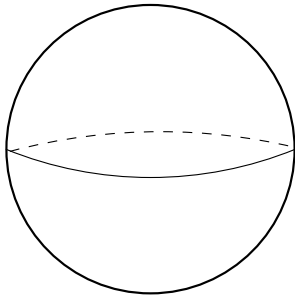# 2 Manifolds

After the invention of special relativity, Einstein tried for a number of years to invent a Lorentz-invariant theory of gravity, without success. His eventual breakthrough was to replace Minkowski spacetime with a curved spacetime, where the curvature was created by (and reacted back on) energy and momentum. Before we explore how this happens, we have to learn a bit about the mathematics of curved spaces. First we will take a look at manifolds in general, and then in the next section study curvature. In the interest of generality we will usually work in n dimensions, although you are permitted to take $n = 4$ if you like.

A manifold (or sometimes "differentiable manifold") is one of the most fundamental concepts in mathematics and physics. We are all aware of the properties of n-dimensional Euclidean space, $\mathbf{R}^n$, the set of n-tuples $(x^1, \ldots, x^n)$. The notion of a manifold captures the idea of a space which may be curved and have a complicated topology, but in local regions looks just like $\mathbf{R}^n$. (Here by "looks like" we do not mean that the metric is the same, but only basic notions of analysis like open sets, functions, and coordinates.) The entire manifold is constructed by smoothly sewing together these local regions. Examples of manifolds include:

- $\mathbf{R}^n$ itself, including the line ($\mathbf{R}$), the plane ($\mathbf{R}^2$), and so on. This should be obvious, since $\mathbf{R}^n$ looks like $\mathbf{R}^n$ not only locally but globally.

- The n-sphere, S^n. This can be defined as the locus of all points some fixed distance from the origin in $\mathbf{R}^{n+1}$. The circle is of course S^1, and the two-sphere S^2 will be one of our favorite examples of a manifold.

- The n-torus T^n results from taking an n-dimensional cube and identifying opposite sides. Thus T^2 is the traditional surface of a doughnut.
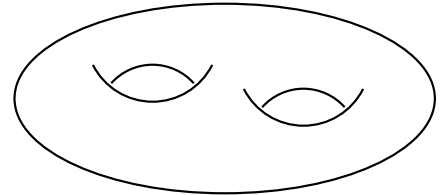

identify opposite sides
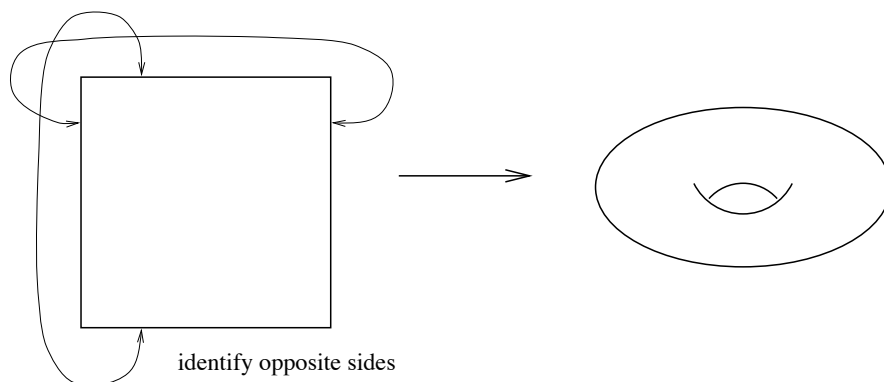
- A Riemann surface of genus g is essentially a two-torus with g holes instead of just one. S^2 may be thought of as a Riemann surface of genus zero. For those of you who know what the words mean, every "compact orientable boundaryless" two-dimensional manifold is a Riemann surface of some genus.
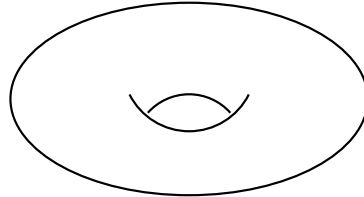

genus 0   genus 1   genus 2

- More abstractly, a set of continuous transformations such as rotations in $\mathbf{R}^n$ forms a manifold. Lie groups are manifolds which also have a group structure.

- The direct product of two manifolds is a manifold. That is, given manifolds M and M' of dimension n and n', we can construct a manifold $M \times M'$, of dimension $n + n'$, consisting of ordered pairs (p, p') for all $p \in M$ and $p' \in M'$.

With all of these examples, the notion of a manifold may seem vacuous; what isn't a manifold? There are plenty of things which are not manifolds, because somewhere they do not look locally like $\mathbf{R}^n$. Examples include a one-dimensional line running into a two-dimensional plane, and two cones stuck together at their vertices. (A single cone is okay; you can imagine smoothing out the vertex.)
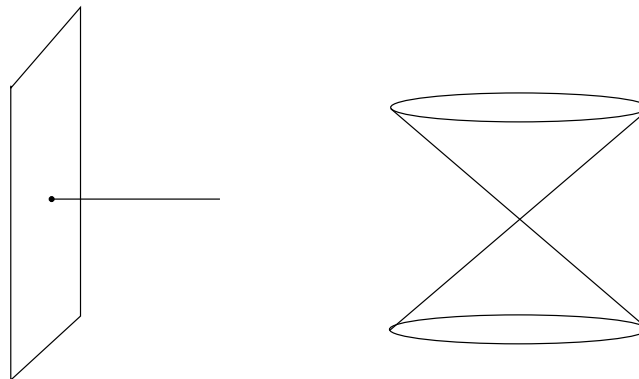

We will now approach the rigorous definition of this simple idea, which requires a number of preliminary definitions. Many of them are pretty clear anyway, but it's nice to be complete.

The most elementary notion is that of a **map** between two sets. (We assume you know what a set is.) Given two sets M and N, a map $\phi : M \to N$ is a relationship which assigns, to each element of M, exactly one element of N. A map is therefore just a simple generalization of a function. The canonical picture of a map looks like this:
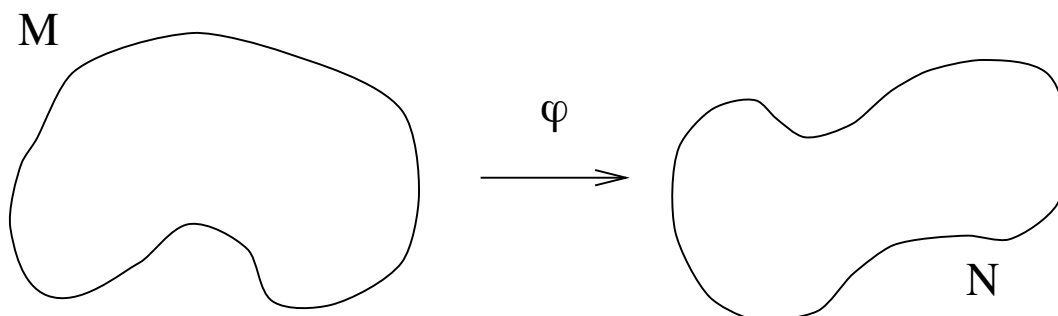

Given two maps $\phi : A \to B$ and $\psi : B \to C$, we define the **composition** $\psi \circ \phi : A \to C$ by the operation $(\psi \circ \phi)(a) = \psi(\phi(a))$. So $a \in A$, $\phi(a) \in B$, and thus $(\psi \circ \phi)(a) \in C$. The order in which the maps are written makes sense, since the one on the right acts first. In pictures:
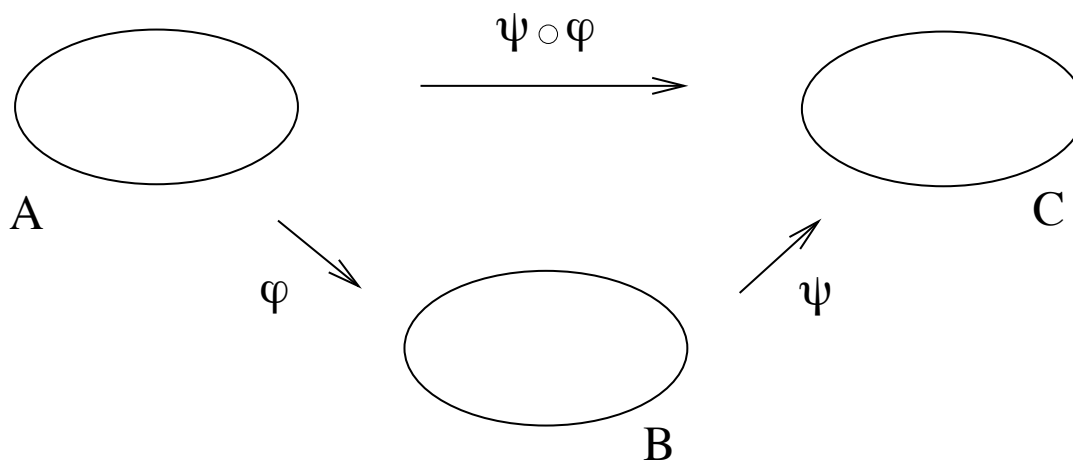
A map ϕ is called **one-to-one** (or "injective") if each element of N has at most one element of M mapped into it, and **onto** (or "surjective") if each element of N has at least one element of M mapped into it. (If you think about it, a better name for "one-to-one" would be "two-to-two".) Consider a function $\phi : \mathbf{R} \to \mathbf{R}$. Then $\phi(x) = e^x$ is one-to-one, but not onto; $\phi(x) = x^3 - x$ is onto, but not one-to-one; $\phi(x) = x^3$ is both; and $\phi(x) = x^2$ is neither.

