{Ricci}_{ij}] = \begin{pmatrix} 1 & 0 \\ 0 & \sin(\theta)^2 \end{pmatrix} = [g_{ij}]$$
and so
$$[\mathrm{Ricci}_i^j] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
and $R = \sum_i \mathrm{Ricci}_i^i = 2$.

3.11 The second Bianchi identity

There exists an other identity for the curvature, which is very useful for example in general relativity:

Proposition 3.11.1 (Second Bianchi identity)

$$\nabla_X R(Y,Z,V,W) + \nabla_Y R(Z,X,V,W) + \nabla_Z R(X,Y,V,W) = 0 \,.$$

Proof. We prove the identity in operator form, that is

$$\nabla_X R(Y,Z) + \nabla_Y R(Z,X) + \nabla_Z R(X,Y) = 0 \,.$$

From the definition of the covariant derivative, we have

$$\nabla_X R(Y,Z) = [\nabla_X, R(Y,Z)] - R(\nabla_X Y, Z) - R(Y, \nabla_X Z) \,.$$

Plug in the definition of $R(Y,Z) = [\nabla_Y, \nabla_Z] - \nabla_{[Y,Z]}$, to get

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

(In the last step, we used $[\nabla_X, \nabla_{[Y,Z]}] = R(X,[Y,Z]) - \nabla_{[X,[Y,Z]]} = 0$.) When summing up three such terms obtained by cyclic permutation of X, Y, Z, the terms involving ∇ disappear because of the Jacobi identity. Using that ∇ is torsion free: $\nabla_X Y - \nabla_Y X = [X,Y]$, we see that also the terms involving R disappear. □

Remark. In coordinates, the second Bianchi identity is

$$R^l_{ijk;m} + R^l_{jmk;i} + R^l_{mik;j} = 0 \,.$$

If we write the Riemann tensor in matrix form $[R_{ij}]^l_k = R^l_{ijk}$, then this can be rewritten as

$$R_{ij;m} + R_{jm;i} + R_{mi;j} = 0 \,.$$

In order to give a more intuitive notion of the Bianchi identity, we introduce now normal coordinates.

Appendix. Natural coordinates near a point $x \in M$ can often simplify the computations and proofs. These coordinates are called **normal coordinates** or **geodesic coordinates** or **Riemannian coordinates**.

Proposition 3.11.2 (Existence of normal coordinates) *Let (M,g) be a pseudo Riemannian manifold and $x \in M$. There exists a coordinate system near x such that $g_{ij}(x) = \mathrm{Diag}(c_1, \dots, c_n)$ with $c_i \in \{-1, 1\}$ and $\Gamma^k_{ij}(x) = 0$.*

Proof. Take a coordinate system, such that $x = 0$. Define in a neighborhood of x the map

$$\phi(x) = x + \frac{1}{2} \nabla_x x \,,$$

or in coordinates

$$\phi(x)^i = x^i + \sum_{j,k} \frac{1}{2} \Gamma^i_{jk}(0) x^j x^k \,.$$

We compute

$$(D\phi(x)u)^i = u^i + \sum_{j,k} \Gamma^i_{jk}(0)x^k u^j$$

and especially $D\phi(0)u = u$. Furthermore,

$$(D^2\phi(x)(u,v))^i = \sum_{j,k} \Gamma^i_{jk}(0)u^k u^j \ ,$$

so that $(D^2\phi(0))^l_{jk} = \Gamma^l_{jk}(0)$. Because in the coordinates $\phi(x)$, the Christoffel symbols are

$$\Gamma^{l'}_{j'k'} = \sum_{j,k,l} \Gamma^l_{jk} D\phi^{k'}_k D\phi^j_{j'} D\phi^l_{l'} \ - \sum_{j,k} D^2\phi^{l'}_{jk} D\phi^j_{j'} D\phi^k_{k'}$$

and $D\phi^j_i = \delta^j_i$ at the point $x = 0$, we have

$$\Gamma^{l'}_{j'k'} = 0 \ .$$

(The formula for the transformation of the Christoffel symbols follows from the axioms of the connection in coordinates $\nabla_{X_i} X_j = \sum_k \Gamma^k_{ij} X_k$:

$$\nabla_{\sum_i A^i_{i'} X_i} \sum_j A^j_{j'} X_j = \sum_{k'} \Gamma^{k'}_{i'j'} \sum_k A^k_{k'} X_k \ .$$

The left hand side is

$$\sum_i A^i_{i'} \sum_{j,k} A^j_{j'} \Gamma^k_{ij} X_k + \sum_{i,j,k} A^i_{i'} A^j_{j'i} X_j$$

where $A^j_{j'i} = \partial_i A^j_{j'}$, so that

$$\Gamma^{k'}_{i'j'} = \sum_{i,j,k} A^i_{i'} A^j_{j'} A^{k'}_k \Gamma^k_{ij} + \sum_k A^k_{j'i'} A^{k'}_k \ .$$

Using

$$0 = \partial_{i'}(\sum_k A^k_{j'} A^{k'}_k) = \sum_k (A^k_{j'i'} A^{k'}_k + A^k_{j'} A^{k'}_{ki'}) \ ,$$

we have

$$\sum_k A^k_{j'i'} A^{k'}_k = -\sum_k A^k_{j'} A^{k'}_{ki'} = -\sum_{ij} A^{k'}_{ij} A^{i'}_i A^{j'}_j \ .$$

This gives the identity

$$\Gamma^{k'}_{i'j'} = \sum_{i,j,k} A^i_{i'} A^j_{j'} A^{k'}_k \Gamma^k_{ij} - \sum_{ij} A^{k'}_{ij} A^{i'}_i A^{j'}_j$$

we have used above.) $\square$

Remark. One of the main advantages of these normal coordinates is, that at a given point, the covariant derivative becomes the usual partial derivative.

We give now a more intuitive view ("proof") of the second Biancchi identity which also shows, why physisics consider general relativity as a gauge field theory: consider a little cube in the manifold M. The cube is attached at $x \in M$ and has the sides $\delta x^k, \delta x^l$ and δx^m located at some point $x \in M$. Given a vector $U^j \in T_x M$, a parallel transport around the $\delta x^k - \delta x^l$ plaquette attached to x gives a change of the vector U by an amount

$$\delta U^i = R^i_{klj} A^j \delta x^k \delta x^l \ .$$

The sum of the parallel transports around this plaquette $P_{kl}(x)$ and the parallel plaquette $P_{kl}(x+\delta x^m)$ is

$$\delta U^i = \frac{\partial R^i_{klj}}{\partial x^m} A^j \delta x^k \delta x^l \delta x^m \; .$$

If we sum up all parallel transports around the cube, then each parallel transport around an edge has been added twice with opposit sign. That is

$$\frac{\partial R^i_{klj}}{\partial x^m} + \frac{\partial R^i_{lmj}}{\partial x^k} + \frac{\partial R^i_{mkj}}{\partial x^l} = 0 \; .$$

Since we have used normal coordinates, this is the Bianchi identity.

3.12 General relativity

History. Einstein and the mathematician Grossmann realized between 1909 and 1913 that the metric of space-time seems to depend on the amount of matter in the region in question (the letter g for the metric stands for gravity).

Definition. For defining the **energy momentum tensor** of matter, one consideres matter as a fluid with velocity vector field $u = dx/d\tau \in \Gamma(TM)$, (where τ is the proper time) and pressure field $p \in C^\infty(M)$ and density $\rho \in C^\infty(M)$ is defined as

$$T^{ij} = (\rho + p/c^2) u^i u^j - p g^{ij} \; ,$$

where c is the speed of light. One requires from physical conservation laws (momentum, mass and energy) that

$$T^{kj}_{;k} = 0 \; .$$

Example. If the mass distribution with density ρ and pressure p is not moving, then

$$[T^{ij}] = -\mathrm{Diag}(c^2\rho, p, p, p) \, .$$

History. During 1914-1915, Einstein made several attempts to find the correct relationship between the metric tensor and matter. The guess

$$g^{ij} = \kappa T^{ij}$$

does not reduce to the Newtonian limit and is so not good. Einstein published a guess

$$R^{ij} = \kappa T^{ij}$$

which was discarted because $\sum_k R^{ik}_{;k} \neq 0$. In the same year, he modified the equation to

$$\boxed{G^{ij} = R^{ij} - \frac{1}{2} R g^{ij} = \kappa T^{ij} \; .}$$

Definition. This equation is now called **Einstein's field equation**. The left hand side is a symmetric tensor called the **Einstein tensor**.

The fact that $G^{ij}_{;j} = 0$ follows from the second Bianchi identity and the symmetries of the curvature tensor: contract

$$R^l_{ijk;m} + R^l_{jmk;i} + R^l_{mik;j} = 0$$

over j and l to get

$$\mathrm{Ricci}_{ik;m} - \mathrm{Ricci}_{im;k} + \sum_j R^j_{mik;j} = 0$$

Now raise i and contract with k to get

$$\mathrm{Scal}_{;m} - \sum_k \mathrm{Ricci}^k_{m;k} - \sum_j \mathrm{Ricci}^j_{m;j} = 0$$

which gives

$$\mathrm{Scal}_{;m} - 2\sum_k \mathrm{Ricci}^k_{m;k} = 0 \; .$$

A manifold, for which the Riccy tensor is a multiple of the Riemann tensor is called a **Einstein manifold**. A special case is Ricci = 0, which correspond to vaccuum solutions of the Einstein equations. In homework 8, you have verified that the Schwarzschild manifold is an example of such an Einstein manifold.

Remark. The Einstein field equations can be seen as the equations saying that g is a critical point of the Hilbert action functional $g \mapsto \int_M R(g)\; d\mu(g)$ where μ is the natural volume form associated to g and $R(g)$ is the scalar curvature. Hilbert showed (also in 1915) that the condition $\partial/\partial g \phi = 0$ leads to the Einstein equations.

A final remark. General relativity is for the mathematician a special case of Riemannian geometry, where many interesting problems need further research:

• **Solving the Einstein equations**. This problem is a problem in partial differential equations. Given the energy momentum tensor of matter T on a manifold M the problem is to find a metric g, which satisfies the Einstein equations. As the above remark tells, this is also a problem in the calculus of variations, since one has to find a critical point of a functional. There are simple situations like a rotating binary system, where one still has not yet found the solutions.

• **Conceptional problems in general relativity**. General relativity deals either with the problem of finding the metric g from the energy momentum tensor T or with the problem of finding geodesic curves in a given Lorentzian manifold (M, g). It is most of the time a fair simplification to disregard the influence the gravitation of the test particle (which is moving on the geodesics). This is similar to other classical theories:
- in **classical mechanics**, where one treats the evolution of a multiparticle system like the solar system approximatively and through perturbation approaches (even if one knows today mathematically that this is **nonsense** since the Taylor series one writes down are **proven** to be not convergent in general). Even the mathematics of the classical Newtonian three body problem is too complex for todays mathematics.
- in **quantum mechanics**, quantum mechanical multiparticle systems are also treated perturbatively (even the Helium atom can not yet been solved completely without simplifications). Quantum mechanical many body problem lead even under simplifications to very difficult problems.
- in **electro magnetism**, one either deals with a given current j and solves the Maxwell equations $dF = 0, d^*F = j$, to get the fields F or one deals with a given field F and finds the solution of a charged testparticle. In the second problem, one disregards the field given by the test particle.

Coming back to general relativity: the problem of two massive bodies "two body problem" **is not even an issue of the theory** (one nevertheless treats the problem perturbatively). Assume, the metric is given, then each of the two bodies should move on geodesics of (M, g). Their motion defines an energy momentum tensor T in space-time M which then defines the metric g through the Einstein equations. But where do we begin? Since we do not know the metric, we do not know the motion of the two bodies and can not determine the energy momentum tensor. A physisist who deals with this problem either fixes the orbits of the bodies in a Newtonian approximation (maybe with corrections due to gravitational radiation) and tries then to solve the Einstein equations $G(g) = T$, where T is now given. Even this problem is unsolved (see previous problem). The full problem can not even be asked in the theory!

• **Finding special solutions of the Einstein equations with interesting topology**. Popular are **worm holes** (Einstein-Rosen 1935, Wheeler 1955, Morris-Thorne 1988) compact regions G in a Lorentzian manifold (M, g), which have a topological simple boundary (that is δG is diffeomorphic to a sphere S^3) but for which the interior of G has a nontrivial topology (that is $\mathrm{int}(G)$ is not simply connected). Such solutions can have intriguing properties like the existence of closed timelike curves "time machines". It was once under discussion (Einstein-Rosen) that elementary particles are just special solutions of the Einstein equations (little black holes). A good try since one knows "no hair theorem" that a black hole is determined by its mass, charge and angular momentum. This idea is nevertheless discarted because elementary particles like the neutrino behave as if they were points. The recent new interest in worm hole (Morris-Thorne 1988) came because it became reasonable to have worm holes, which are "traversable", in the sense that there are no event horizons. The approach was to take an interesting geometry (M, g), compute the Einstein tensor G and look, whether it is reasonable to have a energy momentum tensor T which satisfies the $G = T$.

• **Geodesic flow**. The geodesic flow on a given Riemannian manifold is a Hamiltonian system which is interesting from the dynamical system point of view. It is known for a long time already, that such a dynamical system is "chaotic" (a sloppy expression for dozens of mathematical properties) if the curvature $R(x)$ is strictly negative for all points in the manifold. It is only in the last 10 years Mathematicians have found that there are metrics on any manifold (like the sphere), for which the flow is ergodic.

• **Physics in the Planck scale**. The laws of general relativity break down in the **Planck scale** similar as the laws of classical mechanics have in the small to be replaced by classical quantum mechanics or field theories are replaced by quantum field theories. The problem is that the experimental investigation of this region is not yet possible (if the current progress in accelerator physics is extrapolated, then one still had to wait until the year 2200, until one reaches this domain). There is therefore plenty of time for mathematics to develop and test new ideas. The impossibility to remove ideas through experiments has lead to too many approaches. Even if one sticks to one approach of quantisation, like functional integrals, there are still to many possibilities and mathematical problems in trying to "quantize gravity". It seems that one is not only lacking of ideas today but also of **lacking of tools to remove ideas**!

Chapter 4

Exercices with solutions

Week 1

Ma109a, O. Knill, **September 1995**

Week 1

Due: Friday, October 6, 1995

Topics: Manifolds, Diffeomorphisms

Only the problems with a star (*) have to be done. If you have time, look also at the rest of the problems.

1) (*) Topic: Manifolds. Show that the n dimensional torus $\mathbf{T}^n = S^1 \times \ldots \times S^1$ is a manifold. Construct an explicit atlas. Hint: Use that $\mathbf{T}^n$ is a product manifold.

2) (*) Topic: Manifolds. Show that the k-dimensional sphere can not be made into a manifold where the differential structure is generated by an atlas with a single chart. Hint: S^k is compact.

3) Topic: Manifolds. Show that the set $\{(x, y) \in \mathbf{R}^2 \mid xy = a\}$ is a manifold if and only if $a \neq 0$. Construct an explicit atlas in the case $a \neq 0$.

4) (*) Topic: Manifolds. A real $2n \times 2n$ matrix A is called symplectic, if $A^T J A = J$, where $J = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}$. (The $n \times n$ matrix $\mathbf{1}$ is the identity matrix and the $n \times n$ matrix $\mathbf{0}$ is the matrix with all zeros.) Show that the set of symplectic 2×2 matrices is coinciding with $SL(2, \mathbf{R})$ and is a manifold.

5) Topic Manifolds. The Möbius strip is the open set $M = \{(\phi, t) \mid \phi \in [0, 2\pi], t \in (-1, 1)\}$ in the plane, where the identification $(0, t) \sim (2\pi, -t), \forall t \in (-1, 1)$ is done on eht boundary. Give an atlas, which makes M to a manifold.

6) (*) Topic: Diffeomorphisms. Given the map $f : \mathbf{R}^3 \to \mathbf{R}^2$ given by $f(x_1, x_2, x_3) = (\sin(x_1) x_2^3 x_3, x_3^2 x_1^2)$. Compute $D^k f(x)$ for $k = 1, 2, 3$ and state explicitely, from where to where the map $D^k f(x)$ goes. Write the Taylor expansion of f at 0 up to order 3.

7) Topic: Diffeomorphisms. Show that the manifold $\mathbf{R}^k$ is diffeomomorphic to the manifold $S^k \setminus \{P\}$, where P is a point on the sphere S^k. Give an explicit formula for the diffeomorphism in the case $k = 1$.

8) (*) Topic: Diffeomorphisms. Show that a smooth bijective map needs not to be a diffeomorphism. Hint: Give an example $f : \mathbf{R} \to \mathbf{R}$.

9) Topic: Diffeomorphisms. Let M, F be manifolds. Show that the projection $\pi : M \times F \to M, \pi(x, y) = x$ is a smooth map.

10) Topic: Diffeomorphisms. Let M_1, M_2, N_1, M_2 be manifolds. Show that if M_1 is diffeomorphic to N_1 and M_2 is diffeomorphic to N_2, then $M_1 \times M_2$ is diffeomorphic to $N_1 \times N_2$.

Ma109a, O. Knill, **October 1995**

Week 1 (Solutions)

Topics: Manifolds, Diffeomorphisms

1) (*) Topic: Manifolds. Each of the one dimensional spheres S^1 is a manifold with the atlas $(U_1, \phi_1), (U_2, \phi_2)$, where the ϕ_1, ϕ_2 are the stereographic projections from the north and south pole. The torus $\mathbf{T}^n$ is the set $\mathbf{T}^n = \{x = (x_1, x_2, \dots, x_n) \mid x_i \in S^1\}$. We construct an atlas on $\mathbf{T}^n$ by taking for each finite sequence $I = (i_1, \dots, i_n)$ the open set $U_I = U_{i_1}^{(1)} \times U_{i_2}^{(2)} \times \dots \times U_{i_n}^{(n)}$ and the map $\phi_I(x) = (\phi_{i_1}^{(1)}(x_1), \phi_{i_2}^{(2)}(x_2), \dots, \phi_{i_n}^{(n)}(x_n))$ from $U_I \subset \mathbf{T}^n$ to $\mathbf{R}^n$.

