Notes on “General Relativity” (Wald, 1984)

Robert B. Scott,\textsuperscript{1,2*}

\textsuperscript{1}Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin,
Austin, Texas, USA
\textsuperscript{2}Department of Physics, University of Brest,
Brest, France

\*To whom correspondence should be addressed; E-mail: rscott@ig.utexas.edu.

March 10, 2012
Part I

Fundamentals
Chapter 1

Introduction
Chapter 2

Manifolds and Tensor Fields

2.5 Problems

Problem 1 (a) Show that the overlap functions, discussed in § 2.1 in the proof that the 2-sphere is a manifold, are smooth.

The overlap functions are

\[ f_i^\pm \circ (f_j^\pm)^{-1} \]

where \( f_i^\pm \) is the projection, for example:

\[ f_1^+(x, y, z) = (y, z) \]

on the domain \( x > 0 \). Let’s construct a particular overlap function. For

\[ f_2^+(x, y, z) = (x, z) \]

This projects the hemisphere \( y > 0 \) onto the \( x - z \) plane. Its inverse takes points in the \( x - z \) plane and projects them back onto the 2-sphere surface:

\[ (f_2^+)^{-1}(x, z) = (x, +\sqrt{1 - x^2 - z^2}, z) \]

where the positive root corresponds to the domain of \( f_2^+ \). The composite with \( f_1^+ \) will project these points onto the \( y - z \) plane.

\[ f_1^+ \circ (f_2^+)^{-1}(x, z) = (\sqrt{1 - x^2 - z^2}, z) \]
We must show that this function is $C^\infty$.

**Problem 1 (b)** Construct two coordinate systems that cover the 2-sphere.

We use spherical coordinates being careful to avoid the singularities at the poles. But remember that the charts must map from open subsets of the manifold into open subsets of $\mathbb{R}^n$, so we cannot just use the two hemispheres since the Equator would be omitted.

For the first coordinate system, $\psi_1$, use spherical coordinates and avoid the poles by restricting $\pi/6 < \theta < 5\pi/6$. Then the subset of the manifold $S^2$ that’s covered is

$$O_1 = \left\{ (x^1, x^2, x^3) \in S^2 \mid -\frac{\sqrt{3}}{2} < x^3 < \frac{\sqrt{3}}{2} \right\}$$

So the first chart

$$\psi_1: O_1 \rightarrow U_1 \subset \mathbb{R}^2$$

is specifically

$$\psi_1(x^1, x^2, x^3) = (\theta, \phi) \quad (2.1)$$

with

$$\theta = \arccos(z)$$
$$\phi = \arctan(y/x) \quad (2.2)$$

To fill the holes near the poles, we construct another coordinate system with a similar chart, $\psi_2$. We simply replace $x^3$ with $x^1$ so that the poles of the second spherical coordinate system lie on the $x$-axis. These cover

$$O_2 = \left\{ (x^1, x^2, x^3) \in S^2 \mid -\frac{\sqrt{3}}{2} < x^1 < \frac{\sqrt{3}}{2} \right\}$$

and we see our coordinates cover the manifold.

It seems to me that we should also check the *regularity* of the coordinates, (Faber, 1983, §1.3), although it wasn’t clear to me that Wald discusses this. This amounts to inverting the function (2.1),