The set M is known as the **domain** of the map ϕ, and the set of points in N which M gets mapped into is called the **image** of ϕ. For some subset $U \subset N$, the set of elements of M which get mapped to U is called the **preimage** of U under ϕ, or $\phi^{-1}(U)$. A map which is

e^x

x

one-to-one,
not onto

$x^3 - x$

x

onto, not
one-to-one

x^3

x

both

x^2

x

neither
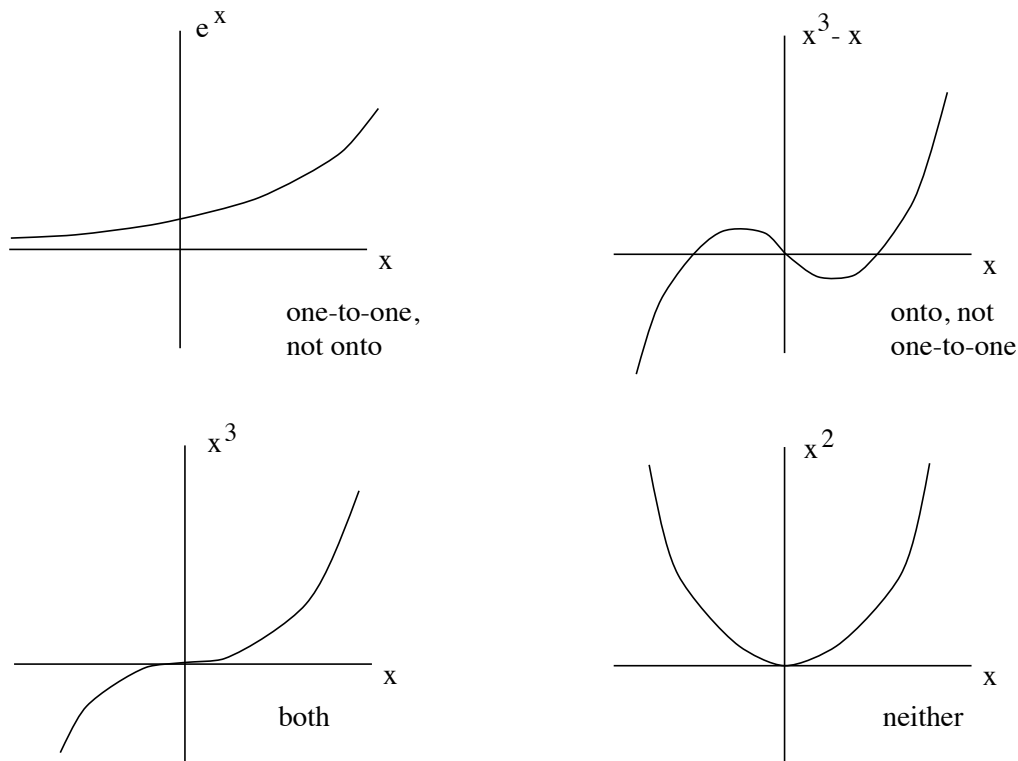
both one-to-one and onto is known as **invertible** (or "bijective"). In this case we can define the **inverse map** $\phi^{-1} : N \to M$ by $(\phi^{-1} \circ \phi)(a) = a$. (Note that the same symbol ϕ^{-1} is used for both the preimage and the inverse map, even though the former is always defined and the latter is only defined in some special cases.) Thus:

M

φ

N

φ^{-1}
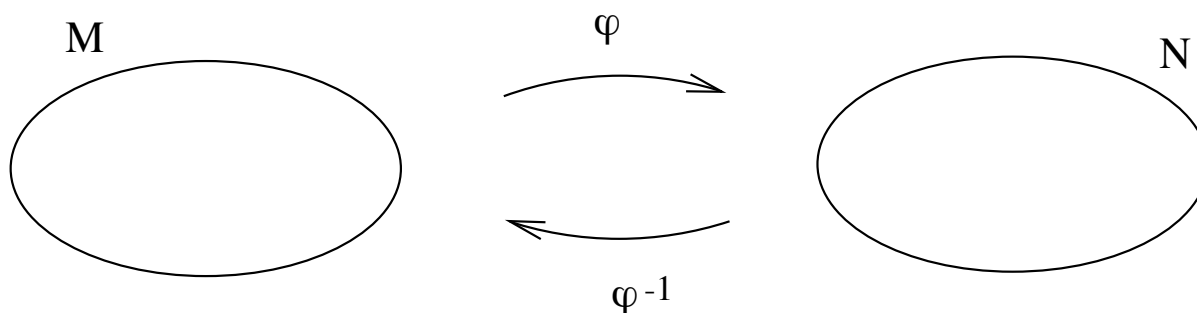
The notion of **continuity** of a map between topological spaces (and thus manifolds) is actually a very subtle one, the precise formulation of which we won't really need. However the intuitive notions of continuity and differentiability of maps $\phi : \mathbf{R}^m \to \mathbf{R}^n$ between Euclidean spaces are useful. A map from $\mathbf{R}^m$ to $\mathbf{R}^n$ takes an m-tuple $(x^1, x^2, \ldots, x^m)$ to an n-tuple $(y^1, y^2, \ldots, y^n)$, and can therefore be thought of as a collection of n functions ϕ^i of

m variables:

$$
\begin{aligned}
y^1 &= \phi^1(x^1, x^2, \ldots, x^m) \\
y^2 &= \phi^2(x^1, x^2, \ldots, x^m) \\
&\quad\vdots \\
y^n &= \phi^n(x^1, x^2, \ldots, x^m) \ .
\end{aligned}
\tag{2.1}
$$

We will refer to any one of these functions as C^p if it is continuous and p-times differentiable, and refer to the entire map $\phi : \mathbf{R}^m \to \mathbf{R}^n$ as C^p if each of its component functions are at least C^p. Thus a C^0 map is continuous but not necessarily differentiable, while a C^∞ map is continuous and can be differentiated as many times as you like. C^∞ maps are sometimes called **smooth**. We will call two sets M and N **diffeomorphic** if there exists a C^∞ map $\phi : M \to N$ with a C^∞ inverse $\phi^{-1} : N \to M$; the map ϕ is then called a diffeomorphism.

Aside: The notion of two spaces being diffeomorphic only applies to manifolds, where a notion of differentiability is inherited from the fact that the space resembles $\mathbf{R}^n$ locally. But "continuity" of maps between topological spaces (not necessarily manifolds) can be defined, and we say that two such spaces are "homeomorphic," which means "topologically equivalent to," if there is a continuous map between them with a continuous inverse. It is therefore conceivable that spaces exist which are homeomorphic but not diffeomorphic; topologically the same but with distinct "differentiable structures." In 1964 Milnor showed that S^7 had 28 different differentiable structures; it turns out that for $n < 7$ there is only one differentiable structure on S^n, while for $n > 7$ the number grows very large. $\mathbf{R}^4$ has infinitely many differentiable structures.

One piece of conventional calculus that we will need later is the **chain rule**. Let us imagine that we have maps $f : \mathbf{R}^m \to \mathbf{R}^n$ and $g : \mathbf{R}^n \to \mathbf{R}^l$, and therefore the composition $(g \circ f) : \mathbf{R}^m \to \mathbf{R}^l$.
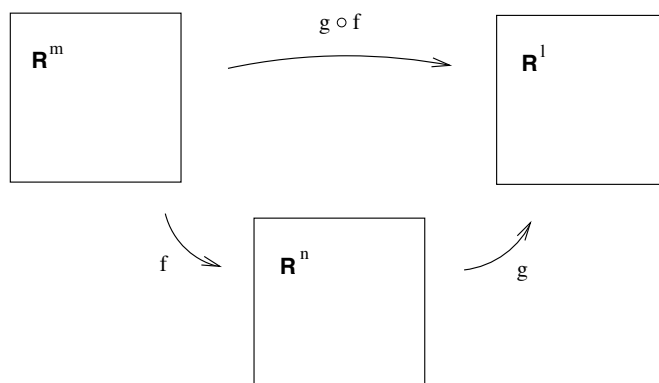
We can represent each space in terms of coordinates: x^a on $\mathbf{R}^m$, y^b on $\mathbf{R}^n$, and z^c on $\mathbf{R}^l$, where the indices range over the appropriate values. The chain rule relates the partial

derivatives of the composition to the partial derivatives of the individual maps:

$$\frac{\partial}{\partial x^a}(g \circ f)^c = \sum_b \frac{\partial f^b}{\partial x^a}\frac{\partial g^c}{\partial y^b} \ . \tag{2.2}$$

This is usually abbreviated to

$$\frac{\partial}{\partial x^a} = \sum_b \frac{\partial y^b}{\partial x^a}\frac{\partial}{\partial y^b} \ . \tag{2.3}$$

There is nothing illegal or immoral about using this form of the chain rule, but you should be able to visualize the maps that underlie the construction. Recall that when $m = n$ the determinant of the matrix $\partial y^b/\partial x^a$ is called the **Jacobian** of the map, and the map is invertible whenever the Jacobian is nonzero.