2) (*) Topic: Manifolds. Assume S^n has one chart (U, ϕ). Then $U = S^n$ is compact. Because ϕ is continuous, $\phi(U) \subset \mathbf{R}^n$ is compact. Because ϕ^{-1} is continuous from $\phi(U)$ to U, the set $\phi(U)$ is also open. The only open and closed set in $\mathbf{R}^n$ is the emptyset. This is not possible: because S^n is not empty, it is not possible that $\phi : S^n \to \emptyset$ is a bijection.

3) (*) Topic: Manifolds. The condition $A^T J A = J$ is equivalent with $ac - bd = \det(A) = 1$ as Mathematica

```
A={{a,b},{c,d}}; J={{0,1},{-1,0}}; Transpose[A].J.A
{{0, -(b c) + a d}, {b c - a d, 0}}
```

shows. The set of symplectic 2×2 matrices is therefore the set of $SL(2, \mathbf{R})$ of 2×2 matrices with determinant 1. Consider the manifold $M(2, \mathbf{R})$ of all real 2×2 matrices. Since it is a finite dimensional vector space, it is trivially a manifold. Consider the map $f : M \to \mathbf{R}, f(A) = \det(A), f(a, b, c, d) = ac - bd$. The map f has the derivative $Df(A)u = (c, -d, a, -b) \cdot u$ which is a linear map from $\mathbf{R}^4 \sim M(2, \mathbf{R})$ to $\mathbf{R}$. The rank of this map is maximal 1 if and only if $Df(x)$ is not vanishing or equivalently, if A is not vanishing. Since $f^{-1}(1)$ does not contain the zero matrix, 1 is a regular point. The lemma in the course shows that $SL(2, \mathbf{R}) = f^{-1}(1)$ is a 3 dimensional manifold.

4) (*) Topic: Diffeomorphisms. Write $f(x) = f(x_1, x_2, x_3) = (\sin(x_1)x_2^3 x_3, x_3^2 x_1^2) = (f_1(x), f_2(x))$.
a) The first derivative is a linear map from $\mathbf{R}^3$ to $\mathbf{R}^2$. We write D_i for $\partial/\partial x_i$.

$$\begin{aligned} Df(x)u &= (D_1 f_1(x)u_1 + D_2 f_1 u_2 + D_3 f_1 u_3, D_1 f_2(x)u_1 + D_2 f_2 u_2 + D_3 f_2 u_3) \\ &= (\cos(x_1)x_2^3 x_3 u_1 + \sin(x_1)3x_2^2 x_3 u_2 + \sin(x_1)x_2^3 u_3, 2x_1 x_3^2 u_1 + 2x_3 x_1^2 u_3) . \end{aligned}$$

Also here (and later) one could get assisted by Mathematica:

```
g={Sin[x1] x2^3 x3,x3^2 x1^2};
Diff[f_,{n1_,n2_,n3_}]:=D[D[D[f,{x1,n1}],{x2,n2}],{x3,n3}];
Dg=Diff[g,{1,0,0}] u1 + Diff[g,{0,1,0}] u2 + Diff[g,{0,0,1}] u3;
```

b) The second derivative is a bilinear map from $\mathbf{R}^3 \times \mathbf{R}^3$ to $\mathbf{R}^2$. We can write $D^2f(x) = (D^2f_1(x), D^2f_2(x))$.

$$\begin{aligned} D^2f_1(x)(u^{(1)},u^{(2)}) &= -\sin(x_1)x_2^3x_3u_1^{(1)}u_1^{(2)} + \sin(x_1)6x_2x_3u_2^{(1)}u_2^{(2)} \\ &+ \cos(x_1)3x_2^2x_3(u_1^{(1)}u_2^{(2)} + u_2^{(1)}u_1^{(2)}) \\ &+ \cos(x_1)x_2^3(u_1^{(1)}u_3^{(2)} + u_3^{(1)}u_1^{(2)}) \\ &+ \sin(x_1)3x_2^2(u_2^{(1)}u_3^{(2)} + u_3^{(1)}u_2^{(2)}) . \end{aligned}$$

$$D^2f_2(x)(u^{(1)},u^{(2)}) = 2x_3^2u_1^{(1)}u_1^{(2)} + 2x_1^2u_3^{(1)}u_3^{(2)} + 4x_1x_3(u_1^{(1)}u_3^{(2)} + u_3^{(1)}u_1^{(2)}) .$$

c) The third derivative $D^3(x)$ is a multilinear map from $\mathbf{R}^3 \times \mathbf{R}^3 \times \mathbf{R}^3$ to $\mathbf{R}^2$. Again, we decompose $D^3f(x) = (D^3f_1(x), D^3f_2(x))$ and since I have to type everything in TEX, I give only the first part:

$$\begin{aligned} D^3f_1(x)(u^{(1)},u^{(2)},u^{(3)}) &= -\cos(x_1)x_2^2x_3u_1^{(1)}u_1^{(2)}u_1^{(3)} + \sin(x_1)6x_3u_2^{(1)}u_2^{(2)}u_2^{(2)} \\ &+ -\sin(x_1)3x_2^2x_3(u_1^{(1)}u_2^{(2)}u_1^{(3)} + u_2^{(1)}u_1^{(2)}u_1^{(3)} + u_1^{(1)}u_1^{(2)}u_2^{(3)}) \\ &+ \cos(x_1)6x_2x_3(u_1^{(1)}u_2^{(2)}u_2^{(3)} + u_2^{(1)}u_1^{(2)}u_2^{(3)} + u_2^{(1)}u_2^{(2)}u_1^{(3)}) \\ &- \sin(x_1)x_2^3(u_1^{(1)}u_1^{(2)}u_3^{(3)} + u_1^{(1)}u_3^{(2)}u_1^{(3)} + u_3^{(1)}u_1^{(2)}u_1^{(3)}) \\ &+ \sin(x_1)6x_2(u_2^{(1)}u_2^{(2)}u_3^{(3)} + u_2^{(1)}u_3^{(2)}u_2^{(3)} + u_3^{(1)}u_2^{(2)}u_2^{(3)}) \\ &+ \cos(x_1)3x_2^2(u_1^{(1)}u_2^{(2)}u_3^{(3)} + u_1^{(1)}u_3^{(2)}u_2^{(3)} + \cdots + u_3^{(1)}u_2^{(2)}u_1^{(3)}) . \end{aligned}$$

For the Taylor expansion at $x = 0$, we get with $Df(0)(u) = (0,0)$, $D^2f(0)(u^{(1)},u^{(2)}) = (0,0)$, $D^3f(0)(u^{(1)},u^{(2)},u^{(3)}) = (0,0)$ the expansion

$$\begin{aligned} f(0+y) &= 0 + Df(0)y + \frac{D^2f(0)}{2!}(y,y) + \frac{D^3(f(0))}{3!}(y,y,y) + R(0,y)(y,y,y,y) \\ &= R(0,y)(y,y,y,y) , \end{aligned}$$

where $||R(0,y)(y,y,y,y)|| \leq C||y||^4$ for some constant C.

5) (*) Topic: Diffeomorphisms. The map $x \mapsto x^3$ is bijective from $\mathbf{R}$ to $\mathbf{R}$. However, the inverse $g(x) = x \mapsto \operatorname{sign}(x)|x|^{1/3}$ is not differentiable at zero, because $\lim_{x\to 0}|Dg(x)| \to \infty$.

Week 2

Ma109a, O. Knill, **October 1995**

Week 2

Due: Friday, October 13, 1995

Topics: Compactnesss, Embedding and immersion, Theorems of Sard and Whitney

1) (*) **Topic: Critical points, regular points, critical values, regular values.**
What is the set of critical points, critical values, regular points and regular values of the following maps:
a) $f : \mathbf{R}^2 \to \mathbf{R}^2$, $f(x,y) = (x^2, 2xy)$,
b) $f : \mathbf{R}^2 \to \mathbf{R}$, $f(x,y) = (x^2 + y^2)^2$,
c) $f : \mathbf{R}^4 \to \mathbf{R}, f(x) = 1$,
d) $f : \mathbf{R}^3 \to \mathbf{R}^3, f(x) = x$.

2) **(*) Topic: Immersions, Embeddings.**
a) Give an example which shows that an injective immersion does not need to be an embedding.
b) Prove that an injective immersion of a compact manifold is an embedding.
c) Give an example of an immersion of the circle in the plane, which is not an embedding.

3) **Topic: Compactness.**
A **theorem of Tychonov** states that a product of compact sets is compact. Prove the following special case of this theorem. If $(X_1, d_1), \dots, (X_n, d_n)$ are compact metric spaces, then the product space $(X_1 \times \dots \times X_n, d)$ with the metric $d(x,y) = \sum_i d(x_i, y_i)$ is compact too.
Hint. A metric space (X, d) is compact, if and only if every sequence x_n in X has an accumulation point x: for every $\epsilon > 0$, there exists n, such that $d(x_n, x) < \epsilon$.

4) **(*) Topic: Compactness.**
We have shown in class that a manifold is paracompact, a property which allowed us to construct a partition of unity. However, not all manifolds are compact. Which of the following manifolds are compact and why:
a) $M = \mathbf{R}^n$, $n > 0$,
b) $M = S^n$, the n-dimensional sphere.
c) $M = SL(2, \mathbf{R}) = \{A \mid 2 \times 2 \text{ matrices with determinant } 1\}$
d) $M = SO(3, \mathbf{R}) = \{A \mid \text{ orthogonal } 3 \times 3 \text{ matrices with determinant } 1\ \}$,
e) $M = \mathbf{T}^n = S^1 \times \dots \times S^1$, the n-dimensional torus.
Hint. You can use the fact that a subset A of $\mathbf{R}^k$ is compact if and only if it both bounded (it is contained in a ball of finite radius) and closed (the complement is open).

5) **Topic: Measure zero sets.**
Show that the set of rational numbers has measure zero in the set of real numbers.

6) **Topic: Theorem of Sard.**
Prove that the set of critical values of a smooth function $f : \mathbf{R} \to \mathbf{R}$ has measure zero. In class, we have constructed the **devil staircase function** $g(x)$, which has the Cantor set $C \subset [0,1]$ as the set of critical values. Show that the function $f(x,y) = g(x) + g(y)$ from $\mathbf{R}^2 \to \mathbf{R}$ has the set $[0,2]$ as the set of critical values.
Hint. You have to verify that $C + C = [0,2]$.

7) **(*) Topic: Theorem of Sard.**
Show that a k-dimensional plane in $\mathbf{R}^n$ has measure zero if $k < n$.

8) **(*) Topic: Theorem of Sard.**
A manifold M is called **simply connected**, if every closed smooth curve $t \mapsto \gamma(t)$ in M can be continuously deformed to a point: For each γ, there exists a continuous map $F : S^1 \times [0,1] \to M$ and a point $P \in M$ such that $(t,0) \mapsto F(t,0) = \gamma(t)$ and $F(t,1) = P$, for all $t \in S^1$. Show that S^k is simply connected if $k \geq 2$.
Hint. A closed curve in S^k is given by a smooth map $f : S^1 \to S^k$. Use Sard to find a point, $P \in S^k$, $P \notin f(S^1)$ and use then stereographic projection.

9) **Topic: Partition of Unity.**
Construct an explicit partition of unity of the circle S^1 with two functions g_1, g_2. Hint: Construct first (see the lemma in real analysis) an explicit partition of unity of $\mathbf{R}$ with infinitely many functions h_i such that $h_i(x + 2\pi) = h_{i+2}$ for all i. Construct then the set of functions $g_i : S^1 \to \mathbf{R}$ by $g_i((\cos(\phi), \sin(\phi)) := h_i(\phi)$ which satisfy $g_i(x) = g_{i+2}(x)$.

10) **Topic: Theorem of Whitney, Immersion.**
The **Klein bottle** is a two dimensional compact manifold obtained by identifying the boundaries of a cylinder in opposite direction. It is can be realized as the square $[0,1] \times [0,1]$, with the following identifications of points $(0,t) \sim (1, 1-t), t \in [0,1]$, $(s,0) \sim (s,1), s \in [0,1]$. Draw an immersion of the Klein bottle in $\mathbf{R}^3$ and argue (not prove), why it is not possible to embed it in $\mathbf{R}^3$. Show with Whitney's theorem that one can embed the Klein bottle in $\mathbf{R}^5$. Can you embed it in $\mathbf{R}^4$?

Ma109a, O. Knill, **October 1995**

Week 2 (Solutions)

Due: Friday, October 13, 1995

1) (*) a) $f : \mathbf{R}^2 \to \mathbf{R}^2$, $f(x,y) = (x^2, 2xy)$,

$$Df(x)(u) = (2xu_1 + 0u_2, 2yu_1 + 2xu_2) = Au ,$$

with

$$A = \begin{pmatrix} 2x & 0 \\ 2y & 2x \end{pmatrix} .$$

A point x is a critical point if and only if A is not surjective which is equivalent that the determinant $4x^2$ is vanishing and so $x = 0$. So, $C = \{x = 0\}$. The set of critical values is $\{(0,0)\}$. The complement is the set of regular values.
b) $f : \mathbf{R}^2 \to \mathbf{R}$, $f(x,y) = (x^2 + y^2)^2$,

$$Df(x)(u) = (2(x^2 + y^2)2xu_1, 2(x^2 + y^2)2yu_2)$$

A point is a critical point if and only if $Df(x) : \mathbf{R}^2 \to \mathbf{R}$ is not surjective which is equivalent that $(2(x^2 + y^2)2x, 2(x^2 + y^2)2y) = (0,0)$ or $(x,y) = (0,0)$. The set of critical values is $\{f(0,0)\} = \{0\}$.
c) $f : \mathbf{R}^4 \to \mathbf{R}, f(x) = 1$
Because $Df(x) = 0$, every point is a critical point and the image $\{1\}$ is the set of critical values.
d) $f : \mathbf{R}^3 \to \mathbf{R}^3, f(x) = x$. Since $Df(x)u = u$, there are no critical points and no critical values.

2) **(*) Topic: Immersions, Embeddings.**
a) Take the map f from $M = (0,1)$ to $N = \mathbf{R}^2$ given by $f(t) = (\sin(2\pi t), -\sin(4\pi t))$. You can produce a nice plot with Mathematica:

```
ParametricPlot[{Sin[2Pi t], - Sin[4Pi t]}, {t,0,1}]
```

The image of this map is a figure eight, a compact set, which can not be homeomorphic to the set $(0,1)$ which is not compact.

b) An injective immersion of a compact manifold M in a manifold N is an embedding:
Proof. We know that $Df(x)$ is injective at all points and that f is injective and that the image $f(M)$ is compact. We have to show that f^{-1} is continuous.

Given a sequence $y_n \to y$, with $y_n = f(x_n), y = f(x) \in f(M)$. By the inverse function theorem, there exists a neighborhood U of x such that $f : U \to f(U)$ is a diffeomorphism. By the injectivity of f, the compact set $K = f(M \setminus U)$ does not contain x and there exists a neighborhood V of y, which is disjoint from $f(M \setminus U)$. Therefore, if $y_n \in V$, then $f^{-1}(y_n) \in U$. Take now a sequence of neighborhoods $U_k \subset U$ of x with $\bigcap_k U_k = \{x\}$. Denote by V_k the corresponding neighborhoods of y which have emtpy intersection with $f(M \setminus U_k)$. Given any k. For large enough n, $y_n \in V_k$ so that $x_n \in U_k$. This shows that $x_n \to x$.

c) Take any closed curve $f : S^1 \to \mathbf{R}^2$ which has a selfintersection.

3) **(*) Topic: Compactness.**
a) $M = \mathbf{R}^n$, $n > 0$, not compact because not bounded.
b) $M = S^n$, the n-dimensional sphere. Is bounded and closed: for any $x \notin S_n$, $||x|| \neq 1$, the open set $\{y \in \mathbf{R}^n ||y - x|| < (1 - ||x||)/2$ is also not in S_n which shows that the complement of S^n is open.
c) $M = SL(2, \mathbf{R}) = \{A \mid 2 \times 2$ matrices with determinant $1\}$ is not bounded. Since the sequence $x_n = \text{Diag}(n, n^{-1}) \in SL(2, \mathbf{R})$ has no accumulation point, the manifold M is not compact.
d) $M = SO(3, \mathbf{R}) = \{A \mid$ orthogonal 3×3 matrices with determinant 1 $\}$ is bounded since $|A_{ij}| \leq 1$ and closed: if $AA^T \neq \mathbf{1}$, then also $BB^T \neq \mathbf{1}$ for B near enough to A.
e) $M = \mathbf{T}^n = S^1 \times \ldots \times S^1$, the n-dimensional torus. is compact since S^1 is compact and the theorem of Tychonov says that the product of compact spaces is compact.

4) **(*) Topic: Theorem of Sard.**
To see that the k-dimensional plane in $\mathbf{R}^n$ has measure zero if $k < n$, write it as an image of a smooth map $f : \mathbf{R}^k \to \mathbf{R}^n$ given by $f(x_1, \ldots, x_k) = P + \sum_i x_i e_i$, where P is a point in the plane and e_i are k linearly independent vectors in the plane. Every point $x \in \mathbf{R}^k$ is a critical point because $Df(x)$ has only a k-dimensional range. Therefore, $f(\mathbf{R}^k)$ is the set of critical values. The theorem of Sard tells that this set has measure zero.

5) **(*) Topic: Theorem of Sard.**
Given a closed smooth curve $t \mapsto \gamma(t)$. This is a map from S^1 to S^k. Every point in S^1 is a critical point and the curve $f(S^1)$ has by Sard measure zero in S^k. There exists therefore a point $Q \in S^k$ which is not in the curve. Consider the stereographic projection ϕ_Q from this point which is a smooth diffeomorphism from $S^k \setminus \{Q\}$ to $\mathbf{R}^k$. Consider the map $F(s,t) = \phi_Q(1-s)\phi_Q^{-1}\gamma(t)$. For $s = 0$, $F(0,t) = \gamma_t$ and for $s = 1$, $F(1,t) = \phi_P^{-1}(0) = Q$. We have proven that γ is contractible.

Week 3

Ma109a, O. Knill, **October 1995**

Week 3

Due: Friday, October 20, 1995

Topics: Tensors, Forms, Tensor product, Exterior product

1) **Topic: Tensors.**
Given three vectors v_1,v_2,v_3 in $E = \mathbf{R}^3$. Show that $(v_1, v_2, v_3) \to \det(V)$ is a tensor, where V is a 3×3 matrix with rows v_i. Verify that it is an antisymmetric tensor and so a form.

