A. MANIFOLDS

This course is about manifolds. An m-manifold is a space that looks locally like Euclidean space $\mathbb{R}^m$.

Examples of manifolds include a circle $\mathbb{S}^1$, a sphere $\mathbb{S}^2$ or a torus $T^2 = \mathbb{S}^1 \times \mathbb{S}^1$ or any surface in $\mathbb{R}^3$. These generalize for instance to an m-sphere $\mathbb{S}^m \subset \mathbb{R}^{m+1}$ or an m-torus $T^m = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$.

In multivariable calculus, one studies (smooth) m-submanifolds of $\mathbb{R}^n$; these have several equivalent characterizations (locally being level sets or images of smooth functions). These are the motivating examples of (smooth) manifolds. Indeed, we will see any manifold can be embedded as a submanifold of some (high-dimensional) $\mathbb{R}^n$. But it is important to give an abstract definition of manifolds, since they usually don't arise as submanifolds.

The idea of manifolds is that they are spaces on which one can do analysis (derivatives, integrals, etc.). This means we are talking here not simply about topological manifolds, but about smooth (differentiable) manifolds. This is a distinction we explain soon.

A1. Topological manifolds

Definition A1.1. A *topology* on a set X is a collection $\mathcal{U}$ of subsets of X (called the *open* subsets) such that

- $\varnothing, X \in \mathcal{U}$,
- $U, V \in \mathcal{U} \implies U \cap V \in \mathcal{U}$,
- $\{U_\alpha\} \subset \mathcal{U} \implies \bigcup_\alpha U_\alpha \in \mathcal{U}$.

All other topological notions are defined in terms of open sets.

Definition A1.2. A subset $A \subset X$ is *closed* if its complement $X \smallsetminus A$ is open. Any subset $Y \subset X$ naturally becomes a topological space with the *subspace topology*: $\{U \cup Y : U \subset X \text{ open}\}$. That is, the open sets in Y are exactly the interestions of Y with open sets in X.

Definition A1.3. A space X is *connected* if $\varnothing$ and X are the only subsets that are both open and closed. A space is *compact* if every open cover has a finite subcover. (If we talk about a subset $Y \subset X$ being connected or compact, etc., we mean with respect to the subspace topology.)

Definition A1.4. A map $f\colon X \to Y$ between topological spaces is *continuous* if the preimage of any open set in Y is open in X. A continuous bijection $f\colon X \to Y$ whose inverse is also continuous is called a *homeomorphism* – an equivalence of topological spaces.

Definition A1.5. Usually, one specifies a topology not by listing all open sets, but by giving a *base* $\mathcal{B}$. This is a collection of "basic" open sets sufficient to generate the topology: an arbitrary $U \subset X$ is defined to be open if and only if it is a union of sets from $\mathcal{B}$. The requirements on $\mathcal{B}$ to form a base are (1) that $\mathcal{B}$ covers X and (2) that intersections $B_1 \cap B_2$ of two basic open sets are open, that is, for any $x \in B_1 \cap B_2$, there exists $B_3 \in \mathcal{B}$ with $x \in B_3 \subset B_1 \cap B_2$. These conditions are familiar from metric spaces (X, d), where the open balls $B_\varepsilon(x) := \{y : d(x, y) < \varepsilon\}$ form a base for the metric topology. The topological spaces that will arise for us are all metrizable, meaning the topology arises from some metric. In particular, the standard topology on $\mathbb{R}^n$ comes, of course, from the Euclidean inner product (scalar product) $\langle x, y\rangle = x \cdot y = \sum x^i y^i$ via the metric $d(x, y) = |x - y| = \sqrt{\langle x - y, x - y\rangle}$.

Definition A1.6. A space X is *Hausdorff* if any two distinct points $x \neq y \in X$ have disjoint (open) neighborhoods. It is *regular* if given a nonempty closed set $A \subset X$ and a point $x \in X \smallsetminus A$, there are disjoint (open) neighborhoods of A and x. (These are just two examples of the many "separation axioms" in point-set topology.) A space X is *second countable* if there is a countable base for the topology.

Metric spaces are Hausdorff and regular (take metric neighborhoods of radius $d(x, A)/2$). Eucldiean space is second countable (take balls with rational centers and radii). The importance of these notions is clear from the Urysohn metrization theorem, which says that X is separable (that is, has a countable dense subset) and metrizable if and only if it is Hausdorff, regular and second countable.

Definition A1.7. We say a space M is *locally homeomorphic* to $\mathbb{R}^m$ if each $p \in M$ has an open neighborhood U that is homeomorphic to some open subset of $\mathbb{R}^m$. If $\varphi\colon U \to \varphi(U) \subset \mathbb{R}^m$ denotes such a homeomorphism, then we call (U, φ) a *(coordinate) chart* for M. An *atlas* for M is a collection $\{(U_\alpha, \varphi_\alpha)\}$ of coordinate charts which covers M, in the sense that $\bigcup U_\alpha = M$.

Clearly we can rephrase the definition to say M is locally homeomorphic to $\mathbb{R}^m$ if and only if it has an atlas of charts. Less obvious (but an easy exercise) is that it is equivalent to require each $p \in M$ to have a neighborhood homeomorphic to $\mathbb{R}^m$.

Although one might expect this to be a good topological definition of an abstract manifold, it turns out that there are some pathological examples that we would like to rule out. Certain properties from point-set topology are not automatically inherited. For instance, examples like the line $\mathbb{R}$ with the origin doubled (or with all $x \geq 0$ doubled) fail to be Hausdorff. The "long line" (obtained by gluing uncountably many unit intervals) fails to be second countable – it is sequentially compact but not compact. The "Prüfer surface" is separable but not second countable. (Note there are also much weirder examples, for instance in papers of Alexandre Gabard.) For technical reasons, we also prefer manifolds to have at most countably many components, as is guaranteed by second countability.

Thus we are led to the following:

Definition A1.8. A *topological m-manifold* is a second-countable Hausdorff space $M = M^m$ that is locally homeomorphic to $\mathbb{R}^m$.

Regularity then follows, ensuring that our manifolds are metrizable spaces. Indeed, we will later put a (Riemannian) metric on any (smooth) manifold. It is also straightforward to check various other local properties: A topological manifold M is locally connected, locally compact, normal and paracompact (defined later when we need it). Being separable and locally compact, it is also globally the union of countably many compact subsets.

Note: For $m \neq n$, it is easy to see there is no diffeomorphism $\mathbb{R}^m \to \mathbb{R}^n$ It is also true that there is no homeomorphism, but this requires the tools of algebraic topology like homology theory. Any $\mathbb{R}^m$ is contractible, so they all have the same (trivial) homology. The trick is to first remove a point. Then $\mathbb{R}^m \smallsetminus \{0\}$ contracts to $\mathbb{S}^{m-1}$, and spheres of different dimension have different homology. This was the start of topological dimension theory, and shows that every (nonempty) manifold has a well-defined dimension.

Examples A1.9.

- $\mathbb{R}^m$ is an m-manifold (with a single chart).
- An open subset $U \subset M^m$ of an m-manifold is itself an m-manifold (restricting charts to U).
- Any smooth surface $M^2 \subset \mathbb{R}^3$ is a 2-manifold. (Get a chart around $p \in M$ by projecting orthogonally to T_pM.)
- Other surfaces – like polyhedral surfaces – are also topological manifolds.
- More generally, any smooth m-submanifold in $\mathbb{R}^n$ is an m-manifold. (We will consider such examples in general later.)
- $\mathbb{R}P^m := \mathbb{S}^m/\pm = \{\text{lines through 0 in } \mathbb{R}^{m+1}\}$ is an m-manifold called real projective space.
- $M^m \times N^n$ is an $(m+n)$-manifold (using product charts).
- For any smooth surface $M^2 \subset \mathbb{R}^3$, the tangent bundle $TM = \{(p, v) : p \in M, v \in T_pM\}$ is a 4-manifold

One of our first tasks will be to define T_pM for an abstract smooth manifold M^m; in general we will find that it is a m-dimensional vector space and that these can be put together to form a $2m$-manifold, the tangent bundle.

In some cases it is important to consider also manifolds with boundary, modeled on the halfspace

$$H^m := \{(x^1, x^2, \dots, x^m) \in \mathbb{R}^m : x^1 \leq 0\},$$

whose boundary is $\partial H^m = \{x^1 = 0\} \cong \mathbb{R}^{m-1}$.

Definition A1.10. An *m-manifold with boundary* is then a second-countable Hausdorff space locally homeomorphic to H^m. If a point $p \in M$ is mapped to the boundary in one chart, then this is true in every chart. Such points form the *boundary* $\partial M \subset M$ of M; it is an $(m-1)$-manifold (without boundary, and perhaps empty). The complement $M \smallsetminus \partial M$ is called the *interior* and is an m-manifold (without boundary).

We will use manifolds with boundary later when we study integration and Stokes' theorem. Until then, we will basically neglect them, with the understanding that all our theory extends in the "obvious" way. The following terminology is standard even if confusing at first: a *closed manifold* is a compact manifold without boundary.

Let us look at the lowest dimensions. A 0-manifold is a countable discrete set. Equivalently, we can say a connected 0-manifold is a point. In general, of course, any manifold is a countable union of connected components, so it makes sense to classify connected manifolds. It is not hard to show that any connected 1-manifold is (homeomorphic to) either $\mathbb{R}^1$ or $\mathbb{S}^1$. If we allow manifolds with boundary, there are just two more examples: the compact interval, and the ray or half-open interval.

While the complete classification of noncompact surfaces is known, we will only consider the compact case. We have seen the examples $\mathbb{S}^2$, T^2 and $\mathbb{R}P^2$. It turns out that any connected closed surface is a connect sum of these: either an orientable surface Σ_g of genus g or a nonorientable surface N_h. (And if we allow compact surfaces with boundary, then all we get are these examples with some number k of open disks removed, denoted $\Sigma_{g,k}$ or $N_{h,k}$.)

The uniformization theorem, a classical result in complex analysis and Riemann surface theory, implies that any surface admits a metric of constant Gauss curvature. Indeed, the sphere and the projective plane have spherical ($K \equiv 1$) metrics, the torus and Klein bottle have euclidean (flat, $K \equiv 0$) metrics, and all other closed surfaces have hyperbolic ($K \equiv -1$) metrics.

Guided by this, Bill Thurston conjectured a method to understand compact 3-manifolds (with boundary). In 2003, Grigory Perelman proved this "geometrization conjecture", establishing that any 3-manifold can be cut into pieces, each of which admits one of eight standard geometries. There is interesting work remaining to be done to better understand the case of hyperbolic 3-manifolds.

In dimensions four and higher, there is in some sense no hope of classifying manifolds. Given any finite group presentation, one can build a closed 4-manifold with that fundamental group. Since the word problem is known to be undecidable, it is impossible in general to decide whether 4-manifolds are homeomorphic. Much interesting research thus restricts attention to the case of simply connected manifolds (with trivial fundamental group).

In certain other ways, higher dimensions are easier to understand. One reason is that two generic 2-disks will have empty intersection in dimensions five and above. Thus, for instance, the Poincaré conjecture was first proved in these dimensions.

A2. Smooth structures

If (U, φ) and (V, ψ) are two charts for a manifold M^m, then

$$\psi \circ \varphi^{-1} \colon \varphi(U \cap V) \to \psi(U \cap V)$$

is a homeomorphism between open sets in $\mathbb{R}^m$, called a *change of coordinates* or *transition function*. The inverse homeomorphism is of course $\varphi \circ \psi^{-1}$.

Since the transition functions are maps between Euclidean spaces, we know how to test how smooth they are. Suppose $U \subset \mathbb{R}^m$ is open and $f \colon U \to \mathbb{R}^n$. To say f is C^0 just means that it is continous. If f is differentiable at each

$p \in U$, then its derivative is a function $Df: U \to \mathbb{R}^{n\times m}$. We say f is C^1 if Df is continuous. By induction, we say f is C^r if Df is C^{r-1}, that is, if $D^r f$ is continuous. If f has (continuous) derivatives of all orders, we say it is C^∞. If a C^∞ map f is real analytic, meaning that its Taylor series around any $p \in U$ converges to f, then we say f is C^ω.

Definition A2.1. Fix $r \in \{0, 1, 2, \dots, \infty, \omega\}$. We say two charts (U, φ) and (V, ψ) for a manifold M^m are C^r-*compatible* if the transition functions $\psi \circ \varphi^{-1}$ and $\varphi \circ \psi^{-1}$ are C^r maps. (They are then inverse C^r diffeomorphisms.) A C^r*-atlas* for M is a collection of C^r-compatible coordinate charts which covers M. A C^r*-structure* on M is a *maximal* C^r-atlas, that is an atlas $\mathcal{U} = \{(U_\alpha, \varphi_\alpha)\}$ such that any coordinate chart (V, ψ) which is compatible with all the $(U_\alpha, \varphi_\alpha)$ is already contained in $\mathcal{U}$. A C^r*-manifold* is a topological manifold M^m with a choice of C^r-structure. A *chart* for a smooth manifold will mean a chart in the given smooth structure (unless we explicitly refer to a "topological chart").

Of course the case $r = 0$ is trivial: any atlas is C^0 and the C^0-structure is the set of all possible topological charts. (In this case, of course, one should use the term "homeomorphism" instead of "C^0-diffeomorphism".)

This course is about smooth manifolds, where we use the word "smooth" to mean C^∞. When we say "manifold" we will mean smooth manifold unless we explicitly say otherwise. Of course many of our results will be valid even for lower degrees of smoothness (usually C^1 or C^2 or C^3 would suffice) but we will not attempt to keep track of this.

Lemma A2.2. *Any C^r-atlas is contained in a unique maximal one.*

Proof. Given an atlas $\mathcal{U} = \{(U_\alpha, \varphi_\alpha)\}$, let $\mathcal{V}$ be the collection of all charts (V, ψ) that are compatible with every $(U_\alpha, \varphi_\alpha)$. We just need to show that $\mathcal{V}$ is a C^r-atlas, that is that any charts (V_1, ψ_1) and (V_2, ψ_2) are compatible with each other. But any $p \in V_1 \cap V_2$ is contained in some U_α, and on $\psi_1(V_1 \cap V_2 \cap U_\alpha)$ we can write

$$\psi_2 \circ \psi_1^{-1} = (\psi_2 \circ \varphi_\alpha^{-1}) \circ (\varphi_\alpha \circ \psi_1^{-1}). \quad \square$$

At the end of this proof, we implicitly use three properties:

- The composition of two C^r functions is C^r.
- The restriction of a C^r function to an open subset is C^r.
- A map that is C^r is some neighborhood of each point in U is C^r on U.

Without getting into formal details, these properties mean that the class of C^r diffeomorphisms form a *pseudogroup* (of homeomorphisms on the topological space $\mathbb{R}^m$).

Although we have only defined C^r structures, other kinds of structures on manifolds arise from other pseudogroups. For instance, a *projective* or *Möbius* structure on M arises from an atlas where the transtion functions are all projective or Möbius transformations (respectively). An *orientation* on M arises from an atlas where all transition functions are orientation-preserving. (We will return to this later. Note that it is easy to tell if a diffeomorphism is orientation-preserving; for C^0 manifolds one needs homology theory.)

Given a C^r-structure on any manifold, for any $s \le r$, by the lemma it extends to a unique C^s-structure. On the manifold $\mathbb{R}^m$, the standard C^r structure arises from the atlas $\{(\mathbb{R}^m, \mathrm{id})\}$ consisting of a single chart. If we let $\mathcal{U}^r$ denote the collection of all charts C^r-compatible with this one, then we have

$$\mathcal{U}^0 \supset \mathcal{U}^1 \supset \mathcal{U}^2 \supset \cdots \supset \mathcal{U}^\infty \supset \mathcal{U}^\omega.$$

`End of Lecture 12 Oct 2015`

The point of a smooth structure is to know which mappings are smooth. Suppose $f: M^m \to N^n$ is a (continuous) map between two smooth manifolds. Given $p \in M$, we can find (smooth) charts (U, φ) around $p \in M$ and (V, ψ) around $f(p) \in N$. Then the composition

$$\psi \circ f \circ \varphi^{-1}: \varphi(U \cap f^{-1}(V)) \to \mathbb{R}^n$$

is called the expression of f in these coordinates. (Writing $(x^1, \dots, x^m)$ for a typical point in $\varphi(U) \subset \mathbb{R}^m$ and $(y^1, \dots, y^n)$ for a typical point in $\psi(V) \subset \mathbb{R}^n$, then we can think of $\psi \circ f \circ \varphi^{-1}$ very explicitly as n real-valued functions, giving the y^j as a function of $(x^1, \dots, x^m)$.)

Now it is easy to define smoothness:

Definition A2.3. The map f is *smooth* if for each p we can find charts (U, φ) and (V, ψ) as above such that $\psi \circ f \circ \varphi^{-1}$, the expression of f in these coordinates, is smooth (as a map between euclidean spaces). (If M and N are only C^r-manifolds, then it makes sense to ask if $f: M \to N$ is C^s for $s \le r$ but not for $s > r$.) A *diffeomorphism* $f: M \to N$ between two smooth manifolds is a homeomorphism such that both f and f^{-1} are smooth. The set of all smooth maps $M \to N$ is denoted by $C^\infty(M, N)$; we write $C^\infty(M) := C^\infty(M, \mathbb{R})$.

Exercise A2.4. If $f: M^m \to N^n$ is smooth, then its expression $\psi \circ f \circ \varphi^{-1}$ in *any* (smooth) coordinate charts is smooth.

Note two special cases: if $M = \mathbb{R}^m$ (of course with the standard smooth structure), then we can take $\varphi = \mathrm{id}$ and thus consider $\psi \circ f$; if $N = \mathbb{R}^n$ then we can take $\psi = \mathrm{id}$ and consider $f \circ \varphi^{-1}$. For a map $f: \mathbb{R}^m \to \mathbb{R}^n$, we take $\varphi = \mathrm{id}$ and $\psi = \mathrm{id}$ and see that our new definition of smoothness agrees with the one we started with for maps between euclidean spaces.

The basic constructions of new manifolds from old – open subsets and products – can be adapted to the smooth setting.

If $U \subset M^m$ is an open subset of a smooth manifold, then we can restrict the smooth structure on M to a smooth structure on U (which we already know is a topological manifold). In particular, each chart (V, ψ) for M gives a chart $(V \cap U, \psi|_{V\cap U})$ for U. (If we talk about a smooth map on an open subset U of a smooth manifold M, then we implicitly mean smooth with respect to this restricted structure.)

Suppose we have a cover $\{U_\alpha\}$ of M and a map $f\colon M \to N$. Then f is smooth if and only if its restriction to each U_α is smooth. (This is a version of the pseudogroup property above.)

If M^m and N^n are smooth manifolds, then it is a straightforward exercise to put a smooth structure on the manifold $M \times N$. (Hint: Use only product charts $(U \times V, \varphi \times \psi)$ obtained from *smooth* charts (U, φ) and (V, ψ).)

A3. Exotic smooth structures

Suppose $h\colon M^m \to N^m$ is a homeomorphism between topological manifolds (so that M and N are really the "same" topological manifold). Then h can be used to move other structures between M and N. A trivial example would be a real-valued function $f\colon N \to \mathbb{R}$; it can be "pulled back" to give the real-valued function $f \circ h$ on M.

Of interest to us is the case of a C^r-structure $\mathcal{U}$ on N (a maximal atlas). We can use the homeomorphism h to pull it back to give a C^r-structure $h^*(\mathcal{U})$ on M: the pull-back of a chart $(U, \varphi) \in \mathcal{U}$ is the chart $(h^{-1}(U), \varphi \circ h)$ for M. Almost by definition, $h\colon M \to N$ is then a C^r-diffeomorphism from $(M, h^*(\mathcal{U}))$ to $(N, \mathcal{U})$.

Since M and N are homeomorphic, they are really the same topological manifold, and we might as well be considering self-homeomorphisms $h\colon M \to M$. If $h = \mathrm{id}$ then clearly $h^*(\mathcal{U}) = \mathcal{U}$; more generally this is true any time h is a diffeomorphism from the smooth manifold $(M, \mathcal{U})$ to itself.

But suppose h is a homeomorphism that is not a diffeomorphism. Then $h^*(\mathcal{U})$ is a *distinct* smooth structure on the manifold M. Consider a couple of examples on the line $M = \mathbb{R}$, starting with its standard smooth structure $\mathcal{U}$; the pull-back structure $h^*(\mathcal{U})$ is the one generated by the single coordinate chart $(\mathbb{R}, h)$. If $h\colon x \mapsto x^3$ then h is smooth but its inverse is not, so with respect to $h^*(\mathcal{U})$ it is easier for maps into M to be smooth, but harder for maps from M to be smooth. (The reverse is true of course if we start with $x \mapsto \sqrt[3]{x}$.) If on the other hand $h\colon x \mapsto 2x + |x|$, then neither h nor its inverse is smooth. (Note that in all these examples, the meaning of smoothness changes only near 0.)

Such examples are weird, but in fact they are all trivial. As we noted above, h is always a diffeomorphism from $(M, h^*(\mathcal{U}))$ to $(M, \mathcal{U})$. Thus the two smooth manifolds are diffeomorphic to each other – really the same smooth manifold. We have merely put on strange eyeglasses – the map h – to relabel the points of M.

More interesting is the question of existence of "exotic" smooth structures – can two different (nondiffeomorphic) smooth manifolds have a homeomorphism between them (meaning that their underlying topological manifolds are the same). There are still many interesting open questions here, especially in dimension 4. The following facts are known:

- Up to diffeomorphism, there is a unique smooth structure on any topological manifold M^m in dimension $m \le 3$. Up to diffeomorphism, there is a unique smooth structure on $\mathbb{R}^m$ for $m \ne 4$.
- The *Hauptvermutung* (known by that name even in English) of geometric topology (formulated 100 years ago) suggested that every topological manifold should have a unique piecewise linear (PL) structure – essentially given combinatorially by a triangulation – and a unique smooth structure. This is now known to be false.
- Every smooth manifold has a (PL) triangulation. For every dimension $m \ge 4$, there are topological m-manifolds that admit no triangulation – and in particular no PL or smooth structures. (For $m = 4$ this has been known since the 1980s, but for $m > 4$ it was just proven in 2013!)
- There are uncountably many different smooth structures on $\mathbb{R}^4$. It is unknown if there is any exotic smooth structure on $\mathbb{S}^4$.
- In higher dimensions, some things get easier. In dimensions $m \ge 7$, for instance, there are exotic spheres $\mathbb{S}^m$, but these form a well-understood finite group (e.g., there are 28 for $m = 7$). In general, the differences between smooth and PL manifolds (and to some extent between PL and topological manifolds) can be analyzed for $m \ge 5$ by means of algebraic topology.
- For compact, simply connected topological 4-manifolds, Freedman showed how to use invariants from algebraic topology to check when they are homeomorphic. In most (but not all) cases we know which of these topological manifolds admit smooth structures; it is not known how to classify the smooth structures when they do exist.

Especially since we know there are exotic spheres in certain dimensions, it is important to say what we mean by the standard sphere $\mathbb{S}^m$ as a smooth manifold. The "right" answer is that it inherits a smooth structure as a smooth submanifold of $\mathbb{R}^{m+1}$, but since we haven't developed that theory yet, we use explicit charts. Any "obvious" atlas will give the same standard smooth sphere, for instance the two charts of stereographic projection:

$$U_\pm = \mathbb{S}^m \smallsetminus \{\pm e_{m+1}\}, \quad \varphi_\pm(\mathbf{x}, z) = \frac{\mathbf{x}}{1 \mp z}, \;\; \mathbf{x} \in \mathbb{R}^m, \;\; z \in \mathbb{R},$$

or the $2m + 2$ charts of orthogonal projection:

$$U_{\pm j} = \{x \in \mathbb{S}^m \subset \mathbb{R}^{m+1} : \operatorname{sgn} x^j = \pm 1\},$$
$$\varphi_{\pm j}(x) = (x^1, \ldots, \widehat{x^j}, \ldots, x^{m+1}).$$

It is a good exercise to check that all these charts are C^∞-compatible.

With our basic constructions, we then get many further examples of smooth manifolds, like the m-torus $T^m = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ (a product of circles) or the n^2-dimensional matrix group $GL_n\mathbb{R} \subset \mathbb{R}^{n \times n}$ (an open subset).

A4. Smooth maps, rank, immersions

Suppose $f\colon M^m \to N^n$ is smooth, and $g = \psi \circ f \circ \varphi^{-1}$ is its expression in some local coordinates (U, φ) around $p \in M$ and (V, ψ) around $f(p) \in N$. Then we define the *rank* of f at p to be the rank of g at $x = \varphi(p)$, that is, the rank of the Jacobian derivative matrix $(\partial g^j/\partial x^i)$ there.

That this is well-defined (independent of coordinates) follows from the chain rule: a change of coordinates (on one side or the other) would multiply the Jacobian by an invertible square matrix (on one side or the other), leaving the rank unchanged. Note that the rank can be at most $\min(m, n)$.

The rank theorem from multivariable calculus can be restated most nicely for smooth manifolds. (When stated for Euclidean spaces it needs to mention diffeomorphisms.)

Theorem A4.1 (Rank Theorem). *Suppose $f\colon M^m \to N^n$ is a smooth map of constant rank k. Then for each $p \in M$ there are coordinate neighborhoods (U, φ) of p and (V, ψ) of $f(p)$ such that $\psi \circ f \circ \varphi^{-1}$ is the orthogonal projection map*

$$(x^1, \dots, x^m) \mapsto (x^1, \dots, x^k, 0, \dots, 0),$$

where of course there are $n - k$ zeros at the right. (Note: if we want, we can require that $\varphi(U) = B_1(0)$ and $\psi(V) = B_1(0)$.)

In particular, maps of maximum rank are important. We say f is a *submersion* if f has constant rank $n \le m$. We say f is an *immersion* if f has constant rank $m \le n$. For $m = n$ these notions coincide. A map $f\colon M^m \to N^m$ is a diffeomorphism if and only if it is bijective and has constant rank m.