These basic definitions were presumably familiar to you, even if only vaguely remembered. We will now put them to use in the rigorous definition of a manifold. Unfortunately, a somewhat baroque procedure is required to formalize this relatively intuitive notion. We will first have to define the notion of an open set, on which we can put coordinate systems, and then sew the open sets together in an appropriate way.

Start with the notion of an **open ball**, which is the set of all points x in $\mathbf{R}^n$ such that $|x - y| < r$ for some fixed $y \in \mathbf{R}^n$ and $r \in \mathbf{R}$, where $|x - y| = [\sum_i(x^i - y^i)^2]^{1/2}$. Note that this is a strict inequality — the open ball is the interior of an n-sphere of radius r centered at y.
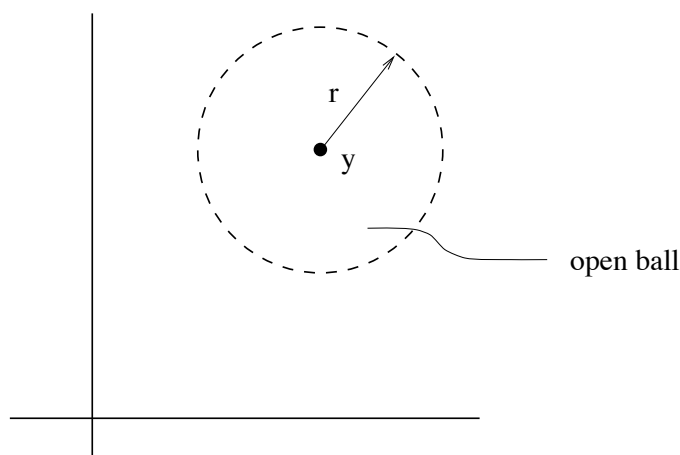
An **open set** in $\mathbf{R}^n$ is a set constructed from an arbitrary (maybe infinite) union of open balls. In other words, $V \subset \mathbf{R}^n$ is open if, for any $y \in V$, there is an open ball centered at y which is completely inside V. Roughly speaking, an open set is the interior of some $(n-1)$-dimensional closed surface (or the union of several such interiors). By defining a notion of open sets, we have equipped $\mathbf{R}^n$ with a topology — in this case, the "standard metric topology."

A **chart** or **coordinate system** consists of a subset U of a set M, along with a one-to-one map $\phi : U \to \mathbf{R}^n$, such that the image $\phi(U)$ is open in $\mathbf{R}$. (Any map is onto its image, so the map $\phi : U \to \phi(U)$ is invertible.) We then can say that U is an open set in M. (We have thus induced a topology on M, although we will not explore this.)
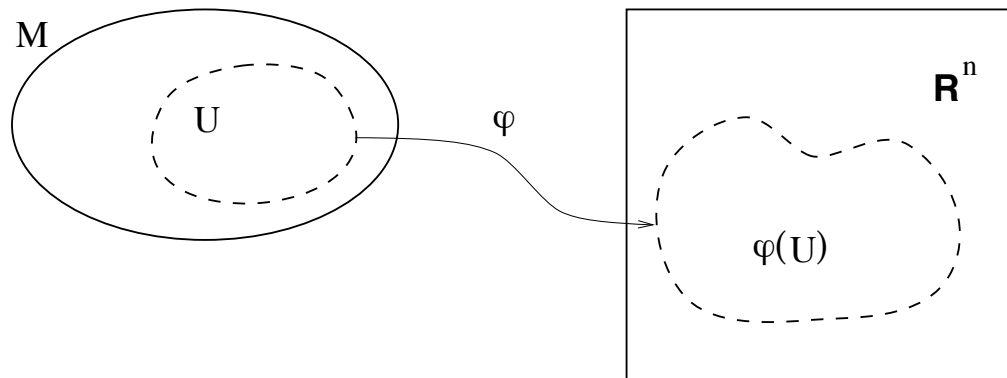
A C^∞ **atlas** is an indexed collection of charts $\{(U_\alpha, \phi_\alpha)\}$ which satisfies two conditions:

1. The union of the U_α is equal to M; that is, the U_α cover M.

2. The charts are smoothly sewn together. More precisely, if two charts overlap, $U_\alpha \cap U_\beta \neq \emptyset$, then the map $(\phi_\alpha \circ \phi_\beta^{-1})$ takes points in $\phi_\beta(U_\alpha \cap U_\beta) \subset \mathbf{R}^n$ *onto* $\phi_\alpha(U_\alpha \cap U_\beta) \subset \mathbf{R}^n$, and all of these maps must be C^∞ where they are defined. This should be clearer in pictures:

these maps are only defined on the shaded regions, and must be smooth there.
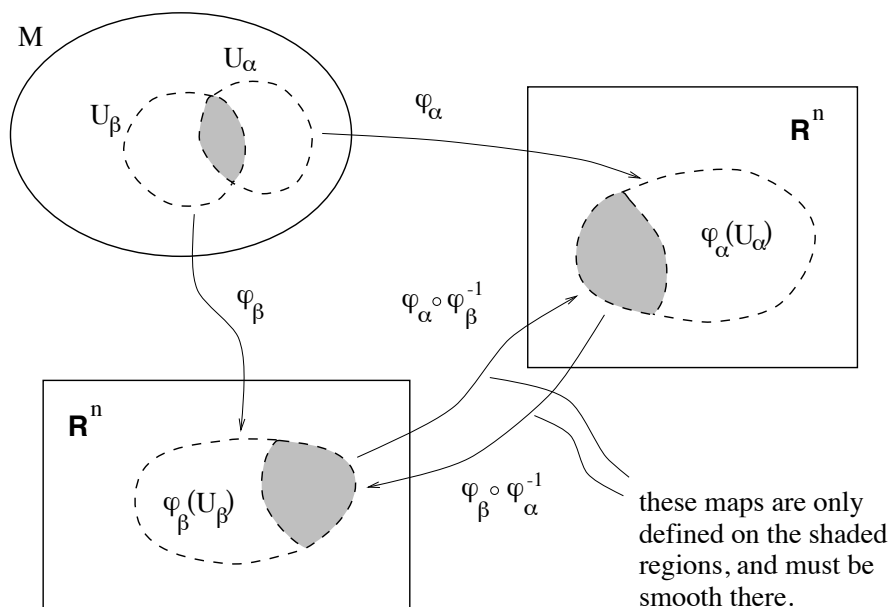
So a chart is what we normally think of as a coordinate system on some open set, and an atlas is a system of charts which are smoothly related on their overlaps.

At long last, then: a C^∞ n-dimensional **manifold** (or n-manifold for short) is simply a set M along with a "maximal atlas", one that contains every possible compatible chart. (We can also replace C^∞ by C^p in all the above definitions. For our purposes the degree of differentiability of a manifold is not crucial; we will always assume that any manifold is as differentiable as necessary for the application under consideration.) The requirement that the atlas be maximal is so that two equivalent spaces equipped with different atlases don't count as different manifolds. This definition captures in formal terms our notion of a set that looks locally like $\mathbf{R}^n$. Of course we will rarely have to make use of the full power of the definition, but precision is its own reward.

One thing that is nice about our definition is that it does not rely on an embedding of the manifold in some higher-dimensional Euclidean space. In fact any n-dimensional manifold can be embedded in $\mathbf{R}^{2n}$ ("Whitney's embedding theorem"), and sometimes we will make use of this fact (such as in our definition of the sphere above). But it's important to recognize that the manifold has an individual existence independent of any embedding. We have no reason to believe, for example, that four-dimensional spacetime is stuck in some larger space. (Actually a number of people, string theorists and so forth, believe that our four-dimensional world is part of a ten- or eleven-dimensional spacetime, but as far as GR is concerned the 4-dimensional view is perfectly adequate.)

Why was it necessary to be so finicky about charts and their overlaps, rather than just covering every manifold with a single chart? Because most manifolds cannot be covered with just one chart. Consider the simplest example, S^1. There is a conventional coordinate system, $\theta : S^1 \to \mathbf{R}$, where $\theta = 0$ at the top of the circle and wraps around to 2π. However, in the definition of a chart we have required that the image $\theta(S^1)$ be open in $\mathbf{R}$. If we include either $\theta = 0$ or $\theta = 2\pi$, we have a closed interval rather than an open one; if we exclude both points, we haven't covered the whole circle. So we need at least two charts, as shown.
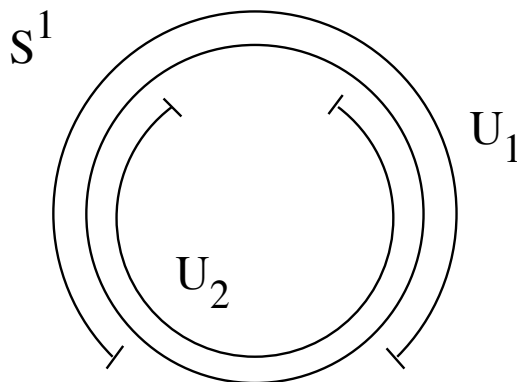
A somewhat more complicated example is provided by S^2, where once again no single