2) (*) **Topic: Tensors.**
a) Given a tensor $f \in T_0^1(E)$. Let f^i be the components with respect to a basis $\{e_i\}_{i=1}^n$ so that $f = \sum_i f^i e_i$. Assume $\tilde{e}_i = Ae_i$ is an other basis in E, where the coordinate change is given by a linear invertible map $A : E \to E$. Compute the components $\tilde{f}^k$ of the tensor f in the basis $\{\tilde{e}_i\}_{i=1}^n$.
b) Let f be a tensor in $T_1^0(\mathbf{R}^n)$ having the components f_i with respect to a basis $\{e_i\}_{i=1}^n$ so that $f = \sum_{i=1}^n f_i e^i$, where e^i is the dual basis. Compute the components $\tilde{f}_k$ of the tensor f in the same basis $\{\tilde{e}_i\}_{i=1}^n$ as in a).

3) (*) **Topic: Tensors.**
a) Given a tensor $f \in T_0^2(E)$. Let f^{ij} be the components with respect to a basis $\{e_i\}_{i=1}^n$ so that $f = \sum_{i,j=1}^n f^{ij} e_i \otimes e_j$. Assume $\tilde{e}_i = Ae_i$ is an other basis, where A is an invertible linear map $E \to E$. Compute the components $\tilde{f}^{kl}$ of the tensor f in the basis $\{\tilde{e}_i\}_{i=1}^n$.
b) Let f be a tensor in $T_2^0(\mathbf{R}^n)$ having the components f_{ij} with respect to a basis $\{e_i\}_{i=1}^n$, so that $f = \sum_i f_{ij} e^i \otimes e^j$, where e^i is the dual basis. Compute the components $\tilde{f}_{kl}$ of the tensor f in the same basis $\{\tilde{e}_i\}_{i=1}^n$ as in a).
c) Let f be a tensor in $T_1^1(\mathbf{R}^n)$ having the components f_i^j with respect to a basis $\{e_i\}_{i=1}^n$ so that $f = \sum_{i,j=1}^n f_i^j e^i \otimes e_j$, where e^i is the dual basis. Compute the components $\tilde{f}_k^l$ of the tensor f in the same basis $\{\tilde{e}_i\}_{i=1}^n$ as in a).

4) **Topic: Tensors.**
Formulate and prove the general transformation rule for a tensor $f \in T_q^p(E)$ as done in the last two questions.

5) **Topic: Forms.**
Let $E = \mathbf{R}^n$. Show that $\Lambda^p(E)$ is isomorphic to $\Lambda^{n-p}(E)$ by giving an explicit isomorphism.

6) **Topic: Tensor product .**
Give an example which shows that the tensor product is not commutative.

7) (*) **Topic: Tensor product .**
Let E be the 3-dimensional vector space. Given the two tensors $f = f^1 e_1 + f^2 e_2 + f^3 e_3 = e_1 + 2e_2 + 3e_3, g = g^1 e_1 + g^2 e_2 + g^3 e_3 = 4e_1 + 5e_2 + e_3 \in T_0^1(E)$. Compute $f \otimes g$.

8) (*) **Topic: Exterior product.**
a) $E = \mathbf{R}^n$, $n \geq 3$. Let $\alpha \in \Lambda^2(E)$ and $\beta \in \Lambda^1(E)$. Compute $\alpha \wedge \beta(u_1, u_2, u_3)$.
b) Let $E = \mathbf{R}^n$. Show that for $\alpha \in \Lambda^3(E)$, $\alpha \wedge \alpha = 0$. Give an example $\beta \in \Lambda^2(E)$ for which $\beta \wedge \beta \neq 0$.

9) (*) **Topic: Exterior product.**
Given $\alpha^1, \dots, \alpha^n \in \Lambda^1(E)$. Let $\{e_i\}_{i=1}^n$ be a basis in E. Define the matrix $A_j^i = \alpha^i(e_j)$. Verify

$$(\alpha^1 \wedge \alpha^2 \dots \wedge \alpha^n)(e_1, \dots, e_n) = \det(A) \, .$$

10) **Topic: Exterior product.**
Show that in general, not every q-form can be written as an exterior product $f_1 \wedge f_2 \wedge \dots \wedge f_q$ of 1-forms.

Ma109a, O. Knill, **October 1995**

Week 3 (Solutions)

Topics: Tensors, Forms, Tensor product, Exterior product

1) (*) **Topic: Tensors.**
We have seen that if the basis transforms as $\tilde{e}_i = Ae_i$, then the dual basis transforms as $\tilde{e}^k = Be^k = \sum_j B_j^k e^j$, where $B = (A^T)^{-1}$.

a)
$$\tilde{f}^k = f(\tilde{e}^k) = f(Be^k) = f(\sum_i B_i^k e^i) = \sum_i B_i^k f(e^i) = \sum_i B_i^k f^i \, .$$

b)
$$\tilde{f}_k = f(\tilde{e}_k) = f(Ae_k) = f(\sum_i A_k^i e_i) = \sum_i A_k^i e_i f(e^i) = \sum_i A_k^i f_i \, .$$

2) (*) **Topic: Tensors**.

a)

$$\tilde{f}^{kl} = f(\tilde{e}^k, \tilde{e}^l) = f(Be^k, Be^l) = f(\sum_i B_i^k e^i, \sum_j B_j^l e^j) = \sum_{i,j} B_i^k B_j^l f(e^i, e^j) = \sum_{i,j} B_i^k B_j^l f^{ij} .$$

b)

$$\tilde{f}_{kl} = f(\tilde{e}_k, \tilde{e}_l) = f(Ae_k, Ae_l) = f(\sum_i A_k^i e_i, \sum_j A_l^j e_j) = \sum_{i,j} A_k^i A_l^j f(e_i, e_j) = \sum_{i,j} A_k^i A_l^j f_{ij} .$$

c)

$$\tilde{f}_l^k = f(\tilde{e}^k, \tilde{e}_l) = f(Be^k, Ae_l) = f(\sum_i B_i^k e^i, \sum_j A_l^j e_j) = \sum_{i,j} B_i^k A_l^j f(e^i, e_j) = \sum_{i,j} B_i^k A_l^j f_j^i .$$

3) (*) **Topic: Tensor product** . Use the bilinearity of the tensor product to compute

$$\begin{aligned} f \otimes g &= (e_1 + 2e_2 + 3e_3) \otimes (4e_1 + 5e_2 + e_3) \\ &= 4e_1 \otimes e_1 + 5e_1 \otimes e_2 + e_1 \otimes e_3 \\ &+ 8e_2 \otimes e_1 + 10e_2 \otimes e_2 + 2e_2 \otimes e_3 \\ &+ 12e_3 \otimes e_1 + 15e_3 \otimes e_2 + 3e_3 \otimes e_3 . \end{aligned}$$

4) (*) **Topic: Exterior product**.

a) $E = \mathbf{R}^n$, $n \geq 3$. Write $\alpha = \sum_{i<j} \alpha_{ij} e^i \wedge e^j$ and $\beta = \sum_k \alpha_k e^k$.

$$\alpha \wedge \beta(u_1, u_2, u_3) = \alpha(u_1, u_2)\beta(u_3) - \alpha(u_1, u_3)\beta(u_2) + \alpha(u_2, u_3)\beta(u_1) .$$

b) From $\alpha \wedge \alpha = (-1)^9 \alpha \wedge \alpha = -\alpha \wedge \alpha$ follows $\alpha \wedge \alpha = 0$. Take $\alpha = e^1 \wedge e^2 + e^3 \wedge e^4$. Then $\alpha \wedge \alpha = 2(e^1 \wedge e^2 \wedge e^3 \wedge e^4) \neq 0$.

5) (*) **Topic: Exterior product**.

$$(\alpha^1 \wedge \alpha^2 \ldots \wedge \alpha^n)(e_1, \ldots, e_n) = \det(A)$$

is proven by induction with n using the development of the determinant with respect to the n'th row. Let $\beta = \alpha^1 \wedge \alpha^2 \ldots \wedge \alpha^{n-1}$. Then by induction

$$\begin{aligned} \beta \wedge \alpha^n(e_1, \ldots, e_{n-1}) &= \sum_\sigma (-1)^\sigma \beta(e_{\sigma(1)}, \ldots, e_{\sigma(n-1)}) \alpha^n(e_{\sigma(n)}) \\ &= \sum_i \sum_{\sigma, \sigma(n)=i} (-1)^\sigma \beta(e_{\sigma(1)}, \ldots, e_{\sigma(n-1)}) A_i^n \\ &= \sum_i (-1)^{n+i} \det(A^{(n,i)}) A_i^n = \det(A) , \end{aligned}$$

where $A^{(n,i)}$ is the matrix obtained from A_i^n by deleting the n'th row and the i'th column.

Week 4: Midterm

Ma109a, O. Knill, **October 1995**

Midterm

Due: Friday, October 27, 1995

Topics: Manifolds (Chapter 1), Tensors (Chapter 2, Sections 1-4)

Material: There is no restriction in the material. All books, notes and homework can be used. It is only required that everybody makes the midterm by himself/herself.

Time: **4 hours in one sitting**. No credit for work done in **overtime**.

Form: Please use a **blue book** or write on good new paper, which is stapled well together, before turning it in.

Due: Return the paper to the Mailbox of the TA Yue Lei, by 12:00 AM, Friday, October 27, 1995. Note that there is then still no class. Class starts the Monday 30. October again.

Points: there are 6 problems with a total of 150 points. The number for each problem and subproblem is indicated. Note that there are no long computations involved.

1) **Topic: Manifolds** (30 points)

a) (10 points) Show that the n-dimensional torus $\mathbf{T}^n$ and the n-dimensional sphere S^n are diffeomorphic if and only if $n = 1$. Hint. You can without proof use the fact that $\mathbf{T}^n$ is not simply connected, that is, there are closed curves in $\mathbf{T}^n$ which can not be deformed to a point.
b) (10 points) Show that $\mathbf{T}^n$ is not diffeomorphic to $\mathbf{R}^n$.
c) (10 points) Show that $\mathbf{T}^k$ is not diffeomorphic to $\mathbf{T}^n$ for $k < n$. Hint. Find an embedding $f : \mathbf{T}^k \to \mathbf{T}^n$ and use the theorem of Sard.

2) **Topic: Manifolds**. (Total 20 points)

Consider $SU(2)$, the set of all complex 2×2 matrices A such that $A^T\overline{A} = 1$. More precicely, we define for $(a, b, c, d) \in \mathbf{R}^4$, the matrix

$$A(a,b,c,d) = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}$$

and define

$$SU(2) = \{(a,b,c,d) \in \mathbf{R}^4 \mid A^T\,\overline{A} = 1, A = A(a,b,c,d)\ \}$$

where A^T is the transpose of A and $\overline{A}$ is the complex conjugate of A.
a) (10 points) Prove that $SU(2)$ is a manifold by identifying it as the set $f^{-1}(h)$ where $f : \mathbf{R}^4 \to \mathbf{R}$ is smooth and where $h \in \mathbf{R}$ is a regular value for f.
b) (10 points) Prove that $SU(2)$ is diffeomorphic to S^3.

3) **Topic: Manifolds** (Total 30 points)

Let N be a n-dimensional manifold and let M be the manifold $N \times \mathbf{R}^n$. Given a smooth map $H : M \to \mathbf{R}, x = (q,p) \mapsto H(q,p)$. The differential equations

$$\dot{q} = \frac{\partial}{\partial p} H(q,p),\ \dot{p} = -\frac{\partial}{\partial q} H(q,p),$$

describe a so called **Hamiltonian system** on M. (Our solar system is on a classical level described by such a system). Because

$$\begin{aligned} \frac{d}{dt} H(q(t),p(t)) &= \frac{\partial}{\partial q} H(q,p)\dot{q} + \frac{\partial}{\partial p} H(q,p)\dot{p} \\ &= \frac{\partial}{\partial q} H(q,p) \frac{\partial}{\partial p} H(q,p) - \frac{\partial}{\partial p} H(q,p) \frac{\partial}{\partial q} H(q,p) = 0\ , \end{aligned}$$

the **energy surface** $S_h := \{x = (q,p) \in M \mid H(q,p) = h\} \subset M$ is left invariant by the dynamics. Knowing the topology of the set S_h is an important first step to understand the dynamics.

a) (10 points) Show that S_h is a manifold if there is no point $(q,p) \in S_n$ for which $(\dot{q}, \dot{p}) = (0,0)$.

Let from now on the special case, where M is the zylinder $M = \mathbf{T}^1 \times \mathbf{R}$, where $\mathbf{T}^1 = \mathbf{R}/(2\pi\mathbf{Z})$ is the one dimensional circle obtained by identifying points $q, q + 2k\pi$ on $\mathbf{R}$ for $k \in \mathbf{Z}$. The smooth map $H : M \to \mathbf{R}$,

$$H(q,p) = \frac{p^2}{2m} + g \cdot \sin(q)$$

is the energy function of the pendulum, m is the mass of the pendulum and g is the gravitation constant.

b) (10 points) For which energy values $h \in \mathbf{R}$ is the energy surface S_h a manifold?

c) (10 points) For which energy values $h \in \mathbf{R}$ is S_h a connected manifold?
(A manifold M is called **connected**, if any two points can be connected by a path, that is if for any pair of points $x, y \in M$, there exists a smooth curve $\gamma : \mathbf{R} \to M$ satisfying $\gamma(0) = x$ and $\gamma(1) = y$).

4) **Topic: Immersion, Embedding** (Total 30 points)
Determine from each of the following maps from $M = (0, 2\pi)$ to $N = \mathbf{R}^2$ whether they are immersions, embeddings or nothing from both. Prove your claims.

a) (10 points) $t \in (0, 2\pi) \mapsto (\cos(2t), \sin(2t))$.
b) (10 points) $t \in (0, 2\pi) \mapsto (\cos(3t), \sin(5t))$.
c) (10 points) $t \in (0, 2\pi) \mapsto (\cos(2t), \cos(3t))$.

5) **Topic: Linear Tensors, Tensor product, Exterior product**. (Total 20 points).
Let $E = \mathbf{R}^3$
a) (10 points) Compute the tensor product $g \otimes h$ of the tensors

$$g = g^{13} e_1 \otimes e_3 + g^{31} e_3 \otimes e_1 \in T_0^2(E) \; ,$$

$$h = h_{12} e^1 \otimes e^2 + h_{11} e^1 \otimes e^1 \in T_2^0(E) \; .$$

b) (10 points) Compute the exterior product of two tensors $f = dx^1 \wedge dx^2 \in \Lambda^2(E), g = 2dx^2 \wedge dx^3 \in \Lambda^2(E)$.

6) **Topic: Exterior derivative**. (Total 20 points).
Let M be our universe, that is a four dimensional manifold. Given a 1-form $A \in \Lambda^1(M)$

$$A(x) = A_1(x)dx^1 + A_2(x)dx^2 + A_3(x)dx^3 + A_4(x)dx^4$$

on M. (The 1-forms on M are describing an electromagnetic vector potential).
a) (10 points) Compute the exterior derivative $F = dA = \sum_{i<j} F_{ij}(x) dx^i \wedge dx^j$. (The 2-form $F = dA \in \Lambda^2(M)$ describes an eletromagnetic field).
b) (10 points) Compute the exterior derivative dF of the 2-form F. (The equation you get is one group of the **Maxwell equations**).

Ma109a, O. Knill, **October 1995**

Midterm (Solutions)

1) **Topic: Manifolds** (30 points)
a) (10 points) If $n = 1$, then $\mathbf{T}^1 = S^1$ by definition since $\mathbf{T}^n$ was defined as the product $S^1 \times S^1 \times \ldots \times S^1$.
We have shown in an exercice that S^n is simply connected for $n \geq 2$. However, $\mathbf{T}^n$ is not simply connected. If two manifolds are diffeomorphic, then one manifold is simply connected if and only if the other manifold is simply connected.
b) (10 points) $\mathbf{T}^n$ is compact while $\mathbf{R}^n$ is not compact. If two manifolds are diffeomorphic, then either both are compact or both are not compact.
c) (10 points) Assume there $f : \mathbf{T}^k \to \mathbf{T}^n$ is a diffeomorphism. Every point $x \in \mathbf{T}^k$ is a critical point because $Df(x)$ is not surjective (it has rank k). The theorem of Sard tells that the image has measure zero in $\mathbf{T}^n$. It can therefore not be surjective.

2) **Topic: Manifolds**. (Total 20 points)
a) (10 points) The condition $A^T A = 1$ is equivalent to $a^2 + b^2 + c^2 + d^2 = 1$. Define therefore $f(a, b, c, d) = a^2 + b^2 + c^2 + d^2$. Then $SU(2)$ can be identified with the subset $f^{-1}(h)$ of $\mathbf{R}^4$. A point $(q, r, s, t) \in \mathbf{R}^4$ is a regular point if $Df(q, r, s, t) = (2q, 2r, 2s, 2t)$ is not vanishing. This is the case if and only if $q^2 + r^2 + s^2 + t^2 \neq 0$. Every $h \neq 0$ is therefore a regular value and especially $h = 1$. By the lemma proven in class, the set $SU(2)$ is a real manifold.
b) (10 points) The diffeomorphism between $SU(2)$ and S^3 is given by the identification between 2×2 matrices and $\mathbf{R}^4$, that is by the already in a) used map $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, c, d) \in \mathbf{R}^4$.

3) **Topic: Manifolds** (Total 30 points)
Note that the "if and only" claim is not true. The energy surface can be a manifold either with smaller dimension as in question b) below.

a) (10 points) There is no point $(q,p) \in S_h$ for which

$$(\dot{q},\dot{p}) = (\frac{\partial}{\partial p}H(q,p), -\frac{\partial}{\partial q}H(q,p)) = (0,0)$$

if and only if (q,p) is a regular point for the Hamiltonian H. Because $S_h = H^{-1}(h)$, we know by the lemma proven in class that S_h is a manifold if h is a regular value.

b) (10 points) By a), S_h is a $n-1$ dimensional manifold if $(2p, \cos(q)) \neq (0,0)$ that is if (q,p) is disjoint from the two sets S_{-g}, S_g which pass through the two points $(\pi/2, 0), (3\pi/2, 0)$. The first set is consisting of a point, the second set is the set $\{p^2/2 + g\sin(q) = g\}$ which is not a manifold (it is not locally Euclidean near the point $(0,0)$).

c) (10 points) For $h > g$, the manifold S_h contains a circle above the line $p = 0$ and a circle below $p = 0$. It is therefore not connected. For $-g < h < g$, the manifold S_h is a circle and so connected. (Physically, in the first situation $h > g$, where the energy is large, the pendulum rotates around its center either in one direction or in the other direction. In the second situation, the pendulum has not enough energy to turn around and oscillates. In the situation $h = -g$, the pendulum is at rest and can not move.)