Now recall that smooth functions (unlike analytic functions) are quite flexible. Starting with a function like

$$h\colon \mathbb{R} \to \mathbb{R}, \quad h(x) = \begin{cases} 0, & x \le 0, \\ e^{-1/x}, & x > 0, \end{cases}$$

we can use it to give a smoothed step function (0 for $x \le 0$ and 1 for $x \ge 1$), and then a smoothed bump function, which can also be rotated into higher dimensions.

For the next proof, we thus fix a smooth function g on $\mathbb{R}^m$ with $g(x) \equiv 1$ for $|x| \le 1$ and $g(x) \equiv 0$ for $|x| \ge 2$.

Theorem A4.2. *Suppose F and K are disjoint closed subsets of a (smooth) manifold M^m, with K compact. Then there is a smooth function $\sigma\colon M \to \mathbb{R}$ with values in $[0, 1]$ such that $\sigma \equiv 0$ on F and $\sigma \equiv 1$ on K.*

Proof. Each point $p \in K$ has a coordinate neighborhood $U_p \subset M \smallsetminus F$ such that $\varphi(p) = 0$ and $\varphi_p(U_p) = B_3(0) \subset \mathbb{R}^m$. Then we can define a smooth function $g_p\colon M \to \mathbb{R}$ by $g_p := g \circ \varphi_p$ on U_p, extending it to be 0 outside U_p. The open sets $V_p := \varphi_p^{-1}(B_1(0))$ cover $K \subset M \smallsetminus F$ (since there is one for each $p \in K$). By compactness, we get a finite subcover $\{V_{p_1}, \dots, V_{p_k}\}$. Then if we define $\sigma := 1 - \prod_{i=1}^k (1 - g_{p_i})$, it is easy to check $\sigma \equiv 1$ on K and $\sigma \equiv 0$ on F. □

Corollary A4.3. *Suppose $U \subset M^m$ is open, $f\colon U \to \mathbb{R}$ is smooth and $p \in U$ is given. Then there is a neighborhood $V \subset U$ of p and a smooth map $g\colon M \to \mathbb{R}$ such that $g \equiv f$ on V and $g \equiv 0$ outside U.*

Proof. Since M is locally compact, we can choose neighborhoods

$$p \in V_1 \subset \overline{V_1} \subset V_2 \subset \overline{V_2} \subset U$$

with $\overline{V_i}$ compact. (Quite explicitly, we can take any coordinate neighborhood (V_3, φ) with $p \in V_3 \subset U$ and $\varphi(p) = 0$ and $\varphi(V_3) = B_3(0)$; then we set $V_i := \varphi^{-1}(B_i(0))$.) Use the theorem to find σ with $\sigma \equiv 1$ on $\overline{V_1}$ and $\sigma \equiv 0$ outside V_2. Then define g to equal σf on U and to be 0 outside $\overline{V_2}$. (Each of these is a smooth function on an open set; they agree on the intersection. Thus they define a smooth g on the union, which is all of M.) □

A5. Tangent vectors and tangent spaces

We know that the tangent space $T_p\mathbb{R}^n$ at a point $p \in \mathbb{R}^n$ is a copy of the vector space $\mathbb{R}^n$; a vector $v \in T_p\mathbb{R}^n$ can be viewed intuitively as an arrow from p to $p+v$. (Technically, of course, a tangent vector knows where it is based, so we could set $T_p\mathbb{R}^n = \{(p, v) : v \in \mathbb{R}^n\}$, but usually we write just v for (p, v).)

If γ is a curve through $p := \gamma(0)$ in $\mathbb{R}^n$, then its velocity $\gamma'(0)$ is best viewed as a vector in $T_p\mathbb{R}^n$. If M^m is an m-submanifold through $p \in \mathbb{R}^n$, then the tangent space T_pM^m is an m-dimensional linear subspace of $T_p\mathbb{R}^n$, consisting of all velocity vectors to curves lying in M.

For an abstract manifold M^m, its tangent space T_pM^m should still be the collection of velocity vectors to curves through $p \in M$. But of course, there are always many curves with the same tangent vector. One approach would be to define tangent vectors as equivalence classes of curves, but when are two curves equivalent? One could say: "when they agree to first order", but this begs the question.

A good approach is to think about what we use tangent vectors for: to take directional derivatives! If $g\colon \mathbb{R}^n \to \mathbb{R}$ is a real-valued function, and γ is a curve through $p = \gamma(0)$ with velocity $v = \gamma'(0)$ there, then the directional derivative of g is the derivative along γ:

$$\partial_v g = D_p g(v) = \frac{d}{dt}\Big|_{t=0} (g \circ \gamma).$$

Let us think about ∂_v as a map taking g to the real number $\partial_v g \in \mathbb{R}$. This is linear:

$$\partial_v(g + \lambda h) = \partial_v g + \lambda \partial_v h$$

and satisfies the Leibniz product rule:

$$\partial_v(gh) = (\partial_v g)h(p) + (\partial_v h)g(p).$$

Such a map is called a *derivation*. Furthermore, it is local in the sense that $\partial_v g$ only depends on values of g in an arbitrarily small neighborhood of p. A fact we will check later is that there are no other local derivations at $p \in \mathbb{R}^n$ besides these directional derivatives. Thus we can use this as the definition of tangent vector.

So fix a point p in a (smooth) manifold M^m. What is the right domain for a derivation (think of a directional derivative) at $p \in M$? Consider the class

$$C = \bigcup_{U \ni p} C^\infty(U)$$

of all real-valued functions g defined on some open neighborhood of p. If two functions agree on some neighborhood, then they must have the same derivatives at p, so we consider them to be equivalent. More precisely, $g\colon U \to \mathbb{R}$ is *equivalent* to $h\colon V \to \mathbb{R}$ if there is some open $W \ni p$ (with $W \subset U \cap V$) such that $g|_W = h|_W$. An equivalence class is called a *germ* (of a smooth function) at p. The set of germs at p is the quotient space $C/\!\sim =: C^\infty(p)$. If g is a function on a neighborhood of p, we often write simply g for its germ (which might more properly be called $[g]$).

Note that if $g \in C^\infty(p)$ is a germ, we can talk about its value $g(p) \in \mathbb{R}$ at p, but not about its value at any other point. (For $M = \mathbb{R}^m$, a germ g at p also encodes all derviatives at p – that is, the Taylor series of g – but also much more information, since g is not necessarily analytic.)

The set $C^\infty(p)$ of germs is an (infinite dimensional) algebra over $\mathbb{R}$, that is, a vector space with multiplication. (Exercise: check that multiplication of germs makes sense, etc.) We now define a *tangent vector* X_p at $p \in M$ to be a derivation on this algebra. That is, $X_p\colon C^\infty(p) \to \mathbb{R}$ is a linear functional:

$$X_p(g + \lambda h) = X_p g + \lambda X_p h$$

satisfying the Leibniz rule:

$$X_p(gh) = (X_p g)h(p) + (X_p h)g(p).$$

We let T_pM denote the *tangent space* to M at p, that is, the set of all such tangent vectors X_p.

End of Lecture 19 October 2015

Clearly T_pM is a vector space, with the obvious operations

$$(X_p + \lambda Y_p)f := X_p f + \lambda Y_p f.$$

(In fact, this is just exhibiting T_pM as a linear subspace within the abstract dual vector space $C^\infty(p)^*$ of *all* linear functionals on $C^\infty(p)$).

Note that if $U \subset M$ is open with $p \in U$, then $T_pU = T_pM$ since the set $C^\infty(p)$ of germs is the same whether we start with M or U.

Now suppose $f\colon M^m \to N^n$ is a smooth map of manifolds and consider a point $p \in M$ and its image $q := f(p) \in N$. If g is a germ at q, then $g \circ f$ is a germ at p. (Here of course, we really compose f with any of the functions in the equivalence class g.) This gives a map

$$f^*\colon C^\infty(q) \to C^\infty(p), \qquad f^*(g) := g \circ f$$

between these algebras of germs, which we claim is linear, indeed an algebra homomorphism. (Note that the upper star is used to indicate a "pull-back", a map associated to f acting in the opposite direction.)

Like any linear map between vector spaces, f^* induces a dual map between the dual spaces; here we claim this restricts to a map $f_*\colon T_pM \to T_qN$. Working out what the dual map means, we find that for $X_p \in T_pM$ and $g \in C^\infty(q)$ we have

$$(f_*(X_p))(g) = X_p(f^*(g)) = X_p(g \circ f).$$

This linear map f_* is called the *differential* of f at p and we will usually write it as $D_p f$. (Other common notations include $d_p f$ or simply f'.)

Theorem A5.1. *Given a smooth map $f\colon M^m \to N^n$ of manifolds and a point $p \in M$, the construction above induces a linear map $f_* = D_p f\colon T_pM \to T_{f(p)}N$, the differential of f.*

Proof. The many claims we made during the construction are all routine to check. We gives just two examples. To see that $f_*(X_p)$ is actually a tangent vector at $q := f(p)$, we need to check the Leibniz rule:

$$\begin{aligned} f_*(X_p)(gh) &= X_p((gh) \circ f) = X_p((g \circ f)(h \circ f)) \\ &= (X_p(g \circ f))(h(q)) + (X_p(h \circ f))(g(q)) \\ &= (f_*(X_p)g)h(q) + (f_*(X_p)h)g(q). \end{aligned}$$

To see that f_* is linear, we compute:

$$\begin{aligned} f_*(X_p + \lambda Y_p)(g) &= (X_p + \lambda Y_p)(g \circ f) \\ &= X_p(g \circ f) + \lambda Y_p(g \circ f) \\ &= (f_*(X_p) + \lambda f_*(Y_p))(g). \end{aligned}$$

□

It is now a straightforward exercise to check the "functoriality" of the operation $f \mapsto f_*$, that is, the following two properties:

- For $f = \mathrm{id}\colon M \to M$, the maps f^* and f_* are also the identity maps.
- If $h = g \circ f$ (for maps between appropriate manifolds), then $h^* = f^* \circ g^*$ and $h_* = g_* \circ f_*$.

(The second of these is of course the chain rule from calculus.)

Corollary A5.2. *If $f\colon M \to N$ is a diffeomorphism, then for any $p \in M$, the map $D_p f\colon T_pM \to T_{f(p)}N$ is an isomorphism. In particular, if (U, φ) is a coordinate chart for M^m, then $\varphi_*\colon T_pM \to T_{\varphi(p)}\mathbb{R}^m$ is an isomorphism*

Of course, this refers to $T_q\mathbb{R}^m$ in the sense we have just defined for abstract manifolds. It is time to go back and prove the claim we made early on, that there are no derivations on $\mathbb{R}^m$ other than the usual directional derivatives, that is, that $T_p\mathbb{R}^m \cong \mathbb{R}^m$.

We know we have a map $\mathbb{R}^m \to T_p\mathbb{R}^m$ which associates to each $v \in \mathbb{R}^m$ the directional derivative ∂_v at p. This map is clearly linear and is easily seen to be injective. Indeed, if $\pi^i\colon \mathbb{R}^m \to \mathbb{R}$ denotes the projection $p \mapsto p^i$, then $\partial_v \pi^i =$

$v^i = \pi^i(v)$; now two distinct vectors $v \neq w$ must differ in some component $v^i \neq w^i$, meaning $\partial_v \pi^i \neq \partial_w \pi^i$, so $\partial_v \neq \partial_w$. The claim that now remains is just that this map is surjective – there are no other derivations.

Lemma A5.3. *Suppose $X_p \in T_p\mathbb{R}^m$ and $g \in C^\infty(p)$ is constant (in some neighborhood of p). Then $X_p g = 0$.*

Proof. By linearity of X_p it suffices to consider $g \equiv 1$. By the Leibniz rule,

$$X_p(1) = X_p(1 \cdot 1) = X_p(1) \cdot 1 + X_p(1) \cdot 1 = 2X_p(1)$$

which clearly implies $X_p(1) = 0$. □

Our next lemma can be thought of as a version of Taylor's theorem. (Note that one could let B be an arbitrary star-shaped region around p.)

Lemma A5.4. *Let $B := B_\varepsilon(p)$ where $p \in \mathbb{R}^m$ and $\varepsilon > 0$. For any $g \in C^\infty(B)$, we can find a collection of m functions $h^i \in C^\infty(B)$ with $h^i(p) = \frac{\partial g}{\partial x^i}(p)$, such that on B we have*

$$g(x) = g(p) + \sum_i (x^i - p^i)h^i(x).$$

Proof. If we set

$$h^i(x) := \int_0^1 \frac{\partial g}{\partial x^i}(p + t(x - p))\, dt$$

then the desired properties follow from the fundamental theorem of calculus in the form

$$g(x) = g(p) + \int_0^1 \frac{d}{dt} g(p + t(x - p))\, dt,$$

noting that this t-derivative is the directional derivative of g in direction $x - p$. □

Theorem A5.5. *The map $v \mapsto \partial_v$ is a (natural) isomorphism $\mathbb{R}^m \to T_p\mathbb{R}^m$.*

Proof. As noted above, all that remains is to prove surjectivity. Given $X_p \in T_p\mathbb{R}^m$, define $v \in \mathbb{R}^m$ by $v^i := X_p(\pi^i)$. We claim $\partial_v = X_p$. By definition, these agree on (the germs of) the projections π^i. Now suppose $g \in C^\infty(p)$ is any germ. Finding a representative $g \in C^\infty(B_\varepsilon(p))$ for some $\varepsilon > 0$, we can use the second lemma to write

$$g = g(p) + \sum_i (\pi^i - p^i)h^i.$$

Then by the definition of derivation,

$$X_p g = X_p(g(p)) + \sum (X_p \pi^i - X_p p^i)(h^i(p)) + \sum (X_p h^i)(\pi^i(p) - p^i).$$

Here the last sum (which would seem to involve second derivatives of g) vanishes simply because $\pi^i(p) = p^i$. And the terms $X_p p^i$ and $X_p(g(p))$ vanish by the first lemma. Thus we are left with

$$X_p g = \sum_i (X_p \pi^i)(h^i(p)) = \sum v^i \frac{\partial g}{\partial x^i}(p) = \partial_v(g)$$

as desired. □

Note that if $\{e_i\}$ is the standard basis of $\mathbb{R}^m$ (so that $v = \sum v^i e_i$), then $\{\partial_{e_i} = \frac{\partial}{\partial x^i}\}$ is the corresponding standard basis of $T_p\mathbb{R}^m$. Recall that, given a coordinate chart (U, φ) around $p \in M^m$, the differential $D_p\varphi = \varphi_* \colon T_pM \to T_{\varphi(p)}\mathbb{R}^m \cong \mathbb{R}^m$ is an isomorphism. Under this isomorphism, the $\frac{\partial}{\partial x^i}$ correspond to the elements of a basis for T_pM, which we write as

$$\partial_i = \partial_{i,p} := \varphi_*^{-1}\left(\frac{\partial}{\partial x^i}\right).$$

Suppose a function $f \in C^\infty(U)$ has coordinate expression $f \circ \varphi^{-1} \colon \varphi(U) \to \mathbb{R}$, then at $p \in U$ we get

$$\partial_i f = \varphi_*^{-1}\left(\frac{\partial}{\partial x^i}\right)(f) = \frac{\partial}{\partial x^i}\left(f \circ \varphi^{-1}\right).$$

In particular, if we consider the individual components $x^j = \pi^j \circ \varphi$ of the coordinates φ as real-valued functions, we find $\partial_i(\pi^j \circ \varphi) = \partial \pi^j / \partial x^i = \delta_i^j$. We can express any $X_p \in T_pM$ in terms of our basis $\{\partial_i\}$ as follows:

$$X_p = \sum_{i=1}^m (X_p(\pi^i \circ \varphi))\partial_i$$

Consider a smooth map $f \colon M^m \to N^n$ between manifolds, and choose local coordinates (U, φ) around $p \in M$ and (V, ψ) around $q := f(p) \in N$. (We write $x^i = \pi^i \circ \varphi$ and $y^j = \pi^j \circ \psi$.) In these coordinates, f is represented by the map $\psi \circ f \circ \varphi^{-1}$ of Euclidean spaces, or more explicitly as functions $y^j = f^j(x^1, \ldots, x^m)$. Here the derivative is given by the Jacobian matrix

$$J = \left(\frac{\partial y^j}{\partial x^i}\right) = \left(\frac{\partial f^j}{\partial x^i}\right).$$

Let us write $\{\partial_i\}$ as usual for the coordinate frame of T_pM, where $\partial_i = \varphi_*^{-1}(\partial/\partial x^i)$. For T_qN, we use the notation $\tilde\partial_j = \psi_*^{-1}(\partial/\partial y^j)$. We find that J is the matrix of D_pf with respect to these bases. That is,

$$D_pf(\partial_{i,p}) = \sum_j \left(\frac{\partial y^j}{\partial x^i}\right)_{\varphi(p)} \tilde\partial_{j,q},$$

or equivalently, if $X_p = \sum v^i \partial_{i,p}$ and $f_*(X_p) = \sum w^j \tilde\partial_{j,q}$, then we have

$$w^j = \sum_i v^i \left(\frac{\partial y^j}{\partial x^i}\right)_{\varphi(p)}.$$

We should have waited until now to define the rank at $p \in M$ of a map $f \colon M^m \to N^n$. It is simply the rank of the linear map $D_pf \colon T_pM \to T_{f(p)}N$. In coordinates this is, of course, the rank of the Jacobian matrix above, as we defined before.

One special case is when $f = \mathrm{id}_M$ is the identity map. That is, we have overlapping coordinate charts (U, φ) and (V, ψ) for M^m. At any point $p \in U \cap V$, we have two different coordinate bases for T_pM, which we write as $\{\partial_i\}$ (with respect to φ) and $\{\tilde\partial_i\}$ (with respect to ψ). Then the change-of-basis matrix is just the Jacobian matrix of the coordinate expression of id_M, which here is just the transition function $\psi \circ \varphi^{-1}$. (This is the basis for a definition of tangent vectors still popular among physicists: a tangent vector "is" its expression in a coordinate base, with the rules for changing this "covariantly" when we change coordinates.)

As usual, we also consider the special cases where one of the manifolds M or N is (a submanifold of) $\mathbb{R}$. For $\mathbb{R}$ of course we use the standard chart (the identity map), and we write ∂_t for the corresponding basis vector for the tangent space to $\mathbb{R}$ at any point.

A map $\gamma\colon (a,b) \to M^m$ is a *curve* in M. Its tangent vector at $p := \gamma(t) \in M$ is $\gamma'(t) := D_t\gamma(\partial_t) \in T_pM$.

The opposite case is a real-valued function $f \in C^\infty(M)$. For $X \in T_pM$, we have $(D_pf)(X) \in T_{f(p)}\mathbb{R}$, so $(D_pf)(X) = \lambda\partial_t$ for some $\lambda \in \mathbb{R}$. Of course, this λ is just the directional derivative Xf. For instance, in local coordinates, $(D_pf)(\partial_i) = (\partial_i f)\partial_t$. We write $df = d_pf\colon T_pM \to \mathbb{R}$ for the linear map $X \mapsto Xf$. The dual vector space T_p^*M is called the *cotangent space* and its elements are *cotangent vectors* (or *covectors* for short). Thus $d_pf \in T_p^*M$ is the covector given by $df(X) := Xf$; this is just another way to view the differential since we have $D_pf(X) = d_pf(X)\partial_t$.

A6. The tangent bundle

Definition A6.1. The *tangent bundle* $TM = T(M)$ to a smooth manifold M^m is, as a set, the (disjoint) union $TM := \bigcup_{p\in M} T_pM$ of all tangent spaces to M; there is obviously a projection $\pi\colon TM \to M$ with $\pi^{-1}\{p\} = T_pM$. We can equip TM in a natural way with the structure of a smooth $2m$-manifold. Start with a (smooth) atlas for M. Over any coordinate chart (U,φ), there is a a bijection $D\varphi\colon TU \to \varphi(U) \times \mathbb{R}^m \subset \mathbb{R}^{2m}$ sending $\sum_i v^i\partial_i \in T_pU = T_pM$ to $(\varphi(p), v)$. We define the topology on TM by specifying that these maps $D\varphi$ are homeomorphisms; they then form an atlas for TM as a topological $2m$-manifold.

Exercise A6.2. These charts for TM are C^∞-compatible, and thus define a smooth structure on TM.

Note that TM is an example of a vector bundle, which is a special kind of fiber bundle to be defined later. Without going into details, a *fiber bundle* with base B and fiber F is a certain kind of space E with projection $\pi\colon E \to B$ such that the preimage of any point $b \in B$ is isomorphic to F. A trivial bundle is $E = F \times B$ projecting to the second factor. Any fiber bundle is required to be locally trivial in the sense that B is covered by open sets U over which the bundle is trivial ($F \times U$). A *section* of a bundle $\pi\colon E \to B$ is a continuous choice of point in each fiber, that is, a map $\sigma\colon B \to E$ such that $\pi \circ \sigma = \mathrm{id}_B$.

Definition A6.3. A (smooth) *vector field* X on a manifold M^m is a smooth choice of a vector $X_p \in T_pM$ for each point $p \in M$. That is, X is a (smooth) *section* of the bundle $\pi\colon TM \to M$, meaning a smooth map $X\colon M \to TM$ such that $\pi \circ X = \mathrm{id}_M$. We write $\mathcal{X} = \mathcal{X}(M) = \Gamma(TM)$ for the set of all vector fields.

We define addition of vector fields pointwise: $(X+Y)_p = X_p + Y_p$. Similarly, we can multiply a vector field $X \in \mathcal{X}(M)$ by a smooth function $f \in C^\infty(M)$ pointwise: $(fX)_p = f(p)X_p$. That is, the vector fields $\mathcal{X}(M)$ form not just a real vector space, but in fact a module over the ring $C^\infty(M)$ of smooth functions.

Given a vector field X and a function f, we can also define $Xf \in C^\infty(M)$ by $(Xf)(p) := X_pf \in \mathbb{R}$. Note the distinction between the vector field fX (given by pointwise scalar multiplication) and the function Xf (given by directional derivatives of f).

Exercise A6.4. Each of the following conditions is equivalent to the smoothness of a vector field X as a section $X\colon M \to TM$:

- For each $f \in C^\infty(M)$, the function Xf is also smooth.
- If we write $X|_U =: \sum v^i\partial_i$ in a coordinate chart (U,φ), then the components $v^i\colon U \to \mathbb{R}$ are smooth.

`End of Lecture 26 October 2015`

A7. Submanifolds

The canonical example of an m-dimensional submanifold of an n-manifold is $\mathbb{R}^m \times 0 \subset \mathbb{R}^n$, the set of vectors whose last $n-m$ components vanish.

Definition A7.1. Given a manifold N^n, we say a subset $M \subset N$ is an *m-submanifold* if around each point $p \in M$ there is a coordinate chart (U,φ) for N in which M looks like $\mathbb{R}^m \times 0 \subset \mathbb{R}^n$. That is, in such a *preferred chart* we have

$$\varphi(M \cap U) = \varphi(U) \cap (\mathbb{R}^m \times 0).$$

It is straightforward then to check that M (with the subspace topology) is an m-manifold. Indeed, the preferred charts form a C^∞ atlas for M^m.

Two alternative local characterizations – as for submanifolds in $\mathbb{R}^n$ – are then immediate. A submanifold $M^m \subset N^n$ can be described locally (that is, in some neighborhood $U \subset N$ of any point $p \in M$) as

1. the zero level set of a submersion $N^n \to \mathbb{R}^{n-m}$ (here φ composed with projection onto the last $n-m$ coordinates), or
2. the image of an immersion $\mathbb{R}^m \to N^n$ (here the standard inclusion $\mathbb{R}^m \hookrightarrow \mathbb{R}^n$ composed with φ^{-1}).

We now want to consider in more detail the description of submanifolds via immersions. Immersions from $\mathbb{R}^m$ are also known as *regular parametrizations*. Recall that last semester we used such regular parametrizations to describe curves and surfaces in $\mathbb{R}^3$. It is of course important that the parametrization be an immersion, in order to be sure that the image is a smooth submanifold.

Examples A7.2. Consider the following examples of immersions based on smooth plane curves.

1. $t \mapsto (\cos 2\pi t, \sin 2\pi t)$ is a periodic parametrization of a simple closed curve, a 1-submanifold. This immersion is not injective, but becomes injective if we consider the domain to be the circle $\mathbb{R}/\mathbb{Z} \cong \mathbb{S}^1$.

2. If $r\colon \mathbb{R} \to (1,2)$ is strictly monotonic, then
$$t \mapsto (r(t)\cos 2\pi t, r(t)\sin 2\pi t)$$
is an injective immersion whose image is a submanifold, a spiral curve in the plane. Note that this image is not a closed subset of $\mathbb{R}^2$ (because the immersion is not "proper").

3. $t \mapsto (\cos 2\pi t, \sin 4\pi t)$ is again a closed curve, this time a figure-eight. It descends again to the quotient cirle $\mathbb{R}/\mathbb{Z}$, but is not injective even there. The image is not a submanifold.

4. If we restrict this last example to the open interval $(-1/4, 3/4)$, which of course is diffeomorphic to $\mathbb{R}$, we get an injective immersion whose image is still the whole figure-eight curve, not a submanifold.

5. One can build an injective immersion whose image is not even locally connected. For instance, join the "topologist's sine curve", the curve $t \mapsto (1/t, \sin t)$ for $t \geq 2$, to a downward ray in the y-axis, the curve $t \mapsto (0, t)$ for $t \leq 1$, via a smooth intermediate arc for $t \in [1, 2]$.

6. For any slope $\alpha \in \mathbb{R}$, we can project the line $t \mapsto (t, \alpha t)$ of slope α from $\mathbb{R}^2$ to the quotient torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$. For $\alpha = p/q \in \mathbb{Q}$, this gives a periodic curve, that is, a circle submanifold in T^2 called a (p, q)-torus knot. For irrational α on the other hand, the immersion is injective but its image X is dense in T^2 and thus is not a submanifold (in our sense).