chart will cover the manifold. A Mercator projection, traditionally used for world maps, misses both the North and South poles (as well as the International Date Line, which involves the same problem with θ that we found for S^1.) Let's take S^2 to be the set of points in $\mathbf{R}^3$ defined by $(x^1)^2 + (x^2)^2 + (x^3)^2 = 1$. We can construct a chart from an open set U_1, defined to be the sphere minus the north pole, via "stereographic projection":
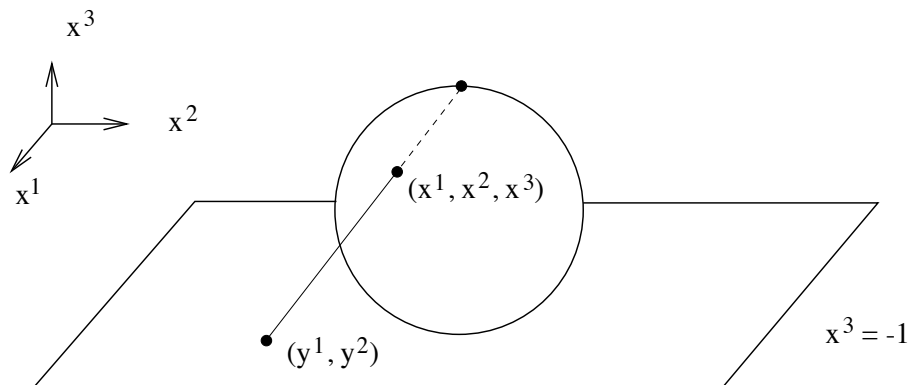
Thus, we draw a straight line from the north pole to the plane defined by $x^3 = -1$, and assign to the point on S^2 intercepted by the line the Cartesian coordinates (y^1, y^2) of the appropriate point on the plane. Explicitly, the map is given by

$$\phi_1(x^1, x^2, x^3) \equiv (y^1, y^2) = \left(\frac{2x^1}{1 - x^3} \ , \ \frac{2x^2}{1 - x^3} \right) \ . \tag{2.4}$$

You are encouraged to check this for yourself. Another chart (U_2, ϕ_2) is obtained by projecting from the south pole to the plane defined by $x^3 = +1$. The resulting coordinates cover the sphere minus the south pole, and are given by

$$\phi_2(x^1, x^2, x^3) \equiv (z^1, z^2) = \left(\frac{2x^1}{1 + x^3} \ , \ \frac{2x^2}{1 + x^3} \right) \ . \tag{2.5}$$

Together, these two charts cover the entire manifold, and they overlap in the region $-1 < x^3 < +1$. Another thing you can check is that the composition $\phi_2 \circ \phi_1^{-1}$ is given by

$$z^i = \frac{4y^i}{[(y^1)^2 + (y^2)^2]} \ , \tag{2.6}$$

and is C^∞ in the region of overlap. As long as we restrict our attention to this region, (2.6) is just what we normally think of as a change of coordinates.

We therefore see the necessity of charts and atlases: many manifolds cannot be covered with a single coordinate system. (Although some can, even ones with nontrivial topology. Can you think of a single good coordinate system that covers the cylinder, $S^1 \times \mathbf{R}$?) Nevertheless, it is very often most convenient to work with a single chart, and just keep track of the set of points which aren't included.

The fact that manifolds look locally like $\mathbf{R}^n$, which is manifested by the construction of coordinate charts, introduces the possibility of analysis on manifolds, including operations such as differentiation and integration. Consider two manifolds M and N of dimensions m and n, with coordinate charts ϕ on M and ψ on N. Imagine we have a function $f : M \to N$,
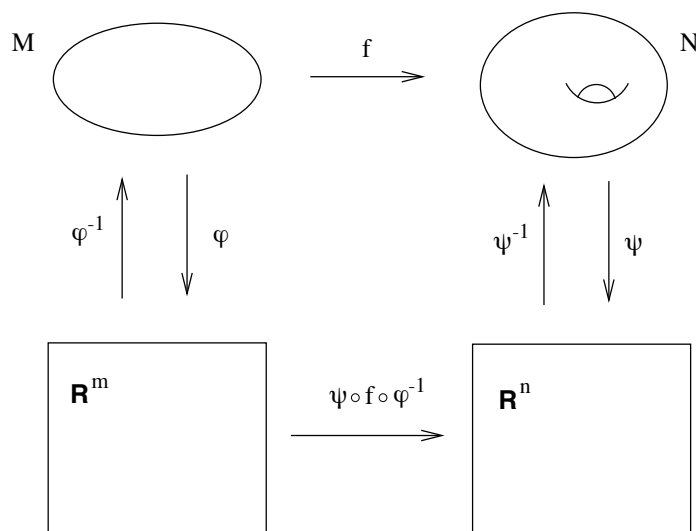
Just thinking of M and N as sets, we cannot nonchalantly differentiate the map f, since we don't know what such an operation means. But the coordinate charts allow us to construct the map $(\psi \circ f \circ \phi^{-1}) : \mathbf{R}^m \to \mathbf{R}^n$. (Feel free to insert the words "where the maps are defined" wherever appropriate, here and later on.) This is just a map between Euclidean spaces, and all of the concepts of advanced calculus apply. For example f, thought of as an N-valued function on M, can be differentiated to obtain $\partial f / \partial x^\mu$, where the x^μ represent $\mathbf{R}^m$. The point is that this notation is a shortcut, and what is really going on is

$$\frac{\partial f}{\partial x^\mu} \equiv \frac{\partial}{\partial x^\mu}(\psi \circ f \circ \phi^{-1})(x^\mu) \; . \tag{2.7}$$

It would be far too unwieldy (not to mention pedantic) to write out the coordinate maps explicitly in every case. The shorthand notation of the left-hand-side will be sufficient for most purposes.

Having constructed this groundwork, we can now proceed to introduce various kinds of structure on manifolds. We begin with vectors and tangent spaces. In our discussion of special relativity we were intentionally vague about the definition of vectors and their relationship to the spacetime. One point that was stressed was the notion of a tangent space — the set of all vectors at a single point in spacetime. The reason for this emphasis was to remove from your minds the idea that a vector stretches from one point on the manifold to another, but instead is just an object associated with a single point. What is temporarily lost by adopting this view is a way to make sense of statements like "the vector points in

the x direction" — if the tangent space is merely an abstract vector space associated with each point, it's hard to know what this should mean. Now it's time to fix the problem.

Let's imagine that we wanted to construct the tangent space at a point p in a manifold M, using only things that are intrinsic to M (no embeddings in higher-dimensional spaces etc.). One first guess might be to use our intuitive knowledge that there are objects called "tangent vectors to curves" which belong in the tangent space. We might therefore consider the set of all parameterized curves through p — that is, the space of all (nondegenerate) maps $\gamma : \mathbf{R} \rightarrow M$ such that p is in the image of γ. The temptation is to define the tangent space as simply the space of all tangent vectors to these curves at the point p. But this is obviously cheating; the tangent space T_p is supposed to be the space of vectors at p, and before we have defined this we don't have an independent notion of what "the tangent vector to a curve" is supposed to mean. In some coordinate system x^μ any curve through p defines an element of $\mathbf{R}^n$ specified by the n real numbers $dx^\mu/d\lambda$ (where λ is the parameter along the curve), but this map is clearly coordinate-dependent, which is not what we want.

Nevertheless we are on the right track, we just have to make things independent of coordinates. To this end we define $\mathcal{F}$ to be the space of all smooth functions on M (that is, C^∞ maps $f : M \rightarrow \mathbf{R}$). Then we notice that each curve through p defines an operator on this space, the directional derivative, which maps $f \rightarrow df/d\lambda$ (at p). We will make the following claim: *the tangent space T_p can be identified with the space of directional derivative operators along curves through p.* To establish this idea we must demonstrate two things: first, that the space of directional derivatives is a vector space, and second that it is the vector space we want (it has the same dimensionality as M, yields a natural idea of a vector pointing along a certain direction, and so on).

The first claim, that directional derivatives form a vector space, seems straightforward enough. Imagine two operators $\frac{d}{d\lambda}$ and $\frac{d}{d\eta}$ representing derivatives along two curves through p. There is no problem adding these and scaling by real numbers, to obtain a new operator $a\frac{d}{d\lambda} + b\frac{d}{d\eta}$. It is not immediately obvious, however, that the space closes; *i.e.*, that the resulting operator is itself a derivative operator. A good derivative operator is one that acts linearly on functions, and obeys the conventional Leibniz (product) rule on products of functions. Our new operator is manifestly linear, so we need to verify that it obeys the Leibniz rule. We have

$$
\begin{aligned}
\left(a\frac{d}{d\lambda} + b\frac{d}{d\eta} \right)(fg) &= af\frac{dg}{d\lambda} + ag\frac{df}{d\lambda} + bf\frac{dg}{d\eta} + bg\frac{df}{d\eta} \\
&= \left(a\frac{df}{d\lambda} + b\frac{df}{d\eta} \right)g + \left(a\frac{dg}{d\lambda} + b\frac{dg}{d\eta} \right)f \; . \quad (2.8)
\end{aligned}
$$

As we had hoped, the product rule is satisfied, and the set of directional derivatives is therefore a vector space.