4) **Topic: Immersion, Embedding** (Total 30 points)
a) (10 points) $f(t) = (\cos(2t), \sin(2t))$ is an immersion because

$$Df(t)(u) = 2(-\sin(2t)u, \cos(2t)u) \neq (0,0) .$$

It is not an embedding, because the map f is not injective: $f(t) = f(t+\pi)$.

b) (10 points) $t \in (0, 2\pi) \mapsto (\cos(3t), \sin(5t))$ is an immersion because $(-3\sin(3t), 5\cos(5t)) \neq 0$. It is not an embedding because the map is not injective: $f(\pi/6) = f(5\pi/6)$.

c) (10 points) $t \in (0, 2\pi) \mapsto (\cos(2t), \cos(3t))$ is neither an immersion nor an embedding, $Df(t) = (0,0)$ for $t = 0$ and we have therefore no immersion. A map which is no immersion can also not be an embedding.

5) **Topic: Linear Tensors, Tensor product, Exterior product**. (Total 20 points).
a) (10 points)

$$\begin{aligned}
(g^{13}e_1 \otimes e_3 + g^{31}e_3 \otimes e_1 \quad & \otimes \quad (h_{12}e^1 \otimes e^2 + h_{11}e^1 \otimes e^1) \\
& = \quad g^{13}h_{12}e_1 \otimes e_3 \otimes e^1 \otimes e^2 + g^{13}h_{11}e_1 \otimes e_3 \otimes e^1 \otimes e^1 \\
& + \quad g^{31}h_{12}e_3 \otimes e_1 \otimes e^1 \otimes e^2 + g^{31}h_{11}e_3 \otimes e_1 \otimes e^1 \otimes e^1
\end{aligned}$$

b) (10 points) The exterior product of two tensors $f = dx^1 \wedge dx^2 \in \Lambda^2(E), g = 2dx^2 \wedge dx^3 \in \Lambda^2(E)$ is vanishing because $dx^2 \wedge dx^2 = 0$.

6) **Topic: Exterior derivative**. (Total 20 points).
a) (10 points) $F = dA = \sum_{i<j} F_{ij}(x) dx^j \wedge dx^i$. By definition $F_{ij}(x) = (\frac{\partial A_i(x)}{\partial x^j} - \frac{\partial A_j(x)}{\partial x^i})$. This means

$$\begin{aligned}
F \quad = \quad & \frac{\partial A_2(x)}{\partial x^1} - \frac{\partial A_1(x)}{\partial x^2} dx^1 \wedge dx^2 + \frac{\partial A_3(x)}{\partial x^1} - \frac{\partial A_1(x)}{\partial x^3} dx^1 \wedge dx^3 \\
+ \quad & \frac{\partial A_4(x)}{\partial x^1} - \frac{\partial A_1(x)}{\partial x^4} dx^1 \wedge dx^4 + \frac{\partial A_3(x)}{\partial x^2} - \frac{\partial A_2(x)}{\partial x^3} dx^2 \wedge dx^3 \\
+ \quad & \frac{\partial A_4(x)}{\partial x^2} - \frac{\partial A_2(x)}{\partial x^4} dx^2 \wedge dx^4 + \frac{\partial A_4(x)}{\partial x^3} - \frac{\partial A_3(x)}{\partial x^4} dx^3 \wedge dx^4 \; .
\end{aligned}$$

b) (10 points) Since $d^2 = 0$, we have $dF = d^2 A = 0$. (If we write $-E^1 = F_{12}, -E^2 = F_{13}, -E^3 = F_{14}$ (electric field) and and $B^1 = F_{34}, -B^2 = F_{24}, B^3 = F_{23}$ (magnetic field), then $dF = 0$ is equivalent to the two Maxwell equations $\text{curl}(E) + \frac{d}{dt}B = 0$ and $\text{div}(B) = 0$, where curl and div are with respect to the space coordinates x^2, x^3, x^4 and where we wrote $t = x^1$ for the time coordinate.)

Week 5

Ma109a, O. Knill, **October 1995**

Week 5

Due: Friday, November 3, 1995

Topics: Vector fields, Tensor fields, differential form, exterior derivative

1) (*) **Topic: Tensor fields, Riemannian metric**
Let M be a manifold and let $g \in T_2^0(M)$ be a tensor field on M which is **symmetric** $g(x)(u,v) = g(x)(v,u), \forall u, v \in T_x M$ and **nondegenerate**: $g(x)(u,u) \geq 0$ and $g(x)(u,u) = 0$ if and only if $u = 0$. The pair (M, g) is called a Riemannian manifold. On a Riemannian manifold, one can find an isomorphism between the cotangent bundle $T^*M = T_1^0(M)$ and the tangent bundle $TM = T_0^1(M)$ by defining $v^\flat(x)(u) = g(x)(v(x), u)$. This is called **lowering of indices**. The inverse map is denoted by $v \mapsto v^\sharp$ and is called **raising of indices**. Verify that $w = v^\flat$ satisfies in coordinates $w_i(x) = \sum_j g_{ij}(x) v^i(x)$ and that $v = w^\sharp$ satisfies in coordinates $v^i(x) = \sum_j g^{ij}(x) w_j(x)$, where the matrix $B(x) = g^{ij}(x)$ is the inverse matrix of $A(x) = g_{ij}(x)$.

2) (*) **Topic: Vector fields, Gradient field.**
Let $M = S^2$ be the two dimensional sphere $S = \{x_1^2 + x_2^2 + x_3^2 = 1\}$. Let $f : M \to \mathbf{R}$ be the function $f(x) = x_3$. Let Y be a vector field on M given by $Y(x) = \text{grad}(f)^\flat(x) = df(x)$ and $\flat$ is defined as in question 1) with respect to the metric $g_{ij} = \delta_{ij}$. Let ϕ_t be the flow defined by this vector field: that is $t \mapsto \phi_t(x)$ is a curve through x so that $d/dt\phi_t(x) = Y(\phi_t(x))$. Show that for all $x \in M$, there exists $y \in M$ with $\phi_t(x) \to y$ for $t \to \infty$. Hint. Consider the function $t \mapsto f(\phi_t(x))$ from $\mathbf{R} \to \mathbf{R}$ and prove that it is monotone.

3) (*) **Topic: Exterior derivative.**
Let $M = \mathbf{R}^3$ and $f \in \Lambda^0(M)$ be given by $f(x,y,z) = x^2 + 3y^3 z$. Compute the 1-form $g = df \in \Lambda^1(M)$ and the 2-form $dg \in \Lambda^2(M)$.

4) (*) **Topic: Exterior derivative.**
Let $f(x) = f_1(x)dx^1 + f_2(x)dx^2 + f_3(x)dx^3$ be a 1- form on a three dimensional manifold. Compute $g = df$.

5) (*) **Topic: Exterior derivative.**
This is a proof of Corollary 2.4.2 in the notes:
Define

$$\tilde{d}\alpha(u_0, \dots, u_p) = \sum_{j=0}^{p} (-1)^j D\alpha(x)(u_j)(u_0, \dots, \hat{u}_j, \dots, u_p) .$$

Show
(i) $\tilde{d}\alpha \in \Lambda^{p+1}(M)$ if $\alpha \in \Lambda^p(M)$. That is verify that $\tilde{d}(\alpha)(x)$ is an antisymmetric tensor for $x \in M$.
(ii) Verify that $\tilde{d}$ is an anti-derivation: for $\alpha \in \Lambda^p(M)$ and $\beta \in \Lambda^q(M)$, one has

$$\tilde{d}(\alpha \wedge \beta) = \tilde{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \tilde{d}\beta .$$

(iii) Verify $(\tilde{d} \circ \tilde{d})\alpha = 0$.
(iv) Show that $\tilde{d}$ is local.
(v) Conclude $\tilde{d}\alpha = d\alpha$, where d is the exterior derivative.

Ma109a, O. Knill, **November 1995**

Week 5 (Solutions)

1) (*) **Topic: Tensor fields, Riemannian metric**
a) In order to rewrite $v^\flat(x)(u) = g(x)(v(x), u)$ in coordinates, write $u = \sum_i u^i e_i$ and $v(x) = \sum_i v^i(x) e_i$, $g(x) = \sum_{i,j} g_{ij}(x) e^i \otimes e^j$ and $v^\flat(x) = \sum_i v_i^\flat(x) e^i$. We get

$$\begin{aligned} \sum_j v_j^\flat(x) u^j &= \sum_j v_j^\flat(x) e^j (\sum_i u^i e_i) \\ &= v^\flat(x) u = g(x)(v(x), u) = \sum_{i,j} g_{ij}(x) v^i(x) u^j \\ &= \sum_j (\sum_i g_{ij}(x) v^i(x)) u^j \end{aligned}$$

and so $v_j^\flat = \sum_i g_{ij}(x) v^i(x)$.
b) Write the relation $v_j^\flat = \sum_i g_{ij}(x) v^i(x)$ as $v^\flat = [g]v$, we get by inversion the other relation $v = [g]^{-1} v^\flat$. In coordinates, $v^i(x) = \sum_j g^{ij}(x) w_j(x)$.

2) (*) **Topic: Vector fields, Gradient field**.
Consider the function $t \mapsto f(\phi_t(x))$ from $\mathbf{R} \to \mathbf{R}$. The derivative $\frac{d}{dt} f(\phi_t(x))$ is by the chain rule

$$\begin{aligned} \sum_i (\frac{\partial}{\partial x^i} f)(\phi_t(x)) \dot{\phi}_t^i(x) &= df(\phi_t(x))(Y(\phi_t(x))) = df(\phi_t(x))(df^\sharp(\phi_t(x))) \\ &= \sum_{i,j} g_{ij}(x) df^i(\phi_t(x)) df^j(\phi_t(x)) = \sum_i df^i(\phi_t(x))^2 \geq 0 \; . \end{aligned}$$

(The special form of the tensor g was not important. We needed only $g(v, v) > 0$ for $v \neq 0$.)
Since $\frac{d}{dt} f(\phi_t(x)) > 0$, $f(\phi_t)$ is monotonically increasing until the maximal value $f(\phi_t) = 1$ is reached. That is $y = (0, 0, 1) \in M$ is the north pole.

Remark. The above calculation shows, it does not matter, what form the metric has. For any **gradient flow**, from which we have just seen an example, the function f is always monotone. One can use this flow in order to find local maxima or critical points of a functional f.

3) (*) **Topic: Exterior derivative**.
Given $f(x, y, z) = x^2 + 3y^3 z$. We have $g(x, y, z) = 2x dx^1 + 3y^2 dx^2 + 3y^3 dx^3$ and so $dg(x, y, z) = 0$.

4) (*) **Topic: Exterior derivative**.

$$df(x) = (\frac{\partial f_2}{\partial x^1} - \frac{\partial f_1}{\partial x^2}) dx^1 \wedge dx^2 + (\frac{\partial f_3}{\partial x^1} - \frac{\partial f_1}{\partial x^3}) dx^1 \wedge dx^3 + (\frac{\partial f_3}{\partial x^2} - \frac{\partial f_2}{\partial x^3}) dx^2 \wedge dx^3 \; .$$

5) (*) **Topic: Exterior derivative**.
(i) $\tilde{d}$ is clearly a multilinear tensor field in $\Gamma(T_q^0(M))$. In order to show that it is a differential form, we have to verify the anti-symmetry:

$$\tilde{d}\alpha(x)(u_{\sigma(0)}, \dots, u_{\sigma(p)}) = (-1)^\sigma \tilde{d}\alpha(x)(u_0, \dots, u_p) \; .$$

That is, we have to show that $d\alpha(x)(u_0, u_1, \dots, u_k, u_{k+1}, \dots, u_p) = -d\alpha(x)(u_0, u_1, \dots, u_{k+1}, u_k, \dots, u_p)$ for any choice $k < k+1$, because any permutation can be written as a product of such permutations. Write

$$\begin{aligned}
\tilde{d}\alpha(x)(u_0, \dots, u_p) &= \sum_{j \neq k,k+1}^{p} (-1)^j D\alpha(x)(u_j)(u_0, \dots, \hat{u}_j, \dots, u_p) \\
&+ (-1)^k D\alpha(x)(u_j)(u_0, \dots, \hat{u}_k, \dots, u_p) \\
&+ (-1)^{k+1} D\alpha(x)(u_j)(u_0, \dots, \hat{u}_{k+1}, \dots, u_p) .
\end{aligned}$$

The last two summands change sign, when k and $k+1$ are interchanged. The first sum $\sum_{j \neq k,k+1}^{p}(\dots)$ on the right hand side changes sign, when $k, k+1$ are interchanged, because α was antisymmetric.

(ii) The claim

$$\tilde{d}(\alpha \wedge \beta) = \tilde{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \tilde{d}\beta .$$

is proven by putting in the definition for $\tilde{d}$ and the formula for the wedge product you verified in an earlier homework: use the notation S^*_{p+q+1} for the set of permutations σ satisfying $\sigma(0) < \dots < \sigma(p)$ and $\sigma(p+1) < \dots < \sigma(p+q)$ and S^{**}_{p+q+1} for the set of permutations σ satisfying $\sigma(1) < \dots < \sigma(p)$ and $\sigma(0) < \sigma(p+1) < \dots < \sigma(p+q)$. Adding the right hand sides of the two equations

$$(\tilde{d}\alpha) \wedge \beta(u_0, \dots, u_{p+q}) = \sum_{j=1}^{p} \sum_{\sigma \in S^*_{p+q+1}} (-1)^j (-1)^\sigma D\alpha(u_j)(u_{\sigma(0)}, \dots, \hat{u}_j, \dots, u_{\sigma(p)}) \beta(u_{\sigma(p+1)}, \dots, u_{\sigma(p+q)}) .$$

$$\begin{aligned}
(-1)^p(\alpha \wedge \tilde{d}\beta)(u_0, \dots, u_{p+q}) &= \sum_{\sigma \in S^{**}_{p+q+1}} (-1)^\sigma \alpha(u_{\sigma(1)}, \dots, u_{\sigma(p)}) \\
&\quad (\sum_{j=1}^{q} (-1)^{p+j} D\beta(u_j)(u_{\sigma(0)}, u_{\sigma(p+1)}, \dots, \hat{u}_{\sigma(p+j)}, \dots, u_{\sigma(p+q)})) .
\end{aligned}$$

gives the right hand side of the equation

$$\begin{aligned}
\tilde{d}(\alpha \wedge \beta)(u_0, \dots, u_{p+q}) &= \sum_{\sigma \in S^*_{p+q+1}} \sum_{j=1}^{p} (-1)^\sigma (-1)^j D\alpha(u_j)(u_{\sigma(0)}, \dots, u_{\sigma(p)}) \beta(u_{\sigma(p+1)}, \dots u_{\sigma(p+q)}) \\
+ \quad \sum_{\sigma \in S^{**}_{p+q+1}} \sum_{j=p+1}^{p+q} & (-1)^\sigma (-1)^j \alpha(u_{\sigma(1)}, \dots, u_{\sigma(p+1)}) D\beta(u_j)(u_{\sigma(0)}, u_{\sigma(p+1)}, \dots, \hat{u}_j, \dots, u_{\sigma(p+q)}) .
\end{aligned}$$

(iii)

$$\begin{aligned}
(\tilde{d} \circ \tilde{d})\alpha(x)(u_0, \dots, u_{p+1}) &= \sum_{i \neq j}^{p+1} (-1)^{i+j} D^2\alpha(x)(u_i, u_j)(u_0, \dots, \hat{u}_i, \dots, \hat{u}_j, \dots, u_{p+1}) \\
&= \sum_{j<i} (-1)^{i+j} D^2\alpha(x)(u_i, u_j)(u_0, \dots, \hat{u}_i, \dots, \hat{u}_j, \dots, u_{p+1}) \\
&+ \sum_{i<j} (-1)^{i+j-1} D^2\alpha(x)(u_i, u_j)(u_0, \dots, \hat{u}_i, \dots, \hat{u}_j, \dots, u_{p+1}) = 0 .
\end{aligned}$$

(iv) $\tilde{d}$ is local because the derivative D is local and a sum of local functions is local.

(v) From the theorem in the notes, we know that $\tilde{d}\alpha = d\alpha$, where d is the exterior derivative.

Week 6

Ma109a, O. Knill, **November 1995**

Week 6

Due: Friday, November 10, 1995

Topics: Integration of forms, Theorem of Stokes, closed and exact forms

1) (*) **Topic: Integration**
Consider the 1-form

$$\alpha = \alpha_1 dx^1 + \alpha_2 dx^2 = ((x^1)^2 + 7x^2)dx^1 + (-x^1 + x^2 \sin((x^2)^5))dx^2$$

on the manifold $M = \mathbf{R}^2$. Compute the integral $\int_c \alpha$ over the 1-chain $c = 2\sigma_1 + 3\sigma_2 - \sigma_3$, where the simplices σ_i are defined as

$$\begin{aligned}
\sigma_1 &: \Delta^1 = [0,1] \to M,\ \sigma_1(t) = (t, 0) \\
\sigma_2 &: \Delta^1 = [0,1] \to M,\ \sigma_2(t) = (1-t, t) \\
\sigma_3 &: \Delta^1 = [0,1] \to M,\ \sigma_3(t) = (0, 1-t)\ .
\end{aligned}$$

2) (*) **Topic: Exact and closed forms, Theorem of Stokes.**
A p-form α is called **closed**, if $d\alpha = 0$. It is called **exact**, if $\alpha = d\beta$ for some $(p-1)$-form β.
a) Show that the 1-form

$$\alpha = \alpha_1 dx^1 + \alpha_2 dx^2 = (2x^1 + x^2\cos(x^1x^2))dx^1 + (x^1\cos(x^1x^2))\ dx^2$$

is closed.
b) Show that α is exact, by finding a 0-form β such that $\alpha = d\beta$.
c) Let c be the chain $c = \sigma_1 + \sigma_2 + \sigma_3$ where the σ_i are as in problem 1. Show that there exists a 2-chain b such that $c = \delta b$.
d) Compute $\int_c \alpha$ over the chain $c = \sigma_1 + \sigma_2 + \sigma_3$ defined in c) and where α is the one form $\alpha = \alpha_1 dx^1 = x^2 dx^1$.
Note: In the whole question, expressions like x^2 mean coordinates $x = \sum_i x^i e_i$ and not exponents.

