Chapter 2

Manifolds

Problem Set #2: 2.1, 2.3, 2.5, 2.7, 2.9 (Due Monday, Oct. 14th)

2.1 Maps and basis

The special theory relativity is dealing with a very special space-time known as Minkowski space or (flat space-time). After ten years of trials and errors the theory was generalized by Einstein to deal with more general space-times. In the more general theory of relativity, known as general relativity, the space-time itself is a dynamical variable that moves (or deforms) according to equations of motion. Unfortunately, there is no easy way to understand physics of general relativity without learning the basics of differential geometry. The central objects in differential geometry are the manifolds (or more precisely differentiable manifolds), but their introduction is impossible without a general discussion of maps.

Map \( \phi \) from set \( M \) to set \( N \), denoted as

\[
\phi : M \rightarrow N, \tag{2.1}
\]

is a relation which assigns to each element of \( M \) an exactly one element of \( N \). Domain is the entire set \( M \) and image (or range) is defined as the set of points of \( N \) to which points of \( M \) are mapped. The most trivial example is the identity map which maps each element to itself

\[
I : A \rightarrow A. \tag{2.2}
\]

The maps are called one-to-one (or injective) if each element of \( M \) is mapped to at most one element in \( N \) and onto (or surjective) map if at least one element of \( M \) is mapped to each element in \( N \) (i.e. the entire set \( N \) is the
image). The invertible (or bijective) map are the maps which are one-to-one and onto. For invertible maps one can also define the inverse map

\[ \phi^{-1} : N \to M \]  

by

\[ \phi^{-1}(\phi(a)) = a. \]

For example, \( \phi(x) = x^2 \) is not one-to-one nor onto, \( \phi(x) = e^x \) is one-to-one but not onto, \( \phi(x) = x^3 - x \) is onto but not one-to-one, and \( \phi(x) = x^3 \) is one-to-one and onto and, thus, invertible \( \phi^{-1}(x) = x^{1/3} \).

**Composition** is a map \( \psi \circ \phi : A \to C \) which is a composition of maps \( \phi : A \to B \) and \( \psi : B \to C \) defined as

\[ \psi \circ \phi \equiv \psi(\phi(a)) \in C \]

for any \( a \in A \). Thus,

\[ \phi^{-1} \circ \phi = I. \]  

For a composition \( g \circ f \) of two maps \( f : \mathbb{R}^m \to \mathbb{R}^n \) and \( g : \mathbb{R}^n \to \mathbb{R}^l \), we can use the chain rule to calculate partial derivatives

\[ \frac{\partial}{\partial x^a} (g \circ f) = \sum_{b} \frac{\partial f^b}{\partial x^a} \frac{\partial g^c}{\partial y^b} \]  

(2.7)

sometimes abbreviated by

\[ \frac{\partial}{\partial x^a} = \sum_{b} \frac{\partial y^b}{\partial x^a} \frac{\partial}{\partial y^b} \]  

(2.8)

where \( x^a \in \mathbb{R}^m \), \( y^b \in \mathbb{R}^n \) and \( z^c \in \mathbb{R}^l \). Note that for \( m = n \) the matrix \( \partial f^b/\partial x^a \) (or \( \partial y^b/\partial x^a \)) is called a *Jacobian*.

A map \( \phi : \mathbb{R}^m \to \mathbb{R}^n \) is a \( C^p \) map if \( p \) of its derivatives exist and continues. In general relativity we are usually interested in infinitely differentiable maps \( C^\infty \) (also called *smooth maps*) between more general manifolds. Then the manifolds \( M \) and \( N \) are *diffeomorphic* if there exist a \( C^\infty \) map \( \phi \) between \( M \) and \( N \) with a \( C^\infty \) inverse. It is a well established experimental fact that the fundamental laws of nature are invariant under diffeomorphism transformation. This is the key observation which generalizes the concept of Lorentz invariance of special relativity to diffeomorphism invariance of general relativity (usually called *general covariance*).

A subset \( U \) of \( M \) together with a one-to-one map \( \phi : U \to \mathbb{R}^n \) is a **chart** (or **coordinate system**) if the image is an open set in \( \mathbb{R}^n \) (i.e. does
not contain its boundary points). Given two charts \((M, \phi : M \to \mathbb{R}^n)\) and \((N, \psi : N \to \mathbb{R}^n)\) and a map \(f : M \to N\) one can define differentiation operation as

\[
\frac{\partial f}{\partial x^\mu} = \frac{\partial}{\partial x^\mu} (\psi \circ f \circ \phi^{-1}) (x^\nu).
\]  

(2.9)

