Geometry for General Relativity

Sam B. Johnson
MIT 8.962, Spring 2016
Professor: Washington Taylor
Dated: 2/19/2016

samj at mit.edu

Contents

1 Purpose

The purpose of these notes is to assemble a simple, self-contained reference for all the math used in General Relativity. The goal is to be concise and clear, being neither too abstract to be accessible nor too concrete to be easily generalized. The intended audience is `8.962` students at MIT in Spring 2016. However, anyone with a general interest and a grounding in 1) physics; 2) multivariable calculus; 3) linear algebra should be able to read these notes. In terms of material covered, the notes begin with the basic definitions of manifolds, then introduce the idea of fiber bundles and connections, and finally connect to gerneral relativity by realizing it as a gauge theory. Along the way, many practical formulas are discussed.

These notes will update as the semester progesses. The current draft is incomplete and contains skeleton sections, which may nonetheless be interesting as previes of upcoming material.

2 Notation

In these notes, I will follow standard mathematical conventions. This section defines standard terms and symbols. It can be skipped by those familiar with math, and/or used as a glossary as necessary.

- A "set" A is an unordered collection of objects, without duplicates. (In other words, adding an object to a set that already contains the object does nothing.) Let A and B be sets.
  - $x \in A$ denotes that the object x belongs to the set A. E.g. if $A = \{1, 2, 3\}$, then $1 \in A$ but $4 \notin A$.
  - Objects that belong to a set are called its **elements**. Thus if $x \in A$, then x is an element of the set A.
  - When writing sets by explicitly listing their elements, one writes the list in curly braces: E.g. $A = \{1, 2, 3\}$.
  - $A \subset B$ denotes that every object in A is also in B. In this case, A is called a **subset** of B.
  - $A \cup B := \{x \text{ such that } x \in A \text{ or } x \in B\}$. $A \cup B$ is called the **union** of A and B.
  - $A \cap B := \{x \text{ such that } x \in A \text{ and } x \in B\}$. $A \cap B$ is called the **intersection** of A and B.
  - The symbol $\emptyset$ is used to denote a set with no elements. I.e. $\emptyset = \{\}$. This immediately implies that $\emptyset \cup A = A$ for any set A and $\emptyset \cap A = \emptyset$ for any set A.
- $\mathbb{R}$ denotes the real line (i.e. the set of real numbers).
- The symbol $|$ generally stands for the phrase "such that." Therefore I might write $x \in \mathbb{R}|x^2 = x$. The only x which satisfy this condition are 0 and 1.

- General spaces (manifolds) will usually be denoted by capital letters M or N. The product $\times$ of these spaces is the space $M \times N := \{(p,q)|p \in M, q \in N\}$.

- The product of n copies of a space M with itself is sometimes written $M \times M \times \cdots \times M = M^n$.

- $\mathbb{C}$ denotes the set of complex numbers, which as a vector space is identical to $\mathbb{R}^2$.

- The symbol $\exists$ is shorthand for the words "there exists." For instance, I might write, for every vector $v \in \mathbb{R}^n$, $\exists$ a vector $w \in \mathbb{R}^n | v + w = 0$.

- The symbol $\forall$ is shorthand for the words "for all." For instance, I might write $\forall v \in \mathbb{R}^n$, $\exists w \in \mathbb{R}^n | v + w = 0$.

- The symbol $:=$ indicates that the quantity on the lhs is *defined* to be the quantity on the rhs. For instance, the 2-sphere S^2 is defined to be the set of points in $\mathbb{R}^3$ of unit distance from the origin. Therefore I can write $S^2 := \{p \in \mathbb{R}^3 || p|^2 = 1\}$.

- A **map** f from set A to set B is a rule that associates to each $x \in A$ a single element $f(x)$ in B. If f is a map from A to B, this is written

  $$f : A \to B$$

  If you like, you can think of map as another word for function.

3 Manifolds = coordinate charts

A manifold is a geometric space that is locally like $\mathbb{R}^n$ for some dimension n. This means if we zoom in close enough to a point on the shape, the shape becomes indistinguishable from $\mathbb{R}^n$. A sphere S^2 (hollow inside; e.g. the material of a balloon) is an example of a 2-dimensional manifold because if you zoom in on any point, it looks like a plane $\mathbb{R}^2$. We are familiar with this fact because we live on the Earth!

This means that around any point, we can put down a system of coordinates, valid in some (possibly small) vicinity, is just a patch of $\mathbb{R}^n$. For instance, we can describe nearly all of the sphere (except the north and south poles) by coordinates $(\theta \in (0, \pi), \phi \in (0, 2\pi))$, which is just an open rectangle in $\mathbb{R}^2$.

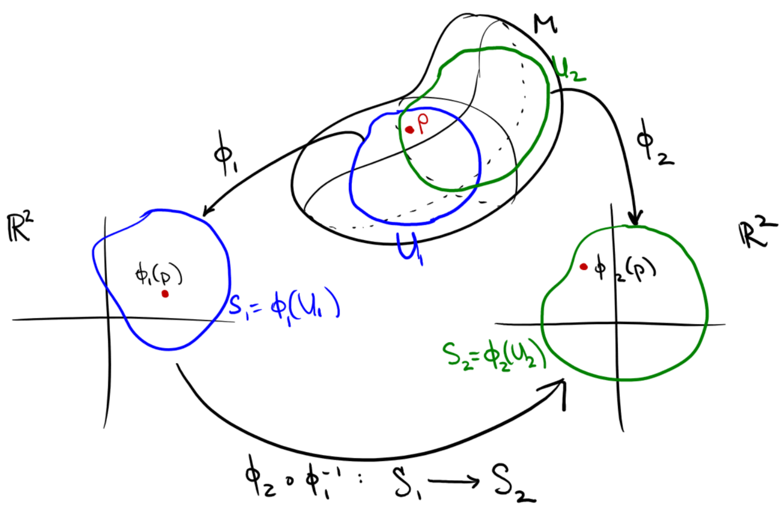
To describe a manifold M of dimension n, the idea is to cover the shape with open subsets $\{U_\alpha\}$, called **coordinate charts**:

- each U_α must be mapped by a continuous, invertible map to an open subset of $\mathbb{R}^n$. Call this map $\phi_\alpha : U_\alpha \to S_\alpha \in \mathbb{R}^n$.