In some other contexts, mainly that of Lie groups, examples like this last one can be considered as submanifolds. A Lie group is smooth manifold with the structure of an algebraic group, where the group operations are smooth maps; we will discuss these later. The torus T^2 is an example of a (compact, 2-dimensional) Lie group (under addition). The dense subset X – consisting of points of the form $(t, \alpha t)$ – is a subgroup; from the point of view of Lie groups this is a 1-dimensional Lie subgroup. Of course $X \subset T^2$ with the subset topology is not a manifold. Instead we simply use the bijective immersion $\mathbb{R} \to X$ to transfer the standard smooth manifold structure from $\mathbb{R}$ to X. (Indeed, any time we have an injective immersion $f\colon M^m \to N^n$, it is a bijection onto its image, and could be used to transfer the topology and smooth structure from M to that image, making f by definition a diffeormorphism, though not to a subspace of N.)

Definition A7.3. A continuous injection $f\colon X \to Y$ of topological spaces is a topological *embedding* if it is a homeomorphism onto its image $f(X)$. A (smooth) *embedding* $f\colon M^m \to N^n$ of manifolds is an immersion that is a topological embedding.

We will show that the image of a smooth embedding is a submanifold (and vice versa); the embedding is then not merely a homeomorphism but indeed a diffeomorphism onto its image.

Above we saw many examples of injective immersions of manifolds which were not embeddings. However, there is one standard result from point-set topology which guarantees that this never happens when M is compact.

Proposition A7.4. *If X is compact and Y is Hausdorff, then any continuous bijection $f\colon X \to Y$ is a homeomorphism.*

For the proof, recall that a map $f\colon X \to Y$ is *open* if the image of every open set $U \subset X$ is open in Y, and f is *closed* if the image of every closed set $A \subset X$ is closed in Y. If f is a bijection, then these notions are equiavalent, and also equivalent to f^{-1} being continuous.

Proof. We need to show f^{-1} is continuous, or equivalently that f is a closed map. So suppose $A \subset X$ is closed; we need to show $f(A)$ is closed in Y. Since X is compact, A is also compact. Since f is continuous, $f(A)$ is then compact. But a compact subset of the Hausdorff space Y is necessarily closed. □

Two examples related to the quotient map $I \to \mathbb{S}^1$ (where $I = [0, 1]$ and the quotient identifies the endpoints $\{0, 1\}$ to a single point) show why the two conditions are necessary. First, we can get a bijection by restricting this map to the noncompact interval $[0, 1)$. Second, we can get a bijection by replacing $\mathbb{S}^1$ by a non-Hausdorff circle with a doubled basepoint (like our line with doubled origin).

Corollary A7.5. *If X is compact and Y is Hausdorff, then any (continuous) injection $f\colon X \to Y$ is a topological embedding.*

Corollary A7.6. *If M is a compact manifold, then any injective immersion $f\colon M^m \to N^n$ is a smooth embedding.*

Now we show that the concepts of smooth embedding and submanifold coincide in the following sense:

Theorem A7.7. *If $f\colon M^m \to N^n$ is an embedding, then $f(M) \subset N$ is a submanifold and $f\colon M \to f(M)$ is a diffeomorphism. If $M^m \subset N^n$ is any submanifold, then the inclusion $i\colon M \hookrightarrow N$ is an embedding.*

Proof. For the first statement, consider a point $p \in M$ and its image $q = f(p) \in f(M) \subset N$. Because f has constant rank $m \leq n$, by the rank theorem, we can find coordinates (U, φ) around $p \in M$ and (V, ψ) around $q \in N$ in which f looks like the embedding $\mathbb{R}^m \hookrightarrow \mathbb{R}^n$. It is tempting to hope that (V, ψ) is the preferred chart we seek in the definition of submanifold – but we have not yet used the fact that f is an embedding and the problem is that other parts of $f(M)$ might enter V, while we want $f(U) = f(M) \cap V$. But since f is an embedding, $f(U)$ is open in $f(M)$ – in the subspace topology from N. By definition of subspace topology, this means there is an open subset $W \subset N$ such that $f(U) = W \cap f(M)$. Now we simply restrict (V, ψ) to

$W \cap V$ and we find these are preferred coordinates showing $f(M)$ as a submanifold of N (around q).

We essentially proved the second statement when we put a smooth structure on the submanifold $M \subset N$: we remarked then that the manifold topology on M was the subspace topology from N, which exactly means the inclusion is a topological embedding. The fact that it is an immersion is also obvious in a preferred coordinate chart. □

Suppose $M^m \subset N^n$ is a submanifold. Then at any $p \in M$ we can view $T_pM \subset T_pN$ in a natural way as a vector subspace (using the injective differential $D_p i$ of the inclusion map i). More generally, if M is described (locally) as the image of a regular parameterization, an immersion $\mathbb{R}^m \supset U \to N$, then T_pM is the image of its differential. If instead M is (locally) the zero set of a submersion $f\colon N \to \mathbb{R}^{n-m}$, then T_pM is the kernel of $D_p f$.

A8. Vector fields and their flows

Suppose $f\colon M^m \to N^n$ is a smooth map and X is a vector field on M. For any point $p \in M$, we can use $f_* = D_p f$ to push a vector $X_p \in T_pM$ to a vector at $f(p) \in N$. If there exists a vector field Y on N such that for each $p \in M$ we have $Y_{f(p)} = D_p f(X_p)$, then we say Y is *f-related* to X. Of course, when f is not injective, it might be impossible to find an f-related vector field; when f is not surjective, Y is not uniquely determined away from $f(M)$. But when f is a diffeomorphism, there clearly is a unique Y that is f-related to any given X, and then we write $Y = f_*(X)$.

If $f\colon M \to M$ is a diffeomorphism it can happen that a vector field X is f-related to itself: $X = f_*(X)$. In this case, we say X is *f-invariant*. As a simple example, consider the radial field $X_p = p$ on $\mathbb{R}^m$. It is invariant under any homothety $f\colon p \mapsto \lambda p$. This may seem like a very special situation, but in fact our goal now, given an arbitrary vector field X, is to construct a one-parameter family of diffeomorphisms $\theta_t\colon M \to M$ under which X is invariant.

Recall that if G is an algebraic group and X is any set, then an *action* θ of G on X is a map $\theta\colon G \times X \to X$, often written as

$$(g, x) \mapsto g \cdot x := \theta_g(x),$$

satisfying the following properties:

$$\theta_e = \mathrm{id}_X, \qquad \theta_{gh} = \theta_g \circ \theta_h.$$

(That is, in the typical group theory notation, $e \cdot x = x$ and $(gh) \cdot x = g \cdot (h \cdot x)$.) Each $\theta_g\colon X \to X$ is a bijection (with inverse $\theta_{g^{-1}}$). The action θ partitions the set X into *orbits*

$$G \cdot x := \{g \cdot x : g \in G\},$$

which are the equivalence classes under the equivalence relation $x \sim g \cdot x$.

We are interested in smooth actions of the (1-dimensional Lie) group $(\mathbb{R}, +)$ on a smooth manifold M^m. Such an action is a smooth map $\theta\colon \mathbb{R} \times M \to M$, again written as $(t, p) \mapsto \theta_t(p)$, satisfying

$$\theta_0 = \mathrm{id}_M, \qquad \theta_s \circ \theta_t = \theta_{s+t}.$$

It follows that each $\theta_t\colon M \to M$ is a diffeomorphism, with inverse θ_{-t}. Note that since $\mathbb{R}$ is an abelian group ($s + t = t + s$), these diffeomorphisms all commute:

$$\theta_t \circ \theta_s = \theta_{s+t} = \theta_s \circ \theta_t.$$

This is often simply called a *one-parameter group action* or a *(global) flow* on M.

Definition A8.1. We say a vector field X is *invariant* under the action if it is invariant under each θ_t, that is, if $(\theta_t)_* X = X$ for all t.

This may seem like a very special situation, but we will see it is quite natural.

Our notation $\theta_t(p) := \theta(t, p)$ emphasizes the diffeomorphisms θ_t obtained by fixing $t \in \mathbb{R}$. If instead, we fix a point $p \in M$, we of course get a curve $\gamma_p\colon \mathbb{R} \to M$ defined by $\gamma_p(t) := \theta_t(p)$. The trace of this curve is the orbit of $p \in M$ under the action θ. It is helpful to rewrite the defining property $\theta_s \circ \theta_t = \theta_{s+t}$ of a flow in terms of these flow curves. For any point $q := \gamma_p(s) = \theta_s(p)$ along γ_p, we find that the curve γ_q is just a reparametrization of γ_p; indeed

$$\gamma_q(t) = \theta_t(q) = \theta_t(\theta_s(p)) = \theta_{t+s}(p) = \gamma_p(s + t).$$

Equivalently, $\gamma_q = \theta_s \circ \gamma_p$:

$$\gamma_q(t) = \theta_s(\theta_t(p)) = \theta_s(\gamma_p(t)).$$

Definition A8.2. The *infinitesimal generator* of the flow θ is the vector field X on M defined by $X_p := \gamma'_p(0)$, the velocity vector of the curve γ_p at $p = \gamma_p(0)$.

`End of Lecture 2 November 2015`

An equivalent way to define the infinitesimal generator X comes from looking at a standard "vertical" vector field on $\mathbb{R} \times M$, defined by

$$T_{(t,p)}(\mathbb{R} \times M) = T_t\mathbb{R} \times T_pM \ni (\partial_t, 0) =: V_{(t,p)}.$$

Then it is easy to check that $X_p = (D_{(0,p)}\theta)(V)$.

Theorem A8.3. *Suppose θ is a flow on M with infinitesimal generator X. Then X is θ-invariant. That is, for any $s \in \mathbb{R}$ and $p \in M$ we have*

$$(\theta_s)_*(X_p) = X_{\theta_s p}.$$

Proof. Write $q = \theta_s(p)$ so that $X_p = \gamma'_p(0)$ and $X_q = \gamma'_q(0)$. Then the desired formula follows immediately from the observation above that $\gamma_q = \theta_s \circ \gamma_p$. □

Corollary A8.4. *If $X_p = 0$ then the curve γ_p is the constant map $\gamma_p(t) \equiv p$. If $X_p \neq 0$, then the curve γ_p is an immersion. If it is not injective on $\mathbb{R}$ then it is s-periodic and injective on $\mathbb{R}/s\mathbb{Z}$ for some $s > 0$.*

Proof. First note that

$$\gamma'_p(t) = X_{\gamma_p(t)} = (\theta_t)_* X_p.$$

Since θ_{t*} is a linear isomorphism, these vectors either remain zero or remain nonzero. Finally, if we have a noninjective curve with $\gamma_p(t + s) = \gamma_p(t)$ for some s and t, then the same holds for this s and every t, that is, γ_p is s-periodic. □

Recall that these curves are the *orbits* of the flow θ and form a partition of M. (Each orbit is topologically an open arc or a closed loop or a single point.)

Note also that the vector field X is θ-related to the vector field $V = (\partial_t, 0)$ on $\mathbb{R} \times M$ since

$$X_{\theta(t,p)} = D_{(t,p)}\theta(V_{(t,p)}).$$

As we have seen, any flow θ has an infinitesimal generator X. What if we start with a vector field X on M: does it generate a flow? We will see that the answer is always yes when M is compact; in general the flow might exist only for small t (depending on p).

Definition A8.5. Given a vector field X on a manifold M, a curve $\gamma\colon J \to M$ (where $J \subset \mathbb{R}$ is some open interval) is called an *integral curve* of X if $\gamma'(t) = X_{\gamma(t)}$ for all $t \in J$.

As we have seen, any $\mathbb{R}$-action θ has an infinitesimal generator X. Then each orbit $\gamma_p(t) = \theta_t(p)$ is an integral curve of X, defined on $J = \mathbb{R}$. In other cases, the integral curve does not exist for all time, since it flows out of M in finite time. For instance, consider the flow $\theta_t(x) = x + te_1$ on $\mathbb{R}^m$, whose infinitesimal generator is $X = \partial_1$. If we replace $\mathbb{R}^m$ by an open subset (like $\mathbb{R}^n \smallsetminus \{0\}$ or $\mathbb{R}^n \smallsetminus B_1(0)$) then we sometimes leave this open subset in finite time.

Think for a minute about dimension $m = 1$. Up to diffeomorphism, there is no difference between reaching the "end" of a finite open interval like $(0, 1)$ and reaching ∞. A classical example is the flow of $t^2\partial_t$ on $\mathbb{R}$, that is, the solution of the ODE $du/dt = u^2$, which blows up in finite time. So it is too much to hope for a global solution in general. But of course, standard theorems on ODEs guarantee local existence and uniqueness of solutions, which can be viewed as integral curves of a vector field. (In an ODE course, you might learn about minimal smoothness conditions for existence and for uniqueness; certainly C^∞ or even C^1 suffices for both.)

Theorem A8.6. *Suppose $U \subset \mathbb{R}^m$ is open and $f\colon U \to \mathbb{R}^m$ is smooth. Then for each $p \in U$, there is a unique solution to the equation $dx/dt = f(x)$ with initial condition $x(0) = p$; it is smooth and is defined on some maximal open time interval $(a_p, b_p) \ni 0$.*

A proof of this basic result (using the Banach fixed-point theorem for contraction mappings) can be found in Boothby. Somewhat more subtle is the "smooth dependence on parameters" as given in the next theorem. (See Conlon's textbook for a proof.) In our version, there are no parameters other than the initial point p.

Theorem A8.7. *Suppose $U \subset \mathbb{R}^m$ is open and $f\colon U \to \mathbb{R}^m$ is a smooth function. For any point $p \in U$ there exists $\varepsilon > 0$, a neighborhood $V \subset U$ of p, and a smooth map*

$$x\colon (-\varepsilon, \varepsilon) \times V \to U$$

satisfying

$$\frac{\partial x}{\partial t}(t, q) = f(x(t, q)), \quad x(0, q) = q$$

for all $t \in (-\varepsilon, \varepsilon)$ and $q \in V$.

Like any local result, this can be transferred immediately to the context of an arbitrary manifold, where its restatement has a more geometric flavor.

Definition A8.8. A *local flow* around $p \in M$ is a map

$$\theta\colon (-\varepsilon, \varepsilon) \times V \to M$$

(for some $\varepsilon > 0$ and some open $V \ni p$) such that $\theta_0(q) = q$ for all $q \in V$, and

$$\theta_t(\theta_s(q)) = \theta_{t+s}(q)$$

whenever both sides are defined. The *flow lines* are the curves $\gamma_q(t) := \theta_t(q)$; the *infinitesimal generator* is the vector field $X_q = \gamma_q'(0)$ tangent to the flow lines.

Theorem A8.9. *Any vector field X on a manifold M has a local flow around any point $p \in M$.*

Note that if we prove this theorem by appealing to the previous theorem on $\mathbb{R}^m$, then the neighborhood V we construct (and even the values of θ) will be in some coordinate chart around p. But this doesn't affect the statement of the theorem.

Theorem A8.10. *On a compact manifold M^m, any vector field X has a global flow.*

Proof. For any $p \in M$ we have a local flow, defined on some $(-\varepsilon_p, \varepsilon_p) \times V_p$. By compactness, finitely many of the V_p suffice to cover M. Let $\varepsilon > 0$ be the minimum of the corresponding (finitely many) ε_p. Then we know that the flow of X exists everywhere for a uniform time $t \in (-\varepsilon, \varepsilon)$. But then, for instance using $\theta_{nt} = \theta_t \circ \cdots \circ \theta_t$, we can construct flows for arbitrary times. □

A9. Lie brackets and Lie derivatives

Definition A9.1. A *Lie algebra* is a vector space $\mathcal{L}$ with an antisymmetric (or skew-symmetric) product

$$\mathcal{L} \times \mathcal{L} \to \mathcal{L}, \qquad (v, w) \mapsto [v, w] = -[w, v]$$

that is bilinear (i.e., linear in v and in w, where it suffices to check one of these:

$$[\lambda v + v', w] = \lambda[v, w] + [v', w],$$

noting this could also be called a distrubutive law) and satisfies the *Jacobi identity*

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0.$$

A trivial example is any vector space with the zero product $[v, w] := 0$. The Jacobi identity may not at first seem intuitive, but in fact there are some familiar nontrivial examples.

Example A9.2. Three-space $\mathbb{R}^3$ with the usual vector cross product $[v, w] := v \times w = v \wedge w$ is a Lie algebra.

Example A9.3. The ordinary matrix product on $\mathbb{R}^{n\times n}$ is bilinear but neither symmetric nor antisymmetric. But the matrix commutator $[A, B] := AB - BA$ is clearly antisymmetric. To check the Jacobi identity, we compute

$$[A,[B,C]] = ABC - ACB - BCA + CBA$$

and then cyclically permute. The bracket notation for Lie algebras comes from this earlier use of brackets for commutators.

Example A9.4. More generally (and more abstractly), suppose V is any vector space, and consider the set $\mathrm{End}(V) := L(V, V)$ of linear endomorphisms (self-maps) on V. Then the commutator

$$[A, B] := A \circ B - B \circ A$$

is again a Lie product.

Now consider the set $\mathcal{X}(M)$ of smooth vector fields on a manifold M^m. As we have observed, this is an (infinite-dimensional) vector space over $\mathbb{R}$ and indeed a module over $C^\infty(M)$, where for $X, Y \in \mathcal{X}$ and $f \in C^\infty$ the vector field $fX + Y$ is defined pointwise:

$$(fX + Y)_p = f(p)X_p + Y_p.$$

But we also recall that a vector field $X \in \mathcal{X}$ gives (or indeed can be viewed as) a map $C^\infty(M) \to C^\infty(M)$ via $f \mapsto Xf$, taking directional derivatives of f in the directions X_p. That is, we can view vector fields as endomorphisms of $C^\infty(M)$:

$$\mathcal{X}(M) \subset \mathrm{End}(C^\infty M).$$

As Lie observed, the commutator product on $\mathrm{End}(C^\infty M)$ in fact restricts to the subspace $\mathcal{X}$:

Theorem A9.5. *The space of vector fields $\mathcal{X}(M)$ is a Lie algebra with the Lie bracket*

$$[X, Y]f := X(Yf) - Y(Xf).$$

Note that the ordinary composition product does not restrict: the mapping $f \mapsto X(Yf)$ is an endomorphism of $C^\infty(M)$ which should be thought of as taking a second derivative of f in particular directions; this does not correspond to a vector field, because second derivatives do not satisfy the Leibniz product rule. But partial derivative commute; in the commutator above the second-order terms cancel, leaving only first-order terms, that is, a vector field $[X, Y]$. To understand why there can be first-order terms remaining, recall the formula for the second derivative of a function f along a curve γ in $\mathbb{R}^n$ passing through $\gamma(0) = p$ with velocity $\gamma'(0) = v$ and acceleration $\gamma''(0) = a$:

$$\left.\frac{d^2}{dt^2}\right|_{t=0} f(\gamma(t)) = D_p^2 f(v, v) + D_p f(a).$$

Proof. We need to check that the endomorphism

$$[X, Y]\colon C^\infty(M) \to C^\infty(M)$$

is a vector field, that is, that it is local and satisfies the Leibniz product rule. But locality – the fact that the value of $[X, Y]f$ at p depends only on the germ of f at p and not on its values elsewhere – is clear, giving

$$[X, Y]_p\colon C^\infty(p) \to \mathbb{R}, \qquad f \mapsto X_p(Yf) - Y_p(Xf).$$

To show $[X, Y]_p \in T_pM$ we now check the Leibniz rule:

$$\begin{aligned}
[X, Y]_p(fg) &= X_p(Y(fg)) - Y_p(X(fg)) \\
&= X_p(fYg + gYf) - Y_p(fXg + gXf) \\
&= (X_pf)(Y_pg) + f(p)X_p(Yg) \\
&\quad + (X_pg)(Y_pf) + g(p)X_p(Yf) \\
&\quad - (Y_pf)(X_pg) - f(p)Y_p(Xg) \\
&\quad - (Y_pg)(X_pf) - g(p)Y_p(Xf) \\
&= f(p)[X, Y]_pg + g(p)[X, Y]_pf
\end{aligned}$$ □

Exercise A9.6. Of course the Lie bracket $[X, Y]$ is $\mathbb{R}$-bilinear, but it is not $C^\infty(M)$-bilinear. Instead we have

$$[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X.$$

(Hint: consider first the case $[fX, Y] = f[X, Y] - (Yf)X$ and then use that twice, with the antisymmetry.)

Lemma A9.7. *Let $\partial_i \in \mathcal{X}(U)$ denote as usual the coordinate basis vector fields on a chart (U, φ). Then their Lie brackets vanish: $[\partial_i, \partial_j] = 0$ for all $i, j = 1, \ldots, m$.*

Proof. Let f be a germ at $p \in U$, and write $\hat f$ for the germ $f \circ \varphi^{-1}$ at $\varphi(p) \in \mathbb{R}^m$. By definition of ∂_i, we have $(\partial_i f)(p) = (\partial \hat f/\partial x^i)(\varphi(p))$. Then for the Lie bracket we get:

$$\begin{aligned}
[\partial_i, \partial_j]_p f &= \partial_i(\partial_j f)(p) - \partial_j(\partial_i f)(p) \\
&= \frac{\partial^2 \hat f}{dx^i\, dx^j}(\varphi(p)) - \frac{\partial^2 \hat f}{dx^j\, dx^i}(\varphi(p)) = 0,
\end{aligned}$$

using the fact that the mixed partials commute. □

Exercise A9.8. Using this lemma and the result of the previous exercise, compute the formula for the Lie bracket $[X, Y]$ in coordinates, if $X = \sum \alpha^i \partial_i$ and $Y = \sum \beta^i \partial_i$ in a chart (U, φ).

Of course, vector fields are used to take derivatives of functions. If $X \in \mathcal{X}(M)$ and $f \in C^\infty(M)$ then $(Xf)(p) = X_pf$ is a directional derivative of f. If γ is an integral curve of X through p and θ its (local) flow, then

$$X_pf = \left.\frac{d}{dt}\right|_{t=0} f(\gamma_p(t)) = \left.\frac{d}{dt}\right|_{t=0} f(\theta_t(p)).$$

Now suppose we want to take a derivative of a vector field Y along a curve γ. The problem is that for each $q = \gamma(t)$, the vector Y_q lives in a different tangent space $T_q(M)$. So we cannot compare these vectors or ask for their rate of change along γ without some sort of additional information. Later in the course, we will introduce the notion of a "connection" (for instance coming from a Riemannian metric), which does allow us to differentiate a vector field along a curve. But Lie suggested a different approach. Suppose we have not just one (integral) curve γ_p but a whole vector field X and its associated local flow around p. Then we can use

$$D_p\theta_t = \theta_{t*}\colon T_pM \to T_{\theta_t p}M$$

to identify the tangent spaces along the integral curve γ_p. In particular, for each t (in the interval $(-\varepsilon,\varepsilon)$ of definition) we have $\theta_{-t*}(Y_{\theta_t p}) \in T_pM$.

Definition A9.9. If $X, Y \in \mathcal{X}(M)$ then the *Lie derivative* L_XY of Y with respect to X is the vector field defined by

$$(L_XY)_p := \frac{d}{dt}\bigg|_{t=0} \theta_{-t*}(Y_{\theta_t p}) \in T_pM.$$

It would be straightforward but tedious to check in coordinates that this is a smooth vector field. For us, that will follow from the theorem below, saying that the Lie derivative is nothing other than the Lie bracket:

$$L_XY = [X,Y] = -[Y,X] = -L_YX.$$

For this we first need the following lemma, a modification of the Taylor-type lemma we used to prove $T_p\mathbb{R}^m \cong \mathbb{R}^m$.

Lemma A9.10. *Suppose a vector field $X \in \mathcal{X}(M)$ has local flow $\theta\colon (-\varepsilon,\varepsilon)\times V \to M$ around $p \in M$. Given any $f \in C^\infty(M)$, there exists a smooth function $g\colon (-\varepsilon,\varepsilon)\times V \to \mathbb{R}$, which we write as $(t,q) \mapsto g_t(q)$, such that*

$$f(\theta_t(q)) = f(q) + tg_t(q), \qquad X_qf = g_0(q).$$

Proof. First we define $h_t(q) := \frac{d}{dt} f(\theta_t(q)) = X_{\theta_t q}f$, and then we set $g_t(q) := \int_{s=0}^1 h_{st}(q)\,ds$. For $t = 0$ this clearly means $g_0(q) = h_0(q) = X_qf$. Using a change of variables and the fundamental theorem of calculus, for arbitrary t we get $tg_t(q) = f(\theta_t(q)) - f(q)$ as desired. □

Theorem A9.11. *For any vector fields X, Y on M, the Lie derivative and Lie bracket coincide, that is, we have $L_XY = [X,Y]$.*

Proof. Suppose $f \in C^\infty(p)$ is a germ at $p \in M$. We want to show $(L_XY)_pf = [X,Y]_pf$. Choose a representative $f \in C^\infty(U)$ for the germ and use the lemma (applied to the manifold U) to find g_t such that $g_0 = Xf$ and $f\circ\theta_t = f + tg_t$, or negating t as we will, $f\circ\theta_{-t} = f - tg_{-t}$. Then starting from the definition of L_XY we find

$$\begin{aligned}
(L_XY)_pf &= \lim_{t\to 0}\frac{1}{t}\Big((\theta_{-t*}Y_{\theta_t p})f - Y_pf\Big)\\
&= \lim_{t\to 0}\frac{1}{t}\Big(Y_{\theta_t p}(f\circ\theta_{-t}) - Y_pf\Big)\\
&= \lim_{t\to 0}\frac{1}{t}\Big(Y_{\theta_t p}(f - tg_{-t}) - Y_pf\Big)\\
&= \frac{d}{dt}\bigg|_{t=0}(Y_{\theta_t p}f) - \lim_{t\to 0} Y_{\theta_t p}g_{-t}\\
&= \frac{d}{dt}\bigg|_{t=0}(Yf)(\theta_t(p)) - Y_pg_0\\
&= X_p(Yf) - Y_p(Xf) = [X,Y]_pf
\end{aligned}$$ □

Exercise A9.12. Suppose $f\colon M^m \to N^n$ is a smooth map and $Y \in \mathcal{X}(N)$ is f-related to $X \in \mathcal{X}(M)$ while Y' is f-related to X'. Then $[Y,Y']$ is f-related to $[X,X']$.