Unfortunately only very few manifolds can be covered with a single chart, but all sufficiently smooth manifolds can be covered with an infinite collection of charts (e.g. one needs at least two charts to cover \(S^1\), but only one chart to cover \(S^1 \times \mathbb{R}\)). A countable collection of charts \(\{\{U_\alpha, \phi_\alpha\}\}\) is an atlas if the charts cover all of the manifold (i.e. \(\cup_\alpha U_\alpha = M\) and the charts are smoothly sewn together (i.e. for any pair of charts \((U_\alpha, \phi_\alpha)\) and \((U_\beta, \phi_\beta)\) the composition maps \(\phi_\alpha \circ \phi_\beta^{-1}\) and \(\phi_\beta \circ \phi_\alpha^{-1}\) are smooth)). Then a smooth \(n\)-dimensional manifold is a set \(M\) with an atlas which contains every possible compatible chart.

The manifolds that we are considering are not embedded in a higher dimensional space, although an arbitrary \(n\)-dimensional manifold can always be embed in \(\mathbb{R}^{2n}\) according to Whitney’s embedding theorem. For example, all of the results of general relativity can be understood from the point of view of four-dimensional manifolds. Then our next task is to construct a tangent space at some point \(p\) using only intrinsic properties of the \(n\)-dimensional manifold \(M\). This can be done by defining a set of direction derivatives \(\frac{d}{dx} \in T_p\) along parametrized curves through \(p\) which act on a space to smooth functions \(f \in F\) where \(f : M \to \mathbb{R}\) and \(\frac{d}{dx} : f \to \mathbb{R}\). From the definition it should be clear that the set of directional derivatives is a vector space (i.e. for any \(\frac{d}{dx}\) and \(\frac{d}{dy}\) their linear combination \(a\frac{d}{dx} + b\frac{d}{dy}\) is also a directional derivative) and it elements obey the product rule

\[
\left(a\frac{d}{d\lambda} + b\frac{d}{d\eta}\right) (fg) = a\frac{d}{d\lambda} (fg) + b\frac{d}{d\eta} (fg) =
\]

\[
= af \frac{d}{d\lambda} g + ag \frac{d}{d\lambda} f + bf \frac{d}{d\lambda} g + bg \frac{d}{d\lambda} f =
\]

\[
= f \left(a\frac{d}{d\lambda} + b\frac{d}{d\lambda}\right) g + g \left(a\frac{d}{d\lambda} + b\frac{d}{d\lambda}\right) f. \tag{2.10}
\]

The \(n\) basis vectors for this vector space \(T_p\) can, for example, be constructed from a given chart with coordinates \(x^\mu\) as \(n\) partial derivatives \(\partial_\mu\) at \(p\). It is now easy to show that the vectors \(\{\partial_\mu\}\) are linearly independent and span all of the vectors space of directional derivatives. Consider a coordinate chart \(\phi : M \to \mathbb{R}^n\), a curve \(\gamma : \mathbb{R} \to M\) and a function \(f : M \to \mathbb{R}\). If \(\lambda\) is a
parameter along $\gamma$, then

\[
\frac{d}{d\lambda} f = \frac{d}{d\lambda} (f \circ \gamma) \quad \text{rewriting using a formal notation}
\]

\[
= \frac{d}{d\lambda} (f \circ \phi^{-1} \circ \phi \circ \gamma)
\]

\[
= \frac{d}{d\lambda} (f \circ \phi^{-1}) \phi_x \quad \text{inserting the identity map}
\]

\[
= \frac{d}{d\lambda} \phi_x f \quad \text{using the chain rule}
\]

\[
= \frac{d\nu^\mu}{d\lambda} \partial _\mu f \quad \text{returning to informal notation. (2.11)}
\]

Since this is true for an arbitrary smooth function $f$ we conclude that the directional derivative

\[
\frac{d}{d\lambda} = \frac{d\nu^\mu}{d\lambda} \partial _\mu
\]

and thus the coordinate basis $\{\partial _\mu\}$ is a good basis for the tangent space $T_p$,

\[
\hat{e}_{(\mu)} = \partial _\mu. (2.13)
\]

which may or may not be orthogonal.