- every point $p \in M$ must belong to at least one of the U_α.

- when two coordinate charts intersect, say $p \in U_{1,2} = U_1 \cap U_2$, we can express p in either coordinate chart. Indeed, $\phi_1(p)$ is p's coordinates in chart 1, and $\phi_2(p)$ is p's coordinates in chart 2. Therefore, if we know the coordinates of p in one chart, we can find them in the other, using

$$\phi_2(p) = (\phi_2 \circ \phi_1^{-1})(\phi_1(p))$$

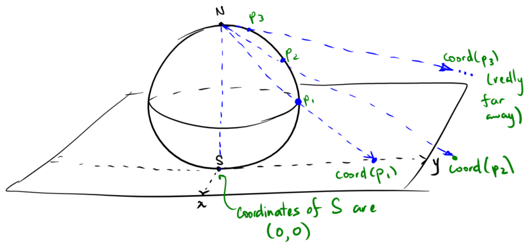


Example: coordinates on S^2. There are many different choices of coordinates on the same shape! But by definition, unless our shape can be continuously deformed to an open subset of $\mathbb{R}^n$ for some n, then we'll need more than one coordinate chart to cover the whole thing!

- "Stereographic projection" (2 coordinate charts) idea: map to complex plane from the north pole. Works for every point but the north pole N. So two coordinate charts: one that projects from the north pole and covers the whole sphere except for N: $U_N = S^2 - N$ and $U_S = S^2 - S$. If we imagine that the sphere lives in 3 dimensions, with its south pole coinciding with the origin of the plane, then $x_N(p) = 2p/(p_z - 2)$. Similarly, $x_S(p) = 2p/p_z$. Of course, in each case, we can calculate $p_z = \sqrt{1 - p_x^2 + p_y^2}$ since we're on the sphere. Therefore, to go from N coordinates to S coordinates, we apply the map

$$\begin{aligned} x_S \circ x_N^{-1}(x_1, x_2) &= x_s(\left(x_1 \frac{4}{|x|^2(1+4/|x|^2)}, x_2 \frac{4}{|x|^2(1+4/|x|^2)}, 1 \pm \sqrt{1 - (\frac{4}{|x|^2(1+4/|x|^2)})^2} \right)) \\ &= 4(x_1, x_2)/|x|^2 \end{aligned}$$

- "Flat projection": (6 coordinate charts) this one is easier. Just consider the sphere S^2 as living in 3d, construct coordinate charts in (x,y), (x,z), and (y,z), such that each



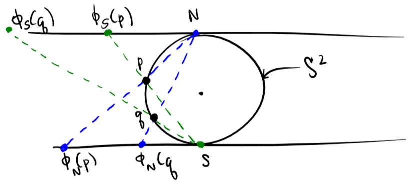


chart just takes those two coordinates of points. Call these $U_{\pm i}$, where i is the omitted coordinate and $\pm$ indicates that we take from the half of the sphere with positive (or negative) values of the coordinate i. To go back to 3d coordinates from say, U_{-x}, we just set $(x, y, z) = (-\sqrt{1 - y^2 - z^2}, y, z)$. We can go from coordinates in this chart (y, z) to e.g. U_{z+} coordinates

$$(x, y) = (-\sqrt{1 - y^2 - z^2}, y)$$

3.1 Maps between manifolds

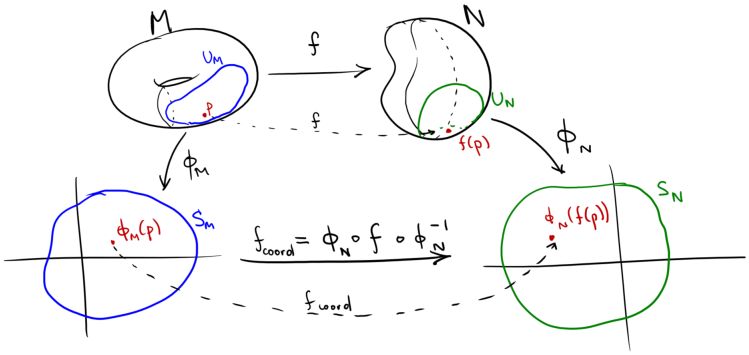
A manifold is called **smooth** if all its transition functions $\phi_\alpha \circ \phi_\beta^{-1} : S_\beta \to S_\alpha$ are smooth, e.g. infinitely differentiable. This is possible to check using standard calculus because the S_α are just subsets of $\mathbb{R}^N$.

We can also consider maps between manifolds $f : M \to N$. Abstractly, such a function is merely a rule that assigns a unique point $f(p)$ to each $p \in M$. Fortunately, we can write f in coordinates if we so choose. How? Well in a neighborhood around $p \in U_\alpha \subset M$, f (if continuous) maps p into a neighborhood of $f(p) \in W_\beta \subset N$. Then just take coordinates on both sides!

$$\begin{aligned} f_{coord} :& \quad S_\alpha \to S_\beta \\ f_{coord} =& \; \phi_\beta \circ f \circ \phi_\alpha^{-1} \end{aligned}$$

A map f between manifolds is **smooth** if its coordinate representations are all smooth.

Summary: we can easily put coordinates on any shape that is locally like $\mathbb{R}^n$ for some n. These shapes are called manifolds. We can define maps between these spaces either abstractly or in coordinates, and go easily between either description. Smoothness properties are all defined with respect to coordinates.