$$\psi^{-1}(\theta, \phi) = (x^1, x^2, x^3)$$
which is written
\[ \mathbf{X}(u, v) = (x, y, z) \]
in the notation of Faber (1983). To demonstrate regularity of the coordinates we must show that
\[
\mathbf{X}_1 \equiv \frac{\partial \mathbf{X}}{\partial u} = \frac{\partial \psi^{-1}}{\partial \theta} \\
\mathbf{X}_2 \equiv \frac{\partial \mathbf{X}}{\partial v} = \frac{\partial \psi^{-1}}{\partial \phi}
\]
are linearly independent. (The notation \( \mathbf{X}_1 \) is also from Faber (1983) but corresponds closely to that of Wald (1984, Eq. (2.2.1)), who writes \( X_1 \).) First we invert \( \psi_1 \), which is just spherical coordinates with \( r = 1 \):
\[
\begin{align*}
x^1 &= \sin \theta \cos \phi \\
x^2 &= \sin \theta \sin \phi \\
x^3 &= \cos \theta
\end{align*}
\]
Next we find the vectors
\[
\mathbf{X}_1 = \frac{\partial \psi^{-1}_1}{\partial \theta} = \begin{pmatrix} \frac{\partial x^1}{\partial \theta} \\ \frac{\partial x^2}{\partial \theta} \\ \frac{\partial x^3}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{pmatrix}
\]
and
\[
\mathbf{X}_2 = \frac{\partial \psi^{-1}_1}{\partial \phi} = \begin{pmatrix} \frac{\partial x^1}{\partial \phi} \\ \frac{\partial x^2}{\partial \phi} \\ \frac{\partial x^3}{\partial \phi} \end{pmatrix} = \begin{pmatrix} -\sin \theta \sin \phi \\ \sin \theta \cos \phi \\ 0 \end{pmatrix}
\]
To be linearly independent in 2D means these vectors \( \mathbf{X}_1 \) and \( \mathbf{X}_2 \) are never parallel. So we can simply check
\[
\mathbf{X}_1 \times \mathbf{X}_2 = \begin{pmatrix} \sin^2 \theta \cos \phi \\ \sin^2 \theta \sin \phi \\ \cos \theta \sin \theta \end{pmatrix}
\]
which has magnitude \( |\mathbf{X}_1 \times \mathbf{X}_2| = \sin \theta > 0 \) on the interval \( \pi/6 < \theta < 5\pi/6 \).
(In fact we knew this immediately from experience with spherical coordinates and that is exactly why we restricted the domain of chart \( \psi_1 \).)
**Problem 2** Prove that any smooth function \( F: \mathbb{R}^n \to \mathbb{R} \) can be written in the form equation (2.2.2).

Let’s first make sense of the identity provided in the hint,

\[
F(x) - F(a) = (x - a) \int_0^1 F'[t(x - a) + a] dt \tag{2.5}
\]

Let’s take the antiderivative of the integrand \( F' \),

\[
\int F'[t(x - a) + a] dt = \frac{1}{(x - a)} F[t(x - a) + a]
\]

which is easily verified by taking the derivative of both sides with respect to \( t \). And then simply substitute this into the RHS:

\[
F(x) - F(a) = (x - a) \int_0^1 F'[t(x - a) + a] dt \\
= F[t(x - a) + a]_{t=1}^{t=0} \\
= F(x) - F(a). \tag{2.6}
\]

So the identity clearly holds when \( x \in \mathbb{R} \). Now let’s extend it to \( x \in \mathbb{R}^n \). First define a vector

\[
y^\mu = t(x^\mu - a^\mu)
\]

Then by the chain rule,

\[
\frac{dF}{dt} = \sum \frac{\partial F}{\partial y^\mu} \frac{dy^\mu}{dt} = \sum F'(x^\mu - a^\mu)
\]

And

\[
F(x) - F(a) = \int_0^1 \frac{dF}{dt} dt \\
= \int_0^1 \sum F'(x^\mu - a^\mu) dt \\
= \sum (x^\mu - a^\mu) \int_0^1 F' dt \\
= \sum (x^\mu - a^\mu) H_\mu \tag{2.7}
\]
where

\[ H_{\mu} \equiv \int_{0}^{1} F' dt. \]

So it’s not clear why Wald suggests using induction.

Problem 2 Verify that the commutator, defined by equation (2.2.14), satisfies the linearity and Leibnitz properties, and hence defines a vector field.

Let’s start with linearity (in the arguments). Let

\[ f = ag + bh \]

where \( a, b \in \mathbb{R} \) are constants and \( g, h \in \mathcal{F} : M \to \mathbb{R} \) are \( C^\infty \) functions from \( M \) into \( \mathbb{R} \). Then by the properties of tangent vectors introduced on p. 15, and the definition of the commutator (2.2.14), we have

\[
[v, w](ag + bh) = v[w(ag + bh)] - w[v(ag + bh)] \\
= v[a w(g) + b w(h)] - w[a v(g) + b v(h)] \\
= a v[w(g)] + b w[v(h)] - a w[v(g)] + b w[v(h)] \\
= a [v, w](g) + b [v, w](h) , \text{ regrouped terms, used definition (2.2.14)}
\]

(2.8)

And we see that the commutator obeys linearity in the arguments \( i.e. \) property one on p. 15.