Is it the vector space that we would like to identify with the tangent space? The easiest way to become convinced is to find a basis for the space. Consider again a coordinate chart with coordinates x^μ. Then there is an obvious set of n directional derivatives at p, namely the partial derivatives ∂_μ at p.
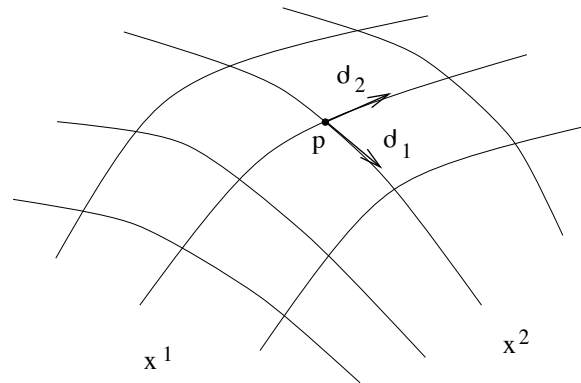
We are now going to claim that the partial derivative operators $\{\partial_\mu\}$ at p form a basis for the tangent space T_p. (It follows immediately that T_p is n-dimensional, since that is the number of basis vectors.) To see this we will show that any directional derivative can be decomposed into a sum of real numbers times partial derivatives. This is in fact just the familiar expression for the components of a tangent vector, but it's nice to see it from the big-machinery approach. Consider an n-manifold M, a coordinate chart $\phi : M \rightarrow \mathbf{R}^n$, a curve $\gamma : \mathbf{R} \rightarrow M$, and a function $f : M \rightarrow \mathbf{R}$. This leads to the following tangle of maps:
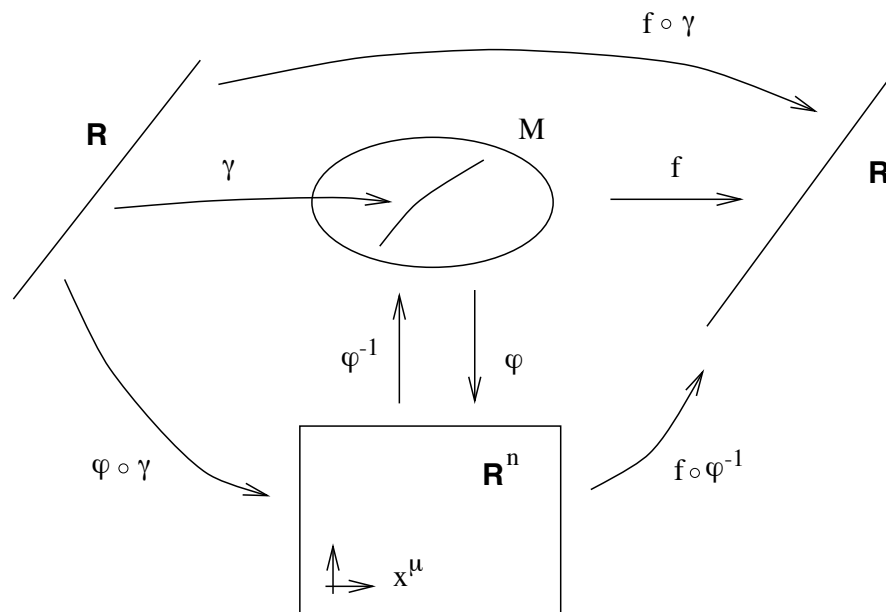
If λ is the parameter along γ, we want to expand the vector/operator $\frac{d}{d\lambda}$ in terms of the

partials ∂_μ. Using the chain rule (2.2), we have

$$
\begin{aligned}
\frac{d}{d\lambda}f &= \frac{d}{d\lambda}(f \circ \gamma) \\
&= \frac{d}{d\lambda}[(f \circ \phi^{-1}) \circ (\phi \circ \gamma)] \\
&= \frac{d(\phi \circ \gamma)^\mu}{d\lambda} \frac{\partial(f \circ \phi^{-1})}{\partial x^\mu} \\
&= \frac{dx^\mu}{d\lambda}\partial_\mu f \ .
\end{aligned}
\tag{2.9}
$$

The first line simply takes the informal expression on the left hand side and rewrites it as an honest derivative of the function $(f \circ \gamma) : \mathbf{R} \to \mathbf{R}$. The second line just comes from the definition of the inverse map ϕ^{-1} (and associativity of the operation of composition). The third line is the formal chain rule (2.2), and the last line is a return to the informal notation of the start. Since the function f was arbitrary, we have

$$
\frac{d}{d\lambda} = \frac{dx^\mu}{d\lambda}\partial_\mu \ .
\tag{2.10}
$$

Thus, the partials $\{\partial_\mu\}$ do indeed represent a good basis for the vector space of directional derivatives, which we can therefore safely identify with the tangent space.

Of course, the vector represented by $\frac{d}{d\lambda}$ is one we already know; it's the tangent vector to the curve with parameter λ. Thus (2.10) can be thought of as a restatement of (1.24), where we claimed the that components of the tangent vector were simply $dx^\mu/d\lambda$. The only difference is that we are working on an arbitrary manifold, and we have specified our basis vectors to be $\hat{e}_{(\mu)} = \partial_\mu$.

This particular basis $(\hat{e}_{(\mu)} = \partial_\mu)$ is known as a **coordinate basis** for T_p; it is the formalization of the notion of setting up the basis vectors to point along the coordinate axes. There is no reason why we are limited to coordinate bases when we consider tangent vectors; it is sometimes more convenient, for example, to use orthonormal bases of some sort. However, the coordinate basis is very simple and natural, and we will use it almost exclusively throughout the course.

One of the advantages of the rather abstract point of view we have taken toward vectors is that the transformation law is immediate. Since the basis vectors are $\hat{e}_{(\mu)} = \partial_\mu$, the basis vectors in some new coordinate system $x^{\mu'}$ are given by the chain rule (2.3) as

$$
\partial_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\partial_\mu \ .
\tag{2.11}
$$

We can get the transformation law for vector components by the same technique used in flat space, demanding the the vector $V = V^\mu \partial_\mu$ be unchanged by a change of basis. We have

$$
\begin{aligned}
V^\mu \partial_\mu &= V^{\mu'}\partial_{\mu'} \\
&= V^{\mu'}\frac{\partial x^\mu}{\partial x^{\mu'}}\partial_\mu \ ,
\end{aligned}
\tag{2.12}
$$

and hence (since the matrix $\partial x^{\mu'}/\partial x^{\mu}$ is the inverse of the matrix $\partial x^{\mu}/\partial x^{\mu'}$),

$$V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^{\mu}} V^{\mu} \; . \tag{2.13}$$

Since the basis vectors are usually not written explicitly, the rule (2.13) for transforming components is what we call the "vector transformation law." We notice that it is compatible with the transformation of vector components in special relativity under Lorentz transformations, $V^{\mu'} = \Lambda^{\mu'}{}_{\mu} V^{\mu}$, since a Lorentz transformation is a special kind of coordinate transformation, with $x^{\mu'} = \Lambda^{\mu'}{}_{\mu} x^{\mu}$. But (2.13) is much more general, as it encompasses the behavior of vectors under arbitrary changes of coordinates (and therefore bases), not just linear transformations. As usual, we are trying to emphasize a somewhat subtle ontological distinction — tensor components do not change when we change coordinates, they change when we change the basis in the tangent space, but we have decided to use the coordinates to define our basis. Therefore a change of coordinates induces a change of basis:
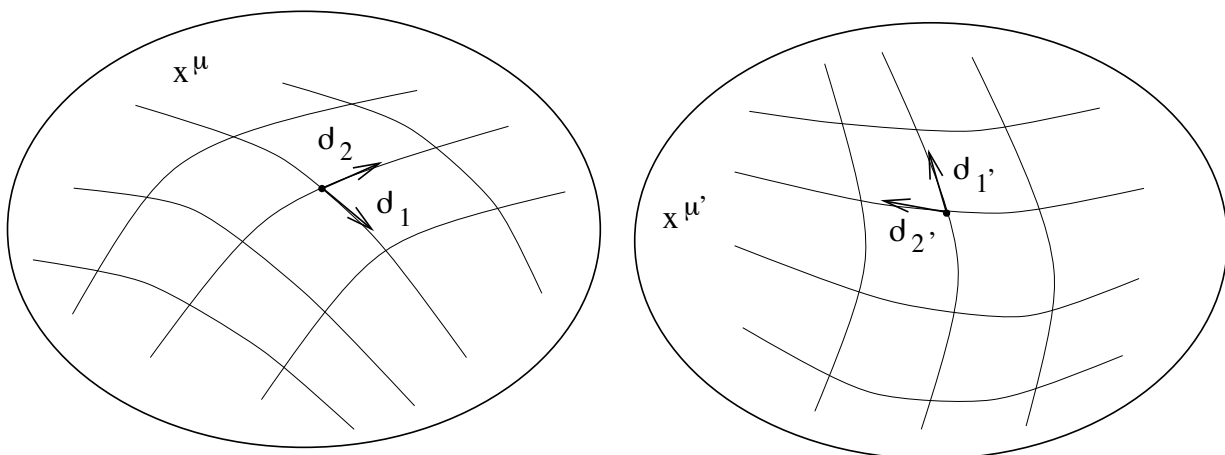
Having explored the world of vectors, we continue to retrace the steps we took in flat space, and now consider dual vectors (one-forms). Once again the cotangent space T_p^* is the set of linear maps $\omega : T_p \to \mathbf{R}$. The canonical example of a one-form is the gradient of a function f, denoted $\mathrm{d}f$. Its action on a vector $\frac{d}{d\lambda}$ is exactly the directional derivative of the function:

$$\mathrm{d}f\left(\frac{d}{d\lambda}\right) = \frac{df}{d\lambda} \; . \tag{2.14}$$

It's tempting to think, "why shouldn't the function f itself be considered the one-form, and $df/d\lambda$ its action?" The point is that a one-form, like a vector, exists only at the point it is defined, and does not depend on information at other points on M. If you know a function in some neighborhood of a point you can take its derivative, but not just from knowing its value at the point; the gradient, on the other hand, encodes precisely the information

necessary to take the directional derivative along any curve through p, fulfilling its role as a dual vector.