Exercises:

- (easy) Compute the identity map $\mathbb{I} : S^2 \to S^2$ that takes every point p on the sphere to itself: $\mathbb{I}(p) = p$. Note that this map is defined independent of coordinates. However, show that it is smooth by writing it in coordinates by using on the first S^2 flat projection and on the second S^2 sterographic projection. I.e., compute the function $\mathbb{I}_{coord}$:$\mathrm{S}_N \to S_{-z}$, which is just a function from $\mathbb{R}^2$ to the open disc in $\mathbb{R}^2$. Check that this function is indeed differentiable at all points where it is defined.

- (easy) Consider the map $-\mathbb{I} : S^2 \to S^2$ that takes a point p to $-p$ (assuming the sphere is centered on the origin. Do this in both sets of coordinates we discussed!

- (regular) Convince yourself (and a friend!) that stereographic projection works for all spheres S^n and requires only two coordinate charts in each case. (Project from one point and its antipodal point.)

- (!advanced!) This problem is about a special map called the Hopf map between spheres, $\eta : S^3 \to \mathrm{S}^2$. The map η is defined as follows: first picture S^3 as the unit sphere in $\mathbb{R}^4 = \mathbb{C}^2$. Then $S^3 = \{(z_0, z_1) \in \mathbb{C}^2 | |z_0|^2 + |z_1^2| = 1\}$. We can also think of the two-sphere S^2 as living in $\mathbb{C} \times \mathbb{R} = \mathbb{R}^3$. Given this, we can define η as follows:

$$\eta(z_0, z_1) = (2z_0\bar{z}^1, |z_0| -^2 |z_1|^2) \tag{3.1}$$

  Your mission: check that η maps to $S^2 \subset C \times \mathbb{R}$. Moreover, check that η is smooth by writing it in flat projection coordinates on both spheres and differentiating.

4 Vector fields, one-form fields, general tensors; maps; Stoke's theorem

Now that we understand generally *what* a manifold is and how to describe maps between manifolds in coordinates, we can discuss the many other objects that can live on such a manifold. The first of these is a vector field.

4.1 Tangent spaces are different at different points!

The idea of a tangent space is very simple. Think of a curvy manifold M of dimension n; each point $p \in M$ has a unique vector space of dimension n tangent to p. (See the picture.) Thinking either abstractly or visually, by imagining e.g. a two-dimensional surface $M \subset \mathbb{R}^3$, it becomes clear that *tangent spaces at different points on M are* **different** *spaces, e.g. different linear subspaces of* $\mathbb{R}^3$.

4.2 Vectors are directional derivatives

The **tangent space at p** is denoted by T_pM and is defined to be the linear vector space with basis $\{\partial_i|_p\}$, the partial derivatives (in coordinates) of functions at p. Therefore a vector is of the form

$$V = V^i \partial_i$$

Why this definition? This specifies an actual direction on M at p. (See the picture.) Moreover, once we make this definition, it tells us precisely how the coefficients V^i must change when we change coordinates in order for the actual vector not to change.

$$\begin{aligned} V &= V^i \partial_i \\ &= V^{j'} \partial_{j'} \\ &= V^{j'} \frac{\partial x^i}{\partial x^{j'}} \partial_{x^i} \\ \Rightarrow \quad V^{j'} &= \frac{\partial x^{j'}}{\partial x^i} V^i \end{aligned}$$

A **vector field** is a (smoothly varying) choice of vector for each $p \in M$, i.e. $V = V(p)$. In coordinates, one can imagine $V = V^i(x)\partial_i|_x$.

4.3 One-forms are differentials

A one-form ω at $p \in M$ is an element of the dual vector space T_p^*M of T_pM. Of course, like any two vector spaces of the same dimension, they are isomorphic, but no canonical isomorphism between them exists. The dual space W^* to a vector space W is defined to be $\omega : W \to \mathbb{R}$ the set of linear functions from W to the reals R. By linearity, any such functional ω is uniquely defined by the values it takes on a basis of W. $\omega(\partial_i) = \omega_i$.

It is convenient to introduce **differentials**, which are infinitesimal displacements; these form a basis for the space of one-forms, due to their transformation properties. Given a coordinate system $\{x^1, ..., x^n\}$, one gets differentials $\{dx^1, ..., dx^n\}$. In fact, one could simply *define* the differentials to be the functionals in T_p^*M such that $dx^i(\partial_j) = \delta^i_j$, in which case they are automatically a basis for T_p^*M. Based on this definition, or by heuristics, these transform as

$$dx^{j'} = \frac{\partial x^{j'}}{\partial x^i} dx^i$$

This is just as it should be: the differentials are expressed as

$$\begin{aligned} \omega &= \omega_i dx^i \\ \text{so:} \quad \omega_{j'} &= \frac{\partial x_i}{\partial x^{j'}} \omega_i \\ \text{compare to:} \quad V^{j'} &= \frac{\partial x^{j'}}{\partial x^i} V^i \end{aligned}$$

Given a one-form $\omega \in T_p^*M$, I've said that it is a linear functional on T_pM. Thus, we ought to be able to calculate $\omega(V) \in \mathbb{R}$ for any vector $V \in T_pM$. With these definitions, it is easy:

$$\begin{aligned} \omega(V) &= \omega_i dx^i(V^j \partial_j) \\ &= V^j \omega_i dx^i(\partial_j) \\ &= V^j \omega_i \delta^i_j \\ &= V^i \omega_i \end{aligned}$$

Notice that becuase of the opposite transformation properties of of V^i and ω_i, this quantity $\omega(V)$ is independent of coordinates; it is a scalar quantity.

Important point: A vector field $V(x)$ does not change under coordinate transformations! Only its components change! Similarly, a one-form ω does not change either! Only its components change. The total expression $\omega_i dx^i$ is defined in such a way that the dx^i transforms in an oppposite way to the ω_i, so the total expression does not change at all under coordinate transformations!

A **one-form field**, usually just called a one-form, is a smoothly varying choice of $\omega(x) \in T_x^*M$ for all x in M. E.g. the components in coordinates $\omega_i(x)$ should be differentiable functions.