3) ( ) **Topic: De Rham cohomology (Preparation for material coming later in class).**
a) Verify that the set of closed p-forms is a vector space in $\Lambda^p(M)$.
b) Verify that the set of exact p-forms is a vector space in $\Lambda^p(M)$.
c) Verify that the set of exact p-forms is a vector space in the set of closed p-forms.
d) Consider in the set of closed forms the equivalence relation $\alpha \sim \beta \Leftrightarrow \alpha + \gamma = \beta$, for an exact p-form γ. Verify that the set of equivalence classes forms a vector space. (This vector space is denoted by $H^p(M, \mathbf{R}) = \{\text{closed forms}\}/\{\text{exact forms}\}$ and it is called the p-th **de Rham cohomology group**).

4) (*) **Topic: Theorem of Green in the plane.**
Given a vector field $v = v^1e_1 + v^2e_2 = x^1e_1 + (x^2 + x^1)e_2$ in the plane $M = \mathbf{R}^2$ equiped with the flat metric $g_{ij}(x) = \delta_{ij}$ (see Week 5 homework). By index lowering, we can associate to v the 1-form $\alpha = v^\flat$, given by $\alpha(x) = \alpha_1(x)dx^1 + \alpha_2(x)dx^2$, where $\alpha_i(x) = \sum_j g_{ij}(x)v^j(x) = v^i(x)$.
a) Compute $d\alpha$.
b) Let c by the 2-chain $c = \sigma$, where σ is the simplex

$$\sigma : \Delta^2 = \{x = (x^0, x^1, x^2) \in \mathbf{R}^3 \mid x^0 + x^1 + x^2 = 1, x^i \geq 0\} \to \mathbf{R}^2$$

given by $\sigma(x) = (x^1, x^2)$. Compute

$$\int_c d\alpha \; .$$

c) Compute $\int_{\delta c} \alpha$ directly.
d) Compute $\int_{\delta c} \alpha$ with the theorem of Stokes using b).

5) (*) **Topic: Gradient, Divergence, Curl.**
Recall. Let $v = (v^1, v^2, v^3)$ be a vector field in $\mathbf{R}^3$. We associate with this vector field a 1-form α by index lowering as in the previous problem: $\alpha = v^\flat = \alpha_1 dx^1 + \alpha^2 dx^2 + \alpha_3 dx^3$.
Definition. To any 1-form $\beta = \beta_1 dx^1 + \beta_2 dx^2 + \beta_3 dx^3$ we can associate the 2-form $*\beta = \beta_3 dx^1 \wedge dx^2 + \beta_1 dx^2 \wedge dx^3 - \beta_2 dx^1 \wedge dx^3$. Denote also the inverse operation with $*$ so that $** \alpha = \alpha$.
Definition. The curl(v) of a vector field v is defined as a vector field w which satisfies

$$w^\flat = *dv^\flat \; .$$

Definition. The **divergence** div(v) of a vector field v is defined as the function $\mathrm{div}(v) = *d*v^\flat$.
Definition. Given a function f on $\mathbf{R}^3$, define the **gradient** field as the vector field $\mathbf{grad}(f) = df^\sharp$.
a) Verify that div curl$(v) = 0$.
b) Verify that curl grad$(f) = 0$.

6) (*) **Topic: Fundamental theorem of calculus, Stokes, Divergence theorem.**
a) Let σ be a 1-simplex in $\mathbf{R}^3$ and let α be a 0-form in $\mathbf{R}^3$. Let g be the tensor in $T_2^0(\mathbf{R}^3)$ as in the last problem. Use the theorem of Stokes to verify

$$\int_0^1 \mathrm{grad}^b(f)(\sigma(t)))(D\sigma(t)) \, dt = f(\sigma(1)) - f(\sigma(0)) \; .$$

Hint. In Calculus, a 1-simplex σ is a curve γ and one writes $\int_0^1 \mathrm{grad}(f)(\gamma(t) \cdot \gamma'(t) \; dt$ for the line integral on the left hand side, where $v \cdot w$ is the scalar product between vectors which is in our case more accuratly replaced by $g(v, w)$.
b) Let σ be a 2-simplex in $\mathbf{R}^3$ and α be a 1-form in $\mathbf{R}^3$. Use the theorem of Stokes to verify the theorem of Stokes in $\mathbf{R}^3$, that is $\int_{\sigma(\Delta^2)} \mathrm{curl}(v) = \int_{\delta\sigma(\Delta^2)} v \; dO$.
Hint. Associate to α a vector field v and to the 2-form $d\alpha$ a vector field $w = \mathrm{curl}(v)$. Identify the integral of the 2-form $d\alpha$ over a 2-chain as a surface integral of curl(v) through the surface $\sigma(\Delta^2)$, so that the theorem of Stokes becomes the usual theorem of Stokes in $\mathbf{R}^3$.
c) Let σ be a 3-simplex in $\mathbf{R}^3$ and α a 2-form in $\mathbf{R}^3$. Use the theorem of Stokes to verify the divergence theorem $\int_{\delta\sigma} v = \int_\sigma \mathrm{div}(v)$.
Hint. Associate to α a vector field v and to the 3-form $d\alpha$ the function div(v), so that the relation $\int_\sigma d\alpha = \int_{d\sigma} \alpha$ becomes the divergence theorem.

Ma109a, O. Knill, **November 1995**

Week 6 (Solutions)

The aim of this homework was to let you see that the integration of k-forms on k dimensional chains is a generalisation of line integrals, surface integrals, volume integrals you know from calculus and that the theorems of Green, Stokes, Gauss are generalized by the theorem of Stokes for chains. The integration in calculus is not satisfactory. For example, in three dimensions, too many spaces are identified: like the space of vectors, the space of one-forms, the space of two-forms. This makes the extension to higher dimensions impossible without a change of approach. Historically, it is a bad accident that the language of differential forms (Cartan calculus) was invented only in this century, partly forced by the fact that general relativity needed a theorem of Stokes also in four dimensional space.

1) Since many of you asked questions about this problem, there is here a (maybe too) detailed solution:

The **integration of a** 1-**form** α over a 1-simplex $\sigma : \Delta \to M$ was defined as $\int_\sigma \alpha = \int_\Delta \sigma^*\alpha$, which is an integral of a 1-form over a subset of $\mathbf{R}$ which is just usual integration over $\Delta = [0,1] \subset \mathbf{R}$. Given $\sigma(t) = (x(t), y(t))$, then the **pull-back** $(\sigma^*\alpha)(t)(u)$ was defined as $\alpha(\sigma(t))(D\sigma(t)u)$ for any $u \in T_t\Delta$. Note that $A(t) = D\sigma(t) : \mathbf{R} \to \mathbf{R}^2$ is given by the 1×2 matrix $\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix}$. To get the pull-back of $\alpha(x) = \alpha_1(x)dx^1 + \alpha_2(x)dx^2$ and $\sigma^*\alpha(t) = \beta(t)\,dt$ in coordinates, use $dx^1(e_2) = 0$ and $dx^2(e_1) = 0$ and $dx^1(e_1) = 1$ and $dx^2(e_2) = 1$ to get

$$\begin{aligned}
\beta(t)(u) &= \alpha(\sigma(t))(A(t)u) \\
&= \alpha_1(\sigma(t))dx^1(A(t)u) + \alpha_2(\sigma(t))dx^2(A(t)u) \\
&= \alpha_1(\sigma(t))dx^1((A(t)u)^1e_1 + (A(t)u)^2e_2) + \alpha_2(\sigma(t))dx^2((A(t)u)^1e_1 + (A(t)u)^2e_2) \\
&= \alpha_1(\sigma(t))((A(t)u)^1) + \alpha_2(\sigma(t))((A(t)u)^2) \\
&= [\alpha_1(\sigma(t))\dot{x}(t) + \alpha_2(\sigma(t))\dot{y}(t)]u\ .
\end{aligned}$$

We obtain therefore

$$\int_\sigma \alpha = \int_0^1 \alpha_1(\sigma(t))\dot{x}(t) + \alpha_2(\sigma(t))\dot{y}(t)\ dt\ ,$$

which is the usual integral of a curve $\sigma : \Delta = [0,1] \to \mathbf{R}^2$ from calculus. We get $D\sigma_1(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $D\sigma_2(t) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $D\sigma_3(t) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. Therefore

$$\int_{\sigma_1} \alpha = \int_0^1 \alpha_1(\sigma(t)) \cdot 1 + \alpha_2(\sigma(t)) \cdot 0\ dt = \int_0^1 t^2\ dt = \frac{t^3}{3}|_0^1 = \frac{1}{3}\ .$$

$$\begin{aligned}
\int_{\sigma_2} \alpha &= \int_0^1 \alpha_1(\sigma(t)) \cdot (-1) + \alpha_2(\sigma(t)) \cdot 1\ dt \\
&= -\int_0^1 (-t)^2 + 7t^2\ dt + \int_0^1 (-t) + t\sin(t^5)\ dt \\
&= -\frac{1}{3} - \frac{7}{2} - \frac{1}{2} + \int_0^1 t\sin(t^5)\ dt \\
&= -\frac{13}{3} + \int_0^1 t\sin(t^5)\ dt\ .
\end{aligned}$$

$$\int_{\sigma_3} \alpha = \int_0^1 -\alpha_2(\sigma(t))\ dt = -\int_0^1 (1-t)\sin((1-t)^5)\ dt$$

so that (one integral can not be solved elementarily (a property of most integrals))

$$\begin{aligned}
\int_c \alpha &= 2\frac{1}{3} + 3[-\frac{13}{3} + \int_0^1 t\sin(t^5)\ dt] + \int_0^1 (1-t)\sin((1-t)^5)\ dt \\
&= -\frac{37}{3} + 4\int_0^1 t\sin(t^5)\ dt\ .
\end{aligned}$$

The integral can be written as a series: $1/7 - 1/102 + 1/3240 - 1/186480 \ldots$ by making a Taylor expansion and integrating each term.

2)

$$\begin{aligned} d\alpha &= \frac{\partial}{\partial x^2}(2x^1 + x^2\cos(x^1x^2))dx^2 \wedge dx^1 + \frac{\partial}{\partial x^1}(x^1\cos(x^1x^2))\, dx^1 \wedge dx^2 \\ &= [\cos(x^1x^2) - x^1x^2\sin(x^1x^2)]dx^2 \wedge dx^1 \\ &\quad +[\cos(x^1x^2) - x^1x^2\sin(x^1x^2)]dx^1 \wedge dx^2 \\ &= 0\,. \end{aligned}$$

b) $\beta = (x^1)^2 + x^1x^2\cos(x^1x^2)$ satisfies $\alpha = d\beta$ so that α is exact.

c) Let b be the simplex $b : \Delta^2 = \{(x_0, x_2, x_2) \mid x_0 + x_1 + x_2 = 0, x_0, x_1, x_2 \geq 0\} \to \mathbf{R}^2$. From the definition of the boundary we have $\delta b = c = \sigma_1 + \sigma_2 + \sigma_3$. (It is here more convenient to describe Δ^2 in homogeneous coordinates.)

d) Use Stokes

$$\int_c \alpha = \int_b d\alpha = -\int_b 1 dx^1 \wedge dx^2 = -\int_b 1 = -\frac{1}{2}\,,$$

because $\int_b 1 = 1/2$ is the area of the triangle $b(\Delta^2)$.

4) a)

$$d\alpha = (\frac{\partial v^2(x)}{\partial x^1} - \frac{\partial v^1(x)}{\partial x^2})dx^1 \wedge dx^2\,.$$

b) Consider $\sigma : \Delta^2 \to \sigma(\Delta^2)$. By definition,

$$\int_c d\alpha = \int_{\Delta^2} \sigma^* d\alpha\,.$$

In order to compute this integral, let U be the triangle $U = \sigma(\Delta^2) \subset \mathbf{R}^2$. Define a map $\phi : U \to \Delta^2$ as the inverse of σ. Then

$$\int_{\Delta^2} \sigma^* d\alpha = \int_U \phi^*\sigma^* d\alpha = \int_U d\alpha \ = \int_U (\frac{\partial v^2(x)}{\partial x^1} - \frac{\partial v^1(x)}{\partial x^2})dx^1 dx^2\,,$$

because $\int_U f(x)dx^1 \wedge dx^2 = \int_U f(x)dx$.

c) δc is a sum of one dimensional simplices $\sigma_1 + \sigma_2 + \sigma_3$ and we can associate to it a path γ in the plane. By problem 1), integration of a 1-form α over such a simplex is the line integral of the vector field $x \mapsto v(x) = \alpha^\flat(x)$ over the path γ

$$\int_{\delta c} \alpha = \int_\gamma v\, ds\,.$$

d) Using Stokes, we get

$$\int_{\delta c} \alpha = \int_c d\alpha$$

The left hand side is by c) the line integral $\int_\gamma v(\gamma(t)) \cdot \dot\gamma(t)\, dt$, where γ is the boundary of the region $c(\Delta^2) \subset \mathbf{R}^2$. The right hand side is by b)

$$\int_U (\frac{\partial v^2(x)}{\partial x^1} - \frac{\partial v^1(x)}{\partial x^2})dx^1 dx^2\,.$$

5) By definition

$$\mathrm{curl}(v) = (*dv^\flat)^\sharp \; .$$
$$\mathrm{div}(v) = *d*v^\flat$$
$$\mathrm{grad}(f) = (df)^\sharp$$

a) We compute, using $**\alpha = (-1)^{p(3-p)}\alpha = \alpha$ and $dd = 0$ and $\sharp^{-1} = \flat$

$$\mathrm{div}\ \mathrm{curl}(v) = *d**dv^\flat = *ddv^\flat \; .$$

b)

$$\mathrm{curl}\ \mathrm{grad}(f) = (*ddf)^\sharp = 0 \; .$$

6) See also Abraham-Marsden-Ratiu p. 504-510. The translation between the different languages is the following: (it is convenient to take the following three boxes as **definitions** of the line integral, surface integral or volume integral):

The **line integral** of a vector field v over a curve γ is

$$\int_\gamma v \cdot ds = \int_\sigma v^\flat \; ,$$

where $\sigma(t) = \gamma(t)$ is the one dimensional simplex. The right hand side is the integral of a one-form over a one simplex.

The **surface integral** of a vector field v through a surface S is

$$\int_S v \cdot dO = \int_S (*v^b) = \int_\sigma (*v^b) \; ,$$

where $\sigma : \Delta^2 \to \mathbf{R}^3$ is the simplex assocated to the surface. The right hand side is an integral of a two-form over a two dimensional simplex.

The **volume integral** of a function field f over a region G in $\mathbf{R}^3$ is

$$\int_G f \; dV = \int_\sigma *f \; ,$$

where σ is the simplex assocated with the region G. The right hand side is an integral of a three-form over a three dimensional simplex.

a) The fundamental theorem of calculus is obtained from the Stokes for forms (identify the curve γ with a one dimensional simplex σ)

$$\begin{aligned} \int_\gamma \mathrm{grad}(f)\; ds &= \int_\sigma \mathrm{grad(f)}^\flat \\ &= \int_\sigma (df^\sharp)^\flat \\ &= \int_\sigma df = \int_{\delta\sigma} f \\ &= f(\sigma(1)) - f(\sigma(0)) \; . \end{aligned}$$

b) Similarly, the theorem of Stokes is obtained from the theorem of Stokes for forms: (identify the surface S with a two dimensional simplex σ)

$$\begin{aligned}
\int_S \mathrm{curl}(v) \ \cdot dO &= \int_\sigma *(\mathrm{curl}(v))^\flat \\
&= \int_\sigma *(*d(v^\flat)) \\
&= \int_\sigma d(v^\flat) \\
&= \int_{\delta\sigma} v^\flat \\
&= \int_{\delta S} v \ ds \ .
\end{aligned}$$

c) Also the theorem of Gauss is obtained from the theorem of Stokes for forms: (identify the region G with a three dimensional simplex σ)

$$\begin{aligned}
\int_G \mathrm{div}(v) \ dV &= \int_\sigma *\mathrm{div}(v) \\
&= \int_\sigma **d*v^\flat \\
&= \int_\sigma d*v^\flat \\
&= \int_{\delta\sigma} *v^\flat \\
&= \int_{\delta G} v \ \cdot dO \ .
\end{aligned}$$

Week 7

Ma109a, O. Knill, **November 1995**

Week 7

Due: Friday, November 17, 1995

Topics: Riemannian manifolds, change of metric under transformations, Hodge star operation

1) (*) **Topic: Flat manifold in spherical coordinates**
   a) Compute the coordinates of the metric tensor in $\mathbf{R}^3$ in spherical coordinates (r, θ, ϕ). Hint: You have to verify that
   $$[g_{ij}](x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2(\theta) \end{pmatrix} .$$
   b) Let $\phi : S^2 = \{||x|| = 1\} \to \mathbf{R}^3$ be the incusion $\phi(x) = x$. Compute the pull-back metric $\phi^* g$ on S^2.

2) ( ) **Topic: Schwarzschild metric**
   Let (M, g) be the Schwarzschild manifold. That is $M = \{x = (t, r, \theta, \phi) \mid r \neq 2m\}$ and
   $$[g_{ij}] = \begin{pmatrix} -(1 - 2m/r) & 0 & 0 & 0 \\ 0 & (1 - 2m/r)^{-1} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2(\theta)) \end{pmatrix} .$$

(We assume here $c = 1$). Find the lengths of the following vectors $u = u(x), v = v(x), w = w(x)$ in the tangent space $T_x M$ and the angles between them: $u = (1,0,0,0), v = (0,1,0,0), w = (1,(1-2m/r),0,0)$.

3) (*) **Topic: Lorentz transformations**
Verify that the linear transformations $u \mapsto Au$ (Lorenz boost) and $u \mapsto Bu$ (Rotation) given by

$$A = \begin{pmatrix} \cosh(\phi) & \sinh(\phi) & 0 & 0 \\ \sinh(\phi) & \cosh(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) & 0 \\ 0 & -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

are preserving the length of a vector in the Lorenz manifold (M,g), where $M = \mathbf{R}^4$ and $[g] = \mathrm{Diag}(-1,1,1,1)$.