The basis vectors in a new coordinate system are given by

\[
\partial _{\mu'} = \frac{\partial \zeta^\mu}{\partial \zeta^{\mu'}} \partial _\mu. (2.14)
\]

and, similar to Lorentz transformations in flat space-time, the matrix

\[
\frac{\partial \zeta^\mu}{\partial \zeta^{\mu'}} (2.15)
\]

or its inverse

\[
\left( \frac{\partial \zeta^\mu}{\partial \zeta^{\mu'}} \right)^{-1} = \frac{\partial \xi^{\mu'}}{\partial \zeta^\mu} (2.16)
\]

can be used to transform from one coordinate basis to another, e.g.

\[
V^\mu' = \frac{\partial \xi^{\mu'}}{\partial \zeta^\mu} V^\mu. (2.17)
\]

Given a tangent space of vectors $T_p$ at a point $p$ one can introduce a dual space of one-forms $T^*_p$ as a space of maps

\[
\omega : T_p \to \mathbb{R}. (2.18)
\]

For example, the gradient of a function $df$ is one-form whose action on a directional derivative is

\[
df \left( \frac{d}{d\lambda} \right) = \frac{df}{d\lambda}. (2.19)
\]
Then the coordinate basis for a dual space can be constructed from coordinate basis of the original vector space by
\[ dx^\mu (\partial_\nu) \equiv \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu \] (2.20)
with transformation properties for basis one-forms
\[ dx'^\mu = \frac{\partial x'^\mu}{\partial x^\mu} dx^\mu \] (2.21)
and for components of one-forms
\[ \omega_{\mu'} = \frac{\partial x^\mu}{\partial x'^\mu} \omega_\mu. \] (2.22)

## 2.2 Transformations

More general tensors are transformed as in a flat space-time with Lorentz transformations replaced by general coordinate transformations,
\[ T^{\mu_1...\mu_k}_{\nu_1...\nu_l} = \frac{\partial x^{\mu_1}}{\partial x^{\mu_1'}} \frac{\partial x^{\mu_2}}{\partial x^{\mu_2'}} \cdots \frac{\partial x^{\mu_k}}{\partial x^{\mu_k'}} T_{\nu_1...\nu_l}^{\mu_1...\mu_k}. \] (2.23)

As an example consider a (0, 2) tensor
\[ S_{\mu\nu} = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \] (2.24)
where \( x = x^1 \) and \( y = x^2 \), or in terms of basis one-forms as
\[ S = S_{\mu\nu} (dx^\mu \otimes dx^\nu) = x dx \otimes dx + 1 dy \otimes dy = x dx^2 + dy^2. \] (2.25)

To find the same tensor in another coordinate system defined by, e.g.
\[ x' = x^{1/3} \]
\[ y' = e^{x+y} \] (2.26)
(2.27)
we can also take a gradient derivative, e.g.
\[ x = (x')^3 \]
\[ y = \ln(y') - (x')' \] (2.28)
(2.29)
to obtain the basis one-forms
\[ dx = 3 (x')^2 \, dx' \]
\[ dy = (y')^{-1} \, dy' - 3 (x')^2 \, dx'. \] (2.30)
(2.31)
By substituting (2.26), (2.30) and (2.31) into (2.25) we obtain

\[ S = 9 \left( x' \right)^7 \, dx'^2 + \left( y' \right)^{-1} \, dy' - 3 \left( x' \right)^2 \, dx' \]

\[ = 9 \left( \left( x' \right)^7 + \left( x' \right)^4 \right) \, dx^2 + \left( y' \right)^{-2} \, dy' - 3 \left( y' \right)^{-1} \, dx' \, dy' - 3 \left( y' \right)^{-1} \, \left( x' \right)^2 \, dy' \, dx' \]

or in a matrix form

\[ S = \begin{pmatrix}
9 \left( \left( x' \right)^7 + \left( x' \right)^4 \right) & -3 \left( y' \right)^{-1} \left( x' \right)^2 \\
-3 \left( y' \right)^{-1} \left( x' \right)^2 & \left( y' \right)^{-2}
\end{pmatrix}. \tag{2.33}
\]

As was mentioned earlier the partial derivatives of vectors are not legitimate tensors and do not transform correctly under coordinated transformation

\[ \frac{\partial}{\partial x'^\mu} W_\nu = \frac{\partial x^\mu}{\partial x'^\nu} \frac{\partial}{\partial x^\mu} \left( \frac{\partial x^\nu}{\partial x'^\nu} W_\nu \right) = \]

\[ = \frac{\partial x^\mu}{\partial x'^\mu} \frac{\partial x^\nu}{\partial x'^\nu} \frac{\partial}{\partial x^\mu} W_\nu + \frac{\partial x^\mu}{\partial x'^\mu} W_\nu \frac{\partial}{\partial x'^\nu} \left( \frac{\partial x^\nu}{\partial x'^\nu} \right) = \]

\[ = \frac{\partial x^\mu}{\partial x'^\mu} \frac{\partial x^\nu}{\partial x'^\nu} \frac{\partial}{\partial x^\mu} W_\nu + W_\nu \frac{\partial^2 x^\nu}{\partial x'^\mu \partial x'^\nu}. \tag{2.34}
\]

with an exception of flat space-time where the second term contains a derivative of Lorentz transformation and thus vanishes. But if one uses the (antisymmetric) exterior derivative instead the second (symmetric) term must vanish which is exactly what one needs for \( dW \) to transform as a \((0, 2)\) tensor.