Because the bases $\{\partial_i\} \leftrightarrow \{dx^j\}$ are dual, we can also simply define $\partial_i(dx^j) = \delta^j_i$.

4.4 Tensors are multi-linear functionals of vectors and one-forms

We could think of one-forms or co-vectors as linear functionals from $T_pM \to \mathbb{R}$. We can do much more. A **tensor** is a linear map from $T_xM \otimes T_xM \otimes \cdots \otimes T_xM \otimes T_x^*M \otimes \cdots T_x^*M$ to $\mathbb{R}$.

- Aside: what is a tensor product of vector spaces V (dim n) and W (dim m)? This is written $V \otimes W$. If V has basis $\{e_i\}$ and W has basis $\{f_j\}$, then the new vector space has basis $\{e_i \otimes f_j\}_{i \in [1,n], j \in [1,m]}$.

- Thus, $V \otimes W$ is $n \times m$-dimensional.
- Note well! In the vector space $V \otimes V$, the basis element $e_i \otimes e_j \neq e_j \otimes e_i$ for $i \neq j$. (This often confuses people.)
- The tensor product acts on pairs of vectors as follows: $v \otimes w = v^i w^j e_i \otimes f_j$.

Basically, though, you can just think of a tensor as an object with U slots that accept one-forms and L slots that accept vectors. When each of the L slots is passed a vector as an argument, and each of the U slots is passed a one-form, a tensor T spits back a number in $\mathbb{R}$ With some rules:

$$\begin{aligned} T(aV_1, V_2, ...V_L, \omega_1, \omega_2, ..., \omega_U) &= aT(V_1, V_2, ...V_L, \omega_1, \omega_2, ..., \omega_U) \\ T(V_1 + V_1', V_2, ...V_L, \omega_1, \omega_2, ..., \omega_U) &= T(V_1, V_2, ...V_L, \omega_1, \omega_2, ..., \omega_U) + T(V_1', V_2, ...V_L, \omega_1, \omega_2, ..., \omega_U) \end{aligned}$$

And these properties of linearity hold for each argument, so if I chose to replace ω_3 by $a\omega_3 + b\omega_3'$, I would get $aT(\omega_3) + bT(\omega_3')$ (all other arguments suppressed.)

Since the tensor T is linear in each argument, it is determined by its value on every combination of basis elements of each vector space. So we can calculate T on any combination of vectors/one-forms as long as we know the following $n^{(L+U)}$ numbers:

$$T_{i_1 i_2 \ldots i_L}{}^{j_1 j_2 \ldots j_U} = T(\partial_{i_1}, \partial_{i_2}, ..., \partial_{i_L}, dx^{j_1}, dx^{j_2}, ..., dx^{j^U})$$

Each i_1 etc. can take values from 1 to n (the dimension of M), so there are total $n^{(U+L)}$ such numbers that characterize at tensor T. Note that U and L can be any integers greater than or equal to zero. In fact,

$$T = T_{i_1 i_2 \ldots i_L}{}^{j_1 j_2 \ldots j_U} dx^{i_1} \otimes dx^{i_2} \otimes \cdots \otimes dx^{i_L} \otimes \partial_{j_1} \otimes \partial_{j_2} \otimes \cdots \otimes \partial_{j_U}$$

It is easy to check the above statement by using the relations $dx^j(\partial_i) = \partial_i(dx^j) = \delta_i^j$. All there is to it is evaluating T on some vectors and one-forms, $T(V_1, V_2, \ldots, V_L, \omega_1, \omega_2, \ldots, \omega_U)$ and expanding each argument in its basis. Then, you will see that the ∂_i's of V's get eaten by the dx^is and vice versa for the dx^j's of the ωs.

- When a tensor with only lower indices is symmetric in all its arguments, $dx_i \otimes dx^j \otimes \cdots$ may be replaced with its symmetrized version instead

  $$dx^{i_1} dx^{i_2} \cdots dx^{i_N} := \frac{1}{N!} \sum_{\sigma \in S_N} dx^{i_{\sigma(1)}} \otimes dx^{i_{\sigma(2)}} \otimes \cdots \otimes dx^{i_{\sigma(N)}}$$

  The notation is chosen to indicatethat ordering does not matter (as in regular multiplucation).

- When a tensor with only lower indices is antisymmetric in any two indices, its $dx^{i_1} \otimes dx^{i_2} \otimes \cdots \otimes dx^{i_N}$ can be freely replaced with the *anti*symmetrized version:

$$dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_N} := \frac{1}{N!} \sum_{\sigma \in S_N} (-1)^\sigma dx^{i_{\sigma(1)}} \otimes dx^{i_{\sigma(2)}} \otimes \cdots \otimes dx^{i_{\sigma(N)}}$$

  The symbol $(-1)^\sigma$ is -1 if the permutation σ can be implemented with an odd number of transpositions (swaps of two items), and $+1$ if it can be implemented with an even number.

Summary: vectors $V(x)$ and one-forms $\omega(x)$ are the simplest objects that can live on manifolds other than scalar-valued functions $f(x)$. Given a coordinate chart $\{x^i\}$, $\{\partial_i\}$ form a basis for the tangent space T_pM in which vectors live; and $\{dx^i\}$ form a basis for the cotangent space T_p^*M in which one-forms live. By definition $dx^j(\partial_i) = \partial_i(dx^j) = \delta_i^j$; this directly implies that one-forms and vectors transform in opposite ways (i.e. with inverse matrices).

4.5 Caution: a note on tensor product spaces

It is very important to notice that our definition of tensor product $V \otimes W$ of vector spaces V and W was quite carefully worded. The statement is that if V (dim n) has basis $\{e_1, e_2, \ldots, e_n\}$, and W (dim m) has basis $\{f_1, f_2, \ldots, f_m\}$, then $V \otimes W$ is *defined* to be the space with basis $\{(e_i, f_j)\}_{i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}}$. It is clear that there are $n \times m$ basis elements of $V \otimes W$ and therefore that $V \otimes W$ has dimension nm.