4) (*) **Topic: Eddington-Finkelstein metric, Black holes I**
Let (M,g) be the Schwarzschild manifold from problem 1). It describes the metric of space-time near a star or a planet of mass $\tilde{m}$. If you look at the metric tensor g, then it blows up at $r = 2m = 2G \cdot \tilde{m}/c^2 = 2G\tilde{m}$, the **Schwarzschild radius**. If we want to see, what happens, when the star is smaller than the Schwarzschild radius, we need to introduce a new coordinate system. Define the map $\phi : \mathbf{R}^4 \to \mathbf{R}^4$ by

$$\phi(t,r,\theta,\phi) = (v,r,\theta,\phi)$$

where

$$v = t + r + 2m \log \left|\frac{r}{2m} - 1\right| .$$

Verify that the push-forward metric $h = \phi_* g$ is given by

$$[h_{ij}] = \begin{pmatrix} -(1-\frac{2m}{r}) & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\sin^2(\theta)) \end{pmatrix}.$$

It is called **Eddington-Finkelstein metric** and is only singular at $r = 0$.

5) (*) **Topic: Light can not escape black holes**
If we restrict the Schwarzschild manifold (M,h) with a subspace $\{\phi = \mathrm{const}_1, \psi = \mathrm{const}_2\}$, we get the manifold N. Let $\phi : N \to M$ be the embedding. Verify that the pull-back $\phi^* h$ of the Eddington-Finkelstein metric h is

$$[k_{ij}] = [\phi^* h_{ij}] = \begin{pmatrix} -(1-\frac{2m}{r}) & 1 \\ 1 & 0 \end{pmatrix} .$$

Given a point x in the pseudo Riemannian manifold (N,k). The set of null vectors in $T_x N$ is called the **light cone** at x. This name comes from the fact that a radially moving photon, which is described by a path $\tau \mapsto \gamma(\tau)$ in N, satisfies always $D\gamma(\tau)$ in the light cone.
a) Verify that the light cone at a point $x = (v,r)$ in N is given by the union of

$$\begin{aligned} L_{in}(t,r) &= \{\lambda \cdot (0,1) \mid \lambda \in \mathbf{R}\} \subset T_x N, \\ L_{out}(t,r) &= \{\lambda \cdot (2/(1-\frac{2m}{r}),1) \mid \lambda \in \mathbf{R}\} \subset T_x N . \end{aligned}$$

b) Given a world line $\tau \mapsto \gamma(\tau)$ of an **incoming photon**, that is γ satisfies $D\gamma(\tau) \in L_{in}(x)$. Verify that $v(\tau) = A$ is constant.
c) Given a world line $\tau \mapsto \gamma(\tau) = (v(\tau), r(\tau))$ of an **out-going photon** defined by the property that $D\gamma(\tau) \in L_{out}$. Verify that $v(\tau) = 2r(\tau) + 4m \log|r(\tau) - 2m| + B$ for some constant B.
d) Use $b), c)$ to verify the following picture. An incoming radially falling incoming photon can

enter from the region $r > 2m$ into the region $r < 2m$ and reach the singularity $r = 0$. An outgoing photon which is in the region $r < 2m$ at some time t_0, stays in the region $r < 2m$ for time $t > t_0$. An out-going photon which is in the region $r > 2m$ at time t_0 escapes to infinity. In other words, light is trapped inside a black hole.

6) (*) **Topic: Hodge star operation, Maxwell equations**
Consider the Lorenz manifold $M = \mathbf{R}^4$ with the metric $[g_{ij}] = \mathrm{Diag}(-1, 1, 1, 1)$.

$$\begin{aligned} F(x) &= -E^1(x)dx^1 \wedge dx^2 - E^2(x)dx^1 \wedge dx^3 - E^3(x)dx^1 \wedge dx^4 \\ &+ B^3(x)dx^2 \wedge dx^3 - B^2(x)dx^2 \wedge dx^4 + B^1(x)dx^3 \wedge dx^4 \end{aligned}$$

be a 2-form on the Lorenz manifold (M, g), where the electric field $E(x) = (E^1(x), E^2(x), E^3(x))$ and magnetic field $B(x) = (B^1(x), B^2(x), B^3(x))$ are given. We have verified already in the midterm, that $dF = 0$ follows from $F = dA$. With $(t, y^1, y^2, y^3) = (x^1, x^2, x^3, x^4)$, the equation $dF = 0$ is equivalent to the **first group of Maxwell equations**

$$\boxed{\begin{aligned} \mathrm{div}(B) &= 0 \\ \mathrm{curl}(E) &= -\partial B/\partial t\ , \end{aligned}}$$

where div and curl are with respect to the space coordinates $y = (y^1, y^2, y^3)$.
a) Define $d^*\alpha = *d*\alpha$. Verify that d^* is a map from $\Lambda^p(M)$ to $\Lambda^{p-1}(M)$ which satisfies $d^* \circ d^* = 0$.
b) Given the electromagnetic field tensor $F \in \Lambda^2(M)$ as above. Compute d^*F.
c) Let $j(x) = j_1(x)dx^1 + j_2(x)dx^2 + j_3(x)dx^3 + j_4(x)dx^4 = \rho(x)dx^1 + i^1(x)dx^2 + i^2(x)dx^3 + i^3(x)dx^4$ be a 1-form describing the scalar field $\rho(x)$ (time dependent charge) and three dimensional vector field $i = (i^1, i^2, i^3)$ (time dependent current). Verify that $d^*F = j$ is equivalent to the **second group of Maxwell equations**

$$\boxed{\begin{aligned} \mathrm{div}(E) &= \rho \\ \mathrm{curl}(B) - \partial E/\partial t &= i\ . \end{aligned}}$$

Ma109a, O. Knill, **November 1995**

Week 7 (Solutions)

1) We have $g = \Phi^* h$, where h is the flat metric on $\mathbf{R}^3$ and where

$$\Phi(y) = \Phi(r, \theta, \phi) = \begin{pmatrix} r\cos(\phi)\cos(\theta) \\ r\sin(\phi)\cos(\theta) \\ r\sin(\theta) \end{pmatrix} .$$

By definition, $\Phi^*h(x)(u, v) = h(\Phi(x))(D\Phi u, D\Phi v)$ so that g_{ij} is the standard scalar product of the i'th row e_i in $D\Phi$ with the j'th row e_j in $D\Phi$ (see the notes for e_i). That is

$$[g_{ij}](x) = h(e_i(x) \cdot e_j(x)) = [g_{ij}](x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2(\theta) \end{pmatrix} .$$

b) We have $\Phi(\phi, \theta) = (1, \phi, \theta)$ and so

$$D\Phi = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

We get $D\Phi(\phi,\theta)(u_1,u_2) = (0,u_1,u_2)$ and for $i = 1,2$,

$$\begin{aligned}
[\Phi^* g(x)]_{ij} &= g(1,\phi,\theta)(D\Phi(x)e_i, D\Phi(x)e_j) \\
&= [g(1,\phi,\theta)]_{ij} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1\sin^2(\theta) \end{pmatrix} .
\end{aligned}$$

3) Given a vector $u = (u^1,\dots,u^4)$ in T_xM. Its length in the Lorentz metric is

$$||u|| = \sqrt{-(u^1)^2 + (u^2)^2 + (u^3)^2 + (u^4)^2} .$$

We have

$$Au = \begin{pmatrix} \cosh(\phi)u^1 + \sinh(\phi)u^2 \\ \sinh(\phi)u^1 + \cosh(\phi)u^2 \\ u^3 \\ u^4 \end{pmatrix} ,$$

and

$$Bu = \begin{pmatrix} u^1 \\ \cos(\phi)u^2 + \sin(\phi)u^3 \\ -\sin(\phi)u^2 + \cos(\phi)u^3 \\ u^4 \end{pmatrix} ,$$

Using the identity $\cosh^2(\phi) - \sinh^2(\phi) = 1$ we get

$$||Au||^2 = -\cosh^2(\phi)(u^1)^2 - \sinh^2(\phi)(u^2)^2 + \sinh^2(\phi)(u^1)^2 + \cosh^2(\phi)(u^2)^2 + (u^3)^2 + (u^4)^2 = ||u||^2 .$$

Using the identity $\cos^2(\phi) + \sin^2(\phi) = 1$, we have

$$||Bu||^2 = -(u^1)^2 + \cos^2(\phi)(u^2)^2 + \sin^2(\phi)(u^3)^2 + \sin^2(\phi)(u^2)^2 + \cos^2(\phi)(u^3)^2 + (u^4)^2 = ||u||^2 .$$

4) The derivative of $\phi(x) = \phi(t,r,\theta,\phi) = (t + r + 2m\log|\frac{r}{2m} - 1|, r, \theta, \phi)$ is

$$A(x) = D\phi(x) = \begin{pmatrix} 1 & 1 + 1/(\frac{r}{2m} - 1) & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

If $B = (A^{-1})^T$, then the push-forward metric $h = \phi_* g$ is defined by $h(y)(u,v) = h(\phi^{-1}x)(Bu,Bv)$. That is, in coordinates: $h_{kl} = \sum_{i,j} B^i_k B^j_l g_{ij}$, where g is the Schwarzschild metric. This gives

$$[h_{ij}] = \begin{pmatrix} -(1-\frac{2m}{r}) & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\sin^2(\theta)) \end{pmatrix} .$$

5) Let $\phi : N \to M$ be the embedding of N in M. Then $D\phi$ is a constant 2×4 matrix

$$A = D\phi(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} .$$

That is Au sets the last two coordinates of u to zero.
We have $(\phi^* h)(u,v) = h(Au,Av)$ and so

$$[k_{ij}] = [\phi^* h_{ij}] = \begin{pmatrix} -(1-\frac{2m}{r}) & 1 \\ 1 & 0 \end{pmatrix} .$$

a) Given a point $x = (v, r)$ in N, the light cone at x is the set of vectors (u, w) such that

$$k((u,w),(u,w)) = (1 - \frac{2m}{r})u^2 + 2uw = u((1 - \frac{2m}{r}) + 2w) = 0 \; .$$

That is, either $u = 0$, which gives

$$L_{in}(t,r) = \{\lambda \cdot (0,1) \mid \lambda \in \mathbf{R}\} \subset T_x N$$

or if $u \neq 0$, $2w = (1 - \frac{2m}{r})u$, that is

$$L_{out}(t,r) = \{\lambda \cdot (2/(1 - \frac{2m}{r}), 1)\}$$

b) If $D\gamma(\tau) = (u, v) \in (0, \lambda\mathbf{R})$, then $u = \dot{v} = 0$ and $v(\tau)$ is constant.
c) If $v(\tau) = 2r(\tau) + 4m \log |r(\tau) - 2m| + B$, then

$$\frac{dv(\tau)}{d\tau} = 2\frac{dr(\tau)}{d\tau} + \frac{4m}{(r(\tau) - 2m} \frac{dr(\tau)}{(d\tau)} \; .$$

That is $u = 2v + 4m/(r - 2m)w$ and so $(u, w) \in L_{out}$.
d) We have to analyze the two vector fields, one of the incoming and one of the outgoing photon. The incoming photon satisfies $v = const$ and r decreases constantly. The vector field of the outgoing photon is becoming vertical at the Schwarzschild radius and does therefore not cross it. The best is to make a picture (you can find it in almost any book about general relativity for example: J.Foster, J.D. Nightingale "A short course in general relativity", p. 155.

6) a) $d^*\alpha = *d**d*\alpha = \pm * dd * \alpha = 0$ follows from $dd = 0$.
b) From

$$\begin{aligned} F(x) &= -E^1(x)dx^1 \wedge dx^2 - E^2(x)dx^1 \wedge dx^3 - E^3(x)dx^1 \wedge dx^4 \\ &+ B^3(x)dx^2 \wedge dx^3 - B^2(x)dx^2 \wedge dx^4 + B^1(x)dx^3 \wedge dx^4 \end{aligned}$$

we get (write also $e^i = dx^i$)

$$\begin{aligned} *(e^1 \wedge e^2) = -e^3 \wedge e^4 \; , *e^3 \wedge e^4 = (e^1 \wedge e^2) \\ *(e^2 \wedge e^3) = e^1 \wedge e^4 \; , *e^1 \wedge e^4 = -(e^2 \wedge e^3) \\ *(e^1 \wedge e^3) = e^2 \wedge e^4 \; , *e^2 \wedge e^4 = -(e^1 \wedge e^3) \end{aligned}$$

(note that $** \alpha = -\alpha$ for a 2-form in the Lorenz metric)

$$\begin{aligned} *F(x) &= E^1 dx^3 \wedge dx^4 - E^2(x)dx^2 \wedge dx^4 + E^3(x)dx^2 \wedge dx^3 \\ &+ B^3(x)dx^1 \wedge dx^4 - B^2(x)dx^1 \wedge dx^3 + B^1(x)dx^1 \wedge dx^2 \; . \end{aligned}$$

c) Using $**\alpha = \alpha$ for a 1 form, we get (write also $e^i = dx^i$

$$\begin{aligned} e^1 &= -*e^2 \wedge e^3 \wedge e^4 \ , \\ e^2 &= -*e^1 \wedge e^3 \wedge e^4 \ , \\ e^3 &= *e^1 \wedge e^2 \wedge e^4 \ , \\ e^4 &= -*e^1 \wedge e^2 \wedge e^3 \end{aligned}$$

and so

$$\begin{aligned} *d*F &= (-\partial_2 E_1 - \partial_3 E_2 - \partial_4 E_3)dx^1 \\ &\quad +(-\partial_1 E_1 + \partial_3 B_3 - \partial_4 B_2)dx^2 \\ &\quad +(-\partial_1 E_2 + \partial_4 B_2 - \partial_2 B_4)dx^3 \\ &\quad +(-\partial_1 E_3 + \partial_2 B_4 - \partial_3 B_3)dx^4 \end{aligned}$$

so that by comparison of coefficients, $d^*F = j = (-\rho, i_1, i_2, i_3) = (\rho, i_1, i_2, i_3)^\flat$ is indeed the second group of Maxwell equations:

$$\begin{aligned} \mathrm{div}(E) &= \rho \\ \mathrm{curl}(B) - \partial E/\partial t &= i \ . \end{aligned}$$

Week 8

Ma109a, O. Knill, **November 1995**

Week 8 (Last homework)

Due: Monday, November 27, 1995 (after thanksgiving)

Topics: Riemannian manifolds, connections, geodesics, curvature Einstein equations

1) (*) **Topic: Christoffel symbols**
We have defined the Christoffel symbols on a Riemannian manifold (M, g) by $\Gamma^i_{jk} = \sum_l g^{il} g(\nabla_j \partial_k, \partial_l)$. Suppose, we define $\tilde\Gamma$ by $\nabla_i \partial_j = \sum_k \tilde\Gamma^k_{ij} \partial_k$. Prove that $\Gamma^k_{ij} = \tilde\Gamma^k_{ij}$.

2) ( ) **Topic: Christoffel symbols**.
Compute the Christoffel symbols for the two dimensional manifold (M, g), where M is the two dimensional sphere and where g is the metric in the week 7 homework.

3) (*) **Topic: Connection**
Verify that if $t \in \Gamma(T^p_q(M))$ and $s \in \Gamma(T^{p'}_{q'}(M))$, then

$$\nabla_X (t \otimes s) = (\nabla_X t) \otimes s + t \otimes (\nabla_X s) \ .$$

4) (*) **Topic: Covariant derivative**
Write down the general formula for the covariant derivative in coordinates

$$t^{j_1, \dots, j_p}_{i_1, i_2, \dots, i_q; i}$$

using the last question. Hint. Use

$$(t \otimes u)^{j_1, \dots, j_p, l_1, \dots, l_s}_{i_1, \dots, i_q, k_1, \dots, k_r} = t^{j_1, \dots, j_p}_{i_1, \dots, i_q} u^{l_1, \dots, l_s}_{k_1, \dots, k_r}$$

and use induction in p and in q.

5) ( ) **Topic: Covariant derivative**
Verify
$$g_{ij;k} = 0$$
and from that Ricci's theorem
$$g^{ij}_{;k} = 0 \ .$$

6) (*) **Topic: Einstein equations, Curvature**.
Let (M, g) be a pseudo Riemannian manifold. Let Ricci $\in \Gamma(T^0_2(M))$ be the Ricci tensor in (M, g) and let R be the scalar curvature. One defines the **Einstein tensor**
$$G = \text{Ricci} - \frac{1}{2} R g \ .$$
The **Einstein equation** of general relativity is
$$G = 8\pi T \ ,$$
where T is the energy-momentum tensor of matter. If $T = 0$ (vaccuum), the **vaccum Einstein equation** is $G = 0$.

Show that $G = 0$ is equivalent to Ricci $= 0$ if the manifold has dimension 4.
Hint. For Ricci $= 0 \Rightarrow G = 0$ we give no hint. For the other way, write $G = 0$ in coordinates $\text{Ricci}_{ij} = 1/2Rg_{ij} = 0$, raise one coordinate $\text{Ricci}^j_i = 1/2R\delta^j_i$, and form the trace of this equation.

7) (*) **Schwarzschild metric**. (This exercice needs either patience, that is hours of calculation (with the risk of making a mistake), or a trick of special normal coordinates still with nasty calculation (again with the risk of doing a mistake), or a symbolic Mathematica computation which takes a few seconds. Since I chose the third possibility myself, I can lean you my Mathematica source code, which you can find on my Web page or below. To get the credit for this problem, you have to print out the intermediate results, Mathematica gives for the Christoffel symbols and for the Curvature tensor).

Problem. Verify that the Schwarzschild metric g in $\mathbf{R}^4$ satisfies the vaccum Einstein equation. That is, compute
$$2\Gamma^i_{jk} = \sum_l g^{il}(\partial_j g_{kl} + \partial_k g_{lj} - \partial_l g_{jk}) \ .$$
and then
$$R^i_{kmn} = \partial_m \Gamma^i_{kn} - \partial_n \Gamma^i_{km} + \sum_j \Gamma^j_{kn}\Gamma^i_{jm} - \sum_j \Gamma^j_{km}\Gamma^i_{jn}$$
and finally
$$\text{Ricci}_{ik} = \sum_j R^j_{jik}$$
and (using the last problem) verify that $G = 0$.