Just as the partial derivatives along coordinate axes provide a natural basis for the tangent space, the gradients of the coordinate functions x^μ provide a natural basis for the cotangent space. Recall that in flat space we constructed a basis for T_p^* by demanding that $\hat{\theta}^{(\mu)}(\hat{e}_{(\nu)}) = \delta_\nu^\mu$. Continuing the same philosophy on an arbitrary manifold, we find that (2.14) leads to

$$\mathrm{d}x^\mu(\partial_\nu) = \frac{\partial x^\mu}{\partial x^\nu} = \delta_\nu^\mu \ . \tag{2.15}$$

Therefore the gradients $\{\mathrm{d}x^\mu\}$ are an appropriate set of basis one-forms; an arbitrary one-form is expanded into components as $\omega = \omega_\mu\, \mathrm{d}x^\mu$.

The transformation properties of basis dual vectors and components follow from what is by now the usual procedure. We obtain, for basis one-forms,

$$\mathrm{d}x^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\, \mathrm{d}x^\mu \ , \tag{2.16}$$

and for components,

$$\omega_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\omega_\mu \ . \tag{2.17}$$

We will usually write the components ω_μ when we speak about a one-form ω.

The transformation law for general tensors follows this same pattern of replacing the Lorentz transformation matrix used in flat space with a matrix representing more general coordinate transformations. A (k, l) tensor T can be expanded

$$T = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l}\partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_k} \otimes \mathrm{d}x^{\nu_1} \otimes \cdots \otimes \mathrm{d}x^{\nu_l} \ , \tag{2.18}$$

and under a coordinate transformation the components change according to

$$T^{\mu_1' \cdots \mu_k'}{}_{\nu_1' \cdots \nu_l'} = \frac{\partial x^{\mu_1'}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\mu_k'}}{\partial x^{\mu_k}}\frac{\partial x^{\nu_1}}{\partial x^{\nu_1'}} \cdots \frac{\partial x^{\nu_l}}{\partial x^{\nu_l'}}T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} \ . \tag{2.19}$$

This tensor transformation law is straightforward to remember, since there really isn't anything else it could be, given the placement of indices. However, it is often easier to transform a tensor by taking the identity of basis vectors and one-forms as partial derivatives and gradients at face value, and simply substituting in the coordinate transformation. As an example consider a symmetric $(0, 2)$ tensor S on a 2-dimensional manifold, whose components in a coordinate system $(x^1 = x, x^2 = y)$ are given by

$$S_{\mu\nu} = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \ . \tag{2.20}$$

This can be written equivalently as

$$
\begin{aligned}
S &= S_{\mu\nu}(\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu) \\
&= x(\mathrm{d}x)^2 + (\mathrm{d}y)^2 \ ,
\end{aligned}
\tag{2.21}
$$

where in the last line the tensor product symbols are suppressed for brevity. Now consider new coordinates

$$
\begin{aligned}
x' &= x^{1/3} \\
y' &= e^{x+y} \ .
\end{aligned}
\tag{2.22}
$$

This leads directly to

$$
\begin{aligned}
x &= (x')^3 \\
y &= \ln(y') - (x')^3 \\
\mathrm{d}x &= 3(x')^2 \, \mathrm{d}x' \\
\mathrm{d}y &= \frac{1}{y'} \, \mathrm{d}y' - 3(x')^2 \, \mathrm{d}x' \ .
\end{aligned}
\tag{2.23}
$$

We need only plug these expressions directly into (2.21) to obtain (remembering that tensor products don't commute, so $\mathrm{d}x' \, \mathrm{d}y' \neq \mathrm{d}y' \, \mathrm{d}x'$):

$$
S = 9(x')^4[1 + (x')^3](\mathrm{d}x')^2 - 3\frac{(x')^2}{y'}(\mathrm{d}x' \, \mathrm{d}y' + \mathrm{d}y' \, \mathrm{d}x') + \frac{1}{(y')^2}(\mathrm{d}y')^2 \ ,
\tag{2.24}
$$

or

$$
S_{\mu'\nu'} = \begin{pmatrix} 9(x')^4[1 + (x')^3] & -3\frac{(x')^2}{y'} \\ -3\frac{(x')^2}{y'} & \frac{1}{(y')^2} \end{pmatrix} \ .
\tag{2.25}
$$

Notice that it is still symmetric. We did not use the transformation law (2.19) directly, but doing so would have yielded the same result, as you can check.

For the most part the various tensor operations we defined in flat space are unaltered in a more general setting: contraction, symmetrization, etc. There are three important exceptions: partial derivatives, the metric, and the Levi-Civita tensor. Let's look at the partial derivative first.

The unfortunate fact is that the partial derivative of a tensor is not, in general, a new tensor. The gradient, which is the partial derivative of a scalar, is an honest $(0, 1)$ tensor, as we have seen. But the partial derivative of higher-rank tensors is not tensorial, as we can see by considering the partial derivative of a one-form, $\partial_\mu W_\nu$, and changing to a new coordinate system:

$$
\begin{aligned}
\frac{\partial}{\partial x^{\mu'}}W_{\nu'} &= \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial}{\partial x^\mu}\left(\frac{\partial x^\nu}{\partial x^{\nu'}}W_\nu\right) \\
&= \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\nu}{\partial x^{\nu'}}\left(\frac{\partial}{\partial x^\mu}W_\nu\right) + W_\nu\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^{\nu'}} \ .
\end{aligned}
\tag{2.26}
$$

The second term in the last line should not be there if $\partial_\mu W_\nu$ were to transform as a $(0, 2)$ tensor. As you can see, it arises because the derivative of the transformation matrix does not vanish, as it did for Lorentz transformations in flat space.

On the other hand, the exterior derivative operator d does form an antisymmetric $(0, p+1)$ tensor when acted on a p-form. For $p = 1$ we can see this from (2.26); the offending non-tensorial term can be written

$$W_\nu \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\mu} \frac{\partial x^\nu}{\partial x^{\nu'}} = W_\nu \frac{\partial^2 x^\nu}{\partial x^{\mu'} \partial x^{\nu'}} \ . \tag{2.27}$$

This expression is symmetric in μ' and ν', since partial derivatives commute. But the exterior derivative is defined to be the antisymmetrized partial derivative, so this term vanishes (the antisymmetric part of a symmetric expression is zero). We are then left with the correct tensor transformation law; extension to arbitrary p is straightforward. So the exterior derivative is a legitimate tensor operator; it is not, however, an adequate substitute for the partial derivative, since it is only defined on forms. In the next section we will define a covariant derivative, which can be thought of as the extension of the partial derivative to arbitrary manifolds.

The metric tensor is such an important object in curved space that it is given a new symbol, $g_{\mu\nu}$ (while $\eta_{\mu\nu}$ is reserved specifically for the Minkowski metric). There are few restrictions on the components of $g_{\mu\nu}$, other than that it be a symmetric $(0, 2)$ tensor. It is usually taken to be non-degenerate, meaning that the determinant $g = |g_{\mu\nu}|$ doesn't vanish. This allows us to define the inverse metric $g^{\mu\nu}$ via

$$g^{\mu\nu} g_{\nu\sigma} = \delta^\mu_\sigma \ . \tag{2.28}$$

The symmetry of $g_{\mu\nu}$ implies that $g^{\mu\nu}$ is also symmetric. Just as in special relativity, the metric and its inverse may be used to raise and lower indices on tensors.

It will take several weeks to fully appreciate the role of the metric in all of its glory, but for purposes of inspiration we can list the various uses to which $g_{\mu\nu}$ will be put: (1) the metric supplies a notion of "past" and "future"; (2) the metric allows the computation of path length and proper time; (3) the metric determines the "shortest distance" between two points (and therefore the motion of test particles); (4) the metric replaces the Newtonian gravitational field ϕ; (5) the metric provides a notion of locally inertial frames and therefore a sense of "no rotation"; (6) the metric determines causality, by defining the speed of light faster than which no signal can travel; (7) the metric replaces the traditional Euclidean three-dimensional dot product of Newtonian mechanics; and so on. Obviously these ideas are not all completely independent, but we get some sense of the importance of this tensor.

In our discussion of path lengths in special relativity we (somewhat handwavingly) introduced the line element as $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$, which was used to get the length of a path.

Of course now that we know that $\mathrm{d}x^\mu$ is really a basis dual vector, it becomes natural to use the terms "metric" and "line element" interchangeably, and write

$$ds^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \ . \tag{2.29}$$

(To be perfectly consistent we should write this as "g", and sometimes will, but more often than not g is used for the determinant $|g_{\mu\nu}|$.) For example, we know that the Euclidean line element in a three-dimensional space with Cartesian coordinates is

$$ds^2 = (\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2 \ . \tag{2.30}$$

We can now change to any coordinate system we choose. For example, in spherical coordinates we have

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \ , \end{aligned} \tag{2.31}$$

which leads directly to

$$ds^2 = \mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2 + r^2\sin^2\theta\,\mathrm{d}\phi^2 \ . \tag{2.32}$$

Obviously the components of the metric look different than those in Cartesian coordinates, but all of the properties of the space remain unaltered.