These basis elements are not commonly written as doubles (i.e. ordered pairs) as I have done above. Instead, the basis element (e_i, f_j) is usually written $e_i \otimes f_j$. This is nothing more than a change of notation. However, it is convenient to extend this operation to act on vectors. So instead of just taking tensor products of spaces, we will also be able to take the tensor product of two vectors, one from each space. Take $v = v^i e_i \in V$ and $w = w^j f_j \in W$. Then I *define* their tensor product as follows:

$$\begin{aligned} v \otimes w &= v^i w^j e_i \otimes w_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} v^i w^j e_i \otimes w_j \end{aligned}$$

(In the second line, I just explicitly wrote the summations that are usually implied by the Einstein summation convention.) From this definition, it is clear that this operation $\otimes$ on vectors is distributive and associative in the same way that regular multiplication is. However, it is not commutative like ordinary multiplucation of real numbers; rather it is not commutative, like multiplication of matrices, for instance.

Now we can ask: is it true that $V \otimes W = \{v \otimes w \mid \forall v \in V, \forall v \in W\}$? If true, this would say that every element in the tensor product space $V \otimes W$ can be written as $v \otimes w$ for the appropriate choice of $v \in V$ and $w \in W$. It turns out that this is false. Because this confusion is so easy and so important, I want to pause to explain exactly why this reasonable sounding statement is false.

Fact: $V \otimes W \neq \{v \otimes w \mid \forall v \in V, \forall v \in W\}$. In fact,

$$V \otimes W \supsetneq \{v \otimes w \mid \forall v \in V, \forall v \in W\}$$

The set on the right, which I'll call $V \oplus W$, is actually a *strict subset* of $V \otimes W$ which has dimension $n + m$. In the case of $T^*M \otimes T^*M$ on a 4-dimensional manifold M, this is an 8-dimensional subset of the 16-dimensional space $T^*M \otimes T^*M$.

So why are these two sets unequal? I'll give two proofs.

Proof 1 This first proof constructs an explicit counterexample – I will construct a vector $z \in V \otimes W$ that *cannot* be written as $v \otimes w$ for *any* combination of $v \in V$ and $w \in W$. It will suffice to prove it in the interesting case where each space is two-dimensional. Take for instance $V = W = \mathbb{C}^2$. We can think of this as the state-space of a two-spin electron (up to normalization of the states). Then the tensor product space $V \otimes W = \mathbb{C}^2 \otimes \mathbb{C}^2$ can be thought of as the space of a two-electron system. Let's write the basis elements as $|-\rangle$ and $|+\rangle$. Now consider the following state Q in $V \otimes W = \mathbb{C}^2 \otimes \mathbb{C}^2$:

$$Q := |+\rangle \otimes |-\rangle - |-\rangle \otimes |+\rangle$$

(To be a normalized state, we would need to divide this whole thing by $\sqrt{2}$, but that is immaterial for this discussion.) Can this state be written as $v \otimes w$ for two one-electron states v and w? No. We can see this by trying and failing to write it this way. Take $v = a|+\rangle + b|-\rangle$ and $w = c|+\rangle + d|-\rangle$. Then

$$v \otimes w = ac(|+\rangle \otimes |+\rangle) + ad(|+\rangle \otimes |-\rangle) + bc(|-\rangle| \otimes |+\rangle) + bd(|-\rangle \otimes |-\rangle)$$

By definition, $|+\rangle \otimes |+\rangle$ is linearly independent from $|+\rangle \otimes |-\rangle$. (Indeed, all four terms above involve a coefficient multiplying vector that is linearly independent from all the others in this tensor product space. That is why I did not further collect terms after distributing.) Because Q has no term proportional to $|+\rangle \otimes |+\rangle$, we must set either $a = 0$ or $c = 0$. But now we're already in trouble. Our state will now look like either $b(c|-\rangle \otimes |+\rangle + d|-\rangle \otimes |-\rangle)$ or $d(a|+\rangle \otimes |-\rangle + b|-\rangle \otimes |-\rangle)$. Neither of these is the state we want, since both terms have $|-\rangle$ as the first factor. So it is impossible to write Q as $v \otimes w$. This is a very important fact. If you've ever heard about quantum entanglement, then consider this: it would be impossible if Q could be written as $v \otimes w$ for v and w single-electron states!

Proof 2: Dimension counting. This is easy, but not as general as our above proof, since $2 + 2 = 2 \times 2$, so dimension counting alone will be sufficient to conclude that $V \otimes W \neq V \oplus W$ when V and W are two-dimensional. Just take components and write out the matrix of coefficients, e.g. in four dimensions, so $v = (v_1, v_2, v_3, v_4)$, $w = (w_1, w_2, w_3, w_4)$:

$$v \otimes w = \begin{pmatrix} v_1w_1 & v_1w_2 & v_1w_3 & v_1w_4 \\ v_2w_1 & v_2w_2 & v_2w_3 & v_2w_4 \\ v_3w_1 & v_3w_2 & v_3w_3 & v_3w_4 \\ v_4w_1 & v_4w_2 & v_4w_3 & v_4w_4 \end{pmatrix}$$

As you can see, this matrix has 16 entries, but only eight degrees of freedom. The point is that there would be 16 degrees of freedom if I were allowed to choose a different w for every *component* of v, but I am restricted to choosing, for any v, a w.

NB: A final note: this section was added to clarify a confusion that arose in Monday's recitation 3/14/2016. It is possible that only I was confused, and not any students. However, this point is so fundamental that I thought it was necessary to emphasize

$$V \otimes W \neq \{v \otimes w \mid \forall v \in V, \forall v \in W\}$$

since it would be a real shame if even one student were confused on this point.