Hint. Use Mathematica on a Macintosh, PC or Sun or Next workstation and run the following few lines. Identify first from each variable, what it does and whether it is deduced correctly from the formulas you have in the text. I simple method to get also the intermediate results on a Unix machine is to load down the program, store it as "*schwarzschild.m*" on your computer and type in "*math* < *schwarzschild.m* > *results*". In "results", you find all the necessary information. Clean this file by editing the mess (removing unnecessary lines and labeling essential things) and print it out "*lpr results* and attach it to your paper.

```
g={{ -(1-2m/r),     0          ,           0      ,          0      },
   {        0 ,     (1-2m/r)^(-1),         0      ,          0      },
   {        0 ,     0          ,           r^2    ,          0      },
   {        0 ,     0          ,           0      , r^2 Sin[theta]^2 }};
ginverse={{-1/(1-2m/r)    , 0          , 0      , 0                 },
          {0              , 1 - 2*m/r  , 0      , 0                 },
          {0              , 0          , r^(-2) , 0                 },
          {0              , 0          , 0      , 1/(r^2 Sin[theta]^2)}};
x={t,r,theta,phi};
GDown[j_,k_,i_]:=1/2(D[g[[k,i]],x[[j]]] + D[g[[i,j]],x[[k]]] - D[g[[j,k]],x[[i]]]);
GDown=Table[GDown[j,k,i],{j,4},{k,4},{i,4}]
G[j_,k_,i_]:=Sum[ginverse[[i,l]] GDown[[j,k,l]],{l,4}];
G=Table[G[j,k,i],{j,4},{k,4},{i,4}]
Riemann[i_,k_,m_,n_]:=D[G[[k,n,i]],x[[m]]]-D[G[[k,m,i]],x[[n]]]+
                 Sum[G[[k,n,j]] G[[j,m,i]]-G[[k,m,j]] G[[j,n,i]],{j,4}]
Ricci[i_,k_]:=Sum[Riemann[j,i,k,j],{j,4}];
Ricci=Table[Ricci[i,k],{i,4},{k,4}]
Simplify[%]
```

8) ( ) **Kerr-Newman metric**. If you like letting a computer do proofs for you, take for example the Kerr-Newman metric (see for example the new and nice book of Ciufolini and Wheeler "Gravitation and Inertia" p.41) and verify that it satisfies the Einstein equations by modifying the above program.

Ma109a, O. Knill, **November 1995**

Week 8 (Solutions)

1) Replacing $\nabla_j \partial_k = \sum_m \tilde{\Gamma}^m_{jk} \partial_m$ in $\Gamma^i_{jk} = \sum_l g^{il} g(\nabla_j \partial_k, \partial_l)$, using the linearity of g and the definition $g(\partial_k, \partial_l) = g_{kl}$ as well as $\sum_l g^{il} g_{kl} = \delta^i_k$ gives

$$\begin{aligned}
\Gamma^i_{jk} &= \sum_l g^{il} \tilde{\Gamma}^m_{jk} g(\partial_m, \partial_l) \\
&= \sum_{l,m} \tilde{\Gamma}^l_{jk} g^{il} g_{kl} \\
&= \sum_m \tilde{\Gamma}^l_{jk} \delta^i_k \\
&= \tilde{\Gamma}^i_{jk} \, .
\end{aligned}$$

3) By the definition of the covariant derivative

$$\begin{aligned}
\nabla_X (t \otimes s)(Y^1, \dots, Y^{p+p'}, X_1, \dots, X_{q+q'}) &= X.(t \otimes s)(Y^1, \dots, Y^{p+p'}, X_1, \dots, X_{q+q'}) \\
&\quad - \sum_{i=1}^{p+p'} (t \otimes s)(Y^1, \dots, \nabla_X Y^i, \dots, Y^{p+p'}, X_1, \dots, X_{q+q'}) \\
&\quad - \sum_{j=1}^{q+q'} (t \otimes s)(Y^1, \dots, Y^{p+p'}, X_1, \dots, \nabla_X X_j, \dots X_{q+q'}) \, .
\end{aligned}$$

We split up each of the two sums into two sums and use the definition of the tensor product and the Leibniz rule $X.(f\cdot g) = (X.f)\cdot g + f\cdot (X.g)$ to get

$$\begin{aligned}
\nabla_X(t\otimes s)(\cdots) &= (X.t(Y^1,\dots,Y^{p+p'}))s(X_1,\dots,X_{q+q'})\\
&+t(Y^1,\dots,Y^{p+p'}))(X.s(X_1,\dots,X_{q+q'})\\
&-\sum_{i=1}^{p} t(Y^1,\dots,\nabla_X Y^i,\dots,Y^p,X_1,\dots,X_q)s(Y^{p+1},\dots,Y^{p+p'},X_1,\dots,X_{q+q'})\\
&-\sum_{i=p+1}^{p+p'} t(Y^1,\dots,Y^p,X_1,\dots,X_q)s(Y^{p+1}\dots,\nabla_X Y^i,\dots,Y^{p+p'},X_1,\dots,X_{q+q'})\\
&-\sum_{j=1}^{q} t(Y^1,\dots,Y^p,X_1,\dots,\nabla_X X_j,\dots,X_q)s(Y^{p+1}\dots,Y^{p+p'},X_1,\dots,X_{q+q'})\\
&-\sum_{j=p+1}^{q+q'} t(Y^1,\dots,Y^p,X_1,\dots,X_q)s(Y^{p+1},\dots,Y^{p+p'},X_1,\dots,\nabla_X X_j,\dots,X_{q+q'})\,.
\end{aligned}$$

The first expression plus the first sum plus the third sum are

$$(\nabla_X t)(Y^1,\dots,Y^p,X_1,\dots,X_q)s(Y^{p+1},\dots,Y^{p+p'},X_{q+1},\dots X_{q+q'})\,.$$

The second expression together with the second and fourth sum are

$$t(Y^1,\dots,Y^p,X_1,\dots,X_q)\nabla_X s(Y^{p+1},\dots,Y^{p+p'},X_{q+1},\dots X_{q+q'})\,.$$

4) In order to prove the formula

$$\begin{aligned}
t^{j_1\dots j_p}_{i_1\dots i_q;i} &= \frac{\partial}{\partial x^i}t^{j_1\dots j_p}_{i_1\dots i_q}\\
&+\sum_l \Gamma^{j_1}_{il}t^{l\dots j_p}_{i_1\dots i_q} + \dots + \sum_l \Gamma^{j_1}_{il}t^{j_1\dots l}_{i_1\dots i_q}\\
&-\sum_l \Gamma^{l}_{ii_1}t^{j_1\dots j_p}_{l\dots i_q} - \dots - \sum_l \Gamma^{l}_{ii_q}t^{j_1\dots j_p}_{i_1\dots l}\ ,
\end{aligned}$$

we assume that it is known for tensors of the form T^p_q and deduce from it the claim for tensors of the form T^{p+1}_q and T^p_{q+1}. We use also the formulas for $p=1,q=0$: (see a proposition in the text)

$$u^j_{;i} = \frac{\partial}{\partial x^i}u^j + \sum_l \Gamma^j_{il}u^l$$

and for $p=0,q=1$ (from the definition)

$$u_{j;i} = \frac{\partial}{\partial x^i}u_j - \sum_l \Gamma^l_{ji}u_l\ .$$

Since every tensor t of the form T^{p+1}_q can be written as a sum of tensors of the form $s\otimes u$, where $u\in T^1_0$ and $s\in T^p_q$, we can assume without loss of generality that $t=s\otimes u$. Since

$$t^{j_1\dots j_{p+1}}_{i_1\dots i_q} = s^{j_1\dots j_p}_{i_1\dots i_q}u^{j_{p+1}}$$

we get by Leibniz

$$t^{j_1\dots j_{p+1}}_{i_1\dots i_q;i} = s^{j_1\dots j_p}_{i_1\dots i_q;i}u^{j_{p+1}} + s^{j_1\dots j_p}_{i_1\dots i_q}u^{j_{p+1}}_{;i}\ .$$

Plugging in the induction assumption for s and the known formula for u gives the claim for T^{p+1}_q. Similarily, one can show the other induction step.

6) Assume first Ricci $= 0$. Since the scalar curvature is a trace of the Ricci tensor, we have also Scal $= 0$ and so also the Einstein tensor $G = 0$.
For the other direction, if $G = 0$, then we have (by the hint)

$$\mathrm{Ricci}_{ij} = \frac{1}{2} R g_{ij} .$$

Raising one coordinate (multiply with g^{jk} and sum over j) gives

$$\mathrm{Ricci}_i^k = \frac{1}{2} R \delta_i^k .$$

If we take the trace, we get on the left hand side the scalar curvature R and on the right hand side $\frac{1}{2}R \cdot 4 = 2R$. This equality $R = 2R$ is only possible if $R = 0$. But $R = 0$ and $G = 0$ together imply by definition of G that Ricci $= 0$.

7) If you were running the program, you obtained first the coefficients for the Christoffel symbols Γ_{ijk}, then the coefficients of the Christoffel symbols Γ_{ij}^k, then Riemann curvature tensor R_{jkl}^i and finally the Ricci tensor Ricci_{ik} which vanishes after the simplification command.

Chapter 5

Appendices

In the appendices, we add definitions and results from other topics like topology, measure theory, linear algebra, real analysis or differential equations.

Topology

Definition. A **topological space** $(X, \mathcal{O})$ is a set X and a collection $\mathcal{O}$ of subsets of X which is closed under finite intersections and arbitrary unions and such that $\emptyset, X \in \Omega$. Topological space

Notion. The sets in $\mathcal{O}$ are called **open sets.** The complement of an open set is called **closed.** Open set
For an arbitrary set A in X, denote with $\overline{A}$ its **closure**, the interection of all closed sets which contain A. Denote also with $\mathrm{int}(A)$ the **interiour** of A, that is the union of all open subsets of A. Closed set Closure of a set
Given a point $x \in X$. A set V which contains $U \in \mathcal{O}$ with $x \in U$ is also called a **neighborhood** of x. Interiour Neighborhood

Examples.
1) If d is a metric on a set X, (this means that d is a function from $X \times X \to [0, \infty)$, which satisfies $d(x,y) = d(y,x), d(x,y) = 0 \Leftrightarrow x = y$ and $d(x,z) \leq d(x,y) + d(y,z)$), then $\mathcal{O} = \{U \subset X \mid \forall x \in X, \exists r > 0, B_r(x) = \{y \mid d(x,y) < r\} \subset U\}$ is a topology. The metric space (X, d) defines so a topological space. Metric space
2) If $\mathcal{O}$ is the set of all subsets of X, then $(X, \mathcal{O})$ is called a **discrete topological space.** Discrete topological space

Definition. Let $(X, \mathcal{O}), (Y, \mathcal{P})$ be topological spaces. A map $f : X \to Y$ is called **continuous**, if $f^{-1}(U) \in \mathcal{O}$ for all $U \in \mathcal{P}$. If f is continuous, bijective (one to one and onto) and if f^{-1} is also continuous, then f is called a **homeomorphism**. A map $f : X \to \mathbf{R}$ is often called a **function**. The **support** of a function f is the closure of the set $f^{-1}(\mathbf{R} \setminus \{0\})$. Continuous map Homeomorphism Support of a function

Example. If $(X, d), (Y, \delta)$ are metric spaces, then f is continuous if and only if $d(x_n, x) \to 0$ implies $\delta(f(x_n), f(x)) \to 0$.

Definition. A topological space $(X, \mathcal{O})$ is called **Hausdorff**, if for two different points $x, y \in X$, there exist disjoint sets $U, V \in \mathcal{O}$ with $x \in U$ and $y \in V$. Hausdorff space

Example. Every metric space is a Hausdorff space because two points x, y are seperated by $B_r(x), B_r(y)$ if $r < d(x,y)/3$.

Definition. A **basis** $\mathcal{B}$ of a topological space $(X, \mathcal{O})$ is a subset of $\mathcal{O}$ such that every $U \in \mathcal{O}$ can be written as a union of elements in $\mathcal{B}$. Basis

Example. Given a metric space (X, d), then $\mathcal{B} = \{B_r(x) \mid x \in X, r > 0\}$ is a basis.

Definition. A topological space $(X, \mathcal{O})$ is called **second countable** if the topology has a countable basis.

Second countable

Example. Every metric space (X, d) such that X contains a dense set is second countable. Especially, $(\mathbf{R}^n, d(x, y) = ||x - y||)$ is second countable.

Definition. A set of sets $\mathcal{U} \subset \mathcal{O}$ is called an **open cover** if $\bigcup_{U \in \mathcal{U}} U = X$. If $\mathcal{U}$ and $\mathcal{V}$ are open covers and $\mathcal{V} \subset \mathcal{U}$, then $\mathcal{V}$ is called a **subcover** of $\mathcal{U}$. A set of subsets $\mathcal{U} \subset \mathcal{O}$ is called **locally finite**, if every point has a neighborhood U, such that $\mathcal{O} \cap U$ is a finite set of sets. Open cover Subcover Locally finite

Definition. A topological space $(X, \mathcal{O})$ is called **compact** if every open cover has a finite subcover. Compact

Examples. A metric space (X, d) is compact if every sequence x_n in X has an accumulation point in X. A subset in $\mathbf{R}^n$ is compact if and only if it is bounded and closed. If $(X_i, \mathcal{O}_i)$, for $i = 1, 2$ are compact topological spaces, then the product space is compact too.

Definition. A topological space $(X, \mathcal{O})$ is called **locally compact** if every point has a neighborhood, which is contained in a compact set. Locally compact

Example. $\mathbf{R}^n$ is locally compact but not compact.

Definition. A topological space $(X, \mathcal{O})$ is called **paracompact** if every open cover has a countable, locally finite subcover. Paracompact

Definition. A subset A of a topological space $(X, \mathcal{O})$ is called **dense** if every $U \in \mathcal{O}$ contains an element of A. Dense set

Measure theory

Definition. $\mathcal{A}$ is a σ- **algebra** if $\mathcal{A}$ is a set of subsets of Ω satisfying Sigma algebra

$$\boxed{\begin{array}{l} \Omega \in \mathcal{A}, \\ A \in \mathcal{A} \to A^c \in \mathcal{A}, \\ A_n \in \mathcal{A} \Rightarrow \bigcup_{n \in \mathbf{N}} A_n \in \mathcal{A} \end{array}}$$

The pair $(\Omega, \mathcal{A})$ is called a **measure space**.

Measure space

Definition. $\mu : \mathcal{A} \to \mathbf{R}$ is a **measure** if Measure

$$\boxed{\begin{array}{l} \mu(A) \in [0, \infty] \ \ \forall A \in \mathcal{A}, \\ A_n \in \mathcal{A} \text{ disjoint} \Rightarrow \mu(\bigcup_n A_n) = \sum_n \mu(A_n) \ (\sigma\text{-additivity}) \end{array}}$$

The measure is called **finite** if $\mu(\Omega) < \infty$. In this case, it can be normalized to be a probability measure $\mu(\Omega) = 1$. Finite measure

Example. Let Ω be a bounded closed subset of $\mathbf{R}^n$ and let $\mathcal{A}$ be the smallest σ-algebra $\mathcal{A}$ called **Borel σ-algebra** which contains all open sets in Ω. There exists a measure μ on Ω such that $\mu(\prod_{i=1}^n [a_i, b_i]) = \prod_{i=1}^d (b_i - a_i)$ for all cubes $\prod_{i=1}^n [a_i, b_i] \subset \Omega$. This measure is called the **Lebesgue measure**.

Lebesgue measure

Definition. A map f from one measure space $(\Omega, \mathcal{A})$ to an other measure space $(\Delta, \mathcal{B})$ is called **measurable** if $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$. Especially, if $\Delta \subset \mathbf{R}$ and $\mathcal{B}$ is the Borel σ-algebra on $\mathbf{R}$, then f is called a **bounded measurable map**.

Measurable map

Definition. Denote by $\mathcal{S}$ the set of measureable maps which take finitely many values: $\mathcal{S} = \{f = \sum_{i=1}^n \alpha_i \cdot 1_{A_i}$ with $\alpha_i \in \mathbf{R}\}$ Define for $f \in \mathcal{S}$ the **integral**

Integral

$$\int_\Omega f \, d\mu := \sum_{a \in f(\Omega)} a \cdot \mu\{X = a\} .$$

For a general bounded measureable map $f : \Omega \to \mathbf{R}^+$, we define the **integral**

$$\int_\Omega f \, d\mu = \sup_{g \in \mathcal{S}} \int_\Omega g \, d\mu .$$

For a general function f define $\int f = \int f^+ - \int f^-$, where $f^+(x) = \max(f(x), 0)$ and $f^-(x) = -(-f)^+(x)$.

Definition. Given two measure spaces $(\Omega_1, \mathcal{A}_1)$, $i = 1, 2$ with finite measures μ_i. Define on $\Omega \times \Omega_2$ the σ-algebra $\mathcal{A}$ as the smallest σ-algebra which contains all sets $A_1 \times A_2$ with $A_i \in \mathcal{A}_i$. There exists on $(\Omega, \mathcal{A})$ a finite measure $\mu = \mu_1 \otimes \mu_2$ which satisfies $\mu(A_1 \times A_2) = \mu_1(A_1) \times \mu_2(A_2)$. It is called the **product measure**.

Product measure

Theorem of Fubini: Let f be a bounded measureable map on $(\Omega = \Omega_1 \times \Omega_2, \mathcal{A}_1 \times \mathcal{A}_2)$. Then

$$\int_\Omega f \, d\mu = \int_{\Omega_1} f(x, y) \, d\mu_1(x) \, d\mu_2(y)$$

Linear algebra

Definition. Let E be a finite dimensional vector space and let E^* be the dual space. The **scalar product** between a vector $v \in E^*$ and a vector $w \in E$ is a map from $E^* \times E \to \mathbf{R}$ defined as $v \cdot w = v(w)$.

Scalar product

Property. In coordinates, the scalar product is given as follows. If $v = \sum v_i e^i$, where $v_i = v(e_i)$ and $w = \sum_j w^j e_j$ with $w^j = e^j(w)$, then

$$v \cdot w = v(w) = \sum_{i,j} v_i w^j e^i(e_j) = \sum_j v_j w^j .$$

Definition. Given a basis $\{e_i\}$ in E and the dual basis $e^i \in E^*$. A linear map A defines a **matrix** $A_i^j = e^j \cdot Ae_i$, where A_j^i is the entry in the i'th row and j'th column.