End of Lecture 9 Nov 2015

A10. Vector bundles

When we defined the tangent bundle TM of a manifold M^m, we mentioned that it is a specific example of a smooth vector bundle over M, with fibers the tangent spaces T_pM. In general, a bundle over a *base space* M consists of a *total space* E with a projection $\pi\colon E \to M$. When there is no confusion, we often refer to E as the bundle. The *fiber* over $p \in M$ is simply the preimage $E_p := \pi^{-1}\{p\}$. If $S \subset M$, we write E_S for the *restriction* of E to S, that is, the bundle $\pi|_S\colon E_S = \pi^{-1}(S) \to S$. If $\pi\colon E \to M$ and $\pi'\colon E' \to M$ are two bundles, then $\varphi\colon E \to E'$ is *fiber-preserving* if $\pi'\circ\varphi = \pi$, that is, if $\varphi(E_p) \subset E'_p$ for each p.

Of course, we only call $\pi\colon E \to M$ a bundle if the fibers are all isomorphic (in an appropriate sense, to some F); and if E is locally trivial, locally looking like a product with F. We now give a precise definition for the case of interest here.

Definition A10.1. A *(smooth) vector bundle* of *rank k* is a map $\pi\colon E^{m+k} \to M^m$ of manifolds such that

- each fiber is a vector space of dimension k,
- each point in M has a *trivializing neighborhood* U, meaning there is a fiber-preserving diffeomorphism $E_U \to U\times\mathbb{R}^k$ that is a vector space isomorphism on each fiber.

Exercise A10.2. If $M^m \subset N^n$ is a submanifold and $\pi\colon E \to N$ is a vector bundle of rank k over N, then the restriction E_M is a vector bundle of rank k over M.

Remark A10.3. The tangent bundle TM is a rank-m bundle over M^m. Any coordinate chart (U,φ) is a trivializing neighborhood where the fiber-preserving diffeomorphism is

$$(\pi, D\varphi)\colon X_p \mapsto (p, D_p\varphi(X_p)).$$

The restriction $D_p\varphi\colon T_pM \to \mathbb{R}^m$ to each fiber is indeed linear.

Suppose $\pi\colon E \to M$ is a vector bundle. We can cover M by open sets U that are (small enough to be) both coordinate charts for the manifold M and trivializing neighborhoods for the bundle E. That is, we have diffeomorphisms $\varphi\colon U \to \varphi(U) \subset \mathbb{R}^m$ and $\psi\colon E_U \to U\times\mathbb{R}^k$, the latter being fiber-preserving and linear on each fiber. Composing these gives a diffeomorphism

$$(\varphi, \mathrm{id})\circ\psi\colon E_U \to \varphi(U)\times\mathbb{R}^k \subset \mathbb{R}^{m+k},$$

which is a coordinate chart for the manifold E.

Definition A10.4. A *section* of a vector bundle $\pi\colon E \to M$ is a smooth map $\sigma\colon M \to E$ such that $\pi\circ\sigma = \mathrm{id}_M$, meaning $\sigma(p) \in E_p$ for each $p \in M$. The space of all (smooth) sections is denoted $\Gamma(E)$.

Vector fields, for example, are simply sections of the tangent bundle: $\mathcal{X}(M) = \Gamma(TM)$. As in that example, $\Gamma(E)$ is always a module over $C^\infty(M)$, using pointwise addition and scalar multiplication: if $\sigma, \tau \in \Gamma(E)$ and $f \in C^\infty M$, then $\sigma + f\tau$ is defined pointwise by

$$(\sigma + f\tau)_p = \sigma_p + f(p)\tau_p \in E_p.$$

Sometimes we talk about local sections $\sigma \in \Gamma(E_U)$ that are not defined globally on all of M but only on a subset $U \subset M$. A trivialization of E over U is equivalent to a *frame*, that is, a set of k sections $\sigma_i \in \Gamma(E_U)$ such that at each $p \in U$, the $\sigma_i(p)$ form a basis for E_p.

Operations on vector spaces yield corresponding operations (acting fiberwise) on vector bundles. For instance, if $E \to M$ and $F \to M$ are two vector bundles over M (of ranks k and l, respectively) then their *direct sum* (or *Whitney sum*) $E \oplus F \to M$ is a vector bundle of rank $k+l$, where we have $(E \oplus F)_p = E_p \oplus F_p$ fiberwise. Any neighborhood which trivializes both E and F will trivialize their sum.

A11. Dual spaces and one-forms

We next turn to various constructions on a single vector space V. Even though much of what we say could extend to arbitrary spaces, we assume V is a real vector space of finite dimension k; later it will be a tangent space to a manifold.

The *dual space* $V^* := L(V, \mathbb{R})$ is defined to be the space of all linear functionals $V \to \mathbb{R}$ (also called *covectors*). The dual space V^* is also k-dimensional.

As we have mentioned before, a linear map $L\colon V \to W$ induces a dual linear map $L^*\colon W^* \to V^*$ in the opposite direction, defined naturally by $(L^*\sigma)(v) = \sigma(Lv)$. This construction is functorial in the sense that $\mathrm{id}^* = \mathrm{id}$ and $(L \circ L')^* = (L')^* \circ L^*$. One can check that L is surjective if and only if L^* is injective, and vice versa.

While there is no natural isomorphism $V \to V^*$, any basis $\{e_1, \ldots, e_k\}$ for V determines a *dual basis* $\{\omega^1, \ldots, \omega^k\}$ for V^* by setting

$$\omega^i(e_j) = \delta^i_j = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

The covector ω^i is the functional that gives the i^{th} component of a vector in the basis $\{e_i\}$, that is, $v = \sum \omega^i(v)e_i$, which we could also write as $v^i = \omega^i(v)$. Similarly, $\sigma = \sum \sigma(e_i)\omega_i$.

There is a natural isomorphism $V \to V^{**}$, where $v \in V$ induces the linear functional $\sigma \mapsto \sigma(v)$ on V^*. Given a linear map $L\colon V \to W$, we have $L^{**} = L$.

Applying duality to each fiber E_p of a vector bundle gives the *dual bundle* E^* with $(E^*)_p = (E_p)^*$. It is trivialized over any trivializing neighborhood for E, as one sees by choosing a frame and taking the dual frame.

Applying duality to T_pM gives the *cotangent space* T_p^*M to M at p. These fit together to form the *cotangent bundle*, which (just like the tangent bundle) is trivialized over any coordinate neighborhood.

A (smooth) section ω of the cotangent bundle is called a *covector field* or more often a *(differential) one-form*. We write $\Omega^1(M) = \Gamma(T^*M)$ for the space of all sections. A one-form $\omega \in \Omega^1(M)$ acts on a vector field $X \in \mathcal{X}(M)$ to give a smooth function $\omega(X) \in C^\infty(M)$ via $(\omega X)(p) = \omega_p(X_p) \in \mathbb{R}$. In a coordinate chart (U, φ) we have the basis vector fields ∂_i; taking the dual basis pointwise we get the basis one-forms dx^i satisfying $dx^i(\partial_j) = \delta^i_j$. Any one-form $\sigma \in \Omega^1(U)$ can be written as $\sigma = \sum \sigma_i dx^i$, where the components $\sigma_i = \sigma(\partial_i) \in C^\infty U$ are smooth functions.

An important way to construct one-forms is as the differentials of functions. If $f \in C^\infty M$ then $D_p f\colon T_pM \to T_{f(p)}\mathbb{R}$, under the identification $T_t\mathbb{R} \cong \mathbb{R}$, can be thought of as a covector df_p at p. If $X \in \mathcal{X}(M)$ is a vector field, then $df(X) = Xf$, meaning $df_p(X_p) = X_p f$ for each $p \in M$. While $X_p f$ depends on the germ of f at p, it only depends on df at p: the covector df_p encodes exactly all the directional derivatives of f at p.

Note that the notation dx^i we used above for the coordinate basis one-forms in a coordinate chart (U, φ) is consistent: these are indeed the differentials of the coordinate component functions $x^i := \pi^i \circ \varphi\colon U \to \mathbb{R}$. In coordinates, we have $df = \sum(\partial_i f)dx^i$, where we recall that $\partial_i f$ is the i^{th} partial derivative of the coordinate expression $f \circ \varphi^{-1}$ of f.

An interesting example of a one-form is the form $d\theta$ on $\mathbb{S}^1$. We cover $\mathbb{S}^1$ with coordinate charts of the form

$$(\cos\theta, \sin\theta) \mapsto \theta \in (\theta_0, \theta_0 + 2\pi).$$

On each such chart, we write $d\theta$ for the coordinate basis one-form (which would also be called dx^1); then we note that on the overlaps, these forms $d\theta$ agree (independent of the omitted point θ_0).

Unlike for most manifolds, the tangent bundle $T\mathbb{S}^1$ is (globally) trivial; thus $T^*\mathbb{S}^1$ is also trivial. Any one-form is written as $f d\theta$ for some smooth function f. The form $d\theta$ can be thought of as dual to the vector field $(-\sin\theta, \cos\theta)$ on $\mathbb{S}^1$, which is ∂_1 in any of the charts above.

The notation $d\theta$ is slightly confusing, since this one-form is not globally the differential of any smooth function on $\mathbb{S}^1$. Thus on $\mathbb{S}^1$, this is a one-form which is "closed" but not "exact", meaning that $d\theta$ looks like a differential locally but not globally. This shows, in a sense we may explore later, that the space $\mathbb{S}^1$ has nontrivial "first cohomology", that is, that it has a one-dimensional loop.

One forms are in some sense similar to vector fields, but we will see later how they (as well as differential forms of higher degree) are often more convenient. This is mainly because, while a map $f\colon M \to N$ does not in general act on vector fields, it can be used to pull one-forms on N back to M. To see this, note that

$$f_* = D_p f\colon T_pM \to T_{f(p)}N$$

at each $p \in M$ induces a dual map

$$f^* := (D_p f)^*\colon T^*_{f(p)}N \to T^*_pM.$$

Given $\omega \in \Omega^1(N)$, we define $f^*\omega \in \Omega^1(M)$ by $(f^*\omega)_p = f^*(\omega_{f(p)}) \in T^*_pM$.

A special case is the restriction $\omega|_M$ of a form $\omega \in \Omega^1 N$ to a submanifold $M^m \subset N^n$, which is simply the pullback under the inclusion map. For any $X_p \in T_pM \subset T_pN$ at any $p \in M \subset N$ we of course simply have $\omega|_M(X_p) = \omega(X_p)$.

A12. Bilinear forms and Riemannian metrics

A bilinear map $b\colon V \times V \to \mathbb{R}$ is called a *bilinear form* on V. If $\{e_1, \dots, e_k\}$ is a basis for V, then b is given by $b_{ij} = b(e_i, e_j)$: if $v = \sum v^i e_i$ and $w = \sum w^i e_i$ then $b(v, w) = \sum_{i,j} b_{ij} v^i w^j$. The bilinear form b is called *symmetric* if $b(w, v) = b(v, w)$ (i.e., if $b_{ij} = b_{ji}$) and *antisymmetric* or *alternating* if $b(w, v) = -b(v, w)$ (i.e., if $b_{ij} = -b_{ji}$). Any bilinear form b can be uniquely decomposed $b = b^+ + b^-$ into a symmetric part b^+ and an antisymmetric part b^-, defined by $2b^\pm(v, w) = b(v, w) \pm b(w, v)$. Given a linear map $L\colon V \to W$, we can pull back any bilinear form b on W to a bilinear form L^*b on V, defined by $(L^*b)(v, v') = b(Lv, Lv')$.

We will consider antisymmetric differential forms later. For now we restrict attention to symmetric bilinear forms on V. These are in one-to-one correspondance with quadratic forms $q\colon V \to \mathbb{R}$: of course $q(v) := b(v, v)$ depends only on the symmetric part of b and we can recover b from q via the formula

$$2b(v, w) := q(v + w) - q(v) - q(w).$$

(Note that much of our discussion would fail for vector spaces over fields of characteristic 2, where $2 = 0$.)

A symmetric bilinear form b (or the associated quadratic form q) is called *positive semidefinite* if $q(v) = b(v, v) \geq 0$ for all $v \in V$. It is called *positive definite* if $q(v) = b(v, v) > 0$ for all $v \neq 0$. A positive definite form on V is also called an *inner product* (or *scalar product*) on V. An inner product is what we need to define the geometric notions of *length* (or *norm*) $\|v\| := \sqrt{b(v, v)}$ and *angle*

$$\angle(v, w) := \arccos \frac{b(v, w)}{\|v\|\,\|w\|}$$

between vectors in V. The pullback L^*b of a positive definite form is always positive semidefinite, but it is positive definite if and only if L is injective.

Of course the standard example of an inner product is the Euclidean inner product $b(v, w) = \sum v^i w^i$ on $\mathbb{R}^m$.

The quadratic forms on V form a vector space $Q(V)$ of dimension $\binom{m+1}{2}$. The positive definite forms form an open convex cone in this vector space, whose closure consists of all positive semidefinite forms. (A convex cone is a set closed under taking positive linear combinations.)

Again, we can apply this construction to the fibers of any vector bundle E. If E has rank k, then $Q(E)$ has rank $\binom{k+1}{2}$. In case of the tangent bundle TM, we get a vector bundle $Q(TM)$ of rank $\binom{m+1}{2}$. A positive definite section $g \in \Gamma(Q(TM))$ is called a *Riemannian metric* on M. It consists of an inner product $\langle X_p, Y_p \rangle := g_p(X_p, Y_p)$ on each tangent space T_pM, which lets us measure length and angles between tangent vectors at any $p \in M$.

In a coordinate chart (U, φ), the metric g is given by components $g_{ij} := g(\partial_i, \partial_j) \in C^\infty U$ so that

$$g\Big(\sum \alpha^i \partial_i, \sum \beta^j \partial_j\Big) = \sum_{i,j} g_{ij} \alpha^i \beta^j.$$

The matrix (g_{ij}) is of course symmetric and positive definite at each $p \in U$.

The standard Riemannian metric g on the manifold $\mathbb{R}^m$ comes from putting the standard Euclidean inner product on each $T_p\mathbb{R}^m = \mathbb{R}^m$. That is, in the standard chart $(\mathbb{R}^m, \mathrm{id})$ we have $g_{ij} = \delta_{ij}$.

If $f\colon M^m \to N^n$ is a smooth map, then we can pull back sections of $Q(TN)$ to sections of $Q(TM)$ in the natural way:

$$(f^*g)(X_p, Y_p) := g(f_*X_p, f_*Y_p).$$

If g is a Riemannian metric, then of course f^*g will be positive semidefinite at each $p \in M$, but it will be a Riemannian metric if and only if f is an immersion (meaning that D_pf is injective for every $p \in M$, and in particular $m \leq n$). Again, an important special case of this pull-back metric is when f is the inclusion map of a submanifold; then we speak of restricting the Riemannian metric g on N to $g|_M$ on the submanifold $M \subset N$.

In particular, the standard metric on $\mathbb{R}^n$ restricts to give a Riemannian metric on any submanifold $M^m \subset \mathbb{R}^n$. Last semester, we studied the case $m = 2$, $n = 3$, and called this metric $g(v, w) = \langle v, w \rangle$ the first fundamental form.

`End of Lecture 16 Nov 2015`

A13. Partitions of unity

The sections of any vector bundle $E \to M$ form a vector space. In particular, there is always the zero section $\sigma_p = 0 \in E_p$. An interesting question is whether there is a nowhere vanishing section. A trivial bundle $M \times \mathbb{R}^r$ of course has constant nonzero sections. Sometimes it turns out that the tangent bundle TM is trivial – this happens for instance for $\mathbb{S}^1$ or more generally the m-torus T^m. Other times, there is no nonvanishing section of TM. For instance the Hopf index theorem shows this is the case for any closed orientable surface M^2 other than the torus T^2. Over $\mathbb{S}^1$ there is also a nontrivial line bundle, whose total space is topologically a Möbius strip. This bundle has no nonvanishing section.

From this point of view, it might be surprising that the bundle $Q(TM)$ has nonvanishing sections for any manifold M^m, indeed sections which are positive definite everywhere. In other words, any manifold M can be given a Riemannian metric. This follows from the fact that any manifold can be embedded in $\mathbb{R}^n$ for sufficiently large n, or can be proven more directly by taking convex combinations of standard metrics in different coordinate charts. Either of these approaches requires the technical tool of a partition of unity, a collection of locally supported functions whose sum is everywhere one. This gives a general method for smoothly interpolating between different local definitions. We do not want to get into questions of summing infinite sequences; thus we impose a local finiteness condition.

Definition A13.1. A collection $\{S_\alpha\}$ of subsets of a topological space X is called *locally finite* if each $p \in X$ has a neighborhood that intersects only finitely many of the S_α.

Definition A13.2. The *support* $\operatorname{supp} f$ of a function $f: X \to \mathbb{R}^n$ is the closure of the set where $f \neq 0$.

Lemma A13.3. *If $f_\alpha : M \to \mathbb{R}$ are smooth functions such that $\{\operatorname{supp} f_\alpha\}$ is locally finite, then $\sum f_\alpha$ defines a smooth function $M \to \mathbb{R}$.*

Proof. The local finiteness means that each $p \in M$ has a neighborhood U which meets only a finite number of the $\operatorname{supp} f_\alpha$. On U, the sum $\sum f_\alpha$ is thus a sum of a fixed finite collection of smooth functions, hence smooth. □

Note that saying the collection of supports $\operatorname{supp} f_\alpha$ is locally finite is stronger than saying each $p \in M$ is contained in finitely many $\operatorname{supp} f_\alpha$, which in turn is stronger than saying only finitely many f_α are positive at each $p \in M$. This would suffice to evaluate $\sum f_\alpha(p)$ as a finite sum at each p. The stronger conditions ensure that the sum is a smooth function.

Definition A13.4. A (smooth) *partition of unity* on a manifold M is a collection of functions $\psi_\alpha : M \to \mathbb{R}$ such that

- $\psi_\alpha \geq 0$,
- $\{\operatorname{supp} \psi_\alpha\}$ is locally finite,
- $\sum \psi_\alpha \equiv 1$.

A trivial example is the single constant function 1. The interest in partitions of unity comes from examples where the support of each ψ_α is "small" in some prescribed sense.

Definition A13.5. Given an open cover $\{U_\alpha\}$, a partition of unity $\{\psi_\beta\}$ is *subordinate* to the cover $\{U_\alpha\}$ if for each β there exists $\alpha = \alpha(\beta)$ such that the support $\operatorname{supp} \psi_\beta$ is contained in U_α.

Note that if we want, we can then define a new partition of unity $\{\overline{\psi}_\alpha\}$ also subordinate to $\{U_\alpha\}$ and now indexed by the same index set. Simply set $\overline{\psi}_\alpha$ to be the sum of those ψ_β for which $\alpha = \alpha(\beta)$. (Note that this is not the sum of all ψ_β supported in U_α.)

To give the flavor of results about partitions of unity, we start with the easy case of a compact manifold. All manifolds have a related property called paracompactness, which will be enough to extend this result.

Proposition A13.6. *Given any open cover $\{U_\alpha\}$ of a compact manifold M^m, there exists a partition of unity subordinate to this cover.*

Proof. For each $p \in M$, we have $p \in U_\alpha$ for some $\alpha = \alpha(p)$, and we can choose a smooth nonnegative function f_p supported in U_α with $f_p > 0$ on some neighborhood $V_p \ni p$. Since M is compact, a finite subcollection $\{V_{p_1}, \ldots, V_{p_k}\}$ covers M. Then $f := \sum_i f_{p_i}$ is a positive smooth function on M, so we can define a finite collection of smooth functions $\psi_i := f_{p_i}/f$. These form a (finite) partition of unity, subordinate to the given cover. □

Definition A13.7. An open cover $\{V_\beta\}$ is a *refinement* of another open cover $\{U_\alpha\}$ if each V_β is contained in some U_α. A space X is *paracompact* if every open cover has a locally finite refinement.

Example A13.8. The cover $\{(-n, n) : n \in \mathbb{N}^+\}$ of $\mathbb{R}$ is not locally finite, but any covering of $\mathbb{R}$ by bounded open sets is a refinement, so for instance $\{(k-1, k+1) : k \in \mathbb{Z}\}$ is a locally finite refinement.

A standard result in point set topology says that any second countable, locally compact Hausdorff space is paracompact. It is also true that any metric space is paracompact. Some authors replace "second countable" by "paracompact" in the definition of manifold, which makes no difference except for allowing uncountably many components. (The long line, for instance, is not paracompact.) It is known that a topological space X admits a continuous partition of unity subordinate to any given open cover if and only if X is paracompact and Hausdorff. We will, however, explicitly prove what we need for manifolds.

Lemma A13.9. *Every manifold M has a countable base consisting of coordinate neighborhoods with compact closure.*

Proof. Start with any countable base $\{B_i\}$ and let $\mathcal{B}$ be the subcollection of those B_i that are contained in some coordinate neighborhood and have compact closure. Now suppose we are given an open subset $W \subset M$ and a point $p \in W$. Choose a coordinate chart (U, φ) around p such that

$$U \subset W, \qquad \varphi(p) = 0, \qquad B_2(0) \subset \varphi(U)$$

and set $V := \varphi^{-1}(B_1(0))$. Then V has compact closure. Since $\{B_i\}$ is a base, we have $p \in B_i \subset V \subset U$ for some i. But then this B_i is also contained with compact closure in the coordinate neighborhood U; thus $B_i \in \mathcal{B}$. Since p and U were arbitrary, this shows $\mathcal{B}$ is a base. □

Lemma A13.10. *Every manifold M has a "compact exhaustion", indeed a nested family of subsets*

$$\varnothing \neq W_1 \subset \overline{W}_1 \subset W_2 \subset \overline{W}_2 \subset \cdots$$

with W_k open and $\overline{W}_k$ compact, whose union is M.

Proof. Choose a countable base $\{B_i\}$ as in the last lemma. We will choose $1 = i_1 < i_2 < \cdots$ and set $W_k := \bigcup_{i=1}^{i_k} B_i$. These automatically have compact closure and are nested. We just need to choose each i_k large enough that $W_k \supset \overline{W}_{k-1}$. But this is possible, since $\overline{W}_{k-1}$ is compact and thus covered by some finite collection of the B_is. Finally, since $i_k \geq k$ we have that $W_k \supset B_k$ so $\bigcup W_k \supset \bigcup B_k = M$. □

Corollary A13.11. *Given a manifold M, we can find countable families of subsets $K_i \subset O_i \subset M$, for $i \in \mathbb{N}^+$, with K_i compact, O_i open, $\bigcup K_i = M$, and $\{O_i\}$ locally finite.*

Proof. Using the nested W_i from the lemma, simply set $K_i := \overline{W}_i \setminus W_{i-1}$ and $O_i := W_{i+1} \setminus \overline{W}_{i-2}$ (where we take $W_0 = \varnothing = W_{-1}$). The local finiteness follows from the fact that any p is contained in some $W_{j+1} \setminus \overline{W}_{j-1}$, which meets only four of the O_i. □

Corollary A13.12. *Any manifold M is paracompact.*

Proof. Suppose $\{U_\alpha\}$ is an open cover. Choose O_i and K_i as in the last corollary. For each i, the compact set K_i is covered by the sets $O_i \cap U_\alpha$, and thus by a finite subcollection, which we name O_i^j for $j = 1, \ldots, k_i$. The union of these finite collections, over all i, is a locally finite refinement. □

We are now set up to adapt the construction of partitions of unity from the compact case to the general case.

Theorem A13.13. *Given any covering $\{U_\alpha\}$ of a manifold M, there exists a partition of unity $\{\psi_i\}$ subordinate to this covering.*

Proof. Find $K_i \subset O_i$ as above. Fixing i, for each $p \in K_i$ we have $p \in U_\alpha \cap O_i$ for some $\alpha = \alpha(p)$. Choose a smooth nonnegative f_p with support in $U_\alpha \cap O_i$ such that $f_p > 0$ on some neighborhood $V_p \ni p$. Finitely many of these neighborhoods cover the compact set K_i. Now letting i vary, we have a countable family of bump functions f_i^j, whose supports form a locally finite family. Thus dividing by their well-defined, positive, smooth sum gives a partition of unity. □

Note that if $K \subset M$ is compact, then for any partition of unity $\{\psi_i\}$ for M, only finitely many ψ_i have support meeting K. (Each $p \in K$ has a neighborhood meeting only finitely many $\operatorname{supp}\psi_i$; by compactness finitely many such neighborhoods cover K.)

Now we turn to some applications of these ideas. Note that if $\{\psi_\alpha\}$ is a partition of unity subordinate to a cover $\{U_\alpha\}$ and we have functions $f_\alpha \in C^\infty(U_\alpha)$, then $\psi_\alpha f_\alpha$ defines a smooth function on M supported in U_α. Then $\sum \psi_\alpha f_\alpha$ makes sense as a locally finite sum of smooth functions. The same works for sections of any vector bundle $E \to M$: local sections $\sigma_\alpha \in \Gamma(E_{U_\alpha})$ can be combined to get a global section $\sum \psi_\alpha \sigma_\alpha \in \Gamma(E)$.