Exercises:

- Prove the statement I made about tensor components!
- Vector fields on manifolds M not equal to $\mathbb{R}^n$ have very intriguing behavior, which frequently reveals some properties about M. Here's a classic example called "You can't comb a hedgehog!"
  - First, remember that $S^n := \{\vec{x} \in \mathbb{R}^{n+1} | \, |\vec{x}|^2 = 1\}$
  - Now, try to convince yourself (or a friend) that you every vector field on S^2 has to be zero at at least one point. Think geometrically!
  - Construct an explicit vector field on S^1 that doesn't vanish anywhere. Now construct such a vector field for S^{2n+1} (i.e. odd-dimensional spheres).
  - Now, if you're ready for more, try to prove that any vector field on a sphere S^{2n} (integer n) has to vanish at at least one point!

  If we imagine a Hedgehog with long hair growing everywhere, then we could try to comb the hair so it is laying flat everywhere (tangent to the hedgehog). But then this makes the hair-field a tangent vector field to something that's basically S^2, and if the hair is long everywhere, then this vector field would not vanish anywhere. So–what we just proved says this is impossible! What would happen if we performed this experiment on a very hairy hedgehog? Well, basically in at least one place, the hair would be sticking straight out from the Hedghog (i.e. the hair-field would have zero component tangent to the hedgehog).

  The above exercises are telling us something very interesting about spheres–vector fields on spheres behave very differently from vector fields on $\mathbb{R}^n$.

5 Derivatives on manifolds

5.1 Without metrics

There is a special kind of tensor known as a differential form. These are amenable to both integration and differentiation. Moreover, these operations can always be performed, *even without a metric!*

A **differential form** is a tensor with only lower indices, antisymmetric under swapping any two. Differential forms are allowed to have any number of lower indices, including zero; the number of indices is called the **rank** or **degree** of the form, and a form of degree k is sometimes referred to as a *k-form.* (0-forms are just functions on M.) According to our previous conventions, we can write a k-form as:

$$\omega_{(k)} = \omega_{i_1 i_2 \dots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k}$$

Fact: There are $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ linearly independent k-forms on an n-dimensional manifold M. The reason is that in a differential like $dx^i \wedge dx^j$, i and j must be distinct (by antisymmetry) and also this is equal to $-dx^j \wedge dx^i$. Therefore i can take any value in $\{1, ..., n\}$, but once this value is chosen, j must be one of the others. Finally, the two orders are not independent; so in this simple example, the number of independent 2-forms is $n(n-1)/2$. By noticing that all reorderings of indices are related by (possibly) minus signs, you can convince yourself that the number of independent k-forms is just the number of ways to pick k *distinct* indices from n, neglecting order.

- **Important corollary:** there is only $1 = \binom{n}{n}$ n-form on an n-dimensional manifold.

- **Another important corollary:** there are no k-forms with $k > n$ on an n-dimensional manifold. By antisymmetry, no index can appear twice in a differential $dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_{n+1}}$, but clearly at least one index would have to appear twice in this hypothetical differential, since we only have n to choose from!

The **exterior derivative** is the first (and most important!) derivative we will meet. It can always be defined in a coordinate-invariant way. The exterior derivative is denoted by "d." Acting on a function, it produces a one-form that behaves much like the gradient.

$$f \xrightarrow{d} df = (\partial_i f) dx^i$$

This is a 1-form (with coefficients $\partial_i f$). So it should be able to eat up a vector and return a number. Let's see what happens when we feed it a vector:

$$\begin{aligned} df(V) &= \partial_i f dx^i (V^j \partial_j) \\ &= V^j \partial_i f dx^i(\partial_j) \\ &= V^j \partial_i f \delta^i_j \\ &= V^i \partial_i f \end{aligned}$$

So the one-form *df* eats a vector V and produces a number that is the directional derivative of f in the direction of V. Just like the gradient! So d can take a 0-form and make a 1-form. It can do more...

Let $\bigwedge^k(T^*M)$ denote the space of all $k-form$ fields on M. Then the **exterior derivative** takes

$$d : \bigwedge^{k}(T^*M) \to \bigwedge^{k+1}(T^*M)$$

In other words, $d\omega_{(k)} = (d\omega)_{(k+1)}$: d can take a k-form and produce a $(k+1)$-form – a differential form with one more index. How? Easy!

$$d\omega = \partial_i \omega_{j_1 j_2 \cdots j_k} dx^i \wedge dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_k}$$

Important: $d^2 = 0$. What do I mean? I just mean that I can apply d twice. Starting with a k-form ω, I can get a (k+1)-form $d\omega$, and I could try to take d of this to get yet another form $d(d\omega) = d^2\omega$. This form will always be zero (i.e. all of its coefficients are zero in any basis!). Why?

$$\begin{aligned} d^2\omega &= d(\partial_i \omega_I dx^i \wedge dx^I) \\ &= \partial_j \partial_i \omega_I dx^j \wedge dx^i \wedge dx^I \\ &= 0 \end{aligned}$$

The reason the second line vanishes is that $\partial_j \partial_i = \partial_i \partial_j$ is symmetric under exchanging i and j, wheras $dx^j \wedge dx^i \wedge dx^I$ is antisymmetric under this exchange. This means that when I sum over i and j, every term in the sum will appear twice but with opposite signs. Zero!

Integrals It turns out that a $k-form$ is the only thing we can integrate over a k-dimensional submanifold! That's it! Now are you beginning to think they're important? I thought so. So let's take an n-form, which has the same degree as the manifold M and integrate it over M. First I'll tell you how this integral is defined. Then I'll tell you why this definition makes sense. The integral is defined to be:

$$\int_M \omega = \int dx^1 dx^2 \cdots dx^d \omega_{1 \cdots d}$$

This happens to make sense because it is invariant under coordinate changes. This fact, in turn, relies on the transformation properties of n-forms.