One of the most important tensors in general relativity is the metric \( g_{\mu \nu} \) which is a symmetric \((0, 2)\) tensor with non-vanishing determinant. The last property allows us to define an inverse metric \( g^{\mu \nu} \) by

\[ g^{\mu \nu} g_{\mu \sigma} = \delta^\mu_\sigma \tag{2.35}
\]

which is together with the metric itself can be used to raise and lower indices. For a given metric the quantity that is used to measure distances is a line element defined by

\[ ds^2 \equiv g_{\mu \nu} dx^\mu dx^\nu. \tag{2.36}
\]

For example, a line element for a Euclidean three-dimensional space is

\[ ds^2 = dx^2 + dy^2 + dz^2 \tag{2.37}
\]

or in spherical coordinates defined by

\[ x = r \sin \theta \cos \phi \tag{2.38}
\]

\[ y = r \sin \theta \sin \phi \tag{2.39}
\]

\[ z = r \cos \theta \tag{2.40}
\]
the line element is given by
\[ ds^2 = dr^2 + r^2d\theta + r^2 \sin^2 \theta d\phi^2. \] (2.41)

The simplest example of a space with curvature is a two-sphere whose metric is given by
\[ ds^2 = d\theta^2 + \sin^2 \theta d\phi^2 \] (2.42)
which one cannot put in a canonical form
\[ ds^2 = dx^2 + dy^2 \] (2.43)
in some new coordinates. The same is true in 3 + 1 dimensions as it is not always possible to perform coordinate transformations to put the metric in a canonical form
\[ ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \] (2.44)
everywhere on the manifold, but it is always possible to find such a transformation locally. In other words one can always define a local coordinates system \textit{(local Lorentz frame)} in which the metric is flat and moreover its first derivative vanishes. However, the second derivative does not vanish and by Taylor expanding both sides of
\[ g_{\mu\nu}' = \frac{\partial x^\mu}{\partial x'^\lambda} \frac{\partial x'^\nu}{\partial x^\lambda} g_{\mu\nu} \] (2.45)
to the second order one can show that there are 20 parameters to be specified (related to 20 components of the Riemann tensor).

Since vectors \( V \) can be thought of as a directional derivative, one can use it to map smooth functions \( f \) to produce smooth functions, \( V(f) \). An anti-symmetric composition of such maps is called a \textbf{commutator} or \textbf{Lie bracket},
\[ [X, Y] \equiv X(Y(f)) - Y(X(f)). \] (2.46)
or in components
\[ [X, Y]^\mu = X^\lambda \partial_\lambda Y^\mu - Y^\lambda \partial_\lambda X^\mu. \] (2.47)
Despite of the fact that the partial derivatives are involved one can check that the resulting vector is legitimate \((1, 0)\) tensor by preforming coordinate transformations. Another object that must be modified to transform as a tensor is the Levi-Civita tensor defined as
\[ \epsilon_{\mu_1 \ldots \mu_4} = \sqrt{|g|} \epsilon_{\mu_1 \ldots \mu_4}. \] (2.48)
where the Levi-Civita symbol \( \epsilon \) is called a tensor density. In general, tensor density is a quantity which would transform as a tensor once multiplied by
a determinant of the metric raised to some power. For example, a volume element

\[ d^n x \equiv dx^0 \wedge ... \wedge dx^{n-1} \]  

(2.49)

is a density tensor and not a true tensor, but

\[ \sqrt{|g|} d^n x = \sqrt{|g|} dx^0 \wedge ... \wedge dx^{n-1} \]  

(2.50)

is a tensor or actually an \( n \)-from. Then we can write the Stokes’s theorem on an arbitrary \( n \)-dimensional manifold as

\[ \int_M d\omega = \int_{\partial M} \omega \]  

(2.51)

where \( \omega \) is an \( n-1 \)-from and \( d\omega \) is an \( n \)-from.

Note that

\[ g_{\mu'\nu'} = \left( \frac{\partial x^\mu}{\partial x^\mu'} \right) \left( \frac{\partial x^{\nu'}}{\partial x^{\nu}} \right) \eta_{\mu\nu} \]

\[ g = \det \left( \frac{\partial x^\mu}{\partial x^\mu'} \right) \det \left( \frac{\partial x^{\nu'}}{\partial x^{\nu}} \right) \det \eta \]

\[ \det \left( \frac{\partial x^\mu}{\partial x^\mu'} \right) = \sqrt{-g} = \sqrt{|g|} \]

(2.52)

and thus one can think of \( \sqrt{|g|} \) as a determinant of a Jacobian transformation matrix.