Perhaps this is a good time to note that most references are not sufficiently picky to distinguish between "dx", the informal notion of an infinitesimal displacement, and "$\mathrm{d}x$", the rigorous notion of a basis one-form given by the gradient of a coordinate function. In fact our notation "ds^2" does not refer to the exterior derivative of anything, or the square of anything; it's just conventional shorthand for the metric tensor. On the other hand, "$(\mathrm{d}x)^2$" refers specifically to the $(0,2)$ tensor $\mathrm{d}x \otimes \mathrm{d}x$.

A good example of a space with curvature is the two-sphere, which can be thought of as the locus of points in $\mathbf{R}^3$ at distance 1 from the origin. The metric in the (θ, ϕ) coordinate system comes from setting $r = 1$ and $\mathrm{d}r = 0$ in (2.32):

$$ds^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2 \ . \tag{2.33}$$

This is completely consistent with the interpretation of ds as an infinitesimal length, as illustrated in the figure.

As we shall see, the metric tensor contains all the information we need to describe the curvature of the manifold (at least in Riemannian geometry; we will actually indicate somewhat more general approaches). In Minkowski space we can choose coordinates in which the components of the metric are constant; but it should be clear that the existence of curvature
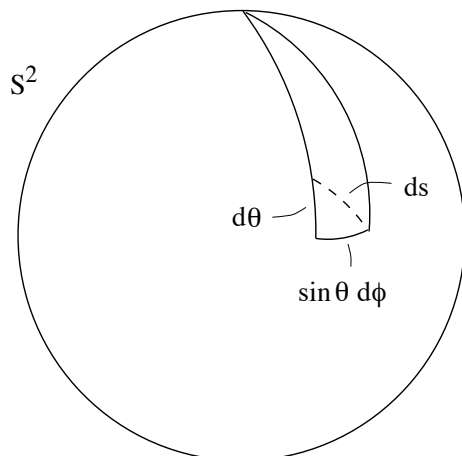
is more subtle than having the metric depend on the coordinates, since in the example above we showed how the metric in flat Euclidean space in spherical coordinates is a function of r and θ. Later, we shall see that constancy of the metric components is sufficient for a space to be flat, and in fact there always exists a coordinate system on any flat space in which the metric is constant. But we might not want to work in such a coordinate system, and we might not even know how to find it; therefore we will want a more precise characterization of the curvature, which will be introduced down the road.

A useful characterization of the metric is obtained by putting $g_{\mu\nu}$ into its **canonical form**. In this form the metric components become

$$g_{\mu\nu} = \text{diag}\,(-1, -1, \ldots, -1, +1, +1, \ldots, +1, 0, 0, \ldots, 0)\,, \tag{2.34}$$

where "diag" means a diagonal matrix with the given elements. If n is the dimension of the manifold, s is the number of $+1$'s in the canonical form, and t is the number of -1's, then $s - t$ is the **signature** of the metric (the difference in the number of minus and plus signs), and $s + t$ is the **rank** of the metric (the number of nonzero eigenvalues). If a metric is continuous, the rank and signature of the metric tensor field are the same at every point, and if the metric is nondegenerate the rank is equal to the dimension n. We will always deal with continuous, nondegenerate metrics. If all of the signs are positive ($t = 0$) the metric is called **Euclidean** or **Riemannian** (or just "positive definite"), while if there is a single minus ($t = 1$) it is called **Lorentzian** or **pseudo-Riemannian**, and any metric with some $+1$'s and some -1's is called "indefinite." (So the word "Euclidean" sometimes means that the space is flat, and sometimes doesn't, but always means that the canonical form is strictly positive; the terminology is unfortunate but standard.) The spacetimes of interest in general relativity have Lorentzian metrics.

We haven't yet demonstrated that it is always possible to but the metric into canonical form. In fact it is always possible to do so at some point $p \in M$, but in general it will

only be possible at that single point, not in any neighborhood of p. Actually we can do slightly better than this; it turns out that at any point p there exists a coordinate system in which $g_{\mu\nu}$ takes its canonical form and the first derivatives $\partial_\sigma g_{\mu\nu}$ all vanish (while the second derivatives $\partial_\rho \partial_\sigma g_{\mu\nu}$ cannot be made to all vanish). Such coordinates are known as **Riemann normal coordinates**, and the associated basis vectors constitute a **local Lorentz frame**. Notice that in Riemann normal coordinates (or RNC's) the metric at p looks like that of flat space "to first order." This is the rigorous notion of the idea that "small enough regions of spacetime look like flat (Minkowski) space." (Also, there is no difficulty in simultaneously constructing sets of basis vectors at every point in M such that the metric takes its canonical form; the problem is that in general this will not be a *coordinate* basis, and there will be no way to make it into one.)

We won't consider the detailed proof of this statement; it can be found in Schutz, pp. 158-160, where it goes by the name of the "local flatness theorem." (He also calls local Lorentz frames "momentarily comoving reference frames," or MCRF's.) It is useful to see a sketch of the proof, however, for the specific case of a Lorentzian metric in four dimensions. The idea is to consider the transformation law for the metric

$$g_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu} \ , \tag{2.35}$$

and expand both sides in Taylor series in the sought-after coordinates $x^{\mu'}$. The expansion of the old coordinates x^μ looks like

$$x^\mu = \left(\frac{\partial x^\mu}{\partial x^{\mu'}} \right)_p x^{\mu'} + \frac{1}{2} \left(\frac{\partial^2 x^\mu}{\partial x^{\mu'_1} \partial x^{\mu'_2}} \right)_p x^{\mu'_1} x^{\mu'_2} + \frac{1}{6} \left(\frac{\partial^3 x^\mu}{\partial x^{\mu'_1} \partial x^{\mu'_2} \partial x^{\mu'_3}} \right)_p x^{\mu'_1} x^{\mu'_2} x^{\mu'_3} + \cdots \ , \tag{2.36}$$

with the other expansions proceeding along the same lines. (For simplicity we have set $x^\mu(p) = x^{\mu'}(p) = 0$.) Then, using some extremely schematic notation, the expansion of (2.35) to second order is

$$
\begin{aligned}
(g')_p &+ (\partial' g')_p \, x' + (\partial' \partial' g')_p \, x'x' \\
&= \left(\frac{\partial x}{\partial x'} \frac{\partial x}{\partial x'} g \right)_p + \left(\frac{\partial x}{\partial x'} \frac{\partial^2 x}{\partial x' \partial x'} g + \frac{\partial x}{\partial x'} \frac{\partial x}{\partial x'} \partial' g \right)_p x' \\
&\quad + \left(\frac{\partial x}{\partial x'} \frac{\partial^3 x}{\partial x' \partial x' \partial x'} g + \frac{\partial^2 x}{\partial x' \partial x'} \frac{\partial^2 x}{\partial x' \partial x'} g + \frac{\partial x}{\partial x'} \frac{\partial^2 x}{\partial x' \partial x'} \partial' g + \frac{\partial x}{\partial x'} \frac{\partial x}{\partial x'} \partial' \partial' g \right)_p x'x'
\end{aligned} \tag{2.37}
$$

We can set terms of equal order in x' on each side equal to each other. Therefore, the components $g_{\mu'\nu'}(p)$, 10 numbers in all (to describe a symmetric two-index tensor), are determined by the matrix $(\partial x^\mu / \partial x^{\mu'})_p$. This is a 4×4 matrix with no constraints; thus, 16 numbers we are free to choose. Clearly this is enough freedom to put the 10 numbers of $g_{\mu'\nu'}(p)$ into canonical form, at least as far as having enough degrees of freedom is concerned.

(In fact there are some limitations — if you go through the procedure carefully, you find for example that you cannot change the signature and rank.) The six remaining degrees of freedom can be interpreted as exactly the six parameters of the Lorentz group; we know that these leave the canonical form unchanged. At first order we have the derivatives $\partial_{\sigma'} g_{\mu'\nu'}(p)$, four derivatives of ten components for a total of 40 numbers. But looking at the right hand side of (2.37) we see that we now have the additional freedom to choose $(\partial^2 x^\mu / \partial x^{\mu'_1} \partial x^{\mu'_2})_p$. In this set of numbers there are 10 independent choices of the indices μ'_1 and μ'_2 (it's symmetric, since partial derivatives commute) and four choices of μ, for a total of 40 degrees of freedom. This is precisely the amount of choice we need to determine all of the first derivatives of the metric, which we can therefore set to zero. At second order, however, we are concerned with $\partial_{\rho'} \partial_{\sigma'} g_{\mu'\nu'}(p)$; this is symmetric in ρ' and σ' as well as μ' and ν', for a total of $10 \times 10 = 100$ numbers. Our ability to make additional choices is contained in $(\partial^3 x^\mu / \partial x^{\mu'_1} \partial x^{\mu'_2} \partial x^{\mu'_3})_p$. This is symmetric in the three lower indices, which gives 20 possibilities, times four for the upper index gives us 80 degrees of freedom — 20 fewer than we require to set the second derivatives of the metric to zero. So in fact we cannot make the second derivatives vanish; the deviation from flatness must therefore be measured by the 20 coordinate-independent degrees of freedom representing the second derivatives of the metric tensor field. We will see later how this comes about, when we characterize curvature using the Riemann tensor, which will turn out to have 20 independent components.