Fact: Under a coordinate change, the measure $d^d x$ transforms with the absolute value of the Jacobian determinant:

$$d^d x \to d^d x' = \left| \det(\frac{\partial x^{j'}}{\partial x^i}) \right| d^d x$$

Under a coordinate change, $\omega_{1\ldots d}$ transforms with the inverse of the Jacobian determinant:

$$\begin{aligned}
\omega'_{j_1 j_2 \ldots j_n} &= \prod_{k=1}^{n} \left(\frac{\partial x^{i_k}}{\partial x^{j_k}} \right) \omega_{i_1 i_2 \ldots i_n} \\
\omega'_{12\ldots n} &= \prod_{k=1}^{n} \left(\frac{\partial x^{i_k}}{\partial x^{k}} \right) \omega_{i_1 i_2 \ldots i_n} \\
&= \prod_{k=1}^{n} \left(\frac{\partial x^{\sigma(k)}}{\partial x^{k}} \right) (-1)^{\sigma} \omega_{12\ldots n} \\
\omega'_{12\ldots n} &= \det \left(\frac{\partial x^{i}}{\partial x^{j'}} \right) \omega_{12\ldots n}
\end{aligned}$$

In the second to last line, we used my favorite formula for the determinant. Now we may compare the two transformation rules to obtain:

$$\begin{aligned}
\int d^n x \omega_{123\ldots n} \to \int d^n x' \omega'_{123\ldots n} \\
&= \int d^n x \left| \det(\frac{\partial x^{j'}}{\partial x^i}) \right| \det \left(\frac{\partial x^{i}}{\partial x^{j'}} \right) \omega_{12\ldots n} \\
&= \pm \int d^n x \omega_{123\ldots n}
\end{aligned}$$

The sign is $+1$ if the coordinate transformation has positive definite Jacobian, i.e. if the transformation is orientation-preserving; -1 otherwise. We will use these properties to calculate (signed) areas/volumes/&c. later on.

Stokes' Theorem: Get excited. Stokes' theorem is the most general statement of the fundamental theorem of calculus.[1] I hope you're excited, because it's already here. All you need to know is that ∂M denotes the boundary of M (see figures) and is generically an $n-1$ dimensional manifold. With this definition, here's the theorem:

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

[1]Yes, there are other versions, but pretty much everything else requires more structure, so in this sense I claim that this is the most general. Of course you can also try to extend to infinite-dimensional manifolds &c.

On the left side, ω is an $(n-1)$-form integrated over the $(n-1)$-dimensional manifold ∂M. On the right-hand side, we have the n-form $d\omega$ integrated over the n-dimensional manifold M. So both integrals make sense. Incredibly, they're equal.

- **Comment on orientations:** Stoke's theorem only works if M and ∂M have compatible orientations. What this means is that we choose an ordered basis $\{e_i\}$ for the tangent space T_pM, and choose the orientation on ∂M such that when e_1 points outwards, the remaining $\{e_i\}_{i=2}^n$ form an ordered basis for $T_p\partial M$. This defines an orientation.

- **Observation 1:** If $M = [0,1]$ is the 1-dimensional manifold (the unit interval), then $\partial M = \{0\} \cup \{1\}$, i.e. the 0-dimensional manifold that is the union of the points 0 and 1. We can integrate a 0-form (just a function, f) on the 0-dimensional manifold. 0 has orientation -1 and 1 has orientation 1. Therefore

$$\begin{aligned}\int_{\partial M} f = f(1) - f(0) &= \int_M df \\ &= \int_0^1 dt \partial_t f(t)\end{aligned}$$

  So this really does contain the fundamental theorem of calculus!

- Proving Stokes' theorem is acutally fairly boring; all the work is already done in the fundamental theorem. The proof is quite similar to the proof of the divergence theorem or the easy Stokes' theorem, which are both special cases of this general theorem. We will see this below.

5.2 With metrics – some formulas

- The metric g_{ij} is a tensor that takes 2 vector arguments; therefore it is a tensor with two *lower* indices. It also satisfies the linearity property of a tensor: $g(aX_1 + bX_2, Y) = ag(X_1, Y) + bg(X_2, Y)$ (and similarly for Y). It also has the extra property of being symmetric.

- We can turn a vector into a one-form with the metric. How? Easy: given a vector X, the object $g(X, \cdot)$ is an object that accepts one vector (in its unoccupied second slot) and return a real number. This is the definition of a one-form! In coordinates, of course, this is just $V_i = g_{ij}V^j$. Abstractly this is called $\omega = \flat V$.

- We can raise indices with the inverse metric: $\omega^i = g^{ij}\omega_j$. It is one line to show that the inverse metric transforms oppositely from the metric. Therefore the inverse metric may be thought of as an object with two slots each accepting a one-form, that returns a real number. Abstractly this is called $V = \sharp\omega$.

- Hodge star: we can also perform an operation that uniquely relates k-forms to (n-k)-forms (on an n-dimensional manifold). Note that this is possible in principle because $\binom{n}{k} = \binom{n}{n-k}$. The operation on a form ω is denoted $\star\omega$.

– With a metric, we can choose an (ordered) orthonormal basis for one-forms $\{de_1, ..., de_k\}$. Then

$$\star(de_1 \wedge de_2 \wedge \cdots \wedge de_k) := de_{k+1} \wedge \cdots \wedge de_n$$

– One can see without too much extra work that this implies the following for unnormalized bases such as dx^i's.

$$(\star\omega)_{i_1 i_2 \ldots i_{n-k}} = \frac{\sqrt{g}}{(n-k)!} \omega^{j_1 j_2 \ldots j_k} \epsilon_{j_1 j_2 \ldots j_k i_1 i_2 \ldots i_{n-k}}$$

Let's use some of these properties to turn Stokes' theorem back into things we can recognize. First we need to figure out how to write curl and divergence.

- div V $= (\star d \star)(\flat V)$. Therefore
- curl V $= \star d(\flat V)$
- $\nabla^2 f = (\star d \star) df$
- Divergence theorem
- The **volume form** and integration of "functions." What I keep emphasizing is that you can only integrate n-forms on an n-dimensional manifold. But what about all those times we integrate functions over space, over spheres, etc. How can we do that? Well, we integrate against a so-called volume form $dV(x)$, a special form on m that is defined to denote a tiny piece of (signed) volume on M.