Theorem A13.14. *Any manifold M^m admits a Riemannian metric.*

Proof. Let g_0 denote the standard (flat) Riemannian metric on $\mathbb{R}^m$. Let $\{(U_\alpha, \varphi_\alpha)\}$ be an atlas for M. In each chart the pullback $g_\alpha := \varphi_\alpha^*(g_0)$ is a Riemannian metric on U_α. Now let $\{\psi_\alpha\}$ be a partition of unity subordinate to $\{U_\alpha\}$. We can consider each $\psi_\alpha g_\alpha$ as a global section of $Q(TM)$, supported of course in U_α. Then $\sum \psi_\alpha g_\alpha$ is a Riemannian metric on M, since locally near any point it is a convex combination of finitely many Riemannian metrics g_α. □

An alternative proof of the existence of Riemannian metrics simply uses the fact that any manifold M^m can be embedded in $\mathbb{R}^n$ for large enough n.

It is not too hard to show that $n = 2m+1$ actually suffices – a generic projection from higher dimensions to dimension $2m + 1$ will still give an embedding – but we omit such discussions. Harder is the *Whitney trick* used to get down to $n = 2m$. Most manifolds actually embed in $\mathbb{R}^{2m-1}$ – the only exceptions (besides S^1) are closed nonorientable manifolds of dimension $m = 2^k$, like closed nonorientable surfaces.

We will restrict to compact manifolds and not attempt to get an optimal n. Rather than using a partition of unity directly, we will repeat the easy proof of the compact case, using some of the functions involved in the construction directly.

The theorems are also true in the noncompact case; the proof uses decompositions like the K_i and O_i above but requires knowing that each compact piece can be embedded in the same dimension, say in $\mathbb{R}^{2m+1}$.

Theorem A13.15. *Any compact manifold M^m can be embedded in some Euclidean space $\mathbb{R}^n$.*

Proof. For each point $p \in M$, find a nonnegative function $f_p \colon M \to \mathbb{R}$ with $f \equiv 1$ in some neighborhood $V_p \ni p$ and with support in a coordinate chart (U_p, φ_p). By compactness, a finite number of the V_p suffice to cover M. Call these points $p_1, \ldots, p_k$ and simply use the indices $1, \ldots k$ for the associated objects. Define a map $g \colon M \to \mathbb{R}^{k(m+1)}$ as follows:

$$g(p) := \big(f_1(p)\varphi_1(p), \ldots, f_k(p)\varphi_k(p), f_1(p), \ldots, f_k(p)\big).$$

On V_i we have $f_i \equiv 1$, so the i^{th} "block" in g equals φ_i, with injective differential. Thus g is an immersion on each V_i, thus on all of M. By compactness, it only remains to show that g is injective. If $g(p) = g(q)$ then in particular we have $f_i(p) = f_i(q)$ for all i. Choose i such that $p \in V_i \subset U_i$. Then $f_i(q) = 1$ implies $q \in U_i$. But then we also have

$$\varphi_i(p) = f_i(p)\varphi_i(p) = f_i(q)\varphi_i(q) = \varphi_i(q).$$

Since φ_i is injective on U_i, it follows that $p = q$. □

A14. Riemannian manifolds as metric spaces

We fix a Riemmanian manifold (M, g), that is, a smooth manifold M^m with a fixed Riemannian metric g. Where convenient, we write $\langle X_p, Y_p\rangle := g_p(X_p, Y_p)$ for the inner product and $\|X_p\| := \sqrt{\langle X_p, X_p\rangle}$ for the length of a tangent vector.

Definition A14.1. Suppose $\gamma \colon [a, b] \to M$ is a piecewise smooth curve. The *length* of γ (with respect to the Riemanian metric g) is

$$\operatorname{len}(\gamma) := \int_a^b \|\gamma'(t)\|\, dt.$$

Note that by the chain rule, this length is invariant under reparametrization. The arclength function along γ is

$$s(t) := \operatorname{len}(\gamma|_{[a,t]}) = \int_a^t \|\gamma'(t)\|\, dt.$$

and – assuming γ is a piecewise immersion – we can reparametrize γ by arclength so that $\|\gamma'\| \equiv 1$.

Note that if the standard Riemannian metric on $\mathbb{R}^n$ is restricted to a submanifold M^m, then the length of a curve γ in M as defined above is the same as its length in $\mathbb{R}^n$ as considered last semester.

Definition A14.2. The *distance* between two points $p, q \in M$ is the infimal length

$$d(p, q) := \inf_{\gamma} \operatorname{len}(\gamma)$$

taken over all piecewise smooth curves γ in M from p to q.

Note that we could easily apply our definition of length to more general curves, say to all rectifiable or Lipschitz curves. Since any curve can be smoothed, in the infimum defining d it is not important whether we allow all rectifiable curves or restrict to smooth curves. We have chosen an option in the middle.

`End of Lecture 23 Nov 2015`

Note that (no matter which smoothness class is chosen) the infimum is not always realized, as one sees for instance if $M = \mathbb{R}^2 \smallsetminus \{0\}$.

The theorem below will show that (M, d) is a metric space compatible with the given topology on M. Of course when M is not connected, points p and q in different components are not connected by any path, so $d(p, q) = +\infty$ by the above definition. It is easiest to use a definition of metric spaces that allows infinite distance. If this is not desired, the following discussion should be restricted to connected manifolds. Note that a connected component of a manifold is automatically path-connected; any pair of points can actually be joined by a smooth path, whose length is necessarily finite.

A few properties of d are immediate. The constant path shows $d(p, p) = 0$. The inverse path shows $d(p, q) = d(q, p)$. Concatenating paths gives the triangle inequality $d(p, r) \le d(p, q) + d(q, r)$. (This is one reason we chose to allow piecewise smooth paths.) That is, we see easily that d is a pseudometric, and to see it is a metric we just need to show that $d(p, q) = 0$ holds only for $p = q$.

Lemma A14.3. *Consider $\mathbb{R}^m$ with the standard Riemannian metric. Then $d(p, q) = \|p - q\|$.*

Proof. Since both sides are clearly translation invariant, it suffices to consider $q = 0$. It is easy to compute the length of the straight path from p to 0 as $\|p\|$. We must show no other path has less length (and may assume $p \ne 0$). So suppose $\gamma(0) = p$ and $\gamma(1) = 0$. We may assume $\gamma(t) \ne 0$ for $t < 1$, since otherwise we replace γ by $\gamma|_{[0,t]}$, which is not longer. Thus for $t < 1$ we can write $\gamma(t) = r(t)\beta(t)$ where $\|\beta(t)\| = 1$ and $r(t) > 0$. We have $r(0) = \|p\|$ and $r(t) \to 0$ as $t \to 1$. Since $\langle \beta, \beta \rangle \equiv 1$, we get $\langle \beta, \beta' \rangle \equiv 0$. The product rule $\gamma' = r'\beta + r\beta'$ then gives

$$\|\gamma'\|^2 = |r'|^2\|\beta\|^2 + r^2\|\beta'\|^2 \ge |r'|^2.$$

Thus

$$\int_0^1 \|\gamma'\|\, dt \ge \int_0^1 |r'|\, dt \ge \left|\int_0^1 r'\, dt\right| = r(0) - r(1) = \|p\|$$

as desired. □

Two norms on a vector space induce the same topology if and only if they are equivalent in the sense that they differ by at most a constant factor. For finite dimensional vector spaces, all norms are equivalent. We sketch a proof of the case we need.

Lemma A14.4. *Any two inner products on $\mathbb{R}^m$ induce equivalent norms.*

Proof. Let $\|v\|$ denote the standard Euclidean norm, and let $\sum g_{ij} v^i w^j$ denote an arbitrary inner product on $\mathbb{R}^m$. On the compact unit sphere $\mathbb{S}^{m-1} = \{v : \|v\| = 1\}$ the other norm $\sqrt{\sum g_{ij} v^i v^j}$ achieves its minimum $c > 0$ and its maximum C. Then by homogeneity, we have

$$c\|v\| \le \sqrt{\sum g_{ij} v^i v^j} \le C\|v\|$$

for all v, as desired. We note that the optimal constants depend smoothly on g. □

Corollary A14.5. *Suppose g is a Riemannian metric on an open set $U \subset \mathbb{R}^m$ and $K \subset U$ is compact. Then there exist constants $0 < c \le C$ such that*

$$c\|v\| \le \sqrt{g(v, v)} \le C\|v\|$$

for all $p \in K$ and all $v \in T_pU = T_p\mathbb{R}^m \cong \mathbb{R}^m$. In particular, for any curve γ in K from p to q we have

$$c\|p - q\| \le c \operatorname{len}_0 \gamma \le \operatorname{len}_g \gamma \le C \operatorname{len}_0 \gamma,$$

where len_0 is the length relative to the standard Euclidean metric and len_g is the length relative to g.

Proof. For each $p \in K$ the lemma gives us $c_p \le C_p$. Assuming we choose the optimal constants at each point, they depend smoothly on g_p thus smoothly on p. By compactness we can set $c := \min_K c_p$ and $C := \max_K C_p$. Integrating the bounds for tangent vectors gives the final statement for any curve γ. □

This corollary gives the key uniformity needed for the following theorem.

Theorem A14.6. *Let (M, g) be a Riemannian manifold. With the distance function d above, it is a metric space (M, d). The metric topology agrees with the given manifold topology on M.*

Proof. We have noted that d is symmetric and satisfies the triangle inequality. We must prove $d(p, q) = 0 \implies p = q$ and show that the topologies agree.

Let D denote the closed unit ball in $\mathbb{R}^m$. Given $p \ne q$ in M, we can find coordinates (U, φ) around p such that $\varphi(p) = 0$, $\varphi(U) \supset D$ and $q \notin \varphi^{-1}(D)$. On D use the last corollary to get c, C comparing $(\varphi^{-1})^* g$ with the standard metric. Any path from p to q must first leave $\varphi^{-1}D$. Its g-length is at least the g-length of this initial piece, which is the ${\varphi^{-1}}^* g$-length of its image α. Since α connects 0 to ∂D, it has Euclidean length at least 1, so ${\varphi^{-1}}^* g$-length at least c. Since this is true for any γ, we find $d(p, q) \ge c > 0$.

We have just seen that a Euclidean ball in a coordinate chart contains a small metric ball. Thus open sets in the manifold topology are open in the metric topology. To get the converse, consider again coordinates around p with $D \subset \varphi(U)$, and find C as in the corollary. If $\varphi(q)$ is in the ε-ball around $0 = \varphi(p)$, then they can be joined by a path of Euclidean length less than $\varepsilon < 1$. Thus p and q can be joined by a path of g-length less than εC. □

B. Differential Forms

We have already seen one-forms (covector fields) on a manifold. In general, a k-form is a field of alternating k-linear forms on the tangent spaces of a manifold. Forms are the natural objects for integration: a k-form can be integrated over an oriented k-submanifold. We start with tensor products and the exterior algebra of multivectors.

B1. Tensor products

Recall that, if V, W and X are vector spaces, then a map $b\colon V \times W \to X$ is called *bilinear* if

$$b(v + v', w) = b(v, w) + b(v', w),$$
$$b(v, w + w') = b(v, w) + b(v, w'),$$
$$b(av, w) = ab(v, w) = b(v, aw).$$

The function b is defined on the set $V \times W$. This Cartesian product of two vector spaces can be given the structure of a vector space $V \oplus W$, the direct sum. But a bilinear map $b\colon V \times W \to X$ is completely different from a linear map $V \oplus W \to X$.

The tensor product space $V \otimes W$ is a vector space designed exactly so that a bilinear map $b\colon V \times W \to X$ becomes a linear map $V \otimes W \to X$. More precisely, it can be characterized abstractly by the following "universal property".

Definition B1.1. The *tensor product* of vector spaces V and W is a vector space $V \otimes W$ with a natural bilinear map $V \times W \to V \otimes W$, written $(v, w) \mapsto v \otimes w$, with the property that any bilinear map $b\colon V \times W \to X$ factors uniquely through $V \otimes W$. That means there exists a unique linear map $L\colon V \otimes W \to X$ such that $b(v, w) = L(v \otimes w)$.

This does not yet show that the tensor product exists, but uniqueness is clear: if X and Y were both tensor products, then each defining bilinear map would factor through the other – we get inverse linear maps between X and Y, showing they are isomorphic.

Note that the elements of the form $v \otimes w$ must span $V \otimes W$, since otherwise L would not be unique. If $\{e_i\}$ is a basis for V and $\{f_j\}$ a basis for W then bilinearity gives

$$\Big(\sum_i v^i e_i\Big) \otimes \Big(\sum_j w^j f_j\Big) = \sum_{i,j} v^i w^j\, e_i \otimes f_j.$$

Clearly then $\{e_i \otimes f_j\}$ spans $V \otimes W$ – indeed one can check that it is a basis. This is a valid construction for the space $V \otimes W$ – as the span of the $e_i \otimes f_j$ – but it does depend on the chosen bases. If $\dim V = m$ and $\dim W = n$ then we note $\dim V \otimes W = mn$.

A much more abstract construction of $V \otimes W$ goes through a huge infinite dimensional space. Given any set S, the *free vector space* on S is the set of all formal finite linear combinations $\sum a_i s_i$ with $a_i \in \mathbb{R}$ and $s_i \in S$. (This can equally well be thought of as the set of all real-valued functions on the set S which vanish outside some finite subset.) For instance, if S has k elements this gives a k-dimensional vector space with S as basis.

Given vector spaces V and W, let F be the free vector space over the set $V \times W$. (This consists of formal sums $\sum a_i(v_i, w_i)$ but ignores all the structure we have on the set $V \times W$.) Now let $R \subset F$ be the linear subspace spanned by all elements of the form:

$$(v + v', w) - (v, w) - (v', w),$$
$$(v, w + w') - (v, w) - (v, w'),$$
$$(av, w) - a(v, w), \qquad (v, aw) - a(v, w).$$

These correspond of course to the bilinearity conditions we started with. The quotient vector space F/R will be the tensor product $V \otimes W$. We have started with all possible $v \otimes w$ as generators and thrown in just enough relations to make the map $(v, w) \mapsto v \otimes w$ be bilinear.

The tensor product is commutative: there is a natural linear isomorphism $V\otimes W \to W\otimes V$ such that $v\otimes w \mapsto w\otimes v$. (This is easiest to verify using the universal property – simply factor the bilinear map $(v, w) \mapsto w \otimes v$ through $V \otimes W$ to give the desired isomorphism.)

Similarly, the tensor product is associative: there is a natural linear isomorphism $V \otimes (W \otimes X) \to (V \otimes W) \otimes X$. Note that any trilinear map from $V \times W \times X$ factors through this triple tensor product $V \otimes W \otimes X$.

Of special interest are the *tensor powers* of a single vector space V. We write $V^{\otimes k} := V \otimes \cdots \otimes V$. If $\{e_i\}$ is a basis for V, then $\{e_{i_1} \otimes \cdots \otimes e_{i_k}\}$ is a basis for $V^{\otimes k}$. In particular if V has dimension m, then $V^{\otimes k}$ has dimension m^k. There is a natural k-linear map $V^k \to V^{\otimes k}$ and any k-linear map $V^k \to W$ factors uniquely through $V^{\otimes k}$.

One can check that the dual of a tensor product is the tensor product of duals: $(V \otimes W)^* = V^* \otimes W^*$. In particular, we have $(V^*)^{\otimes k} = (V^{\otimes k})^*$. The latter is of course the set of linear functionals $V^{\otimes k} \to \mathbb{R}$, which as we have seen is exactly the set of k-linear maps $V^k \to \mathbb{R}$.

Definition B1.2. A *graded algebra* is a vector space A decomposed as $A = \bigoplus_{k=0}^{\infty} A_k$ together with an associative bilinear multiplication operation $A \times A \to A$ that respects the grading in the sense that the product $\omega \cdot \eta$ of elements $\omega \in A_k$ and $\eta \in A_\ell$ is an element of $A_{k+\ell}$. Often we consider graded algebras that are either commutative or anticommutative. Here *anticommutative* has a special meaning: for $\omega \in A_k$ and $\eta \in A_\ell$ as above, we have $\omega \cdot \eta = (-1)^{k\ell}\, \eta \cdot \omega$.

Example B1.3. The *tensor algebra* of a vector space V is

$$\otimes_* V := \bigoplus_{k=0}^{\infty} V^{\otimes k}.$$

Here of course $V^{\otimes 1} \cong V$ and $V^{\otimes 0} \cong \mathbb{R}$. Note that the tensor product is graded, but is neither commutative nor anticommutative.

B2. Exterior algebra

We now want to focus on antisymmetric tensors, to develop the so-called *exterior algebra* or *Grassmann algebra* of the vector space V.

Just as we constructed $V \otimes V = V^{\otimes 2}$ as a quotient of a huge vector space, adding relators corresponding to the rules for bilinearity, we construct the exterior power $V \wedge V = \Lambda_2 V$ as a further quotient. In particular, letting $S \subset V \otimes V$ denote span of the elements $v \otimes v$ for all $v \in V$, we set $V \wedge V := (V \otimes V)/S$. We write $v \wedge w$ for the image of $v \otimes w$ under the quotient map. Thus $v \wedge v = 0$ for any v. From

$$(v + w) \wedge (v + w) = 0$$

it then follows that $v \wedge w = -w \wedge v$. If $\{e_i : 1 \le i \le m\}$ is a basis for V, then

$$\{e_i \wedge e_j : 1 \le i < j \le m\}$$

is a basis for $V \wedge V$.

Higher exterior powers of V can be constructed in the same way, but formally, it is easiest to construct the whole *exterior algebra* $\Lambda_* V = \bigoplus \Lambda_k V$ at once, as a quotient of the tensor algebra $\otimes_* V$, this time by the two-sided ideal generated by the same set $S = \{v \otimes v\} \subset V \otimes V \subset \otimes_* V$. This means the span not just of the elements of S but also of their products (on the left and right) by arbitrary other tensors. Elements of $\Lambda_* V$ are called *multivectors* and elements of $\Lambda_k V$ are more specifically *k-vectors*.

End of Lecture 30 Nov 2015

Again we use $\wedge$ to denote the product on the resulting (still graded) quotient algebra. This product is called the *wedge product* or more formally the *exterior product*. We again get $v \wedge w = -w \wedge v$ for $v, w \in V$. More generally, for any $v_1, \ldots, v_k \in V$ and any permutation $\sigma \in \Sigma_k$ of $\{1, \ldots, k\}$, this implies

$$v_{\sigma 1} \wedge \cdots \wedge v_{\sigma k} = (\operatorname{sgn} \sigma)\, v_1 \wedge \cdots \wedge v_k.$$

A special case is the product of a k-vector α with an ℓ-vector β where we use a cyclic permutation to get the anticommutative law $\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha$.

If $\{e_i : 1 \le i \le m\}$ is a basis for V, then

$$\{e_{i_1 \cdots i_k} := e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \le i_1 < \cdots < i_k \le m\}$$

is a basis for $\Lambda_k V$. In particular, $\dim \Lambda_k V = \binom{m}{k}$; we have $\Lambda_0 V = \mathbb{R}$ but also $\Lambda_m V \cong \mathbb{R}$, spanned by $e_{12\cdots m}$. For $k > m$ there are no antisymmetric tensors: $\Lambda_k V = 0$. The exterior algebra has $\dim \Lambda_* V = \sum_{k=0}^m \binom{m}{k} = 2^m$. The *determinant* has a natural definition in terms of the exterior algebra: if we have m vectors $v_j \in V$ given in terms of the basis $\{e_i\}$ as $v_j = \sum_i v^i_j e_i$ then

$$v_1 \wedge \cdots \wedge v_m = \det(v^i_j) e_{12\cdots m}.$$

(The components of the wedge product of k vectors v_i are given by the various $k \times k$ minor determinants of the matrix (v^i_j).)

The exterior powers of V with the natural k-linear maps $V^k \to \Lambda_k V$ are also characterized by the following universal property. Given any alternating k-linear map $V^k \to X$ to any vector space X, it factors uniquely through $\Lambda_k V$. That is, alternating k-linear maps from V^k correspond to linear maps from $\Lambda_k V$. (One can also phrase the universality for all k together in terms of homomorphisms of anticommutative graded algebras.)

So far we have developed everything abstractly and algebraically. But there is a natural geometric picture of how k-vectors in $\Lambda_k V$ correspond to k-planes (k-dimensional linear subspaces) in V. More precisely, we should talk about *simple* k-vectors here: those that can be written in the form $v_1 \wedge \cdots \wedge v_k$. We will see that, for instance, $e_{12} + e_{34} \in \Lambda_2 \mathbb{R}^4$ is not simple.

A nonzero vector $v \in V$ lies in a unique oriented 1-plane (line) in V; two vectors represent the same oriented line if and only if they are positive multiples of each other. Now suppose we have vectors $v_1, \ldots, v_k \in V$. They are linearly independent if and only if $0 \ne v_1 \wedge \cdots \wedge v_k \in \Lambda_k V$. Two linearly independent k-tuples $(v_1, \ldots, v_k)$ and $(w_1, \ldots, w_k)$ represent the same oriented k-plane if and only if the wedge products $v_1 \wedge \cdots \wedge v_k$ and $w_1 \wedge \cdots \wedge w_k$ are positive multiples of each other, that is, if they lie in the same ray in $\Lambda_k V$. (Indeed, the multiple here is the ratio of k-areas of the parallelepipeds spanned by the two k-tuples, given as the determinant of the change-of-basis matrix for the k-plane.)

We let $G_k(V)$ denote the set of oriented k-planes in V, called the *(oriented) Grassmannian*. Then the set of simple k-vectors in $\Lambda_k V$ can be viewed as the cone over $G_k(V)$. (If we pick a norm on $\Lambda_k V$, say induced by an inner product on V, then we can think of $G_k(V)$ as the set of "unit" simple k-vectors, say those arising from an orthonormal basis for some k-plane.)

(Often, especially in algebraic geometry, one prefers to work with the *unoriented* Grassmannian $G_k(V)/\pm$. It is most naturally viewed as lying in the projective space

$$P(V) := (V \smallsetminus \{0\})/(\mathbb{R} \smallsetminus \{0\}).$$

In algebraic geometry one typically also replaces $\mathbb{R}$ by $\mathbb{C}$ throughout.)

If we give V an inner product, then any k-plane has a unique orthogonal $(m-k)$-plane. This induces an isomorphism between $G_k V$ and $G_{m-k} V$. It extends to a linear, norm-preserving isomorphism

$$\star \colon \Lambda_k V \to \Lambda_{m-k} V$$

called the Hodge star operator. (Recall that both these spaces have the same dimension $\binom{m}{k}$.) If v is a simple k-vector, then $\star v$ is a simple $(m-k)$-vector representing the orthogonal complement. In particular, if $\{e_i\}$ is an oriented orthonormal basis for V, then

$$\star(e_1 \wedge \cdots \wedge e_k) = e_{k+1} \wedge \cdots \wedge e_m$$

and similarly each other vector in our standard basis for $\Lambda_k V$ maps to a basis vector for $\Lambda_{m-k} V$, possibly with a minus sign.

Classical vector calculus in three dimensions uses the Hodge star implicitly: instead of talking about bivectors and trivectors, we introduce the cross product and triple product:

$$v\times w := \star(v\wedge w), \qquad [u,v,w] := \langle u, v\times w\rangle = \star(u\wedge v\wedge w).$$

But even physicists noticed that such vectors and scalars transform differently (say under reflection) than ordinary vectors and scalars, and thus refer to them as pseudovectors and pseudoscalars.

For $\dim V = m$, we can use these terms as follows:

- *scalars* are elements of $\mathbb{R} = \Lambda_0 V$,
- *vectors* are elements of $V = \Lambda_1 V$,
- *pseudovectors* are elements of $\star V = \Lambda_{m-1} V$, and
- *pseudoscalars* are elements of $\star\mathbb{R} = \Lambda_m V$.

Of course, these are in a sense the easy cases. For these k, any k-vector is simple. We can identify both G_1V and $G_{m-1}V$ as the unit sphere in $V = \Lambda_1 V \cong \Lambda_{m-1}V$. For $2 \le k \le m-2$ on the other hand, not all k-vectors are simple, and G_kV has lower dimension than the unit sphere in $\Lambda_k V$. Indeed, it can be shown that the set of simple k-vectors (the cone over G_kV) is given as the solutions to a certain set of quadratic equations called the Grassmann–Plücker relations. For instance $\sum a^{ij}e_{ij} \in \Lambda_2\mathbb{R}^4$ is a simple 2-vector if and only if

$$a^{12}a^{34} - a^{13}a^{24} + a^{14}a^{23} = 0.$$

This shows that $G_2\mathbb{R}^4$ is a smooth 4-submanifold in the unit sphere $\mathbb{S}^5 \subset \Lambda_2\mathbb{R}^4$.

If we choose an inner product on V, then thinking about how oriented orthonormal bases for a k-plane and its orthogonal complement fit together, we see that we can identify $G_kV = SO(m)/(SO(k)\times SO(m-k))$. In particular, it is a smooth manifold of dimension $k(m-k)$.

B3. Differential forms

Many textbooks omit discussion of multivectors and consider only the dual spaces. (This is presumably because the abstract definition of tensor powers and then exterior powers as quotient spaces seems difficult.) Recall that vector subspaces and quotient spaces are dual operations, in the sense that if $Y \subset X$ is a subspace, then the dual $(X/Y)^*$ of the quotient can be naturally identified with a subspace of X^*, namely with the *annihilator* Y^o of X^*, consisting of those linear functionals on X that vanish on Y:

$$(X/Y)^* \cong Y^o \subset X^*.$$

Using this, we find that

$$\Lambda^k V := (\Lambda_k V)^* \subset (V^{\otimes k})^*$$

is the subspace of those k-linear maps $V^k \to \mathbb{R}$ that are *alternating*.