The final change we have to make to our tensor knowledge now that we have dropped the assumption of flat space has to do with the Levi-Civita tensor, $\epsilon_{\mu_1 \mu_2 \cdots \mu_n}$. Remember that the flat-space version of this object, which we will now denote by $\tilde{\epsilon}_{\mu_1 \mu_2 \cdots \mu_n}$, was defined as

$$\tilde{\epsilon}_{\mu_1 \mu_2 \cdots \mu_n} = \begin{cases} +1 \text{ if } \mu_1 \mu_2 \cdots \mu_n \text{ is an even permutation of } 01 \cdots (n-1) \ , \\ -1 \text{ if } \mu_1 \mu_2 \cdots \mu_n \text{ is an odd permutation of } 01 \cdots (n-1) \ , \\ 0 \text{ otherwise} \ . \end{cases} \tag{2.38}$$

We will now define the **Levi-Civita symbol** to be exactly this $\tilde{\epsilon}_{\mu_1 \mu_2 \cdots \mu_n}$ — that is, an object with n indices which has the components specified above *in any coordinate system*. This is called a "symbol," of course, because it is not a tensor; it is defined not to change under coordinate transformations. We can relate its behavior to that of an ordinary tensor by first noting that, given some $n \times n$ matrix $M^\mu{}_{\mu'}$, the determinant $|M|$ obeys

$$\tilde{\epsilon}_{\mu'_1 \mu'_2 \cdots \mu'_n} |M| = \tilde{\epsilon}_{\mu_1 \mu_2 \cdots \mu_n} M^{\mu_1}{}_{\mu'_1} M^{\mu_2}{}_{\mu'_2} \cdots M^{\mu_n}{}_{\mu'_n} \ . \tag{2.39}$$

This is just a true fact about the determinant which you can find in a sufficiently enlightened linear algebra book. If follows that, setting $M^\mu{}_{\mu'} = \partial x^\mu / \partial x^{\mu'}$, we have

$$\tilde{\epsilon}_{\mu'_1 \mu'_2 \cdots \mu'_n} = \left| \frac{\partial x^{\mu'}}{\partial x^\mu} \right| \tilde{\epsilon}_{\mu_1 \mu_2 \cdots \mu_n} \frac{\partial x^{\mu_1}}{\partial x^{\mu'_1}} \frac{\partial x^{\mu_2}}{\partial x^{\mu'_2}} \cdots \frac{\partial x^{\mu_n}}{\partial x^{\mu'_n}} \ . \tag{2.40}$$

This is close to the tensor transformation law, except for the determinant out front. Objects which transform in this way are known as **tensor densities**. Another example is given by the determinant of the metric, $g = |g_{\mu\nu}|$. It's easy to check (by taking the determinant of both sides of (2.35)) that under a coordinate transformation we get

$$g(x^{\mu'}) = \left|\frac{\partial x^{\mu'}}{\partial x^{\mu}}\right|^{-2} g(x^{\mu}) . \tag{2.41}$$

Therefore g is also not a tensor; it transforms in a way similar to the Levi-Civita symbol, except that the Jacobian is raised to the -2 power. The power to which the Jacobian is raised is known as the **weight** of the tensor density; the Levi-Civita symbol is a density of weight 1, while g is a (scalar) density of weight -2.

However, we don't like tensor densities, we like tensors. There is a simple way to convert a density into an honest tensor — multiply by $|g|^{w/2}$, where w is the weight of the density (the absolute value signs are there because $g < 0$ for Lorentz metrics). The result will transform according to the tensor transformation law. Therefore, for example, we can define the Levi-Civita tensor as

$$\epsilon_{\mu_1\mu_2\cdots\mu_n} = \sqrt{|g|}\,\tilde{\epsilon}_{\mu_1\mu_2\cdots\mu_n} . \tag{2.42}$$

It is this tensor which is used in the definition of the Hodge dual, (1.87), which is otherwise unchanged when generalized to arbitrary manifolds. Since this is a real tensor, we can raise indices, etc. Sometimes people define a version of the Levi-Civita symbol with upper indices, $\tilde{\epsilon}^{\mu_1\mu_2\cdots\mu_n}$, whose components are numerically equal to the symbol with lower indices. This turns out to be a density of weight -1, and is related to the tensor with upper indices by

$$\epsilon^{\mu_1\mu_2\cdots\mu_n} = \ \text{sgn}(g)\frac{1}{\sqrt{|g|}}\,\tilde{\epsilon}^{\mu_1\mu_2\cdots\mu_n} . \tag{2.43}$$

As an aside, we should come clean and admit that, even with the factor of $\sqrt{|g|}$, the Levi-Civita tensor is in some sense not a true tensor, because on some manifolds it cannot be globally defined. Those on which it can be defined are called **orientable**, and we will deal exclusively with orientable manifolds in this course. An example of a non-orientable manifold is the Möbius strip; see Schutz's *Geometrical Methods in Mathematical Physics* (or a similar text) for a discussion.

One final appearance of tensor densities is in integration on manifolds. We will not do this subject justice, but at least a casual glance is necessary. You have probably been exposed to the fact that in ordinary calculus on $\mathbf{R}^n$ the volume element $d^n x$ picks up a factor of the Jacobian under change of coordinates:

$$d^n x' = \left|\frac{\partial x^{\mu'}}{\partial x^{\mu}}\right| d^n x . \tag{2.44}$$

There is actually a beautiful explanation of this formula from the point of view of differential forms, which arises from the following fact: *on an n-dimensional manifold, the integrand is properly understood as an n-form.* The naive volume element $d^n x$ is itself a density rather than an n-form, but there is no difficulty in using it to construct a real n-form. To see how this works, we should make the identification

$$d^n x \leftrightarrow \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} \ . \tag{2.45}$$

The expression on the right hand side can be misleading, because it looks like a tensor (an n-form, actually) but is really a density. Certainly if we have two functions f and g on M, then $\mathrm{d}f$ and $\mathrm{d}g$ are one-forms, and $\mathrm{d}f \wedge \mathrm{d}g$ is a two-form. But we would like to interpret the right hand side of (2.45) as a coordinate-dependent object which, in the x^μ coordinate system, acts like $\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1}$. This sounds tricky, but in fact it's just an ambiguity of notation, and in practice we will just use the shorthand notation "$d^n x$".

To justify this song and dance, let's see how (2.45) changes under coordinate transformations. First notice that the definition of the wedge product allows us to write

$$\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} = \frac{1}{n!} \tilde{\epsilon}_{\mu_1 \cdots \mu_n} \mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_n} \ , \tag{2.46}$$

since both the wedge product and the Levi-Civita symbol are completely antisymmetric. Under a coordinate transformation $\tilde{\epsilon}_{\mu_1 \cdots \mu_n}$ stays the same while the one-forms change according to (2.16), leading to

$$
\begin{aligned}
\tilde{\epsilon}_{\mu_1 \cdots \mu_n} \mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_n} &= \tilde{\epsilon}_{\mu_1 \cdots \mu_n} \frac{\partial x^{\mu_1}}{\partial x^{\mu_1'}} \cdots \frac{\partial x^{\mu_n}}{\partial x^{\mu_n'}} \mathrm{d}x^{\mu_1'} \wedge \cdots \wedge \mathrm{d}x^{\mu_n'} \\
&= \left| \frac{\partial x^\mu}{\partial x^{\mu'}} \right| \tilde{\epsilon}_{\mu_1' \cdots \mu_n'} \mathrm{d}x^{\mu_1'} \wedge \cdots \wedge \mathrm{d}x^{\mu_n'} \ .
\end{aligned}
\tag{2.47}
$$

Multiplying by the Jacobian on both sides recovers (2.44).

It is clear that the naive volume element $d^n x$ transforms as a density, not a tensor, but it is straightforward to construct an invariant volume element by multiplying by $\sqrt{|g|}$:

$$\sqrt{|g'|} \, \mathrm{d}x^{0'} \wedge \cdots \wedge \mathrm{d}x^{(n-1)'} = \sqrt{|g|} \, \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} \ , \tag{2.48}$$

which is of course just $(n!)^{-1} \epsilon_{\mu_1 \cdots \mu_n} \mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_n}$. In the interest of simplicity we will usually write the volume element as $\sqrt{|g|} \, d^n x$, rather than as the explicit wedge product $\sqrt{|g|} \, \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1}$; it will be enough to keep in mind that it's supposed to be an n-form.

As a final aside to finish this section, let's consider one of the most elegant and powerful theorems of differential geometry: Stokes's theorem. This theorem is the generalization of the fundamental theorem of calculus, $\int_b^a dx = a - b$. Imagine that we have an n-manifold

M with boundary ∂M, and an $(n-1)$-form ω on M. (We haven't discussed manifolds with boundaries, but the idea is obvious; M could for instance be the interior of an $(n-1)$-dimensional closed surface ∂M.) Then $\mathrm{d}\omega$ is an n-form, which can be integrated over M, while ω itself can be integrated over ∂M. Stokes's theorem is then

$$\int_M \mathrm{d}\omega = \int_{\partial M} \omega \ . \tag{2.49}$$

You can convince yourself that different special cases of this theorem include not only the fundamental theorem of calculus, but also the theorems of Green, Gauss, and Stokes, familiar from vector calculus in three dimensions.