Now let’s consider property two on p. 15, Leibnitz rule.

\[
[v, w](gh) = v[w(gh)] - w[v(gh)] \\
= v[g w(h) + h w(g)] - w[g v(h) + h v(g)] \\
= v(g) w(h) + v(h) w(g) + g v[w(h)] + h v[w(g)] - w(g) v(h) \\
- w(h) v(g) - g w[v(h)] - h w[v(g)] \\
= h [v, w](g) + g [v, w](h) , \text{ regrouped terms, used definition (2.2.14)}
\]

(2.9)

And we see that the commutator obeys the Leibnitz rule \( i.e. \) property two on p. 15. And thus the commutator of two vector fields is also a vector field.
Part II

Advanced Topics
Chapter 10

Initial Value Formulation

Wants to argue that GR admits a well posed initial value problem. For him, this means that it should permit “reasonable specification” of the initial conditions, and that Einstein’s equations uniquely predict the future evolution. [emphasis added].

Then he adds two additional conditions for what it takes for the Einstein Equations to be well posed. Amazingly, we considers sensitivity to initial conditions to render the problem ill-posed:

First, in an appropriate sense, “small changes” in initial data should produce only correspondingly “small changes” in the solution over any fixed compact region of spacetime. If this property were not satisfied, the theory would lose essentially all predictive power, since initial conditions can be measured only to finite accuracy. It is generally assumed that the pathological behaviour that would result from the failure of this property does not occur in physics.

(Wald, 1984, p. 244)

A second condition for the initial value problem to be well-posed is that changes to the initial conditions in a region $S$ should only lead to changes in the casual future of $S$, that is $J^+(S)$. Otherwise, we would have faster than light signals.

He will be expressing the Einstein equations as 2nd order hyperbolic systems. (Other expressions are possible.)
10.1 Initial Value Formulation of Particles and Fields

Gives the example of \( n \) particles governed by Newtonian mechanics, a system that admits an initial value formulation.

Gives the Klein-Gordon equations as another example. Argues that if the initial conditions are specified with analytic functions, then the Klein-Gordon system admits an initial value formulation – that is we can specify arbitrary initial conditions of \( \phi \) and its first temporal derivative on a hyper surface of constant time \( \Sigma_0 \). Then the solution is uniquely determined for later times. He shows that in principle the solution at later times can be written as a Taylor series about the initial time, and one can in principle build all terms in the Taylor series. (In essence because the governing equation gives one the 3rd derivative in terms of the first, the 4th in terms of the 2nd, etc.

The Cauchy-Kolwalski, proven in say Courant and Hilbert, states that the power series \[\text{[sounds like a Taylor series to me]}\] solution has a “finite radius of convergence”. In the theorem they state there is an open neighbourhood \( O \) of the initial time \( t_0 \) hyper surface such that within \( O \) there is a unique analytical solution of the Klein-Gordon equations.

According to Wald, the Cauchy-Kolwalski theorem is not enough for our purposes for two reasons. (1) It doesn’t guarantee the that the map from (analytical) initial conditions to the solution at finite time will be continuous. (2) When the initial conditions are given by analytic functions, then the initial conditions over the entire initial time \( t_0 \) hyper surface are determined by the initial conditions within an open region [because from the information within the open region one could form all the derivatives and then build the Taylor series and describe the analytical function everywhere on the hyper surface.] But then if we change the initial conditions within some open region we must in fact change the initial conditions everywhere. Apparently this violates the causality requirement! Wald says therefore that we must consider initial conditions given by non-analytical functions! [This is amazing to me! Personally, I think it means that you cannot arbitrarily change initial conditions.].
Part III

Appendices
Chapter 11
Topological Spaces

Compact: Every open cover of a set has a finite sub cover. The open interval (0, 1) in the standard topology of the reals is not compact since it has an open cover provided by the sets $O_n = (1/n, 1)$ with $n = 2, 3, \ldots$ that doesn’t allow a finite subcover.

THEOREM A.8 A subset of $R^n$ is compact if and only if it is closed and bounded.
Bibliography