While it is easy to construct the wedge product on multivectors as the image of the tensor product under the quotient map, the dual wedge product on Λ^*V requires constructing a map to the alternating subspace. For $\omega, \eta \in \Lambda^1 V = V^*$ we set

$$\omega\wedge\eta := \omega\otimes\eta - \eta\otimes\omega.$$

More generally, for $\omega \in \Lambda^k V$ and $\eta \in \Lambda^\ell V$ we use an alternating sum over all permutations $\sigma \in \Sigma_{k+\ell}$:

$$(\omega\wedge\eta)(v_1,\dots,v_{k+\ell}) := \frac{1}{k!\ell!}\sum_\sigma (\operatorname{sgn}\sigma)\,\omega(v_{\sigma 1},\dots,v_{\sigma k})\,\eta(v_{\sigma(k+1)},\dots,v_{\sigma(k+\ell)}).$$

The factor is chosen so that if $\{e_i\}$ is a basis for V and $\{\omega^i\}$ is the dual basis for $\Lambda^1 V = V^*$ then

$$\{\omega^{i_1\cdots i_k} := \omega^{i_1}\wedge\cdots\wedge\omega^{i_k}\}$$

is the basis of $\Lambda^k V$ dual to the basis $\{e_{i_1\cdots i_k}\}$ for $\Lambda_k V$.

Putting these spaces together, we get an anticommutative graded algebra

$$\Lambda^* V := \bigoplus_{k=0}^m \Lambda^k V.$$

Again the dimension of each summand is $\binom{m}{k}$ so the whole algebra has dimension 2^m.

If $L\colon V \to W$ is a linear map, then for each k we get an induced map $L^*\colon \Lambda^k W \to \Lambda^k V$ defined naturally by

$$L^*\omega(v_1,\dots,v_k) = \omega(Lv_1,\dots,Lv_k).$$

Of course, we have introduced these ideas in order to apply them to the tangent spaces T_pM to a manifold M^m. We get dual bundles $\Lambda_k TM$ and $\Lambda^k TM$ of rank $\binom{m}{k}$.

Definition B3.1. A *(differential) k-form* on a manifold M^m is a (smooth) section of the bundle $\Lambda^k TM$. We write $\Omega^k M = \Gamma(\Lambda^k TM)$ for the space of all k-forms, which is a module over $C^\infty M = \Omega^0 M$. Similarly we write $\Omega^* M = \Gamma(\Lambda^* TM) = \bigoplus \Omega^k M$ for the *exterior algebra* of M.

If $\omega \in \Omega^k M$ is a k-form, then at each point $p \in M$ the value $\omega_p \in \Lambda^k T_pM$ is an alternating k-linear form on T_pM or equivalently a linear functional on $\Lambda_k T_pM$. That is, for any k vectors $X_1,\dots,X_k \in T_pM$ we can evaluate

$$\omega_p(X_1,\dots,X_k) = \omega(X_1\wedge\cdots\wedge X_k) \in \mathbb{R}.$$

In particular, ω_p naturally takes values on (weighted) k-planes in T_pM; as we have mentioned, k-forms are the natural objects to integrate over k-dimensional submanifolds in M.

If $f\colon M^m \to N^n$ is a smooth map and $\omega \in \Omega^k N$ is a k-form, then we can pull back ω to get a k-form $f^*\omega$ on M defined by

$$(f^*\omega)_p(X_1,\dots,X_k) = \omega_{f(p)}((D_pf)X_1,\dots,(D_pf)X_k).$$

(Of course this vanishes if $k > m$.) As a special case, if $f\colon M \to N$ is the embedding of a submanifold, then $f^*\omega = \omega|_M$ is the *restriction* of ω to the submanifold M, in the sense that we consider only the values of $\omega_p(X_1,\dots,X_k)$ for $p \in M \subset N$ and $X_i \in T_pM \subset T_pN$.

Exercise B3.2. Pullback commutes with wedge product in the sense that

$$f^*(\omega \wedge \eta) = (f^*\omega) \wedge (f^*\eta)$$

for $f\colon M \to N$ and $\omega, \eta \in \Omega^* N$.

In a coordinate chart (U, φ) we have discussed the coordinate bases $\{\partial_i\}$ and $\{dx^i\}$ for T_pM and T_p^*M, respectively, the pullbacks under φ of the standard bases on $\mathbb{R}^m$. Similarly,

$$\left\{dx^{i_1} \wedge \cdots \wedge dx^{i_k} : 1 \le i_1 < \cdots < i_k \le m\right\}$$

forms the standard coordinate basis for k-forms; any $\omega \in \Omega^k(M)$ (or more properly its restriction to U) can be expressed uniquely as

$$\omega|_U = \sum_{i_1<\cdots<i_k} \omega_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

for some smooth functions $\omega_{i_1\cdots i_k} \in C^\infty U$. To simplify notation, we often write this as $\omega|_U = \sum_I \omega_I dx^I$ in terms of the *multi-index* $I = (i_1, \dots, i_k)$.

B4. Exterior derivative

Zero-forms are of course just scalar fields, that is, smooth functions. We have also already considered one-forms, which are simply covector fields. In particular, given $f \in \Omega^0 M$, its differential $df \in \Omega^1 M$ is a one-form with $df(X) = Xf$ for any vector field X. We now want to generalize this to define for any k-form ω its *exterior derivative*, a $(k+1)$-form $d\omega$.

Definition B4.1. An *antiderivation* on the graded algebra $(\Omega^* M, \wedge)$ is a linear map $D\colon \Omega^* M \to \Omega^* M$ satisfying the following version of the Leibniz product rule for $\omega \in \Omega^k M$ and $\eta \in \Omega^\ell M$:

$$D(\omega \wedge \eta) = (D\omega) \wedge \eta + (-1)^k \omega \wedge (D\eta).$$

To remember the sign here, it can help to think of D as behaving like a one-form when it "moves past" ω.

Proposition B4.2. *Any antiderivation on $\Omega^* M$ is a local operator in the sense that if $\omega = \eta$ on an open set U then $D\omega = D\eta$ on U.*

Proof. By linearity it suffices to consider the case $\eta \equiv 0$, $\omega \in \Omega^k M$. Given any $p \in U$, we can find a function $f \in C^\infty M$ supported in U with $f(p) = 1$. Then $f\omega \equiv 0$ on M and it follows that

$$0 = D(f\omega) = (Df) \wedge \omega + f \wedge (D\omega).$$

At p this gives $0 = Df \wedge 0 + 1(D\omega)_p = (D\omega)_p$ as desired. □

`End of Lecture 7 Dec 2015`

Theorem B4.3. *For any manifold M^m, the differential map $d\colon \Omega^0 M \to \Omega^1 M$ has a unique $\mathbb{R}$-linear extension to an antiderivation $d\colon \Omega^* M \to \Omega^* M$ satisfying $d^2 = d \circ d = 0$. This antiderivation has degree 1 in the sense that it sends $\Omega^k M$ to $\Omega^{k+1} M$; it is called the* exterior derivative.

Proof. First suppose $g, f^i \in C^\infty M$ so that $g\, df^1 \wedge \cdots \wedge df^k \in \Omega^k M$. The two conditions on d together automatically imply that

$$d(g\, df^1 \wedge \cdots \wedge df^k) = dg \wedge df^1 \wedge \cdots \wedge df^k \in \Omega^{k+1} M.$$

In a coordinate chart (U, φ) of course every k-form ω can be expressed as a sum of terms of this form. The proposition above shows we can work locally in such a chart. Thus we know the exterior derivative (if it exists) must be given in coordinates by

$$\begin{aligned} d\Big(\sum_I \omega_I dx^I\Big) &= \sum_I d\omega_I \wedge dx^I = \sum_I \sum_i \partial_i \omega_I\, dx^i \wedge dx^I \\ &= \sum_I \sum_i \partial_i \omega_I\, dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}. \end{aligned}$$

(Note that terms here where $i = i_j$ will vanish; for the other terms, reordering the factors in this last wedge product – to put i in increasing order with the i_js and thus obtain a standard basis element – will introduce signs.)

Now a straightforward calculation shows that the operator d defined by this formula really is an antiderivation locally:

$$\begin{aligned} &d((a\, dx^I) \wedge (b\, dx^J)) \\ &\quad= d(ab) \wedge dx^I \wedge dx^J = ((da)b + a(db)) \wedge dx^I \wedge dx^J \\ &\quad= (da \wedge dx^I) \wedge (b\, dx^J) + (-1)^k (a\, dx^I) \wedge (db \wedge dx^J), \end{aligned}$$

where $I = (i_1, \dots, i_k)$ is a k-index. Clearly this antiderivation has degree 1 as claimed.

Now since d is determined uniquely, if we have overlapping charts (U, φ) and (V, ψ), then on $U \cap V$ we must get the same result evaluating d in either chart. Finally, since the exterior algebra operations $+$ and $\wedge$ are defined pointwise, to check that d is an antiderivation and $d^2 = 0$ it suffices that we know these hold locally. □

Proposition B4.4. *The pullback of forms under a map $f\colon M^m \to N^n$ commutes with the exterior derivative. That is, for $\omega \in \Omega^* N$ we have $d(f^*\omega) = f^*(d\omega)$.*

Proof. It suffices to work locally around a point $p \in M$. Let (V, ψ) be coordinates around $f(p)$. By linearity we can assume $\omega = a\, dy^{i_1} \wedge \cdots \wedge dy^{i_k}$ in these coordinates. For $k = 0$ we have $\omega = a \in C^\infty N$. For any $X_p \in T_pM$ we have

$$\begin{aligned} (f^*da)(X_p) &= (da)(f_*X_p) = (f_*X_p)a \\ &= X_p(f^*a) = (df^*a)(X_p). \end{aligned}$$

Note that if $(f^1, \dots, f^n) = \psi \circ f$ is the coordinate expression of f (on some neighborhood of p) then the formula above gives $f^*(dy^i) = df^i$. Since pullback commutes with wedge products, for $k > 0$ we then have

$$f^*\omega = (f^*a)\, df^{i_1} \wedge \cdots \wedge df^{i_k}$$

and so

$$\begin{aligned} d(f^*\omega) &= d(f^*a) \wedge df^{i_1} \wedge \cdots \wedge df^{i_k} \\ &= f^*(da) \wedge df^{i_1} \wedge \cdots \wedge df^{i_k} \\ &= f^*(da \wedge dy^{i_1} \wedge \cdots \wedge dy^{i_k}) = f^*(d\omega). \qquad \square \end{aligned}$$

The *contraction* of a form with a vector field (also known as interior multiplication) has a seemingly trivial definition: if $\omega \in \Omega^k M$ and $X \in \mathcal{X}(M)$ then $\iota_X\omega \in \Omega^{k-1}$ is given by

$$\iota_X\omega(X_2, \ldots, X_k) := \omega(X, X_2, \ldots, X_k).$$

First note that this is a purely pointwise operation, so we could define it on $\Lambda^k V$ for a single vector space – even proving the next proposition at that level – but we won't bother. (It is the adjoint of the operator on $\Lambda_k V$ given by left multiplication by X.)

Next note that for a 1-form, $\iota_X(\omega) = \omega(X) \in \Omega^0 M$. For a 0-form $f \in \Omega^0 M = C^\infty M$ we set $\iota_X f = 0$.

Proposition B4.5. *For any X, the operation ι_X is an antiderivation on $\Omega^* M$ of degree -1 whose square is zero.*

Proof. It is clear that $\iota_X \circ \iota_X = 0$ since

$$\iota_X\iota_X\omega(\ldots) = \omega(X, X, \ldots) = 0.$$

The antiderivation property is

$$\iota_X(\omega \wedge \eta) = (\iota_X\omega) \wedge \eta + (-1)^k \omega \wedge (\iota_X\eta)$$

for $\omega \in \Omega^k M$; we leave the proof as an exercise. $\square$

We will later discuss Cartan's Magic Formula, relating this contraction to exterior and Lie derivatives.

B5. Volume forms and orientation

An *orientation* on an m-dimensional vector space V is a choice of component of $\Lambda^m V \smallsetminus \{0\} \cong \mathbb{R} \smallsetminus \{0\}$, that is a choice of a nonzero m-form in $\omega \in \Lambda^m V$ (up to positive real multiples). If V is oriented by ω, then an ordered basis $\{e_1, \ldots, e_m\}$ for V is said to be *positively oriented* if $\omega(e_1, \ldots, e_m) > 0$. Often an orientation on V is defined through such a basis (to avoid the machinery of the exterior algebra).

A *volume form* on a manifold M is a nowhere vanishing m-form $\omega \in \Omega^m M$. We say M is *orientable* if it admits a volume form. (The Möbius strip and the Klein bottle are examples of nonorientable 2-manifolds.) An *orientation* of M is a choice of volume form, up to pointwise multiplication by positive smooth functions $\lambda > 0 \in C^\infty M$. This is the same as a continuous choice of orientations of the tangent spaces T_pM. A connected orientable manifold has exactly two orientations.

The standard orientation on $\mathbb{R}^m$ is given by $dx^1 \wedge \cdots \wedge dx^m$, so that $\{e_1, \ldots, e_m\}$ is an oriented basis for each T_pM.

An equivalent definition of orientation (analogous to that of smooth structures) is through a coherently oriented atlas for M. Here two charts (U, φ) and (V, ψ) are *coherently oriented* if the transition function $\varphi \circ \psi^{-1}$ is an orientation-preserving diffeomorphism of $\mathbb{R}^m$.

Suppose now M^m is an oriented Riemannian manifold. At any $p \in M$ there is a unique $\Omega_p \in \Lambda^m T_pM$ such that $\Omega_p(e_1, \ldots, e_m) = +1$ for any oriented orthonormal basis $\{e_1, \ldots, e_m\}$ for T_pM. These fit together to give the *Riemannian volume form* $\Omega \in \Omega^m M$. In terms of the Hodge star, we have $\Omega = \star 1$.

Given an oriented coordinate chart (U, φ) then at any $p \in U$ we have the coordinate basis $\{\partial_i\}$ for T_pM but can also choose an oriented orthonormal basis $\{e_k\}$. Then of course for some matrix $A = (a_i^k)$ we have $\partial_i = \sum_k a_i^k e_k$. Since $\langle e_k, e_\ell \rangle = \delta_{k\ell}$, we get

$$g_{ij} = \langle \partial_i, \partial_j \rangle = \Big\langle \sum a_i^k e_k, \sum a_j^\ell e_\ell \Big\rangle = \sum_k a_i^k a_j^k.$$

As a matrix equation, we can write $(g_{ij}) = A^T A$, which implies $\det(g_{ij}) = (\det A)^2$. Since both bases are positively oriented, we know $\det A > 0$, so $\det A = +\sqrt{\det g}$. (Note that while abbreviating $\det(g_{ij})$ as $\det g$ is common, it unfortunately hides the fact that this is an expression in particular coordinates.)

Now we compute

$$\Omega_p(\partial_1, \ldots, \partial_m) = (\det A)\, \Omega_p(e_1, \ldots, e_m) = \det A = \sqrt{\det g}.$$

Equivalently, we have the coordinate expression

$$\Omega = \sqrt{\det g}\; dx^1 \wedge \cdots \wedge dx^m.$$

On an oriented Riemannian manifold (M^m, g), any m-form ω is a multiple $\omega = f\Omega = \star f$ of the volume form Ω, with $f \in C^\infty M$.

B6. Integration

We will base our integration theory on the Riemann integral. Recall that given an arbitrary real-valued function f on a box $B = [a_1, b_1] \times \cdots \times [a_m, b_m] \subset \mathbb{R}^m$ we define upper and lower Riemann sums over arbitrary partitions into small boxes – the function f is *Riemann integrable* if these have the same limiting value, which we call

$$\int_B f\, dx^1 \cdots dx^m.$$

Recall also that $A \subset \mathbb{R}^m$ has (Lebesgue) *measure zero* if for each $\varepsilon > 0$ there is a covering of A by countably many boxes of total volume less than ε. The image of a set of measure zero under a diffeomorphism (or indeed under any locally Lipschitz map) again has measure zero. Thus we can also speak of subsets of measure zero in a manifold M.

Given a function $f\colon D \to \mathbb{R}$ with $D \subset \mathbb{R}^m$, we define its extension by zero $\bar f\colon \mathbb{R}^m \to \mathbb{R}$ by setting $\bar f = f$ on D and $\bar f = 0$ elsewhere. Lebesgue proved the following: A bounded function $f\colon D \to \mathbb{R}$ defined on a bounded domain $D \subset \mathbb{R}^m$ is Riemann integrable if and only if f is *continuous almost everywhere*, meaning that its set of discontinuities has measure zero.

For instance, the characteristic function χ_D is Riemann integrable if D is bounded and its boundary ∂D has measure zero. Then we call D a *domain of integration*.

Because a continuous function f on a compact set is bounded, we find: If $U \subset \mathbb{R}^m$ is open and $f\colon U \to \mathbb{R}$ has compact support in U, then f is Riemann integrable.

We write $\Omega^k_c M \subset \Omega^k M$ for the subspace of k-forms with compact support. (If M is compact, then of course $\Omega^k_c = \Omega^k$.)

Definition B6.1. If $\omega \in \Omega^m_c U$ is an m-form with compact support in $U \subset \mathbb{R}^m$ then of course we can write uniquely $\omega = f\,dx^1 \wedge \cdots \wedge dx^m$. We define

$$\int_U \omega = \int_U f\,dx^1 \wedge \cdots \wedge dx^m := \int_U f\,dx^1 \cdots dx^m.$$

Note that we use the standard basis element for $\Lambda^m\mathbb{R}^m$ here. Otherwise we have for instance $\int f\,dx^2 \wedge dx^1 = -\int f\,dx^1\,dx^2$.

Lemma B6.2. *If $\varphi\colon U \to V$ is a diffeomorphism of connected open sets in $\mathbb{R}^m$ and ω an m-form with compact support in V, then*

$$\int_U \varphi^*\omega = \pm \int_V \omega,$$

where the sign depends on whether φ is orientation-preserving or not.

Proof. Use x^i for the standard coordinates on U and y^j for those on V. Then $\omega = f\,dy^1 \wedge \cdots \wedge dy^m$ for some function f. Writing $\varphi^i = y^i \circ \varphi$, the Jacobian matrix of φ is $J := (\partial\varphi^i/\partial x^j)$. We have $d\varphi^i = \varphi^* dy^i$ and so

$$d\varphi^1 \wedge \cdots \wedge d\varphi^m = \det J\,dx^1 \wedge \cdots \wedge dx^m.$$

Thus

$$\int_U \varphi^*\omega = \int_U (f \circ \varphi)\,d\varphi^1 \wedge \cdots \wedge d\varphi^m \qquad (1)$$

$$= \int_U (f \circ \varphi)\,\det J\,dx^1 \wedge \cdots \wedge dx^m. \qquad (2)$$

On the other hand, the standard change-of-variabes formula says

$$\int_V \omega = \int_V f\,dy^1 \cdots dy^m = \int_U (f \circ \varphi)\,|\det J|\,dx^1 \wedge \cdots \wedge dx^m.$$

Since U is connected, $\det J$ has a constant sign, depending on whether φ is orientation-preserving. □

Now suppose M^m is an oriented manifold, and $\omega \in \Omega^m_c M$ is a compactly supported m-form. Then we will define $\int_M \omega \in \mathbb{R}$.

First consider a single (oriented) chart (U, φ) and assume $\omega \in \Omega^m_c U$. Then we define

$$\int_U \omega := \int_{\varphi(U)} (\varphi^{-1})^*\omega.$$

We claim this is independent of φ: if (U, ψ) is another oriented chart, then using the diffeomorphism $\varphi \circ \psi^{-1}$ we find

$$\int_{\varphi(U)} (\varphi^{-1})^*\omega = \int_{\psi(U)} (\varphi \circ \psi^{-1})^*(\varphi^{-1})^*\omega = \int_{\psi(U)} (\psi^{-1})^*\omega.$$

In general, we choose a partition of unity $\{f_\alpha\}$ subordinate to an oriented atlas $\{(U_\alpha, \varphi_\alpha)\}$. For any $\omega \in \Omega^m_c M$, note that $\omega = \sum_\alpha f_\alpha\omega$ is a *finite* sum and each summand has compact support in the respective U_α. We define

$$\int_M \omega := \sum_\alpha \int_{U_\alpha} f_\alpha\omega.$$

We just need to check this is independent of the choice of atlas and partition of unity.

So suppose $\{g_\beta\}$ is a partition of unity subordinate to another oriented atlas $\{(V_\beta, \varphi_\beta)\}$. Then we have

$$\sum_\alpha \int_{U_\alpha} f_\alpha\omega = \sum_\alpha \int_{U_\alpha} f_\alpha \sum_\beta g_\beta\omega$$

$$= \sum_\alpha \sum_\beta \int_{U_\alpha} f_\alpha g_\beta\omega = \sum_\alpha \sum_\beta \int_{U_\alpha \cap V_\beta} f_\alpha g_\beta\omega.$$

But by symmetry, we see that the last expression also equals $\sum_\beta \int_{V_\beta} g_\beta\omega$, as desired.

Note: If $-M$ denotes the manifold M with opposite orientation, then we have $\int_{-M} \omega = -\int_M \omega$

Note: for $m = 0$, a compact oriented 0-manifold is a finite collection of points with signs ± 1: we write $M = \sum p_i - \sum q_j$. (Here we cannot use charts to test orientation.) The integral of a zero-form (function) $f\colon M \to \mathbb{R}$ is defined to be $\int_M f = \sum_i f(p_i) - \sum_j f(q_j)$.

We have developed this theory for smooth forms, partly just because we have no notation for possibly discontinuous sections of $\Lambda^m TM$. As long as ω is bounded, vanishes outside some compact set and is continuous almost everywhere, we can repeat the calculations above with no changes to define $\int_M \omega$.

`End of Lecture 14 Dec 2015`

On an oriented Riemannian manifold M (or any manifold with a specified volume form Ω), we define the volume integral of a function $f \in C^\infty M$ as

$$\int_M f\,d\,\mathrm{vol} := \int_M f\Omega = \int_M \star f.$$

Note that if we switch orientation, the volume form on $-M$ is $-\Omega$, so the volume integral is independent of orientation: $\int_{-M} f\,d\,\mathrm{vol} = \int_M f\,d\,\mathrm{vol}$.

For a domain $D \subset M$ (compact with boundary of measure zero) we define its volume to be

$$\mathrm{vol}\,D := \int_D 1\,d\,\mathrm{vol} = \int_D \Omega := \int_M \chi_D\Omega \geq 0.$$

The volume of the manifold is $\mathrm{vol}(M) := \int_M 1\,d\,\mathrm{vol} = \int_M \Omega$. This works directly if M is compact; for a noncompact manifold we can take a limit over an appropriate compact exhaustion and reach either a finite value or $+\infty$.

B7. Manifolds with boundary

Suppose M^m is an manifold with boundary; its boundary ∂M is an $(m-1)$-manifold. At $p \in \partial M \subset M$ we see that $T_p\partial M \subset T_pM$ is a hyperplane, cutting T_pM into two parts, consisting of the *inward*- and *outward-pointing* vectors at p.

An orientation on M induces an orientation on ∂M as follows. Suppose $(v, v_1, \dots, v_{m-1})$ is an oriented basis for T_pM, where v is outward-pointing and $v_i \in T_p\partial M$. Then $(v_1, \dots, v_{m-1})$ is by definition an oriented basis for $T_p\partial M$. (There are four obvious possible conventions here – either and inward- or outward-pointing vector could be put either before or after the basis for $T_p\partial M$. Our convention works best for Stokes' Theorem.

Equivalently, suppose the orientation of M is given by a volume form Ω, and we pick a vector field $X \in \mathcal{X}(M)$ which is outward-pointing along ∂M. Then the contraction $\iota_X(\Omega)$ restricted to ∂M is a volume form on the boundary which defines its orientation.

B8. Stokes' Theorem

Suppose M^m is an oriented manifold with boundary and ω is an $(m-1)$-form with compact support on M. Stokes' Theorem then says $\int_M d\omega = \int_{\partial M} \omega$. We see $d^2 = 0$ is dual to the condition that $\partial(\partial M) = \varnothing$:

$$0 = \int_M d^2\eta = \int_{\partial M} d\eta = \int_{\partial\partial M} \eta = 0.$$

Stokes' Theorem is quite fundamental, and can be used for instance to define $d\omega$ for nonsmooth forms, or ∂M for generalized surfaces M.

Remark B8.1. Of course in $\int_{\partial M} \omega$, the integrand is really the restriction or pullback $\omega|_{\partial M} = i^*\omega$ of ω to ∂M. This is now a top-dimensional form on the $(m-1)$-manifold ∂M.

When M is a manifold without boundary ($\partial M = \varnothing$) of course Stokes' Theorem reduces to $\int_M d\omega = 0$. It turns out that on a connected orientable closed manifold M^m, an m-form η can be written as $d\omega$ for some ω if and only if $\int_M \eta$ vanishes; we will return to such questions after proving the theorem.

Stokes himself would probably not recognize this generalized version of his theorem. The modern formulation in terms of differential forms is due mainly to Élie Cartan. The classical cases are those in low dimensions. For M an interval ($m = 1$), we just have the fundamental theorem of calculus; for a domain in $\mathbb{R}^2$, we have Green's theorem; for a domain in $\mathbb{R}^3$, we have Gauss's divergence theorem; and for a surface with boundary in $\mathbb{R}^3$ we have the theorem attributed to Stokes.

These special cases are of course normally formulated not with differential forms and the exterior derivative, but with gradients of functions, and divergence and curl of vector fields. More precisely, on any Riemannian manifold, we use the inner product to identify T_pM and T_p^*M and thus vector fields with one-forms. The gradient ∇f of a function $f \in C^\infty M$ is the vector field corresponding in this way to df. In particular, for any vector field X, we have

$$g(\nabla f, X) = \langle \nabla f, X\rangle = df(X) = Xf.$$

On $\mathbb{R}^3$ we further use the Hodge star to identify vectors with pseudovectors and thus one-forms with two-forms, and to identify scalars with pseudoscalars and thus zero-forms with three-forms. Then div, grad and curl are all just the exterior derivative. Explicitly, we identify both the one-form $p\,dx + q\,dy + r\,dz$ and the two-form $p\,dy \wedge dz + q\,dz \wedge dx + r\,dx \wedge dy$ with the vector field $p\partial_x + q\partial_y + r\partial_z$, and the three-form $f\,dx \wedge dy \wedge dz$ with the function f. Then $d\colon \Omega^0 \to \Omega^1$ is the gradient as above, $d\colon \Omega^1 \to \Omega^2$ is the curl, and $d\colon \Omega^2 \to \Omega^3$ is the divergence.