Matrix

Property. In coordinates,

$$(Av)^j = \sum_i A_i^j v^i ,$$

because $e^j \cdot Ae_i = A_i^j$ and by linearity $e^j \cdot Av = e^j \cdot A \sum_i v^i e_i = \sum_i A_i^j v^i$. We have also

$$Ae_i = \sum_{j=1}^n A_i^j e_j ,$$

because for each e^k, the left hand side is $e^k \cdot Ae_i = A_i^k$, while the right hand side is $e^k \cdot \sum_{j=1}^n A_i^j e_j = \sum_{j=1}^n A_i^j e^k \cdot e_j = A_i^k$.

Definition. If $A : E \to E$ is a linear map, then the **adjoint map** $A^T : E^* \to E^*$ is defined by the relation $v \cdot Aw = A^T v \cdot w$. Adjoint map

Property. This means $(A^T v)_j = A^T v \cdot e_j = v \cdot Ae_j = \sum_i v_i e^i \cdot Ae_j = \sum_i A_j^i v_i$.

Lemma 5.0.1 (Transformation of the dual basis) *Let e_i be a basis in E. Let $\tilde{e}_i = Ae_i$ be a new basis, where A is an invertible map. Then $(A^T)^{-1}e^j$ is the dual basis of the basis $\tilde{e}_j$.*

Proof. Let $B : E^* \to E^*$ be the linear map in E^* such that $Be^j = \tilde{e}^j$ is the dual basis. We have

$$\delta_i^j = \tilde{e}^j \cdot \tilde{e}_i = Be^j \cdot Ae_i = A^T Be^j \cdot e_i \, .$$

Therefore $A^T Be^j = e^j$ for all j so that $A^T B$ is the identity and $B = (A^T)^{-1}$. □

Using this lemma, coordinate transformations of any tensor can be computed. Let us solve Problem 2a) of Week 3:

We have seen that if the basis transforms as $\tilde{e}_i = Ae_i$, then the dual basis transforms as $\tilde{e}^k = B^T e^k = \sum_j B_j^k e^j$, where $B = A^{-1}$.

$$\tilde{f}^k = f(\tilde{e}^k) = f(B^T e^k) = f(\sum_i B_i^k e^i) = \sum_i B_i^k f(e^i) = \sum_i B_i^k f^i \, .$$

Remark. We compute here the tensor in a new basis, that is we move the basis, which corresponds to a **pull-back** of the tensor: Pull-back of a tensor

$$A^* f(e^1, \dots, e^p, e_1, \dots, e_q) = f(A^{-1}e^1, \dots, A^{-1}e^q, Ae_1, \dots, Ae_q) \, .$$

This is different from **push-forward** the tensor with a map $A : E \to E$ which is defined as Push-forward

$$A_* f(e^1, \dots e^p, e_1, \dots, e_q) = f(A^T e^1, \dots, A^T e^q, A^{-1}e_1, \dots, A^{-1}e_q) \, .$$

Definition. Let E be a n dimensional vector space and $M(n, \mathbf{R})$ the algebra of $n \times n$ matrices. For $A \in M(n, \mathbf{R})$, the **determinant** is defined as Determinant

$$\det(A) = \sum_{\sigma \in S_n} (-1)^\sigma A_{1,\sigma(i)} \, \cdots \, A_{n,\sigma(n)} \, ,$$

where σ runs over all $n!$ permutations in the group S_n of all permutations of n elements and where $(-1)^\sigma$ is the sign of the permutation. (We wrote here A_{ij} instead of A_j^i).

Properties of the determinant are: $\det(AB) = \det(A)\det(B)$, which implies for invertible A, $\det(A^{-1}) = \det(A)^{-1}$ $\det(1) = 1$. If λ_i are the eigenvalues of A, then $\det(A) = \prod_{i=1}^n \lambda_i$. If $\det(A) \neq 0$, then A is invertible and $A^{-1} = Ad(A)(\det A)^{-1}$, where $Ad(A)_{ij}^T = (-1)^{i+j} A^{\hat{i}\hat{j}}$, where $A^{\hat{i}\hat{j}}$ is the matrix obtained from A by deleting the i'th row and j'th column. This formula for the inverse is equivalent to the development $\det(A)\delta_{jk} = \sum_{i=1}^n Ad_{ji} A_{ik}$, where j is arbitrary.

Definition. Let $A : E \to F$ be a linear map from a vector space E to a vector space F. The **kernel** of A is the set $\{v \in E \mid Av = 0\}$. The **image** or **range** of A is the set $\{Av \in F \mid v \in E\}$.

Real analysis

Definition. Given a map $f : \mathbf{R}^n \to \mathbf{R}^m$. Denote by $Df(x)$ the linear map $Df(x)h = \lim_{\epsilon \to 0} f(x + \epsilon h) - f(x))/\epsilon$ from $\mathbf{R}^n$ to $\mathbf{R}^m$ if the limit exists for each h. It is called the **derivative**.

Derivative

Notation. Given a basis in $\mathbf{R}^n$ and $\mathbf{R}^m$, we write $f(x) = f(x_1, \dots, x_n) = (f_1(x), \dots f_m(x))$. Then Df, the **Jacobian** of f is a $m \times n$ matrix defined by

Jacobian

$$\begin{pmatrix} \partial_1 f_1 & \dots & \partial_n f_1 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \partial_1 f_m & \dots & \partial_n f_n \end{pmatrix} ,$$

where we wrote ∂_i as a short hand for $\partial/(\partial x_i)$.

Important properties:

> The **chain rule**: $D(f \circ g)(x) = Df(g(x))Dg(x)$.

Chain rule

> The **inverse function theorem** if $Df(x)$ is an isomorphism, then f is a local diffeomorphism at x: there exists a neighborhood U of x, such that $f : U \to f(U)$ is a diffeomorphism.

Inverse function theorem

Definition. A map $f : U \subset \mathbf{R}^m \to \mathbf{R}^n$ is called **differentiable** in U, if $Df(x)$ exists for all $x \in U$. Denote by Df the map $U \times \mathbf{R}^m \to \mathbf{R}^n$ given by $Df(x, u) = Df(x)u$. If this map is continuous, we say it is **continuously differentiable** and we write $f \in C^1(U, \mathbf{R}^n)$. If the map $Df : U \times \mathbf{R}^m, (x, u) \to Df(x)u$ is continuously differentiable, then f is defined to be in $C^2(U, \mathbf{R}^n)$. Inductively if $D_x^{k-1} f : U \times (\mathbf{R}^m)^k \to \mathbf{R}^n$ is continuously differentiable, then we write $f \in C^k(U, \mathbf{R}^n)$. If $f \in C^k(U, \mathbf{R}^n)$ for all k, then f is called **smooth**.

Differentiable map

The k**'th derivative** of a function $f : \mathbf{R}^m \to \mathbf{R}^n$ is defined as follows. For vectors $u^{(i)} = (u_1^{(i)}, \dots, u_m^{(i)}) \in \mathbf{R}^m$ and $f(x) = (f_1(x), \dots, f_n(x))$,

Smooth map

$$\boxed{D^k f(x)(u^{(1)}, \dots, u^{(k)})_l = \sum_{i_1, i_2, \dots, i_k = 1}^{m} \frac{\partial}{\partial x_{i_1}} \frac{\partial}{\partial x_{i_2}} \cdots \frac{\partial}{\partial x_{i_k}} f_l(x) u_{i_1}^{(1)} \dots u_{i_k}^{(k)} \; , l = 1, \dots, n \; .}$$

Note that $D^k f(x)$ is linear in each of the coordinates $u^{(j)}$ and so a **multilinear map** from $(\mathbf{R}^m)^k = \mathbf{R}^m \times \mathbf{R}^m \dots \times \mathbf{R}^m \to \mathbf{R}^n$. Sometimes, you will encounter the notation $D_i = \frac{\partial}{\partial x_i}$.

k-th derivative

Example. Given a map $f : \mathbf{R}^m \to \mathbf{R}$, then $Df(x) = (\partial_1 f, \dots, \partial_m f)$ is the **gradient** and

$$D^2 f(x) = \begin{pmatrix} \partial_1 \partial_1 f & \dots & \partial_m \partial_1 f \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \partial_1 \partial_n f & \dots & \partial_m \partial_n f \end{pmatrix}$$

in the sense that $D^2 f(x)(u, v) = (u^T, D^2 f(x) v)$, where $(\cdot, \cdot)$ is the scalar product (here $\mathbf{R}^m$ is identified with its dual space $(\mathbf{R}^m)^*$ by $u \mapsto u^T$. See the linear algebra appendix for the notion of the scalar product)

If $f : \mathbf{R}^m \to \mathbf{R}^n$ is k times differentiable at x, then, for small enough $h \in \mathbf{R}^m$, **Taylor's theorem** holds:

Multilinear map

$$\boxed{f(x+h) = f(x) + Df(x)h + \frac{D^2f(x)}{2!}(h,h) + \ldots + \frac{D^kf(x)}{k!}(h,\ldots,h) + R(x,h)(h,h,\ldots,h)\ ,}$$

where $R(x,u)$ is an rest term which satisfies $R(x,0) = 0$ and can explicitly be given as

$$R(x,u) = \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!}(D^kf(x+tu) - D^kf(x))\,dt\ .$$

Remark. The notion of derivative is in some books done as follows: denote by $L^1(\mathbf{R}^m, \mathbf{R}^n)$ the set of all linear maps from $\mathbf{R}^m \to \mathbf{R}^n$. Define inductively $L^k(\mathbf{R}^m, \mathbf{R}^n) = L^1(\mathbf{R}^m, L^{k-1}(\mathbf{R}^m, \mathbf{R}^n)$, which is the space of all multilinear maps from $\mathbf{R}^m \times \ldots \times \mathbf{R}^m \to \mathbf{R}^n$. Given a function $f : U \subset \mathbf{R}^m \to \mathbf{R}^n$, then D^kf is a function from U to $L^k(\mathbf{R}^m, \mathbf{R}^n)$ defined inductively by $D^kf = D(D^{k-1}f)$.

Differetial equations

Taylor's theorem

Definition. Let (X,d) be a metric space. A map $\phi : X \to X$ is called a **contraction**, if there exists $\lambda < 1$ such that $d(\phi(x), \phi(y)) \leq \lambda \cdot d(x,y)$ for all $x, y \in X$. The constant λ is called a **contractivity factor** of ϕ.

[Banach's fixed point theorem]
A contraction ϕ in a complete metric space has exactly one fixed point in X.

Proof. (i) By induction, one shows that for any pair $x, y \in X$

$$d(\phi^n(x), \phi^n(y)) \leq \lambda^n \cdot d(x,y)\ .$$

(ii) Using the triangle inequality, we get for all $x \in X$,

$$d(x, \phi^n x) \leq \sum_{k=0}^{n-1} d(\phi^k x, \phi^{k+1} x) \leq \sum_{k=0}^{n-1} \lambda^k d(x, \phi(x)) \leq \frac{1}{1-\lambda} \cdot d(x_0, x_1)\ .$$

(iii) For all $x \in X$, $x_n = \phi^n(x)$ is a Cauchy sequence and has by completeness of X a limit $\tilde{x}$. Indeed, by (i),(ii),

$$d(x_n, x_{n+k}) \leq \lambda^n \cdot d(x, x_k) \leq \lambda^n \cdot \frac{1}{1-\lambda} \cdot d(x, x_1)\ .$$

(iv) The limit $\tilde{x}$ does not depend on x: assume there are two limits $\tilde{x}, \tilde{y}$. They are both invariant under ϕ and satisfy

$$d(\tilde{x}, \tilde{y}) = d(\phi(\tilde{x}), \phi(\tilde{y})) \leq \lambda \cdot d(\tilde{x}, \tilde{y})\ .$$

This is only possible if $\tilde{x} = \tilde{y}$. □

Contraction

Notation. Let $\mathcal{X}$ be a Banach space (a complete normed vector space and especially a complete metric space with metric $d(x,y) = ||x-y||$) and I an interval in $\mathbf{R}$. Denote by $B_r(x)$ the ball $\{y \in \mathcal{X} \mid ||x-y|| < r\}$ around x of radius r in the space $\mathcal{X}$.

Banach's fixed point theorem

Let $X = \overline{B_r(x_0)} \subset \mathcal{X}$ and assume ϕ is a differentiable map $\mathcal{X} \to \mathcal{X}$. If for all $x \in X$, $||D\phi(x)|| \leq |\lambda| < 1$ and

$$||\phi(x_0) - x_0|| \leq (1-\lambda) \cdot r$$

then ϕ has exactly one fixed point in X.

Proof. The condition $||x - x_0|| < r$ implies that

$$||\phi(x) - x_0|| \leq ||\phi(x) - \phi(x_0)|| + ||\phi(x_0) - x_0|| \leq \lambda r + (1-\lambda) r = r \, .$$

The map ϕ maps therefore the ball X into itself. Banach's fixed point theorem applied to the complete metric space X and the contraction ϕ implies the result. □

Banch fixed point theorem

Definition. Let f be a mapping $I \times \mathcal{X} \to \mathcal{X}$. A differentiable mapping $u : \mathcal{J} \to \mathcal{X}$ of an open ball $J \subset I$ in $\mathcal{X}$ is called a **solution of the differential equation**

$$\dot{x} = f(t, x)$$

if we have for all $t \in J$ the relation

$$\dot{u}(t) = f(t, u(t)) \, .$$

[Cauchy-Picard Existence theorem]
Let $f : I \times \mathcal{X} \to \mathcal{X}$ be continuous in the first coordinate and differentiable in the second.
For every $(t_0, x_0) \in I \times \mathcal{X}$, there exists an open interval $J \subset I$ with midpoint t_0, such that on J, there exists exactly one solution of the differential equation $\dot{x} = f(t, x)$.

Proof. There exists an interval $J(t_0, a) = (t_0 - a, t_0 + a) \subset I$ and a ball $B(x_0, b)$, such that

$$M = \sup\{||f(t,x)|| \mid (t,x) \in J(t_0,a) \times B(x_0,b)\}$$

as well as

$$k = \sup\{\frac{||f(t,x_1) - f(t,x_2)||}{||x_1 - x_2||} \mid (t,x_1),(t,x_2) \in J(t_0,a) \times B(x_0,b), x_1 \neq x_2\}$$

are finite. Define for $r < a$ the Banach space

$$\mathcal{X}_r = C(\overline{J}(t_0,r),\mathcal{X}) = \{y : \overline{J}(t_0,r) \to \mathcal{X}, \ y \text{ continuous}\}$$

with norm

$$||y|| = \sup_{t \in \overline{J}(t_0,r)} ||y(t)||$$

Let $V_{r,b}$ be the open ball in $\mathcal{X}_r$ with radius b around the constant mapping $t \mapsto x_0$. For every $y \in V_{r,b}$ we define

$$\phi(y) : t \mapsto x_0 + \int_{t_0}^t f(s, y(s)) ds$$

which is again an element in $\mathcal{X}_r$. We prove now, that for r small enough, ϕ is a contraction. A fixed point of ϕ is then a solution of the differential equation $\dot{x} = f(t,x)$, which exists on $J = J_r(t_0)$. For two points $y_1, y_2 \in V_r$, we have by assumption

$$||f(s,y_1(s)) - f(s,y_2(s))|| \leq k \cdot ||y_1(s) - y_2(s)|| \leq k \cdot ||y_1 - y_2||$$

for every $s \in \overline{J}_r$. Thus, we have

$$\begin{aligned} ||\phi(y_1) - \phi(y_2)|| &= ||\int_{t_0}^{t} f(s, y_1(s)) - f(s, y_2(s))\, ds|| \\ &\leq \int_{t_0}^{t} ||f(s, y_1(s)) - f(s, y_2(s))||\, ds \\ &\leq kr \cdot ||y_1 - y_2|| \,. \end{aligned}$$

On the other hand, we have for every $s \in \overline{J}_r$

$$||f(s, y(s))|| \leq M$$

and so

$$||\phi(x_0) - x_0|| = ||\int_{t_0}^{t} f(s, x_0(s))\, ds|| \leq \int_{t_0}^{t} ||f(s, x_0(s))||\, ds \leq M \cdot r \,.$$

We can apply the above lemma, if $kr < 1$ and $Mr < b(1 - kr)$. This is the case, if $r < b/(M + kb)$. By choosing r small enough, we can get the contraction rate as small as we wish. $\square$

Bibliography

[1] R. Abraham, J.E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis and Applications.* Applied Mathematical Sciences, 75. Springer Verlag, New York etc., second edition, 1988.

[2] R. Bott and L. W. Tu. *Differential forms in algebraic topology.* Graduate texts in mathematics. Springer Verlag, 1991.

[3] I. Chavel. *Eigenvalues in Riemannian Geometry.* Pure and applied mathematics. Academic Press Inc., Orlando, 1984.

[4] I. Ciufolini and J.A. Wheeler. *Gravitation and inertia.* Princeton Series in Physics, 1995.

[5] B.A. Dubrovin, A.T.Fomenko, and S.P.Novikov. *Modern Geometry-Methods and Applications Part I,II,III.* Graduate Texts in Mathematics. Springer Verlag, New York, 1985.

[6] A. Fomenko. *Visual Geometry and Topology.* Springer-Verlag, Berlin, 1993.

[7] S. Gallot, D.Hulin, and J.Lafontaine. *Riemannian geometry (second edition).* Springer-Verlag, Berlin, 1990.

[8] V. Guillemin and A. Pollack. *Differential topology.* Prentice-Hall, Inc., New Jersey, 1974.

[9] M.W. Hirsch. *Differential topology.* Graduate texts in mathematics. Springer-Verlag, Berlin, 1976.

[10] J.D. Nightingale J. Foster. *A short course in general relativity.* Springer-Verlag, 1995.

[11] S. Lang. *Differential and Riemannian Manifolds.* Graduate texts in mathematics. Springer-Verlag, Berlin, third edition edition, 1995.

[12] J. Milnor. *Morse theory*, volume 51 of *Annals of Mathematics Studies.* Princeton University press, Princeton, New Jersey, 1963.

[13] J. Milnor. *Topology from the differential viewpoint.* University of Virginia Press, Charlottesville, Va, 1965.

[14] C.W. Misner, K.S. Thorne, and J.A. Wheeler. *Gravitation.* Freeman, 1973.

[15] M. Spivak. *Calculus on Manifolds.* Addison-Wesley publishing company, 1965.

[16] F. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate texts in mathematics.* Springer, New York, 1983.