  In the absence of a metric, there is no preferred way to choose a volume form. There is only one rule: the volume form $dV = dV(p)$ should never vanish anywhere! This also means that if I (smoothly) choose an ordered basis $\{e_1, \ldots, e_n\}$ for T_pM for each $p \in M$, then $dV(e_1(p), e_2(p), \ldots, e_n(p))$ is either positive or negative for all p. (In other words, it cannot change sign.) This also implies that a choice of volume form implies a choice of orientation on a manifold. A basis for T_pM is oriented positively iff $dV_p(e_1(p), e_2(p), \ldots, e_n(p)) > 0$.

  In the presence of a metric, there's obviously a natural choice for dV. First take an ordered, orthonormal basis $\{e_1, \ldots e_n\}$ for each $p \in M$. I.e. the vector fields $e_i(p)$ satisfy $g_p(e_i(p), e_j(p)) = \delta_{ij}$. Now, given such a basis for T_pM, we can easily find a corresponding orthonormal basis for T_p^*M. Define $de_i := \flat e_i = g(e_i, \cdot)$ is a one-form, and this happens to be dual to e_i in the normal co-vector sense: $de_i(e_i) = 1$. So the dual basis is in this case coincides with the basis of flatted vectors. Either way, we now have a perfectly good orthonormal basis $\{de_i\}$ for T_p^*M, and we can simply define dV:

$$\begin{aligned} dV &:= de_1 \wedge de_2 \wedge \cdots \wedge de_n \\ &= \pm\sqrt{g} dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n \end{aligned}$$

What's with the second line? It's just a different way to write the same form, but which is written directly in terms of the coordinates and thereby saves us the trouble of having to find coordinate expressions for the de_i (which might be easy, but could also be very time-consuming.) How to see it's true? Just notice that over any point p, we can choose local coordinates $\{x^1, \cdots, x^n\}$ that diagonalize the metric at p. This allows us to write $g_p = \mathrm{diag}(g_{11}, g_{22}, \ldots, g_{nn})$. Then $\det(g) = g_{11}g_{22}\cdots g_{nn} =: \prod_{i=1}^n g_{ii}$. But now, what are the normalized one-forms in these diagonalizing coordinates? $de^i = \frac{1}{\sqrt{g^{ii}}}dx^i = \sqrt{g_{ii}}dx^i$. Thus

$$\begin{aligned} dV &= \sqrt{g_11}dx^1 \wedge \sqrt{g_{22}}dx^2 \wedge \cdots \wedge \sqrt{g_{nn}}dx^n \\ &= \sqrt{g_{11}g_{22}\cdots g_{nn}}dx^1 \wedge \cdots \wedge dx^n \\ &= \sqrt{\mathrm{g}}dx^1 \wedge \cdots \wedge dx^n \end{aligned}$$

By the way, an even slicker way to see this is just to realize that the Jacobian of a coordinate transformation that transforms an arbitrary basis into an orthonormal basis must be $\sqrt{\det(g)}$ and then apply what we know about transofmation properties of n-forms on n-diensional manifolds. (They transform with a power of the determinant of the Jacobian.)

For our purposes, then, we will very frequently be integrating functions over manifolds (and in GR, we always have a metric) so we will see, over and over again, the formula (definition, if you will):

$$\begin{aligned} \int_M f(x) &= \int_M dV(x) f(x) \\ &= \int d^n x \sqrt{g(x)} f(x) \end{aligned} \tag{5.1}$$

Exercises

- Do the integral $dx \wedge dy$ over the hemisphere $z \geq 0$ in whatever coordinates you like. Relate this to a vector field
- Prove the formula for
- Prove the following formulas

6 Fiber bundles and connections

So far we've seen a lot of formulas. Some even involve the metric. But we haven't yet seen made a *connection (–get it?)* to curvature and/or Christoffel symbols / covariant derivatives. Now we'll fix that.

6.1 Tangent bundle is a manifold

We can stitch all the tangent spaces together into a smooth manifold of dimension $2n$. In a chart U_α on M, we can put $2n$ coordinates $\phi_\alpha(x, V) = (x^1, \ldots, x^n, V^1(x)\partial_{x^1}|_x, \ldots, V^n(x)\partial_{x^n}|_x)$. It is an easy (but useful) exercise to derive the transformation rule $\phi_\beta \circ \phi_\alpha^{-1} : S_\alpha \to S_\beta$ describing how coordinates in the patch U_α are related to coordinates in the patch U_β. (This will simply reproduce are usual law for vector component transformation properties.

6.2 Bundles with any fiber

We have seen how, if we have a vector space of dimension n T_xM for each point $x \in M$, we can stitch these spaces smoothly together into a smooth shape of dimension $2n = \dim(M) + \dim(T_xM) = n+n$. But we can do a much more general procedure. Given any fixed topological shape (or, more simply, any vector space) of fixed dimension m, if we have one of these objects F_x for each $x \in M$, then it is possible to stitch these together to form a shape

$$FM = \coprod_x F_xM \tag{6.1}$$

where the dimension of FM is now $\dim(M) + \dim(F_x) = n + m$.

Here are a few examples in the figure. For instance, we can get the cylinder $S^1 \times \mathbb{R}^1$ by placing the same one-dimensional vector space $\mathbb{R}^1$ at each point over S^1. If we break the rule that the fiber F has to be the same topological space over every point x, then we can make 2-sphere S^2 by fibering $F_x = S^1_{R^2=1-x^2}$ over $x \in [-1, 1]$. (Try putting coordinates on this space; the simple ones will work everywhere except at -1 and 1)

To

6.3 Holonomy, parallel transport, covariant derivative: all just forms of a "connection"

6.4 "Groups" and holonomy

7 Gauge theories and general relativity; curvature

7.1 Physics motivation: E & M is a U(1) gauge theory

7.2 Gauge theories in general

7.3 Gauge fields are connection one-forms

7.4 Field strength tensors are curvature tensors

7.5 Preview: General relativity is a gauge theory

8 General Relativity as a gauge theory