Our version of Stokes' theorem is (as mentioned above) certainly not the most general. For instance, we could easily allow "manifolds with corners", like compact domains with piecewise smooth boundaries. (It should be clear that the divergence theorem in $\mathbb{R}^3$ is valid for a cube as well as a sphere.)

Theorem B8.2 (Stokes). *Suppose M^m is an oriented manifold with boundary and ω is an $(m-1)$-form on M with compact support. Then*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Proof. Both sides are linear and integrals are defined via partitions of unity. In particular

$$d\omega = \sum d(f_\alpha\omega) = \Big(d\sum f_\alpha\Big)\omega + \sum f_\alpha\,d\omega = \sum f_\alpha\,d\omega,$$

we see that it suffices to consider the case when ω is compactly supported inside one oriented coordinate chart (U, φ). We may also assume that $\varphi(U) = \mathbb{R}^m$ or $\varphi(U) = H^m$, depending on whether U is disjoint from ∂M or not. Since the statement of the theorem is invariant under pullback by a diffeomorphism, we have shown it suffices to consider the cases (a) $M = \mathbb{R}^m$ and (b) $M = H^m$.

After scaling, we can assume that ω is compactly supported within the cube (a) $Q := (-1, 0)^m$ or (b) $Q := (-1, 0] \times (-1, 0)^{m-1}$. In either case, we write

$$\omega = \sum_{j=1}^m (-1)^{j-1}\omega^j\,dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^m$$

with $\operatorname{supp}\omega^j \subset Q$, so that

$$d\omega = \sum \frac{\partial\omega^j}{\partial x^j}\,dx^1 \wedge \cdots \wedge dx^m,$$

meaning that

$$\int_M d\omega = \sum \int_Q \frac{\partial\omega^j}{\partial x^j}\,dx^1 \cdots dx^m.$$

Now for each j we have

$$\int_Q \frac{\partial \omega^j}{\partial x^j} = \int_{-1}^0 \cdots \int_{-1}^0 \Big(\int_{-1}^0 \frac{\partial \omega^j}{\partial x^j}\, dx^j \Big) dx^1 \cdots \widehat{dx^j} \cdots dx^m.$$

By the fundamental theorem of calculus, the inner integral in parentheses equals $\omega^j(\dots,0,\dots) - \omega^j(\dots,-1,\dots)$. Since ω has compact support in Q, this vanishes for $j > 1$. In case (a) it vanishes even for $j = 1$, completing the proof that $\int_M d\omega = 0$.

In case (b) we have obtained

$$\int_{H^m} d\omega = \int_{-1}^0 \cdots \int_{-1}^0 \omega^1(0, x^2, \dots, x^m)\, dx^2 \cdots dx^m.$$

Now consider the restriction of ω to ∂H^m, the pullback under the inclusion map i. Since $i^* dx^1 = 0$ we immediately get

$$i^*\omega = \omega^1\, dx^2 \wedge \cdots \wedge dx^m.$$

Comparing this to the formula for $\int_M d\omega$ shows we are done. □

B9. De Rham cohomology

Definition B9.1. We say a k-form ω on M^m is *closed* if $d\omega = 0$; we say ω is *exact* if there is a $(k-1)$-form η such that $d\eta = \omega$. For clarity, write $d_k := d|_{\Omega^k} \colon \Omega^k \to \Omega^{k+1}$. We write $B^k(M)$ for the space of exact forms and $Z^k(M)$ for the space of closed forms. That is, $Z^k = \ker d_k$ and $B^k = \operatorname{Im} d_{k-1}$.

Since by definition $d^2 = 0$, it is clear that exact forms are closed. (Algebraically, we have $B^k \subset Z^k \subset \Omega^k$.) An interesting question is to what extent the converse fails to be true. The answer is measured by the *de Rham cohomology* $H^k(M) := Z^k/B^k$, the quotient vector space (over $\mathbb{R}$). A typical element is the equivalence class $[\omega] = \{\omega + d\eta\}$ of a closed k-form ω.

If we consider all degrees k together, we set

$$Z := Z^0 \oplus \cdots \oplus Z^m = \ker d, \qquad B := B^0 \oplus \cdots \oplus B^m = \operatorname{Im} d.$$

Defining

$$H := Z/B = H^0 \oplus \cdots \oplus H^m$$

we find this *cohomology ring* is not just a vector space but indeed an algebra under the wedge product. To check the details, start by noting that if ω' is closed, then

$$(\omega + d\eta) \wedge \omega' = \omega \wedge \omega' + d(\eta \wedge \omega').$$

An important theorem in the topology of manifolds says that this cohomology agrees with other standard definitions, in particular that it is dual to singular homology. (This is defined via cycles of simplices modulo boundaries, and can be thought of as counting loops or handles in dimension k.) The key here is Stokes' Theorem: a closed form integrates to zero over any boundary, so closed forms can be integrated over homology classes. Furthermore an exact form integrates to zero over any cycle.

Theorem B9.2. *If M^m is an orientable closed manifold with n components, then $H^0(M) \cong \mathbb{R}^n$.*

Proof. Note that $B^0 = 0$ so $H^0 = Z^0$, which is the space of functions with vanishing differential. But these are just the locally constant functions, so it is clear this space is n-dimensional. □

For orientable closed manifolds M^m, Poincaré duality (related to the Hodge star operation) gives a connection between co/homology in complementary dimensions. As an example, if such a manifold has n components, then $H^m(M) \cong \mathbb{R}^n$. We prove the dimension is at least this big.

Theorem B9.3. *If M^m is an orientable closed manifold with n components, then $H^m(M)$ has dimension at least n.*

Proof. Denote the components by M_i. By Stokes, integration $\omega \mapsto \int_{M_i} \omega$ over each component gives a map $\Omega^m = Z^m \to \mathbb{R}$ vanishing on B^m, and thus a map $H^m \to \mathbb{R}$; together these give a map to $\mathbb{R}^n$. Choosing a Riemannian metric on M_i, its volume form has positive integral; these n forms show that our map $H^m \to \mathbb{R}^n$ is surjective. □

B10. Lie derivatives

Earlier we defined the Lie derivative of a vector field Y with respect to a vector field X. This is a derivative along the integral curves of X, where we use (pushforwards under) the flow φ_t of X to move vectors of Y between different points along these curves.

The Lie derivative of a differential k-form ω is defined in the same way, except that the pushforward under φ_{-t} is replaced by a pullback under φ_t. That is, we define:

$$(L_X\omega)_p := \frac{d}{dt}\Big|_{t=0} \varphi_t^* \omega_{\varphi_t(p)} = \frac{d}{dt}\Big|_{t=0} (\varphi_t^*\omega)_p.$$

Note that this is again a k-form. In the particular case of $k = 0$ where $\omega = f \in C^\infty M$ we can ignore the pullback – $L_X f$ is simply the derivative of f along the integral curve, that is, $L_X f = Xf$.

Proposition B10.1. *The Lie derivative L_X on forms satisfies the following properties:*

1. *it is a derivation on $\Omega^* M$, that is, an $\mathbb{R}$-linear map satisfying*

$$L_X(\omega \wedge \eta) = (L_X\omega) \wedge \eta + \omega \wedge (L_X\eta);$$

2. *it commutes with the exterior derivative, that is,*

$$L_X(d\omega) = d(L_X\omega);$$

3. *it satisfies the "product" formula – for a k-form ω applied to k vector fields $Y_i \in \mathfrak{X}(M)$ we have*

$$L_X(\omega(Y_1, \dots, Y_k)) = (L_X\omega)(Y_1, \dots, Y_k) + \sum_{i=1}^k \omega(Y_1, \dots, L_X Y_i, \dots, Y_k).$$

Proof. 1. This follows directly from the fact that pullback commutes with wedge product and from the product rule for d/dt:

$$\begin{aligned}(L_X(\omega\wedge\eta))_p &= \frac{d}{dt}\Big|_{t=0}(\varphi_t^*(\omega\wedge\eta))_p\\ &= \frac{d}{dt}\Big|_{t=0}(\varphi_t^*\omega)_p\wedge(\varphi_t^*\eta)_p\\ &= \left(\frac{d}{dt}\Big|_{t=0}(\varphi_t^*\omega)_p\right)\wedge\eta_p+\omega_p\wedge\left(\frac{d}{dt}\Big|_{t=0}(\varphi_t^*\eta)_p\right)\\ &= (L_x\omega)_p\wedge\eta_p+\omega_p\wedge(L_X\eta)_p.\end{aligned}$$

2. This follows from the fact that d is linear and commutes with pullback:

$$\begin{aligned}L_Xd\omega &= \frac{d}{dt}\Big|_{t=0}\varphi_t^*d\omega=\frac{d}{dt}\Big|_{t=0}d\varphi_t^*\omega\\ &= d\Big(\frac{d}{dt}\Big|_{t=0}\varphi_t^*\omega\Big)=dL_X\omega.\end{aligned}$$

3. The proof follows (as for the product rule for d/dt) from a clever splitting of one difference quotient into two or more. We will write out the proof only for $k=1$, considering $\omega(Y)$. We find

$$\begin{aligned}L_X(\omega(Y))_p &= \lim_{t\to0}\tfrac{1}{t}\Big(\omega_{\varphi_tp}(Y_{\varphi_tp})-\omega_p(Y_p)\Big)\\ &= \lim_{t\to0}\tfrac{1}{t}\Big(\omega_{\varphi_tp}(Y_{\varphi_tp})-\omega_p(\varphi_{-t*}Y_{\varphi_tp})\Big)\\ &\quad+\lim_{t\to0}\tfrac{1}{t}\Big(\omega_p(\varphi_{-t*}Y_{\varphi_tp})-\omega_p(Y_p)\Big).\end{aligned}$$

Here the second limit clearly gives

$$\omega_p\Big(\frac{d}{dt}\Big|_{t=0}\varphi_{-t*}Y_{\varphi_tp}\Big)=\omega_p(L_XY).$$

For the first limit, we can rewrite the first term as $(\varphi_t^*\omega_{\varphi_tp})\big(\varphi_{-t*}Y_{\varphi_tp}\big)$, so that both terms are applied to the same vector. The limit becomes

$$\lim_{t\to0}\frac{\varphi_t^*(\omega_{\varphi_tp})-\omega_p}{t}(\varphi_{-t*}Y_{\varphi_tp}),$$

where the form clearly limits to $(L_X\omega)_p$ and the vector to Y_p.

□

`End of Lecture 4 Jan 2016`

Since the Lie derivatives of functions and vector fields are known, we can rewrite the product formula as a formula for $L_X\omega$ as follows:

$$\begin{aligned}&(L_X\omega)(Y_1,\dots,Y_k)\\ &\quad= X(\omega(Y_1,\dots,Y_k))-\sum_{i=1}^k\omega(Y_1,\dots,[X,Y_i],\dots,Y_k).\end{aligned}$$

Proposition B10.2 (Cartan's magic formula). *For any vector field X we have $L_X=d\iota_X+\iota_Xd$.*

Proof. We know that L_X is a derivation commuting with d. Since $d^2=0$, it is easy to check the right-hand side also commutes with d. Furthermore it is a derivation: for $\omega\in\Omega^kM$ we get

$$\begin{aligned}&d\iota_X(\omega\wedge\eta)+\iota_Xd(\omega\wedge\eta)\\ &\quad= d((\iota_X\omega)\wedge\eta)+(-1)^kd(\omega\wedge\iota_X\eta)\\ &\qquad+\iota_X((d\omega)\wedge\eta)+(-1)^k\iota_X(\omega\wedge d\eta)\\ &\quad= (d\iota_X\omega)\wedge\eta+(-1)^k(\iota_X\omega)\wedge(d\eta)+\cdots\\ &\quad= (d\iota_X\omega)\wedge\eta+(\iota_Xd\omega)\wedge\eta+\omega\wedge(d\iota_X\eta)+\omega\wedge(\iota_Xd\eta).\end{aligned}$$

Thus if the formula holds for ω and η, it also holds for $\omega\wedge\eta$ and for $d\omega$. By linearity and locality, this means it is enough to check it for 0-forms:

$$(d\iota_X+\iota_Xd)f=\iota_Xdf=(df)(X)=Xf=L_Xf.\qquad\square$$

Proposition B10.3. *Suppose X and Y are vector fields on M^m and ω is a 1-form. Then*

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

Proof. We use Cartan's Magic Formula and the product rule for $L_X\omega$:

$$\begin{aligned}d\omega(X,Y) &= (\iota_Xd\omega)(Y)\\ &= (L_X\omega)(Y)-(d\iota_X\omega)(Y)\\ &= X(\omega(Y))-\omega([X,Y])-d(\omega(X))(Y)\\ &= X(\omega(Y))-\omega([X,Y])-Y(\omega(X)).\qquad\square\end{aligned}$$

Note that by linearity and locality it suffices to consider $\omega=f\,dg$. So an alternate proof simply computes each term for this case, getting for instance $X\omega(Y)=X(f\,dg(Y))=X(fYg)=(Xf)(Yg)+fXYg$.

Theorem B10.4. *Suppose $\omega\in\Omega^k(M^m)$ is a k-form and $X_0,\dots,X_k\in\mathcal{X}(M)$ are $k+1$ vector fields. Then*

$$\begin{aligned}&(d\omega)(X_0,\dots,X_k)\\ &\quad=\sum_{0\le i\le k}(-1)^iX_i(\omega(X_0,\dots,\widehat{X_i},\dots,X_k))\\ &\quad+\sum_{0\le i<j\le k}(-1)^{i+j}\omega([X_i,X_j],X_0,\dots,\widehat{X_i},\dots,\widehat{X_j},\dots,X_k)).\end{aligned}$$

Note that the case $k=0$ is simply $df(X)=Xf$, and the case $k=1$ is the last proposition. The general proof by induction on k is left as an exercise; the hint is to use Cartan's magic formula as in the proof of the proposition to write

$$\begin{aligned}&(d\omega)(X_0,\dots,X_k)\\ &\qquad=(L_{X_0}\omega)(X_1,\dots,X_k)-(d\iota_{X_0}\omega)(X_1,\dots,X_k).\end{aligned}$$

C. RIEMANNIAN GEOMETRY

To take derivatives of a vector field along a curve requires comparing tangent spaces at different points. The Lie derivative uses diffeomorphisms to do this, which is not entirely satisfactory since we need not just a curve but a whole vector field.

Another approach is through connections or covariant derivatives. In particular, there is a natural connection on any Riemannian manifold, which is the starting point for studying its geometry.

C1. Submanifolds in Euclidean space

Suppose $M^m \subset \mathbb{R}^n$ is a submanifold. A map $X\colon M \to T\mathbb{R}^n$, $p \mapsto X_p \in T_p\mathbb{R}^n$ is called an $\mathbb{R}^n$-valued *vector field along M*. Of course $T_p\mathbb{R}^n \cong \mathbb{R}^n$, so we can identify X with a function $\tilde{X}\colon M \to \mathbb{R}^n$.

But $T_p\mathbb{R}^n$ also has an orthogonal decomposition (with respect to the standard Euclidean inner product) into spaces tangent and normal to M:

$$T_p\mathbb{R}^n = T_pM \oplus N_pM.$$

We let $\pi^\parallel$ and $\pi^\perp$ denote the orthogonal projections onto these subspaces, so that $X_p = \pi^\parallel X_p + \pi^\perp X_p$.

Now if $\gamma\colon [a,b] \to M$ is a curve (embedded) in M, then we have the function $\tilde{X} \circ \gamma\colon [a,b] \to \mathbb{R}^n$ and can take its derivative. We can view this derivative as an $\mathbb{R}^n$-valued function on the 1-submanifold $\gamma \subset M \subset \mathbb{R}^n$ instead of on $[a,b]$ (technically we compose with γ^{-1}). Again such a map to $\mathbb{R}^n$ can be identified with an $\mathbb{R}^n$-valued vector field along γ (viewing its value at each point p as lying in $T_p\mathbb{R}^n$). We call this vector field the derivative dX/dt of X along γ.

Both the original field X and its derivative dX/dt can be decomposed (via $\pi^\parallel$ and $\pi^\perp$) into parts tangent and normal to M. These decompositions are not in any definite relation to each other. Consider for instance vector fields along a surface in $\mathbb{R}^3$ as we studied last semester. The derivatives of the unit normal vector field are tangent vectors; the derivatives of tangent vector fields will usually have both tangent and normal components.

Definition C1.1. Suppose X is a smooth vector field on $M^m \subset \mathbb{R}^n$ and γ is a curve in M. Then the vector field

$$\frac{DX}{dt} := \pi^\parallel\Big(\frac{dX}{dt}\Big)$$

along γ, which is tangent to M, is called the *covariant derivative* of X along γ.

Note that we only need X to be defined along γ. Note also that we could apply this definition to any $\mathbb{R}^n$-valued field X, but there is little reason to do so – our goal is to focus on the geometry of M. Indeed, we will see that this covariant derivative can be defined in a way depending only on the Riemannian metric on M and independent of the particular embedding $M \subset \mathbb{R}^n$.

Example C1.2. Consider the round sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ and let $\gamma(t) := (\cos t, \sin t, 0)$ be the equator parametrized by arclength. Consider the vector field X along γ given by the tangent vector

$$X_{\gamma(t)} := \gamma'(t) = (-\sin t, \cos t, 0).$$

Since $dX/dt = \gamma''(t) = -\gamma(t)$ is normal to $\mathbb{S}^2$, we find $DX/dt \equiv 0$.

In general, a parametrized curve γ on M is called a *geodesic* if its velocity vector field $X = \dot\gamma$ satisfies $DX/dt \equiv 0$. On the sphere, the geodesics are exactly the great circles parametrized at constant speed.

Now we want to work out coordinate expressions for the covariant derivative. So let (U,φ) be a coordinate chart for $M^m \subset \mathbb{R}^n$ and write $V := \varphi(U) \subset \mathbb{R}^m$. Write $\{u^i : i = 1,\dots,m\}$ for the coordinates on $\mathbb{R}^m$. Because M is embedded in $\mathbb{R}^n$, we can also write the inverse map

$$\varphi^{-1} =: \psi = (\psi^1,\dots,\psi^n)\colon V \to U \subset M \subset \mathbb{R}^n$$

explicitly in coordinates. (Here we use $\{x^\alpha : \alpha = 1,\dots,n\}$ for the coordinates on $\mathbb{R}^n$ and have $\psi^\alpha = x^\alpha \circ \psi$.) The standard coordinate frame for TU is of course given by

$$\partial_i = \psi_*\Big(\frac{\partial}{\partial u^i}\Big) = \sum_\alpha \frac{\partial\psi^\alpha}{\partial u^i}\frac{\partial}{\partial x^\alpha}.$$

A curve γ in M will be given in coordinates as

$$\gamma(t) = \psi(u^1(t),\dots,u^m(t))$$

for some real-valued functions $u^i(t)$.

A vector field Y (tangent to M) along γ can be expressed in the coordinate basis as

$$Y_{\gamma(t)} = Y(t) = \sum_i b^i(t)\partial_i$$

for some real-valued functions $b^i(t)$. Its derivative and covariant derivative along γ are then

$$\frac{dY}{dt} = \sum_i \frac{db^i}{dt}\partial_i + b^i\frac{d\partial_i}{dt},$$
$$\frac{DY}{dt} = \sum_i \frac{db^i}{dt}\partial_i + b^i\frac{D\partial_i}{dt}.$$

To compute the covariant derivative $D\partial_i/dt$ of the coordinate basis vectors, we recall that a time derivative along γ is a directional derivative in direction $\dot\gamma$, so we get

$$\frac{D\partial_i}{dt} = \pi^\parallel\Big(\frac{d}{dt}\sum_\alpha \frac{\partial\psi^\alpha}{\partial u^i}\frac{\partial}{\partial x^\alpha}\Big) = \sum_\alpha\sum_j \frac{\partial^2\psi^\alpha}{\partial u^j\,\partial u^i}\frac{du^j}{dt}\pi^\parallel\Big(\frac{\partial}{\partial x^\alpha}\Big).$$

Here the $\partial/\partial x^\alpha$ are the standard basis vectors in $\mathbb{R}^n$. Their tangent parts can of course be expressed in the coordinate basis:

$$\pi^\parallel\Big(\frac{\partial}{\partial x^\alpha}\Big) = \sum_k c^k_\alpha\partial_k$$

for some smooth functions $c_\alpha^k \in C^\infty(U)$.

We now define the so-called *Christoffel symbols*

$$\Gamma_{ij}^k := \sum_\alpha \frac{\partial^2\psi^\alpha}{\partial u^j\,\partial u^i} c_\alpha^k,$$

noting the symmetry $\Gamma_{ij}^k = \Gamma_{ji}^k$. We have $\Gamma_{ij}^k \in C^\infty(U)$ for $1 \le i, j, k \le m$.

Using these, the expression above for the covariant derivative of ∂_i becomes

$$\frac{D\partial_i}{dt} = \sum_{j,k} \Gamma_{ij}^k \frac{du^j}{dt}\partial_k.$$

We can consider in particular the covariant derivative along a u^j-coordinate curve, where $u^j = t$ and each other u^i is constant. We write this as

$$\frac{D\partial_i}{\partial u^j} = \sum_k \Gamma_{ij}^k \partial_k.$$

That is, the Christoffel symbol Γ_{ij}^k is the ∂_k component of the covariant derivative of ∂_i in direction ∂_j.

We can now return to the general case of the covariant derivative of Y along γ; our formula becomes

$$\frac{DY}{dt} = \sum_k \Big(\frac{db^k}{dt} + \sum_{i,j} \Gamma_{ij}^k b^i \frac{du^j}{dt}\Big)\partial_k.$$

Note here that we don't see the coordinates in $\mathbb{R}^n$ at all; the vector field Y and curve γ on M are expressed in the standard instrinsic ways in the coordinate chart (U,φ). The embedding of $M \subset \mathbb{R}^n$ enters only in the computation of the Christoffel symbols Γ_{ij}^k, and our goal is to show these really only depend on the Riemannian metric induced on M by the embedding.

Now suppose $Y = \sum b^k\partial_k$ is a vector field defined on all of M (rather than just along γ) – its components b^k are now functions on U. We note that the covariant derivative DY/dt at a point $p = \gamma(t_0)$ doesn't depend on the whole curve γ but only on its velocity vector $X_p := \dot\gamma(t_0)$ there. In particular, if we set $a^j := du^j/dt$ then $X_p = \sum a^j\partial_j$, and the time derivative db^k/dt appearing in the formula is the directional derivative $X_p(b^k)$.

To emphasize this viewpoint, we introduce new notation and write this covariant derivative of Y at p in the direction X_p as $\nabla_{X_p}Y$. If X and Y are vector fields on M, we write $\nabla_X Y$ for the vector field whose value at p is $\nabla_{X_p}Y$. The formulas above mean that if $X = \sum a^j\partial_j$ and $Y = \sum b^k\partial_k$ in some coordinate chart, then

$$\nabla_X Y = \sum_{j,k} \Big(a^j(\partial_j b^k) + \sum_i \Gamma_{ij}^k b^i a^j\Big)\partial_k.$$

We have thus defined a *connection*, meaning an operation

$$\nabla\colon \mathcal{X}(M)\times\mathcal{X}(M) \to \mathcal{X}(M), \qquad \nabla\colon (X,Y) \mapsto \nabla_X Y.$$

This is clearly bilinear (over $\mathbb{R}$) and we claim it satisfies the following four properties:

- it is C^∞-linear in X:

$$\nabla_{fX}Y = f\nabla_X Y;$$

- it satisfies a product rule in Y:

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

- it is symmetric in the following sense:

$$\nabla_X Y - \nabla_Y X = [X,Y] = L_X Y;$$

- it is compatible with the Riemannian metric:

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

The first two properties are easily verified. The symmetry is equivalent to the fact that $\Gamma_{ij}^k = \Gamma_{ji}^k$. The metric property is left as an exercise.

C2. Connections

Let us now move to a very general situation. Suppose E is a vector bundle over a manifold M. A connection ∇ on E allows us to take covariant derivatives of sections of E. These are directional derivatives in the direction of some vector field $X \in \mathcal{X}(M)$ and are again sections of the same bundle E. That is, given a section $\sigma \in \Gamma(E)$, its covariant derivative (with respect to ∇) in direction X is the section $\nabla_X\sigma \in \Gamma(E)$. The formal definition is as follows:

Definition C2.1. Given a vector bundle $E \to M$, a *connection* on E is a bilinear map $\nabla\colon \mathcal{X}(M)\times\Gamma(E)\to\Gamma(E)$, written $(X,\sigma)\mapsto\nabla_X\sigma$, which is $C^\infty(M)$-linear in X and satisfies a product rule for σ:

$$\nabla_{fX}\sigma = f\nabla_X\sigma, \qquad \nabla_X(f\sigma) = (Xf)\sigma + f\nabla_X\sigma.$$

We call $\nabla_X\sigma$ the *covariant derivative* of σ.

Note that the tensoriality ($C^\infty M$-linearity) implies that the dependence on X is pointwise: $(\nabla_X\sigma)_p$ depends only on X_p and can be written as $\nabla_{X_p}\sigma$. This covariant derivative of course depends on more than just σ_p, but as for the other derivatives we have studied, the product rule means that the definition is local: if σ and τ have the same germ at p (that is, agree in some open neighborhood U) then $\nabla_{X_p}\sigma = \nabla_{X_p}\tau$. The trick is again to pick a bump function f supported within U with $f \equiv 1$ on some smaller neighborhood of p, so then $f\sigma = f\tau$. We calculate

$$\nabla_{X_p}\sigma = 1\nabla_{X_p}\sigma + 0\sigma_p = f(p)\nabla_{X_p}\sigma + (X_p f)\sigma_p = \nabla_{X_p}(f\sigma)$$

with the same for τ. Indeed, it suffices that σ and τ agree locally along some curve γ with $\dot\gamma(0) = X_p$; this can perhaps most easily be seen in coordinates as below.

There are many other ways to rephrase this definition, for instance in terms of sections of various induced bundles. For any fixed σ, we can consider $\nabla\sigma$ as a map taking a vector field X to the section $\nabla_X\sigma$. But the pointwise dependence on X means that this acts pointwise as a linear map $T_pM \to E_p$. That is, $\nabla\sigma$ can be viewed as a section of the bundle $L(TM, E) = E \otimes T^*M$; such a section can be called a vector-valued one-form. In this picture, the connection ∇ is a map from $\Gamma(E)$ to $\Gamma(E \otimes T^*M)$. While this is the approach taken in many books, we will stick to our more down-to-earth approach.

`End of Lecture 11 Jan 2016`

We have already seen one example of a connection: the one on TM induced by an embedding $M \to \mathbb{R}^n$, which satisfied not only the properties in this definition but also two further properties. As in that case, any connection can be expressed in coordinates via Christoffel symbols.

Suppose U is a coordinate neighborhood for M and a trivializing neighborhood for E, with $\{\partial_i : 1 \le i \le \dim M\}$ the coordinate frame for TM and $\{e_a : 1 \le a \le \operatorname{rk} E\}$ a frame for E. Then a connection ∇ is expressed in coordinates by the *Christoffel symbols* Γ^b_{ia} defined by $\nabla_{\partial_i} e_a = \sum_b \Gamma^b_{ia} e_b$, so that in general for $X = \sum v^i \partial_i$ and $\sigma = \sum \sigma^a e_a$ we have

$$\nabla_X \sigma = \sum_{i,b} v^i \Big(\partial_i \sigma^b + \sum_a \Gamma^b_{ia} \sigma^a \Big) e_b.$$

Any collection of smooth functions Γ^b_{ia} describes a connection.

Any connection induces a notion of *parallel transport* along a curve. If γ is a curve (from p to q) in M, then a section σ of E (defined at least on γ – more formally a section of γ^*E) is said to be *parallel* along γ (with respect to the connection ∇) if $\nabla_{\dot\gamma(t)} \sigma \equiv 0$ along γ. This corresponds to a first-order ODE , which has a unique solution given any initial value. That is, given any $\sigma_p \in E_p$, there is a unique extension to a parallel section σ along γ. In particular, looking at its value σ_q at the endpoint q, we get a linear map $P_\gamma \colon E_p \to E_q$ called *parallel transport* along γ (w.r.t. ∇).

If γ is a closed curve – a loop based at p – then we call $P_\gamma \colon E_p \to E_p$ the *holonomy* of ∇ around γ. Note that this need not be the identity; instead it demonstrates the curvature of the connection ∇.

The tangent space to E at each point has a natural *vertical subspace* of dimension k at each point: the tangent space to the fiber E_p or equivalently the kernel of the differential of the projection $E \to M$. Another way to view a connection is as a choice of a complementary *horizontal subspace* of dimension m. A section σ is parallel along γ if $D\sigma(\dot\gamma)$ lies in these horizontal subspaces.

C3. The Levi-Civita Connection

Specializing to the case of connections on the tangent bundle $E = TM$, we can compare $\nabla_X Y$ and $\nabla_Y X$. It is too much to hope that these are the same for any vector fields X and Y – the behavior when we replace X by fX is different. But this kind of effect is captured also in the Lie bracket of the vector fields. We define the *torsion* of the connection as $T(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y]$. This expression is $C^\infty M$-linear in each of its arguments. The connection ∇ is said to be *symmetric* or *torsion-free* if $T(X, Y) = 0$ for all X and Y, that is, if $\nabla_X Y - \nabla_Y X = [X, Y]$. (This is of course one of the properties we observed for the connection induced from an embedding $M \subset \mathbb{R}^n$.) In terms of Christoffel symbols in a coordinate basis (where $[\partial_i, \partial_j] = 0$), we find that ∇ is torsion-free if and only if $\Gamma^k_{ij} = \Gamma^k_{ji}$.

On a Riemannian manifold (M, g) we can also ask whether a connection ∇ on TM is compatible with the metric. A *metric connection* is one satisfying

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

One interpretation of this equation is as saying that the metric tensor g is parallel with respect to ∇. Just as we saw for the Lie derivative, a connection on one bundle naturally induces connections on the dual bundle and its tensor powers such that product rules hold. In particular, we could define the covariant derivative $\nabla_X g$ via

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

Then clearly ∇ is a metric connection if and only if for all X we have $\nabla_X g = 0$.

We will now show that any Riemannian manifold has a unique torsion-free metric connection ∇; this is called the *Levi-Civita connection*. Note that we have already constructed such a connection on any manifold $M \subset \mathbb{R}^n$ embedded in Euclidean space. We give a proof due to Koszul, using the fact that a vector field $\nabla_X Y$ on a Riemannian manifold is specified by its inner products with arbitrary vector fields Z.

Theorem C3.1. *Any Riemannian manifold (M, g) has a unique Levi-Civita connection, characterized by the Koszul formula*

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

Proof. Because the metric is fixed, we write $g(\cdot, \cdot)$ as $\langle \cdot, \cdot \rangle$. The uniqueness amounts to checking that any Levi-Civita connection does satisfy the Koszul formula. We use the metric property to expand each of the first three terms; the first (for instance) becomes $\langle \nabla_X Y, Z\rangle + \langle \nabla_X Z, Y\rangle$. We use the symmetry to expand each of the last three terms; the first (for instance) becomes $\langle \nabla_X Y, Z\rangle - \langle \nabla_Y X, Z\rangle$. Adding everything we find that most terms cancel out; we are left with $2\,\langle \nabla_X Y, Z\rangle$.

It remains to show that the formula does define a Levi-Civita connection. First, we claim that the right-hand side is tensorial (meaning $C^\infty M$-linear) in Z:

$$\begin{aligned} &X\langle Y, fZ\rangle + Y\langle X, fZ\rangle - fZ\langle X, Y\rangle \\ &\qquad + \langle [X, Y], fZ\rangle - \langle [X, fZ], Y\rangle - \langle [Y, fZ], X\rangle \\ &= (Xf)\langle Y, Z\rangle + fX\langle Y, Z\rangle + (Yf)\langle X, Z\rangle + fY\langle X, Z\rangle \\ &\qquad - fZ\langle X, Y\rangle + f\langle [X, Y], Z\rangle \\ &\qquad - f\langle [X, Z], Y\rangle - (Xf)\langle Z, Y\rangle \\ &\qquad - f\langle [Y, Z], X\rangle - (Yf)\langle Z, X\rangle \\ &= f\Big(X\langle Y, Z\rangle + Y\langle X, Z\rangle - Z\langle X, Y\rangle \\ &\qquad + \langle [X, Y], Z\rangle - \langle [X, Z], Y\rangle - \langle [Y, Z], X\rangle \Big) \end{aligned}$$

This means for any fixed X and Y, the right hand side is $\omega(Z)$ for some one-form ω. But using the metric g, this one-form equivalent to a vector field W defined by $2g(W,Z) = \omega(Z)$.

This construction $\nabla\colon (X,Y) \mapsto W$ is clearly bilinear in X and Y. The facts that it is tensorial in X and satisfies the product rule in Y are verified by calculuations similar to the one above. The the Koszul formula defines a connection ∇. What remains to show is that it is symmetric and compatible with g.

To check the symmetry, note that the right-hand side of the Koszul formula is symmetric in X and Y except for the term $\langle [X,Y], Z\rangle$. Thus

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle - 2\langle \nabla_Y X, Z\rangle &= \langle [X,Y], Z\rangle - \langle [Y,X], Z\rangle \\ &= 2\langle [X,Y], Z\rangle. \end{aligned}$$

Since this holds for all Z, we conclude the connection is torsion-free: $\nabla_X Y - \nabla_Y X = [X,Y]$.

To check the metric property, note that the right-hand side is antisymmetric in Y and Z except for the term $X\langle Y,Z\rangle$. Thus

$$2\langle \nabla_X Y, Z\rangle + 2\langle Y, \nabla_X Z\rangle = 2X\langle Y,Z\rangle$$

as desired. □

Now we want to consider what the Levi-Civita connection looks like in coordinates. We know the Christoffel symbols for a torsion-free connection will be symmetric: $\Gamma^k_{ij} = \Gamma^k_{ji}$. In terms of the components $g_{ij} := g(\partial_i, \partial_j)$ of the metric tensor, we can express the metric property of ∇ as follows:

$$\begin{aligned} \partial_k g_{ij} &= \partial_k g(\partial_i, \partial_j) \\ &= g(\nabla_{\partial_k}\partial_i, \partial_j) + g(\partial_i, \nabla_{\partial_k}\partial_j) \\ &= \sum_\ell \Gamma^\ell_{ki} g_{\ell j} + \Gamma^\ell_{kj} g_{\ell i}. \end{aligned}$$

We can express this more simply in terms of another form of Christoffel symbols. If we define

$$\Gamma_{ijk} := \sum_\ell \Gamma^\ell_{ij} g_{k\ell} = g(\nabla_{\partial_i}\partial_j, \partial_k),$$

then we get $\partial_k g_{ij} = \Gamma_{kij} + \Gamma_{kji}$. Using the symmetry $\Gamma_{ijk} = \Gamma_{jik}$, we can solve this system to give

$$2\Gamma_{ijk} = \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}.$$

Writing $(g^{k\ell})$ for the matrix inverse of (g_{ij}), we have

$$\Gamma^k_{ij} = \sum_\ell g^{k\ell}\Gamma_{ij\ell} = \sum_\ell \frac{g^{k\ell}}{2}\Big(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}\Big).$$

C4. Parallel transport and holonomy

Suppose ∇ is any metric connection on a Riemannian manifold M, and γ is a smooth curve from p to q in M with velocity vector $X = \dot\gamma$. Suppose Y and Z are parallel fields along γ, meaning $\nabla_X Y = 0 = \nabla_X Z$. Then the metric condition implies

$$X\langle Y,Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle = 0,$$

that is, $\langle Y,Z\rangle$ is constant along γ. In particular, the length of any parallel field is constant, as is the angle between two parallel fields.

It follows that parallel fields along γ with respect to two different metric connections will thus differ from one another by some rotation – the torsion-free Levi-Civita connection is in some sense the one for which parallel fields rotate the least.

`End of Lecture 18 Jan 2016`

The metric condition means parallel transport $P_\gamma\colon T_pM \to T_qM$ is an orthogonal transformation between these inner product spaces; the holonomy around any loop based at p is an element of $O(T_pM)$. (If M is orientable, then the holonomy actually lives in $SO(T_pM)$. On an oriented surface, for instance, the holonomy around any loop is rotation by some angle θ.) Note that it is easy to extend the notion of parallel transport to the case of a piecewise smooth curve γ, simply by transporting along each smooth piece in order.

Consider the example of the round sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. Look at a right-angled triangle; at a lune of angle α; at a general triangle. In each case we find the holonomy is 2π minus the sum of the exterior angles, which also equals the enclosed area.

These are special cases of the Gauss–Bonnet theorem: for any disk D on any surface, the holonomy around ∂D is

$$2\pi - \mathrm{TC}(\partial D) = \int_D K\, dA.$$

On the sphere we have $K \equiv 1$ so the right-hand side is just the area.

The Gauss curvature at a point p on a surface can be measured by measuring the holonomy angle around a small loop based at p and dividing by the area of the loop.

C5. Riemannian curvature

The idea of Riemannian curvature is that given a two-plane in T_pM, the holonomy around an infinitesimal loop in this plane will give an infinitesimal rotation of T_pM. The two-plane is specified by a two-vector, and the infinitesimal rotation is given by an operator on T_pM saying in which direction each vector moves.

So suppose we have vector fields X and Y near $p \in M$. We consider parallel transport for time s along X followed by time t along Y, and compare this with going the other way around. Of course if $[X,Y] \neq 0$ this isn't even a closed loop, but let's assume $[X,Y] = 0$. Then the holonomy around this loop will be approximately st times what we call the curvature $R(X,Y)$.

In general, of course we need to correct by $[X,Y]$. Recall that this Lie bracket is the commutator of directional derivatives:

$$0 = X(Yf) - Y(Xf) - [X,Y]f.$$

This inspires the definition of the *Riemannian curvature operator*

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

Lemma C5.1. *On any Riemannian manifold the curvature operator $R(X,Y)Z$ is tensorial – its value at p depends only on X_p, Y_p and Z_p. In particular, $R(X_p,Y_p)$ is a linear operator on T_pM.*

The proof proceeds by checking that

$$R(fX,Y)Z = R(X,fY)Z = R(X,Y)(fZ) = fR(X,Y)Z,$$

which follows from the product rules we have for the Lie bracket and covariant derivative. The details are left as an exercise.

The definition implies directly that $R(X,Y) = -R(Y,X)$. This antisymmetry means that we can think of this linear operator as depending on $X \wedge Y$.

On an inner product space like T_pM, a linear operator is equivalent to bilinear form. Hence we get the *Riemannian curvature tensor*

$$R(X,Y,Z,W) := \langle R(X,Y)Z, W\rangle.$$

In coordinates (U,φ), with respect to the coordinate frame $\{\partial_i\}$, the curvature has components given by

$$R(\partial_k,\partial_\ell)\partial_i =: \sum R_{i\ k\ell}^{j}\partial_j,$$
$$\langle R(\partial_k,\partial_\ell)\partial_i,\partial_j\rangle =: R_{ijk\ell} = \sum_m g_{jm}R_{i\ k\ell}^{m}.$$

Theorem C5.2. *The Riemannian curvature satisfies the following symmetries:*

(1) $R(X,Y) = -R(Y,X)$,
(2) $\langle R(X,Y)Z,W\rangle = -\langle R(X,Y)W,Z\rangle$,
(3) $R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$,
(4) $\langle R(X,Y)Z,W\rangle = \langle R(Z,W)X,Y\rangle$.

Note that we have already proved the antisymmetry (1). The further antisymmetry (2) is equivalent to saying that $R(X,Y)$ is an infinitesimal rotation.

All four symmetries involve different permutations of the vector fields X, Y, Z, W, and are related to each other. (A more sophisticated approach would study them in terms of representations of the symmetric group S_4.) For instance, it is easy to see that, given (4), properties (1) and (2) are equivalent.

Instead, we start by observing that (4) is an algebraic consequence of the first three. For this, write (3) as

$$\langle R(X,Y)Z,W\rangle + \langle R(Y,Z)X,W\rangle + \langle R(Z,X)Y,W\rangle = 0.$$

Then cyclically permute $XYZW$ to get

$$\langle R(Y,Z)W,X\rangle + \langle R(Z,W)Y,X\rangle + \langle R(W,Y)Z,X\rangle = 0.$$

Add these two and subtract the remaining two cyclic permutations. Using the antisymmetries (1) and (2), the result follows.

End of Lecture 25 Jan 2016

Proof. It remains to show properties (2) and (3). By tensoriality, it suffices to prove (3) for the commuting basis vector fields $X = \partial_i$, $Y = \partial_j$, $Z = \partial_k$. We will abbreviate $\nabla_i := \nabla_{\partial_i}$. First note that $R(\partial_i,\partial_j)\partial_k = \nabla_i(\nabla_j\partial_k) - \nabla_j(\nabla_i\partial_k)$. Thus the sum of three terms can be written as

$$\nabla_i(\nabla_j\partial_k - \nabla_k\partial_j) + \nabla_j(\nabla_k\partial_i - \nabla_i\partial_k) + \nabla_k(\nabla_i\partial_j - \nabla_j\partial_i).$$

Because the connection is torsion-free, each of the expressions in parentheses is a Lie bracket like $[\partial_j,\partial_k]$, but these all vanish.

For (2) it also suffices to consider $X = \partial_i$, $Y = \partial_j$. Since the symmetric part of a bilinear form in determined by its associated quadratic form, to show the antisymmetry (2) it suffices to prove

$$0 = \langle R(\partial_i,\partial_j)Z,Z\rangle = \langle \nabla_i(\nabla_j Z) - \nabla_j(\nabla_i Z), Z\rangle.$$

That is, it suffices to prove that $\langle \nabla_i(\nabla_j Z), Z\rangle$ is symmetric in i and j. To do so, consider second derivatives of the function $\langle Z,Z\rangle$:

$$\begin{aligned}\partial_j(\partial_i\langle Z,Z\rangle) &= \partial_j(2\langle Z,\nabla_i Z\rangle)\\ &= 2\langle Z,\nabla_j(\nabla_i Z)\rangle + 2\langle \nabla_j Z,\nabla_i Z\rangle.\end{aligned}$$

The last term is clearly symmetric, and the left-hand side is symmetric since $[\partial_i,\partial_j] = 0$, so we are done. □

The antisymmetry properties (1) and (2) mean that the curvature tensor really can and should be thought of as a bilinear form on the space of two-vectors:

$$S(X\wedge Y, Z\wedge W) := -\langle R(X,Y)Z,W\rangle$$

(extended by bilinearity to nonsimple two-vectors). Property (4) is then simply the symmetry of S:

$$S(X\wedge Y, Z\wedge W) = S(Z\wedge W, X\wedge Y);$$

this symmetry of course holds for arbitrary two-vectors, not just simple ones. In these terms, property (3) gets no simpler:

$$S(X\wedge Y, Z\wedge W) + S(Y\wedge Z, X\wedge W) + S(Z\wedge X, Y\wedge W) = 0.$$

If X, Y is an orthonormal basis for a two-plane $\Pi \subset T_pM$, then $K(\Pi) := S(X\wedge Y, X\wedge Y) = R(X,Y,Y,X)$ is called the *sectional curvature* of Π. It turns out that this equals the Gauss curvature of the “flattest” surface $N \subset M$ with $T_pN = \Pi$, say foliated by geodesics through p tangent to Π.

Since any symmetric bilinear form is determined by the associated quadratic form, it is not surprising that the sectional curvatures of two-planes determine R completely – but note that here we are considering only simple two-vectors. The following lemma (applied to the difference of two possible tensors with the same sectional curvatures) takes care of this problem.

Lemma C5.3. *Suppose S is a symmetric bilinear form on $\Lambda^2 V$ satisfying*

$$S(X\wedge Y, Z\wedge W) + S(Y\wedge Z, X\wedge W) + S(Z\wedge X, Y\wedge W) = 0.$$

If $S(X\wedge Y, X\wedge Y) = 0$ for all X and Y, then $S = 0$.

Proof. First compute

$$\begin{aligned} 0 &= S\big((X+Y)\wedge Z, (X+Y)\wedge Z\big) \\ &= S(X\wedge Z, Y\wedge Z) + S(Y\wedge Z, X\wedge Z) \\ &= 2S(X\wedge Z, Y\wedge Z). \end{aligned}$$

Now using this we get

$$\begin{aligned} 0 &= S\big(X\wedge(Z+W), Y\wedge(Z+W)\big) \\ &= S(X\wedge Z, Y\wedge W) + S(X\wedge W, Y\wedge Z). \end{aligned}$$

Then we use this to show

$$\begin{aligned} S(Y\wedge Z, X\wedge W) = S(X\wedge W, Y\wedge Z) &= -S(X\wedge Z, Y\wedge W) \\ &= S(Z\wedge X, Y\wedge W). \end{aligned}$$

That is, S is invariant under a cyclic permutation of XYZ. But we have assumed the sum of all three cyclic permutations is zero, so we find $S = 0$. □

Note that actually there is a formula with 16 terms giving

$$\begin{aligned} &6S(X\wedge Y, Z\wedge W) \\ &\quad = S\big((X+Z)\wedge(Y+W), (X+Z)\wedge(Y+W)\big) \\ &\qquad - S\big((X+Z)\wedge Y, (X+Z)\wedge Y\big) + \cdots . \end{aligned}$$

C6. The exponential map

Any tangent vector $X_p \in T_pM$ determines a unique geodesic starting at p with velocity X; this is the solution to a second order ODE and will exist at least for some time $\delta > 0$.

More precisely, we can consider the equation for a geodesic $x(t)$ in a coordinate chart (U, φ). In terms of the Christoffel symbols we get

$$\ddot{x}^k + \sum_{i,j} \Gamma^k_{ij}(x)\dot{x}^i\dot{x}^j = 0.$$

The existence theorem we have used before then says: For any $q \in U$ there is a neighborhood $V \ni q$ and $\varepsilon, \delta > 0$ such that, given any initial conditions $p \in V$ and $X_p \in T_pM$ with $\|X_p\| < \varepsilon$ (meaning of course $x(0) = phi(p)$ and $\dot{x}(0) = \varphi_* X_p$), a unique solution as above exists within U for $|t| < \delta$, and it depends smoothly on the initial conditions.

Of course if we rescale X_p to λX_p, the geodesic is still the same curve, simply parametrized at λ times the speed. Thus there is a tradeoff between ε and δ; the statement above is also true say for $\delta = 2$. Note also that we can piece together geodesic arcs and the resulting curve is still a geodesic, so we can apply the existence results independent of any coordinate chart.

Suppose the geodesic γ starting at p with velocity X_p exists for at least unit time. Then we define $\exp(X_p) := \gamma(1)$ to be the point at time 1 along this geodesic. For any Riemannian manifold M, this *exponential map* is defined on some open neighborhood W of the zero-section of TM and is a smooth map $\exp\colon W \to M$. Note that $t \mapsto \exp(tX_p)$ is the parametrized geodesic with constant speed $\|X_p\|$; the length of this curve from p to $\exp(X_p)$ is $\|X_p\|$.

We use the notation $\exp_p$ for the restriction of exp to $T_pM \cap W$. Since the differential $D_p \exp_p$ is the identity map, for small $\varepsilon > 0$, $\exp_p$ is a diffeormorphism from the ε-ball in T_pM to a neighborhood $B_\varepsilon(p)$ called a *geodesic ball*. Choosing an orthonormal basis for T_pM to identify it with $\mathbb{R}^m$, the map $\exp_p^{-1}\colon U \to T_pM = \mathbb{R}^m$ is called *normal coordinates* around p on the *normal neighborhood U*.

Similar considerations show that the map $(\pi, \exp)\colon X_p \mapsto (p, \exp X_p)$ is a local diffeomorphism $TM \to M \times M$. For any $p \in M$ and $\varepsilon > 0$, we deduce the existence of a neighborhood N such that any two points in N are joined by a unique geodesic of length less than ε.

In normal coordinates around p, all geodesics through p are the images of straight lines through the origin. Furthermore, at p we have $g_{ij}(0) = \delta_{ij}$ and $\Gamma^k_{ij} = 0$. The image $S_r := \{\exp_p X_p : \|X_p\| = r\}$ of a sphere in T_pM is called a *geodesic sphere* around $p \in M$. It is easy to check that geodesics through p meet each of these sphere orthogonally and that (for small $r < r'$) any curve from S_r to $S_{r'}$ has length at least $r' - r$. Thus geodesics are locally shortest curves.

Note that for small r, the geodesic sphere $S_r(p)$ is topologically a sphere and is the boundary of the geodesic ball $B_r(p)$, which is a topological ball. For larger r, the exponential map may still exist but no longer be a diffeomorphism; these spheres and balls will start to overlap and intersect themselves. The geodesic ball $B_r(p)$ is always the metric ball in (M, d), the set of points at distance less than r from p.

So far, we have only discussed local existence of geodesics. A manifold is called *geodesically complete* if every geodesic can be extended indefinitely, that is, if exp is defined on all of TM. The Hopf–Rinow theorem says this happens if and only if the metric space (M, d) is metrically complete. In particular, every compact manifold is geodesically complete.

C7. Ricci and scalar curvatures

If $\Pi \subset T_pM$ is a two-plane, its image under the exponential map $\exp_p$ is locally a two-dimensional submanifold through p. One can show that its Gauss curvature at p is the sectional curvature $K(\Pi)$. (See for instance Boothby, Theorem VIII.4.7.)

Suppose $X \in T_pM^m$ is a unit vector. We define the Ricci curvature in direction X to be the average sectional curvature of two-planes including X:

$$\operatorname{Ric}(X, X) = (m-1)\operatorname{ave}_{\Pi \ni X} K(\Pi).$$

A manifold with constant Ricci curvature is called an Einstein manifold, because of the way this condition arises in general relativity.

The Ricci curvature is a quadratic form; the associated symmetric bilinear form on T_pM is called the Ricci tensor; if $\{e_i\}$ is an orthonormal frame, we have

$$\mathrm{Ric}(X, Y) = \sum_{i=1}^{m} S(X \wedge e_i, Y \wedge e_i).$$

In coordinates, $\mathrm{Ric}_{ij} = \sum_k R^k_{i\,kj}$. On an Einstein manifold, $\mathrm{Ric}(X, Y) = cg(X, Y)$.

For $m = 3$, but not in higher dimensions, the Ricci curvature determines the full Riemannian curvature tensor.

`End of Lecture 1 Feb 2016`

The Ricci flow $dg_{ij}/dt = -2\,\mathrm{Ric}_{ij}$ is a nonlinear heat flow of Riemannian manifolds which tries to smooth out the Ricci curvature. Perelman's proof of Thurston's geometrization conjecture (including the Poincaré conjecture) used Ricci flow on 3-manifolds.

The average Ricci curvature in all directions at p gives the *scalar curvature* $S(p) = \sum g^{ij} R_{ij}$.

The images under the exponential map of round balls and spheres in T_pM are called geodesic balls and spheres – $B_r(p)$ is the set of all points at distance less than r from p.

The scalar curvature measures the volume growth rate of geodesic balls or spheres around p – positive curvature means the volume grows more slowly than in Euclidean space, which negative curvature means it grows faster.

The Ricci curvature $\mathrm{Ric}(X, X)$ measures in a similar sense the rate of spreading of geodesics emerging from p in directions near X.